Faculty of Mathematical Studies

MA204 Real Analysis: Metric Spaces

Index: Click on the bullet next to the section you want

- **Introduction**
- **Prerequisites: revision**
- **The Bolzano-Weierstrass Theorem**
- **Definition of a metric space**
- **Examples of metric spaces**
- **Convergence and Continuity**
- **Cauchy sequences**
- **Completeness**
- **The Topology of Metric Spaces**
- **The Cantor ternary set**
- **Open and Closed sets: theorems and Examples**
- **Compactness**
- **Continuous functions**

Introduction: Aims

Measuring distance has been part of human activity for millennia. Certainly the development of agriculture made it necessary, and in particular the annual Nile flood made land measurement a necessity in ancient Egypt. It is said that a (3,4,5) triangle was used to measure right angles in that context.

In co-ordinate geometry we measure the distance between two points in 2 or 3 dimensions by using Pythagoras' Theorem, and we then generalise the algebraic expression which arises and use it to define a measure of distance in n-dimensional space.

In the development of modern mathematics it was realised that we measure distance in other contexts than simply points in Euclidean space. In approximating functions we are concerned with measures of accuracy, and want to be able to describe how close two functions are in some sense, perhaps by looking at their graphs in some way.

The developments in topology around the turn of the century and subsequently gave rise to the theory of Metric Spaces, in which the fundamental features of distance measurements are formulated as basic axioms, and theorems developed which can then be applied to many examples involving distance.

The aim of this section of the course is to study some of the basic theory of metric spaces, to explore some examples, both familiar and unfamiliar, and to classify metric spaces through some of their properties.

Prerequisites: Revision

MA201, in particular the section on limits of sequences of real numbers.

Definition 1 A sequence (a_n) of real numbers is said to *converge, with limit a*, if

$$\forall \varepsilon > 0, \exists N \in \mathbf{N}, \forall n \geq N, |a_n - a| < \varepsilon \ .$$

Definition 2 A function $f(x)$ is said to have limit l as $x \to a$ if

$$\forall \varepsilon > 0, \exists \delta, \forall x, 0 < |x-a| < \delta \ \Rightarrow |f(x) - l| < \varepsilon \, .$$

Definition 3 A function $f(x)$ is continuous at $x = a$ if $\lim_{x \to a} f(x) = f(a)$, i.e.

$$\forall \varepsilon > 0, \exists \delta, \forall x, |x-a| < \delta \ \Rightarrow |f(x) - f(a)| < \varepsilon \, .$$

Look back in your MA201 notes at the examples and tutorial sheets which you were given to help you understand these definitions.

One of the most important results about sequences of real numbers is the Bolzano-Weierstrass Theorem, on the **following page**.

Theorem 1 The Bolzano-Weierstrass Theorem
Every bounded sequence S of real numbers contains a convergent subsequence.

Proof
The idea of the proof is to successively bisect the interval, choosing a point each time in such a way that the resulting sequence converges.

Suppose without loss of generality that the sequence is contained in the interval [0,1].

We define three sequences $(a_n),(b_n),(x_n)$ inductively as follows. Let $a_1 = 0, b_1 = 1,$ and let x_1 be any term from the sequence S. Now if the interval [½,1] contains infinitely many terms of the sequence then let $a_2 = \frac{1}{2}, b_2 = 1,$ and let x_2 be any term of the sequence S in this interval which has not already been selected. Otherwise we let $a_2 = 0, b_2 = \frac{1}{2}$ with x_2 again any term of the sequence S in this interval which has not already been selected.

Now let us suppose that we have specified a_n and b_n with $b_n - a_n = 1/2^{n-1}$, with $[a_n, b_n]$ containing infinitely many terms of the sequence S and with $x_n \in [a_n, b_n]$ a member of the sequence S not previously selected. We now perform the inductive step.

If the interval $\left[\frac{a_n + b_n}{2}, b_n\right]$ contains infinitely many terms of the sequence then let $a_{n+1} = \frac{a_n + b_n}{2}, b_{n+1} = b_n$ and let $x_{n+1} \in [a_{n+1}, b_{n+1}]$ be any term of the sequence S in this interval which has not already been selected. Otherwise we let $a_{n+1} = a_n, b_{n+1} = \frac{a_n + b_n}{2}$ and again let $x_{n+1} \in [a_{n+1}, b_{n+1}]$ be any term of the sequence S in this interval which has not already been selected. We now have $b_{n+1} - a_{n+1} = 1/2^n$.

Now the sequences (a_n) and (b_n) are monotonic and bounded. They therefore converge. Since $b_n - a_n = 1/2^{n-1}$ they have the same limit a. Since $x_n \in [a_n, b_n]$ we have $|x_n - a| \le |x_n - a_n| + |a_n - a| \le 1/2^{n-1} + |a_n - a| \to 0$ as $n \to \infty$.

So the sequence (x_n) is a convergent subsequence of S.

Definition of a Metric Space

We shall now give the definition of a metric space, followed by some examples.

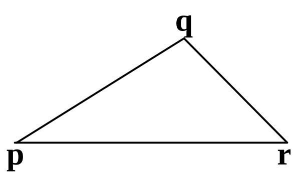
Definition 4 A metric space (X, d) consists of a set X together with a "distance function" d on X, i.e. a function $d: X \times X \to \mathbf{R}$ satisfying

(a) $\forall p, q \in X,\ d(p, q) \geq 0$;
(b) $d(p, q) = 0 \Leftrightarrow p = q$;
(c) $\forall p, q \in X, d(p, q) = d(q, p)$;
(d) $\forall p, q, r \in X, d(p, r) \leq d(p, q) + d(q, r)$.

(a) says that distance is never negative. (b) then says that distance is always positive unless the two points are identical.

(a) and (b) should clearly be satisfied by any reasonable measure of distance, while (c) says that distance is the same in either "direction".

Axiom (d) is normally called the triangle inequality, and with any given formula for the distance function is usually the most difficult to verify. Geometrically it expresses the simple relationship



Examples of Metric Spaces

1. $X = \mathbf{R}$; $d(x, y) = |x - y|$.

2. $X = \mathbf{C}$; $d(z, w) = |z - w|$.

3. $X = \mathbf{R}^n$; If $x = (x_1, x_2, \ldots, x_n)$ and $y = (y_1, y_2, \ldots, y_n)$ we define the Euclidean, or Pythagorean, metric by $d(x, y) = \sqrt{\sum_{i=1}^{n} (x_i - y_i)^2}$.

We need to give an algebraic proof of the triangle inequality, as follows.

Suppose $z = (z_1, z_2, \ldots, z_n)$. We have to prove that

$$\sqrt{\sum_{i=1}^{n} (x_i - z_i)^2} \le \sqrt{\sum_{i=1}^{n} (x_i - y_i)^2} + \sqrt{\sum_{i=1}^{n} (y_i - z_i)^2}.$$

Let $a_i = x_i - y_i$ and $b_i = y_i - z_i$. Then $x_i - z_i = a_i + b_i$. So we need to show that $\sqrt{\sum (a_i + b_i)^2} \le \sqrt{\sum a_i^2} + \sqrt{\sum b_i^2}$. This is equivalent to proving that $\left(\sum a_i b_i\right)^2 \le \sum a_i^2 \sum b_i^2$.

This is known as the Cauchy-Schwarz Inequality, proved as follows.

We consider the quadratic expression in t below, which is non-negative for all t.

$$\sum (a_i t + b_i)^2 = \sum a_i^2 t^2 + 2\sum a_i b_i t + \sum b_i^2 = t^2 \sum a_i^2 + 2t \sum a_i b_i + \sum b_i^2.$$

Since this is non-negative for all t the discriminant must be negative, which proves the inequality. To investigate when equality occurs, we note that this is the case where the discriminant is zero, corresponding to the quadratic having just one real root. So there is a number k for which $\sum (a_i t + b_i)^2 = 0$, i.e. $b_i = -k a_i$ for all i.

4. The discrete metric on any set X defined by $d(x, y) = \begin{cases} 1 & \text{if } x \ne y, \\ 0 & \text{if } x = y. \end{cases}$

5. $X = \mathbf{R}^n$, $d(x, y) = \sum |x_i - y_i|$.

6. $X = \mathbf{R}^n$, $d(x, y) = \max_{1 \le i \le n} |x_i - y_i|$.

more examples are on the next page

7. X is the set of continuous real valued functions on $[a,b]$.
$d(f,g) = \sup\{|f(x) - g(x)| : a \le x \le b\}$.

8. X is the set of bounded real valued functions on $[a,b]$.
$d(f,g) = \sup\{|f(x) - g(x)| : a \le x \le b\}$.

9. X is the set of continuous real valued functions on $[a,b]$. $d(f,g) = \int_a^b |f - g|$.
Note that if <u>continuous</u> is replaced by <u>bounded</u> we no longer have a metric.

10. X is the set of continuous real valued functions on $[a,b]$. $d(f,g) = \sqrt{\int_a^b (f - g)^2}$.

11. X is the set of all bounded sequences. $d((a_i),(b_i)) = \sum_{i=1}^{\infty} \frac{|a_i - b_i|}{2^i}$.

12. $X = \mathbf{R}^2$, $d((x_1, y_1),(x_2, y_2)) = \begin{cases} |y_1 - y_2| & \text{if } x_1 = x_2, \\ |x_1 - x_2| + |y_1| + |y_2| & \text{if } x_1 \ne x_2. \end{cases}$ (The Vineyard metric).

Theorem 2 If d is a metric on a set X then so is $d/(1+d)$.

Note This enables us to define a bounded metric on any set.

Proof
The only property which is non-trivial is the triangle inequality. The expression $t/(1+t)$ is increasing for $t > 0$ (do the algebra, or sketch a graph - click **here** for details). So using the triangle inequality for d itself gives

$$\frac{d(x,z)}{1+d(x,z)} \le \frac{d(x,y)+d(y,z)}{1+d(x,y)+d(y,z)}$$
$$= \frac{d(x,y)}{1+d(x,y)+d(y,z)} + \frac{d(y,z)}{1+d(x,y)+d(y,z)}$$
$$\le \frac{d(x,y)}{1+d(x,y)} + \frac{d(y,z)}{1+d(y,z)}.$$

Convergence and Continuity

Both these ideas, when in the familiar setting of real numbers and functions of a single real variable, involve only the notion of distance between numbers, and so will immediately generalise to arbitrary metric spaces.

Definition 5 A sequence (x_n) of points in a metric space (X,d) is said to converge to $x \in X$ if and only if $d(x_n, x) \to 0$ as $n \to \infty$. Note that $d(x_n, x) \to 0$ as $n \to \infty$ is a statement about a sequence of real numbers, for which the definition of convergence has already been established. Adapting **Definition 1**, the more general definition here becomes

$$\forall \varepsilon > 0, \exists N \in \mathbf{N}, \forall n \geq N, d(x_n, x) < \varepsilon$$

Examples

13. $X = \mathbf{R}^3$ with the **Euclidean metric**. We prove that $(x_n, y_n, z_n) \to (x, y, z)$ if and only if $x_n \to x$ and $y_n \to y$ and $z_n \to z$. Firstly suppose that $(x_n, y_n, z_n) \to (x, y, z)$. This means that $d((x_n, y_n, z_n), (x, y, z)) \to 0$. In other words

$$\sqrt{(x_n - x)^2 + (y_n - y)^2 + (z_n - z)^2} \to 0.$$

We then have the following inequalities

$$\sqrt{(x_n - x)^2 + (y_n - y)^2 + (z_n - z)^2} \geq \sqrt{(x_n - x)^2} = |x_n - x|,$$
$$\sqrt{(x_n - x)^2 + (y_n - y)^2 + (z_n - z)^2} \geq \sqrt{(y_n - y)^2} = |y_n - y|,$$
$$\sqrt{(x_n - x)^2 + (y_n - y)^2 + (z_n - z)^2} \geq \sqrt{(z_n - z)^2} = |z_n - z|,$$

which prove that $x_n \to x$ and $y_n \to y$ and $z_n \to z$.

Now suppose that $x_n \to x$ and $y_n \to y$ and $z_n \to z$. Then the algebra of limits together with the continuity of the square root function at zero show that

$$\sqrt{(x_n - x)^2 + (y_n - y)^2 + (z_n - z)^2} \to 0.$$

More examples on the next page

14. Consider the **discrete metric** on any set X. Suppose that $x_n \to x$. Using the definition of limit, and taking $\varepsilon = \frac{1}{2}$ tells us that $\exists N \in \mathbf{N}, \forall n \geq N, d(x_n, x) < \frac{1}{2}$, which is only possible if $\exists N \in \mathbf{N}, \forall n \geq N, d(x_n, x) = 0,$ since in the discrete metric distances are either 0 or 1. So the sequence must be identically equal to x from some point onwards.

[These two examples demonstrate that convergence depends on the metric. There will be many sequences convergent in the **Euclidean metric** but not in the discrete metric.]

15. X is the set of bounded real valued functions on [0,1].

$$d(f, g) = \sup\{|f(x) - g(x)| : 0 \leq x \leq 1\}.$$

We let $f_n(x) = x^n$ and $f(x) = \begin{cases} 0 & \text{if } x < 1, \\ 1 & \text{if } x = 1. \end{cases}$

For any value of n, $\sup|f_n(x) - f(x)| = 1$. Therefore $f_n \not\to f$.

The Definition of Continuity

Definition 6 $f:(X,d_X)\to(Y,d_Y)$ is said to be <u>continuous</u> at $a\in X$ if and only if

$$\forall\varepsilon>0,\exists\delta>0,\forall x\in X, d_X(x,a)<\delta\Rightarrow d_Y(f(x),f(a))<\varepsilon.$$

Theorem 3

The following are equivalent.

(i) f is continuous at a,

(ii) for all sequences for which $\lim_{n\to\infty}a_n=a, \lim_{n\to\infty}f(a_n)=f(a)$.

Proof

Let (a_n) be any sequence with limit a. Then $\exists N,\forall n\geq N, d_X(a_n,a)<\delta.$ from the definition of continuity above we conclude that $\forall n\geq N, d_Y(f(a_n),f(a))<\varepsilon.$ this shows that (i) implies (ii).

To prove the converse we prove the equivalent statement that if f is not continuous at a then there exists a sequence (a_n) for which $\lim_{n\to\infty}a_n=a$, but $\lim_{n\to\infty}f(a_n)\neq f(a)$.
Negating the definition of continuity above gives

$$\exists\varepsilon>0,\forall\delta>0,\exists x\in X, d_X(x,a)<\delta \text{ and } d_Y(f(x),f(a))\geq\varepsilon.$$

Taking $\delta=\frac{1}{n}$ enables us to say that $\exists a_n\in X, d_X(a_n,a)<\delta$ and $d_Y(f(a_n),f(a))\geq\varepsilon$, i.e. there exists a sequence (a_n) for which $\lim_{n\to\infty}a_n=a$, but $\lim_{n\to\infty}f(a_n)\neq f(a)$.

Examples: Applications of Theorem 3

16. To prove that the function which is 0 on the rationals and 1 on the irrationals is discontinuous everywhere, we simply note that given any irrational *x* we can produce a sequence of rationals with *x* is its limit, and similarly, given a rational *y* we can produce a sequence of irrationals with *y* as limit.

Click **here** for some revision notes relating to this example

17. We use this result whenever we solve an equation numerically using an iterative process of the form $x_{n+1}=F(x_n)$, using the argument that if *F* is continuous and if the sequence (x_n) has a limit *x* then $x=F(x)$.

Click **here** for an example of such an iteration

Cauchy Sequences

The definitions and examples for convergence have all needed knowledge of the limiting value, or at least a good guess. In this section we consider a criterion for convergence which depends only on the sequence itself.

Definition 7 A sequence (x_n) is said to be a Cauchy Sequence if and only if

$$\forall \varepsilon > 0, \exists N \in \mathbf{N}, \forall m \geq N, \forall n \geq N, d(x_m, x_n) < \varepsilon.$$

So a sequence is a Cauchy Sequence provided that the terms are all arbitrarily close to one another eventually.

Click **here** for further discussion of this definition.

Theorem 4 A Cauchy sequence (x_n) of real numbers has a real limit.

Proof
We shall use the completeness axiom for the real numbers, namely that a non-empty set of real number which is bounded above has a least upper bound (supremum).

We first prove that a Cauchy sequence is bounded. Take $\varepsilon = 1$. Then taking $n = N$ in the definition we see that $\forall m \geq N, d(x_m, x_N) < 1$.

Now let $P = \min\{x_1, x_2, \ldots x_{N-1}, x_N - 1\}$ and $Q = \max\{x_1, x_2, \ldots x_{N-1}, x_N + 1\}$.

Then for all n we have $P \leq x_n \leq Q$, i.e. the sequence is bounded

Now by the **Bolzano-Weierstrass Theorem** the sequence (x_n) has a convergent subsequence (x_{n_i}), with limit x, say.

Intuitively, because the subsequence converges to x, eventually its terms will be arbitrarily close to x. Then because the entire sequence is a Cauchy sequence, this ensures that eventually the terms will all be arbitrarily close to one another, and hence arbitrarily close to terms of the subsequence, and so arbitrarily close to x itself.

Mathematically we formulate this argument as follows.
Let $\varepsilon > 0$. Then $\exists K_1, \forall n_i \geq K_1, d(x_{n_i}, x) < \varepsilon/2$.

Since (x_n) is a Cauchy sequence, $\exists K_2, \forall m \geq K_2, \forall n \geq K_2, d(x_n, x_m) < \varepsilon/2$.

Let $N = \max(K_1, K_2)$. Let x_{n_k} be a member of the subsequence for which $n_k \geq N$.
Then we have $\forall n \geq N, d(x_n, x) \leq d(x_n, x_{n_k}) + d(x_{n_k}, x) < \varepsilon/2 + \varepsilon/2 = \varepsilon$.
Thus the entire sequence (x_n) converges to x.

Completeness

Definition 8 A metric space is said to be complete if every Cauchy sequence has a limit.

Examples

18. $\mathbf{R}^n$, with either the Euclidean metric, or those of examples 5 and 6, is complete. We use the Bolzano-Weierstrass Theorem in the proof, noting that in **R** all three metrics are identical.

So let $(x^{(i)})$ be a Cauchy sequence in $\mathbf{R}^n$, where $x^{(i)}$ is written in terms of co-ordinates as $x^{(i)} = (x_1^{(i)}, x_2^{(i)}, \ldots, x_n^{(i)})$.

Now given $\varepsilon > 0, \exists N \in \mathbf{N}, \forall i \geq N, \forall j \geq N, d(x^{(i)}, x^{(j)}) < \varepsilon.$

It follows that for $1 \leq k \leq n,\ \ \forall m \geq N, \forall n \geq N, \left|x_k^{(i)} - x_k^{(j)}\right| < \varepsilon.$ So each sequence of coordinates is a Cauchy sequence, and so has a limit x_k.

We therefore have $x^{(i)} \to x = (x_1, x_2, \ldots, x_n).$

19. The set **Q** of rational numbers is not complete. Let $x_n = 0.1010010001\ldots1\underbrace{00\ldots0}_{n\text{ zeros}}1$.

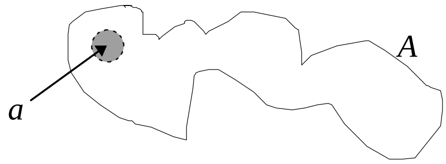
The limit of this sequence in **R** is not a rational number, and so is not in **Q**, which is therefore not complete.

The Topology of Metric Spaces

In this section we shall discuss certain kinds of sets of points within a metric space which have properties analogous to those of open intervals, and closed intervals. We shall use these to generalise the notion of continuity for functions. The classification of sets according to these properties is important in much of advanced pure mathematics.

In this section some of the examples use countable and uncountable sets. For some revision notes click **here**.

Definition 9 A point a is said to be an interior point of a set A in a metric space (X,d) if there is a ball surrounding a lying entirely within A. Symbolically $\exists \varepsilon > 0, \forall x, d(x,a) < \varepsilon \Rightarrow x \in A$.



Note In $\mathbf{R}^2$ a "ball" is a filled-in disk. In $\mathbf{R}^3$ a "ball" is a solid sphere. In $\mathbf{R}$ a "ball" is an interval, in the definition having the form $(a-\varepsilon, a+\varepsilon)$. The term "ball" is intended to be understood generically to apply to any metric space.

Definition 10 A subset A of a metric space (X,d) is said to be an open set if all its points are interior points.

Examples

20. Let A denote the set of all points (x,y) in the plane for which x is positive. Then A is an open set. Given any such point (x,y), let $\varepsilon = \frac{x}{2} > 0$. Then the disc centred at (x,y) with radius ε is entirely contained within A.

21. Let $A = \{(x,y,z) \in \mathbf{R}^3 : x^2 + y^2 \le 1\}$. Some points of A are interior points, for example the origin. But others are not, for example the point $a = (1,0,0)$. Any ball surrounding this must be of the form $\{(x,y,z) \in \mathbf{R}^3 : (x-1)^2 + y^2 + z^2 < \varepsilon\}$. The point $(1+\sqrt{\varepsilon/2}, 0, 0)$ belongs to this ball, but not to the set A.

22. In $\mathbf{R}$ the set of rational numbers does not form an open set. In fact no point is an interior point, for any interval centred on a rational contains some irrationals.
Click **here** if you need further explanation of this example.

23. In $\mathbf{R}^2$ the set of lines through the origin whose slopes are integers form a set, none of whose points are interior points.

Further definitions and examples are on the next page.

Sometimes we are interested in focusing on the set of interior points of a set.

Definition 11 Given a set A, the interior of A, denoted by A^{o}, is the set of all interior points of A.

Definition 12 A point a is said to be a limit point (or point of accumulation or cluster point) of a set A in a metric space (X,d) if every ball surrounding a contains a point of A other than a itself. Symbolically

$$\forall \varepsilon > 0, \exists x \in A, 0 < d(x,a) < \varepsilon.$$

Sometimes limit points belong to a set and sometimes not. For example in $\mathbf{R}$ both 0 and 1 are limit points of the open interval (0,1), but neither belongs to the interval. The point ½ is a limit point of the set (it is also an interior point - the two categories are not disjoint.

Definition 13 A subset A of a metric space (X,d) is said to be a closed set if it contains all its limit points.

Examples

24. We have already seen that the interval (0,1) in $\mathbf{R}$ is not a closed set, since the limit points 0 and 1 do not belong to the set.

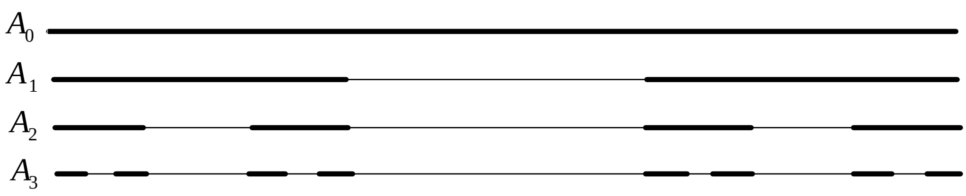
25. In $\mathbf{R}$ the set A of points satisfying $x \geq 1$ is closed. In fact every point of the set is a limit point. Also if we let y be any point not in A, then $y < 1$, and the interval centred on y with radius $|y-1|/2$ contains no points of A. Therefore y is not a limit point of A. In fact this argument shows that the complement of A is an open set.

26. In $\mathbf{R}^2$ let A be the set of points (x,y), both of whose co-ordinates are irrational. This set has no interior points. Every point of the set is a limit point. Every point of the complement is also a limit point. So A is neither open nor closed, and the complement of A is neither open nor closed.

27. Let $A = \{1, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \ldots\}$. A has just one limit point, namely 0. 0 is not a member of A and so A is not closed. The set $A \cup \{0\}$ is a closed set.

The Cantor Ternary Set

28. The Cantor Ternary Set This is a standard example which is often used in many areas of Topology. In outline, we take the closed interval [0,1], remove the middle open third, remove the middle open third of the two intervals left, and so on.



We can characterise this set by considering the "decimal" expansion of numbers in [0,1], but in scale 3 instead of in scale 10. The set A_1 consists of all numbers which have an expansion whose first place is not 1 (we include the number 1/3, which has the non-terminating scale 3 expansion $0\cdot0222\ldots$ So this set consists of two closed intervals. The set A_2 consists of those numbers without a 1 in the first two places. The set A_3 consists of those numbers without a 1 in the first three places, and so on. We then let $A=\bigcap_{n\geq 0} A_n$. A consists of all numbers which have an expansion in scale 3 without a 1 in any position.

Now the **diagonal argument** which Cantor used to establish the **uncountability** of the real numbers can also be used to show that the Cantor Ternary Set A is uncountably infinite. We can also calculate the length of A_n, and show that it tends to 0. Therefore A contains no intervals, and therefore its interior is empty. Now we can prove that the complement of A is open, as follows.

Since $A=\bigcap_{n\geq 0} A_n$, **De Morgan's law** gives $A^c=\left(\bigcap_{n\geq 0} A_n\right)^c=\bigcup_{n\geq 0} A_n^c$. The set A_n^c consists of 2^n-1 open intervals, and is therefore an open set. Any point belonging to A^c must belong to A_n^c for some n, and is therefore an interior point.

Every point of A is a limit point, for if $a=0\cdot a_1a_2a_3\ldots$, where $a_i\neq 1$ for any i, then $a=\lim_{n\to\infty} 0.a_1a_2\ldots a_{i-1}b_ia_{i+1}\ldots$ where $b_i=0$ if $a_i=2$, and $b_i=2$ if $a_i=0$. Moreover, every point of A is a limit point of the complement, for $a=\lim_{n\to\infty} 0.a_1a_2\ldots a_{n-1}101$, and $0.a_1a_2\ldots a_{n-1}101\notin A$.

Now A is closed, but instead of giving a proof we shall deduce the fact from the next theorem.

Some theorems and examples on open and closed sets are on the next page.

Theorem 5 In a metric space (X,d), a set is closed if and only if its complement is an open set.

Proof Suppose that A is a closed subset of X, and let $a \in A^c$. Then a is not a limit point of A. Negating the definition of limit point **Definition 12** gives

$$\exists \varepsilon > 0, \forall x \in A, d(x,a) \geq \varepsilon.$$

So a is an interior point and therefore A^c is open.

Now suppose that A is an open subset of X, and let a be a limit point of A^c. So $\forall \varepsilon > 0, \exists x \in A^c, 0 < d(x,a) < \varepsilon$ i.e. $\forall \varepsilon > 0, \exists x \notin A, 0 < d(x,a) < \varepsilon$. So a is not an interior point of A, and because A is open a does not belong to A. So $a \in A^c$, and so A^c is closed.

Theorem 6 The union of any collection of open sets is an open set.

Proof Let $\{A_\alpha\}$ denote an arbitrary collection of open sets, and let $A = \bigcup_\alpha A_\alpha$. Let a be an arbitrary point of A.

Then $\exists \alpha, a \in A_\alpha$. A_α is open, so $\exists \varepsilon > 0, \forall x, d(x,a) < \varepsilon \Rightarrow x \in A_\alpha$.

But $A_\alpha \subseteq A$, and so $\forall x, d(x,a) < \varepsilon \Rightarrow x \in A$. Therefore A is open.
Click **here** for a discussion of the notation $\{A_\alpha\}$ if you don't understand it.

Corollary The intersection of any collection of closed sets is a closed set.

Proof Apply Theorem 5 to the result of Theorem 6. Click **here** for details.

more theorems and examples on the nest page

Theorem 7 The intersection of two open sets is an open set.

Proof
Let A and B denote two open sets, and let $c \in A \cap B$. A and B are open sets so

$$\exists \varepsilon_1, \forall x, d(x,c) < \varepsilon_1 \Rightarrow x \in A \quad \text{and} \quad \exists \varepsilon_2, \forall x, d(x,c) < \varepsilon_2 \Rightarrow x \in B.$$

Let $\varepsilon = \min(\varepsilon_1, \varepsilon_2)$. Then $\forall x, d(x,c) < \varepsilon \Rightarrow x \in A \cap B$. Therefore $A \cap B$ is open.

Corollary 1 The intersection of any finite collection of open sets is an open set.

Proof By induction on the number of sets. Click **here** for details.

Corollary 2 The union of any finite collection of closed sets is a closed set.

Proof Apply the result of Theorem 5 to Theorem 7. Click **here** for details.

Examples

29. In $\mathbf{R}^2$ with the **Euclidean metric** let $A_n = \{(x,y) : d((x,y),(0,0)) < 1/n\}$. Each A_n is an open set. But $\bigcap_{n \in \mathrm{N}} A_n = \{(0,0)\}$, which is not an open set.

30. Consider the complement of example 29. Click **here** for details.

Definition 14 Given a set A, the <u>closure of A</u>, denoted by $\overline{A}$, consists of the set A together with all its limit points.

Definition 15 The <u>boundary</u> of a set A is defined to be $\overline{A} \cap \overline{A^c}$.

Examples

31. In $\mathbf{R}^2$ let $A = \{(x, y) : x > 0\}$. The boundary of A is the y-axis.

32. In $\mathbf{R}^2$ let $A = \{(x, y) : x \text{ and } y \text{ are both rational}\}$. The boundary of A is the whole of $\mathbf{R}^2$.

33. In $\mathbf{R}$ let $A = \{1, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \ldots\}$. The boundary of A is $A \cup \{0\}$.

Compactness

Definition 16 A set A in a metric space is said to be compact if, given any infinite sequence of points of A, there is a convergent subsequence whose limit is in A. [Sometimes this concept is called sequential compactness, to contrast with a more abstract definition of compactness which is used in non-metric topological spaces.]

Theorem 8

1. A compact set in a metric space is closed and bounded.
2. A closed and bounded subset A of $\mathbf{R}^n$ is compact.

Part 2 is known as the Heine-Borel Theorem, originally a theorem about **R**.

Proof 1(a) Let A be a compact subset of a metric space (X,d). Let a be a limit point of A. Let $A_n = \left\{x \in X : 0 < d(x,a) < \frac{1}{n}\right\}$.

Because a is a limit point of A, $\exists a_n \in A_n, a_n \in A$. A is compact, and so the sequence of points (a_n) has a convergent subsequence whose limit is in A.

But the limit of the whole sequence (a_n) is the point a, and so $a \in A$, and therefore A is closed.

1(b) Let A be a subset of a metric space (X,d), and suppose that A is not bounded. We shall prove that A is not compact.

Let a be a point in A. Since A is unbounded, $\forall n \in N, \exists a_n \in A, d(a_n,a) > n$. This sequence (a_n) is unbounded.

Consider an arbitrary subsequence (b_n). We shall prove that this cannot converge.

Now because it is a subsequence, $b_n = a_p$ for some $p \geq n$, and so $d(b_n,a) > n$. Now let b_n be an arbitrary member of the sequence and let $k = d(b_n,a)$. Now choose an integer $m > k+1$.
Then $d(b_m,a) > m$, and $d(b_m,b_n) \geq d(b_m,a) - d(b_n,a) > k+1-k = 1$.

This proves that the sequence (b_n) is not a **Cauchy sequence**, so it cannot be convergent. Therefore A is not compact.

The proof of part 2 is on the next page.

2. This says that the converse of 1 is true in $\mathbf{R}^n$.

Let $x^{(m)} (m = 1,2,\ldots)$ be a sequence of points of A, written in terms of co-ordinates as $x^{(m)} = \left(x_1^{(m)}, x_2^{(m)}, \ldots, x_n^{(m)}\right)$. Since A is bounded, the sequence of first co-ordinates forms a bounded sequence of real numbers, which by the **Bolzano-Weierstrass theorem** has a convergent subsequence $x_1^{m_1}, x_1^{m_2}, \ldots$ with limit x_1 say.

Now consider the subsequence of $x^{(m)}$ where m takes only the values $m_1, m_2, \ldots$ The sequence of second co-ordinates for this sequence of points again forms a bounded sequence and so it has a convergent subsequence with limit x_2 say.

We continue this process, finishing with a subsequence of the sequence of n-th co-ordinates with a limit x_n. Restricting m to this subsequence then provides a subsequence of $x^{(m)}$ with limit $(x_1, x_2, \ldots, x_n)$.

Note

The converse of 1. is not true in general. For example if we consider an infinite set X with the **discrete metric**, then X is bounded, since all points are at distance 1 from any given point a. Every set is both open and closed. However a sequence (x_n) of distinct points of X has no convergent subsequence, since again all the points are at distance 1 from one another. We can prove a partial result however.

Theorem 9 A closed subset of a compact set is compact.

Proof

Let A be a compact set and let B be a closed subset of A. Let (b_n) be a sequence of points of B.

Since (b_n) is also a sequence of points of A, there is a subsequence converging to a point $a \in A$.

Since all the members of the subsequence belong to B, a is a limit point of B and therefore $a \in B$ since B is closed.

Hence B is compact.

Continuous Functions

We shall now characterise continuity in terms of open sets.

Theorem 10 The function $f:(X,d_X)\to(Y,d_Y)$ is continuous if and only if the inverse image of every open set in Y is an open set in X.

Proof We first explain the idea of inverse image. This is defined and symbolised by

$$f^{-1}(B)=\{x\in X: f(x)\in B\}.$$

Click **here** for some examples of inverse images if you need them.

Suppose that f is continuous at every point $a\in X$ in the sense of **Definition 6**. Let B be any open set in Y. If the inverse image of B is empty then it is an open set. If not, let a be any point in the inverse image of B.

For some $\varepsilon>0$, $S=\{y:d_Y(y,f(a))<\varepsilon\}\subseteq B$.

Since f is continuous at a, $\exists\delta, 0<d_X(x,a)<\delta\Rightarrow d_Y(f(x),f(a))<\varepsilon$, which tells us that $\exists\delta, 0<d_X(x,a)<\delta\Rightarrow f(x)\in S$, i.e. $\exists\delta, 0<d_X(x,a)<\delta\Rightarrow x\in f^{-1}(B)$. So a is an interior point of $f^{-1}(B)$ and so $f^{-1}(B)$ is open in X.

Now to prove the converse, we assume that the inverse image of every open set in Y is open in X, and prove that f is continuous (as in definition 6) at any point $a\in X$.

Let $\varepsilon>0$ and let $B=\{y\in Y:d_Y(y,f(a))<\varepsilon\}$. Then B is an open subset of Y and so its inverse image is open in X.

The point a belongs to this inverse image and is an interior point, so $\exists\delta>0, d_X(x,a)<\delta\Rightarrow x\in f^{-1}(B)$.

From the definition of the set B we deduce that $d_X(x,a)<\delta\Rightarrow d_Y(f(x),f(a))<\varepsilon$, so that **Definition 6** is satisfied, and f is therefore continuous.

Theorem 11 Let $f:(X,d_X)\to(Y,d_Y)$ be a continuous function, and let A be a compact subset of X. Then $f(A)$ is a compact subset of Y.

Proof Let (c_n) be a sequence of points of $f(A)$. Then $\forall n,\exists a_n\in A, c_n=f(a_n)$. The sequence (a_n) has a convergent subsequence, (b_n) say, converging to a point $b\in A$, since A is compact. Let $d_n=f(b_n)$. Then (d_n) is a subsequence of (c_n). Since f is continuous and $b_n\to b$, it follows by **Theorem 3(ii)** that $f(b_n)\to f(b)$, i.e. the sequence (d_n) converges to a point in $f(A)$, so $f(A)$ is compact.

More theorems on compactness and continuity on the next page

Theorem 12
Let $f:(X,d_X)\to(\mathbf{R},d)$ (where d denotes the Euclidean metric) be a continuous function, and let A be a compact subset of X. Then f is bounded on A and attains its bounds.

Proof
Theorem 11 proves that $f(A)$ is compact, so it is bounded by **Theorem 8.1**. Let a be the supremum of the set $f(A)$, i.e. $a=\sup\{f(x):x\in A\}$. From the definition of supremum there is a sequence (x_n) of points of A with $a=\lim_{n\to\infty} f(x_n)$. The set $f(A)$ is compact and is therefore closed by **Theorem 8.1**, so $a\in f(A)$, i.e. $\exists x\in A, a=f(x)$. So f attains its supremum. Similarly f attains its infimum.

Note that this theorem was proved in MA201 for real-valued functions of a real variable.

Theorem 12 If $f:(X,d_X)\to(Y,d_Y)$ is a continuous bijection and X is compact then the inverse function f^{-1} is continuous.

Proof
We shall use the idea of continuity from **Theorem 10**. Note that inverse images under f^{-1} are the same as direct images under f, so we have to show that the image of any open subset of X is an open subset of Y.

Since X is compact, by **Theorem 10** the set $f(X)$ is compact. Since f is a bijection, $f(X)=Y$, and so Y is compact. Let A be an open subset of X. The complement A^c is closed, and so by **Theorem 8.1** is compact. The image $f(A^c)$ is therefore compact. It is therefore closed, and so its complement is open.

Now because f is a bijection, we have $(f(A^c)^c)=f(A)$, so $f(A)$ is open.

Note If X is not compact the result is not necessarily true, even when f is a continuous bijection, as the following example shows.

Take X and Y to be the real number system. For X we use the **discrete metric**, and for Y the **Euclidean metric**. Let f be the identity. Now every subset of X is open in the discrete metric, and so the inverse image of every open subset of Y is open in X. Thus f is continuous. However, if we consider a singleton set, say $\{0\}$, it is open in X, but its image $\{0\}$ is not open in Y, so that f^{-1} is not continuous.

The function $f(t)=\dfrac{t}{1+t}$ is increasing for $t > 0$.

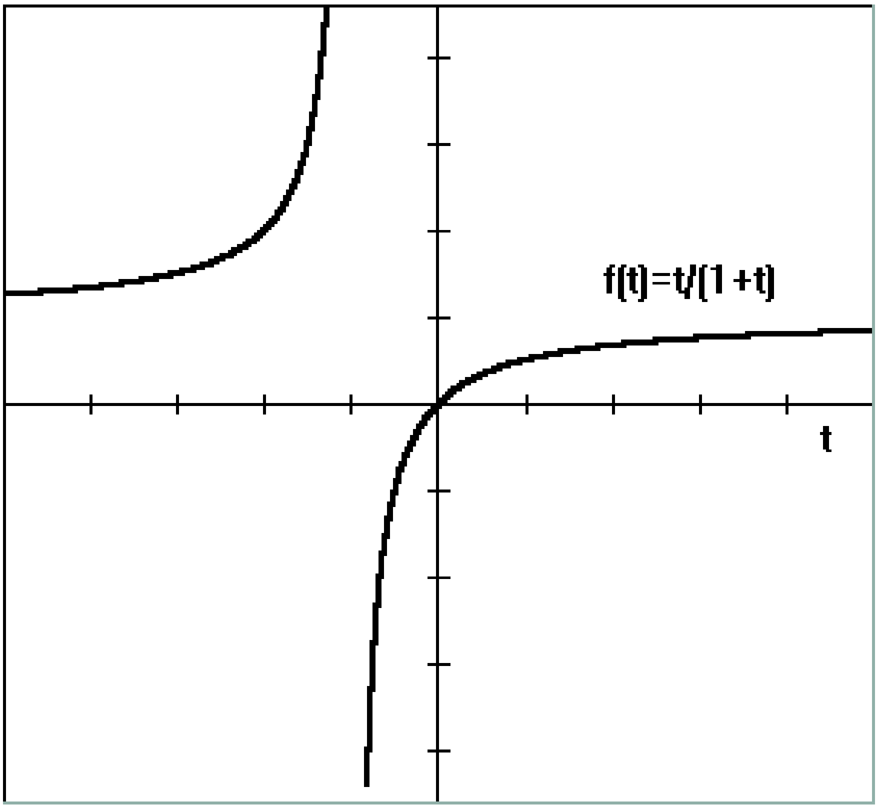
Algebraic Proof:

$$\frac{a}{1+a}-\frac{b}{1+b}=\frac{a-b}{(1+a)(1+b)}>0 \text{ when } a>b>0.$$

Calculus Proof:

$$f'(t)=\frac{1}{(1+t)^2}>0 \text{ for all } t\neq -1.$$

Graphical demonstration:



Any irrational number is a limit of a sequence of rational numbers

Any rational number is a limit of a sequence of irrational numbers

Suppose x is an irrational number, with decimal expansion $x = [x] + \sum_{i=1}^{\infty} \frac{x_i}{10^i}$.

Let $x_n = [x] + \sum_{i=1}^{n} \frac{x_i}{10^i}$. x_n is the decimal expansion of x truncated after n places.

Then (x_n) is a sequence of rational numbers converging to the irrational number x.

Now suppose y is a rational number. Then $y_n = y + \frac{\sqrt{2}}{n}$ defines a sequence of irrational numbers converging to the rational number y.

An Example of an Iteration

Consider the iteration

$x_{n+1} = F(x_n) = \frac{1}{2}\left(x_n + \frac{2}{x_n}\right)$. So $F(x) = \frac{1}{2}\left(x + \frac{2}{x}\right)$, which is continuous for $x \neq 0$.

If $x > 0$ then $F(x) > 0$. So if we take $x_0 > 0$ then (x_n) is a sequence of positive numbers.

Therefore ***if*** $x_n \to x$ as $x \to \infty$ then $F(x_n) \to F(x)$ since F is continuous. Also $x_{n+1} \to x$ since (x_{n+1}) is the same as the sequence (x_n) but with the "labels" n simply shifted by 1.

Thus $x = F(x) = \frac{1}{2}\left(x + \frac{2}{x}\right)$, which gives $x^2 = 2$.

So ***if*** the sequence has a limit, that limit must be equal to $\sqrt{2}$.

Note: This does not prove that the limit x exists, only that ***if*** it does exist then it must be equal to $\sqrt{2}$. To prove that the limit does exist we prove that the sequence (x_n) is monotonic and bounded. Refer to your notes for MA201 for this.

As an illustration of this note, take $x_0 = 1$ and $x_{n+1} = -x_n$. Then the sequence thus defined is $1, -1, 1, -1, \cdots$ which has no limit. The principle used above says that ***if*** the sequence were to have a limit, it would have to satisfy $x = -x$, i.e., $x = 0$, which is clearly false.

A Note on Cauchy Sequences

Notice that the condition for a Cauchy sequence requires that terms which are arbitrarily far apart along the sequence must eventually get close. This means that when *m* and *n* are sufficiently large, x_m and x_n are close, even though *m* and *n* may be far apart.

The weaker condition, that successive terms get close, i.e. that $x_{n+1} - x_n \to 0$ as $n \to \infty$, is not the same as the Cauchy condition, and neither is it sufficient to guarantee convergence.

This can be illustrated by taking $x_n = \ln(n)$. We then have

$$x_{n+1} - x_n = \ln(n+1) - \ln(n) = \ln\left(\frac{n+1}{n}\right) = \ln\left(1 + \frac{1}{n}\right) \to \ln(1) = 0 \text{ as } n \to \infty.$$

But $x_n = \ln(n)$ does not have a limit, in fact $x_n \to \infty$ as $n \to \infty$.

Moreover $x_m - x_n = \ln(m) - \ln(n) = \ln\left(\frac{m}{n}\right)$, and if we take $m = kn$, where k is a large integer, then $x_m - x_n = \ln(k)$, which can be made as large as you like by choosing k sufficiently large.

Intervals always contain irrational numbers

Let a be a rational number and let $(a-h, a+h)$ be an interval centred on a.

Now if h is irrational then the irrational number $a+\frac{h}{2}$ is within the interval.

If h is rational then the irrational number $a+\frac{h}{\sqrt{2}}$ is within the interval.

So in both cases the interval contains and irrational number.

Further, if h is irrational then for $n = 2, 3, 4, \ldots$ the irrational number $a+\frac{h}{n}$ is within the interval.

Also if h is rational then for $n = 1, 2, 3, \ldots$ the irrational number $a+\frac{h}{n\sqrt{2}}$ is within the interval.

So no matter how small h is the interval contains infinitely many irrational numbers.

De Morgan's Laws

These are two of the laws of the algebra of sets. the basic versions are:

$$(A \cup B)^c = A^c \cap B^c; \quad (A \cap B)^c = A^c \cup B^c$$

They are identical in form with corresponding laws in Boolean algebra and in propositional logic.

In fact the laws hold for arbitrary unions and intersections, proved below.

1. To prove that $\left(\bigcup_\alpha A_\alpha\right)^c = \bigcap_\alpha (A_\alpha)^c$.

$x \in \left(\bigcup_\alpha A_\alpha\right)^c$	
$\Leftrightarrow not\left(x \in \bigcup_\alpha A_\alpha\right)$	definition of complement
$\Leftrightarrow not(\exists \alpha, x \in A_\alpha)$	definition of union
$\Leftrightarrow \forall \alpha, x \notin A_\alpha$	negation of quantifier
$\Leftrightarrow \forall \alpha, x \in (A_\alpha)^c$	definition of complement
$\Leftrightarrow x \in \bigcap_\alpha (A_\alpha)^c$	definition of intersection

2. To prove that $\left(\bigcap_\alpha A_\alpha\right)^c = \bigcup_\alpha (A_\alpha)^c$.

$x \in \left(\bigcap_\alpha A_\alpha\right)^c$	
$\Leftrightarrow not\left(x \in \bigcap_\alpha A_\alpha\right)$	definition of complement
$\Leftrightarrow not(\forall \alpha, x \in A_\alpha)$	definition of intersection
$\Leftrightarrow \exists \alpha, x \notin A_\alpha$	negation of quantifier
$\Leftrightarrow \exists \alpha, x \in (A_\alpha)^c$	definition of complement
$\Leftrightarrow x \in \bigcup_\alpha (A_\alpha)^c$	definition of union

Index notation for arbitrary collections of sets

You are probably familiar with the notation $(S_n): n = 1, 2, \cdots$ to denote a sequence of sets. The positive integer variable *n* acts as an <u>index</u> for the collection of sets.

Associated with this are notations such as $\bigcup_{n=1}^{k} S_n$ and $\bigcup_{n=1}^{\infty} S_n$ to denote finite and infinite unions respectively.

However, we might have a collection of sets which cannot be indexed by the positive integers. For example consider the set $S_x = \{(x, y) : y \in \mathbf{R}\}$. This is a line parallel to the *y*-axis through the point on the *x*-axis specified by the number *x*. this gives us a collection of lines $\{S_x : x \in \mathbf{R}\}$. We can then consider a set specified by $\bigcup_{x \in \mathbf{R}} S_x$, which is in fact the whole *x*-*y* plane. In fact a collection of sets can be indexed by the natural numbers **N**, the real numbers **R**, or indeed any set **A** if we have a set S_α defined for each $\alpha \in \mathbf{A}$. We can then use the notation $\bigcup_{\alpha \in \mathbf{A}} S_\alpha$ or simply $\bigcup_{\alpha} S_\alpha$ to denote the union of all the sets.

To show that the intersection of any collection of closed sets is a closed set, using Theorems 5 and 6.

Let $\{A_\alpha\}$ denote an arbitrary collection of closed sets, and let $A = \bigcap_\alpha A_\alpha$.

Then by De Morgan's law $A^c = \bigcup_\alpha (A_\alpha)^c$. A_α is closed and so $(A_\alpha)^c$ is open by Theorem 5. Therefore by Theorem 6 A^c is open. Finally by Theorem 5 A is closed.

Proof of Theorem 7, Corollary 1

Let P_n stand for the statement: The intersection of any collection of n open sets is an open set.

P_1 is trivially true.

Now let $A_1, A_2, \cdots, A_{n+1}$ be an arbitrary collection of $n+1$ open sets.

Let $B = A_1 \cap A_2 \cap \cdots \cap A_n$.

By the inductive hypothesis B is open.

Then $A_1 \cap A_2 \cap \cdots \cap A_n \cap A_{n+1} = B \cap A_{n+1}$, and by Theorem 7 the intersection of two open sets is open.

Hence the result by mathematical induction.

Proof of Theorem 7, Corollary 2

Let $A_1, A_2, \cdots, A_n$ by an arbitrary finite collection of closed sets, and let $A = \bigcup_{i=1}^{n} A_i$.

By De Morgan's law $A^c = \bigcap_{i=1}^{n} (A_i)^c$. By Theorem 5 each A_i^c is an open set, and by Corollary 1 to theorem 7 A^c is therefore open. So by Theorem 5 A is closed.

An infinite union of closed sets need not be closed

In example 29 $A_n = \left\{(x,y): d\big((x,y),)0,0)\big) < \frac{1}{n}\right\}$

So $A_n^c = \left\{(x,y): d\big((x,y),)0,0)\big) \geq \frac{1}{n}\right\}$ which is a closed set (its complement is an open disc centre the origin).

$\bigcup_{n=1}^{\infty} A_n^c = \mathbf{R}^2 - \{(0,0)\}$ (the plane with the origin removed) which is not a closed set.

Some Revision Notes on Countable and Uncountable Sets

The basic idea is that of a 1-1 correspondence, or bijection, between two sets. This can be defined as follows:
A function $f : A \to B$ is said to be a bijection from A to B if

(a) $f(A) = B$, i.e. $\forall b \in B, \exists a \in A, f(a) = b$.
(b) $\forall x, y \in A, f(x) = f(y) \Rightarrow x = y$.

Note that f^{-1} is a bijection from B to A.

Two sets A and B are said to be equivalent if there is a bijection between them. This defines an equivalence relation.

A set S is said to be countable (or countably infinite) if there is a bijection between S and the set **N** of natural numbers: $f : \mathbf{N} \to S$. Now we often denote $f(n)$ by a_n, and we can think of a set being countable if it can be written as a sequence in this way.

Examples

1. Consider the set E of all even positive integers. The function $f : \mathbf{N} \to E$ defined by $f(n) = 2n$ is a bijection, so that E is countable.
Note that one characterisation of an infinite set is that there exists a bijection between the set and some proper subset, as in this example.

2. Consider the set **Z** of all integers. A bijection can be exhibited diagrammatically as follows

1	2	3	4	5	6	7	8	9	10 . . .
0	1	-1	2	-2	3	-3	4	-4	5 . . .

A formula can be given as follows

$$f(n) = \begin{cases} \frac{n}{2}; & n \text{ even} \\ -\frac{n-1}{2}; & n \text{ odd} \end{cases}$$

or as a "single expression"

$$f(n) = \frac{(-1)^n}{4}\left(2n - 1 + (-1)^n\right).$$

3. Consider the set **Q** of all rational numbers. The function $f : \mathbf{N} \to \mathbf{Q}$ defined by $f(n) = n$ defines a 1-1 mapping from **N** <u>into</u> **Q**. The function $g : \mathbf{Q} \to \mathbf{N}$ defined by $f\left(\frac{p}{q}\right) = 2^p 3^q$ defines a 1-1 mapping from **Q** <u>into</u> **N**. A theorem known as the Cantor-Bernstein Theorem tells us that under these circumstances there exists a bijection between **N** and **Q**.
Some of you will have seen a diagrammatic illustration of this fact using a two-dimensional array of fractions.

The real numbers are uncountable

An infinite set of numbers which cannot be put into 1-1 correspondence with the natural numbers is said to be uncountable. We show that the real numbers form an uncountable set

The proof of this given here is known as the Cantor Diagonal Argument. It is a proof by contradiction.

The precise form in which we establish the result is to prove that is f is any function from **N** to **R** which is 1-1 then it cannot be onto. In other words, given and arbitrary 1-1 function $f:\mathbf{N}\to\mathbf{R}$ then $\exists x\in\mathbf{R},\forall n\in\mathbf{N}, f(n)\neq x$.

Given an arbitrary such f we shall write the real number $f(n)$ in the form of its unique non-terminating decimal expansion. So we have

$$\begin{aligned} f(1) &= n_1.a_{11}a_{12}a_{13}a_{14}\ldots \\ f(2) &= n_2.a_{21}a_{22}a_{23}a_{24}\ldots \\ f(3) &= n_3.a_{31}a_{32}a_{33}a_{34}\ldots \\ f(4) &= n_4.a_{41}a_{42}a_{43}a_{44}\ldots \\ &\vdots \end{aligned}$$

We now define a sequence of integers (x_n) as follows:

$$x_n = \begin{cases} 3 \text{ if } a_{nn}\neq 3 \\ 5 \text{ if } a_{nn}=3 \end{cases}$$

We now define a real number x as follows:

$$x = 0.x_1x_2x_3x_4\ldots$$

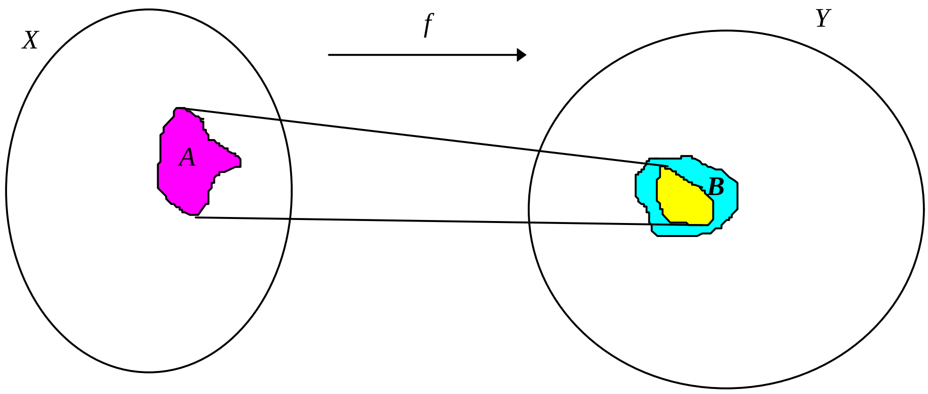
Because of the uniqueness of non-terminating decimal expansions we have $\forall n\in\mathbf{N}, f(n)\neq x$.

Thus f cannot be a bijection, and so **R** is uncountable.

A sort of generalisation of this is that if S is an arbitrary set, then there cannot be a bijection between S and the set T of all subsets of S. In some sense T is “infinitely larger” than S.

Some notes on inverse images

Given a function $f : X \to Y$ and a subset B of X we can think of the inverse image of B as the set of all x in X which map <u>into</u> B. The diagram gives an illustration of this idea, and the examples which follow help to clarify it.



The image of A is a subset of B.

Examples

We shall use functions $f : \mathbf{R} \to \mathbf{R}$ as illustrations. In the table which follows there are several cells with ???. You should try to fill them in yourselves and then make up some more.

$f(x)$	A	$f(A)$	B	$f^{-1}(B)$
$2x+3$	[1,2]	[5,7]	[1,5]	[0,1]
x^2	[-1,2)	[0,4)	(1,4)	$(-2,-1)\cup(1,2)$
x^2	(-3,2]	[0,9)	(-1,0)	$\varnothing$
$\ln x$	(1,2)	$(\ln 1, \ln 2)$	[2,3]	???
$\sin x$	$[-10\pi, 10\pi]$	[-1,1]	{0}	$\{n\pi : n \in \mathbf{Z}\}$
$\sin x$	$(-6\pi, 6\pi)$	???	{-1,1}	???
x^4	(-1,2]	???	[-1,1]	???