University of New England

School of Science and Technology

MATH101

ALGEBRA AND DIFFERENTIAL CALCULUS

Lecture Notes Part 1

Trimester 1, 2015

©University of New England

CRICOS Provider No: 00003G

Contents

Preface

Mathematics today is a vast enterprise. Advances and breakthroughs have been painstakingly built on the structure(s) erected by earlier mathematicians. The history of mathematics is quite different from the that of other human endeavours. In other fields, previously held views are typically extended or proved wrong with each advance there is a process of correction and extension. "Only in mathematics is there no significant correction – only extension".

The work of Euclid has certainly been extended many times. Euclid, however, has not been corrected – his theorems are valid today and for all time! The other remarkable thing about mathematics is its extraordinary utility in describing and quantifying the world around us. Mathematics is the language of the sciences, both natural and social. This forces mathematics to be abstract, since it must embrace theories from physics, economics, chemistry, psychology, etc. Mathematics is so widely applicable precisely because of — not despite — its intrinsic abstractness.

MATH101 is the first half of the MATH101/102 sequence, which lays the foundation for all further study and application of mathematics and statistics, presenting an introduction to differential calculus, integral calculus, algebra, differential equations and statistics, providing sound mathematical foundations for further studies not only in mathematics and statistics, but also in the natural and social sciences.

Achieving this, requires a brief, preliminary foray into the basics of mathematics, because much of the material requires a high degree of abstract reasoning, rather than rote learning of computational techniques.

A rigorous approach to the basics provides a deeper understanding of the whole structure. The assumptions upon which the structure is built are thereby clarifed, with both the scope and limitations of the intellectual framework made readily understandable. Moreover, this deeper understanding, does not come at the expense of applicability. Quite the contrary!

One consequence of providing sound fundamentals is that there is considerable time devoted to matters whose importance and applicability is not immediately obvious. But such study of these fundamental areas of mathematics is also stimulating. If you enjoy puzzles here is an "intellectual game" par excellence. A game played within a rigid framework of rules, but with unlimited scope for creativity in the search for problems and the solutions to problems.

This is the first of three parts of the lecture notes which together constitute the unit material for MATH101. These notes were originally prepared by Chris Radford and have been revised by Shusen Yan and others.

The reader is invited and encouraged to point out any mistakes (s)he finds. I hope you enjoy the challenges the unit offers and that you experience a sense of achievement at the end.

Bea Bleile
UNE

Lecture 1.1 Mathematical Language and Proof

Introduction

What is research in mathematics? A mathematician would answer "proving theorems". The language and etiquette of mathematics has evolved over a long period of time. Terms such as theorem, axiom, definition and proof have a universal and well understood meaning. It is these conventions and definitions we want to examine in this lecture.

Basic Terminology

Statement: A statement is a sentence which is either true or false (according to some previously accepted criteria). Statements do not include exclamations, questions or orders. A statement cannot be true and false at the same time.

A statement is *simple* when it cannot be broken down into other statements. A statement is *compound* when it contains more than one simple statement.

⧫**Example** "I will have a BBQ and it will rain" is a compound statement consisting of two simple statements. "I will have a BBQ" and "It will rain".

Definition: A definition is a statement of the precise meaning of a word, phrase, mathematical symbol or concept.

In trying to understand a piece of mathematics it is important to have a good working understanding of the initial definitions. You need to look at examples which satisfy and do not satisfy the definitions.

Theorem: A theorem is a mathematical statement that can be proved true by a chain of logical argument based on assumptions that are given or implied in the statement of the theorem.

A theorem will give a deeper insight into the structure of a piece of mathematics.

Lemma: A lemma is a preliminary theorem read in the proof of another theorem.

Some theorems can have proofs that are long and intricate. It is useful in such cases to break the proof into intermediate steps which are separated out as lemmas which lead into the main result.

> *Corollary: A corollary is a theorem that is a natural consequence of a preceding theorem. Generally, a corollary will follow in a relatively easy and straight-forward way from the previous theorem or proposition.*

> *Proof: A proof of a proposition, theorem, lemma or corollary is a sequence of logical reasoning. It is based on the given assumptions or hypothesis and aims to establish the truth of the proposition or theorem.*

Basic Techniques of Proof

Implication

Consider the following compound statement: "If I pass MATH101 then I will do MATH 102". Under what circumstances is this statement true or false?

Let's have a look at the two simple statements making up this compound statement.

A. I pass MATH101.

B. I will do MATH102.

For statement A there are two possibilities,

1. I **do** pass MATH101 (A is true).
2. I fail MATH101 (A is false).

There are also two possibilities for B,

1. I will enrol in MATH102 (B is true).
2. I will not enrol in MATH102 (B is false).

So we have to consider four possibilities:

	A	B
1.	True	True
2.	True	False
3.	False	True
4.	False	False

As well as the truth or otherwise of A and B, individually, we must consider the truth of the original compound statement which takes the form of an *implication*. This compound statement is "If A then B". We enlarge our table above to include a column for "If A then B". We must decide for each of the four cases if the compound entry is True or False.

Case 1: I pass MATH101 and enrol in MATH 102. The implication is True.

Case 2: I pass 101 and do not enrol in 102. Implication is False.

Case 3: I do not pass 101 but still enrol in 102. My original statement is not a lie, it is not a falsehood. Implication is True.

Case 4: I do not pass 101 and do not enrol in 102. Again, my original statement is not a lie or a falsehood. Implication is True.

	A	B	If A then B
1.	True	True	True
2.	True	False	False
3.	False	True	True
4.	False	False	True

There is only one case when the implication is false, ie. when A is true and B is false.

The implication "If A then B" is true if we can prove that it is impossible to have A true and B false at the same time.

This means if we assume "A implies B" to be true and we also assume that A is true then we must conclude that B is true.

The statement "if A then B" is equivalent to the statement "A is a sufficient condition for B" and to the statement "B is a necessary condition for A".

Notice that in a statement of the form "A implies B" the hypothesis, part A, is clearly distinguished from the conclusion, part B. A *direct proof* of a mathematical implication, "A implies B", proceeds on the assumption that the hypothesis A is true. The proof will work via a series of logically connected steps to obtain the conclusion B.

♦**Example** *Prove that the sum of two prime numbers larger than 2 is an even number.*

Solution We formulate this as an implication:

"If p and q are prime numbers larger than 2 then $p+q$ is even".

The hypothesis "p and q are prime numbers larger than 2" assumes that we know what prime numbers are. A prime number is a natural or counting number divisible only by itself and one – 2, 3, 5, 7, 11, 13, 17, 19, 23 are prime numbers.

So we are to assume p and q are prime numbers bigger than 2. We must then construct a series of arguments which lead directly to the conclusion that $p+q$ is divisible by 2 (i.e. $p+q$ is even).

Firstly, we see that p and q must be odd. We must be careful about this point because 2 is prime but even! However, for primes greater than 2 we note that a prime cannot be divisible by 2; so it must have remainder 1 after division by 2. So there must be counting numbers n and m such that

$$\begin{aligned} p &= 2n+1 \\ q &= 2m+1 \end{aligned}$$

This really just says that p and q are odd. We can now move to our conclusion,

$$\begin{aligned} p+q &= (2n+1)+(2m+1) \\ &= 2n+2m+2 \\ &= 2(n+m+1). \end{aligned}$$

We conclude that $p+q$ is divisible by 2 and so $p+q$ is even. □

This is a direct proof as we moved directly from the (assumed true) hypothesis, "p and q are prime numbers larger than 2" directly to the conclusion "$p+q$ is an even number".

Proof by Contradiction

Recall that the implication "If A then B" can only be false if we have A true and B false. So if we were to assume B to be false and *prove*, as a consequence, that A is necessarily false, then we have proved the implication is true. We would have shown that the only possible case with B false is A false which *contradicts* the assumed truth of the hypothesis. This is proof by contradiction.

In general one tries to avoid using a proof by contradiction. If a direct proof is straightforward then this is to be preferred – a direct proof usually provides more insight into the mathematical structure at hand.

♦**Example** *Prove that if p is a prime number larger than 2 then $p+1$ is not prime.*

Solution We formulate this statement as an implication with hypothesis "p is a prime number larger than 2" and conclusion "$p+1$ is not prime".

We will use a proof by contradiction. Assume the conclusion is false. We need the negative of the statement "$p+1$ is not prime". This means we must assume that "$p+1$ *is* prime".

So our starting assumptions are now

"p is a prime number larger than 2" (hypothesis)

"$p+1$ is prime" (negative of the conclusion).

We are now free to use any means at our disposal to find a contradiction based on these two assumptions. If we can find such a contradiction then we have established the truth of the original implication.

We will use the result established in the previous example.

We have two primes p and $p+1$ both larger than 2, so we know that $p+(p+1)$ must be even (this is the result of our earlier exercise). But this is clearly false as $p+(p+1) = 2p+1$ is an odd number.

We have a contradiction. Our proof is complete, our assumption that the conclusion is false cannot hold, the conclusion must be true. □

♠ Exercises 1

1. Construct a direct proof for the last example, above.

2. Use a direct proof technique to prove that if $(a+b)^2 = a^2 + b^2$ for all real numbers b, then $a = 0$.

*3. Let n be a counting number. If $2^n - 1$ is a prime number prove that n is also prime. Use a proof by contradiction to establish the truth of this statement.

Lecture 1.2 Important Types of Theorems and Proof

In your mathematical reading you will find that there are certain types of theorem whose statement and proof follow a standard pattern. Our aim here is to look at a couple of the more important examples.

Equivalence or "If and Only If" Theorems.

The statement of theorems of this type takes one of the following possible forms

"A is true if and only if B is true"
"A is false if and only if B is false"
"A is equivalent to B".

Notice that these forms are all basically equivalent – the truth (or falsity) of A automatically implies the truth (or falsity) of B and vice versa. So to prove such a theorem we have to produce *two* proofs:

one proof for "A implies B" and
one proof for "B implies A".

The first proof says A is sufficient for B (or B is necessary for A). The second proof says B is sufficient for A (or A is necessary for B). In fact another common statement of equivalence is a theorem which takes the following form, "A is a necessary and sufficient condition for B".

♦**Example** *Let n be a counting number. Then n is odd if and only if n^2 is odd.*

Solution Let A be the statement "n is odd" and let B be the statement "n^2 is odd".

We have to prove two implications

1. A implies B
2. B implies A

Firstly, let's examine A implies B. We'll provide a *direct* proof. So we assume A, that is, we assume n is an odd counting number. This means there is a counting number r such that

$$n = 2r + 1$$

(any odd number is one plus an even number). Then we have

$$\begin{aligned} n^2 = (2r+1)^2 &= 4r^2 + 4r + 1 \\ &= 4(r^2 + r) + 1, \end{aligned}$$

which is again an odd number. The first implication has been proved.

We must now tackle "B implies A". This is not easy to do using a direct proof. Try it, you need to assume n^2 is odd and then prove n is odd. Here we will use a proof by contradiction. The conclusion of the implication is that n is odd, so we will *assume* that n is even in order to produce a contradiction to our assumed hypothesis (that n^2 is odd). That is, we assume n^2 is odd (hypothesis) and that n is even and show that this leads to a contradiction.

Well, if n is even there is a counting number m such that $n = 2m$. In which case we have

$$\begin{aligned} n^2 &= (2m)^2 \\ &= 4m^2. \end{aligned}$$

But his means n^2 is even. We have a contradiction with the hypothesis that n^2 was odd. We have a proof by contradiction.

We conclude that A is equivalent to B. $\square$

The way we have written the above proof contains much more verbiage than we would normally expect in a mathematical proof. We needed it to explain what was going on! So let's rewrite the above theorem and proof in a more economical and "cleaner" style.

Theorem *Let n be a counting number. Then n is odd if and only if n^2 is odd.*

Proof First, we assume n is odd. Then $n = 2r + 1$, for some counting number r.

We have

$$\begin{aligned} n^2 &= (2r+1)^2 \\ &= 4r^2 + 4r + 1 \\ &= 4(r^2 + r) + 1, \end{aligned}$$

so that n^2 is odd, as required.

Next, assume n^2 is odd. To prove that n is odd using a proof by contradiction, assume n is even. Then $n = 2m$, for some counting number m.

Thus

$$\begin{aligned} n^2 &= (2m)^2 \\ &= 4m^2, \end{aligned}$$

so that n^2 is even. This contradicts our hypothesis. We conclude that n is odd, completing the proof. ■

This "if and only if" type of theorem need not be restricted to the equivalence of just two statements. A theorem might state the equivalence between three (or more) statements. Something like

"The following statements are equivalent"

1. A
2. B
3. C.

might be found in the mathematical literature. In such a case it is not necessary to prove directly that each statement is equivalent to each of the others. For example, it would suffice to prove the following three implications "A implies B", "B implies C" and "C implies A". Why is this true?

Mathematical Induction

Mathematical induction is used to prove theorems that would otherwise require the checking of a large (usually infinite) number of cases. For example, consider the following claim for any counting number n,

$$1 + 2 + 3 + \ldots + n = \frac{1}{2}n(n+1).$$

It is certainly easy to check the veracity of this statement for, say, the first ten counting numbers. After that it gets a bit tiresome! For example for $n = 6$ we have

$$\begin{aligned} &\text{L.H.S.} &=\;& 1 + 2 + 3 + 4 + 5 + 6 = 21 \\ \text{and}\quad &\text{R.H.S.} &=\;& \tfrac{1}{2}6(6+1) = 3 \times 7 = 21. \end{aligned}$$

Here, L.H.S. is an abbreviation for left hand side of the equation and R.H.S. an abbreviation for right hand side of the equation.

Our statement says the equation is true for all counting numbers n. Just because we have checked a few examples does not mean we can conclude that the general case is true. We cannot go from the particular to the general.

The usual way in which statements like this are proved is via the *principle of mathematical induction*. This technique involves three steps

1. Verify that the statement is true for the smallest number that can be used in the statement.
2. Assume that the statement is true for an arbitrary counting number k. This is called the inductive hypothesis.
3. Using 1 and 2 prove that the statement is true for the next counting number $k+1$. This is called the deductive step.

Suppose that we wish to prove that the statement $P(n)$ is true for all counting numbers n. For example $P(n)$ might be the statement

$$1+2+3+\ldots+n=\frac{1}{2}n(n+1),$$

of our example above. Our three step program would now read

> *Mathematical Induction.*
>
> 1. *Verify that $P(1)$ is true (we have assumed $n=1$ is the lowest value of n for which $P(n)$ makes sense).*
> 2. *Assume that $P(k)$ is true.*
> 3. *Use the assumed truth of $P(k)$ and the verified truth of $P(1)$ to prove $P(k+1)$ is true.*

If we can complete these three steps then we have proved $P(n)$ for all counting numbers n. To see why this is true notice that steps 2 and 3 show the following: If $P(k)$ is true then $P(k+1)$ is also true. Well, we know that $P(1)$ is true (step 1), so $P(2)$ must be true. But if $P(2)$ is true then $P(3)$ must be true. And so on 'ad infinitum'!

We can now finally prove the statement of our working example, that the sum of the first n counting numbers is $\frac{1}{2}n(1+n)$.

Theorem *For any counting number n*

$$1+2+3+\ldots+n=\frac{1}{2}n(1+n).$$

Proof Let $P(n)$ be the statement

$$1+2+3+\ldots+n=\frac{1}{2}n(1+n).$$

We verify that $P(1)$ is true.

L.H.S. $P(1) = 1$ and R.H.S. $P(1) = \frac{1}{2} \cdot 1 \cdot (1+1) = 1$ so $P(1)$ is true.

Now assume $P(k)$ is true (inductive hypothesis). That is, assume

$$P(k): \qquad 1 + 2 + 3 + \ldots + k = \frac{1}{2}k(1+k).$$

Next we prove $P(k+1)$ is true using the assumed truth of $P(k)$. We have

$$\begin{aligned} \text{L.H.S. } P(k+1) &= 1 + 2 + 3 + \ldots + k + (k+1) \\ &= \frac{1}{2}k(1+k) + (k+1), \text{ using } P(k). \end{aligned}$$

And,

$$\begin{aligned} \text{R.H.S. } P(k+1) &= \frac{1}{2}(k+1)(1+(k+1)) \\ &= \frac{1}{2}(k+1)(k+2) \\ &= \frac{1}{2}(k+1)k + \frac{1}{2}(k+1)2 \\ &= \frac{1}{2}k(k+1) + (k+1). \end{aligned}$$

Thus we conclude

$$\text{L.H.S. } P(k+1) = \text{ R.H.S. } \; P(k+1),$$

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

Alternatively, we can rewrite the L.H.S. of $P(k+1)$ until we obtain the R.H.S. using the assumed truth of $P(k)$:

$$\begin{aligned} &1 + 2 + 3 + \ldots + k + (k+1) \\ &= \frac{1}{2}k(1+k) + (k+1) \quad \text{using } P(k) \\ &= \frac{1}{2}k(k+1) + \frac{1}{2}(2(k+1)) \quad \text{as } \frac{1}{2} \times 2 = 1 \\ &= \frac{1}{2}(k(k+1) + 2(k+1)) \quad \text{by factorising} \\ &= \frac{1}{2}((k+2)(k+1)) \quad \text{again by factorising.} \end{aligned}$$

We have established, by the principle of mathematical induction, that $P(n)$ is true for all counting numbers n. ∎

Theorem *If n is a counting number greater than 9, then $2^n > n^3$.*

Proof Let $P(n)$ be the statement $2^n > n^3$.

Now $n > 9$, so we must verify $P(10)$. We have

$$\begin{aligned} \text{L.H.S. } P(10) &= 2^{10} = 1024 \text{ and} \\ \text{R.H.S. } P(10) &= 10^3 = 1000. \end{aligned}$$

We have established the truth of $P(10)$.

Next, assume that $P(k)$ is true, that is

$$P(k), \qquad 2^k > k^3.$$

We now prove $P(k+1)$. We have

$$\begin{aligned} \text{L.H.S. } P(k+1) &= 2^{k+1} \\ &= 2^k \cdot 2 > 2k^3, \text{ using} P(k). \end{aligned}$$

And,

$$\begin{aligned} \text{R.H.S. } P(k+1) &= (k+1)^3 \\ &= k^3 + 3k^2 + 3k + 1. \end{aligned}$$

We note that $3k^2 + 3k + 1 < 3k^2 + 3k^2 + k^2 = 7k^2$, for $k > 3$. So if $k > 9$, we have $k^3 > 7k^2 > 3k^2 + 3k + 1$.

We can thus conclude

$$\text{L.H.S. } P(k+1) = 2^{k+1} > 2k^3 > k^3 + 3k^2 + 3k + 1 = \text{ R.H.S. } P(k+1).$$

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

So by the principle of mathematical induction, $P(n)$ is true for all counting numbers n. ■

♠ Exercises 2

1. Let m and n be counting numbers. Prove that the following statements are equivalent:

   (a) $m > n$.

   (b) $a^m > a^n$, for all real numbers $a > 1$.

   (c) $a^m < a^n$, for all positive real numbers $a < 1$.

2 For all counting numbers n prove that

$$1 + 2 + 2^2 + 2^3 + \ldots + 2^{n-1} = 2^n - 1.$$

3. For all counting numbers $n \geq 1$ prove that

$$1^3 + 2^3 + 3^3 + \ldots + n^3 = \left[\frac{1}{2}n(n+1)\right]^2.$$

4. Let a be a real number with $a > 0$. Prove for any counting number $n > 1$ that $(1+a)^n > 1 + na$.

5*. For all counting numbers $n \geq 1$ prove that $3^{2n} - 1$ is divisible by 8.

Lecture 1.3 Sets and Functions

Basic Terminology

We do not attempt to define the word *set*. Intuitively a set is a *collection, aggregate* or *ensemble* of objects which are called the *elements* or *members* of the set. The set is completely determined by a knowledge of which objects are elements of it.

Capital letters will be used to denote sets and lower case letters will denote elements of the sets. The symbol $\in$ is used to denote membership of a set. So

$$x \in S$$

reads "x is an element of the set S." The statement "x is not an element of S" is abbreviated to

$$x \notin S.$$

Note that we do not exclude the possibility that x is a set in its own right, except that x cannot be S: **We explicitly exclude** $S \in S$.

The statement that a set is completely determined by its elements can be thought of in terms of the following equivalence.

> *If X and Y are sets then $X = Y$ if and only if for all $x \in X$ we have $x \in Y$ and for all $y \in Y$ we have $y \in X$.*

A set may be given by a property which defines the elements of the set. We write

$$S = \{x : \quad \text{“statement involving } x\text{”}\},$$

which reads S is the set of all x such that "statement involving x" holds. For example,

$$\begin{aligned} \mathbb{N} &= \{x : x \text{ is a natural number}\} \\ &= \{0, 1, 2, 3, 4, 5, \ldots\} \end{aligned}$$

is the set of natural numbers. The set of even natural numbers could be written as

$$\{2n : n \in \mathbb{N}\} \quad \text{or} \quad \{m \in \mathbb{N} : m = 2n, \ \text{ for } \quad n \in \mathbb{N}\}.$$

If X and Y are sets and every element of X is also an element of Y, then X is a *subset* of Y. This is denoted symbolically by

$$X \subset Y, \quad \text{or} \quad Y \supset X.$$

So $X \subset Y$ is shorthand for the statement "if $x \in X$ then $x \in Y$". We see that $X = Y$ is equivalent to the two statements $X \subset Y$ and $Y \subset X$. The negation of the statement $X \subset Y$ is written

$$X \not\subset Y,$$

and read as "X is not a subset of Y". A set X is called a *proper* subset of another set Y if $X \subset Y$ but $X \neq Y$.

The *empty set* is the set with no elements. It is denoted by the symbol ϕ. Note that $\phi \subset X$, for any set X.

Operations on Sets

The *intersection* of two sets X and Y, denoted by $X \cap Y$, is defined to be the set of all objects which are elements of both X and Y. In symbols,

$$\boxed{X \cap Y = \{x : x \in X \ \textit{and} \ x \in Y \ \}.}$$

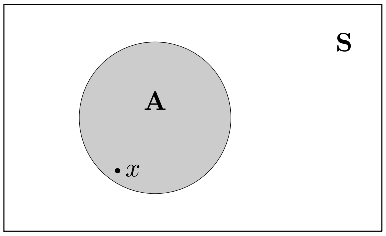
The *union* of sets X and Y is the set of all objects which are elements of X or Y. Note that the "or" is inclusive, that is, x is contained in the union of X and Y if and only if it is in X, in Yor both in X and in Y. Symbolically,

$$\boxed{X \cup Y = \{x : x \in X \ \textit{or} \ x \in Y\}.}$$

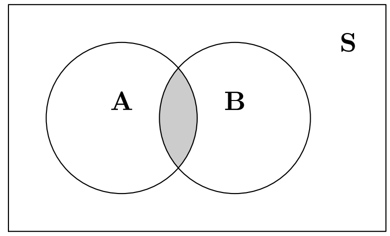
If X is a subset of a set S, then the *complement* of X in S is the set of elements of S which are not in X. If it is understood or explicitly stated in the context what the set S is, then the words "in S" are often omitted. In probability theory S is often understood as "the universal set". We denote the complement of X (in S) by $\bar{X}$,

$$\bar{X} = \{x \in S : x \notin X\}.$$

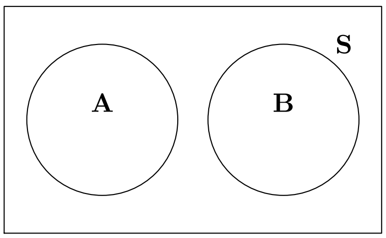
Venn diagrams provide a convenient diagrammatic way of representing the various operations and relations on sets.



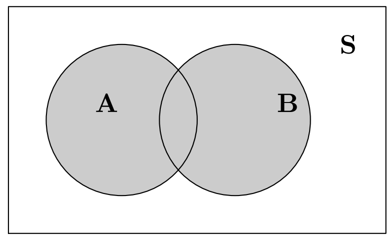
$x \in A$



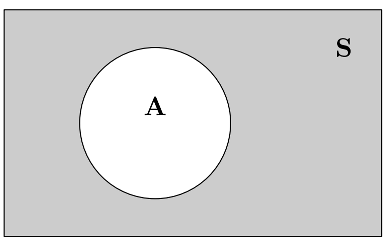
$A \cap B \neq \phi$



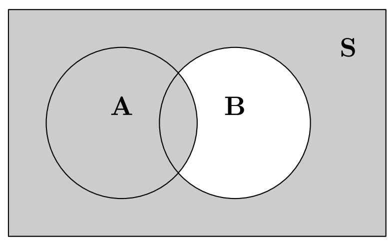
$A \cap B = \phi$



$A \cup B$



$\bar{A}$ (in S)



$A \cup \bar{B}$

Such diagrams provide a simple way of demonstrating the equality of sets. For example, draw the diagrams for $\overline{A \cup B}$ and for $\overline{A} \cap \overline{B}$ to discover that $\overline{A \cup B} = \overline{A} \cap \overline{B}$ (one of the De Morgan laws).

Given two sets X and Y we define their *Cartesian product*, $X \times Y$, to be the set of all ordered pairs the first member of which is in X and the second member of which is in Y.

$$\boxed{X \times Y = \{(x, y) : x \in X, y \in Y\}.}$$

The usual pictorial example of the Cartesian product is that of rectangular coordinates in the plane. The whole plane is identified as the Cartesian product of the two axes.

♦**Example** Find, explicitly, the Cartesian product of the two sets $X = \{1, 7, 9, 21, 66\}$ and $Y = \{2, 3\}$.

Solution The Cartesian product, $X \times Y$, consists of all possible ordered pairs (x, y), where $x \in X$ and $y \in Y$. Hence

$$X \times Y = \{(1,2),\ (1,3),\ (7,2),\ (7,3),\ (9,2),\ (9,3),\ (21,2),\ (21,3),\ (66,2),\ (66,3)\}$$

Notice in this example we can explicitly write down each element of $X \times Y$. We can do this here because the sets X and Y are finite. Of course this is not possible if the sets are infinite. For example, the set of lattice points in the first quadrant of the $x - y$ plane is given by $\mathbb{N} \times \mathbb{N}$ – an infinite set of points (n, m) where $n \in \mathbb{N}$ and $m \in \mathbb{N}$.

Functions

If X and Y are sets a *function* from X to Y (or a function from X into Y, or a function with values in Y) is a rule which associates with each element of X a unique element of Y. The words *mapping* or *map* are often used instead of *function.*

The function "rule" can be specified in many ways, but the important thing is that to any given element of X we associate, somehow, a unique element of Y. We indicate that we have a function or map between sets X and Y by the arrow notation,

$$f : X \longrightarrow Y.$$

For any $x \in X$ the element y of Y associated with x by the function f is denoted $f(x)$. Then $f(x) \in Y$ and we write

$$y = f(x) \quad \text{or} \quad \mathrm{x} \longmapsto \mathrm{y}.$$

This is a more general concept of a function than you have probably met before. The familiar functions of calculus are functions from a subset, D, of the real numbers to the real numbers. Denoting the set of real numbers by $\mathbb{R}$, we obtain

$$f : D \longrightarrow \mathbb{R}.$$

♦**Example** We may use the formula $f(x) = x^2 + x + 1$ to define the function

$$f : \mathbb{R} \longrightarrow \mathbb{R}; \quad x \longmapsto x^2 + x + 1.$$

The *graph* of such a real valued function on D is the set of points (x, y) in the plane such that $y = f(x)$ for $x \in D$. In fact, this idea of graph is easy to generalise to our more abstract setting.

> *The graph of a function $f : X \longrightarrow Y$ is the set of ordered pairs $\{(x, y) \in X \times Y : y = f(x)\}$ or $\{(x, f(x)) \in X \times Y : x \in X\}$*

Another class of functions with which you should be familiar are the *sequences.* A sequence $\{a_n\}$ where $n \in \mathbb{N}$ can be thought of as a function

$$f : \mathbb{N} \longrightarrow \mathbb{R}.$$

This means for each $n \in \mathbb{N}$ the function f associates a real number $f(n)$, which by tradition we denote by a symbol such as a_n. In fact all your familiar functions from calculus (quadratics, exponentials and so on) can be made into sequences if we *restrict* ourselves to the subset $\mathbb{N}$ of $\mathbb{R}$. For example, consider the quadratic

$$f : \mathbb{R} \longrightarrow \mathbb{R}, \text{ with } f(x) = x^2.$$

We may *restrict* this function to the subset of natural numbers, that is, we think of f as acting only on the subset $\mathbb{N}$ of $\mathbb{R}$. This is usually written as $f|_{\mathbb{N}}$, and read as "f restricted to $\mathbb{N}$". The function $f|_{\mathbb{N}}$ has exactly the same function rule as f, it simply acts only on the subset $\mathbb{N}$ of $\mathbb{R}$. This all sounds a bit abstract, but in practice it's very simple, we have

$$f|_{\mathbb{N}} : \mathbb{N} \longrightarrow \mathbb{R}, \text{ with } f|_{\mathbb{N}}(n) = n^2.$$

Similarly, restricting any function $\mathbb{R} \longrightarrow \mathbb{R}$ we obtain a sequence. Our example, $f(x) = x^2$, defines the sequence $\{n^2\}_{n\in\mathbb{N}} = \{1, 2^2, 3^2, 4^2, \ldots\}$.

We will take a more detailed look at sequences later in these notes.

In our general setting with functions $f : X \longrightarrow Y$ we make the following definitions

> *The set X is called the domain and the set Y is called the codomain of the function f. The set $f(X) = \{f(x) : x \in X\}$ is called the range or image of the function.*

We should emphasise the importance of the domain in the definition of a function – functions with different domains are *different* functions. Consider the following two functions,

$$g : \{x \in \mathbb{R} : x \geq 0\} \to \{y \in \mathbb{R} : y \geq 0\}, \quad x \longmapsto x^2,$$
$$f : \mathbb{R} \to \mathbb{R}, \quad x \longmapsto x^2.$$

They are given by the same formula, that is, $f(x) = x^2$ and $g(x) = x^2$, but they are quite different functions – they have different domains. They also have different properties, for example, g is one-to-one (see lecture 8), whereas f is not.

♦**Example** What is the largest subset, D of $\mathbb{R}$, such that

$$f : D \longrightarrow \mathbb{R}, \quad x \longmapsto \sqrt{1 - x^2}$$

is a function?

Solution As $\sqrt{u}$ is defined if and only if $u \geq 0$, we must have $1 - x^2 \geq 0$, that is, $-1 \leq x \leq 1$. Thus $D = \{x : -1 \leq x \leq 1\}$.

Suppose we have two functions $f : X \longrightarrow Y$ and $g : Y \longrightarrow Z$, where X, Y and Z are sets. Is there a natural way to go from X to Z using f and g? Well, we could simply use f to take us from $x \in X$ to $y = f(x) \in Y$ and then use g to take us from y to $g(y) \in Z$. This gives us a new function $X \longrightarrow Z$, the *composition* of f and g.

For functions $f : X \longrightarrow Y$ and $g : Y \longrightarrow Z$ the composition of f and g, denoted $g \circ f$, is defined to be the function

$$g \circ f : X \longrightarrow Z$$

$$g \circ f(x) = g(f(x)), \quad \textit{for} \quad x \in X.$$

Notice the order of f and g in $g \circ f$ – we use the function f first and then we use g. The expression $g(f(x))$ simply means $g(y)$ where $y = f(x)$.

♦**Example** Find $g \circ f(x)$ and $f \circ g(x)$ where $f, g : \mathbb{R} \longrightarrow \mathbb{R}$ are defined by

$$f(x) = x^2 \quad \text{and} \quad g(x) = x^2 + 1.$$

Solution

$$\begin{aligned} g \circ f\,(x) &= g(f(x)) \\ &= g(x^2) \\ &= (x^2)^2 + 1 \\ &= x^4 + 1 \end{aligned}$$

Perhaps this is a little easier to see if we introduce the intermediate step, $y = f(x) = x^2$, then

$$\begin{aligned} g \circ f(x) = g(f(x)) &= g(y) \\ &= y^2 + 1 \\ &= x^4 + 1, \quad \text{as} \quad y = x^2. \end{aligned}$$

Furthermore,

$$\begin{aligned} f \circ g\,(x) &= f(g(x)) \\ &= f(x^2 + 1) \\ &= (x^2 + 1)^2 \\ &= x^4 + 2x^2 + 1 \end{aligned}$$

This shows that, in general, $f \circ g$ is not equal to $g \circ f$. □

♦**Example** For functions $h : \mathbb{N} \longrightarrow \mathbb{R}$ and $g : \mathbb{R} \longrightarrow \mathbb{R}$, where $h(n) = \frac{1}{n}$ and $g(x) = \frac{1}{x^2+1}$, find the equation for $g \circ h(n)$.

Solution

$$\begin{aligned} g \circ h(n) &= g(h(n)) \\ &= g(\frac{1}{n}) \\ &= \frac{1}{(\frac{1}{n})^2 + 1} \\ &= \frac{n^2}{1 + n^2} \end{aligned}$$

□

♠ Exercises 3

1. If $A \subset S$ show that

   (a) $\overline{(\overline{A})} = A$

   (b) $A \cup A = A \cap A = A \cup \phi = A$

2. Use a Venn diagram to demonstrate the truth of the second De Morgan law

$$\overline{A \cap B} = \overline{A} \cup \overline{B}.$$

3. Consider the functions $f : D_f \longrightarrow \mathbb{R}$ and $g : D_g \longrightarrow \mathbb{R}$ given by $f(x) = \sqrt{x^2 - 4}$ and $g(x) = x^2 + 4$.

   (a) What are the largest subsets D_f and D_g of $\mathbb{R}$, such that f and g are functions?

   (b) What are the ranges (images) of f and g?

   (c) Find the equation defining $g \circ f$. What are the domain, codomain and and range of $g \circ f$?

4. Let $f : X \longrightarrow Y$ be a function, let A and B be subsets of X. Prove that

   (a) $f(A \cup B) = f(A) \cup f(B)$.

   (b) $f(A \cap B) \subset f(A) \cap f(B)$.

Lecture 1.4 Numbers

The awareness of numbers, particularly the counting numbers, is in some ways a fairly primitive sensibility – certain birds can distinguish sets containing up to four elements. But it is this concept of number which is at the heart of the origins of mathematics.

Number theory, even today, offers some of the deepest and most challenging problems in mathematics. Our aim over the next couple of lectures will be quite modest. We want to give a brief introduction to those ideas from the theory of number which are required for a rigorous look at the differential calculus.

Integers

We will accept as given the counting numbers $\{1, 2, 3, 4, \ldots\}$ and the natural numbers $\mathbb{N} = \{1, 2, 3, 4, \ldots\}$, stressing only the facts needed for extensions of this number system.

Elements of $\mathbb{N}$ can be added or multiplied to give further elements of $\mathbb{N}$. If $m, n \in \mathbb{N}$ then $m + n \in \mathbb{N}$ and $mn \in \mathbb{N}$. This is called *closure* under addition or multiplication. There is a number 0 for which $0 + m = m + 0 = m$ and there is a number 1 for which $1.m = m.1 = m$, for all $m \in \mathbb{N}$. We call 0 the neutral element with respect to addition and 1 the neutral element with respect to multiplication. We also have an *order* expressed by the symbols $<$ (less than),$>$ (greater than).

To deal, in a more complete fashion, with equations

$$m + x = n; \quad m, n \in \mathbb{N}$$

we will need to introduce more numbers. Notice this equation can only be solved in $\mathbb{N}$ (i.e. with $x \in \mathbb{N}$) if $n > m$. We widen our number system to include the "negatives" of the counting numbers.

The integers $\mathbb{Z} = \{\ldots, -3, -2, -1, 0, 1, 2, 3, \ldots\}$.

Notice that the properties of closure under addition and multiplication still hold for $\mathbb{Z}$, as does the ordering property.

If a and b are integers the equation

$$ax = b.$$

is not, in general, satisfied by an integral value of x. If this equation is to always have a solution we need to again widen our number system to include the *rationals* a/b (with $a, b \in \mathbb{Z},\ b \neq 0$).

Rational Numbers

The set of rational numbers, $\mathbb{Q} = \{\frac{p}{q} : q \neq 0;\ p, q \in \mathbb{Z}\}$.

In the system of rationals the usual operations of arithmetic, addition, subtraction, multiplication and division all apply. The rationals can also be ordered according to their "size" if we define

$$\frac{a}{b} > \frac{c}{d};\ a, b, c, d \in \mathbb{Z} \text{ and } b, d > 0$$

to *mean* $ad > bc$. Note the restriction b and d positive is not really a restriction at all, if $b < 0$ then we can write

$$\frac{a}{b} = \frac{-a}{-b}.$$

The denominator is then positive.

The rationals have an important property not possessed by the integers:

Theorem *Between any two rationals there are infinitely many other rationals.*

Proof The proof is a simple direct "construction".

We simply give an infinite set of rationals between $\frac{a}{b}$ and $\frac{c}{d}$; $\frac{a}{b}, \frac{c}{a} \in \mathbb{Q}$. As noted above we can always take b and d to be positive integers. Then

$$\frac{a + mc}{b + md}$$

lies between a/b and c/d for any positive integer m. There are infinitely many positive integers so we get infinitely many rationals between a/b and c/d. ■

Irrational Numbers

It was realised by the Greeks more than 2000 years ago that there is an incompleteness about the system of rational numbers. The diagonal of a square with sides of unit length has a length which is not a rational number – it is an *irrational* number. In algebraic language the equation for x,

$$x^2 = a, \quad a \in \mathbb{Q},$$

has a rational solution for x only for exceptional values of the rational number a. For our unit square Pythagoras' theorem says that the length of the diagonal, x, is given by

$$x^2 = 1^2 + 1^2 = 2.$$

We will now show that the formal solution to this equation, $x = \sqrt{2}$, cannot be a rational number. We call $\sqrt{2}$ an irrational number.

Theorem *No rational number has square 2.*

Proof This is a very old theorem having been proved first by Pythagoras or one of his school. It is a very simple and economical proof by contradiction.

Assume, on the contrary, that the rational p/q has square 2. We can also assume that p and q are integers having no common factors – if they did have common factors we would simply cancel them from the "fraction". We have

$$\begin{aligned} \frac{p^2}{q^2} &= 2, \text{ or} \\ p^2 &= 2q^2. \end{aligned}$$

We see that p^2 is divisible by 2 so p must be an even integer (see example in lecture 1.2). So we can write $p = 2r$, for some integer r. We now have

$$p^2 = 4r^2 = 2q^2,$$

from which we see that $q^2 = 2r^2$ so that q must be even. This means that p and q have the common factor 2, contradicting our hypothesis. The theorem is proved. ■

It can be quite difficult to determine whether a given number is rational or not. The fact that the number π (the ratio of the circumference of a circle to its diameter) is not rational was only discovered at the end of the 18th century.

The rationals and irrationals together give us the real numbers – all possible points on the number line.

The Real Numbers.

The irrationals fill in the "gaps" in the rationals. Given any irrational we can find rational numbers arbitrarily close to it. This process is just the decimal approximation to the infinite decimal expression for the irrational number. This is most easily seen by looking at the irrational $\sqrt{2} = 1.4142\ldots$ (an infinite decimal). We then simply look at successive decimal approximations to $\sqrt{2}$,

$$1, 1.4, 1.41, 1.414, 1.4142, \ldots.$$

Each of these numbers is rational (any finite decimal is rational, e.g. $1.414 = \frac{1414}{1000}$) and they are in increasing order. All of these numbers are less than $\sqrt{2}$. However, we can get arbitrarily close to $\sqrt{2}$.

We can also approximate $\sqrt{2}$ from above

$$2, 1.5, 1.42, 1.415, 1.4143, \ldots\ .$$

This sequence is found by adding a one to the last decimal place of our approximation from below. Notice that this latter sequence is a decreasing sequence of rationals.

The sequence approximating $\sqrt{2}$ from below has no greatest member – we simply go to the next decimal place to get a bigger member of the sequence. The sequence approximating $\sqrt{2}$ from above has no smallest member.

So the irrational number $\sqrt{2}$ can be defined by cutting the rationals into two classes L and R. Where L has no greatest member and R which has no smallest member.

$$\overbrace{\quad 0 \quad\quad 1 \quad\quad}^{L}\sqrt{2}\overbrace{\quad\quad 2 \quad\quad 3 \quad}^{R}$$

This is Dedekind's definition (1872) of the irrationals by a *cut* of the rationals.

The rationals and the reals satisfy the axioms for what is known in algebra as an *ordered field*. The reals are distinguished from the rationals by one further axiom,

> **Dedekind's Axiom.** *Suppose that the system of all real numbers is divided into two classes L and R, every member ℓ of L being less than every member r of R (neither class being empty). Then there is a dividing number ξ with the property that every number less than ξ belongs to L and every number greater than ξ belongs to R. The number ξ may belong to either L or R. If it is in L, it is the greatest member of L. If ξ is in R it is the least member of R.*

Such a division of the real numbers into two classes by means of some rule is called a *Dedekind cut*.

♠ Exercises 4

1. (a) Is $\frac{223}{71}$ greater than $\frac{22}{7}$?

   (b) Is $\frac{265}{153}$ greater than $\frac{1351}{780}$?

2. If $a, b \in \mathbb{R}$ prove, that if $a < b < 0$, then

$$\frac{1}{a} > \frac{1}{b}.$$

3. Prove that if m/n is a rational approximation to $\sqrt{2}$ from below then

$$\frac{m+2n}{m+n}$$

   is a closer approximation from above. Hence write down approximations to $\sqrt{2}$, obtaining two which differ by less than $\frac{1}{10{,}000}$.

4*. The density property of the rationals amounts to saying that there is no rational which is *next* to another. The following plan for arranging the rationals (not in order of magnitude) does assign a definite place to each

$$\frac{1}{1}, \frac{2}{1}, \frac{1}{2}, \frac{3}{1}, \frac{2}{2}, \frac{1}{3}, \frac{4}{1}, \frac{3}{2}, \frac{2}{3}, \frac{1}{4}, \frac{5}{1}, \dots .$$

   What place does $\frac{6}{1}$ occupy? (Counting from the left, of course!)

   Show that p/q occupies the $\{\frac{1}{2}(p+q-1)(p+q-2)+q\}$th place. (Notice that each rational occurs infinitely often in this scheme; e.g. 2 appears as $\frac{2}{1}, \frac{4}{2}, \frac{6}{3}, \dots$.)

Lecture 1.5 Some Properties of Real Numbers

Basic Arithmetic

We assume the familiar properties of the real numbers: For all $a, b, c \in \mathbb{R}$

- $a + b = b + a$ (commutativity of addition);
- $(a + b) + c = a + (b + c)$ (associativity of addition);
- $0 + a = a + 0 = a$ (0 is the neutral element with respect to addition);
- There exists $y \in \mathbb{R}$ such that $a + y = y + a = 0$ (existence of additive inverses);
- $ab = ba$ (commutativity of multiplication);
- $(ab)c = a(bc)$ (associativity of multiplication);
- $1a = a1 = a$ (1 is the neutral element with respect to multiplication);
- There exists $z \in \mathbb{R}$ such that $az = za = 1$ (existence of multiplicative inverses).
- $(a + b)c = ac + bc$ (distributivity).
- One and only one of
$$a > b,\ a = b,\ a < b$$
is true.
- If $a > b$ and $b > c$, then $a > c$.
- If $a > b$, then $a + c > b + c$.
- If $a > b$ and $c > 0$, then $ac > bc$.

These properties are essentially the axioms for an ordered field. If we add the Dedekind axiom of the last lecture then we have the axioms defining the real numbers.

From these basic rules or axioms all the well-known and well-used properties of real numbers follow.

For example, defining

$$a^n = \underbrace{a \cdot a \ldots a}_{n \text{ times}} \text{ and (for } a \neq 0) \ a^0 = 1, \ a^{-n} = \frac{1}{a^n},$$

we obtain the usual statements for exponentials of a

$$\begin{aligned} a^m \cdot a^n &= a^{m+n} \\ (a^m)^n &= a^{mn} \\ (ab)^n &= a^n b^n, \end{aligned}$$

with $m, n \in \mathbb{Z}$. These index rules can, in fact, be extended to $m, n \in \mathbb{R}$.

We can also impose a "distance measure" on $\mathbb{R}$, this is the *absolute value* of a real number – it measures the distance of that number from zero.

$$|a| = \begin{cases} a, & \text{if } \ a \geq 0 \\ -a, & \text{if } \ a \leq 0. \end{cases}$$

Theorem *The absolute value has the following properties:*

1. $|a| \geq 0$ *for all* $a \in \mathbb{R}$ *and* $|a| = 0$ *if and only if* $a = 0$.
2. $|ab| = |a| \cdot |b|$, *for all* $a, b \in \mathbb{R}$.
3. $|a|^2 = a^2$, *for all* $a \in \mathbb{R}$.
4. $|a + b| \leq |a| + |b|$, *for all* $a, b \in, \mathbb{R}$.
5. $|a - b| \geq ||a| - |b||$, *for all* $a, b \in \mathbb{R}$.

Proof The first three properties are trivial consequences of the definition and are left as exercises.

To prove (4) note that if $a + b \leq 0$ then

$$\begin{aligned} |a + b| &= -(a + b) \\ &= -a - b \\ &= -a + (-b) \\ &\leq |a| + |b|, \end{aligned}$$

as $-a \leq |a|$ and $-b \leq |b|$. On the other hand, if $a + b \geq 0$ then

$$\begin{aligned} |a + b| &= a + b \\ &\leq |a| + |b|. \end{aligned}$$

And we have the required property.

To prove (5) note that

$$|a| = |(a-b)+b| \leq |a-b| + |b|,$$

so that

$$|a-b| \geq |a| - |b|.$$

Similarly,

$$|b| = |(b-a)+a| \leq |b-a| + |a|,$$

so that

$$|b-a| = |a-b| \geq |b| - |a|.$$

Combine the two inequalities for $|a-b|$,

$$|a-b| \geq \pm(|a| - |b|).$$

So we have (5),

$$|a-b| \geq ||a| - |b||.$$

■

Bounded Sets of Numbers

Consider the following sets of numbers:

1. All prime numbers.
2. All integers greater than 1000 which are perfect squares.
3. All rational numbers x, with $1 \leq x \leq 2$.
4. All real numbers x such that $1 \leq x \leq 2$.
5. All real numbers x such that $1 < x < 2$.

We first observe that all these sets are *infinite*. The examples 1 and 2 might give the impression that an infinite set must contain elements which are unboundedly large. However, examples 3, 4 and 5 correct that false impression!

The sets of examples 4 and 5 are of an important type, each is called an *interval*. Example 4 in which the end points, $x = 1$ and 2, are members of the set is called a *closed interval* and denoted by $[1, 2]$. In general, the closed interval from a to b with $a \leq b$ is the set

$$[a, b] = \{x \in \mathbb{R} : a \leq x \leq b\}.$$

The set in example 5, in which the end points are excluded from the set, is called an *open interval* and denoted by $(1, 2)$. In general, the open interval from a to b with $a \leq b$ is the set

$$(a, b) = \{x \in \mathbb{R} : a < x < b\}.$$

If a set consists of finitely many numbers then clearly we can distinguish a greatest (and least) element of that set. However, if S is an infinite set of numbers there may or may not be a greatest element.

For example, in the set $S = [-1, 1]$ the number 1 is greater than any other member of S. The set S also has a least member, -1.

On the other hand the set $S = (-1, 1)$ does not have a largest member. Given any element of S we can always find another member of S which is larger. Suppose $k \in S = (-1, 1)$, then $-1 < \frac{1}{2}(k+1) < 1$ so $\frac{1}{2}(k+1) \in S$. But, $\frac{1}{2}(k+1) > k$.

What both sets do have in common is the fact that they are *bounded* above (and below). That is, there is a number K (not necessarily in the set S) such that $x \leq K$ for all $x \in S$.

Let S be a set of real numbers. If there is a number K such that for every $x \in S$, $x \leq K$, we say S is bounded above. K is an upper bound for S. *Similarly if there is a $k \in \mathbb{R}$ such that $x \geq k$ for all $x \in S$ we say S is bounded below. k is a lower bound for S.*

A set which is both bounded above and below is said to be bounded.

♦**Example** Is the set $\{\frac{1}{2}, \frac{2}{3}, \frac{3}{4}, \ldots, \frac{n}{n+1}, \ldots\}$ bounded or unbounded?

Solution The set is clearly bounded. The number $\frac{1}{2}$ serves as a lower bound – or any number less than $\frac{1}{2}$. The number 1 (or any number greater than 1) serves as an upper bound. Note that the set does have a least member, $\frac{1}{2}$. However, the set does not have a greatest member. The elements of the set form an increasing sequence which is bounded above by 1. □

If K is an upper bound of a set $S \subseteq \mathbb{R}$ then any number greater than K is also an upper bound. If we want to make the sharpest possible statement by confining S as closely as possible then we would aim at getting the *least upper bound*. That is a number K which is an upper bound, but such that $K - \varepsilon$ is not an upper bound for any positive number ε, no matter how small ε is. Or, stated another way

$$K \geq x \quad \text{for all} \quad x \in S \text{ and for any } \varepsilon > 0 \text{ there is an } s \in S \text{ such that } s > K - \varepsilon.$$

In a similar fashion we may seek a *greatest lower bound* k with the properties

$$k \leq x \ \text{ for all } \ x \in S \text{ and for any } \varepsilon > 0 \text{ there is an } s \in S \text{ such that } s < k + \varepsilon.$$

The least upper bound is usually called the supremum and we write $K =$ *sup* S*, read as* K *is the supremum of the set* S*.*

The greatest lower bound is usually called the infimum, and we write $k =$ *inf* S*.*

Using the Dedekind axiom we can prove that *a (non-empty) set of real numbers which is bounded above must have a supremum.* And in a similar way that *a nonempty set of real numbers which is bounded below must have an infimum.* We will not give the proof here – although we now certainly have the background to tackle it. As illustrations consider the sets

1. $S_1 = \{x \in \mathbb{Q} : 0 \leq x \leq \frac{1}{2}\}$
2. $S_2 = \{x \in \mathbb{Q} : x^2 < 2\}$
3. $S_3 = \{\frac{1}{2}, \frac{2}{3}, \frac{3}{4}, \ldots, \frac{n}{n+1}, \ldots\}$.

For S_1 the least upper bound is $\frac{1}{2}$, that is, sup $S_1 = \frac{1}{2}$. Note $\frac{1}{2}$ is also the greatest member of S_1.

The set S_2 has a supremum of $\sqrt{2}$, i.e. sup $S_2 = \sqrt{2}$. In this case sup $S_2 \notin S_2$.

Finally, for the set S_3 the supremum is 1, sup $S_3 = 1$. Again, sup $S_3 \notin S_3$.

We also have inf $S_1 = 0$, inf $S_2 = -\sqrt{2}$ and inf $S_3 = \frac{1}{2}$.

♠ Exercises 5

1. If $a, b \in \mathbb{R}$ and
$$a < b + \varepsilon, \quad \text{for any} \quad \varepsilon > 0,$$
prove $a \leq b$.

   [Hint: try a proof by contradiction. The negative of the statement $a \leq b$ is $a > b$.]

2. Show, for any $a, b \in \mathbb{R}$,
$$\begin{aligned}
\text{maximum } \{a, b\} &= \frac{a + b + |a - b|}{2} \\
\text{minimum } \{a, b\} &= -\text{ maximum } \{-a, -b\} \\
&= \frac{a + b - |a - b|}{2}
\end{aligned}$$

3. Find (if they exist) the supremum and infimum of the following sets
   (a) $\{1, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \ldots, \frac{1}{n}, \ldots\}$
   (b) $\{1.1, (1.1)^2, (1.1)^3, \ldots, (1.1)^n, \ldots\}$
   (c) $\{\sqrt{2}, \sqrt{\sqrt{2}}, \sqrt{\sqrt{\sqrt{2}}}, \ldots\}$

4. Given two bounded subsets $A, B \subseteq \mathbb{R}$ prove
$$\sup\ (A \cup B) = \text{maximum } \{\sup\ A, \sup B\}.$$

Lecture 1.6 Complex Numbers

Our system of real numbers $\mathbb{R}$ would appear to be perfectly adequate for mathematics and indeed most applications. We feel comfortable with the fact that $\mathbb{R}$ can be represented as the real number line every point of which is in $\mathbb{R}$. We think of rulers and other measuring devices as segments or intervals of the real number line. Why would we want to extend further our system of numbers?

Well, curiosity would be one reason. If it is possible let's try it! On the other hand there are sound mathematical reasons why we might want to extend our number system. The extension to the *complex numbers* was first carried out by Gauss in about 1795. Complex numbers are indispensable in modern mathematics and physics. In Quantum theory in particular complex numbers are used in a fundamental way.

One motivation for the extension of the number system is to note that only non-negative numbers have square roots. In all other respects there is complete symmetry between positive and negative numbers. Another way of stating this is to note that the equation

$$x^2 = a, \quad \text{with} \quad a \in \mathbb{R},$$

has a solution $x \in \mathbb{R}$ if and only if $a \geq 0$. If we write an equation such as

$$x^2 = -1$$

then $x \notin \mathbb{R}$. So there exist "numbers", satisfying perfectly reasonable equations, which are not in $\mathbb{R}$.

We extend the field of real numbers to the field of *complex numbers* by adjoining a new element, i, to the set. This element i is defined as a solution to the equation

$$i^2 = -1.$$

It is traditional to write $i = \sqrt{-1}$ as the solution; note that $x^2 = -1$ then has two possible solutions $x = \pm i$.

We also include in our set of complex numbers all numbers which can be formed from i and $\mathbb{R}$ by use of the operations $+$ and $-$. The elements of this field of complex numbers then take the form $a + ib$, where $a, b \in \mathbb{R}$.

$$\boxed{\textit{The Complex Numbers } \mathbb{C} = \{a + ib : a, b \in \mathbb{R}\}.}$$

All the axioms for multiplication and addition of the real number system listed in the last lecture apply to $\mathbb{C}$. However, the order axioms involving $<$ and $>$ no longer apply. The real numbers can be included in $\mathbb{C}$ as a subset, namely $\mathbb{R}$ can be identified with the complex numbers of the form $a + i0$. For this reason we write $a + i0 = a$.

So we may think of complex numbers as numbers of the form $a + ib$, with a and b real. We refer to the a as the *real part* and b as the *imaginary part*.

Any complex number z can be written as

$$z = a + ib, \text{ with } a, b \in \mathbb{R}.$$

$$\begin{aligned} \textit{We write Re } (z) &= a \textit{ for the real part of } z \textit{ and} \\ \textit{Im } (z) &= b \textit{ for the imaginary part of } z. \end{aligned}$$

Addition of Complex Numbers

For two complex numbers $z_1 = a_1 + ib$, and $z_2 = a_2 + ib_2$, with $a_1, a_2, b_1, b_2 \in \mathbb{R}$, we have

$$z_1 + z_2 = (a_1 + ib_1) + (a_2 + ib_2) = (a_1 + a_2) + i(b_1 + b_2).$$

Notice that

$$\begin{aligned} \text{Re } (z_1 + z_2) &= a_1 + a_2 = \text{Re}(z_1) + \text{Re}(z_2) \\ \text{and Im } (z_1 + z_2) &= b_1 + b_2 = \text{Im}(z_1) + \text{Im}(z_2). \end{aligned}$$

♦Example

(a) $(1 + 3i) + (3 + i) = 4 + 4i = 4(1 + i)$

(b) $(1 + 3i) - (3 + i) = -2 + 2i = 2(-1 + i)$

(c) $(\pi + i) - (1 + \sqrt{2}i) = (\pi - 1) + (1 - \sqrt{2})i$

(d) $\left(\frac{1}{2} + \frac{1}{3}i\right) + \left(\frac{1}{4} - \frac{1}{6}i\right) = \frac{3}{4} + \frac{1}{6}i$

Multiplication of Complex Numbers

Multiplication for elements of $\mathbb{C}$ follows the same field axioms as multiplication of real numbers. Let's start with an example.

♦**Example** What is $(1+i)\cdot(1-3i)$?

Solution

We simply expand the brackets,

$$\begin{aligned}(1+i)\cdot(1-3i) &= 1\cdot(1-3i)+i\cdot(1-3i)\\ &= 1-3i+i-i\cdot 3i\\ &= 1-3i+i-3i^2\\ &= 1-2i-3\cdot(-1),\ \text{as } i^2=-1\\ \text{i.e. } (1+i)\cdot(1-3i) &= 4-2i.\end{aligned}$$

□

Now for the general case

For complex numbers z_1 and z_2 (as above) we have $$z_1z_2=(a_1+ib_1)(a_2+ib_2)=a_1a_2-b_1b_2+i(a_1b_2+a_2b_1).$$

You should **not** bother memorising this formula, instead you simply multiply out the brackets as per our example. In the general case we have

$$\begin{aligned}z_1z_2=(a_1+ib)(a_2+ib_2) &= a_1(a_2+ib_2)+ib_1)\\ &= a_1a_2+ia_1b_2+ib_1a_2+i^2b_1b_2\\ &= a_1a_2-b_1b_2+i(a_1b_2+b_1a_2),\ \text{as } i^2=-1.\end{aligned}$$

So the policy is expand everything in sight, use $i^2=-1$ and then collect terms into two groups those with an i and those without an i. The terms without an i give the real part of the product and the coefficient of the i gives the imaginary part. So,

$$\begin{aligned}\text{Re }(z_1z_2) &= a_1a_2-b_1b_2 \quad\text{and}\\ \text{Im }(z_1z_2) &= a_1b_2+b_1a_2.\end{aligned}$$

One thing worth emphasising Im $(a+ib)=b$ is a real number, it is the coefficient of i.

♦Example

(a)
$$\begin{aligned}(3+4i)(6+i) &= 3\cdot(6+i)+4i\cdot 6+i)\\ &= 18+3i+24i+4i^2\\ &= 18+3i+24i-4\\ &= 14+27i.\end{aligned}$$

(b)
$$\begin{aligned}(2-7i)(3-2i) &= 2(3-2i)-7i(3-2i)\\ &= 6-4i-21i-14\\ &= -8-25i.\end{aligned}$$

(c)
$$\begin{aligned}(\sqrt{2}+i\sqrt{3})(1-i) &= \sqrt{2}(1-i)+i\sqrt{3}(1-i)\\ &= \sqrt{2}-i\sqrt{2}+i\sqrt{3}+\sqrt{3}\\ &= (\sqrt{2}+\sqrt{3})+i(\sqrt{3}-\sqrt{2}).\end{aligned}$$

(d)
$$\begin{aligned}(\sqrt{2}-i)^2 &= (\sqrt{2})^2-2\sqrt{2}i+i^2\\ &= 2-2\sqrt{2}i-1\\ &= 1-2\sqrt{2}i.\end{aligned}$$

Equality of Complex Numbers

To specify a complex number we must give two real numbers, the real and imaginary parts. So two complex numbers are equal if and only if their real and imaginary parts are equal (respectively).

> *For complex numbers $z_1 = a_1+ib_1$ and $z_2 = a_2+ib_2$ we have $z_1 = z_2$ if and only if $a_1 = a_2$ and $b_1 = b_2$.*

♦Example Find all complex numbers for which

$$z^2 = -3+4i.$$

Solution We write $z = x+iy$, with x and y real. Substituting into the equation we have

$$\begin{aligned} z^2 = (x+iy)^2 &= -3+4i\\ \text{i.e. } x^2-y^2+i2xy &= -3+4i.\end{aligned}$$

Now equate real and imaginary parts – remember the complex number on the left is equal to that on the right if and only if their real and imaginary parts are (respectively) equal. We get

$$x^2 - y^2 = -3 \quad \text{and} \quad 2xy = 4.$$

From the second of these equations we have

$$y = \frac{2}{x},$$

which we substitute into the first equation. This gives

$$x^2 - \left(\frac{2}{x}\right)^2 = -3.$$

Multiplying this equation through by x^2 gives as

$$\begin{aligned} x^4 - 4 &= -3x^2 \\ \text{i.e.} \quad x^4 + 3x^2 - 4 &= 0. \end{aligned}$$

This is a quadratic in x^2, we factorise

$$(x^2 + 4)(x^2 - 1) = 0,$$

so that $x^2 = 1$ or $x^2 = -4$. But x must be real, so we cannot have $x^2 = -4$. So we conclude $x^2 = 1$, which gives $x = \pm 1$. We found earlier that $y = 2/x$, so we have two possible solutions

$$(x, y) = (1, 2) \text{ or } (-1, -2).$$

Giving two possible complex numbers z,

$$z = 1 + 2i \quad \text{or} \quad z = -1 - 2i.$$

□

Division by Complex Numbers

If we have two complex numbers $z_1 = a_1 + ib_1$ and $z_2 = a_2 + ib_2 \neq 0$ how do we write

$$\frac{z_1}{z_2} = \frac{a_1 + ib_1}{a_2 + ib_2}$$

in the standard complex number form, i.e. as $a + ib$? To answer this question note that what we really have to do is get rid of the complex number in the denominator. Now recall that

$$(x + iy)(x - iy) = x^2 + y^2$$

is a real (positive) number. So if we multiply the denominator by $a_2 - ib_2$ we will get a real number. But to keep the equality we will have to multiply the numerator by $a_2 - ib_2$ as well. Here is the calculation

$$\begin{aligned}\frac{z_1}{z_2} = \frac{a_1 + ib_1}{a_2 + ib_2} &= \frac{a_1 + ib_1}{a_2 + ib_2} \cdot \frac{a_2 - ib_2}{a_2 - ib_2} \\ &= \frac{(a_1 + ib_1)(a_2 - ib_2)}{(a_2)^2 + (b_2)^2} \\ &= \frac{a_1 a_2 + b_1 b_2 + i(b_1 a_2 - a_1 b_2)}{(a_2)^2 + (b_2)^2}\end{aligned}$$

So the real and imaginary parts are

$$\mathrm{Re}\left(\frac{z_1}{z_2}\right) = \frac{a_1 a_2 + b_1 b_2}{(a_2)^2 + (b_2)^2} \text{ and } \mathrm{Im}\left(\frac{z_1}{z_2}\right) = \frac{b_1 a_2 - a_1 b_2}{(a_2)^2 + (b_2)^2}.$$

♦**Example** Write $\frac{1+i}{3-2i}$ in the form $a + ib$.

Solution

$$\begin{aligned}\frac{1+i}{3-2i} = \frac{1+i}{3-2i} \cdot \frac{3+2i}{3+2i} &= \frac{(1+i)(3+2i)}{3^2 + 2^2} \\ &= \frac{3 + 2i + 3i - 2}{9 + 4} \\ &= \frac{1 + 5i}{13} \\ \text{i.e.} \qquad \frac{1+i}{3-2i} &= \frac{1}{13} + i\frac{5}{13}\end{aligned}$$

□

Solving Real Quadratic Equations

We can now solve all quadratic equations with real coefficients. For example we solve

$$z^2 + z + 1 = 0.$$

Using the usual quadratic formula

$$\begin{aligned} z &= \frac{-1 \pm \sqrt{1^2 - 4.1.1}}{2} \\ &= \frac{-1 \pm \sqrt{-3}}{2}.\end{aligned}$$

Clearly the solutions are complex, we write them in the standard $a + ib$ format. We note that

$$\sqrt{-3} = \sqrt{-1 \times 3} = \sqrt{-1} \cdot \sqrt{3} = i\sqrt{3}.$$

So the solutions to the quadratic are

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

♠ Exercises 6

1. Express each of the following complex numbers in the form $x + iy$.

(a)	$(2-i)(3+2i)$	(b)	$(6+5i)(2+7i)$
(c)	$(3-2i)^2$	(d)	i^3
(e)	$\dfrac{2-i}{1+i}$	(f)	$\dfrac{2}{3+i} - \dfrac{1+i}{1-i}$
(g)	$\dfrac{1-i}{(2+i)^2}$	(h)	i^7

2. Show that $\dfrac{1+\sin\theta + i\cos\theta}{1+\sin\theta - i\cos\theta} = \sin\theta + i\cos\theta$.

3. Solve the following equations for z, writing your solution in the form $a + ib$

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

   (b) $z^2 + 2i + 5 = 0$

   (c) $5z^2 - 4z + 1 = 0$.

4. Find all solutions of the equation

$$z^2 = 6 - 8i.$$

Lecture 1.7 Complex Numbers (continued)

Complex Conjugation

> *If $z = x + iy$ (with x and y real) is a complex number we define the complex conjugate of z by $\overline{z} = x - iy$.*

Notice that all we have to do to get the complex conjugate of a complex number is to replace the imaginary part by its negative.

♦Example

(a) If $z = 3 + 2i$ then $\overline{z} = 3 - 2i$.

(b) If $z = 27 - 5i$ then $\overline{z} = 27 + 5i$.

(c) If $z = 5$ then $\overline{z} = 5$.

(d) If $z = 6i$ then $\overline{z} = -6i$.

Theorem *If $z_1, z_2 \in \mathbb{C}$ then $\overline{z_1 z_2} = \overline{z}_1 \overline{z}_2$.*

Proof Let $z_1 = x_1 + iy_1$, and $z_2 = x_2 + iy_2$. Then $z_1 z_2 = (x_1x_2 - y_1y_2) + i(y_1x_2 + y_2x_1)$.

So $\overline{(z_1 z_2)} = (x_1x_2 - y_1y_2) - i(y_1x_2 + y_2x_1)$.

Now,

$$\begin{aligned}
\overline{z}_1 &= x_1 - iy_1, \quad \text{and} \quad \overline{z}_2 = x_2 - iy_2 \quad \text{so} \\
\overline{z}_1\overline{z}_2 &= (x_1 - iy_1)(x_2 - iy_2) \\
&= (x_1x_2 - y_1y_2) - i(y_1x_2 + y_2x_1), \quad \text{as} \quad (-i)^2 = -1.
\end{aligned}$$

Comparing our expressions for $\overline{z_1 z_2}$ and $\overline{z}_1\overline{z}_2$ we conclude that $\overline{z_1 z_2} = \overline{z}_1\overline{z}_2$, as required. ■

In a similar vein we also have:

Theorem If z_1 and z_2 are complex numbers then

(a) $\overline{z_1 + z_2} = \overline{z}_1 + \overline{z}_2$

(b) $\overline{\left(\dfrac{z_1}{z_2}\right)} = \dfrac{\overline{z}_1}{\overline{z}_2}$, for $z_2 \neq 0$.

Proof Exercise!

Taken together these two theorems tell us that if we wish to take the complex conjugate of an expression involving complex numbers all we have to do is write down the expression with each complex number replaced by its conjugate.

♦Example

For $z_1, z_2, z_3 \in \mathbb{C}$ we have

$$\overline{\left(\frac{z_1^2 + 2z_1 z_2 + z_2 + 1 - i}{z_3}\right)} = \frac{\overline{z}_1^2 + 2\overline{z}_1\overline{z}_2 + \overline{z}_2 + 1 + i}{\overline{z}_3},$$

for $z_3 \neq 0$.

In the previous lecture we mentioned the method for writing a complex number $\frac{z_1}{z_2}$ (with $z_1, z_2 \in \mathbb{C}$) in the form $x + iy$. If $z_2 = x_2 + iy_2$ we had to multiply the expression by $\frac{x_2 - iy_2}{x_2 - iy_2}$. We now see that the method is simply multiplying the expression by $\frac{\overline{z}_2}{\overline{z}_2}$. The method works because $z_2\overline{z}_2$ is always a *real* number. This is easily seen $z_2\overline{z}_2 = (x_2 + iy_2)(x_2 - iy_2) = (x_2)^2 + (y_2)^2$.

For any complex number z we have

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

♦Example

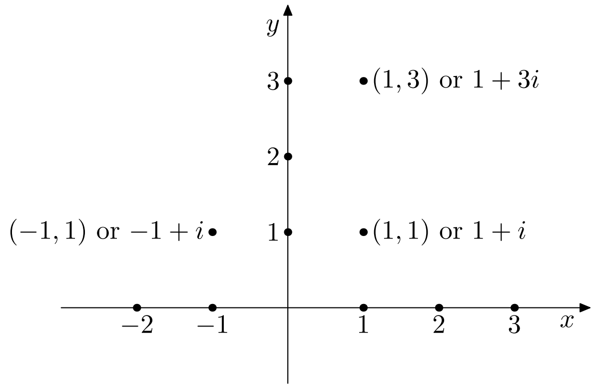
(a) For $z = 2 + 3i$ we have $z\overline{z} = 2^2 + 3^2 = 13$

(b) If $z = 2$ then $z\overline{z} = 2^2 = 4$

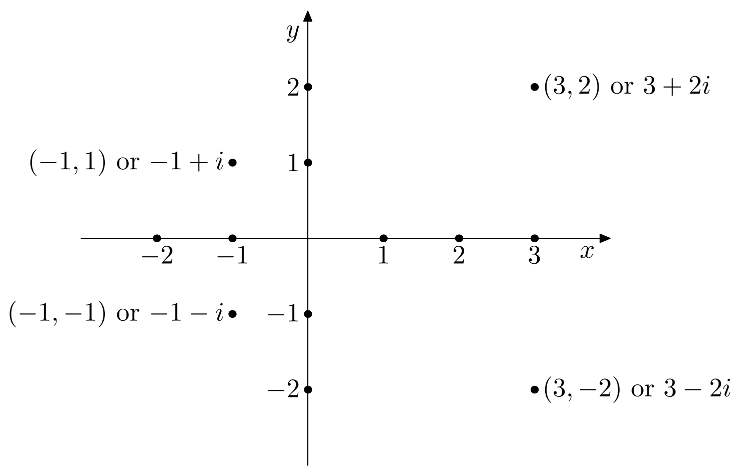
(c) If $z = -3i$ then $z\overline{z} = (-3)^2 = 9$.

Argand Diagrams

A complex number has two parts, its real part and its imaginary part. We can think of a complex number $z = x + iy$ $(x, y \in \mathbb{R})$ as an ordered pair (x, y). In this way we see that a complex number can be thought of as a point in the plane – we simply plot the point (x, y) in the xy plane. The resulting picture is called an Argand diagram. We represent z by the point (x, y) in the plane.



The real numbers (numbers with no imaginary part) simply lie on the x-axis. The pure imaginary numbers (those with no real part) lie along the y-axis.



The complex conjugate of a complex number represented on the Argand diagram.

The distance of a complex number $z = x + iy$ from the origin on an Argand diagram is

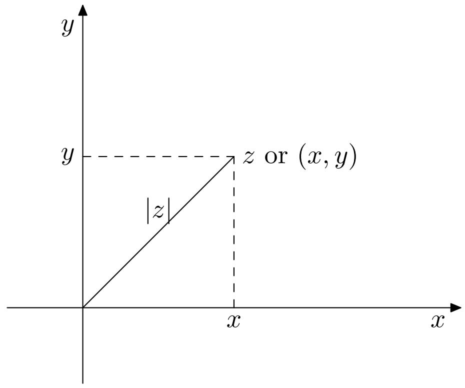
$$\sqrt{x^2 + y^2} = \sqrt{z\overline{z}}.$$

Notice that when z is real (ie $y = 0$) we have the distance as

$$\sqrt{x^2} = |x|.$$

This motivates the definition

If $z = x + iy \in \mathbb{C}$ we define the modulus of z (or the absolute value of z) to be $|z| = \sqrt{x^2 + y^2} = \sqrt{z\overline{z}}$.



♦Example

(a) If $z = 2 - 3i$ then $|z| = \sqrt{2^2 + (-3)^2} = \sqrt{13}$.

(b) If $z = 3 - 4i$ then $|z| = \sqrt{3^2 + (-4)^2} = 5$.

(c) If $z = 3$ then $|z| = 3$.

(d) If $z = -10$ then $|z| = 10$.

(e) If $z = 2i$ then $|z| = \sqrt{2^2} = 2$.

Theorem

If $z_1, z_2 \in \mathbb{C}$ then

(i) $|z_1 z_2| = |z_1||z_2|$.

(ii) $\left|\dfrac{z_1}{z_2}\right| = \dfrac{|z_1|}{|z_2|}$, for $z_2 \neq 0$.

Proof Both follow from the earlier theorem on the complex conjugate. We will prove (ii) and leave (i) as an exercise. For (ii) we have

$$\begin{aligned}\left|\frac{z_1}{z_2}\right| = \sqrt{\left(\frac{z_1}{z_2}\right)\overline{\left(\frac{z_1}{z_2}\right)}} = \sqrt{\frac{z_1}{z_2}.\frac{\overline{z}_1}{\overline{z}_2}} &= \sqrt{\frac{z_1\overline{z}_1}{z_2\overline{z}_2}} \\ &= \frac{\sqrt{z_1\overline{z}_1}}{\sqrt{z_2\overline{z}_2}} \\ &= \frac{|z_1|}{|z_2|}.\end{aligned}$$

■

Notice that, in general, $|z_1 + z_2| \neq |z_1| + |z_2|$. For example for $z_1 = 1 + i$ and $z_2 = 1 - i$ we get

$$|z_1 + z_2| = |2| = 2.$$

And

$$\begin{aligned} |z_1| + |z_2| &= |1+i| + |1-i| \\ &= \sqrt{1^2 + 1^2} + \sqrt{1^2 + (-1)^2} \\ &= \sqrt{2} + \sqrt{2} \\ &= 2\sqrt{2}, \end{aligned}$$

so $|z_1| + |z_2| > |z_1 + z_2|$, in this case. In fact, we have

Theorem If $z_1, z_2 \in \mathbb{C}$ then

$$|z_1 + z_2| \leq |z_1| + |z_2|.$$

Proof Notice that $|z_1 + z_2|$ and $|z_1| + |z_2|$ are both non-negative so we can prove the inequality by proving

$$|z_1 + z_2|^2 \leq (|z_1| + |z_2|)^2.$$

This is what we will now do.

$$\begin{aligned} |z_1 + z_2|^2 &= (z_1 + z_2)\overline{(z_1 + z_2)} \\ &= (z_1 + z_2)(\overline{z}_1 + \overline{z}_2) \\ &= z_1\overline{z}_1 + z_2\overline{z}_2 + z_1\overline{z}_2 + \overline{z}_1 z_2. \\ \text{And} \quad (|z_1| + |z_2|)^2 &= (\sqrt{z_1\overline{z}_1} + \sqrt{z_2\overline{z}_2})^2 \\ &= z_1\overline{z}_2 + z_2\overline{z}_2 + 2\sqrt{z_1\overline{z}_1}\sqrt{z_2\overline{z}_2.} \end{aligned}$$

Hence,

$$\begin{aligned} (|z_1| + |z_2|)^2 - |z_1 + z_2|^2 &= 2\sqrt{z_1\overline{z}_1}\sqrt{z_2\overline{z}_2} - z_1\overline{z}_2 - \overline{z}_1 z_2 \\ &= 2\sqrt{(z_1\overline{z}_2)(\overline{z}_1 z_2)} - z_1\overline{z}_2 - \overline{z}_1 z_2. \end{aligned}$$

Now $Z = z_1\overline{z}_2$ is a complex number so we can write $z_1\overline{z}_2 = Z = X + iY$, with X, Y real.

Then $\sqrt{(z_1\overline{z}_2)(\overline{z}_1 z_2)} = \sqrt{Z\overline{Z}} = \sqrt{X^2 + Y^2}$ and $z_1\overline{z}_2 + \overline{z}_1 z_2 = 2X$. So

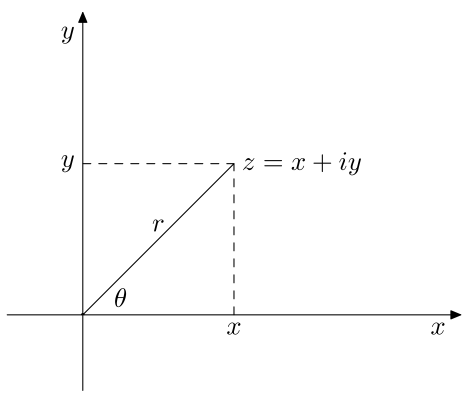
$$(|z_1| + |z_2|)^2 - |z_1 + z_2|^2 = 2\sqrt{X^2 + Y^2} - 2X \geq 2\sqrt{X^2} - 2X \geq 0.$$

Thus
$$(|z_1| + |z_2|)^2 - |z_1 + z_2|^2 \geq 0$$
and our result follows. ■

This result is also known as the triangle in equality. It can be thought of geometrically as stating the old fact that the length of one side of a triangle is less than the sum of the lengths of the other two sides.

De Moivre's Theorem

To motivate de Moivre's theorem we first introduce the polar form of a complex number.



From our diagram we have $z = x + iy = r\cos\theta + ir\sin\theta$.

$$\boxed{\text{The polar form of a complex number is } z = r(\cos\theta + i\sin\theta), |z| = r.}$$

If we wanted to look at powers of z we would need to look at expressions of the form

$$\begin{aligned} z^n &= [r(\cos\theta + i\sin\theta)]^n \\ &= r^n(\cos\theta + i\sin\theta)^n. \end{aligned}$$

What can we say about $(\cos\theta + i\sin\theta)^n$? It is here that the remarkable de Moivre's theorem enters.

Theorem *If $\theta \in \mathbb{R}$ then*

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

for n any rational number.

Proof See the exercises!

♦Example

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

(b) $i = \cos\frac{\pi}{2} + i\sin\frac{\pi}{2}$.

(c)
$$\begin{aligned} 1 &= \cos 0 + i \sin 0 \\ &= \cos(2\pi) + i\sin 2\pi. \end{aligned}$$

Notice in this last example we have a phenomenon typical of trigonometric functions – they are periodic,

$$\begin{aligned} \cos(\theta + 2\pi) &= \cos\theta \\ \sin(\theta + 2\pi) &= \sin\theta \ . \end{aligned}$$

In fact for any integer m,

$$\cos(\theta + 2m\pi) + i\sin(\theta + 2m\pi) = \cos\theta + i\sin\theta.$$

♠ Exercises 7

1. For each of the following complex numbers write down the complex conjugate and modulus
   (a) $6 + 2i$
   (b) $1 - 3i$
   (c) $\frac{1+i}{\sqrt{2}}$
   (d) $\frac{1}{1+i}$
   (e) $\frac{2-3i}{1-i}$
   (f) i.
2. Let $z_1, z_2 \in \mathbb{C}$. Show that
$$|z_1 + z_2|^2 + |z_1 - z_2|^2 = 2|z_1|^2 + 2|z_2|^2.$$
3. For $z_1, z_2 \in \mathbb{C}$ prove that $\overline{\left(\dfrac{z_1}{z_2}\right)} = \dfrac{\overline{z}_1}{\overline{z}_2}$, for $z_2 \neq 0$.
4. Use the addition formulae for the sine and cosine functions to deduce that
$$(\cos\theta + i\sin\theta)(\cos\phi + i\sin\phi) = \cos(\theta + \phi) + i\sin(\theta + \phi).$$

5. By induction, or otherwise, prove de Moivres theorem

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

for $n \in \mathbb{N}$.

6.* Extend de Moivre's theorem taking n to be (i) a negative integer and then (ii) $n = p/q$, a rational number.

Lecture 1.8 Functions on $\mathbb{R}$

In lecture 3 we introduced the idea of a function or mapping, f, consisting of set, A, called the domain, a set B, called the codomain, and a rule which assigns to every element $a \in A$ a unique element $b \in B$. We write

$$f : A \longrightarrow B$$

and

$$f : a \longmapsto b \quad \text{or} \quad \text{f(a) = b.}$$

In lecture 3 we also met the idea of composition of two functions f and g. For

$$\begin{array}{rcll} f : A & \longrightarrow & B, & a \longmapsto b \qquad \text{and} \\ g : B & \longrightarrow & C, & b \longmapsto c \end{array}$$

we define

$$g \circ f : A \longrightarrow C, \quad a \longmapsto c$$

by

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

Before specialising to functions $f : D \longrightarrow \mathbb{R}$ with $D \subseteq \mathbb{R}$, we examine two general concepts. Firstly, the idea of a function being *one-one*:

> *A function $f : A \longrightarrow B$ is one-one or injective if, for $a, a' \in A$, $f(a) = f(a')$ implies $a = a'$.*

Another way of saying that f is injective is to say that, for every $b \in B$ there is at most one $a \in A$ with $f(a) = b$. Or that, whenever $a_1 \neq a_2$, with $a_1, a_2 \in A$, then $f(a_1) \neq f(a_2)$.

Our second concept of this type is that of *surjectivity*:

> *The function $f : A \longrightarrow B$ is said to be onto or surjective if for every $b \in B$ there is at least one $a \in A$ such that $b = f(a)$.*

An alternative statement is to say that f is surjective if $f(A) = B$, that is, the codomain B is equal to the range of f.

$\boxed{\textit{A function which is both injective and surjective is called bijective.}}$

♦Example

(a)
$$\begin{aligned} f : \mathbb{N} &\longrightarrow \mathbb{R} \\ f : n &\longmapsto f(n) = \frac{1}{1+n}. \end{aligned}$$

This function is one-one (injective) but **not** surjective.

(b)
$$\begin{aligned} f : \mathbb{R} &\longrightarrow \mathbb{R} \\ f(x) &= x, \ x \in \mathbb{R}. \end{aligned}$$

This function is both injective and surjective – so its' bijective.

(c)
$$\begin{aligned} f : \mathbb{R} &\longrightarrow \mathbb{R} \\ f(x) &= x^2. \end{aligned}$$

The function is neither injective nor surjective. <u>Note</u>, however, that

$$\begin{aligned} g : \mathbb{R} &\longrightarrow [0, \infty) \\ f(x) &= x^2. \end{aligned}$$

is surjective (of course its' still not injective).

(d)
$$\begin{aligned} f : \mathbb{R} &\longrightarrow \{0, 1\} \\ f(x) &= \begin{cases} 0, & \text{when } x \text{ is irrational,} \\ 1, & \text{when } x \text{ is rational.} \end{cases} \end{aligned}$$

The function is clearly not injective. It is surjective.

(e)
$$\begin{aligned} f : [-1, 1] &\longrightarrow [0, 1] \\ f(x) &= \sqrt{1 - x^2}. \end{aligned}$$

The function is not injective but it is surjective. Note however that

$$\begin{aligned} g : [0, 1] &\longrightarrow [0, 1] \\ f(x) &= \sqrt{1 - x^2} \end{aligned}$$

is both injective and surjective and thus bijective.

Real Valued Functions on $\mathbb{R}$

We now specialise our study to consider only functions $f : D \longrightarrow \mathbb{R}$, with $D \subseteq \mathbb{R}$. All our previous discussions of things such as domain, codomain, range, composition, being injective, surjective and bijective apply to these more specialised functions.

The concepts of injective, surjective and bijective are now particularly easy to deal with. They are basically about how many "x-values" correspond to a given "y-value" – one or none for injective; one or more for surjective; and one and only one for bijective. The easiest way to visualise this is to draw the graph of the function, the question now becomes how many times does a horizontal line drawn at the given y-value hit the graph? Try it on the following examples.

♦**Example** Decide whether the following functions are injective or surjective justifying your statements.

(a) $f : \mathbb{N} \longrightarrow \mathbb{R}, \quad n \longmapsto n^2 - 9$

(b) $f : [-1, 1] = \{x \in \mathbb{R} : -1 \leq x \leq 1\} \longrightarrow \mathbb{R}, \quad x \longmapsto \sqrt{1 - x^4}$

(c) $f : \mathbb{R} \setminus \{1\} \longrightarrow \mathbb{R} \setminus \{0\}, \quad x \longmapsto \dfrac{1}{x - 1}$

Solution

(a) For $n_1, n_2 \in \mathbb{N}$ with $f(n_1) = f(n_2)$ we obtain $n_1^2 - 9 = n_2^2 - 9$. Thus $n_1^2 = n_2^2$, so that $n_1 = n_2$ as $n_1, n_2 \geq 0$. Hence f is injective. As there is no natural number n with $f(n) = 2$, the function is not surjective.

(b) As $f(-x) = f(x)$ for $x \in [-1, 1]$, the function is not injective. Note that the function is well–defined as $1 - x^4 \geq 0$ for all $x \in [-1, 1]$. Since the square root is always greater than or equal to 0, the function is not surjective.

(c) Take $x_1, x_2 \in \mathbb{R} \setminus \{1\}$ with $f(x_1) = f(x_2)$. Then

$$\begin{aligned} \frac{1}{x_1 - 1} = \frac{1}{x_2 - 1} \quad &\Longleftrightarrow \quad x_2 - 1 = x_1 - 1 \\ &\Longleftrightarrow \quad x_2 = x_1 \end{aligned}$$

and hence f is injective. Now take $y \in \mathbb{R}, y \neq 0$. Then

$$\begin{aligned} y = \frac{1}{x - 1} \quad &\Longleftrightarrow \quad y(x - 1) = 1 \\ &\Longleftrightarrow \quad x - 1 = \frac{1}{y} \\ &\Longleftrightarrow \quad x = \frac{1}{y} + 1 = \frac{1 + y}{y} \end{aligned}$$

showing that f is surjective. Thus f is bijective.

$\square$

We should also recall the definition of the *graph* of a function $f : D \longrightarrow \mathbb{R}$ as the set of ordered pairs $\{(x, y) \in \mathbb{R} \times \mathbb{R} : y = f(x) \text{ for some x} \in \text{D}\}$. We identify the Cartesian product $\mathbb{R} \times \mathbb{R}$ as our usual $x - y$ plane, $\mathbb{R}^2$. So the set $\{(x, y) \in \mathbb{R} \times \mathbb{R} : y = f(x), x \in D\}$ is "identified" with our usual notion of the plot of $y = f(x)$. That is a very wordy way of saying that our usual idea of drawing the graph of a function coincides (in the case $f : D \longrightarrow \mathbb{R}$) with our more general idea of a graph.

Your past experience with graph drawing will have taught you that the behaviour of the function as $x \longrightarrow \pm\infty$, or as x approaches a singular point of the function, is a most important feature of your sketch. To discuss such behaviour rigourously we need to understand the notion of a limit.

A First Look at Limits

The idea we want to make precise is the following. Suppose we have a function $f : D \longrightarrow \mathbb{R}$ with $D \subseteq \mathbb{R}$ and $c \in \mathbb{R}$. Our question is *does $f(x)$ approach some specific number, L, as x approaches c.* If it does then we say $f(x)$ has a limit at $x = c$. We write this symbolically as

$$f(x) \longrightarrow L \text{ as } x \longrightarrow c,$$

read as "$f(x)$ approaches L as x approaches c". Or as

$$\lim_{x \longrightarrow c} f(x) = L,$$

read as the "limit of $f(x)$ as x approaches c is L". The idea here is that we can make $f(x)$ as close to L as we like simply by taking x sufficiently near to c.

You might like to think of this idea in terms of approximation. We can make the difference $f(x) - L$ as small as we like by simply taking x close enough to c. Well, we can do this provided $f(x)$ *has limit L as x approaches c.* We are saying that I can always make $|f(x) - L|$ smaller than any number you give me simply by taking x close enough to c.

♦**Example** $f(x) = x^2$ has limit 1 as $x \longrightarrow 1$.

Give me any number, as small as you like, say 0.001. Then we can make $|f(x)-1|$ smaller than 0.001 by taking x close enough to 1.

We want $|x^2 - 1| < 0.001$. We can achieve this by making x just less than 1 or just greater than 1.

Let's have a look at the value of x, x_0 say for which $x_0^2 - 1 = 0.001$, so $x_0 = +\sqrt{1.001}$. Because of the absolute value signs $|x^2 - 1|$ we also need the value x_1, $x_1^2 - 1 = -0.001$ i.e. $x_1 = \sqrt{0.999}$. We conclude $|x^2 - 1| < 0.001$ provided

$$x_1 = \sqrt{0.999} < x < x_0 = \sqrt{1.001}.$$

Now $x_0 \doteqdot 1.0005$ and $x_1 \doteqdot 0.9095$ so we can ensure $|x^2 - 1| < 0.001$ by (for example)

$$0.9996 < x < 1.004.$$

No matter what number you give me, lets say it is ε, I can always find x_0, x_1 such that

$$\begin{aligned} |x^2 - 1| &< \varepsilon \text{ provided} \\ x_1 < x &< x_0, \end{aligned}$$

where x_0 and x_1 are near $x = 1$. Such a statement is just about right for the abstract definition of a limit, we'll say more about this in the next lecture.

♦**Example**

$$\lim_{x \longrightarrow 0} 2^x = 1.$$

This just says $2^x \longrightarrow 1$ as $x \longrightarrow 0$. This means that given **any** number, no matter how small, we can always make $|2^x - 1|$ smaller than this number by taking x close enough to 0. For example lets take the number $\varepsilon = 0.01$, the claim is we can make $|2^x - 1| < 0.01$ by choosing x close enough to 0. We don't have to find precisely the values of x for which the inequality is true we just need to find an interval about $x = 0$ for which it is true. Try some values on a calculator.

x	2^x(4 decimal places)	$\|2^x - 1\|$
-0.1	0.9330	0.0670
-0.05	0.9659	0.0341
-0.01	0.9931	0.0069
0.1	1.0718	0.0718
0.05	1.0070	0.0070

From this table of calculated values its clear that it would suffice to take $-0.01 < x < 0.01$ to ensure that

$$|2^x - 1| < 0.01.$$

♦**Example** Sketch the graph of the function $f : \mathbb{R} \setminus \{-1\} \longrightarrow \mathbb{R}, \quad x \longmapsto \dfrac{1}{1+x}$.

Solution. The first thing you notice is that $f(x)$ is "badly behaved" near $x = -1$, it's undefined there. What happens as we approach $x = -1$? Suppose we approach from below, i.e. consider $x = -1 - \delta$ where δ is small and positive. Then

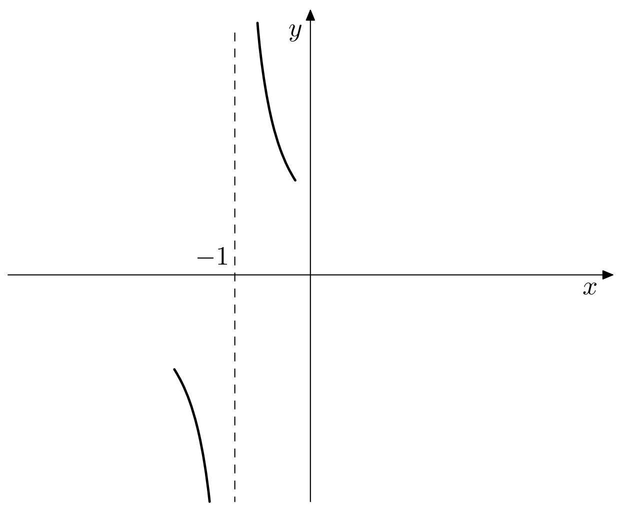
$$f(-1-\delta) = \frac{1}{1+(-1-\delta)} = -\frac{1}{\delta},$$

so if δ is small $\frac{1}{\delta}$ is large. The function is large in magnitude and negative just below $x = -1$. As we approach $x = -1$ from below, $f(x) \longrightarrow -\infty$.

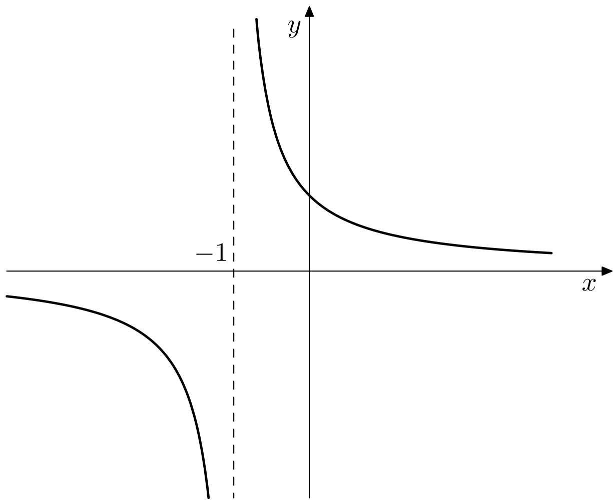
Next let's see what happens when we approach from above, let $x = -1 + \delta$, δ small and positive. Then

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

so $f(x) \longrightarrow +\infty$ as we approach $x = -1$ from above. We have thus far



The behaviour of the function for $x > -1$ or $x < -1$ is easy to determine once we know what happens as $x \longrightarrow \pm\infty$. For x large and positive $\frac{1}{1+x}$ will be **small** and positive. That is $f(x) \longrightarrow 0^+$ as $x \longrightarrow +\infty$. We use the notation $f(x) \longrightarrow 0^+$ to indicate that $f(x)$ is approaching zero from above, i.e. from the positive side. For x large and negative $\frac{1}{1+x}$ will be small and negative. That is $f(x) \longrightarrow 0^-$ as $x \longrightarrow -\infty, f(x)$ approaches 0 from below (i.e. from the negative side). One important point on the graph is where it cuts the y-axis, this is at $x = 0$ and $f(0) = 1$; so the graph cuts the y-axis at $(0, 1)$. This is enough information to give us a reasonably accurate sketch.



$$y = f(x) = \tfrac{1}{1+x}$$

$\square$

♠ Exercises 8

1. What is the largest subset $D \subseteq \mathbb{R}$ such that $f : D \longrightarrow \mathbb{R}$ is a function, where f is given by

   (a) $f(x) = \frac{1}{1+x^2}$
   (b) $f(x) = \sqrt{x^2 - 9}$
   (c) $f(x) = x^3 - 1$

   In each case find the range of f.

2. For each of the functions of question 1 state whether or not the function is injective or surjective.

3. Find a $\delta > 0$ such that $|x^3 - 1| < \frac{1}{10}$ whenever $|x - 1| < \delta$.

   [Hint: $|x - 1| < \delta$ means $1 - \delta < x < 1 + \delta$. You have to find a $\delta > 0$ which ensures that
   $$|x^3 - 1| < \frac{1}{10}.]$$

4. Sketch the graph of $f(x) = \frac{1}{x^2-4}$.

5*. Show that given any $\varepsilon > 0$ we can always find δ such that
   $$\left| \frac{1}{x+1} - 1 \right| < \varepsilon$$
   whenever $|x| < \delta$.

   [Hint: The idea here is to explicitly give a relation which determines the possible values of δ **given** ε. For example you may end up with something like $\delta < 1 + \varepsilon$ (this doesn't work here!).]

Lecture 1.9 Limits

Formal Definition of a Limit

In the last lecture we began to discuss what is meant by a statement such as "$f(x)$ has limit L as x approaches c". The statement

$$\lim_{x \longrightarrow c} f(x) = L$$

means that given any number ε, no matter how small, we can ensure that the distance, $|f(x) - L|$, between $f(x)$ and L, is smaller than ε by taking x close enough to c, that is, by making the distance $|x - x|$ small enough. Let's state this formally.

We will say $\lim\limits_{x \longrightarrow c} f(x) = L$, or $f(x) \longrightarrow L$ as $x \longrightarrow c$ or $f(x)$ has limit L as $x \longrightarrow c$, if, given any $\varepsilon > 0$ there exists a $\delta > 0$ such that $|f(x) - L| < \varepsilon$ whenever $0 < |x - c| < \delta$.

Note that $0 < |x - c| < \delta$ excludes $x = c$. For the limit we are only interested in the behaviour of the function near c, not in $f(c)$. In fact, we don't even require that c is in the domain of f.

To show that $f(x)$ has limit L we have to demonstrate the existence of a $\delta > 0$ for any given ε. The idea is exactly that of the last lecture, you give me any number, no matter how small, call it ε. Then I have to come up with a $\delta > 0$ such that $|f(x) - L| <\in$ whenever $0 < |x - c| < \delta$.

Lemma 1.1 *Take $a, c \in \mathbb{R}$. Then*

(a) $\lim_{x \longrightarrow a} c = c$

(b) $\lim_{x \longrightarrow a} x = a$

Proof

Take $a, c \in \mathbb{R}$ and $\varepsilon > 0$.

(a) As $|c - c| = 0 < \varepsilon$ no matter how far x is from a, we may choose any $\delta > 0$, for example, put $\delta = 231$. Then $|c - c| = 0 < \varepsilon$ whenever $|x - a| < 231$. Thus $\lim_{x \longrightarrow a} c = c$.

(b) Put $\delta = \varepsilon$. Then $|x - a| < \varepsilon$ whenever $|x - a| < \delta = \varepsilon$. Thus $\lim_{x \longrightarrow a} x = a$.
$\square$

♦**Example** Prove that $\dfrac{1}{1+x} \longrightarrow 1$ as $x \longrightarrow 0$.

Solution Let $\varepsilon > 0$ be given.

Our strategy has to be to find a $\delta > 0$, depending on ε, such that if $0 < |x-0| < \delta$ then $\left|\dfrac{1}{1+x} - 1\right| < \varepsilon$.

$$\text{Now we note that } \left|\frac{1}{1+x} - 1\right| = \left|\frac{-x}{1+x}\right| = \frac{|x|}{|1+x|} \leq \frac{|x|}{1-|x|}, \quad \text{for } |x| < 1.$$

So for $|x| < \delta < 1$ we have

$$\left|\frac{1}{1+x} - 1\right| < \frac{\delta}{1-\delta}.$$

[Note we have increased the numerator $|x|$ to δ and decreased the denominator from $1-|x|$ to $1-\delta$. So the fraction $\dfrac{|x|}{1-|x|}$ has been **increased** to $\dfrac{\delta}{1-\delta}$.]

So we have to choose δ so that $\dfrac{\delta}{1-\delta} < \varepsilon$ ie. $\delta < \dfrac{\varepsilon}{1+\varepsilon}$.

So given **any** $\varepsilon > 0$ we have

$$\left|\frac{1}{1+x} - 1\right| < \varepsilon \quad \text{whenever}$$

$$|x-0| < \delta < \frac{\varepsilon}{1+\varepsilon}.$$

We are finished. □

As you can see there is quite an art in getting one of these proofs to work. Remember you are **given** $\varepsilon > 0$ and you have to **find** an appropriate δ such that $|x-c| < \delta$ implies $|f(x) - L| < \varepsilon$. Here is another example:

♦**Example** Prove that $\displaystyle\lim_{x \longrightarrow 0} \left(\frac{1+x}{1-x}\right) = 1$.

Solution Take $x \in \mathbb{R} \setminus \{1\}$ and $\varepsilon > 0$.

And suppose $0 < |x| < \delta$, where $\delta > 0$ is to be determined.

Now $1 + x \le 1 + |x|$ and $1 - x \ge 1 - |x|$, so we have that

$$\frac{1+x}{1-x} \le \frac{1+|x|}{1-|x|}.$$

[**Remember:** You increase the value of a fraction $\frac{a}{b}$ by increasing the numerator a or decreasing the denominator b (or both).]

And as $0 < |x| < \delta$ we have

$$\frac{1+x}{1-x} \le \frac{1+|x|}{1-|x|} < \frac{1+\delta}{1-\delta}.$$

Now, for $|x| < \delta < 1$ we have $|1-x| \ge 1 - |x| > 0$ and hence

$$\frac{1}{1-|x|} \ge \frac{1}{|1-x|}.$$

We obtain

$$\begin{aligned} \left|\frac{1+x}{1-x} - 1\right| &= \left|\frac{x}{1-x}\right| \\ &= \frac{|x|}{|1-x|} \\ &\le \frac{|x|}{1-|x|} \\ &< \frac{\delta}{1-\delta} \end{aligned}$$

To obtain $\dfrac{\delta}{1-\delta} = \varepsilon$ we must have $\delta = \dfrac{\varepsilon}{1+\varepsilon}$. Then $\delta < 1$, so that $\left|\dfrac{1+x}{1-x} - 1\right| < \varepsilon$ whenever $|x| < \delta = \varepsilon/(1+\varepsilon)$.

This is the formal definition of $\lim\limits_{x \longrightarrow 0} \left(\dfrac{1+x}{1-x}\right) = 1$. □

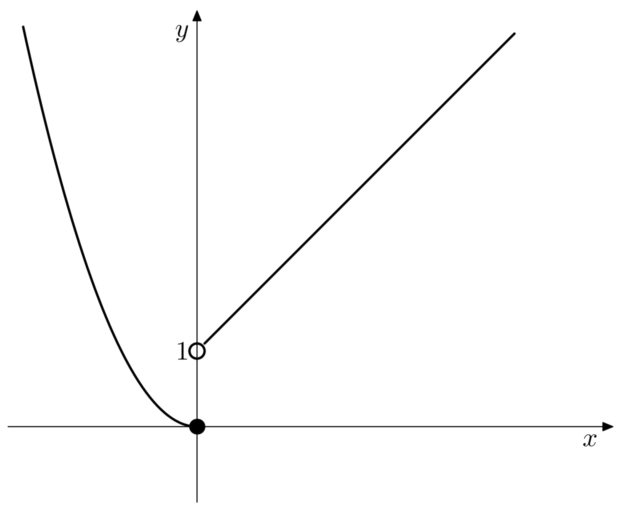
One-Sided Limits

Implicit in our definition of limits, so far, is the notion that we approach the point $x = c$ from either side. That is $f(x) \longrightarrow L$ whether we approach c from above, ie. $x \longrightarrow c^+$, or from below, $x \longrightarrow c^-$. But there are cases where we can approach from only one side. For example, the limit of $\sqrt{x}$ as $x \longrightarrow 0$. This limit, for real valued functions, must be from above, ie. we should really write $x \longrightarrow 0^+$. This is a one-side limit.

♦**Example** Discuss the limit $x \longrightarrow 0$ for the function

$$f(x) = \begin{cases} x^2 & , \quad x \le 0 \\ 1+x & , \quad x > 0. \end{cases}$$

Solution



Notice the jump at $x = 0$. We need to discuss separately the limits from above and below.

From above, it is easy to see

$$\begin{aligned} \lim_{x \longrightarrow 0^+} f(x) &= \lim_{x \longrightarrow 0^+} (1+x) \\ &= 1. \end{aligned}$$

And from below

$$\begin{aligned} \lim_{x \longrightarrow 0^-} f(x) &= \lim_{x \longrightarrow 0^-} x^2 \\ &= 0. \end{aligned}$$

□

Two important one-sides limits are the limits $x \longrightarrow \pm\infty$, if they exist for the given function.

Behaviour of a Function for Large $|x|$

If a function is defined for large values of x it may or may not approach a finite value as $x \longrightarrow \pm\infty$.

If $f(x)$ does have a finite limit, L, as $x \longrightarrow \pm\infty$ then, in terms of our previous understanding, we mean that given any $\varepsilon > 0$, $|f(x) - L| < \varepsilon$, whenever $|x|$ is large enough. Formally:

> *We say $f(x)$ has finite limit L as $x \longrightarrow \pm\infty$, or $\lim\limits_{x\longrightarrow\pm\infty} f(x) = L < \infty$ or $f(x) \longrightarrow L < \infty$ as $|x| \longrightarrow \infty$, if given $\varepsilon > 0$ there exists $X > 0$ such that $|f(x)-L| < \varepsilon$ whenever $|x| > X$.*

All this says is that if you give me any $\varepsilon > 0$ I can find a number $X > 0$ such that $|f(x) - L|$ is smaller than ε whenever $|x| > X$.

⧫Example

Prove that $\dfrac{1}{\sqrt{x}} \longrightarrow 0$ as $x \longrightarrow \infty$.

Solution Let $\varepsilon > 0$ be given. We have to find an $X > 0$ such that $\left|\dfrac{1}{\sqrt{x}}\right| < \varepsilon$ whenever $x > X$. Well this is pretty easy, in this case.

Let $x > X$, with X to be determined. Then $\dfrac{1}{\sqrt{x}} < \dfrac{1}{\sqrt{X}}$, as decreasing the denominator increases the fraction. So $\left|\dfrac{1}{\sqrt{x}}\right| = \dfrac{1}{\sqrt{x}} < \dfrac{1}{\sqrt{X}} = \varepsilon$, for $X = \frac{1}{\varepsilon^2}$. We are finished, we have $\left|\dfrac{1}{\sqrt{x}}\right| < \varepsilon$ whenever $x > \frac{1}{\varepsilon^2}$. $\square$

⧫Example Prove that $\dfrac{\sqrt{x+1}-\sqrt{x-1}}{x} \longrightarrow 0$ as $x \longrightarrow \infty$.

Solution The trick here is to make the troublesome numerator "look better". We rationalize the numerator (i.e. eliminate the square roots),

$$\begin{aligned}\left(\sqrt{x+1}-\sqrt{x-1}\right)\left(\sqrt{x+1}+\sqrt{x-1}\right) &= (x+1)-(x-1), \text{ "difference of squares".}\\ &= 2\end{aligned}$$

Therefore

$$\begin{aligned}\frac{\sqrt{x+1}-\sqrt{x-1}}{x} &= \frac{\sqrt{x+1}-\sqrt{x-1}}{x}\cdot\frac{\sqrt{x+1}+\sqrt{x-1}}{\sqrt{x-1}+\sqrt{x-1}}\\ &= \frac{2}{x\left(\sqrt{x+1}+\sqrt{x-1}\right)}.\end{aligned}$$

It's now clear that this expression tends to zero as $x \longrightarrow \infty$. We need to prove it formally.

We note that for $x > 1$,

$$\sqrt{x+1} + \sqrt{x-1} > 1,$$

so

$$\frac{1}{\sqrt{x+1} + \sqrt{x-1}} < 1.$$

Thus we have

$$\frac{\sqrt{x+1} - \sqrt{x-1}}{x} = \frac{2}{x\left(\sqrt{x+1} + \sqrt{x-1}\right)} < \frac{2}{x}.$$

So for $x > X > 1$, we have

$$\left|\frac{\sqrt{x+1} - \sqrt{x-1}}{x}\right| < \frac{2}{x} < \frac{2}{X} = \varepsilon,$$

for $X = \frac{2}{\varepsilon}$. We are done. $\square$

♠ Exercises 9

1. Give a formal proof that $x^2 \longrightarrow 1$ as $x \longrightarrow 1$.
2. Prove formally that
$$\lim_{x \longrightarrow 0} \frac{1+x^2}{1-x^2} = 1.$$
3. Discuss the limit of $\frac{1}{1-x^2}$ as $x \longrightarrow 1$.
4. Prove that
$$\lim_{x \longrightarrow \infty} \sqrt{\frac{x+1}{x-1}} = 1.$$

Lecture 1.10 Limits and Continuous Functions

In manipulating and calculating limits it is useful to have a set of rules for such things as the limit of a sum or product. The following theorem provides just those rules.

Theorem *Take $f, g : D \longrightarrow \mathbb{R}$ with $D \subseteq \mathbb{R}$. Suppose that*

$$\begin{aligned} \lim_{x \longrightarrow c} f(x) &= L_1 < \infty \quad \textit{and} \\ \lim_{x \longrightarrow c} g(x) &= L_2 < \infty. \end{aligned}$$

Then

(a) $\lim_{x \longrightarrow c} [\alpha f(x) + \beta g(x)] = \alpha L_1 + \beta L_2$ *where* $\alpha, \beta \in \mathbb{R}$.

(b) $\lim_{x \longrightarrow c} [f(x) g(x)] = L_1 L_2$

(c) $\lim_{x \longrightarrow c} \left[\dfrac{f(x)}{g(x)} \right] = \dfrac{L_1}{L_2}$, *for* $L_2 \neq 0$.

Proof

(a) This just says that the "limit of a sum is the sum of the limits".

Let $\varepsilon > 0$ be given. Assume one of α, β is non-zero. Now both f and g have well defined limits as $x \longrightarrow c$ so we know there exist a number δ such that

$$|f(x) - L_1| < \frac{\varepsilon}{|\alpha| + |\beta|} \quad \text{and} \quad |g(x) - L_2| < \frac{\varepsilon}{|\alpha| + |\beta|}$$

whenever $|x - c| < \delta$. Then

$$\begin{aligned} |[\alpha f(x) + \beta g(x)] - (\alpha L_1 + \beta L_2)| &= |\alpha [f(x) - L_1] + \beta [g(x) - L_2]| \\ &\leq |\alpha [f(x) - L_1]| + |\beta [g(x) - L_2]| , \end{aligned}$$

by the triangle inequality. Combining this with our previous equations we have

$$\begin{aligned} |[\alpha f(x) + \beta g(x)] - (\alpha L_1 + \beta L_2)| &\leq |\alpha| |f(x) - L_1| + |\beta| |g(x) - L_2| \\ &< \frac{|\alpha| \varepsilon}{|\alpha| + |\beta|} + \frac{|\beta| \varepsilon}{|\alpha| + |\beta|} = \varepsilon. \end{aligned}$$

Whenever $|x - c| < \delta$.

Which completes the proof of (a).

(b) This is quite an intricate little proof.

Firstly we observe that

$$f(x)g(x) - L_1L_2 = (f(x) - L_1)g(x) + (g(x) - L_2)L_1$$

so by the triangle inequality.

$$(*) \quad |f(x)g(x) - L_1L_2| \leq |f(x) - L_1|\,|g(x)| + |g(x) - L_2|\,|L_1|\,.$$

Now f and g both have well-defined limits as $x \longrightarrow c$, so we have for any $\overline{\varepsilon} > 0$ and some $\delta > 0$

$$|f(x) - L_1| < \overline{\varepsilon} \text{ and } |g(x) - L_2| < \overline{\varepsilon},$$

whenever $|x - c| < \delta$.

From the second of these inequalities we also have

$$-\overline{\varepsilon} < g(x) - L_2 < \overline{\varepsilon},$$

i.e.,

$$L_2 - \overline{\varepsilon} < g(x) < L_2 + \overline{\varepsilon}, \quad \text{whenever} \quad |x - c| < \delta.$$

So, whenever $|x - c| < \delta$, we have

$$|f(x) - L_1| < \overline{\varepsilon}, \ |g(x) - L_2| < \overline{\varepsilon}$$

$$\begin{aligned} \text{and} \quad |g(x)| \quad &< \quad \text{maximum} \quad (|L_2 + \overline{\varepsilon}|\,, |L_2 - \overline{\varepsilon}|) \\ &< \quad |L_2| + \overline{\varepsilon}. \end{aligned}$$

Put all this into (*)

$$|f(x)g(x) - L_1L_2| < \overline{\varepsilon}\,(|L_2| + \overline{\varepsilon}) + \overline{\varepsilon}\,|L_1| = \overline{\varepsilon}\,(\overline{\varepsilon} + |L_1| + |L_2|)\,.$$

This is true for **any** $\overline{\varepsilon} > 0$, in particular for $\overline{\varepsilon}^2 + \overline{\varepsilon}\,(|L_1| + |L_2|) = \varepsilon$, i.e. for

$$\overline{\varepsilon} = \frac{1}{2}\left[-\left(|L_1| + |L_2|\right) + \sqrt{\left(|L_1| + |L_2|\right)^2 + 4\varepsilon}\right]$$

(use the quadratic formula). We have our proof,

$$|f(x)g(x) - L_1L_2| < \varepsilon$$

whenever $|x - c| < \delta$.

(c) The proof is similar to that of (b). See Exercises. ■

♦**Example** Evaluate the following limits

$$\text{(i)}\quad \lim_{x\longrightarrow 0}\frac{3-2x}{x+6} \qquad\qquad \text{(ii)}\quad \lim_{x\longrightarrow\infty}\frac{3-2x}{x+6}$$
$$\text{(iii)}\quad \lim_{x\longrightarrow 0}\frac{x+1}{x-1}\cos x \qquad\qquad \text{(iv)}\quad \lim_{x\longrightarrow\infty}\left(\frac{x+4}{x}\right).$$

Solution

Recall that $\lim_{x\longrightarrow a} c = c$ and $\lim_{x\longrightarrow a} x = a$ for $a, c \in \mathbb{R}$ and use the Theorem above.

(i)

$$\begin{aligned}
\lim_{x\longrightarrow 0}\frac{3-2x}{x+6} &= \frac{\lim_{x\longrightarrow 0}(3-2x)}{\lim_{x\longrightarrow 0}(x+6)} \\
&= \frac{\lim_{x\longrightarrow 0}(3) - \lim_{x\longrightarrow 0}(2x)}{\lim_{x\longrightarrow 0}(x) + \lim_{x\longrightarrow 0}(6)} \\
&= \frac{3 - \lim_{x\longrightarrow 0}(2)\lim_{x\longrightarrow 0}(x)}{0+6} \\
&= \frac{3-0}{6} = \frac{1}{2}.
\end{aligned}$$

(ii) We cannot apply (c) directly as $3-2x \longrightarrow -\infty$ as $x \longrightarrow \infty$. We may assume $x \neq 0$ and rewrite the expression so that we can apply (c):

$$\begin{aligned}
\lim_{x\longrightarrow\infty}\frac{3-2x}{x+6} &= \lim_{x\longrightarrow\infty}\frac{\frac{3}{x}-2}{1+\frac{6}{x}} \\
&= \frac{\lim_{x\longrightarrow\infty}\left(\frac{3}{x}-2\right)}{\lim_{x\longrightarrow\infty}\left(1+\frac{6}{x}\right)} \\
&= \frac{-2}{1} = -2.
\end{aligned}$$

(iii) Apply (b) and (c),

$$\begin{aligned}
\lim_{x\longrightarrow 0}\frac{x+1}{x-1}\cos x &= \frac{\lim_{x\longrightarrow 0}(x+1)}{\lim_{x\longrightarrow 0}(x-1)}\lim_{x\longrightarrow 0}\cos x \\
&= \frac{1}{-1}\cdot 1 \\
&= -1.
\end{aligned}$$

(iv) Again we cannot use (c) directly. We rewrite the expression so that we can apply (c):

$$\lim_{x \longrightarrow \infty} \left(\frac{x+4}{x} \right) = \lim_{x \longrightarrow \infty} \left(1 + \frac{4}{x} \right) = 1.$$

□

Continuity

You will have gained the impression that the common functions are what could reasonably be called *continuous*, that is if you trace along the function by changing x the graph of the function it does not have any breaks. We think, intuitively, of a function being continuous if we can draw its graph without removing the pencil from the paper.

We now refine these ideas and formalise them mathematically. In fact, a little reflection should convince you that for a function to be continuous at $x = c$ we require two things:

1. $f(c)$ is defined, that is, $c \in D$, where $f : D \longrightarrow \mathbb{R}$, and
2. $f(x) \longrightarrow f(c)$ as x approaches c.

But this is nothing more or less than a statement about limits.

A function $f : D \longrightarrow \mathbb{R}, D \subseteq \mathbb{R}$, is continuous at $c \in D$ if $$\lim_{x \longrightarrow c} f(x) = f(c).$$ *We say that f is continuous if it is continuous at all $c \in D$.*

So to formally prove that $f(x)$ is continuous at $x = c$ we need to formally prove that $f(x)$ has limit $f(c)$ at $x = c$.

♦**Example** Prove that the identity function $\mathrm{id}_{\mathbb{R}} : \mathbb{R} \longrightarrow \mathbb{R}, \quad \mathrm{x} \longmapsto \mathrm{x}$ is continuous at x for all $x \in \mathbb{R}$.

Solution Recall that $\lim_{x \longrightarrow c} x = c$. Thus

$$\lim_{x \longrightarrow c} \mathrm{id}_{\mathbb{R}}(\mathrm{x}) = \lim_{\mathrm{x} \longrightarrow \mathrm{c}} \mathrm{x} = \mathrm{c} = \mathrm{id}_{\mathbb{R}}(\mathrm{c}).$$

So we are finished, $\mathrm{id}_{\mathbb{R}}$ is continuous on the whole real line. □

♦**Example** Prove that $f : \mathbb{R} \setminus \{0\} \longrightarrow \mathbb{R}, \quad x \longmapsto \frac{1}{x}$ is continuous at x for all $x \in (0, \infty)$.

Solution We take $x > 0$ and let $\varepsilon > 0$ be given. Then for any $c \in (0, \infty)$,

$$\left|\frac{1}{x} - \frac{1}{c}\right| = \frac{|c - x|}{cx}.$$

Let $|x - c| < \delta$ for some $\delta > 0$ to be determined. Then

$$\left|\frac{1}{x} - \frac{1}{c}\right| < \frac{\delta}{xc}.$$

Now $|x - c| < \delta$ implies

$$-\delta < x - c < \delta,$$

i.e. $c - \delta < x < \delta + c$, we take $\delta \leq \frac{c}{2}$ so that

$$x \geq c - \frac{c}{2} = \frac{c}{2}.$$

Then

$$\left|\frac{1}{x} - \frac{1}{c}\right| < \frac{\delta}{xc} < \frac{\delta}{\frac{c}{2} \cdot c} = \frac{2\delta}{c^2} \leq \varepsilon,$$

for $\delta = \text{minimum}\left(\frac{c}{2}, \frac{\varepsilon c^2}{2}\right)$. We have our formal statement of continuity at any point $c \in (0, \infty)$. □

A trivial consequence of our definition of continuity and our previous theorem on limits is the following.

Theorem *Let $f : D_f \longrightarrow \mathbb{R}$ and $g : D_g \longrightarrow \mathbb{R}$ be continuous at $x = c \in D_f \cap D_g$. Then*

(a) $\alpha f(x) + \beta g(x)$ is continuous at $x = c$ for all $\alpha, \beta \in \mathbb{R}$;

(b) $f(x)g(x)$ is continuous at $x = c$;

(c) $\dfrac{f(x)}{g(x)}$ is continuous at $x = c$ for $g(c) \neq 0$.

Armed with this theorem we can considerably expand our list of continuous functions. All polynomial functions $f : \mathbb{R} \longrightarrow \mathbb{R}$ given by

$$f(x) = a_0 + a_1 x + a_2 x^2 + \ldots + a_n x^n$$

where $n \in \mathbb{N}$ and $a_i \in \mathbb{R},\ i = 0, 1, 2, \ldots$ are continuous.

Furthermore, all rational functions $f : D \longrightarrow \mathbb{R}$ given by

$$f(x) = p(x)/q(x)$$

where $p(x)$ and $q(x)$ are polynomials, are continuous.

For continuity at $x = c$ it is necessary and sufficient that $\lim\limits_{x \longrightarrow c^+} f(x) = f(c) = \lim\limits_{x \longrightarrow c^-} f(x)$. This typically fails when there is a jump in $f(x)$ from one side of $x = c$ to the other.

♦**Example** The following functions are discontinuous:

(a)
$$f : \mathbb{R} \longrightarrow \mathbb{R}, \quad x \longmapsto \begin{cases} x^2, & x < 0 \\ 1 + x, & x > 0 \end{cases}.$$

Here continuity fails at $x = 0$.

(b)
$$f : \mathbb{R} \longrightarrow \mathbb{R}, \quad x \longmapsto [x],$$

where $[x]$ is the greatest integer less than or equal to x. This function has discontinuities at $x = n$ for all $n \in \mathbb{N}$.

♠ Exercises 10

1. Sketch the two functions of the last example, indicating the discontinuities.
2. Evaluate the following limits:
   (a) $\lim\limits_{x \longrightarrow 0} \dfrac{\sqrt{x^2+1} - \sqrt{x+1}}{x}$
   (b) $\lim\limits_{x \longrightarrow -\infty} \dfrac{x^2 - 1}{x^2 + 1}$
   (c) $\lim\limits_{x \longrightarrow \infty} \dfrac{x^2 + x + 1}{2x^2 - x - 1}$
3. Investigate the limit as $x \longrightarrow 1$ of
$$\frac{x^2 - 3x + 2}{x^3 - 3x^2 + 2}.$$
4. Prove formally that $f : (-1, \infty] \to \mathbb{R}, \quad x \longmapsto \dfrac{1}{x+1}$ is continuous.

Lecture 1.11 Continuous Functions

A function $f : D \longrightarrow \mathbb{R}$ is continuous in an open interval $(a, b) \subseteq D$ if it is continuous at each point $x \in (a, b)$.

A function $f : D \longrightarrow \mathbb{R}$ is continuous on a closed interval $[a, b] \subseteq D$ if

(i) $f(x)$ is continuous for each $x \in (a, b)$;

(ii) $\lim\limits_{x \to a^+} f(x) = f(a)$ and $\lim\limits_{x \to b^-} f(x) = f(b)$.

So $f : \mathbb{R} \longmapsto \mathbb{R}, \quad x \longmapsto |x|$, for example, is continuous on $(0, 1)$ (in fact, on $(0, \infty)$), but not continuous on $[0.1]$.

Graphs of continuous functions have all the "nice" properties one would expect of a curve that is drawn on paper without removing your pencil from the page. A good example of such properties is the intermediate value theorem.

The Intermediate Value Property

Theorem *Suppose f is continuous on $[a, b]$ and that $f(a) \neq f(b)$. Then $f(x)$, $x \in (a, b)$, takes every value between $f(a)$ and $f(b)$.*

Remark: As your pencil moves from $f(a)$ to $f(b)$ you must pass over every value (every real number) between $f(a)$ and $f(b)$.

Proof We restrict our function f to $[a, b]$ and suppose that there is a $y \in]f(a), f(b)[\setminus \mathrm{Im(f)}$. Put $A := \{x \in [a, b] \mid f(t) < y \text{ for all } t \in [a, x]\}$.

Since $a \in A$, $A \neq \emptyset$. By the definition of A, b is an upper bound for A. Hence, by the Completeness Axiom for the Real Numbers, A has a supremum, say s. Plainly, $s \in [a, b]$.

As $f(s) \neq y$, either $f(s) < y$, or $f(s) > y$. Put $\varepsilon := |f(s) - y|$, and take $\delta > 0$.

If $f(s) < y$, then $\varepsilon = y - f(s)$.

As $s = \sup A$, it follows from the definition of A that there is an $x \in [s, s + \delta[$ with $f(x) > y$.

Then

$$\begin{aligned}
|f(x) - f(s)| &= f(x) - f(s) && \text{as } f(x) > y > f(s)\\
&= f(x) - y + y - f(s)\\
&= f(x) - y + \varepsilon\\
&> \varepsilon && \text{as } f(x) > y.
\end{aligned}$$

Thus, in this case, f is not continuous at s.

If, on the other hand, $f(s) > y$, then $\varepsilon = f(s) - y$.

Choose $x \in A$ with $s - \frac{\delta}{2} < x$.

By the definition of A, $f(x) < y$. Then

$$\begin{aligned}
|f(x) - f(s)| &= f(s) - f(x) && \text{as } f(x) < y < f(s)\\
&= f(s) - y + y - f(x)\\
&= \varepsilon + y - f(x)\\
&> \varepsilon && \text{as } f(x) < y,
\end{aligned}$$

Thus, in this case, f is not continuous at s. □

Continuity for Trigonometric Functions

We know from the previous lecture that all polynomial functions and all rational functions are continuous.

We would also guess that the trigonometric functions, sine, cosine, tangent, cotangent, secant and cosecant are continuous at all points of their respective natural domains. Just sketch them! Let's formalise this.

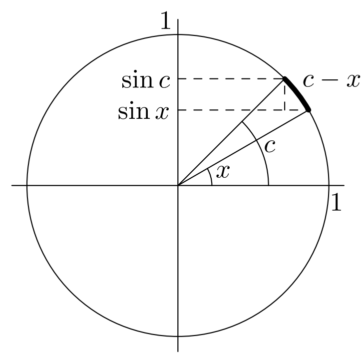
Theorem *The functions*

$$\begin{aligned}
&\sin : \mathbb{R} \longrightarrow \mathbb{R}, \quad x \longmapsto \sin x\\
&\cos : \mathbb{R} \longrightarrow \mathbb{R}, \quad x \longmapsto \cos x\\
&\tan : \mathbb{R} \setminus \{\tfrac{2n+1}{2}\pi : n \in \mathbb{Z}\} \longrightarrow \mathbb{R}, \quad x \longmapsto \tan x\\
&\sec : \mathbb{R} \setminus \{\tfrac{2n+1}{2}\pi : n \in \mathbb{Z}\} \longrightarrow \mathbb{R}, \quad x \longmapsto \sec x
\end{aligned}$$

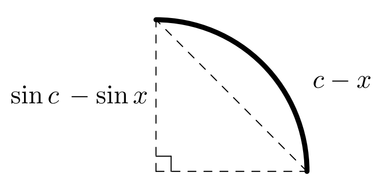
are continuous.

Proof We will only prove continuity of $\sin x$, the other functions can be dealt with in a similar manner. We measure x in radians, $360°$ is 2π radians, $180°$ is π radians and $90°$ is $\frac{\pi}{2}$ radians. In general, $x = \frac{2\pi}{360°}\theta$, where θ is the angle measured in degrees.

In the diagram below we have a unit circle, you need to know that if the angle made by two radii (each one unit) is x radians then the length of the arc of the circle between the ends of the radii is x. The circumference of the circle (the angle in this case is 2π) is just 2π — remember the radius is 1.



From our diagram we isolate the triangle shown below.



The dotted lines form a right-angle triangle. The hypotenuse of this triangle must be greater in length than either of the other two sides. In particular, it is larger than $\sin c - \sin x$. However, this hypotenuse is smaller in length than the circular arc of length $c - x$.

So for $\frac{\pi}{2} \geq x \geq 0$ we have

$$\sin c - \sin x \leq c - x.$$

Thus given any $c \in [0, \frac{\pi}{2}]$ and any $\varepsilon > 0$ we have

$$|\sin c - \sin x| < \varepsilon,$$

whenever $|c - x| < \varepsilon$, for $0 < x < c$.

This proves $\lim\limits_{x \to c^-} \sin x = \sin c$. In a similar way we can prove $\lim\limits_{x \to c^+} \sin x = \sin c$. Which then shows that $\sin x$ is continuous on $[0, \frac{\pi}{2}]$. The result is easy to extend to all values of x. ■

Using the continuity results established thus far with the results on continuity of sums and products of the last lecture we can extend our list of continuous functions in an obvious way to a large list of powers, products, of trigonometric and rational functions.

We are also now in a position to establish many limits involving these functions. A particularly useful way to establish these limits is the *squeezing principle*. It says that, if two functions have limit L as x approaches $c \in \mathbb{R}$, then any function in between must also have limit L as x approaches c.

Theorem (Squeezing Principle)

Let f, g and h be functions whose domains contain the open interval (a, b), such that

$$g(x) \leq f(x) \leq h(x)$$

for all $x \in (a, c)$ and for all $x \in (c, b)$ with $a < c < b$.

If $\lim_{x \to c} g(x) = \lim_{x \to c} h(x) = L$ then

$$\lim_{x \to c} f(x) = L.$$

Proof Let $x \in (a, b)$. From the definition of the limit we know that for any $\varepsilon > 0$ there exists a $\delta > 0$ such that

$$|g(x) - L| < \varepsilon \text{ and } |h(x) - L| < \varepsilon$$

whenever $|x - c| < \delta$. Thus

$$-\varepsilon < g(x) - L < \varepsilon \text{ and } -\varepsilon < h(x) - L < \varepsilon$$

whenever $|x - c| < \delta$. Now

$$g(x) \leq f(x) \leq h(x),$$

so

$$g(x) - L \leq f(x) - L \leq h(x) - L.$$

Combining this with the previous inequalities, we obtain

$$-\varepsilon \leq g(x) - L \leq f(x) - L \leq h(x) - L \leq \varepsilon,$$

so that, $-\varepsilon \leq f(x) - L \leq \varepsilon$, whenever $|x - c| < \delta$.

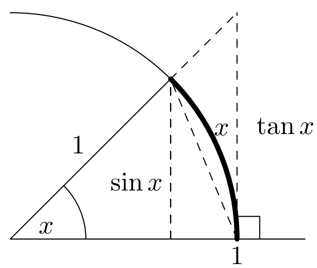
This establishes $\lim_{x \to c} f(x) = L$. □

We now use this principle to establish two very well-known limits.

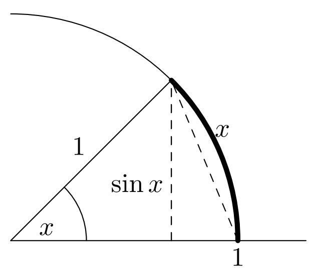
Theorem

$$\text{(a)} \lim_{x\to 0} \frac{\sin x}{x} = 1 \qquad \text{and} \qquad \text{(b)} \lim_{x\to 0} \frac{1-\cos x}{x} = 0.$$

Proof (a) We first establish "squeezing" inequalities for $\frac{\sin x}{x}$. We simply adapt the techniques used in the proof of the continuity of $\sin x$.



As before it is easy to see that $x \geq \sin x$ using the following diagram.



On the other hand comparing the area of the large triangle ($\tan x$ as one side) with the area of the sector of the unit circle subtended by the angle x we have

$$\text{Area of triangle} = \tfrac{1}{2} \cdot 1 \cdot \tan x) \geq \text{Area of sector} = \left(\tfrac{x}{2}\right),$$

so that $\tan x \geq x$. We obtain

$$\tan x \geq x \geq \sin x.$$

Now $x \geq \sin x$ so $\frac{\sin x}{x} \geq 1$ (for $x > 0$). Also $\tan x \geq x$ is just

$$\frac{\sin x}{\cos x} \geq x \quad ,$$

which can be written as

$$\frac{\sin x}{x} \geq \cos x \text{ for } x > 0,\ x \text{ near } 0.$$

Therefore

$$\cos x \leq \frac{\sin x}{x} \leq 1.$$

Notice if $x < 0$ and x is near enough to 0 then

$$\frac{\sin x}{x} = \frac{\sin(-x)}{(-x)} \geq \cos(-x) = \cos x, \quad \text{as } \sin(-x) = -\sin x.$$

The same inequality applies either side of 0.

Now,

$$\lim_{x \to 0} \cos x = 1 \quad \text{and} \quad \lim_{x \to 0} 1 = 1.$$

We can apply the squeezing principle to get the result

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

(b) We use the result from part (a), the identity $\sin^2 x = 1 - \cos^2 x$ and rewrite our expression as follows

$$\begin{aligned}
\frac{1 - \cos x}{x} &= \frac{1 - \cos x}{x} \cdot \frac{1 + \cos x}{1 + \cos x} \\
&= \frac{1 - \cos^2 x}{x(1 + \cos x)} \\
&= \frac{\sin^2 x}{x(1 + \cos x)} = \left[\frac{\sin x}{x}\right]\left[\frac{\sin x}{1 + \cos x}\right].
\end{aligned}$$

Now take limits, remember we have proved the limit of product is the product of the limits, etc.,

$$\begin{aligned}
\lim_{x \to 0}\left(\frac{1 - \cos x}{x}\right) &= \lim_{x \to 0}\left[\frac{\sin x}{x}\right] \lim_{x \to 0}\left[\frac{\sin x}{1 + \cos x}\right] \\
&= (1)\left(\frac{0}{1 + 1}\right) = 0.
\end{aligned}$$

■

♠ Exercises 11

1. Evaluate the following limits:

   (a) $\displaystyle\lim_{h\to 0}\frac{\sin h}{2h}$ (b) $\displaystyle\lim_{x\to 0}\frac{x+\sin x}{x}$

   (c) $\displaystyle\lim_{x\to 0}\frac{1-\cos^2 x}{x}$ (d) $\displaystyle\lim_{y\to 0}\frac{\sin(2y)}{y}$

   (e) $\displaystyle\lim_{x\to 0}\frac{1-\cos x}{\sin x}$ (f) $\displaystyle\lim_{h\to 0}\frac{h}{\tan h}$

   (g) $\displaystyle\lim_{x\to 0^+}\frac{\sin x}{|x|}$ (h) $\displaystyle\lim_{x\to 0^-}\frac{\sin x}{|x|}$.

2. Find the value of the constant k which makes

$$f:\mathbb{R}\longrightarrow\mathbb{R},\quad x\longrightarrow\begin{cases}\frac{\sin 2x}{x}, & x\neq 0\\ k, & x=0\end{cases}$$

   a continuous function at $x=0$.

3. Find $\lim_{x\to 1}\frac{\sin(\pi x)}{x-1}$.

4. Prove that if $f:\mathbb{R}\longrightarrow\mathbb{R}$ is a continuous function then $\sin\circ f$ is also continuous.

Lecture 1.12 More on Continuity

In this lecture we want to confirm the intuition that continuous functions on a closed interval are "nice" functions.

Bounds of a Continuous Function

Theorem *If $f : D \longrightarrow \mathbb{R}$ is continuous on $[a, b] \subseteq D$ then it is bounded on the interval, that is, $\{f(x) : x \in [a, b]\}$ is bounded.*

Remarks It is essential to the truth of the theorem that the interval is closed. The function $\frac{1}{x}$ is continuous in $(0, 1]$, but is certainly **not** bounded as $x \longrightarrow 0$. We can make $\frac{1}{x}$ arbitrarily large by taking x near enough to 0.

Proof We first prove

Lemma 1.2 *If the function $f : \mathbb{R} \to \mathbb{R}$ is continuous at a, then there is an $r > 0$ such that $\{f(x) \mid a - r < x < a + r\}$ is a bounded set of real numbers.*

Proof Since f is continuous at a, and $1 > 0$, there is an $r > 0$ with

$$f(a) - 1 < f(x) < f(a) + 1,$$

whenever $a - r < x < a + r$. □

Now put $A := \{x \in [a, b] \mid f \text{ is bounded on } [a, x]\}$. Since $a \in A$, $A \neq \emptyset$ and b is an upper bound for A. Hence, A has a supremum, $s \in [a, b]$.

Since f is continuous on $[a, b]$, it follows by the Lemma above, that there is an $r > 0$ such that f is bounded on the interval $(s - r, s + r) \cap [a, b]$, and so, in particular, on $[a, s + r)$.

Since $s = \sup A$, this is only possible if $(s, s + r) \cap [a, b] = \emptyset$. Thus $s = b$. □

This is again quite a subtle proof. Read it carefully to get the essential idea of the argument.

Given a function, $f : D \longrightarrow \mathbb{R}$, the supremum of f, $\sup f$, is the supremum of the set $\{f(x) | x \in D\}$, that is, $\sup f$ is the least upper bound of the image of f. Similarly, the infimum of f, $\inf f$, is the infimum of $\{f(x) | x \in D\}$.

♦**Example** If they exist, find sup$f$ and inf$f$ for the following:

(a) $f : (0,1) \longrightarrow \mathbb{R}, \quad x \longmapsto x^2$;

(b) $f : (0,1) \longrightarrow \mathbb{R}, \quad x \longmapsto \frac{1}{x}$;

(c) $f : [0, \frac{\pi}{2}] \longrightarrow \mathbb{R}, \quad x \longmapsto \sin x$.

Solution

(a) Here sup$f$ $= 1$ and inf$f$ $= 0$. Note that there is no $x \in (0,1)$ such that $f(x) =$ sup or $f(x) =$ inf. We say that $f(x)$ does **not** attain its bounds.

(b) Here inf$f(x) = 1$ and sup$f$ does not exist as the function is unbounded above – $\frac{1}{x} \to \infty$ as $x \longrightarrow 0$. Again we see there is no $x \in (0,1)$ such that $f(x) =$ inf$f$. The function does **not** attain its lower bound.

(c) Here sup$f = 1$ and inf$f = 0$. In this case we see that f does attain its bounds – note the interval is closed.

□

We already saw that a continuous function defined on a closed interval is bounded. Thus it has a supremum and an infimum. We now show that it must attain its bounds.

Theorem (The Extreme Value Theorem) *Let $f : [a,b] \longrightarrow \mathbb{R}$ be a continuous function with $M =$ sup f and $m =$ inf$f$. Then there exist $x_1, x_2 \in [a,b]$ such that*

$$f(x_1) = M \text{ and } f(x_2) = m.$$

Proof We prove there exists an $x_1 \in [a,b]$ such that $f(x_1) = M$. The proof for the infinum is similar.

The proof we give is by contradiction – there is another proof using the so-called method of bisection.

Proof by contradiction. Assume there is **no** $x_1 \in [a,b]$ for which $f(x_1) = M$. Then, as M is the supremum,

$$M - f(x) > 0 \text{ for all } x \in [a,b].$$

Since $M - f(x) \neq 0$ for all $x \in [a, b]$ we see, from Lecture 10, that

$$\frac{1}{M - f(x)}$$

is continuous on $[a, b]$.

So, by our previous theorem, $\frac{1}{M - f(x)}$ is bounded on $[a, b]$. We have

$$\frac{1}{M - f(x)} < K, \text{ for some finite } K \in \mathbb{R}, \text{and all } x \in [a, b].$$

Then, as $M - f(x) > 0$,

$$\frac{1}{K} < M - f(x)$$

or

$$f(x) < M - \frac{1}{K} \text{ for } x \in [a, b].$$

But this contradicts $M = \sup f(x)$, M is the **least** upper bound, there cannot be a smaller bound $M - \frac{1}{K}$! We have our proof by contradiction. ■

We have had two reasonably difficult proofs for what would seem to be two intuitively obvious facts!

♦**Example** State whether or not the following functions meet the conditions of the previous two theorems. If the function does, what are its supremum and infinum?

(a) $f : (0, 1] \longrightarrow \mathbb{R}, x \longmapsto \frac{1}{x}$;

(b) $f : [1, 2] \longrightarrow \mathbb{R}, x \longmapsto \frac{1}{x}$;

(c) $f : [-1, 1] \setminus \{0\} \longrightarrow \mathbb{R}, x \longmapsto \frac{1}{x^2}$;

(d) $f : [0, \pi] \longrightarrow \mathbb{R}, x \longmapsto \cos x$.

Solution The hypothesis for both theorems is that f is defined on a closed interval $[a, b]$ and continuous.

(a) Does not satisfy the hypothesis, the interval is not closed.

(b) Satisfies the hypothesis. Note that $\frac{1}{x}$ decreases from $x = 1$ to $x = 2$, so

$$\begin{aligned} \sup\left(\frac{1}{x}\right) &= 1, \text{ on } [1, 2] \text{ and} \\ \inf\left(\frac{1}{x}\right) &= \frac{1}{2}, \text{ on } [1, 2]. \end{aligned}$$

Note that the function attains its bounds on $[1, 2]$.

(c) The function does not satisfy the hypothesis as it is not defined on a closed interval.

(d) The function cos is continuous and restricting a continuous function to a closed interval we obtain another continuous function. Thus the hypothesis is satisfied.

$$\sup f = 1 \text{ and } \inf f = -1.$$

Monotone Functions

Consider a function $f : D \longrightarrow \mathbb{R}, D \subseteq \mathbb{R}$. We say that f is *increasing* if, given any points $x_1, x_2 \in S$ with $x_1 < x_2$, then $f(x_1) \le f(x_2)$. If we have $f(x_1) < f(x_2)$ we say f is *strictly increasing*. Reversing the inequalities on f defines a *decreasing* or *strictly decreasing* function.

> *A function which is either increasing or decreasing is said to be monotone. A function which is strictly increasing or strictly decreasing is said to be strictly monotone.*

It is easy to see that every strictly monotone function must be injective (one-one) but it is by no means true that every injective function has to be strictly monotone.

♦**Example** The following functions are monotone

(a) $\mathrm{id}_{\mathbb{R}} : \mathbb{R} \longrightarrow \mathbb{R}, x \longmapsto x;$

(b) $f : [0, \infty) \longrightarrow \mathbb{R}, x \longmapsto x^2;$

(c) $f : \mathbb{R} \longrightarrow \mathbb{R}, x \longmapsto \begin{cases} x, & x < 0 \\ 0, & 0 \le x < 1; \\ x - 1, & x \ge 1 \end{cases}$

(d) $f : (0, \infty) \longrightarrow \mathbb{R}, x \longmapsto \frac{1}{x}.$

Note that the functions in (a), (b) and (d) are strictly monotone, continuous and injective. The function in example (c) is not strictly monotone, though it is monotone. The example (c) is continuous but not injective. You should draw the graphs of these functions to convince yourself of these facts.

♦**Example** The following functions are not monotone

(a) $f : \mathbb{R} \longrightarrow \mathbb{R}, x \longmapsto |x|$;

(b) $f : \mathbb{R} \longrightarrow \mathbb{R}, x \longmapsto x^2$;

(c) $f : \mathbb{R} \longrightarrow \mathbb{R}, x \longmapsto \begin{cases} x, & x < 0 \\ x-1, & x \geq 0 \end{cases}$;

(d) $f : [0, \pi] \longrightarrow \mathbb{R}, x \longmapsto \sin x$.

The functions in (a), (b) and (d) are all continuous, but they are not injective. Hence they cannot be monotone. The function in (c) is not continuous at $x = 0$ and not monotone. It is, however, injective. In fact, the function in (c) is monotone on each of the intervals $(-\infty, 0)$ and $[0, \infty)$, separately. The same is also true for the other functions – there exist subintervals on which the (restricted) function is monotone.

Now for our main result on monotone functions.

Theorem *Every injective and continuous function $f : [a, b] \longrightarrow \mathbb{R}$ is strictly monotone.*

Proof First we observe that f cannot be constant on any subinterval – if it was there would be points x and y for which $f(x) = f(y)$, which is not possible as f is one-one. In fact it is not hard to see that if f is increasing on a subinterval it must be strictly increasing (again use the fact that f is injective). Similarly, if f is decreasing on a subinterval it must be strictly decreasing on that subinterval. Divide $[a, b]$ into subintervals on which f is either strictly increasing or strictly decreasing. If there is only one subinterval (the whole of $[a, b]$) we are finished. So we assume there are at least two. Hence there is a point $c \in (a, b)$ such that for some $\delta > 0$, the function f is either strictly increasing on $(c - \delta, c)$ and strictly decreasing on $(c, c + \delta)$ or vice versa. Let us assume the former, so

$$\begin{aligned} f(y) &> f(x) \text{ for } x, y \in (c-\delta, c) \text{ and } \; y > x, \\ \text{and } f(y) &< f(x) \text{ for } \; x, y \in (c, c+\delta) \text{ and } y > x. \end{aligned}$$

Now f is continuous and so takes all values between $f(c-\delta)$ and $f(c)$ for $x \in (c-\delta, c)$ – note $f(c) > f(c - \delta)$. Similarly f must take all values between $f(c + \delta)$ and $f(c)$ for $x \in (c, c + \delta)$ – note $f(c) > f(c + \delta)$.

Then either $f(c - \delta) > f(c + \delta)$, in which case there is an $x_1 \in (c, c + \delta)$ such that $f(c - \delta) = f(x_1)$. Or

$$f(c - \delta) < f(c + \delta)$$

in which case there is an $x_2 \in (c-\delta, c)$ such that $f(c+\delta) = f(x_2)$. Or

$$f(c-\delta) = f(c+\delta).$$

However, each of these possibilities contradicts the fact that f is injective. We can thus have only one subinterval – f is either strictly increasing on $[a, b]$ or strictly decreasing on $[a, b]$. It's monotone. □

♠ Exercises 12

1. Given functions $f, g : \mathbb{R} \longrightarrow \mathbb{R}$, suppose f is continuous at $x = a$ whereas g is discontinuous at $x = a$.

   (a) Can one conclude $f + g$ is not continuous at $x = a$?

   (b) Can one conclude fg is not continuous at $x = a$?

2. Repeat question 1 this time assuming both f and g are discontinuous at $x = a$.

3. State whether or not the following functions are monotone. Also decide if the functions are are injective, surjective or bijective. If they exist, find the supremum and infimum in each case and state whether or not the function attains its bounds.

   (a) $f : [0, 3] \longrightarrow \mathbb{R}, \quad x \longmapsto x^3;$

   (b) $f : \mathbb{R} \longrightarrow \mathbb{R}, \quad x \longmapsto \frac{1}{1+x^2};$

   (c) $f : (0, \frac{\pi}{2}) \longrightarrow \mathbb{R}, \quad x \longmapsto \sin(\frac{1}{x});$

   (d) $f : (-1, \infty) \longrightarrow \mathbb{R}, \quad x \longmapsto \frac{1-x}{1+x}.$

4*. Let $f : [a, b] \longrightarrow \mathbb{R}$ be bounded, such that, for every pair of values x_1, x_2 with $a \leq x_1 \leq x_2 \leq b$

$$f\left(\frac{x_1 + x_2}{2}\right) \leq \frac{1}{2}\left[f(x_1) + f(x_2)\right].$$

Prove f is continuous on (a, b).

Lecture 1.13 Sequences

A sequence is simply a function whose domain is $\mathbb{N}$, the natural numbers.

$$\begin{aligned} f : \mathbb{N} &\longrightarrow \mathbb{R} \\ f : n &\longmapsto f(n). \end{aligned}$$

Of course this is not the notation you would be familiar with. The usual way of specifying a sequence is to give the nth term, usually denoted by a_n. In function notation the nth term is just $f(n)$. Clearly any function $f : \mathbb{R} \longrightarrow \mathbb{R}$ can be restricted to the natural numbers to give a sequence.

♦**Example**

(a) $f(x) = \frac{1}{x^2}$, when restricted to $\mathbb{N}$ gives the sequence $a_n = \frac{1}{n^2}$.

(b) $f(x) = \sin x$, when restricted to $\mathbb{N}$ gives the sequence $a_n = \sin n$.

Another common way of specifying a sequence is to write down some of the early terms. But, unless the nth term is given, this does not uniquely specify a sequence.

♦**Example** Examples of notation for sequences

(a) $1, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \ldots, \frac{1}{n}, \ldots$

(b) $\left\{\dfrac{1}{n}\right\}_{n=1}^{\infty} = \left\{1, \dfrac{1}{2}, \dfrac{1}{3}, \ldots, \dfrac{1}{n}, \ldots\right\}$

(c) $f(n) = \frac{1}{n}$.

A sequence need not be infinite, $\{1, 2, 3\}$ is a sequence. If, for each $n \in \mathbb{N}$, we are assigned a number a_n then the sequence is said to be an *infinite sequence*.

For an infinite sequence perhaps the most interesting question is "does the sequence tend to a finite limit as $n \longrightarrow \infty$?".

♦**Example**

(a) $\left\{\dfrac{1}{n}\right\}_{n=1}^{\infty}$ tends to the limiting value 0.

(b) $\dfrac{1}{2}, \dfrac{2}{3}, \dfrac{3}{4}, \ldots, \dfrac{n}{n+1}, \ldots$ approaches the value 1.

Formal Definition of Limits

We can define a limit for a sequence in much the same way as we defined the one-sided limits for $f(x)$ as $x \longrightarrow \pm\infty$.

> *A sequence with nth term a_n is said to have the limit a if, given any $\varepsilon > 0$, there is an N such that*
>
> $$|a_n - a| < \varepsilon \quad \textit{for all} \quad n > N.$$

We write $\lim_{n \longrightarrow \infty} a_n = a$ or $a_n \longrightarrow a$ as $n \longrightarrow \infty$.

Note that if a sequence a_n does have a limit then it must be bounded. For $n > N$ we have

$$-\varepsilon < a_n - a < \varepsilon$$

$$\text{or} \qquad a - \varepsilon < a_n < a + \varepsilon.$$

Our definition of the limit can be abbreviated using some mathematical notation. We introduce two symbols which replace commonly used mathematical phrases.

$\exists \ldots$ stands for the phrase "there exists".

$\forall \ldots$ stands for the words "for all".

So we can write our definition of the limit as,

$a_n \longrightarrow a$ if

given $\varepsilon > 0$ $\exists$ N such that

$$|a_n - a| < \varepsilon, \ \forall \ n > N.$$

The symbols $\exists$ and $\forall$ are not just used in the context of sequences, they may be used in any (mathematical) situation in which the corresponding phrase is appropriate.

If a sequence does **not** tend to a finite limit it does not necessarily diverge to $\pm\infty$ – there are other types of behaviour.

♦Example

(a) The sequence $a_n = (-1)^n$ oscillates.

(b) The sequence $a_n = \sin(n)$ oscillates.

(c) The sequence $a_n = \sqrt{n}$ tends to infinity.

(d) The sequence $a_n = \sin\left(\frac{1}{n}\right)$ approaches 0.

Formally, we will define a diverging sequence as follows.

> *The sequence a_n is said to tend to infinity, or $a_n \longrightarrow \infty$ if given any $A > 0$ (no matter how large) there exists on N such that*
>
> $$a_n > A \text{ for all } n > N.$$

We have a similar definition for a sequence diverging to $-\infty$.

> *The sequence a_n is said to tend to negative infinity, or $a_n \longrightarrow -\infty$ as $n \longrightarrow \infty$ if given $A > 0$ $\exists$ N such that*
>
> $$a_n < -A \; \forall \; n > N.$$

Sequences such as $a_n = (-1)^n$ or $b_n = (-1)^n n$ do not tend to a limit or $\pm\infty$. Such sequences are said to *oscillate*.

From our earlier work on limits we would expect to have a theorem on the limit sums and products of sequences. There is such a theorem.

Theorem *Let $a_n \longrightarrow a$ and $b_n \longrightarrow b$ as $n \longrightarrow \infty$ then*

(a) $\alpha a_n + \beta b_n \longrightarrow \alpha a + \beta b, \;\forall \alpha, \beta \in \mathbb{R}$.

(b) $a_n b_n \longrightarrow ab$

(c) $\dfrac{a_n}{b_n} \longrightarrow \dfrac{a}{b}$, *provided* $b \neq 0$.

Proof The technique of proof is the same as for limits of functions. We will prove (c).

Note that

$$\begin{aligned} \left|\frac{a_n}{b_n} - \frac{a}{b}\right| &= \left|\frac{ba_n - ab_n}{bb_n}\right| \\ &= \left|\frac{b(a_n - a) + a(b - b_n)}{bb_n}\right| \\ &\leq \frac{|b(a_n - a)| + |a(b - b_n)|}{|bb_n|}, \end{aligned}$$

ie.

$$(*) \qquad \left|\frac{a_n}{b_n} - \frac{a}{b}\right| \leq \frac{|b| \cdot |a_n - a| + |a| \cdot |b_n - b|}{|bb_n|}.$$

Now $a_n \longrightarrow a$ and $b_n \longrightarrow b$ so for any given $\varepsilon > 0$ we can find an N such that

$$|a_n - a| < \frac{|b|}{4}\varepsilon, \quad |b_n - b| < \frac{|b|^2}{4|a|}\varepsilon$$

$$\begin{aligned} \text{and} \qquad |b_n| &> \frac{1}{2}|b| \\ \text{for} \qquad n &> N. \end{aligned}$$

If $|a| = 0$ then we do not need the inequality for $|b_n - b|$. Putting these last three inequalities into (*) we have

$$\left|\frac{a_n}{b_n} - \frac{a}{b}\right| < \frac{|b| \cdot \frac{|b|}{4}\varepsilon + |a| \cdot \frac{|b|^2}{4|a|}\varepsilon}{|b|\frac{1}{2}|b|} = \varepsilon.$$

$$\text{for} \qquad n > N.$$

We have formally established $\dfrac{a_n}{b_n} \longrightarrow \dfrac{a}{b}$ as $n \longrightarrow \infty$.

♦**Example** Discuss the behaviour of the following sequences as $n \longrightarrow \infty$:

(a) $\left(\dfrac{n-3}{3n+1}\right)^3$ (b) $\left(\dfrac{3n+1}{n-3}\right)^3$

(c) $\cos n$ (d) $\dfrac{(-1)^n}{n}$

Solution

(a) We need to use the trick of dividing out the n so as to apply our theorem on $\dfrac{a_n}{b_n}$.

$$\begin{aligned} \left(\frac{n-3}{3n+1}\right)^3 &= \left(\frac{1 - \frac{3}{n}}{3 + \frac{1}{n}}\right)^3 \\ &= \frac{\left(1 - \frac{3}{n}\right)^3}{\left(3 + \frac{1}{n}\right)^3}. \end{aligned}$$

Now apply the theorem to get

$$\left(\frac{n-3}{3n+1}\right)^3 \longrightarrow \frac{(1-0)^3}{(3+0)^3} = \frac{1}{27}.$$

(b) Similar to (a), $\left(\frac{3n+1}{n-3}\right)^3 = \left(\frac{3+\frac{1}{n}}{1-\frac{3}{n}}\right)^3 \longrightarrow \frac{3^3}{1^3} = 27.$

(c) The sequence oscillates through all numbers between -1 and $+1$.

(d) The sequence tends to the limit 0,

$$\text{i.e.} \quad \frac{(-1)^n}{n} \longrightarrow 0 \ \text{ as } \ n \longrightarrow \infty.$$

□

Monotone Sequences

If $a_{n+1} \geq a_n$ for all n we call a_n increasing.
If $a_n > a_n$ $\forall$ n we say a_n is strictly increasing. If $a_{n+1} \leq a_n$ $\forall$ n we say a_n is decreasing and if $a_{n+1} < a_n$ we say a_n is strictly decreasing.

A sequence which is increasing or decreasing is called *monotonic*.

Monotonic sequences have the very important property that they must tend to a limit or to $+\infty$ or to $-\infty$. Monotonic sequences cannot oscillate.

We will prove this property for an increasing sequence. This proof is easily adapted to the decreasing case.

Theorem *An increasing sequence either tends to a limit or $+\infty$ as $n \longrightarrow \infty$. An increasing sequence cannot oscillate.*

Proof Let a_n be an increasing sequence.
We have two possibilities either the sequence is bounded above or its not. That is either of the following must hold

1. a number K can be found such that

$$a_n \leq K \quad \forall n$$

or

2. for any number A we can find an N such that $a_n > A$.

Deal with possibility (2) first. As a_n is increasing $a_n > A$ not just for $n = N$ but for all $n \geq N$. So from our definition $a_n \longrightarrow \infty$ as $n \longrightarrow \infty$.

Now for (1). The number K is an upper bound for the set $\{a_1, a_2, \ldots, a_n, \ldots\} = \{a_n\}$, so the set must have supremum a which satisfies

$$\begin{aligned} a_n &\leq a \quad \forall n \text{ and} \\ a_n &> a - \varepsilon \text{ for some particular value } n = N. \end{aligned}$$

As a_n is increasing the second inequality must hold for all $n \geq N$. So we have

$$a_n < a(< a + \varepsilon) \quad \text{and}$$

$$a - \varepsilon < a_n$$

for all $n > N$. We can write this as

$$|a_n - a| < \varepsilon \quad \forall n > N.$$

Which just says a_n has limit a.

♠ Exercises 13

1. Discuss the behaviour of the sequences as $n \longrightarrow \infty$.

$$\begin{array}{lll} \text{(a)}\ \sin\left(\dfrac{1}{n}\right) & \text{(b)}\ n\sin\left(\dfrac{1}{n}\right) & \text{(c)}\ (n!)^{\frac{1}{n}} \\ \text{(d)}\ \dfrac{n^3 - n^2 + n + 1}{10 - 3n + 2n^3} & \text{(e)}\ \dfrac{n^2+1}{n} & \text{(f)}\ \dfrac{2^{n^2}}{n!} \end{array}$$

2. State which of the sequences are monotone. (Are they increasing or decreasing?)

$$\begin{array}{lll} \text{(a)}\ \dfrac{n}{n^2+1} & \text{(b)}\ (-n)^2 & \text{(c)}\ n^n \\ \text{(d)}\ \dfrac{n^2+1}{n} & \text{(e)}\ 2n + (-1)^n & \text{(f)}\ \frac{n^2-1}{n^3-1} \end{array}$$

3. Prove that if $a > 0$ then $a^{\frac{1}{n}} \longrightarrow 1$ as $n \longrightarrow \infty$.

4. In relativistic quantum mechanics the solution of the Dirac equation for the electron in the hydrogen atom gives energy levels

$$E_n = \frac{mc^2}{\sqrt{1 + \frac{\gamma^2/c^2}{\left[n + \sqrt{\kappa^2 - \gamma^2/c^2}\right]}}}$$

where m is the electron mass, γ is a constant related to the electric charge, κ is a constant related to the orbital quantum number and n is the principle quantum number. Does E_n give a monotonic sequence? Does E_n have a limit as $n \longrightarrow \infty$?

5*. Prove that $a_n \longrightarrow 0$ as $n \longrightarrow \infty$ if $\displaystyle\lim_{n\longrightarrow\infty} \frac{a_{n+1}}{a_n} = L$, where $-1 < L < 1$.

Lecture 1.14 Sequences and Series

As a direct consequence of our last theorem (Lecture 13) we have the following useful result.

Theorem *A bounded monotone sequence tends to a limit.*

Proof Direct consequence of previous theorem for an increasing sequence – case 1. The same technique works for a decreasing sequence. Try it! $\square$

The Geometric Sequence

The geometric sequence has nth term $a_n = a^n$ where a is a constant. The behaviour of the sequence as $n \longrightarrow \infty$ depends upon the value of a. There are four main cases to consider.

1. If $a = 1$, $a_n = 1$ for all n. Clearly $\lim\limits_{n \longrightarrow \infty} a_n = 1$! There is also the trivial case $a = 0$ to consider.
2. $a > 1$ let $a = 1 + k, k > 0$ then $(1+k)^n > 1 + nk$ (Exercise 4 Lecture 2). Now as $n \longrightarrow \infty$ $1 + nk \longrightarrow \infty$ and so $a_n = (1+k)^n \longrightarrow \infty$.
3. $0 < a < 1$. Let $a = \dfrac{1}{1+k}$, $k > 0$. Then $\dfrac{1}{a^n} = (1+k)^n > 1 + nk$. So
$$a^n = \frac{1}{(1+k)^n} < \frac{1}{1+nk}$$
Now $\dfrac{1}{1+nk} \longrightarrow 0$ as $n \longrightarrow \infty$ so $a_n = \dfrac{1}{(1+k)^n} \longrightarrow 0$ as $n \longrightarrow \infty$.
4. $a < 0$. Firstly, if $-1 < a < 0$ we use a similar argument to that in 3 to find $a_n \longrightarrow 0$. If $a = -1$ then $a_n = (-1)^n = \pm 1$ and the sequence oscillates finitely.
If $a < -1$ then (using a similar argument to 2) we see that the sequence oscillates infinitely.

Summing up

$$a_n = a^n, \quad a \text{ constant}$$

$$\begin{aligned}
a > 1, &\quad a_n \longrightarrow \infty \\
a = 1 &\quad a_n \longrightarrow 1 \\
-1 < a < 1 &\quad a_n \longrightarrow 0 \\
a = -1 &\quad a_n \text{ oscillates finitely} \\
a < -1 &\quad a_n \text{ oscillates infinitely.}
\end{aligned}$$

Infinite Series

Intuitively we think of an infinite series as a "sum to infinity" of an infinite sequence $\{a_n\}$:

$$a_1 + a_2 + a_3 + \ldots + a_n + \ldots \ .$$

We will use a shorthand notation for such a sum

$$\sum_{n=1}^{\infty} a_n \equiv a_1 + a_2 + a_3 + \ldots + a_n + \ldots \ ;$$

this is the *sigma or summation notation.* We just think of the $\sum$ as standing for sum over all counting number values between the given limits.

We have still not said what we mean by "sum to infinity". This term is fraught with danger and we need to say precisely what we mean by it. We will give the appropriate definition below, but first an illustration of the dangers.

Consider $\displaystyle\sum_{n=1}^{\infty}(-1)^n = -1 + 1 - 1 + 1 - 1 + 1 \ldots$. We might want to "sum" this by grouping terms as follows $(-1+1)+(-1+1)+(-1+1)+\ldots$ to conclude that the sum is zero. This would be **wrong**. To see this we could also group the terms as $-1+(1-1)+(1-1)+\ldots$, to get a sum of -1! There is clearly something wrong with the process of "grouping terms" for such a series.

We will define our infinite sums in terms of a sequence, the *sequence of partial sums.*

Let $\{a_n\}_{n=1}^{\infty}$ be an infinite sequence. We define the nth partial sum S_n of the sequence as

$$\begin{aligned} S_n &= a_1 + 1_2 + a_3 + \ldots + a_n, \quad \textit{or} \\ S_n &= \sum_{r=1}^{n} a_r. \end{aligned}$$

We can now talk about the possible limit of the sequence $\{S_n\}_{n=1}^{\infty}$.

If the sequence of partial sums $\{S_n\}$ has a finite limit as $n \longrightarrow \infty$ we say that the infinite series $a_1+a_2+a_3+\ldots+a_n+\ldots$ is **convergent** *or has a "sum to infinity" and that $S = \lim\limits_{n\longrightarrow\infty} S_n$ is its sum.*

A series which is not convergent is said to be *divergent*.

We think of the words "infinite sum" as applying only in the sense of a convergent limit of the sequence of partial sums.

The Geometric Series

The geometric series is the "sum to infinity" of the geometric sequence. When this sum exists!

Theorem *The infinite geometric series*

$$1 + a + a^2 + \ldots + a^n \ldots$$

converges if and only if $-1 < a < 1$.

Proof The series has first term 1 so the nth term is $a_n = a^{n-1}$. We note that

$$\begin{aligned} S_n &= 1 + a + a^2 + \ldots + a^{n-1} \\ \text{and} \quad aS_n &= a + a^2 + \ldots + a^n. \end{aligned}$$

So $aS_n - S_n = a^n - 1$.

Therefore $S_n = \frac{1-a^n}{1-a}$, for $a \neq 1$, if $a = 1$ $\;S_n = n$. We see that if $a = 1$ then $S_n \longrightarrow \infty$ as $n \longrightarrow \infty$. For $a \neq 1$ we have

$$S_n = \frac{1-a^n}{1-a} = \frac{1}{1-a} - \frac{1}{1-a}a^n.$$

We see that the sequence of partial sums S_n, converges if and only if the sequence a^n converges. But from our section on the geometric sequence we know a^n tends to a finite limit (which is 0) if and only if $-1 < a < 1$. This proves the theorem. Note that we also have the sum to infinity in the case $-1 < a < 1$, it is

$$S_n \longrightarrow \frac{1}{1-a} \quad \text{as} \quad n \longrightarrow \infty.$$

■

The Harmonic Series $\sum \frac{1}{n}$

Theorem *The infinite series*

$$\sum_{n=1}^{\infty} \frac{1}{n} = 1 + \frac{1}{2} + \frac{1}{3} + \ldots$$

diverges.

Proof Consider the following

$$\begin{aligned} \frac{1}{3}+\frac{1}{4} &> 2\times\frac{1}{4}=\frac{1}{2} \\ \frac{1}{5}+\frac{1}{6}+\frac{1}{7}+\frac{1}{8} &> 4\times\frac{1}{8}=\frac{1}{2} \\ \frac{1}{9}+\frac{1}{10}+\ldots+\frac{1}{15}+\frac{1}{16} &> 8\times\frac{1}{16}=\frac{1}{2}, \end{aligned}$$

and so on. The sum of successive blocks of $2, 4, 8, 16\ldots$ terms is greater than $\frac{1}{2}$. Including the first two terms $1+\frac{1}{2}$ we have that the sum of the first $2+2+4+8+16+\ldots+2^{m-1}=2^m$ terms of the series is greater than
$1+\frac{1}{2}+\frac{1}{2}\times(\text{number of blocks}) = 1+\frac{1}{2}+\frac{1}{2}(m-1)=1+\frac{1}{2}m.$
So $S_{2^m} > 1+\frac{1}{2}m$, ie. the partial sums of the first 2^m terms is greater than $1+\frac{1}{2}m$. Clearly, $S_n \longrightarrow \infty$ as $n \longrightarrow \infty$. The series diverges. ■

The Comparison Test

Theorem *If for every $n > N$ (N finite),*

(i) $u_n \geq 0,\ v_n \geq 0$

(ii) $u_n \leq kv_n,\ k>0$ *is a constant,*

then

(a) if $\sum v_n$ *converges* $\sum u_n$ *converges*

or

(b) if $\sum u_n$ *diverges* $\sum v_n$ *diverges.*

Proof

An exercise. Use the fact that the partial sums $S_n = \sum_{r=1}^{n} u_n$ and $T_n = \sum_{r=1}^{n} v_n$ form increasing sequences for $n > N$. Then use the results on monotone sequences.

♦**Example** Discuss the convergence (or divergence) of the following infinite series.

(a) $\displaystyle\sum_{n=1}^{\infty}\frac{1}{\sqrt{n}}$ (b) $\displaystyle\sum_{n=1}^{\infty}\frac{n}{3^n}$

(c) $\displaystyle\sum_{n=1}^{\infty}\frac{n}{n^2+1}$ (d) $\displaystyle\sum_{n=0}^{\infty}\frac{n2^n+1}{5^n+1}$

Solution

(a) For $n \geq 1 \dfrac{1}{\sqrt{n}} \geq \dfrac{1}{n}$

The comparison test then shows that $\sum \frac{1}{\sqrt{n}}$ diverges.

(b) For $n \geq 1 \quad n < 2^n$ so.

$$\frac{n}{3^n} < \frac{2^n}{3^n} = \frac{1}{(1.5)^n}$$

Now $\displaystyle\sum_{n=1}^{\infty} \frac{1}{(1.5)}^n$ is a convergent geometric progression, so by the comparison test $\displaystyle\sum \frac{n}{3^n}$ converges.

(c) $\dfrac{n}{n^2+1} \geq \dfrac{n}{n^2+n^2}$, for $n \geq 1$.

ie. $\dfrac{n}{n^2+1} \geq \dfrac{1}{2} \cdot \dfrac{1}{n}$

$\sum \frac{1}{n}$ is the divergent harmonic series so $\displaystyle\sum \frac{n}{n^2+1}$ diverges by the comparison test.

(d) $\dfrac{n2^n+1}{5^n+1} < \dfrac{n2^n+n2^n}{5^n+1} < \dfrac{n2^{n+1}}{5^n} < \dfrac{2^n \cdot 2^{n+1}}{5^n}$

i.e. $\dfrac{n2^n+1}{5^n+1} < \dfrac{2 \cdot (2^2)^n}{5^n} = 2\dfrac{1}{(1.25)^n}$

So comparison with the convergent geometric series $\displaystyle\sum \frac{1}{(1.25)^n}$ shows that $\displaystyle\sum \frac{n2^n+1}{5^n+1}$ converges.

$\square$

♠ Exercises 14

1. A ball is dropped from a height of 20m. On each bounce the ball returns to a height of $\frac{4}{5}$ of that of the previous bounce.

   What is height of the 3rd bounce?

   Give a formula for the height of the nth bounce.

   Give a formula for the distance travelled by the ball up to and including the nth bounce.

   What is the total distance travelled by the ball from the time it is first dropped till the time it comes to rest?

2. Discuss the convergence or otherwise of the following infinite series:

(a) $\dfrac{1}{3} + \dfrac{2}{5} + \dfrac{3}{7} + \ldots + \dfrac{n}{2n+1} + \ldots$

(b) $1 + \dfrac{1}{1} + \dfrac{1}{2 \cdot 1} + \dfrac{1}{3 \cdot \cdot 1} + \ldots + \dfrac{1}{n!} + \ldots$

(c) $\sum \dfrac{2^n + 1}{3^n + 1}$

(d) $\sum \dfrac{1}{\sqrt{n^2 + 1}}$

(e) $\sum_{k=0}^{\infty} \cos(k\pi)$

(f) $\sum \dfrac{n^2}{2^n}$.

3. Prove that if $0 \leq a_n \leq 1$, then the series $\sum_{n=0}^{\infty} (a_n x^n)$ converges for $0 \leq x < 1$.

4. If

$$a_n = \left(1 + \frac{1}{n}\right)^n \text{ and } b_n = \left(1 - \frac{1}{n}\right)^{-n}$$

prove that a_n is an increasing and b_n a decreasing sequence. Prove further that as $n \longrightarrow \infty$, $b_n - a_n \longrightarrow 0$, and that a_n and b_n lead to the same limit.

(This limit is e the base of the natural logarithms.)

Lecture 1.15 Series

In the previous lecture we showed that the harmonic series $\sum \frac{1}{n}$ diverges. We now generalise this result.

Theorem *The infinite series* $\displaystyle\sum_{n=1}^{\infty} \frac{1}{n^k}$, *$k$ constant, converges if $k > 1$ and diverges if $k \leq 1$.*

Proof We know the series diverges for $k = 1$ so for $k \leq 1$ we can use the comparison test. We have

$$\frac{1}{n^k} \geq \frac{1}{n}, \quad \text{for} \quad k \leq 1.$$

So the series diverges for $k \leq 1$.

For the case $k > 1$ we use the same device we used in proving the divergence of $\sum \frac{1}{n}$. Group terms of the series in blocks of length 2, 4, 8, ..., to get

$$\frac{1}{2^k} + \frac{1}{3^k} < \frac{2}{2^k} = \frac{1}{2^{k-1}}$$

$$\frac{1}{4^k} + \frac{1}{5^k} + \frac{1}{6^k} + \frac{1}{7^k} < \frac{4}{4^k} = \frac{1}{4^{k-1}} = \frac{1}{(2^{k-1})^2}$$

$$\frac{1}{8^k} + \frac{1}{9^k} + \ldots + \frac{1}{15^k} < \frac{8}{8^k} = \frac{1}{8^{k-1}} = \frac{1}{(2^{k-1})^3} \, .$$

The series $1 + \frac{1}{2^{k-1}} + \frac{1}{(2^{k-1})^2} + \frac{1}{(2^{k-1})^3} + \ldots$ is a convergent geometric series, (as a $0 < \frac{1}{2^{k-1}} < 1$, for $k > 1$) . The sequence of partial sums $S_n = \displaystyle\sum_{r=1}^{n} \frac{1}{r^k}$ is then bounded by the sum to infinity of our geometric series ie.

$$S_n < \frac{1}{1 - 2^{k-1}}.$$

Now S_n is an increasing sequence so our theorem on monotone sequences (Lecture 14) tells us S_n must tend to a finite limit. Hence the series $\sum \frac{1}{n^k}$ converges for $k > 1$. We are finished. ■

♦**Example** Discuss the convergence or otherwise of the following series

(a) $\displaystyle\sum \frac{n^2+1}{n^4+1}$

(b) $\sum \frac{1}{\sqrt[3]{n}}$

Solution

(a) Note $\frac{n^2+1}{n^4+1} < \frac{n^2+n^2}{n^4} = \frac{2}{n^2}$.

So, as $\sum \frac{1}{n^2}$, converges (see last theorem) we have from the comparison test that $\sum \frac{n^2+1}{n^4+1}$ converges.

(b) Straightforward application of last theorem with $k = \frac{1}{3}$.

$$\sum \frac{1}{\sqrt[3]{n}} \text{ diverges.}$$

$\square$

Properties of Infinite Series

Theorem *For infinite series the following properties must hold.*

1. *The convergence or divergence of a series is unaffected if a finite number of terms are inserted, or suppressed, or altered.*
2. *If $\sum u_n$ converges then $u_n \longrightarrow 0$ as $n \longrightarrow \infty$.*
3. *If $\sum u_n$ converges to s and $\sum s_n$ converges to t then $\sum(u_n + v_n)$ converges to $s + t$.*
4. *If $\sum u_n$ is convergent then so is any series obtained from $\sum u_n$ by bracketing the terms in any manner. The two series have the same sum.*
5. *If $u_n \geq 0$, for every n, the $\sum u_n$ either converges or diverges to $+\infty$.*

Proof

1. Changing a finite number of terms, say we replace $\sum_{n=1}^{N} a_n$ by some other number A, say. Then $\sum_{n=1}^{\infty} a_n = \sum_{n=1}^{N} a_n + \sum_{n=N+1}^{\infty} a_n$ and the new sum to infinity is

$A + \sum_{n=N+1}^{\infty} a_n$. However $\sum_{n=1}^{\infty} a_n$ converges (or diverges) if and only if $\sum_{n=N+1}^{\infty} a_n$ converges (or diverges) as $\sum_{n=1}^{N} a_n$ is finite. Hence the result.

2. Note that $u_n = S_n - S_{n-1}$ where $S_n = \sum_{r=1}^{n} u_r$ is the n^{th} partial sum.

   However, if $\sum u_n$ converges to s then $S_n \longrightarrow s$ and $S_{n-1} \longrightarrow s$ as $n \longrightarrow \infty$. The result follows.

3. Let $S_n = \sum_{r=1}^{n} u_r$ and $T_n = \sum_{r=1}^{n} v_n$. Then $\lim_{n\to\infty} S_n = s$ and $\lim_{n\to\infty} T_n = t$. However from lecture 13 we know that

$$\lim_{n\to\infty} (S_n + T_n) = s + t.$$

   But, $S_n + T_n = P_n = \sum_{r=1}^{n} (u_n + v_n)$ the n^{th} partial sum of $\sum (u_n + v_n)$. The result is proved.

4. Exercise.

5. With $u_n \geq 0$ $S_n = \sum_{r=1}^{n} u_r$ is an increasing sequence. The result then follows from lecture 13.

■

♦**Example** Discuss the convergence or otherwise of

(a) $\sum \frac{(3^n + 2^n)}{6^n}$

(b) $\sum_{n=1}^{\infty} 2^{10-n}$

Solution

(a) The series $\sum \frac{1}{2^n}$ and $\sum \frac{1}{3^n}$ both converge so the series

$$\begin{aligned} \sum \left(\frac{1}{2^n} + \frac{1}{3^n} \right) &= \sum \left(\frac{3^n + 2^n}{2^n \cdot 3^n} \right) \\ &= \sum \frac{2^n + 3^n}{6^n} \end{aligned}$$

must also converge by our last theorem.

(b)

$$\sum_{n=1}^{\infty} 2^{10-n} = \sum_{n=1}^{9} 2^{10-n} + \sum_{n=10}^{\infty} 2^{10-n}$$

$$\begin{aligned} \sum_{n=10}^{\infty} 2^{10-n} &= 1 + \frac{1}{2} + \frac{1}{2^2} + \frac{1}{2^3} + \dots \\ &= \sum_{n'=0}^{\infty} \frac{1}{2^{n'}}, \quad \text{where } n' = n - 10. \end{aligned}$$

This last series is a convergent geometric series so the original series converges, property 1 of the last theorem. □

The Ratio Test

Theorem *Let $\sum u_k$ be a series with positive terms and suppose that*

$$\rho = \lim_{n \longrightarrow \infty} \frac{u_{n+1}}{u_n}$$

(a) if $\rho < 1$, the series converges;

(b) if $\rho > 1$ the series diverges;

(c) if $\rho = 1$ the series may converge or diverge.

Proof

(a) Now $\dfrac{u_{n+1}}{u_n} \geq 0 \ \forall n \in \mathbb{N}$ so we have $0 \leq \rho < 1$ in this case.

As $\dfrac{u_{n+1}}{u_n} \longrightarrow \rho < 1$ we must have for $n \geq N$, for some N, $\dfrac{u_{n+1}}{u_n} < 1$. So there is an $r < 1$ such that for $n \geq N$, $\dfrac{u_{n+1}}{u_n} < r$. Consider the series $\sum_{n=N}^{\infty} u_n$. Now, using $\dfrac{u_{n+1}}{u_n} < r$, for $n \geq N$, we have

$$\begin{aligned} u_{N+1} &< ru_N \\ u_{N+2} &< ru_{N+1} < r^2 u_N \\ u_{N+3} &< ru_{N+2} < r^3 u_N \\ &\vdots \end{aligned}$$

In general $u_n < r^{n-N} u_N$ for $n \geq N$. The series $\sum_{n=N}^{\infty} r^{n-N} u_N$ is a convergent geometric series (as $0 < r < 1$) so by the comparison test $\sum u_n$ converges.

(b) In this case we have $\dfrac{u_{N+1}}{u_n} \longrightarrow \rho > 1$ as $n \longrightarrow \infty$. So for some N we have

$$\frac{u_{n+1}}{u_n} > 1 \quad \text{for all} \quad n \geq N.$$

From which we have

$$\begin{aligned} u_{N+1} &> u_N \\ u_{N+2} &> u_{N+1} > u_N \\ &\vdots \end{aligned}$$

In general $u_n > u_N$ for all $n \geq N$. However, the series $\displaystyle\sum_{n=N}^{\infty} u_N = u_N \sum_{n=N}^{\infty} 1 = u_N(1+1+1+\cdots)$ clearly diverges. Hence, by the comparison test $\sum u_n$ must diverge.

(c) Consider the ratio test for $u_n = \frac{1}{n^k}$, where $0 < k < \infty$. We have

$$u_{n+1}u_n = \frac{n^k}{(n+1)^k} = \left(\frac{n}{n+1}\right)^k = \frac{1}{\left(1+\frac{1}{n}\right)^k}.$$

But $\dfrac{1}{\left(1+\frac{1}{n}\right)^k} \longrightarrow 1$ as $n \longrightarrow \infty$.

So for $u_n = \frac{1}{n^k}$, $\rho = 1$. However, $\sum u_n$ diverges if $0 < k \leq 1$ and converges for $k > 1$.

■

♦**Example** Determine whether the following series converge or diverge.

$$\text{(a)} \sum_{n=1}^{\infty} \frac{n^n}{n!} \quad \text{(b)} \sum \frac{3^n}{n!} \quad \text{(c)} \sum \frac{5^n}{n^2}$$

Solution

(a) $u_n = \dfrac{n^n}{n!} = \dfrac{n^n}{n(n-1)\dots 2.1}$, so

$$\begin{aligned} \frac{u_{n+1}}{u_n} &= \frac{(n+1)^{n+1}}{(n+1)!} \cdot \frac{n!}{n^n} \\ &= \frac{(n+1)^{n+1}}{(n+1)n^n} \\ &= \frac{(n+1)^n}{n^n} = \left(1+\frac{1}{n}\right)^n \end{aligned}$$

We know from an earlier exercise that $\left(1+\dfrac{1}{n}\right)^n \longrightarrow e > 1$ as $n \longrightarrow \infty$. So the series $\sum \dfrac{n^n}{n!}$ diverges.

(b) $u_n = \dfrac{3^n}{n!}$, $\dfrac{u_{n+1}}{u_n} = \dfrac{3^{n+1}}{(n+1)!}\dfrac{n!}{3^n} = \dfrac{3}{n+1}$. So $\dfrac{u_{n+1}}{u_n} = \dfrac{3}{n+1} \longrightarrow 0$ as $n \longrightarrow \infty$. The series converges.

(c) $u_n = \dfrac{5^n}{n^2}$, $\dfrac{u_{n+1}}{u_n} = \dfrac{5n^2}{(n+1)^2} = 5\left(\dfrac{1}{1+\frac{1}{n}}\right)^2$.

So, $\dfrac{u_{n+1}}{u_n} = \dfrac{5}{\left(1+\frac{1}{n}\right)^2} \longrightarrow 5$ as $n \longrightarrow \infty$. The series diverges. □

♠ Exercises 15

Determine the convergence or otherwise of these following infinite series.

1. $\sum \dfrac{1}{5n^2 - n}$
2. $\sum \dfrac{2}{n^2+n}$
3. $\sum_{k=1}^{\infty} \dfrac{k}{k^{3/2} - \frac{1}{2}}$
4. $\sum \dfrac{n^{10}}{3^k}$
5. $\sum \dfrac{k^2}{3k^3+1}$
6. $\sum \dfrac{\sin^2 k}{k^2}$
7. $\sum \dfrac{k^2}{5^k}$
8. $\sum \dfrac{(n!)^2}{(2n)!}$