G11AN1 ANALYSIS - by Dr J.K. Langley

Section 1. Introduction

This course is about functions of a real variable. Topics will include:
Review of Sequences. Functions, Limits and Continuity. Differentiability. Power Series. Representing Functions by Power Series. Indeterminate Forms. Integration. Improper Integrals.
Suggested reading:
R. Haggarty, Fundamentals of Mathematical Analysis, Addison Wesley
Mathematical Analysis, a Straighforward Approach, K.G. Binmore, CUP.
Calculus, M. Spivak, Addison Wesley.
Mathematical Analysis, T. Apostol, Addison-Wesley (useful for all 3 years).

To establish the aims of the course, we will begin with some examples.

Examples 1.1

(i) This example appeared in G11MA1. Consider the series

$$1 - 1/2 + 1/3 - 1/4 + 1/5 - ....$$

This is an alternating series and, since $1/k$ decreases to 0, the series converges. If the sum is S, then

$$S = (1-1/2) + (1/3-1/4) + ... \quad > \quad 0.$$

However, if we write down the terms in a different order, as

$$1 - 1/2 - 1/4 + 1/3 - 1/6 - 1/8 + 1/5 - 1/10 - 1/12 + 1/7 - .....$$

(ie. one odd term followed by two even terms) then we have

$$1/2 - 1/4 + 1/6 - 1/8 + 1/10 - 1/12 + ...$$

which is

$$(1/2)(\ 1 - 1/2 + 1/3 - 1/4 + ...) \quad = \quad S/2.$$

This shows that series (or infinite sums) can produce surprising phenomena, and so must be handled with care. They cannot be avoided (a) because they are very useful and (b) because even an infinite decimal expansion is in fact a series.

(ii) Consider the differential equation, in which $y' = \mathrm{d}y/\mathrm{d}x$,

$$x^3 y'' \ + \ x^2 y' \ = \ y - x.$$

A method for solving such DEs (which very often works) is to try for a solution of the form

$$y = a_0 \ + \ + a_1 x \ + \ a_2 x^2 \ + \ a_3 x^3 \ + \ .....,$$

that is, a power series solution. *Assuming* such a solution, we differentiate term by term and get

$$y' = a_1 + 2a_2x + 3a_3x^2 + \ldots$$

and

$$y'' = 2a_2 + 3.2a_3x + 4.3a_4x^2 + \ldots\ldots.$$

Putting these together we get

$$x^3y'' + x^2y' = a_1x^2 + 2^2a_2x^3 + 3^2a_3x^4 + \ldots\ldots$$

and we set this equal to

$$y - x = a_0 + a_1x + a_2x^2 + \ldots\ldots - x.$$

This gives

$$a_0 = 0 \,,\, a_1 = 1 \,,\, a_1 = a_2 \,,\, 2^2a_2 = a_3 \,,\, 3^2a_3 = a_4 \,,\ldots\ldots,$$

from which we deduce that $a_k = ((k-1)!)^2$. Thus

$$y = \sum_{k=1}^{\infty} ((k-1)!)^2 x^k.$$

This looks very nice. However, if $x \neq 0$, then the Ratio Test gives
(modulus of $(k+1)$'th term)/(modulus of k'th term) = $k^2 |x| \to \infty$ as $k \to \infty$. Thus the series we have obtained DIVERGES for all $x \neq 0$ and is therefore not much use! The moral of this second example is that technique is not always sufficient by itself - you need also to be able to interpret the solutions you get, and this is one role of analysis. By the way, the series method used above does work for many equations - see later courses!
The aims of analysis can be broadly summarised as follows.
(i) To justify the methods of calculus, and to determine whether some procedure for solving a problem is valid, and to interpret the solutions it gives.
(ii) To study *functions* , and to discover their (often surprising) properties. This course is mainly about (i), while the Complex Analysis course in Semester 3 is more oriented towards (ii). The next idea will be fundamental.

1.2 Upper and Lower Bounds

Let A be a subset of $\mathbb{R}$. We say that $M \in \mathbb{R}$ is an upper bound for A if $x \leqslant M$ for all $x \in A$. We also say that A is bounded above. Similarly, we define lower bounds. A is just called BOUNDED if it is bounded above and below, which is the same as saying that there is some $N \geqslant 0$ such that $|x| \leqslant N$ for all $x \in A$.
Now we discuss the idea of maxima and minima. Let A be a subset of $\mathbb{R}$. We say that X is the maximum of A if $X \in A$ and $x \leqslant X$ for all x in A. This is the same as saying that X is an upper bound for A which lies in A. We write $X = \max A$. Similarly $Y = \min A$ if Y is in A and is a lower bound for A.
Note that a non-empty finite set always has a maximum and a minimum but, for example, the open interval (0,1) is bounded but has no max or min. For if t is in (0,1) then $(1+t)/2$ is in (0,1) and is greater than t.

Hence we need the following idea.

1.3. Least Upper Bounds

Suppose that a subset B of $\mathbb{R}$ has a maximum X. Then no number less than X can be an upper bound for B, so X is the LEAST upper bound of B. Now take a non-empty subset A of $\mathbb{R}$ such that A is bounded above. We've already seen that A might have no maximum, but it is a fundamental property of $\mathbb{R}$ that A has a least upper bound. We express this as the

Continuum Property of the Real Numbers

If A is a non-empty subset of $\mathbb{R}$ which is bounded above, then A has a LEAST UPPER BOUND or SUPREMUM

(denoted l.u.b.A or sup A), that is, a number $s \in \mathbb{R}$ such that

(i) s is an upper bound for A.

(ii) No number less than s is an upper bound for A, i.e. if $t < s$ then t is not an upper bound for A i.e. if $t<s$ then there exists u in A such that $u>t$.

Remarks

1. The empty set has no least upper bound, as every real number is an upper bound for the empty set.

2. The rationals $\mathbb{Q}$ do not have this property. For instance, the set $\{x \in \mathbb{Q} : x^2 < 2\}$ has least upper bound $\sqrt{2}$, but this number is not in $\mathbb{Q}$. Thus $\mathbb{Q}$ "has gaps in it", but $\mathbb{R}$ does not.

3. We cannot prove this Continuum Property of $\mathbb{R}$ just using the algebraic properties of real numbers. It is an AXIOM, or rule assumed without proof. A construction of the real numbers showing that this property holds is very involved and is described in Spivak's book. For this course, we assume this property and are able to deduce from it all the properties of sequences, functions etc. that we expect, plus a few more besides.

Although not a full proof, the following is a reasonably convincing argument for the existence of least upper bounds. Suppose that A is a non-empty set of real numbers, and that A is bounded above. Let S be the set of all real x such that x is an upper bound for A. Then S is not empty. Also, $T = \mathbb{R} \setminus S$ is not-empty, because we can take $y \in A$ and note that $y - 1$ is not an upper bound for A.

Now observe that if $y \in T$ and $x \in S$, then $y < x$. Why? Because if not we must have $y > x$, so that y would be greater than an upper bound of A and so must itself be an upper bound for A. This is a contradiction.

This makes it reasonable to conclude that, since $\mathbb{R}$ has no "gaps" in it, there exists some real s such that every number less than s is in T, and every number greater than s is in S. Is s in S? Yes, because if not then s is not an upper bound for A and so there is some v in A with $s < v$. But then $v > u = (v+s)/2 > s$, which is a contradiction since u is greater than s and so is an upper bound of A.

This is not a complete argument, since we didn't *prove* that s exists. Note, however, that if we only had rational numbers this argument certainly wouldn't work.

Greatest lower bounds: similarly, if A is a non-empty set of real numbers which is bounded below, then A has a greatest lower bound, or infimum, denoted l.u.b A or inf A. Here $s = \inf A$ means

(i) s is a lower bound for A. (ii) if $t>s$ then t is not a lower bound for A i.e. there exists $u \in A$ such that $u < t$.

We use the following convention. If A is not bounded above, we write
$\sup A = +\infty$, and if A is not bounded below, we write $\inf A = -\infty$. Note that this is just a convention, and does not mean that $+\infty, -\infty$ are real numbers (they're not!).
The following fact is useful, and will be proved in Section 2.

Theorem 1.4
Let A be a subset of $\mathbb{R}$. Then a real number s is the least upper bound of A iff the following two conditions both hold.
(i) s is an upper bound for A. (ii) There is a sequence (x_n) such that $x_n \in A$ for all n and $\lim_{n \to \infty} x_n = s$.

This is the right way to think about least upper bounds. The l.u.b. is an upper bound which is also the limit of a sequence of members of the set. It is not hard to prove the following additional fact. If A is a non-empty subset of $\mathbb{R}$ then the sup of A is $+\infty$ iff there is a sequence (x_n) of members of A such that $\lim_{n \to \infty} x_n = +\infty$.

Example 1.5
Let A and B be non-empty bounded above sets of positive real numbers with $\sup A = \alpha$, $\sup B = \beta$ and set

$$C = \{ ab : a \in A, b \in B\}.$$

Then $\sup C = \alpha\beta$.
Proof We first show that $\alpha\beta$ is an upper bound for C. Now if a,b are in A,B respectively then since a, b, α and β are positive, we have $ab \leqslant \alpha b \leqslant \alpha\beta$. Now we just take sequences (x_n) and (y_n) such that $x_n \in A$, $y_n \in B$, $x_n \to \alpha$, $y_n \to \beta$. Now $(x_n y_n)$ is a sequence in C and tends to $\alpha\beta$, by the algebra of limits (also proved in Section 2).
It is also possible to prove this without using sequences, but this is trickier (try it!).

1.6 The triangle inequality
Let x, y be real numbers. Then $|x+y| \leqslant |x| + |y|$ and $|x-y| \geqslant ||x| - |y||$.
Reminder of the Proof
For the first part, just write

$$|x+y|^2 = (x+y)^2 = x^2 + 2xy + y^2 \leqslant x^2 + 2|x||y| + y^2 = (|x| + |y|)^2.$$

For the second part, we can assume WLOG ("without loss of generality" - this means we are not reducing the validity of our proof by this assumption) that $|x| \geqslant |y|$ (if not, we can interchange x and y). Now write

$$|x| = |(x-y) + y| \leqslant |x-y| + |y|$$

so that

$$|x-y| \geqslant |x| - |y| = ||x| - |y||.$$

Section 2. Review of Sequences

Definitions 2.1

A real sequence (x_n) is a non-terminating list of real numbers

$$x_N, x_{N+1}, x_{N+2}, \ldots\ldots,$$

where N is some integer. We use the ( ) to distinguish the SEQUENCE (x_n) from the particular term x_n and the set $\{x_n : n \geqslant N\}$. We are mainly interested in whether a sequence CONVERGES, that is, whether the terms x_n get very close to some real number α as n gets large. We want a definition which expresses this idea in a clear and unambiguous way.

If we consider $a_n = 1/n$, $n \geqslant 1$, then it's fairly obvious that a_n approaches 0 as n gets large, and indeed $|a_{n+1} - 0| < |a_n - 0|$ holds for all $n \geqslant 1$. This might lead us to use the definition:

" The sequence (x_n) converges to α if x_n gets closer and closer to α as n increases."

However, the following example shows that this won't do. Let $b_n = 1/n$ if n is even, and $b_n = 1/n^2$ if n is odd. It's again clear that b_n approaches 0 as n gets large, but it's not true this time that $|b_{n+1} - 0| < |b_n - 0|$ for all $n \geqslant 1$.
So we need a definition which encapsulates the idea that x_n will be close to α, and indeed within any prescribed distance from α, for *all* sufficiently large n, but *not* suggesting that x_{n+1} is closer to α than x_n.

In the G11MA1 course, the following definition was used. (x_n) converges to α if the following is true. If some positive integer k is specified, then $|x_n - \alpha| = 0$ correct to k decimal places, for ALL large enough n. This certainly implies that $|x_n - \alpha| < 10^{-k}$ for all large enough n. This definition has the drawback that is is rather unwieldy for proving theorems, and so for this course, we use another (equivalent) definition.

Definition

The real sequence (x_n) converges to the real number α if the following is true. Given any positive real number ε, we can find an integer n_0 such that $|x_n - \alpha| < \varepsilon$ for all $n \geqslant n_0$.

Note that n_0 is allowed to depend on ε and almost certainly will. The definition is saying that, for any given $\varepsilon > 0$, then $|x_n - \alpha| \geqslant \varepsilon$ can only hold for FINITELY MANY n. It can be interpreted in the following way. Someone gives us a positive number ε, which is the accuracy to which x_n must be approximated by α, and we can find some integer n_0 such that x_n really is within ε of α for all $n \geqslant n_0$. This definition does not mean that $|x_{n+1} - \alpha| < |x_n - \alpha|$ for all n. It just means that if n is large enough, then $|x_n - \alpha|$ will be less than the pre-assigned positive number ε.

The new definition is equivalent to that of G11MA1. For if the new definition holds, then we choose $\varepsilon = 10^{-k-1}$ and we'll have $|x_n - \alpha| < 10^{-k-1}$ for all $n \geqslant$ some n_0, which means that $|x_n - \alpha| = 0$ correct to k decimal places.
On the other hand, if the G11MA1 definition holds, and someone gives us a positive number ε, then we choose a k such that $10^{-k} < \varepsilon$. Now we find an n_0 such that for all $n \geqslant n_0$ we have $|x_n - \alpha| = 0$ correct to k decimal places, and this will certainly mean that $|x_n - \alpha| < 10^{-k} < \varepsilon$ for all $n \geqslant n_0$. The numbers ε considered here will usually be quite small, and it is a "tradition" that a small positive quantity is denoted by ε. Other notations for convergence are

$x_n \to \alpha$ as $n \to \infty$ and $\lim_{n \to \infty} x_n = \alpha$,

both of which just mean that (x_n) converges to α. We don't bother to write $n \to +\infty$ because a sequence is only defined for, say, $n \geqslant N$, so that n can only tend in one direction. A non-convergent sequence is called DIVERGENT.

If you like to think geometrically, then the next idea may help you. The sequence (x_n) converges to the real number α iff the following is true. Given any open interval U with centre α, then $x_n \in U$ for all but finitely many n. To see that this is true you need only write U in the form $(\alpha - \varepsilon\ ,\ \alpha + \varepsilon)$.

Examples

(i) $x_n = 1/n^2$. It is fairly obvious that (x_n) in this case converges to 0. However, to be logically consistent we need to check that our definition is satisfied. If $\varepsilon > 0$ is given we need $|x_n - 0| = 1/n^2 < \varepsilon$, and this is true if $n > \sqrt{1/\varepsilon}$. So if we choose n_0 to be the least integer $> \sqrt{1/\varepsilon}$, then we have $|x_n - 0| < \varepsilon$ for all $n \geqslant n_0$, as required.

(ii) $x_n = (-1)^n$. Again, it's fairly clear that this sequence diverges. Well, suppose that (x_n) does converge, to α. We need to show that there is SOME $\varepsilon > 0$ for which the definition is violated. Try $\varepsilon = 1$. Whether or not $\alpha \geqslant 0$, we are certainly going to have $|x_n - \alpha| \geqslant 1$ for infinitely many n, and this shows that the required condition is not satisfied.

In practice, we do not need to check every sequence so rigorously (pedantically?), but the value of the definition is that it enables us to prove *theorems*.

Lemma 2.2

(i) Let (x_n) be a convergent sequence, with limit α. Then (x_n) is BOUNDED, that is, there exists some $M > 0$ such that $|x_n| \leqslant M$ for all n.

(ii) Let (y_n) be a sequence and let $\beta \in \mathbb{R}$. Then (y_n) converges to β iff (if and only if) the sequence $(y_n - \beta)$ converges to 0.

Proof

(i) Take $\varepsilon = 1$ in the definition of convergence. Then we can find an n_0 such that $|x_n - \alpha| < 1$ for all $n \geqslant n_0$, which implies that $|x_n| \leqslant |\alpha| + 1$ for all $n \geqslant n_0$, by the triangle inequality. Assuming that the sequence starts with x_N, then

$$M \;=\; \max\{|x_N|,\ |x_{N+1}|,\ ,.....,|x_{n_0}|,\ |\alpha| + 1\ \}$$

will do, as $|x_n| \leqslant M$ for ALL n for which the sequence is defined.

Note that the converse to (i) is FALSE, as shown by the example $(-1)^n$, which is bounded but not convergent.

(ii) is obvious.

Lemma 2.3

Let (x_n) and (y_n) be real sequences which converge to 0, and let (z_n) be a bounded sequence. Then

(i) $(x_n z_n)$ converges to 0. (ii) $(x_n + y_n)$ converges to 0.
Part (i) applies, for example, to $(\sin n)/n$, $(1/n)e^{1/n}$.

Proof

(i) Suppose we are given some $\varepsilon > 0$. We must show that we can find an n_0 such that $|x_n z_n| < \varepsilon$ for all $n \geqslant n_0$. Now (z_n) is bounded, so there exists an $M > 0$ such that $|z_n| \leqslant M$ for all n, which makes $|x_n z_n| \leqslant |x_n|M$. So if we can make $|x_n|$ less than ε/M, we are done.

But this we can do using the convergence of (x_n). Now ε/M is a positive number, so there DOES exist some integer n_0 s.t. $n \geqslant n_0$ implies $|x_n| < \varepsilon/M$, which gives $|x_n z_n| < \varepsilon$ as required.

(ii) Again suppose we are given some positive ε. We want to make $|x_n + y_n| < \varepsilon$. Now $|x_n + y_n| \leqslant |x_n| + |y_n|$, by the triangle inequality, so it suffices to make each of $|x_n|$, $|y_n|$ less than $\varepsilon/2$.

Again we can do this. $\varepsilon/2$ is a positive number so there exists n_1 s.t. $|x_n| < \varepsilon/2$ for all $n \geqslant n_1$, and there exists an n_2 (possibly different to n_1) such that $|y_n| < \varepsilon/2$ for all $n \geqslant n_2$. If we set $n_3 = \max\{n_1, n_2\}$, then we have, for all $n \geqslant n_3$, $|x_n + y_n| < 2.\varepsilon/2 = \varepsilon$ as required.

Lemma 2.4

Suppose that (y_n) is a sequence of non-zero real numbers converging to a non-zero limit $\beta \in \mathbb{R}$. Then the sequence $(1/y_n)$ is bounded.

Proof

We know that for n large, y_n is close to β, and we want to show that y_n is not too close to 0. If we can make $|y_n - \beta| < |\beta|/2$ then we must have $|y_n| \geqslant |\beta|/2$. You can see this either by drawing a picture and noting that the distance from 0 to β is $|\beta|$, or by writing

$$|\beta| = |\beta - y_n + y_n| \leqslant |\beta - y_n| + |y_n|.$$

So we proceed as follows. The number $|\beta|/2$ is positive, so there is some n_0 s.t. $n \geqslant n_0$ implies $|y_n - \beta| < |\beta|/2$, which implies that $|y_n| \geqslant |\beta|/2$ so that $|1/y_n| \leqslant 2/|\beta|$. If our sequence starts with y_N then

$$M = \max\{ 1/|y_N|, 1/|y_{N+1}|, ,....., 1/|y_{n_0}|, 2/|\beta| \}$$

is such that $|1/y_n| \leqslant M$ for all $n \geqslant N$.

Now we can prove an important result.

Theorem 2.5 The Algebra of Limits

Suppose that (x_n) , (y_n) are real sequences converging to α, β respectively. Then
(i) $(x_n + y_n)$ converges to $\alpha + \beta$.
(ii) $(x_n y_n)$ converges to $\alpha\beta$.
(iii) For any real λ, the sequence (λx_n) converges to $\lambda\alpha$.
(iv) $(|x_n|)$ converges to $|\alpha|$.
(v) If $\beta \neq 0$, then $(1/y_n)$ converges to $1/\beta$.

Proof By 2.2, $(x_n - \alpha)$ and $(y_n - \beta)$ converge to zero. So by 2.3, $(x_n - \alpha + y_n - \beta)$ converges to 0, and this sequence is $(x_n + y_n - (\alpha + \beta))$. Using 2.2 again, we see that $(x_n + y_n)$ converges to $\alpha + \beta$. This proves (i).

To prove (ii), we need to show that $x_n y_n - \alpha\beta \to 0$ as $n \to \infty$. We write

$$x_n y_n - \alpha\beta = (x_n - \alpha) y_n + \alpha(y_n - \beta).$$

Now $x_n - \alpha$ and $y_n - \beta$ both tend to 0, while (y_n) and α are both bounded. Now the result follows from 2.2 and 2.3.

(iii) Just use (ii), with (y_n) equal to the constant sequence all of whose terms are λ.

(iv) This is easy, as $||x_n| - |\alpha|| \leqslant |x_n - \alpha|$. So to make $||x_n| - |\alpha|| < \varepsilon$, we just need to make $|x_n - \alpha| < \varepsilon$, which we know is true for all large enough n.

(v) This time we first note that we cannot have $y_n = 0$ for infinitely many n and we write

$$(1/y_n) - (1/\beta) = (\beta - y_n)/\beta y_n.$$

We know that $\beta - y_n \to 0$. Also $y_n\beta$ tends to the non-zero limit β^2, and so $(1/\beta y_n)$ is a bounded sequence, by 2.4. So, by 2.3, $(1/y_n) - (1/\beta)$ tends to 0, and we are done, using 2.2.

Corollaries 2.6

(i) A convergent sequence has only one limit.

(ii) Suppose (x_n) converges to β and that α is a real number s.t. $x_n \geqslant \alpha$ for all n. Then $\beta \geqslant \alpha$.

Proof

(i) Suppose that (y_n) converges to a and to b. Then $y_n - a$ tends to 0, and so does $y_n - b$, which implies, by 2.5, that $(y_n - a) - (y_n - b)$ tends to 0 also. But the last sequence is just the constant sequence $(b - a)$. So $b = a$.

(ii) We know that $\lim_{n \to \infty} (x_n - \alpha) = \beta - \alpha$, so that $\lim_{n \to \infty} |x_n - \alpha| = |\beta - \alpha|$. But $x_n - \alpha = |x_n - \alpha|$, so $\beta - \alpha = |\beta - \alpha|$.

Another useful result is

Theorem 2.7. The Sandwich Theorem (or Squeeze Lemma)

Suppose that (a_n), (b_n), (c_n) are sequences s.t. for each n we have $a_n \leqslant b_n \leqslant c_n$, and supppose further that (a_n) and (c_n) converge to $\alpha \in \mathbb{R}$. Then (b_n) converges to α.

Proof

Suppose we are given a positive real number ε. We know there exists some n_1 such that $n \geqslant n_1$ implies that $|a_n - \alpha| < \varepsilon$, so that $a_n \in (\alpha - \varepsilon, \alpha + \varepsilon)$. Similarly there is some n_2 such that $c_n \in (\alpha - \varepsilon, \alpha + \varepsilon)$ for all $n \geqslant n_2$. So if $n \geqslant n_0 = \max\{n_1, n_2\}$ then b_n also lies in $(\alpha - \varepsilon, \alpha + \varepsilon)$ which gives $|b_n - \alpha| < \varepsilon$ as required.

Example 2.8

Suppose $|x| < 1$. Then $\lim_{n \to \infty} x^n = 0$.

Proof

We can write $|x| = 1/y = 1/(1+u)$, where $u > 0$. But for n a positive integer, the binomial theorem gives $(1 + u)^n \geqslant 1 + nu > nu$. So $-1/nu \leqslant x^n \leqslant 1/nu$ and $1/nu \to 0$ as $n \to \infty$. Thus x^n tends to 0, by the sandwich theorem.

This is a convenient point to recall and prove

Theorem 1.4

Let A be a non-empty set of real numbers. Then a real number s is the LUB of A iff the following two conditions both hold.

(i) s is an upper bound for A. (ii) There is a sequence (x_n) such that (x_n) converges to s and $x_n \in A$ for all n.

Proof

Suppose first that conditions (i) and (ii) hold. We have to show that any number $< s$ is not an upper bound for A. So take $t < s$ and put $\varepsilon = (s-t)/2$. Then for all large n we have $|x_n - s| < \varepsilon$, which implies that $x_n > t$, for if $x_n \leqslant t$ we would have $\varepsilon > |s - x_n| = s - x_n \geqslant s - t = 2\varepsilon$. Since x_n is in A, we see that t is not an upper bound for A.

To prove the converse, suppose that $s \in \mathbb{R}$ is the LUB of A. Now $s - 1/n < s$ so that $s - 1/n$ is not an upper bound for A, and so we can choose an $x_n \in A$ s.t. $s - 1/n < x_n < s$. Now (x_n) converges to s by the sandwich theorem.

Theorem 2.9 Existence of *m*'th roots

Let $X > 0$ and let m be a positive integer. Then there exists $Y > 0$ such that $Y^m = X$.

Proof

Let $A = \{ x \geqslant 0 : x^m \leqslant X \}$. Then A is not empty, as $0 \in A$. Also

$$(1+X)^m = 1 + mX + ... > X$$

so $1+X$ is an upper bound for A. Now we use a standard real analysis trick. We let $Y = \sup A$, and we claim that $Y^m = X$.

By Theorem 1.4 we can take a sequence (x_n) such that x_n is in A, which implies that $x_n^m \leqslant X$, and such that $x_n \to Y$. But then $x_n^m \to Y^m$, so that $Y^m \leqslant X$, by 2.6.

To show that $Y^m \geqslant X$, we put $y_n = Y + 1/n$. Then y_n is not in A, so $y_n^m > X$. But $y_n \to Y$, so $y_n^m \to Y^m$, giving $Y^m \geqslant X$.

2.10 Types of Divergence

Consider

$$a_n = (-1)^n \ , \ b_n = n^2 \ , \ c_n = (-1)^n n \ .$$

All three diverge, but (b_n) is better behaved than the others. As n gets large, b_n gets large and positive, while a_n and c_n do not approach anything. We say

$$\lim_{n \to \infty} x_n = +\infty$$

(or (x_n) diverges to $+\infty$) if x_n gets large and positive as n gets large. This means that if we choose any positive real number M, we will have $x_n > M$ for all sufficiently large n. So our precise definition is

Definition The real sequence (x_n) diverges to $+\infty$ if the following is true. Given any positive real number M, we can find an integer n_0 s.t. $x_n > M$ for all $n \geqslant n_0$.

So for any given positive number M, no matter how large M might be, there are only finitely many n such that $x_n \leqslant M$.
This definition does not assume or imply that $x_{n+1} > x_n$. For example, the sequence

$$x_n = n \ (\ n \text{ odd}\ )\ ,\ x_n = n^2 \ (\ n \text{ even}\ )\ ,$$

diverges to $+\infty$ but is not strictly increasing.
We also write $\lim_{n \to \infty} x_n = +\infty$. Similarly we say $\lim_{n \to \infty} x_n = -\infty$ if $\lim_{n \to \infty} (-x_n) = +\infty$.

2.11 Monotone sequences

Let (x_n) be a real sequence. We say (x_n) is
strictly increasing for $n \geqslant N$ if $x_{n+1} > x_n \ \forall n \geqslant N$,
non-decreasing for $n \geqslant N$ if $x_{n+1} \geqslant x_n \ \forall n \geqslant N$,
non-increasing for $n \geqslant N$ if $x_{n+1} \leqslant x_n \ \forall n \geqslant N$,
strictly decreasing for $n \geqslant N$ if $x_{n+1} < x_n \ \forall n \geqslant N$.
If (x_n) is any of the above it is called monotone. For a monotone sequence there are just three possibilities.
(i) (x_n) converges. (ii) (x_n) diverges to $+\infty$. (iii) (x_n) diverges to $-\infty$.
To prove this we need only consider the case where (x_n) is non-decreasing.

Theorem 2.12 The Monotone Sequence Theorem
Let (x_n) be a sequence which is non-decreasing for $n \geqslant N$. Then

$$\lim_{n \to \infty} x_n = \sup\{\ x_n : n \geqslant N\ \}.$$

Proof
Denote the set $\{x_n : n \geqslant N\}$ by A. Suppose first that A is not bounded above, so that $\sup A = +\infty$ by our convention. This means that if we are given some positive number M, then no matter how large M might be, we can find some member of the set A, say x_{n_1}, such that $x_{n_1} > M$. But then, because the sequence is non-decreasing, we have $x_n > M$ for all $n \geqslant n_1$, and this is precisely what we need in order to be able to say that $\lim_{n \to \infty} x_n = +\infty$.

Now suppose that A is bounded above, and let s be the sup. Suppose we are given some positive ε. Then we need to show that $|x_n - s| < \varepsilon$ for all sufficiently large n. But we know that $x_n \leqslant s$ for all n, so we just have to show that $x_n > s - \varepsilon$ for all large enough n.

This we do as follows. The number $s - \varepsilon$ is less than s and so is not an upper bound for A, and so there must be some n_2 such that $x_{n_2} > s - \varepsilon$. But then $x_n > s - \varepsilon$ for all $n \geqslant n_2$, and the proof is complete.

Example 2.13
Determine, for different values of x_0, the behaviour of the sequence (x_n) given by

$$x_{n+1} = 2x_n/(1 + x_n^2).$$

Solution

Obviously if $x_0 = 0$ then x_n is always 0. Consider now the case where x_0 is positive, which certainly implies that x_n is positive for all n. We first identify the possible limits of the sequence. If $x_n \to \alpha$, then so does x_{n+1}, and the algebra of limits gives $2\alpha/(1+\alpha^2) = \alpha$, giving $\alpha = 0, \pm 1$. Obviously if $x_0 = 1$ then $x_n = 1\ \forall n$, so we look at the cases $x_0 \in (0,1)$, $x_0 \in (1,+\infty)$ separately.

If $x_n \in (0,1)$ then $x_{n+1}/x_n = 2/(1+x_n^2) > 1$ so that $x_{n+1} > x_n$. Also $x_{n+1} < 1$, because $2x < x^2+1$ for all $x \neq 1$. So we see that if $x_0 \in (0,1)$ then x_n increases and is bounded above by 1. Thus (x_n) converges. The limit can't be 0 or -1, and so must be 1.

If $x_0 > 1$, we see that $x_1 \in (0,1)$, and so the sequence again converges to 1. Finally, if $x_0 < 0$, we see from the symmetry of the sequence that (x_n) must converge to -1.

2.14 Subsequences

We know that the sequence given by $a_n = (-1)^n$ is divergent. However, if we just take the EVEN n, we get 1,1,1,.... which obviously converges. What we've done is to form a sequence (b_n) given by $b_n = a_{2n}$.
Given a sequence (x_n), a SUBSEQUENCE is a sequence formed by taking some elements of the sequence (x_n), IN THE RIGHT ORDER. Thus, given

$$1, 2, 3, 4, 5, 6, ...$$

then

$$1, 3, 5, 7, 9, ...$$

is a subsequence, but

$$1, 5, 3, 9, 7, ...$$

is not. In general, given a sequence

$$x_N,\ x_{N+1},\ x_{N+2},\ .....,$$

then if we choose integers k_1, k_2, ... s.t.

$$N \leqslant k_1 < k_2 < k_3 < .....,$$

the sequence (y_n) given by $y_n = x_{k_n}$ is a subsequence of (x_n). We saw that $(-1)^n$ has a convergent subsequence. Does the sequence $(\sin n)$? (Answer later.) First we prove

Theorem 2.15

If (x_{k_n}) is a subsequence of (x_n) and $\lim_{n\to\infty} x_n$ exists, then $\lim_{n\to\infty} x_{k_n}$ exists and is the same.

Proof

We just do the case of a finite limit, the other cases being similar. Suppose $x_n \to \alpha$ and we are given some $\varepsilon > 0$. Then we know there is some N such that $|x_n - \alpha| < \varepsilon$ for all $n \geqslant N$. But we can choose some Q such that $k_Q \geqslant N$, and then $k_n \geqslant N$ for all $n \geqslant Q$. Thus $|x_{k_n} - \alpha| < \varepsilon\ \forall n \geqslant Q$, which is what we need.

Now we will answer our question about $(\sin n)$.

Theorem 2.16

Let $x_m, x_{m+1},....$ be any real sequence. Then
(i) If (x_n) is not bounded above, (x_n) has a subsequence with limit $+\infty$.
(ii) If (x_n) is not bounded below, (x_n) has a subsequence with limit $-\infty$.
(iii) If (x_n) is bounded, (x_n) has a convergent subsequence.
Part (iii) is called the Bolzano-Weierstrass theorem.

Proof

(i) Here (x_n) is not bounded above. We first note that for any positive N, there are infinitely many n s.t. $x_n > N$. For if not, then $x_n \leqslant N$ for $n \geqslant n_1$, say, and we have

$$x_n \leqslant \max\{x_m, x_{m+1}, ..., x_{n_1}, N\} \quad \forall n,$$

which is a contradiction. So we construct a subsequence as follows. We choose $k_1 \geqslant m$ s.t. $x_{k_1} > 1$, and $k_2 > k_1$ s.t. $x_{k_2} > 2$, and $k_3 > k_2$ s.t. $x_{k_3} > 3$, and so on. The reason that we specify that $k_2 > k_1$ etc. is to ensure that terms from the original sequence occur in the right order.

(ii) Just apply (i) to the sequence $(-x_n)$.

(iii) Here (x_n) is bounded, say $|x_n| \leqslant M$ for all n. We use a trick, and set

$$z_n = \sup\{ x_p : p \geqslant n \}.$$

Then $z_n \geqslant -M$ for all n, because the x_p are all at least $-M$. Also $z_{n+1} \leqslant z_n$ for all n. This is because $\{x_p : p \geqslant n+1\}$ is a subset of $\{x_p : p \geqslant n\}$ and, if A is a subset of B, then the sup of B is an upper bound for B and so for A, which means that $\sup A \leqslant \sup B$. Thus (z_n) is a non-increasing sequence which is bounded below, and so converges. We denote the limit by S. (S is sometimes called the limsup of (x_n)).
We shall show that (x_n) has a subsequence converging to S. To do this we first make a claim.

Claim

Given any positive real number ε, there are infinitely many p s.t. $|x_p - S| < \varepsilon$.

Proof of Claim

Because ε is positive, we can find some n_1 s.t. $|z_n - S| < \varepsilon$ for all $n \geqslant n_1$, which means that $x_n \leqslant z_n < S + \varepsilon$ for all $n \geqslant n_1$. So $x_p < S + \varepsilon$ for all $p \geqslant n_1$. Also, for all $n \geqslant n_1$, we have $z_n > S - \varepsilon$. This means that $S - \varepsilon$ is not an upper bound for $\{x_p : p \geqslant n\}$, so that there must be some $p \geqslant n$ s.t. $x_p > S - \varepsilon$. In summary, we have proved that for each $n \geqslant n_1$ there exists a $p \geqslant n$ such that $S - \varepsilon < x_p < S + \varepsilon$, and this proves the Claim.

Now we construct our subsequence with limit S. We choose k_1 s.t. $|x_{k_1} - S| < 1$. Then we choose $k_2 > k_1$ s.t. $|x_{k_2} - S| < 1/2$. Now we choose $k_3 > k_2$ s.t. $|x_{k_3} - S| < 1/3$ etc.

So this tells us that ($\sin n$) *does* have a convergent subsequence. How to find it is another story! There is one more concept we need to consider with regard to sequences.

2.17 Cauchy sequences

We can think of a convergent sequence as one in which the elements of the sequence get close to some number, the limit. A Cauchy sequence is one in which the elements of the sequence get close to EACH

OTHER.

The definition is as follows. We say that the real sequence (x_n) is Cauchy if, given any positive number ε, we can find an integer n_1 s.t. $|x_n - x_m| < \varepsilon$ for all n, m which are both $\geqslant n_1$.

Thus we can make the elements of the sequence as close as we need to each other. In fact we have

Theorem A real sequence (x_n) is Cauchy iff it is convergent.

One might therefore ask: why bother with Cauchy sequences? First, because it is often easier to prove that a sequence is Cauchy than to prove it converges (for which you need to know what the limit is). Second, Cauchy sequences become important later in other contexts, where they are not always convergent.

Proof of the Theorem

Suppose first that (x_n) converges to a, and that we are given some $\varepsilon > 0$. If we can make $x_n,\ x_m$ both within $\varepsilon/2$ of a, then they will be within ε of each other. We can do this. $\varepsilon/2$ is positive, so there exists some n_1 such that $n > n_1$ implies $|x_n - a| < \varepsilon/2$. Therefore if $n,m \geqslant n_1$ we have $|x_n - x_m| \leqslant |x_n - a| + |a - x_m| < \varepsilon$.

Now suppose that (x_n) is Cauchy. We first note that (x_n) must be bounded. For there exists an n_2 s.t. $n,m \geqslant n_2$ implies that $|x_n - x_m| < 1$. This tells us that $|x_n - x_{n_2}| < 1$ for all $n \geqslant n_2$, so that $|x_n| \leqslant |x_{n_2}| + 1$ for all $n \geqslant n_2$. Thus if the sequence starts from x_q,

$$M = \max\{|x_q|,\ |x_{q+1}|,\ ,....,\ |x_{n_2-1}|,\ |x_{n_2}|+1\ \}$$

is such that $|x_n| \leqslant M$ for all n.

Now we know that (x_n) is bounded, the Bolzano-Weierstrass theorem gives us a convergent subsequence, (x_{k_n}), say, with limit b, say. Suppose we are now given some positive ε. We know there exists an n_3 s.t. that for all $n,m \geqslant n_3$ we have $|x_m - x_n| < \varepsilon/2$. Choose a $k_Q > n_3$ s.t. $|x_{k_Q} - b| < \varepsilon/2$. This we can do since (x_{k_n}) converges to b. Now we find that for all $m \geqslant n_3$ we have

$$|x_m - b| \leqslant |x_m - x_{k_Q}| + |x_{k_Q} - b| < \varepsilon/2 + \varepsilon/2 = \varepsilon.$$

This tells us that x_m tends to b as m tends to ∞, which is what we need.

This concludes what we need about sequences, but this is a convenient place to talk about countability.

Countability

This is an important idea when deciding how "big" infinite sets are compared to each other. We shall see that $\mathbb{R}$ is a "bigger" set than $\mathbb{Q}$. We use A,B etc. to denote subsets of $\mathbb{R}$, although many of the ideas in this section will work for other types of sets. We say that A is **countable** if either A is empty or there is a sequence (a_n) , $n=1,2,3,....$, which "uses up" A, by which we mean that each $a_n \in A$ and each member of A appears at least once in the sequence.

FACT 1: Any finite set is countable. If $A = \{ x_1 , \ldots , x_N \}$, just put $a_n = x_n$ if $n \leqslant N$, and $a_n = x_N$ if $n > N$.

FACT 2: If $B \subseteq A$ and A is countable, then B is countable. If B is empty, this is obvious. If B is not empty, then nor is A, so take a sequence which uses up A, and delete all entries in the sequence which don't belong to B. We then get either finitely many entries left (so that B is finite), or we get an infinite subsequence which uses up B.

FACT 3: Suppose that A is an infinite, countable set. Then there is a sequence (b_n) , n=1,2,..., of members of A in which each member of of A appears *exactly once*. To see this, suppose that (a_n) , n=1,2,... uses up A. Go through the list, deleting any entry which has previously occurred. So if $a_n = a_j$ for some $j < n$, we delete a_n. The resulting subsequence includes each member of A exactly once. We have thus arranged A into a sequence - first element, second element etc. - hence the name "countable" .

FACT 4: Suppose that A_1 , A_2 , A_3 , ... are countably many countable sets. Then the union $U = \bigcup_{n=1}^{\infty} A_n$, which is the set of all x which each belong to at least one A_n, is countable.
Proof Delete any A_j which are empty, and re-label the rest. Now suppose that the j'th set A_j is used up by the sequence $(a_{j,n})$, n=1,2,..... Write out these sequences as follows:
$a_{1,1}$ $a_{1,2}$ $a_{1,3}$ $a_{1,4}$ ....
$a_{2,1}$ $a_{2,2}$ $a_{2,3}$ $a_{2,4}$ ....
$a_{3,1}$ $a_{3,2}$ $a_{3,3}$ $a_{3,4}$ .... etc.
Now the following sequence uses up all of U. We take $a_{1,1}$, $a_{1,2}$, $a_{2,1}$, $a_{1,3}$, $a_{2,2}$ $a_{3,1}$, $a_{1,4}$,......

FACT 5: The set of positive rational numbers is countable. The reason is that this set is the union of the sets $A_m = \{ p/m : p \in \mathbb{N} \}$, each of which is countable. Similarly, the set of negative rational numbers is countable (union of the sets $B_n = \{ -p/n : p \in \mathbb{N} \}$), and so is $\mathbb{Q}$ (union of these sets and { 0 }).

FACT 6: If A and B are countable sets, then so is the Cartesian product $A \times B$, which is the set of all ordered pairs (a, b), with $a \in A$ and $B \in B$. Here "ordered" means that $(a, b) \neq (b, a)$ unless $a = b$. This is obvious if A or B is empty. Otherwise, if (a_n) uses up A and (b_n) uses up B, then $A \times B$ is the union of the sets $C_n = \{ (a_n , b_m) : m = 1,2,3,... \}$, each of which is countable.

FACT 7: the interval (0, 1) is not countable, and therefore nor are $\mathbb{R}$, $\mathbb{C}$. **Proof** Suppose that the sequence (a_n) , n=1,2,..., uses up (0, 1). Write out each a_j as a decimal expansion not terminating in ...999999' . Suppose this gives
$a_1 = 0.b_{1,1}b_{1,2}b_{1,3}b_{1,4}$.........
$a_2 = 0.b_{2,1}b_{2,2}b_{2,3}b_{2,4}$.........
$a_3 = 0.b_{3,1}b_{3,2}b_{3,3}b_{3,4}$.........
etc. Here each digit $b_{j,k}$ is one of 0,1,2,...9. We make a new number $x = 0.c_1c_2c_3c_4$.... as follows. We look at $b_{n,n}$. If $b_{n,n} = 4$, we put $c_n = 5$, while if $b_{n,n} \neq 4$, we put $c_n = 4$. Now x cannot belong to the list above, for if we had $x = a_m$, then we'd have $c_m = a_{m,m}$, which isn't true.

The above facts give you a useful guide to which sets you can be sure are countable and which definitely aren't. Question: let S be the set of all sequences (q_n) , n=1,2,....., with the entries q_n rational numbers. S can be thought of as $\mathbb{Q} \times \mathbb{Q} \times \mathbb{Q} \times \mathbb{Q} \times$..... . Is S countable?

Answer: no, because every real number in (0, 1) can be identified, by means of its decimal expansion, with a sequence of integers (namely, the digits). So the set of sequences with integer entries is not countable, and so nor is S.

Section 3. Functions, limits and continuity

Definitions 3.1

All functions in this course will be real-valued, each defined on some subset of $\mathbb{R}$. Let A and B be subsets of $\mathbb{R}$. A function $f:A \to B$ is a rule which assigns to each x in A a unique value $f(x)$ in B. We say that A is the DOMAIN of f (the set of points at which f is defined). By the RANGE of f we mean the set $f(A) = \{ f(x) : x \in A\}$ (or the IMAGE of A under f). However, this term "range" is sometimes defined slightly differently.
The GRAPH of f is the set $\{(x, f(x)) : x \in A\}$. Note that this is a subset of $\mathbb{R}\times\mathbb{R}$.
We say that $f:A \to B$ is SURJECTIVE or ONTO if $f(A) = B$ i.e. if for every y in B there is at least one x in A such that $f(x) = y$. We say that f is INJECTIVE or 1-1 (one-one, one-to-one) on A if f takes different values at different points i.e. if the following holds. For all x_1, x_2 in A, $f(x_1) = f(x_2)$ implies that $x_1 = x_2$.

3.2 Some Important Classes of Function

1. powers of x. The powers x^n for $n \in \mathbb{N}$ are defined inductively by $x^0 = 1$, $x^{n+1} = x^n.x$. For n a negative integer, we just put $x^n = (1/x)^{-n}$ (with domain $\mathbb{R}\setminus\{0\}$). Rational powers of x can be defined for $x > 0$ by $x^{p/q} = (x^{1/q})^p$. It is routine to check that this defines $x^{p/q}$ unambiguously. Note that we have not yet defined x^α for α irrational.
2. Polynomials and Rational Functions. If n is a non-negative integer, and $a_0,....,a_n$ are real numbers, then $P(x) = \sum_{k=0}^{n} a_k x^k$ is a polynomial. The degree of P is the largest k for which $a_k \neq 0$. If P and Q are polynomials such that Q is not identically zero, that is, Q has at least one non-zero coefficient, then $R(x) = P(x)/Q(x)$ is a rational function. The domain of R is the set of x for which $Q(x) \neq 0$.

3. Other Functions. The exponential and logarithm functions will be defined later. The trigonometric functions will be used, but properties such as $(\mathrm{d}/\mathrm{d}x)(\sin x) = \cos x$ will be proved in Section 6.

3.3 The limit of a function as x tends to $+\infty$.

Let f be a real-valued function defined on some interval $[B, +\infty)$, and let L be a real number. We say that $f(x) \to L$ as $x \to +\infty$ if $f(x)$ gets close to L as x gets large and positive. This means that $|f(x)-L|$ gets small and, for any positive number (denoted by ε, say) which might be given, we will have $|f(x)-L| < \varepsilon$ for all sufficiently large positive x, that is, for all $x >$ some positive number X, which may depend on ε.
So our precise, unambiguous definition will be:

Definition

Let $L \in \mathbb{R}$. We say that $f(x) \to L$ as $x \to +\infty$, or $\lim_{x \to +\infty} f(x) = L$, if the following is true. For any positive real number ε we can find a positive real number X, which may depend on ε, such that $|f(x)-L| < \varepsilon$ for all $x > X$.

We say that $\lim_{x \to +\infty} f(x) = +\infty$ if the following holds. Given any positive real number M we can find a positive real number X such that $f(x) > M$ for all $x > X$.

These definitions are very close to those for sequences, so that the following theorem is not surprising.

Theorem 3.4

Let $L \in \mathbb{R}$. We have $\lim_{x \to +\infty} f(x) = L$ iff $\lim_{n \to \infty} f(x_n) = L$ for every sequence (x_n) which diverges to $+\infty$.

Similarly, $\lim_{x \to +\infty} f(x) = +\infty$ iff $\lim_{n \to \infty} f(x_n) = +\infty$ for every sequence (x_n) which diverges to $+\infty$.

Proof

Suppose first that $\lim_{x \to +\infty} f(x) = L$, and that $x_n \to +\infty$. Suppose we are given some $\varepsilon > 0$. We need to show that there exists some integer n_0 such that $|f(x_n) - L| < \varepsilon$ for all $n \geqslant n_0$.

Now we know that there is some positive number X such that $|f(x) - L| < \varepsilon$ for all $x > X$. But we also know that $x_n \to +\infty$, so there must be some n_0 such that $x_n > X$ for all $n \geqslant n_0$, and this is all we need.

To prove the converse, suppose that it's not true that $\lim_{x \to +\infty} f(x) = L$. (Note that I do not say: "suppose that $\lim_{x \to +\infty} f(x) \neq L$" - we have to also allow for the case where the limit doesn't exist.)

So we have to look at what it means for our limit definition *not* to hold. It means that there must be *some* positive number ε for which we cannot find any X with the property that $x > X$ implies that $|f(x) - L| < \varepsilon$. So for any X you try, there must be some $x > X$ such that $|f(x) - L| \geqslant \varepsilon$. In particular, for each positive integer n there must be some $x_n > n$ such that $|f(x_n) - L| \geqslant \varepsilon$. Now this sequence (x_n) diverges to $+\infty$, but $(f(x_n))$ doesn't converge to L.

So we've just proved that "the first statement false" implies "the second statement false", so that the second statement implies the first. The proof with L replaced by $+\infty$ is very similar.

3.5 The limit as x tends to $-\infty$.

This is easy to deal with, just by saying that $\lim_{x \to -\infty} f(x)$ will be $\lim_{x \to +\infty} f(-x)$, if the second limit exists.

Obviously we also say that $\lim f(x) = -\infty$ if $\lim(-f(x)) = +\infty$.

3.6 The limit of a function as x tends to $a \in \mathbb{R}$.

Since a real number a can be approached from above or below we first separate these two possibilities. If L is real and f is a real-valued function defined on some interval $(a, B]$ we say that $f(x) \to L$ as x tends to a from above if $|f(x) - L|$ gets small as x approaches a from above, which means that for any positive number ε which might be given, we will have $|f(x) - L| < \varepsilon$ for all x sufficiently close to, but greater than a, that is, for all x in some interval $(a, a + \delta)$, with $\delta > 0$. Here δ is allowed to depend on ε and, if ε is small, then δ will usually also be small. Again, the choice of the greek letter δ to denote a small quantity is "traditional".

Definition

Let a and L be real numbers. We say that $f(x) \to L$ as x tends to a from above, or $\lim_{x \to a+} f(x) = L$, if the following holds. Given any positive number ε, we can find some positive number δ such that $|f(x) - L| < \varepsilon$ for all x such that $a < x < a + \delta$.

We say that $\lim_{x \to a+} f(x) = +\infty$ if the following holds. Given any positive number M, we can find a positive number δ such that $f(x) > M$ for all x such that $a < x < a+\delta$.

It is very important to note that we have not said anything about $f(a)$. The existence or value of $f(a)$ makes NO DIFFERENCE to the existence or value of the limit. A result connecting this definition with sequences is the following.

Theorem 3.7
Let a and L be real numbers. Then $\lim_{x \to a+} f(x) = L$ (respectively, $\lim_{x \to a+} f(x) = +\infty$) iff for every sequence (x_n) which converges to a and satisfies $x_n > a$ for all n, we have $\lim_{n \to \infty} f(x_n) = L$ (respectively $\lim_{n \to \infty} f(x_n) = +\infty$).
Proof
This time we'll do the case of an infinite limit. So suppose that $\lim_{x \to a+} f(x) = +\infty$, and (x_n) converges to a with $x_n > a$ for all n. If we are given some $M > 0$, we have to show that there is some n_0 such that $f(x_n) > M$ for all $n \geqslant n_0$.
Now we know there is some $\delta > 0$ such that $a < x < a+\delta$ implies $f(x) > M$. But then there must be some integer n_0 such that $|x_n - a| < \delta$ for all $n \geqslant n_0$, and this means that $a < x_n < a+\delta$ for all $n \geqslant n_0$, which gives us what we need.
To prove the converse, suppose that it's not true that $\lim_{x \to a+} f(x) = +\infty$. This means that there must be some $M > 0$ for which we cannot find any $\delta > 0$ with the property that $a < x < a+\delta$ implies that $f(x) > M$. So for any positive δ you try, there must be some x in $(a, a+\delta)$ such that $f(x) \leqslant M$. In particular, for each positive integer n, there must be some x_n with $a < x_n < a+1/n$ and $f(x_n) \leqslant M$. Now $x_n \to a$, but $f(x_n)$ doesn't tend to $+\infty$.

Definitions 3.8
We say $\lim_{x \to a-} f(x) = L \in \mathbb{R}$ if the following holds. Given any $\varepsilon > 0$ we can find a $\delta > 0$ such that $|f(x) - L| < \varepsilon$ for all x with $a-\delta < x < a$. Similarly, we say $\lim_{x \to a-} f(x) = +\infty$ if given any $M > 0$ we can find $\delta > 0$ such that $f(x) > M$ for all x such that $a-\delta < x < a$.
There is also a two-sided limit. We say that $\lim_{x \to a} f(x) = L$ (where L can here be finite or infinite) if $\lim_{x \to a+} f(x)$ and $\lim_{x \to a-} f(x)$ both exist and are L.

Theorem 3.9 The Algebra of Limits
Let lim stand for any one of $\lim_{x \to +\infty}$, $\lim_{x \to -\infty}$, $\lim_{x \to a+}$, $\lim_{x \to a-}$, $\lim_{x \to a}$.
Let L,M be real numbers and suppose that $\lim f(x) = L$, $\lim g(x) = M$. Then
(i) $\lim\, (f(x)+g(x)) = L+M$
(ii) $\lim\, (f(x)g(x)) = LM$
(iii) for any real λ, $\lim \lambda f(x) = \lambda L$
(iv) if $M \neq 0$, then $\lim 1/g(x) = 1/M$

(v) $\lim |f(x)| = |L|$.

To prove these, we can just take appropriate sequences. If "lim" means $\lim_{x \to +\infty}$, take any sequence (x_n) which diverges to $+\infty$. Then $f(x_n) \to L$, $g(x_n) \to M$, and all the results follow from the algebra of limits for sequences. This is in fact the main reason for including Theorems 3.4 and 3.7.

Example 3.10 A function with no limits at all

Set $f(x) = 1$ if x is rational and $f(x) = -1$ if x is irrational. This is a perfectly good function but it is worth noting that you cannot draw its graph. If you try to, you end up with what seem to be two horizontal straight lines, which is clearly not allowed.

To see, for instance, that $\lim_{x \to 0+} f(x)$ doesn't exist, just put $x_n = \sqrt{1/n}$. Then x_n tends to 0 from above, but $f(x_n)$ is 1 for infinitely many n and -1 for infinitely many n. In the opposite direction we have:

3.11 A class of functions for which all one-sided limits exist

Let I be any interval (it could be $[a,b]$, $(a,b]$, $(-\infty,b]$, any interval at all), and let f be a real-valued function defined on I. We say:

f is strictly increasing on I if $f(x)<f(y)$ for all x,y in I with $x<y$

f is non-decreasing on I if $f(x) \leqslant f(y)$ for all x,y in I with $x<y$

f is non-increasing on I if $f(x) \geqslant f(y)$ for all x,y in I with $x<y$

f is strictly decreasing on I if $f(x)>f(y)$ for all x,y in I with $x<y$

If any of the above hold, we say that f is monotone on I. Now we show that these functions always have one-sided limits.

Theorem 3.12

Let f be a non-decreasing function on (a,b). Then $\lim_{x \to a+} f(x)$, $\lim_{x \to b-} f(x)$ both exist.

If $a < c < b$, then $\lim_{x \to c-} f(x) \leqslant f(c) \leqslant \lim_{x \to c+} f(x)$.

Similarly, if f is non-decreasing on $(a,+\infty)$ then $\lim_{x \to +\infty} f(x)$ exists.

Proof

These proofs are all easy, once we've decided what the limit should be. For the first part, let

$L = \sup \{ f(x): a < x < b \}$, with the usual convention that L is $+\infty$ if the set is not bounded above. Now it turns out that $\lim_{x \to b-} f(x) = L$.

Why? Suppose first that L is $+\infty$. Then for any $M > 0$ there must be some t in (a,b) such that $f(t) > M$. If we put $\delta = b-t$, then $b-\delta < x < b$ implies $t < x < b$ so that $f(x) \geqslant f(t) > M$, which is exactly what we need.

Similarly, if L is a finite sup, suppose we are given some $\varepsilon > 0$. Because L is a sup, there must be some t in (a,b) with $f(t) > L-\varepsilon$. Again we put $\delta = b-t$, and $b-\delta < x < b$ implies that $f(x) \geqslant f(t) > L-\varepsilon$. But we also have $f(x) \leqslant L$, so $|f(x)-L| < \varepsilon$ for $b-\delta < x < b$.

Similarly, $\lim_{x \to a+} f(x) = \inf \{ f(x): a < x < b \}$. Also,

$\lim_{x \to c-} f(x) = \sup \{ f(x): a < x < c \} \leqslant f(c) \leqslant \inf \{ f(x): c < x < b \} = \lim_{x \to c+} f(x)$.

Note however, that if $g(x) = x$ for $x < 0$ and $g(x) = 1$ for $x \geqslant 0$ then $\lim_{x \to 0} g(x)$ (the two-sided limit) fails to exist.

Definitions 3.13

Let a be a real number. A neighbourhood of a is any open interval (c, d) which contains a.

A real valued function f is called continuous at a if $\lim_{x \to a+} f(x) = \lim_{x \to a-} f(x) = f(a)$.

Obviously if f is continuous at a then f must be defined on a neighbourhood of a.

Example

Set $f(x) = x^2$ if x is rational, and $f(x) = -x^2$ if x is irrational. Then f is continuous at 0. Why? Clearly $f(0) = 0$. Suppose $\varepsilon > 0$ is given. We will have $|f(x) - 0| < \varepsilon$ provided $x^2 < \varepsilon$, and this is true if $-\sqrt{\varepsilon} < x < \sqrt{\varepsilon}$.

This example shows that continuity at a point has nothing to do with being able to draw the graph of the function without a jump. The graph of this function cannot be drawn at all. The example $g(x) = x \sin(1/x)$, $g(0) = 0$ is also interesting.

Theorem 3.14 An alternative definition of continuity

Let f be defined, real-valued, on a neighbourhood of a. Then f is continuous at a if and only if the following is true.

For any given positive ε, we can find a positive δ such that $|f(x) - f(a)| < \varepsilon$ for all x with $|x - a| < \delta$.

(This is sometimes used as the definition of continuity).

Proof

Suppose first that the second statement is true, and we are given some $\varepsilon > 0$. Choose a δ as in the second statement. Then $a - \delta < x < a$ or $a < x < a + \delta$ implies that $|f(x) - f(a)| < \varepsilon$, so that $\lim_{x \to a} f(x) = f(a)$.

To prove the converse, again suppose we are given $\varepsilon > 0$. We know that $\lim_{x \to a+} f(x) = f(a)$ so that there must be some $\rho > 0$ such that $a < x < a + \rho$ implies $|f(x) - f(a)| < \varepsilon$. Similarly there must be some $\sigma > 0$ such that $|f(x) - f(a)| < \varepsilon$ for $a - \sigma < x < a$. Put $\delta = \min\{\rho, \sigma\}$. Then $|x - a| < \delta$ implies that $|f(x) - f(a)| < \varepsilon$.

Theorem 3.15 Algebra of continuous functions

Suppose that f and g are continuous at a. Then so are $f+g, fg, |f|, \lambda f$ (for any real λ) and $1/g$ (if $g(a) \neq 0$).

So a polynomial is continuous on all of $\mathbb{R}$, and a rational function P/Q is continuous at any point where $Q \neq 0$.

These facts just follow from (3.9).

Theorem 3.16

Let lim stand for any of $\lim_{x \to +\infty}$, $\lim_{x \to a+}$ etc. Suppose that $\lim f(x) = b \in \mathbb{R}$, and suppose that g is continuous at b. Then $\lim g(f(x)) = g(b)$.

Proof

We can use sequences again. Suppose, for instance, that $f(x) \to b$ as $x \to +\infty$, and suppose (x_n) is a sequence which diverges to $+\infty$. Then we know that $f(x_n) \to b$. Now if $\varepsilon > 0$ there is some $\delta > 0$ s.t. $|y-b| < \delta$ implies that $|g(y)-g(b)| < \varepsilon$. But we will have $|f(x_n)-b| < \delta$ for all $n \geqslant$ some n_0.
So we've shown that $|g(f(x_n))-g(b)| < \varepsilon$ for all $n \geqslant n_0$. Thus $\lim_{n \to \infty} g(f(x_n)) = g(b)$, which proves the theorem.

Corollary 3.17
If f is continuous at a and g is continuous at $f(a)$ then $g(f)$ is continuous at a.

3.18 Continuity on a closed interval
We often deal with functions which are only defined on a closed interval, and the following definition is convenient.
We say that $f:[a,b] \to \mathbb{R}$ is continuous on $[a,b]$ if f is continuous at each c in (a,b) and
$\lim_{x \to a+} f(x) = f(a), \quad \lim_{x \to b-} f(x) = f(b).$
If this definition is satisfied, the following is true. If (x_n) is a convergent sequence such that $a \leqslant x_n \leqslant b$ for all n, and $\beta = \lim_{n \to \infty} x_n$, then $\lim_{n \to \infty} f(x_n) = f(\beta)$.
Proof
We know that β is in $[a,b]$. Suppose $\varepsilon > 0$ is given. We need to show that there is some n_0 such that $|f(x_n) - f(\beta)| < \varepsilon$ for all $n \geqslant n_0$.
First suppose that $\beta = b$. Then we know there is some $\delta > 0$ such that $b-\delta < x \leqslant b$ implies $|f(x) - f(b)| < \varepsilon$. But we have $|x_n - b| < \delta$ (and hence $b-\delta < x_n \leqslant b$) for all sufficiently large n, and this does it.
The proof for $\beta = a$ is the same. If $\beta = c \in (a,b)$, the proof is easier. We know there is some $\delta > 0$ such that $|x-c| < \delta$ implies $|f(x) - f(c)| < \varepsilon$, and we know that $|x_n - c| < \delta$ for all sufficiently large n.

Now we can prove two important theorems. Both use the existence of suprema.

Theorem 3.19
Let f be continuous, real-valued on $[a,b]$. Then there exist α, β in $[a,b]$ such that

$$f(\alpha) \leqslant f(x) \leqslant f(\beta) \ \forall x \in [a,b],$$

that is, f has a maximum and minimum on $[a,b]$.
Proof
Let $A = \{f(x): x \in [a,b]\}$, and let L and M be the inf and sup of A respectively, with the usual convention if A is not bounded above or not bounded below. Now there must be a sequence of elements y_n of A such that $y_n \to M$ as $n \to \infty$. When M is a real number, this follows from Theorem 1.4, while if M is $+\infty$ it is easy to see that such y_n exist. We choose $x_n \in [a,b]$ such that $f(x_n) = y_n$. Then (x_n) is a sequence in $[a,b]$, and so has, by the Bolzano-Weierstrass theorem, a subsequence (x_{k_n}) converging to a limit β in $[a,b]$. But then $f(x_{k_n}) \to f(\beta)$ as $n \to \infty$ and $f(x_{k_n}) \to M$. Thus $f(\beta)$=M. The proof of the existence of α is the same.

Remarks

1. Theorem 3.19 does not hold for open intervals, as the example of $f(x) = 1/x$ on (0, 1) shows.
2. The function $f:[-1,1] \to \mathbb{R}$ defined by $f(x) = 0$ if x is rational and $f(x) = x$ if x is irrational has no max or min on [-1, 1]. This shows that the hypothesis that f is continuous *is* necessary in Theorem 3.19.

Theorem 3.20 The intermediate value theorem

Let f be continuous on $[a,b]$. If $f(a) < \lambda < f(b)$, or $f(b) < \lambda < f(a)$, then there exists c in (a,b) such that $f(c) = \lambda$.

Proof

We assume that $f(a) < \lambda < f(b)$. In the other case, we can just use $-f$ and $-\lambda$.
Let $A = \{ x \in [a,b]: f(x) \leqslant \lambda \}$. Now A is not empty, as a is in A. Set $c = \sup A$.
First we show that $f(c) \leqslant \lambda$. To see this, we can take a sequence (x_n) such that $x_n \in A$ and $x_n \to c$ as $n \to \infty$. Thus $f(x_n) \leqslant \lambda$, but $f(x_n) \to f(c)$ as $n \to \infty$.
Now we prove that $f(c) \geqslant \lambda$. Now $c < b$, because $f(b) > \lambda$. So we can take the sequence (y_n) given by $y_n = c + 1/n$, and $y_n \in [a,b]$ for all large enough n. Also, y_n is not in A, since $y_n > c$, which means that $f(y_n) > \lambda$. But $f(y_n) \to f(c)$ as $n \to \infty$, which gives $f(c) \geqslant \lambda$.

Theorem 3.21

Let I be any interval (closed, open, half-open etc.) and suppose that f is continuous and one-one on I. Then f is either strictly increasing or strictly decreasing on I.

Proof

We first prove this for I a closed interval $[a,b]$. Suppose that $f(a) < f(b)$. Then we assert that f is strictly increasing on I. Suppose that this is not the case, so that there exist x, y with $a \leqslant x < y \leqslant b$ such that $f(x) \geqslant f(y)$, which implies that $f(x) > f(y)$. We consider two cases.

Case 1 $f(y) < f(a)$

Then by (3.20) there must be some c in (y,b) such that $f(c) = f(a)$, contradicting the fact that f is 1-1.

Case 2 $f(y) \geqslant f(a)$

In this case $f(y) > f(a)$, as $y \neq a$. Moreover, $f(x) > f(y) > f(a)$, which means that there must be some d in (a,x) such that $f(d) = f(y)$.
Thus both cases lead to a contradiction and f must be strictly increasing on I.
If $f(a) > f(b)$ we apply the same argument to $-f$ to see that f is strictly decreasing on I.
Now suppose that we have *any* interval I and f is not strictly increasing or strictly decreasing on I. Then there must exist t, u, v, w in I such that $t < u$, $v < w$, but $f(t) < f(u)$ and $f(v) > f(w)$. Now we just choose a *closed* interval J contained in I such that t, u, v, w all belong to J. By the first part this is impossible.

The converse of this theorem is not true, as an increasing function need not be continuous. However, if the function is also *onto*, we can prove the following. It's convenient to work with open intervals here.

Theorem 3.22

Let I and J be open intervals (not necessarily bounded) and let $f: I \to J$ be non-decreasing and onto. Then f is continuous on I.

Proof

Choose any β in I, and take $\varepsilon > 0$. Because $f(\beta)$ lies in the open interval $J = f(I)$, we can find x_1 and x_2 such that $f(\beta)-\varepsilon < f(x_1) < f(\beta) < f(x_2) < f(\beta)+\varepsilon$. Because f is non-decreasing we must have $x_1 < \beta < x_2$. Now put $\delta = \min \{ \beta - x_1, x_2 - \beta \}$. If $|x-\beta| < \delta$ then $x_1 < x < x_2$, so that $f(x_1) \leqslant f(x) \leqslant f(x_2)$, which implies that $|f(x) - f(\beta)| < \varepsilon$.

Another way to prove Theorem 3.22 is to look back at Theorem 3.12 and to ask what follows if f is *not* continuous at β.

3.23 Inverse functions

Let A be a subset of $\mathbb{R}$ and let $f:A \to \mathbb{R}$ be a function. Let $B = f(A) = \{ f(x) : x \in A\}$. A function $g:B \to A$ is called the inverse function of f if $g(f(x))=x$ for all x in A. We usually write $g = f^{-1}$ (N.B. not the same as $1/f$).

If $g = f^{-1}$ exists then f is one-one on A, for if $f(x_1) = f(x_2)$ then $g(f(x_1)) = g(f(x_2))$ and so $x_1 = x_2$.

Also, if f is one-one on A then clearly g will exist, by defining $g(y)$ to be the unique x in A such that $f(x) = y$.

Theorem 3.24

Let f be one-one and continuous on an open interval I, where I might be (a,b) or $(a,+\infty)$ or $(-\infty,a)$ or $\mathbb{R}$. Then f has an inverse function $g = f^{-1}$ and g is continuous on $f(I)$.

Proof

We know already that f^{-1} exists. Also f is either strictly increasing or strictly decreasing on I.

Suppose first that f is strictly increasing on I. Let L and M be the inf and sup of $f(I)$, respectively. Then $f(I) = (L,M)$, by the intermediate value theorem.

Why? Clearly $f(x) < M$ for all x in I, because if not we could take some $t > x$ in I and we would have $f(t) > M$, which is not allowed. So $f(I)$ is contained in (L,M). Also, if $c \in (L,M)$, we can take points u, v in I such that $f(u) < c < f(v)$, so the value c must be taken.

Set $J = (L,M)$. Now $g = f^{-1}$ exists on J and g is an onto function from J to I. Also g is strictly increasing. Thus g is continuous by 3.22.

If f is strictly decreasing on I, put $h(x) = -f(x)$, and set $J = h(I) = \{-f(x) : x \in I\}$. By the first part there is a continuous function $g:J \to I$ such that $g(h(x)) = x$ for all x in I. We define G on $f(I)$ by $G(y) = g(-y)$. Then G is continuous on $f(I)$ and for each x in I, we have $G(f(x)) = g(-f(x)) = g(h(x)) = x$, so that $G = f^{-1}$.

Section 4 Differentiability

Definition 4.1

We say that the real-valued function f is differentiable at $a \in \mathbb{R}$ if there exists a real number $f'(a)$ such that

$$f'(a) = \lim_{x \to a} \frac{f(x) - f(a)}{x - a}.$$

Remarks

1. f must be defined on a neighbourhood of a for the definition to make sense.

2. We also write $f'(a) = (\mathrm{d}f/\mathrm{d}x)(a)$. Our definition defines a new function f' whose domain is the set of points at which f is differentiable. f' is called the derived function, or derivative, of f. Further, $f''(a) = (f')'(a)$ or, as an alternative notation, $f^{(n+1)}(a) = (f^{(n)})'(a)$, with $f^{(0)} = f$.
3. We can rewrite the definition of differentiability as

$$\frac{f(x)-f(a)}{x-a} = f'(a) + \varepsilon(x),$$

where $\varepsilon(x) \to 0$ as $x \to a$. Putting $\varepsilon(a) = 0$ we get the following.
The real-valued function f is differentiable at a iff there is a function $\varepsilon(x)$ such that:
(i) $\varepsilon(a) = 0$;
(ii) $\varepsilon(x)$ is continuous at a (so that $\lim_{x \to a} \varepsilon(x) = 0$);
(iii) for all x in some neighbourhood U of a we have

$$f(x) = f(a) + f'(a)(x-a) + \varepsilon(x)(x-a).$$

The last formula can be interpreted as follows. To approximate $f(x)$ for x near a, we can use the *linear* function $g(x) = f(a)+f'(a)(x-a)$, and this approximation will be very good if x is close enough to a. Thus differentiability is really about whether you can approximate $f(x)$ by a linear function. The graph of the function g is called the tangent line of f at a. The formulation (iii) also has the advantage that you can generalise it to higher dimensions.

We also see at once from (iii) that $f(x) \to f(a)$ as $x \to a$. So we have proved:

Theorem 4.2 If the real-valued function f is differentiable at a, then f is continuous at a.

The converse is false, as the example $f(x) = |x|$, $a = 0$ shows.

More Examples
1. Define $f(x) = x^2\sin(1/x^2)$ for $x \neq 0$, with $f(0) = 0$. For $x \neq 0$, the product rule and chain rule (see below) give us $f'(x) = 2x\sin(1/x^2) - 2x^{-1}\cos(1/x^2)$. Does $f'(0)$ exist? Yes, because for $x \neq 0$ we have $(f(x)-f(0))/(x-0) = x\sin(1/x^2) \to 0$ as $x \to 0$. So $f'(0) = 0$. Note that $f'(x)$ is not bounded as $x \to 0$ and so not continuous at 0 so $f''(0)$ cannot exist.

2. Set $f(x) = x^2$ for $x > 0$ and $f(x) = x^3$ for $x \leqslant 0$. For $x < 0$, we can write

$\lim_{y \to x} (y^3-x^3)/(y-x) = \lim_{y \to x} (y^2+xy+x^2) = 3x^2$ so $f'(x) = 3x^2$ for $x < 0$. Similarly, $f'(x) = 2x$ for $x > 0$. Also, $f'(0) = 0$ again, as $f(0) = 0$ and $f(x)/x \to 0$ as $x \to 0$.

For $x > 0$ we get $f''(x) = 2$, and for $x < 0$ we get $f''(x) = 6x$. But $f''(0)$ does not exist, as

$\lim_{x \to 0+} (f'(x)-f'(0))/x = 2$, but $\lim_{x \to 0-} (f'(x)-f'(0))/x = 0$.

3. A function continuous everywhere and differentiable nowhere. This example was discovered by Weierstrass. It is $\sum_{n=0}^{\infty} 2^{-n}\cos(21^n \pi x)$!

4. Another example, due to van der Waerden (1930). Let n be a positive integer. For any real number x, define $f_n(x)$ to be the distance from x to the nearest rational number of the form $m/10^n$, with m an integer. Now we define

$$f(x) = \sum_{n=1}^{\infty} f_n(x).$$

Note that $|f_n(x)| < 10^{-n}$ for all x and for all $n \in \mathbb{N}$. So the sum converges, and gives a non-negative function f which is positive at a lot of points. Also f is continuous. Why? Take $\varepsilon > 0$ and $x \in \mathbb{R}$. We can choose N such that $\sum_{n=N+1}^{\infty} 10^{-n} < \varepsilon/4$. So, for all real y, we have $f(y) = \sum_{n=1}^{N} f_n(y) + h(y)$, where $|h(y)| < \varepsilon/4$. But the function $g(y) = \sum_{n=1}^{N} f_n(y)$ is continuous, as it is the sum of finitely many continuous functions. So if y is close enough to x, then $|g(y) - g(x)| < \varepsilon/2$, and this gives $|f(y) - f(x)| < \varepsilon$.

Take any real number x. We show that f is not differentiable at x. For each positive integer q, we choose a new number y_q as follows. x belongs to some interval of the form $[m/10^q, (m+1)/10^q)$, with $m \in \mathbb{Z}$, and this interval can be divided into equal "halves", namely $[m/10^q, (m+1/2)/10^q)$ and $[(m+1/2)/10^q, (m+1)/10^q)$. Now we put $y_q = x \pm 1/10^{q+1}$, with the $\pm$ chosen so that y_q lies in the same "half" as x.

Compare $f_n(x)$ and $f_n(y_q)$. If $n > q$ then since y_q is x shifted through $\pm 10^{-q-1}$, you don't change this distance to the nearest point of form (integer)$/10^n$, so $f_n(x) = f_n(y_q)$. But if $n \leqslant q$ then the nearest point of form (integer)$/10^n$ is the SAME for x and y_q, so $f_n(y_q) - f_n(x) = \pm(y_q - x)$ for $n \leqslant q$. This means that $(f(y_q) - f(x))/(y_q - x) = \sum_{n=1}^{q} \pm 1$. We don't know exactly what this value is, but it is an integer, and it must be odd if q is odd and even if q is even. Now let $q \to \infty$. We find that $y_q \to x$, but $(f(y_q) - f(x))/(y_q - x)$ cannot tend to a finite limit, as it is alternately even and odd!.

Theorem 4.3 The product rule etc.
Suppose that f and g are differentiable at a, and $\lambda \in \mathbb{R}$. Then:
(i) $(f+g)'(a) = f'(a) + g'(a)$; (ii) $(\lambda f)'(a) = \lambda f'(a)$;
(iii) $(fg)'(a) = f'(a)g(a) + f(a)g'(a)$; (iv) if $g(a) \neq 0$, then $(1/g)'(a) = -g'(a)/g(a)^2$.
Proof (i) and (ii) are easy. (iii) As $x \to a$, we have

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

$$= g(x)\left(\frac{f(x) - f(a)}{x-a}\right) + f(a)\left(\frac{g(x) - g(a)}{x-a}\right) \to g(a)f'(a) + f(a)g'(a).$$

(iv) Again, as $x \to a$, we have

$$\frac{(1/g(x)) - (1/g(a))}{x-a} = \frac{g(a) - g(x)}{(x-a)g(x)g(a)} \to -g'(a)/g(a)^2.$$

Theorem 4.4 The chain rule

If g is differentiable at a and f is differentiable at $b = g(a)$, then $h = f(g)$ is differentiable at a and $h'(a) = g'(a)f'(b)$.

Proof

It's convenient to use the formulation (iii) from 4.1. We can write

$$g(x) \quad = \quad g(a) \;+\; (x-a)(\,g'(a) \;+\; \varepsilon(x)\,)$$

where $\varepsilon(x)$ is continuous at a and $\varepsilon(a)=0$. Similarly,

$$f(y) \quad = \quad f(b) \;+\; (y-b)(\,f'(b) \;+\; \rho(y)\,)$$

where $\rho(y)$ is continuous at b and $\rho(b)=0$. We put these together as follows.
If x is sufficiently close to a then $g(x)$ will be close to b (since g is continuous at a) and so

$$h(x)-h(a) \quad = \quad f(g(x))-f(g(a)) \quad = \quad (g(x)-b)(\,f'(b) \;+\; \rho(g(x))\,) \quad =$$

$$= \quad (x-a)(\,g'(a)+\varepsilon(x))(\,f'(b) \;+\; \rho(g(x))\,) \quad = \quad (x-a)g'(a)f'(b) \;+\; (x-a)\delta(x)$$

where

$$\delta(x) \quad = \quad \varepsilon(x)f'(b) \;+\; \varepsilon(x)\rho(g(x)) \;+\; g'(a)\rho(g(x))$$

is continuous at a (using 3.17) with $\delta(a) = 0$. By (iii) of 4.1 this proves what we want.

Corollary 4.5

It follows at once that if f is one-one near a and $f'(a) = 0$, then the inverse function f^{-1} *cannot* be differentiable at $f(a)$, for otherwise we would have $1 = \mathrm{d}x/\mathrm{d}x = f'(a)(\,f^{-1})'(\,f(a)) = 0$. We seek a condition which ensures that f^{-1} *is* differentiable.

Theorem 4.6

Let f be one-one and continuous and real-valued on the open interval I. Suppose that c is in I and that f is differentiable at c with $f'(c) \neq 0$. Then $g = f^{-1}$ is differentiable at $f(c) = d$, and $g'(d) = 1/f'(c)$.

Proof

We know that g is continuous on $f(I)$, by 3.24. Take any sequence (y_n) such that (y_n) converges to d from above or from below. Then there exist points x_n in I such that $f(x_n) = y_n$ and $x_n = g(y_n) \to c$ because g is continuous. Since f is one-one, f is monotone, and x_n tends to c either from above or from below. Thus

$$\frac{g(y_n)-g(d)}{y_n-d} \quad = \quad \frac{x_n-c}{f(x_n)-f(c)} \quad \to \quad 1/f'(c)$$

as $n \to \infty$, which proves that the limit as y tends to d from above or below of $(g(y)-g(d))/(y-d)$ is $1/f'(c)$.

4.7 Local maxima

We say that the real-valued function f has a local maximum at a if there exists a neighbourhood U of a (ie. an open interval containing a) such that $f(x) \leqslant f(a)$ for all x in U. A local minimum is defined similarly. If a is a local maximum or local minimum and f is differentiable at a, then $f'(a) = 0$.

Proof Say a is a local maximum. If x is in U and $x > a$, then $(f(x)-f(a))/(x-a) \leqslant 0$, so $f'(a) \leqslant 0$. Similarly, if x is in U and $x < a$, then $(f(x)-f(a))/(x-a) \geqslant 0$, so $f'(a) \geqslant 0$.
Now we prove a key result.

4.8 The mean value theorem
If $f:[a,b] \to \mathbb{R}$ is continuous on $[a,b]$ and differentiable on (a,b), then there exists c in (a,b) such that

$$f'(c) = \frac{f(b)-f(a)}{b-a}.$$

The special case where $f(a) = f(b)$ and $f'(c) = 0$ is called Rolle's theorem.
Proof
We first prove Rolle's theorem. By Theorem 3.19 there exist α and β in $[a,b]$ such that for all x in $[a,b]$, we have $f(\alpha) \leqslant f(x) \leqslant f(\beta)$. Now if $f(\beta) > f(a)$, then $a < \beta < b$, so that f has a local maximum at β and we can take $c = \beta$. If $f(\alpha) < f(a)$ then f has a local minimum at α and we can take $c = \alpha$. Finally if $f(\alpha) = f(a) = f(\beta)$ then f is constant on (a,b) and $f'(c) = 0$ for every c in (a,b).

To prove the general case, set

$$g(x) = f(x) - (x-a)\left(\frac{f(b)-f(a)}{b-a}\right).$$

Then $g(a) = g(b) = f(a)$, and there must be some $c \in (a, b)$ such that $g'(c) = 0$.

Theorem 4.9
Let f be differentiable on $I = (a,b)$. Then:
(i) f is strictly increasing on I if $f'(x) > 0$ for all x in I:
(ii) f is non-decreasing on I iff $f'(x) \geqslant 0$ for all x in I:
(iii) f is constant on I iff $f'(x) = 0$ for all x in I:
(iv) f is non-increasing on I iff $f'(x) \leqslant 0$ for all x in I:
(v) f is strictly decreasing on I if $f'(x) < 0$ for all x in I:
Proof
(i) and (v) follow straight from the mean value theorem.
(ii) If $f' \geqslant 0$ on I, then f is non-decreasing by the mean value theorem. Conversely, if f is non-decreasing, then $\lim_{x \to c+} (f(x)-f(c))/(x-c)$ will be $\geqslant 0$, so $f'(c) \geqslant 0$. The proof of (iv) is the same.
(iii) Obviously if f is constant then $f' = 0$. Conversely, if f' is always zero, f is constant by the mean value theorem.

Remarks
1. The function $f(x) = x^3$ is strictly increasing but $f'(0) = 0$. Thus (i) is not "iff".
2. If $f'(c) > 0$ at some point c, it DOESN'T follow that f is increasing near c. For example, set $f(x) = x+x^2$ if x is rational, and $f(x) = x$ otherwise. Then $\lim_{x \to 0} f(x)/x = 1$, so $f'(0) = 1 > 0$. But f is not monotone near zero. To see this, choose a small positive rational y and an irrational x which is close to, but greater than, y. Then $f(y) > f(x)$.

The mean value theorem will have many applications later. For now we consider two examples.

Example 1

Show that $f(x) = x/(1+x^2)$ is increasing on $[0,1]$.

This is not obvious, as f is an increasing function divided by an increasing function. But, on (0, 1) we have $f'(x) = (1-x^2)/(1+x^2)^2 \; > \; 0$.

Example 2

Show that $(1+x)^{-1/2} \; > \; 1-x/2$ for $x > 0$.

Setting $g(x) = (1+x)^{-1/2} - 1 + x/2$, we have $g(0)=0$. Also, for x positive,

$$g'(x) = (-1/2)(1+x)^{-3/2} + 1/2 \; = \; (1 \; - \; (1+x)^{-3/2})/2 \; > \; 0.$$

Remark: it might be tempting to try to do this problem *algebraically*, but the method above seems easier.

Section 5 Power Series

A power series is an object of the form

$$F(x) \; = \; \sum_{k=0}^{\infty} a_k (x-c)^k \, ,$$

where the coefficients a_k and the "centre" c are real. These have many applications, for example, in approximating functions such as sin, cos etc. They can also be used for solving differential equations: the method used in Example (ii) of 1.1 *does* often give valid solutions.

The two key questions to ask about power series are: (a) for which x does the series converge? (b) what kind of a function $F(x)$ do we obtain? Is F continuous, differentiable?

Example

If we put $a_k = 1$ and $c = 0$ we get $\sum_{k=0}^{\infty} x^k$. Now $\sum_{k=0}^{n} x^k = (1-x^{n+1})/(1-x)$. If $|x| < 1$ we let $n \to \infty$ and find that $\sum_{k=0}^{\infty} x^k \; = \; 1/(1-x)$. On the other hand, if $|x| \geqslant 1$ then we know that the terms of the series do not tend to zero and so the series must diverge.

To decide where a power series converges we need something called the radius of convergence.

Definitions 5.2

For the series $F(x)$ as in 5.1, we set

$$T_F \; = \; \{ \; t \geqslant 0 : |a_k| t^k \to 0 \; as \; k \to \infty \; \}.$$

Now T_F is non-empty as 0 is in T_F, so we put $R_F = \sup T_F$, with the usual convention that $R_F = +\infty$ if T_F is not bounded above. Now we prove:

Theorem 5.3

If $|x-c| < R_F$, then the series $F(x)$ converges absolutely, that is, $\sum_{k=0}^{\infty} |a_k| |x-c|^k$ converges.

If $|x-c| > R_F$, then the series $F(x)$ diverges.

Proof

Suppose first that $|x-c| > R_F$. Then $t = |x-c|$ does not belong to the set T_F, and so $|a_k||x-c|^k$ does not tend to 0 as $k \to \infty$, which means that the series $F(x)$ cannot possibly converge. (Remember that if the series $\sum_{k=0}^{\infty} B_k$ converges, then $\lim_{k \to \infty} B_k$ must be 0).

Now suppose that $|x-c| < R_F$. Since R_F is a sup, there must be some s in T_F such that $|x-c| < s \leqslant R_F$. So $|a_k|s^k \to 0$ as $k \to \infty$, and so there must be some $M \geqslant 0$ such that $|a_k|s^k \leqslant M$ for all k. But then $|a_k||x-c|^k = |a_k|s^k(|x-c|/s)^k \leqslant M(|x-c|/s)^k$. But we know that the series $\sum_{k=0}^{\infty} M(|x-c|/s)^k$ converges, being M times a convergent geometric series (since $|x-c|/s < 1$).

To actually find the radius of convergence, we often use:

The Ratio Test

Suppose that $\lim_{k \to \infty} |b_{k+1}/b_k|$ exists and is L. If $L < 1$ the series $\sum_{k=0}^{\infty} b_k$ converges absolutely. If $L > 1$ the series diverges. If $L = 1$ the test is not conclusive.

Examples 5.4

Find *all* values of x for which the following power series converge.

(i) $\sum_{k=0}^{\infty} \frac{(x-1)^k}{k+3}$.

Applying the Ratio Test with $b_k = (x-1)^k/(k+3)$, we find that if $x \neq 1$, then $|b_{k+1}/b_k| = |x-1|(k+3)/(k+4) \to |x-1|$ as $k \to \infty$. Thus if $|x-1| < 1$ the power series converges absolutely, and if $|x-1| > 1$ the power series diverges. Obviously the series converges if $x = 1$.

The cases $x-1 = \pm 1$ must be looked at separately. If $x-1 = 1$ we have $\sum_{k=0}^{\infty} 1/(k+3)$ which diverges. Now $x-1 = -1$ gives us $\sum_{k=0}^{\infty} (-1)^k/(k+3)$ and this converges by the alternating series test, as $1/(k+3)$ is decreasing with limit 0.

So we have the interval of convergence $-1 \leqslant x-1 < 1$, or [0, 2).

The reason why the end-points are important is the following, known as *Abel's limit theorem*. If F, as given in (5.1), has radius of convergence R such that $0 < R < +\infty$, and if the power series converges at $c+R$, then $F(x) \to F(c+R)$ as $x \to c+R$ from below. Similarly, if F converges at $c-R$, then $F(x) \to F(c-R)$ as $x \to c-R$ from above. We do not need this theorem here, so we omit the (difficult) proof.

(ii) $\sum_{k=0}^{\infty} k!x^k$. This time $b_k = k!x^k$ and the Ratio Test give, for $x \neq 0$,

$|b_{k+1}/b_k| = (k+1)|x| \to +\infty$ as $k \to \infty$. Thus the power series diverges for all non-zero x. It does, however, converge when $x = 0$ (every power series converges at its centre).

(iii) $\sum_{k=0}^{\infty} \frac{x^k}{k^k}$.

You can use the Ratio Test here, but it's a bit messy. Note that if k is large enough, then $|x|/k < 1/2$. Since $\sum_{k=0}^{\infty} (1/2)^k$ converges, our power series is absolutely convergent for all x.

5.5 Differentiating a Power Series

We want to know where (if at all) the power series $F(x) = \sum_{k=0}^{\infty} a_k (x-c)^k$ is differentiable. If we differentiate each term separately (term-by-term), we get the series

$$G(x) = \sum_{k=1}^{\infty} k a_k (x-c)^{k-1} = \sum_{k=0}^{\infty} (k+1) a_{k+1} (x-c)^k.$$

This, however, does not prove anything in itself. As a first step we prove the following:

Theorem 5.6

With F and G as in 5.5, we have $R_F = R_G$.

Proof

First we show that $R_G \leqslant R_F$. Recall that $R_F = \sup \{ t \geqslant 0 : |a_k| t^k \to 0 \ \ as \ \ k \to \infty \}$. If $R_G = 0$ the conclusion is obvious. Now suppose that $0 < t < R_G$. Then as $k \to \infty$ we have $|(k+1) a_{k+1}| t^k \to 0$, so that $|a_{k+1}| t^k \to 0$, and so does $|a_{k+1}| t^{k+1}$. So $t \leqslant R_F$. Since this holds for any t such that $0 < t < R_G$, we must have $R_G \leqslant R_F$.

Now we show that $R_F \leqslant R_G$. Again if $R_F = 0$ it's obvious. Now suppose that $0 < t < R_F$. Then we can take s such that $t < s < R_F$ and $|a_k| s^k \to 0$ as $k \to \infty$. This is because R_F is a sup. Now we find that

$$|(k+1) a_{k+1}| t^k = |a_{k+1}| s^{k+1} (k+1)(t/s)^k / s.$$

Now, as $k \to \infty$, $|a_{k+1} s^{k+1}| \to 0$. But $(k+1)(t/s)^k/s$ also tends to 0. This is because the series $\sum_{k=0}^{\infty} (k+1)(t/s)^k/s$ is convergent, as is easily checked by the Ratio Test.

So we have shown that $0 < t < R_F$ implies that t belongs to the set T_G (see 5.2) and so $t \leqslant R_G$. Thus $R_F \leqslant R_G$.

Now we can prove:

Theorem 5.7 The differentiation theorem for power series

Suppose that the power series $F(x) = \sum_{k=0}^{\infty} a_k (x-c)^k$ has *positive* radius of convergence $R = R_F$. Then in the open interval $I = (c-R, c+R)$ we have

$$F'(x) = G(x) = \sum_{k=0}^{\infty} (k+1) a_{k+1} (x-c)^k = \sum_{k=1}^{\infty} k a_k (x-c)^{k-1} .$$

In particular F is continuous on I. Further, F has derivatives of all orders on I, and $a_k = F^{(k)}(c)/k!$.

Proof

We assume without loss of generality that $c = 0$. We know that the power series G also has radius of convergence R and that G is obtained from F by differentiating each term $a_k x^k$ separately with respect to x. If we do this to G, we obtain

$$H(x) \;=\; \sum_{k=1}^{\infty}(k+1)ka_{k+1}x^{k-1} \;=\; \sum_{k=2}^{\infty}k(k-1)a_k x^{k-2},$$

and this new series H also has radius of convergence R.
Now suppose that x is in I. Then $|x| < R$, so we choose some S such that $|x| < S < R$. So x is in $J = (-S, S)$. Suppose that y is also in J. Then $F(y) - F(x) \;=\; \sum_{k=0}^{\infty} a_k(y^k - x^k)$.
We apply the mean value theorem. For each $k \geqslant 1$ there is some t_k between x and y such that $y^k - x^k \;=\; (y-x)$ (derivative of the function t^k at t_k) $= (y-x)kt_k^{k-1}$. This gives

$$F(y)-F(x) \;=\; \sum_{k=1}^{\infty}(y-x)ka_k t_k^{k-1} \;=\; (y-x)\sum_{k=1}^{\infty}ka_k x^{k-1} \;+\; (y-x)L(y) \;=\; (y-x)G(x) \;+\; (y-x)L(y)$$

where

$$L(y) \;=\; \sum_{k=1}^{\infty}ka_k(t_k^{k-1}-x^{k-1}) \;=\; \sum_{k=2}^{\infty}ka_k(t_k^{k-1}-x^{k-1}) \;.$$

If we can show that $L(y) \to 0$ as $y \to x$, we will have proved that $F'(x) = G(x)$.
Now we do not know what the t_k actually are, but we do know that they lie between y and x. Now we use the mean value theorem again. For each $k \geqslant 2$ we can find an s_k between t_k and x such that $t_k^{k-1} - x^{k-1} \;=\; (t_k - x)$ (derivative of the function s^{k-1} at s_k) $= (t_k-x)(k-1)s_k^{k-2}$. Thus

$$L(y) \;=\; \sum_{k=2}^{\infty}k(k-1)a_k(t_k-x)s_k^{k-2} \;.$$

We estimate $L(y)$. Now each s_k lies between t_k and x, and so in J, and so $|s_k| \leqslant S$. Also t_k lies between y and x, so $|t_k - x| \leqslant |y-x|$. Therefore

$$|L(y)| \;=\; \lim_{n\to\infty}\Big|\sum_{k=2}^{n}k(k-1)a_k(t_k-x)s_k^{k-2}\Big| \;\leqslant\; \lim_{n\to\infty}\sum_{k=2}^{n}k(k-1)|a_k|\,|t_k-x|\,|s_k|^{k-2}$$

$$\leqslant\; \lim_{n\to\infty}\sum_{k=2}^{n}k(k-1)|a_k|\,|y-x|S^{k-2} \;=\; |y-x|\sum_{k=2}^{\infty}k(k-1)|a_k|S^{k-2} \;.$$

But S is less than the radius of convergence of the power series H, and so $H(S) = \sum_{k=2}^{\infty}k(k-1)a_k S^{k-2}$ is absolutely convergent, and we set $M = \sum_{k=2}^{\infty}k(k-1)|a_k|S^{k-2}$. Then we have proved that if y and x both lie in $J = (-S, S)$ we have $|L(y)| \leqslant |y-x|M \to 0$ as $y \to x$.

So we have proved that for all x in I,

$$F'(x) \;=\; G(x) = \sum_{k=1}^{\infty}ka_k(x-c)^{k-1} \;.$$

Since G also has radius of convergence R, we can differentiate G by the same method, and we obtain, by induction,

$$F^{(p)}(x) \;=\; \sum_{k=p}^{\infty}(x-c)^{k-p}a_k\frac{k!}{(k-p)!}$$

for p in $\mathbb{N}$ and for x in I. Putting $x = c$ we get $F^{(p)}(c) = p!a_p$ as required.

Example 5.8

Find the sum of the series $\sum_{k=1}^{\infty}\frac{k}{2^k}$. For $|x| < 1$ we know that

$$\sum_{k=0}^{\infty}x^k \;=\; F(x) \;=\; \frac{1}{1-x}\ .$$

So by the differentiation theorem we have, for $|x| < 1$,

$$\frac{1}{(1-x)^2} \;=\; F'(x) \;=\; \sum_{k=1}^{\infty}kx^{k-1}\ .$$

So for $|x| < 1$ we have

$$\sum_{k=1}^{\infty}kx^k \;=\; \frac{x}{(1-x)^2}$$

and setting $x = 1/2$ the required sum is $(1/2)/(1-(1/2))^2 = 2$.

Before proceeding further we now deal with the exponential and logarithm functions.

5.9 The exponential function

We define

$$f(x) = \exp(x) \;=\; \sum_{k=0}^{\infty}\frac{x^k}{k!}\ .$$

Obviously this converges if $x = 0$. If $x \neq 0$ we set $b_k = x^k/k!$. Then $|b_{k+1}/b_k| = |x|/(k+1) \to 0$ as $k \to \infty$. So the series has radius of convergence $+\infty$. Thus f is continuous and differentiable on $\mathbb{R}$ and

$$f'(x) \;=\; \sum_{k=1}^{\infty}\frac{x^{k-1}}{(k-1)!} \;=\; \exp(x)\ .$$

Now we prove some facts about exp.

FACT 1

For all real x, we have $\exp(x)\exp(-x)=1$. In particular, $\exp(x)$ is never 0, and so $\exp(x) > 0$ for all real x.

Proof

Clearly $\exp(0)\exp(0) = 1.1 = 1$. Also, by the product rule,

$$\frac{d}{dx}\big(\exp(x)\exp(-x)\big) \;=\; \exp(x)\exp(-x)+\exp(x)\exp(-x)((d/dx)(-x)) \;=\; 0.$$

FACT 2

For all real x and y we have $\exp(x)\exp(y) = \exp(x+y)$.

Proof

We take a real y and keep it fixed, and set

$$g(x) \; = \; \exp(x+y)\exp(-x)\exp(-y) \; .$$

Then $g(0) = 1.1.1 = 1$ and the product and chain rules give

$$g'(x) \; = \; \exp(-x)\exp(-y)\big((d/dx)\exp(x+y)\big) \; - \; \exp(-x)\exp(-y)\exp(x+y) \; = \; 0.$$

FACT 3

We have, for each real λ, $\displaystyle\lim_{x \to +\infty} \frac{\exp(x)}{x^\lambda} \; = \; +\infty.$

Proof

Just choose a positive integer k such that $k > \lambda$. Then for x positive we have $\exp(x) > (x^k)/k!$. But then $(\exp(x))/x^\lambda \; > \; (x^{k-\lambda})/k! \to \infty$ as $x \to +\infty$.

FACT 4

We have $\displaystyle\lim_{x \to -\infty} \exp(x) = 0.$

Proof

Put $y = -x$. Then $\displaystyle\lim_{x \to -\infty} \exp(x) \; = \; \lim_{y \to +\infty} \frac{1}{\exp(y)} \; = \; 0.$

FACT 5

exp is strictly increasing on $\mathbb{R}$ and $\exp(\mathbb{R}) = (0, +\infty)$

Proof

The fact that exp is increasing follows from the fact that the derivative is always positive. Also, if c is positive, then Facts 3 and 4 imply that we can find x and y with $\exp(x) < c < \exp(y)$, and the intermediate value theorem gives us some t between x and y such that $\exp(t) = c$.

Now we define the logarithm, as the inverse function of exp . Thus $\log(\exp(x)) = x$ for all real x, and $\exp(\log(y)) = y$ for all $y > 0$. Our rule for inverse functions gives $(d/dy)(\log y) = 1/($ derivative of exp at $\log y \;) = 1/ \exp(\log y) = 1/y$.

Powers of *x*

We assert first that if α is a rational number, then $x^\alpha \; = \; \exp(\, \alpha \log x \,)$ for all $x > 0$. Why? Write $\alpha = p/q$ where p and q are integers with $q > 0$. Then we know that $(x^\alpha)^q \; = \; x^p$, but there is only one positive number y such that $y^q \; = \; x^p$. This is because the function $h(y) = y^q$ is increasing on $(0, +\infty)$. Now Fact 2 above gives

$$(\; \exp(\, \alpha \log x \,) \;)^q \; = \; \exp(\, q\alpha \log x \,) \; = \; \exp(\, p \log x \,) \; = \; (\; \exp(\, \log x \,) \;)^p \; = \; x^p \; .$$

This proves the assertion.

Now if α is *any* real number, we just set $x^\alpha \; = \; \exp(\, \alpha \log x \,)$ for x positive. Then we find that

$$(d/dx)(x^\alpha) \; = \; \exp(\, \alpha \log x \,)\left(\frac{d}{dx}(\, \alpha \log x \,)\right) \; = \; (x^\alpha)(\alpha/x) = \alpha \exp(\, (\alpha-1) \log x \,) \; = \; \alpha x^{\alpha-1} \; .$$

Now we define the real number e by

$$\mathrm{e} \;=\; \exp(1) \;=\; \sum_{k=0}^{\infty}\frac{1}{k!} \quad .$$

Then for any real x, we have $\mathrm{e}^x = \exp(x\log(\mathrm{e})) = \exp(x)$ as expected.

Examples 5.10

(i) Find the sum of the series $\sum_{k=1}^{\infty}\frac{1}{k2^k}$.

We set

$$h(x) \;=\; \sum_{k=0}^{\infty}\frac{x^{k+1}}{k+1} \;.$$

If we differentiate h term-by-term we get $\sum_{k=0}^{\infty}x^k$, that is, $1/(1-x)$. Now the differentiation theorem tells us that h has radius of convergence 1, and that $h'(x) = 1/(1-x)$ for $|x| < 1$. Set $g(x) = \log(1/(1-x)) - h(x)$. Then $g(0) = 0 - h(0) = 0$. Also, if $|x| < 1$ we have $g'(x) = 1/(1-x) - 1/(1-x) = 0$. So $h(x) = \log(1/(1-x))$ for $|x| < 1$. Putting $x = 1/2$ we get $\log 2 = \sum_{k=0}^{\infty}\frac{1}{(k+1)2^{k+1}} \;=\; \sum_{k=1}^{\infty}\frac{1}{k2^k}$.

(ii) Estimate log(1.1) so that the error has absolute value less than 10^{-3}.

From example (i) we know that for $|x| < 1$, we have $\log(1/(1-x)) = \sum_{k=0}^{\infty}\frac{x^{k+1}}{k+1} = x + x^2/2 + x^3/3 + x^4/4 + ...$ So for $|x| < 1$, we have

$$\log(1+x) \;=\; -\log(1/(1-(-x))) \;=\; x - x^2/2 + x^3/3 - x^4/4 + ...$$

Notice that this is a convergent alternating series, and that the terms have absolute value $|x|^k/k$, which decreases as k increases. So for $0 < x < 1$, we have

$$\log(1+x) \;=\; x - x^2/2 + (x^3/3 - x^4/4) + (x^5/5 - x^6/6) + ... \;>\; x - x^2/2.$$

Also

$$\log(1+x) \;=\; x - x^2/2 + x^3/3 - (x^4/4 - x^5/5) - (x^6/6 - x^7/7) - ..... \;<\; x - x^2/2 + x^3/3.$$

So

$$x - x^2/2 \;<\; \log(1+x) \;<\; x - x^2/2 + x^3/3$$

and the *error* which arises if we approximate $\log(1+x)$ by $x - x^2/2$ is at most $x^3/3$. For $x = 0.1$ this gives an error less than 10^{-3}, and our approximation is 19/200.
This is a very useful trick for estimating functions which are represented by convergent alternating series.

Section 6 The Trigonometric Functions

6.1 Introduction

This section will be omitted from the lectures (and the exam!). Every student (hopefully!) knows that the derivative of sine is cosine, and that

$$\sin x \;=\; x - x^3/3! + x^5/5! - x^7/7! + \ldots.. \tag{1}$$

To prove the first fact just using trigonometry is quite difficult. Consequently, some books use (1) as the *definition* of sine. In this section we will prove both facts, plus some others, using mainly the mean and intermediate value theorems.

Consider the circle T given by $x^2+y^2 = 1$, the area of the enclosed region being (by definition) π. We take a point on T in the first quadrant, given by $(x,y) = (x,\sqrt{1-x^2})$, with $0 < x < 1$. Consider further the region enclosed by the positive x-axis, the circle T and the straight line from (0, 0) to (x,y). Let the area of this region be $t/2$. This corresponds exactly to the angle enclosed being t radians. Now we have

$$x = \cos t \quad , \quad y = \sin t \tag{2}$$

As x increases from 0 to 1, t decreases from $\pi/2$ to 0. Regarding t as a function $t = f(x)$ for $0 < x < 1$, we have

$$t = f(x) = x\sqrt{1-x^2} + 2C(x) \tag{3}$$

where $C(x)$ is the area enclosed by the positive x-axis, the circle T and the straight line from $(x,0)$ to $(x,\sqrt{1-x^2})$.

Now suppose that $0 < u < v < 1$. Then elementary geometric considerations give

$$(v-u)\sqrt{1-v^2} \;<\; C(u) - C(v) \;<\; (v-u)\sqrt{1-u^2}.$$

Dividing through by $(v-u)$ we see that by the intermediate value theorem we have

$$\frac{C(v)-C(u)}{v-u} \;=\; -\sqrt{1-s^2} \tag{4}$$

for some s lying between u and v. We also have (4) if $0 < v < u < 1$ (just interchange u and v). Letting $v \to u$ we see that $s \to u$ and so $C'(u) = -\sqrt{1-u^2}$. Thus, using (3),

$$\frac{\mathrm{d}t}{\mathrm{d}x} \;=\; f'(x) \;=\; (1-x^2)^{1/2} \,+\, x(1/2)(1-x^2)^{-1/2}(-2x) \,-\, 2(1-x^2)^{1/2} \;=\; -(1-x^2)^{-1/2}.$$

Using our rule for inverse functions we therefore have, for $0 < t < \pi/2$,

$$\frac{\mathrm{d}}{\mathrm{d}t}(\cos t) \;=\; \frac{\mathrm{d}x}{\mathrm{d}t} \;=\; -\sqrt{1-x^2} \;=\; -y \;=\; -\sin t.$$

Also, using the chain rule,

$$\frac{\mathrm{d}}{\mathrm{d}t}(\sin t) \;=\; \frac{\mathrm{d}y}{\mathrm{d}t} \;=\; \frac{\mathrm{d}}{\mathrm{d}t}(\sqrt{1-\cos^2 t}) \;=\; (1/2)(1-\cos^2 t)^{-1/2}(-2\cos t)(-\sin t) \;=\; \cos t.$$

We have thus proved the differentiation formulas, but so far only on $(0,\pi/2)$.

6.2 The Power Series
We define

$$S(t) = t - t^3/3! + t^5/5! - t^7/7! + \ldots\ldots$$

$$C(t) = 1 - t^2/2! + t^4/4! - t^6/6! + \ldots\ldots$$

Both series converge absolutely for all t (compare with $e^{|t|}$). The differentiation theorem now gives

$$S'(t) = C(t)\ ,\ C'(t) = -S(t).$$

We claim that $S(t) = \sin t$ for $0 < t < \pi/2$. To see this, put $u = S(t) - \sin t$. Then $u'(t) = C(t) - \cos t$ and $u''(t) = -S(t) + \sin t = -u$. Therefore $u'(t)(u''(t) + u(t)) = 0$ and so $(u'(t))^2 + u^2(t)$ is constant on $(0,\pi/2)$. Letting $t \to 0+$ we see that $S(t)$ and $\sin t$ both tend to 0. Also $C(t) \to 1$ and $\cos t \to 1$. So $(u'(t))^2 + u^2(t) \to 0$ as $t \to 0+$ and so must be 0 for all t on $(0,\pi/2)$. In particular $u(t) = 0$, which is what we wanted to prove.
Moreover, $u'(t) = 0$, so $C(t) = \cos t$ on $(0,\pi/2)$.

Now we want to show that $S(t) = \sin t$ for *all* t. Note first that $S(-t) = -S(t)$, $C(-t) = C(t)$, which means that $S(t) = \sin t$, $C(t) = \cos t$ for $-\pi/2 < t < \pi/2$. For t outside this range, sine and cosine are given by the formulas

$$\sin(t+\pi) = -\sin t\ ,\ \cos(t+\pi) = -\cos t. \tag{5}$$

In fact, the usual convention is to set $\sin(\pi - t) = \sin t$, $\sin(t + 2\pi) = \sin t$, but both these formulas follow from (5). So we just need to check that the relations (5) hold for S and C.

6.3 The Addition Formulas
For all real x and y,

$$S(x+y) = S(x)C(y) + C(x)S(y) \tag{2.1}$$

$$C(x+y) = C(x)C(y) - S(x)S(y) \tag{2.2}$$

We just prove (2.1), the proof of (2.2) being the same. We keep y fixed and set

$$U(x) = S(x+y) - S(x)C(y) - C(x)S(y).$$

Then it is easy to check that $U''(x) + U(x) = 0$ for all x, so that $(U'(x))^2 + U^2(x)$ is constant, by the same argument as above. But $U(0) = U'(0) = 0$, as is easy to check. So $U(x) = 0$ for all x. These addition formulas easily give

$$S(\pi) = 2S(\pi/2)C(\pi/2) = 0,$$

since

$$C(\pi/2) = \lim_{x \to \pi/2-} C(x) = \lim_{x \to \pi/2-} \cos x = 0,$$

and similarly $C(\pi) = -1$. Further, $S(2\pi) = 0, C(2\pi) = 1$, and all the properties such as $S(x+2\pi) = S(x)$ are easy to check. In particular,

$$S(x+\pi) = -S(x)\ ,\ C(x+\pi) = -C(x).$$

This means that $S(t) = \sin t$, $C(t) = \cos t$ for all t. Moreover, we also now know that the formulas (2.1) and (2.2) hold with S,C replaced by sine and cosine respectively.

Section 7 Taylor's Theorem and the Taylor Series

Theorem 7.1 Taylor's Theorem, or the n'th Mean Value Theorem

Suppose that n is a positive integer, and that f is a real-valued function which is n times differentiable on an interval containing the points a and $x \neq a$. Then there exists c lying strictly between a and x (i.e. in (a,x) or in (x,a)) such that

$$f(x) = \sum_{k=0}^{n-1} \frac{f^{(k)}(a)}{k!}(x-a)^k + \frac{(x-a)^n}{n!} f^{(n)}(c).$$

Remarks

1. If we put $n = 1$ we get $f(x) = f(a) + (x-a)f'(c)$ which is the ordinary mean value theorem. 2. In general, c will depend on f, n and x. 3. There are other versions of this, with different forms for the remainder term, but this is probably the easiest to remember!

Proof of Taylor's theorem

We keep x fixed and for y lying between a and x we set

$$H(y) = \left(\sum_{k=0}^{n-1} \frac{(x-y)^k}{k!} f^{(k)}(y) \right) - f(x).$$

Then

$$H'(y) = -\sum_{k=1}^{n-1} \frac{(x-y)^{k-1}}{(k-1)!} f^{(k)}(y) + \sum_{k=0}^{n-1} \frac{(x-y)^k}{k!} f^{(k+1)}(y) = \frac{(x-y)^{n-1}}{(n-1)!} f^{(n)}(y).$$

Now we put

$$G(y) = H(y) - \frac{(x-y)^n}{(x-a)^n} H(a).$$

Then $G(a) = 0$ and $G(x) = H(x) = 0$. So by Rolle's theorem there exists a point c lying between a and x such that $G'(c) = 0$. This gives

$$H'(c) + n\frac{(x-c)^{n-1}}{(x-a)^n} H(a) = 0, \quad \text{and} \quad \frac{(x-c)^{n-1}}{(n-1)!} f^{(n)}(c) + n\frac{(x-c)^{n-1}}{(x-a)^n} H(a) = 0.$$

Therefore $H(a) = -\dfrac{(x-a)^n}{n!} f^{(n)}(c)$ which is what we need.

As an application we prove

Theorem 7.2 The Generalised Second Derivative Test

Suppose that $n \geqslant 2$ and that f is a real-valued function such that $f'(a) = \ldots\ldots = f^{(n-1)}(a) = 0$, and that $f^{(n)}(a)$ exists and is $\neq 0$.

If n is even and $f^{(n)}(a) > 0$, then f has a local minimum at a.

If n is even and $f^{(n)}(a) < 0$, then f has a local maximum at a.

If n is odd then f does not have a local maximum or a local minimum at a.

Proof

We suppose first that $f^{(n)}(a) > 0$ (if not, we can look at $-f$). Then

$$\lim_{x \to a} \frac{f^{(n-1)}(x) - f^{(n-1)}(a)}{x-a} \quad > \quad 0.$$

So there is some $\delta > 0$ such that $f^{(n-1)}(s) > 0$ for all s in $(a, a+\delta)$ and $f^{(n-1)}(s) < 0$ for all s in $(a-\delta, a)$. Now suppose that $0 < |x-a| < \delta$. Then by Taylor's theorem (with n replaced by $n-1$) we have, for some s between a and x,

$$f(x) = f(a) \;+\; \frac{(x-a)^{n-1}}{(n-1)!} f^{(n-1)}(s).$$

If n is odd, this gives $f(x) > f(a)$ if $x > a$ and $f(x) < f(a)$ if $x < a$.

If n is even, we obtain $f(x) > f(a)$ for $x > a$ and for $x < a$. So we have a minimum.

Another application of Taylor's theorem is to estimation.

Example 7.3

Estimate $\cos(0.1)$ so that the error has absolute value less than 10^{-5}.

We could do this using the power series representation for cosine, which gives an alternating series. However, we'll use Taylor's theorem here. With $f(x) = \cos x$ we have, for $n \in \mathbb{N}$,

$$f(x) \;=\; \sum_{k=0}^{n-1} \frac{f^{(k)}(0)}{k!} x^k \;+\; \frac{x^n}{n!} f^{(n)}(s)$$

for some s between 0 and x. Now $|f^{(n)}(s)|$ is certainly $\leqslant 1$. So we need to make $(0.1)^n/n! < 10^{-5}$, and $n = 4$ will do. Our estimate is $\sum_{k=0}^{3} \frac{f^{(k)}(0)}{k!}(0.1)^k \;=\; 199/200$.

7.4 The Taylor Series

Suppose that f is a real-valued function such that *all* the derivatives $f^{(n)}$ exist at a. Then we can form the Taylor series

$$T(x,a) \;=\; \sum_{k=0}^{\infty} \frac{f^{(k)}(a)}{k!} (x-a)^k \,.$$

Remarks

1. The special case where $a = 0$ is called the Maclaurin series of f. 2. $T(x,a)$ is, of course, a power series. 3. The obvious question to ask is: are $T(x,a)$ and $f(x)$ equal? Obviously $T(a,a) = f(a)$.

Example 1

Suppose that $f(x)$ is a polynomial. Let the degree of f be $n-1$. Then $f^{(n)}(x)$ is identically 0, and Taylor's theorem gives

$$f(x) \;=\; \sum_{k=0}^{n-1} \frac{f^{(k)}(a)}{k!}(x-a)^k \;=\; \sum_{k=0}^{\infty} \frac{f^{(k)}(a)}{k!}(x-a)^k \;=\; T(x,a)$$

so that the answer is always yes in this case.

Example 2

Suppose that $f(x) \;=\; \sum_{k=0}^{\infty} c_k(x-a)^k$ is a power series with positive radius of convergence R and centre a. Then we know from Section 5 that $c_k = f^{(k)}(a)/k!$ and so $f(x) = T(x,a)$ for $|x-a| < R$.

Example 3

The functions sine and cosine equal their Maclaurin series for all x, i.e.

$$\sin x = x - x^3/3! + x^5/5! - \ldots.. \;\;,\;\; \cos x = 1 - x^2/4! + x^4/4! - \ldots.. \;\;.$$

This is proved in Section 6. Alternatively, you can use Taylor's theorem with $a = 0$, and f either sine or cosine. We get a remainder term $f^{(n)}(c)x^n/n!$, where c depends on x and n. However, $f^{(n)}(c)$ has absolute value at most 1, and so the remainder tends to 0 as $n \to \infty$.

Example 4

Set $f(0) = 0$ and $f(x) = e^{-1/x^2}$ if $x \neq 0$. If $x \neq 0$ then $f(y) = e^{-1/y^2}$ for all y sufficiently close to x. So we can differentiate repeatedly and it is easy to prove by induction that for $x \neq 0$ and n a positive integer,

$$f^{(n)}(x) \;=\; e^{-1/x^2}\, P_n(1/x),$$

where each P_n is a polynomial. Now, for any integer n,

$$\lim_{x \to 0} \frac{e^{-1/x^2}}{x^n} \;=\; \lim_{y \to +\infty} \frac{y^{n/2}}{e^y} \;=\; 0.$$

Using this, we can prove by induction that $f^{(n)}(0) = 0$. We know this is true for $n = 0$. Assuming it true for n, we get

$$f^{(n+1)}(0) = \lim_{x \to 0} \frac{f^{(n)}(x) - f^{(n)}(0)}{x-0} = \lim_{x \to 0} e^{-1/x^2}\, P_n(1/x)x^{-1} = 0.$$

Therefore the Maclaurin series of f is identically zero, and so does not equal $f(x)$ for any non-zero x. So, in general, the answer to our question above is "not necessarily".

7.5 Multiplying Power Series

Suppose that

$$F(x) \;=\; \sum_{k=0}^{\infty} a_k(x-c)^k \;\;,\;\; G(x) \;=\; \sum_{k=0}^{\infty} b_k(x-c)^k \;\;,$$

are two power series with the *same* centre c, each having radius of convergence at least $R > 0$. Then we know that F and G can be differentiated as many times as we like in $|x-c| < R$, and therefore so can

$H(x) = F(x)G(x)$. Now for such x and $n \in \mathbb{N}$, we have

$$H^{(n)}(c) \ = \ \sum_{k=0}^{n} \frac{n!}{k!(n-k)!} F^{(k)}(c) G^{(n-k)}(c).$$

This is Leibniz' formula for the higher derivatives of FG and is easy to prove by induction. This gives

$$H^{(n)}(c)/n! \ = \ \sum_{k=0}^{n} a_k b_{n-k}$$

so that the Taylor series of H about c is

$$\sum_{n=0}^{\infty} \left(\sum_{k=0}^{n} a_k b_{n-k} \right) (x-c)^n \ , \qquad (*)$$

which is the series you get if you multiply out $F(x)$ times $G(x)$ by long multiplication, gathering up powers of $x-c$.

Theorem
With F and G as above, the series (*) is equal to $H(x) = F(x)G(x)$ for $|x-c| < R$.
To prove this theorem using just real analysis is difficult. Using complex analysis it is very easy and will be proved in Semester 3.

Example 7.5
Find a power series which represents the function $f(x) = (\sin x)/(1-x)$ for $|x| < 1$.
We have $\sin x \ = \ x - x^3/3! + ...$ for all x and $1/(1-x) \ = \ 1 + x + x^2 + x^3 + ...$ for $|x| < 1$. So for $|x| < 1$ we have $f(x) \ = \ x + x^2 + x^3(1 - 1/3!) + x^4(1 - 1/3!) + ...$

Section 8 Indeterminate Forms

8.1 Introduction
Consider the limit

$$\lim_{x \to 1} \frac{x^{16} + x - 2}{x^2 - 1} \ .$$

Both numerator and denominator approach 0 as $x \to 1$. However, the limit may still exist. Such a limit is called an indeterminate form. To develop a quick way to evaluate similar limits, we first need:

Theorem 8.2 Cauchy's mean value theorem
Suppose that f,g are real-valued functions continuous on $[a, b]$ and differentiable on (a, b). Then there exists c in (a, b) such that

$$(f(b) - f(a))g'(c) = (g(b) - g(a))f'(c).$$

The **Proof** consists of just taking the function $h(x) = (f(b) - f(a))(g(x) - g(a)) - (g(b) - g(a))(f(x) - f(a))$. Since $h(a) = h(b) = 0$ we obtain a c with $h'(c) = 0$ from Rolle's theorem.

Theorem 8.3 L'Hôpital's rule, first version

Let "lim" stand for any of $\lim_{x \to a+}$, $\lim_{x \to a-}$, $\lim_{x \to a}$, $\lim_{x \to +\infty}$, $\lim_{x \to -\infty}$. If $\lim f(x) = \lim g(x) = 0$ and

$$\lim \frac{f'(x)}{g'(x)} = L$$

(finite or infinite) exists, then

$$\lim \frac{f(x)}{g(x)}$$

exists and is the same.

Proof

We first consider the case of $\lim_{x \to a+}$. Since $\lim_{x \to a+} \frac{f'(x)}{g'(x)}$ exists, there must be some $\delta > 0$ such that $g'(s) \neq 0$ for $a < s \leqslant a+\delta$, since $f'(s)/g'(s)$ is defined. We set $f(a) = g(a) = 0$ and this makes f,g continuous on $[a, a+\delta]$. We also have $g(x) \neq 0$ for x in $(a, a+\delta]$, for otherwise Rolle's theorem would give us an s between a and x with $g'(s) = 0$.

Now take x such that $a < x < a+\delta$. Then by Theorem 8.2 there is a c_x in (a, x) such that

$$(f(x) - f(a))g'(c_x) = (g(x) - g(a))f'(c_x) .$$

This gives $f(x)/g(x) = f'(c_x)/g'(c_x)$. Now as $x \to a+$ we see that $c_x \to a+$ and so $f(x)/g(x) \to L$.
The proof for $\lim_{x \to a-}$ is the same.

Now we consider the case where "lim" is $\lim_{x \to +\infty}$. Here we set $F(x) = f(1/x)$, $G(x) = g(1/x)$. Then $\lim_{x \to 0+} F(x) = \lim_{x \to 0+} G(x) = 0$. Now

$$L = \lim_{x \to +\infty} \frac{f'(x)}{g'(x)} = \lim_{x \to 0+} \frac{f'(1/x)}{g'(1/x)} = \lim_{x \to 0+} \frac{(-1/x^2)f'(1/x)}{(-1/x^2)g'(1/x)} = \lim_{x \to 0+} \frac{F'(x)}{G'(x)} .$$

By the first part, the last limit is $\lim_{x \to 0+} F(x)/G(x) = \lim_{x \to +\infty} f(x)/g(x)$.

Examples 8.4

1. Consider $\lim_{x \to 1} \frac{x^{16}+x-2}{x^2-1}$. The rule applies and we can look at $\lim_{x \to 1} \frac{16x^{15}+1}{2x} = 17/2$. So the first limit is $17/2$.

2. $\lim_{x \to 0} \frac{x - \sin x}{1 - \cos x}$. Applying the rule, we look at $\lim_{x \to 0} \frac{1-\cos x}{\sin x}$. This is again indeterminate, but we can apply the rule again, and look at $\lim_{x \to 0} \frac{\sin x}{\cos x} = 0$. So the second limit is 0 and so is the first. We could also use power series here. We can write, for $x \neq 0$,

$$\frac{x - \sin x}{1 - \cos x} = (x^3/3! - x^5/5! + \dots)/(x^2/2! - x^4/4! + \dots) = x(1/3! - x^2/5! + \dots)/(1/2! - x^2/4! + \dots) \to 0$$

as $x \to 0$.

3. $\lim\limits_{x \to +\infty} \dfrac{\log(1+e^x)}{x}$. This is slightly different, as numerator and denominator both tend to $+\infty$. We could convert it to the form covered above, but the computations will be very tricky. Instead, we wait for L'Hôpital's second rule (8.5).

4. $\lim\limits_{x \to 1} \dfrac{x^2+1}{x^2-1}$. We CANNOT legitimately apply L'Hôpital's rule here, as the limit is NOT an indeterminate form. This is because $x^2+1 \to 2 \neq 0$ as $x \to 1$. In fact, the required limit does not exist.

5. $\lim\limits_{x \to 0} \dfrac{x}{\sin x}$. Again, we COULD use power series. Applying L'Hôpital's rule, we need to look at $1/(\cos x) \to 1$ as $x \to 0$. So the required limit is 1.

6. $\lim\limits_{x \to 0} \dfrac{x^2 \sin(1/x)}{\sin x}$. This is an indeterminate form. If we apply the rule, we need to look at $\lim\limits_{x \to 0} \dfrac{2x\sin(1/x) - \cos(1/x)}{\cos x}$. However, this limit does not exist. This is because $x\sin(1/x)/\cos x \to 0$ as $x \to 0$, but $\cos(1/x)/\cos x$ has no limit as $x \to 0$.
This does not mean, however, that the required limit does not exist, as the rule says nothing about this case. In fact, since $\sin(1/x)$ is bounded, we see from Example 5 that the required limit is 0.

7. $\lim\limits_{x \to +\infty} (1+1/x)^x$. If we take logarithms, we need to look at $\lim\limits_{x \to +\infty} x\log(1+1/x) = \lim\limits_{y \to 0+} \dfrac{\log(1+y)}{y}$. Applying the rule, we look at $\lim\limits_{y \to 0+} \dfrac{1/(1+y)}{1} = 1$. So the required limit is e, since the exponential function is continuous at 1.

Now we prove the second version of L'Hôpital's rule.

Theorem 8.5

Let "lim" be as in 8.3. Suppose that $\lim f(x)$ is $+\infty$ or $-\infty$ and $\lim g(x)$ is $+\infty$ or $-\infty$.
If $\lim \dfrac{f'(x)}{g'(x)}$ exists and is L (finite or infinite) then $\lim \dfrac{f(x)}{g(x)}$ exists and is L.

Proof

We prove this only for $\lim\limits_{x \to a+}$. For $\lim\limits_{x \to +\infty}$ we can use the same trick as in 8.3.
Since $\lim f'(x)/g'(x)$ is assumed to exist, we see again that there must be some $\delta > 0$ such that $g'(s) \neq 0$ for $a < s \leqslant a+\delta$.
We prove simultaneously the cases $L \in \mathbb{R}$ and $L = +\infty$ (if $L = -\infty$ look at $-f/g$). Take an $\varepsilon > 0$ and an $M > 0$. We know that there is some $\rho > 0$ such that $\rho \leqslant \delta$ and such that $a < y < b = a+\rho$ implies that $f'(y)/g'(y)$ belongs to $(L-\varepsilon/2\ ,\ L+\varepsilon/2)$ (if L is finite) or $(2M\ ,\ +\infty)$ (if $L = +\infty$).
Now suppose that $a < x < b = a+\rho$. Then there exists a y with $x < y < b$ such that

$$(f(b) - f(x))g'(y) = (g(b) - g(x))f'(y) .$$

This gives

$$(f(b) - f(x))/(g(b) - g(x)) = f'(y)/g'(y) .$$

Dividing by $g(b)-g(x)$ is legitimate, for if $g(b) = g(x)$ we would obtain some s with $x < s < b$ and $g'(s) = 0$, which we have ruled out. Thus $(f(x)-f(b))/(g(x)-g(b))$ belongs to $(L-\varepsilon/2\,,\,L+\varepsilon/2)$ or to $(2M\,,\,+\infty)$. Now we write

$$\frac{f(x)}{g(x)} = \frac{f(x)}{f(x)-f(b)}\;\frac{f(x)-f(b)}{g(x)-g(b)}\;\frac{g(x)-g(b)}{g(x)}\;.$$

As $x \to a+$, the first and last terms tend to 1, while the second term lies in $(L-\varepsilon/2\,,\,L+\varepsilon/2)$ or $(2M\,,\,+\infty)$. Therefore if x is close enough to a, then $f(x)/g(x)$ lies in $(L-\varepsilon\,,\,L+\varepsilon)$ or $(M\,,\,+\infty)$, which is exactly what we needed to show.

Examples 8.6

1. $\lim_{x \to +\infty} \frac{\log(1+\mathrm{e}^x)}{x}$. The rule 8.5 tells us to look at $\lim_{x \to +\infty} \frac{\mathrm{e}^x/(1+\mathrm{e}^x)}{1} = 1$. So the required limit is e.

2. $\lim_{x \to 0+} \frac{\log x}{\operatorname{cosec} x}$. We need to look at $\lim_{x \to 0+} \frac{1/x}{-\cot x\,\operatorname{cosec} x} = \lim_{x \to 0+} \frac{-\sin^2 x}{x\cos x}$. This is, using Example 5 of 8.3, equal to 0. So the required limit is 0.

3. $\lim_{x \to +\infty} (\log x)^{1/x}$. We take logarithms and look at $\lim_{x \to +\infty} \frac{\log\log x}{x}$. Applying 8.5 we look at $\lim_{x \to +\infty} \frac{1/(x\log x)}{1} = 0$. The required limit is therefore $\mathrm{e}^0 = 1$.

4. $\lim_{x \to 0+} (\cos x)^{1/x^2}$. We take logarithms and look at $\lim_{x \to 0+} \frac{\log(\cos x)}{x^2}$. The rule tells us to look at

$\lim_{x \to 0+} \frac{-\tan x}{2x} = \lim_{x \to 0+} \frac{-\sec^2 x}{2} = -1/2$. So the required limit is $\mathrm{e}^{-1/2}$.

5. $\lim_{x \to 0+} (\sin x)^{1/\log x}$. We take logarithms and look at $\lim_{x \to 0+} \frac{\log(\sin x)}{\log x}$, and so at $\lim_{x \to 0+} \frac{\cot x}{1/x} = \lim_{x \to 0+} \frac{x\cos x}{\sin x} = 1$. The required limit is therefore e.

Section 9 Integration

Example 9.1

Consider the curve $y = x^2$, $0 \leqslant x \leqslant 1$ and the area A bounded by this curve, the positive x-axis, and the line from (1, 0) to (1, 1) (the area "under the curve"). We determine this area without using calculus. Of course, calculus tells us to expect the answer 1/3.

We proceed by dividing [0, 1] into n equal sub-intervals, each of length $1/n$, this for $n \in \mathbb{N}$. Then the part of the required area between $x = (k-1)/n$ and $x = k/n$ may be enclosed in a rectangle of base $1/n$ and height $(k/n)^2$. Thus

$$A \leqslant \sum_{k=1}^{n} (k/n)^2/n = \left(\sum_{k=1}^{n} k^2\right) n^{-3} = n(n+1)(2n+1)/6n^3\;.$$

(The last fact used is easily proved by induction on n.) Now $n(n+1)(2n+1)/6n^3$ is always greater than

1/3 and tends to 1/3 as $n \to \infty$.

We can also see that the part of the required area between $x = (k-1)/n$ and $x = k/n$ encloses a rectangle of base $1/n$ and height $((k-1)/n)^2$. Thus

$$A \geqslant \left(\sum_{k=1}^{n} (k-1)^2\right) n^{-3} = (n-1)n(2n-1)/6n^3 .$$

The last quantity is always less than 1/3 and tends to 1/3 as $n \to \infty$. Thus A must equal 1/3.

The main idea of this section is to write, for a continuous function f, $F(x) = \int_a^x f(t)\,dt$ and to show that $F'(x) = f(x)$. This will give rise to the familiar method of integrating by using anti-derivatives. To do this, we need to define what we mean by the integral. It may be tempting to define it as the "area under the curve", but this would lead to at least 3 difficulties.

1. How do you know that the area exists? f may be a very messy curve, such as the continuous, nowhere differentiable function in Section 4.
2. What if f changes sign, as, for example, $x\sin(1/x)$ does infinitely often?
3. How would you prove that the integral of $f+g$ is that of f plus that of g?

We shall use a method known as Riemann integration.

Definitions 9.2

Let f be a bounded real-valued function on the closed interval $[a, b] = I$. Henceforth $a < b$ unless otherwise explicitly stated. Assume that $|f(x)| \leqslant M$ for all x in I.

A PARTITION P of I is a finite set $\{ x_0,....,x_n \}$ such that $a = x_0 < x_1 < .... < x_n = b$. The points x_j are called the vertices of P. We say that partition Q of I is a refinement of partition P of I if every vertex of P is a vertex of Q (i.e., crudely speaking, P is a "subset" of Q). For P as above, we define

$M_k(f) = \sup \{ f(x): x_{k-1} \leqslant x \leqslant x_k \} \leqslant M$ and $m_k(f) = \inf \{ f(x): x_{k-1} \leqslant x \leqslant x_k \} \geqslant -M$.

Further, we define the UPPER SUM

$$U(P,f) = \sum_{k=1}^{n} M_k(f)(x_k - x_{k-1})$$

and the LOWER SUM

$$L(P,f) = \sum_{k=1}^{n} m_k(f)(x_k - x_{k-1}) .$$

Notice that $L(P,f) \leqslant U(P,f)$. The reason we require f to be bounded is so that all the m_k and M_k are finite and the sums exist. Notice also that $-M \leqslant m_k \leqslant M_k \leqslant M$ for each k, and so $-M(b-a) \leqslant L(P, f) \leqslant U(P, f) \leqslant M(b-a)$.

Suppose that f is positive on I and that the area A under the curve exists. Reasoning as in 9.1, it is not hard to see that $L(P,f) \leqslant A \leqslant U(P,f)$ for every partition P of I. In our Example 9.1, we had $f(x) = x^2$ and $P = \{ 0, 1/n, \ldots, (n-1)/n, 1 \}$, while $M_k(f) = (k/n)^2$, $m_k(f) = ((k-1)/n)^2$.

Further, if you draw for yourself a simple curve, it is not hard to convince yourself that refining P tends to increase $L(P,f)$ and decrease $U(P,f)$. We prove this as a lemma.

Lemma 9.3

Let f be a bounded real-valued function on $I = [a, b]$.
(i) If P, Q are partitions of I and Q is a refinement of P, then

$$L(P,f) \leqslant L(Q,f) \ , \ U(P,f) \geqslant U(Q,f) \ .$$

(ii) If P_1 and P_2 are any partitions of I, then $L(P_1,f) \leqslant U(P_2,f)$. Thus any lower sum is $\leqslant$ any upper sum.

Proof

(i) We first prove this for the case where Q is P plus one extra point. The general case then follows by adding points one at a time. So suppose that Q is the same as P, except that it has one extra vertex c, where $x_{k-1} < c < x_k$. Then
$U(Q,f) - U(P,f) = (\sup \{ f(x) : x_{k-1} \leqslant x \leqslant c \}) (c - x_{k-1}) + (\sup \{ f(x) : c \leqslant x \leqslant x_k \}) (x_k - c)$ -
$(\sup \{ f(x) : x_{k-1} \leqslant x \leqslant x_k \}) (x_k - x_{k-1})$.
This is using the fact that all other terms cancel. Now comparing the sups we see that $U(Q,f) - U(P,f) \leqslant 0$. The proof for the lower sums is the same idea.
(ii) Here we just set P to be the partition obtained by taking all the vertices of P_1 and all those of P_2. We arrange these vertices in order, and P is a refinement of P_1 and of P_2. Now we can write

$$L(P_1,f) \leqslant L(P,f) \leqslant U(P,f) \leqslant U(P_2,f).$$

Definitions 9.4 The Riemann integral

Let f be bounded, real-valued on $I = [a, b]$ as before, with $|f(x)| \leqslant M$ there. We define the UPPER INTEGRAL of f from a to b as
$\overline{\int}_a^b f(x)\,\mathrm{d}x = \inf \{ U(P,f) : P \text{ a partition of } I \}$.
This exists, because all the upper sums are bounded below by $-M(b-a)$ (see 9.2). Similarly we define the LOWER INTEGRAL
$\underline{\int}_a^b f(x)\,\mathrm{d}x = \sup \{ L(P,f) : P \text{ a partition of } I \}$.
Again this exists, because all the lower sums are bounded above by $M(b-a)$.
Now we define what it means to be Riemann integrable. We say that f is Riemann integrable on I if

$$\underline{\int}_a^b f(x)\,\mathrm{d}x = \overline{\int}_a^b f(x)\,\mathrm{d}x$$

and, if so, we denote the common value by $\int_a^b f(x)\,\mathrm{d}x$.

Notice that the lower integral is always $\leqslant$ the upper integral, because of 9.3, (ii). Also, if f is Riemann integrable and positive on I and the area A under the curve exists, then the fact that $L(P,f) \leqslant A \leqslant U(P,f)$ for every partition P of I implies that the lower integral is $\leqslant A$ and the upper integral is $\geqslant A$, which means that A equals $\int_a^b f(x)\,\mathrm{d}x$. As usual in integration, it does not matter whether you write $f(x)\,\mathrm{d}x$ or $f(t)\,\mathrm{d}t$

etc.

Example 9.5

Define f on $I = [0,\ 1]$ by $f(x) = 1$ if x is rational and $f(x) = 0$ otherwise. Let $P = \{\ x_0, \ldots, x_n\ \}$ be any partition of I. Then clearly $M_k(f) = 1$ for each k, since each sub-interval $[x_{k-1},\ x_k]$ contains a rational number. Thus $U(P,f) = \sum_{k=1}^{n} (x_k - x_{k-1}) = 1$ and so the upper integral is 1. Similarly, we have $m_k(f) = 0$ for each k, all lower sums are 0, and the lower integral is 0.

We aim to prove that continuous functions are Riemann integrable, and to do this we need the following result.

Theorem 9.6 Riemann's criterion

A bounded real-valued f function on $I = [a,\ b]$ is Riemann integrable if and only if the following is true. For any positive number ε which is given, we can find a partition P of I (which may depend on ε) such that $U(P,f) - L(P,f) < \varepsilon$.

Proof

Suppose first that f is Riemann integrable, and that $\varepsilon > 0$. Let $L = \int_a^b f(x)\ \mathrm{d}x$. Since L is the inf of the upper sums, we can find a P_1 such that $U(P_1,f) < L + \varepsilon/2$. Since L is the sup of the lower sums, we can find a P_2 such that $L(P_2,f) > L - \varepsilon/2$. As in (9.3), we now define P to be the partition consisting of all vertices of P_1 and of P_2, so that P is a refinement of both P_j. Now, using 9.3,
$L - \varepsilon/2 < L(P_1,f) \leqslant L(P,f) \leqslant U(P,f) \leqslant U(P_1,f) < L + \varepsilon/2.$

Now we prove the converse, and assume that f is NOT Riemann integrable on I. Then the upper integral must be $>$ the lower integral, and we can write, for some $\varepsilon > 0$,

$$\underline{\int}_a^b f(x)\ \mathrm{d}x = \overline{\int}_a^b f(x)\ \mathrm{d}x - \varepsilon.$$

But every lower sum is $\leqslant$ the lower integral, and every upper sum is $\geqslant$ the upper integral. Thus for every partition P we have $U(P,f) - L(P,f) \geqslant \varepsilon$. So if f is not Riemann integrable, there is some $\varepsilon > 0$ for which Riemann's criterion cannot be satisfied. Therefore if Riemann's criterion can be satisfied for every $\varepsilon > 0$, the function f must be Riemann integrable on I.

Before proving that continuous functions are Riemann integrable, we first deal with the rather easier case of monotone functions.

Theorem 9.7

Suppose that f is a monotone function on $I = [a,\ b]$. Then f is Riemann integrable on I.

Proof

We only deal with the case where f is non-decreasing. The non-increasing case is similar. Now if $f(b) = f(a)$ then f is constant on I and so the result follows from Problem 70. If $f(b) > f(a)$ we proceed as follows. Let $\varepsilon > 0$. If we can find a partition P such that $U(P,f) - L(P,f) < \varepsilon$, then we have proved that f is Riemann integrable on I.

We choose a partition $P = \{ x_0, \ldots, x_n \}$ such that for each k we have $x_k - x_{k-1} < \varepsilon/(f(b) - f(a))$. Now, since f is non-decreasing we have $M_k(f) = f(x_k)$ and $m_k(f) = f(x_{k-1})$. Thus

$$U(P,f) - L(P,f) = \sum_{k=1}^{n} (M_k(f) - m_k(f))(x_k - x_{k-1}) = \sum_{k=1}^{n} (f(x_k) - f(x_{k-1}))(x_k - x_{k-1}) <$$

$$< \sum_{k=1}^{n} (f(x_k) - f(x_{k-1}))\varepsilon/(f(b) - f(a)) = (f(x_n) - f(x_0))\varepsilon/(f(b) - f(a)) = \varepsilon .$$

To handle the case of continuous functions, we need the following.

Theorem 9.8

Let f be a continuous real-valued function on the closed interval $I = [a, b]$ and let $\varepsilon > 0$. Then there exists a $\delta > 0$ such that $|f(x) - f(y)| < \varepsilon$ for all x and y in I such that $|x - y| < \delta$.

Remark

This property is called UNIFORM continuity, because the δ does not depend on the particular choice of x or y. The theorem is NOT true for open intervals, as the example $h(x) = 1/x$, $I = (0, 1)$ shows. To see this, just note that $h(1/n) - h(1/(n-1)) = 1$ for all $n \in \mathbb{N}$, but $|1/n - 1/(n-1)| = 1/n(n-1)$, which we can make as small as we like.

Proof of 9.8

Suppose that $\varepsilon > 0$ and that NO positive δ exists with the property in the statement. Then, for each $n \in \mathbb{N}$, $1/n$ is not such a δ. Thus there are points x_n and y_n in I with $|x_n - y_n| < 1/n$, but with $|f(x_n) - f(y_n)| \geqslant \varepsilon$. Now (x_n) is a sequence in the closed interval I, and so is a bounded sequence, and therefore we can find a convergent subsequence (x_{k_n}), with limit B, say. Since $a \leqslant x_{k_n} \leqslant b$ for each n, we have $B \in I$. Now $|x_{k_n} - y_{k_n}| \to 0$ as $n \to \infty$, and so (y_{k_n}) also converges to B. Since F is continuous on I, we have $f(x_{k_n}) \to f(B)$ as $n \to \infty$ and $f(y_{k_n}) \to f(B)$ as $n \to \infty$, which contradicts the fact that $|f(x_{k_n}) - f(y_{k_n})|$ is always $\geqslant \varepsilon$. This contradiction proves the theorem.

Theorem 9.9

Let f be continuous, real-valued, on $I = [a, b]$ ($a < b$). Then f is Riemann integrable on I.

Proof

We use 9.6 again. Let $\varepsilon > 0$ be given. We choose a $\delta > 0$ such that for all x and y in I with $|x - y| < \delta$ we have $|f(x) - f(y)| < \varepsilon/(b-a)$. We choose a partition $P = \{ x_0, \ldots, x_n \}$ of I such that, for each k, we have $x_k - x_{k-1} < \delta$. Now take a sub-interval $J = [x_{k-1}, x_k]$. We know from Section 3 that there exist c and d in J such that for all x in J we have $f(c) \leqslant f(x) \leqslant f(d)$. This means that $M_k(f) = f(d)$ and $m_k(f) = f(c)$. But $|c - d| < \delta$ and so $M_k(f) - m_k(f) = f(d) - f(c) < \varepsilon/(b-a)$. This holds for each k. Thus

$$U(P,f) - L(P,f) = \sum_{k=1}^{n} (M_k(f) - m_k(f))(x_k - x_{k-1}) < (\varepsilon/(b-a)) \sum_{k=1}^{n} (x_k - x_{k-1}) = \varepsilon .$$

This proves exactly what we need.

Theorem 9.10

Suppose that f and g are Riemann integrable on $I = [a, b]$ and that $f(x) \leqslant g(x)$ for all x in I. Then $\int_a^b f(x)\,\mathrm{d}x \leqslant \int_a^b g(x)\,\mathrm{d}x$. Suppose that h is continuous on $[a, b]$. Then

$$\left|\int_a^b h(x)\,\mathrm{d}x\right| \leqslant \int_a^b |h(x)|\,\mathrm{d}x\,.$$

Proof

The first part is easy. For any partition P of I we have $U(P, f) < U(P, g)$ and so the upper integral of f, which is the inf of the upper sums, is $\leqslant$ the upper integral of g. But the upper integral of f is the same as the integral of f, and the same thing is true of g.

The second part is also easy. Since $U(P, -h) = -L(P, h)$ and $L(P, -h) = -U(P, h)$ for any partition P, we have $\int_a^b -h(x)\mathrm{d}x = -\int_a^b h(x)\mathrm{d}x$. Therefore, $\left|\int_a^b h(x)\,\mathrm{d}x\right|$ is either $\int_a^b h(x)\,\mathrm{d}x$ or $\int_a^b -h(x)\,\mathrm{d}x$ and both of these are, by the first part, less than or equal to $\int_a^b |h(x)|\,\mathrm{d}x$.

Theorem 9.12

(i) If f is continuous on $[a, b]$ and $a < c < b$ then

$$\int_a^c f(x)\,\mathrm{d}x + \int_c^b f(x)\,\mathrm{d}x = \int_a^b f(x)\,\mathrm{d}x\,.$$

Proof

We just take a partition P of $[a, c]$ and a partition Q of $[c, b]$ and combine them to form a partition R of $[a, b]$. Remembering that all upper sums are $\geqslant$ the integral, we have $U(P, f)+U(Q, f) = U(R, f) \geqslant \int_a^b f(x)\mathrm{d}x$. This is true for any partition P of $[a, c]$ and any partition Q of $[c, b]$. Taking the inf over all partitions P of $[a, c]$ we get $\int_a^c f(x)\mathrm{d}x+U(Q, f) \geqslant \int_a^b f(x)\mathrm{d}x$. Now we take the inf over all partitions Q of $[c, b]$ and we get $\int_a^c f(x)\mathrm{d}x+\int_c^b f(x)\mathrm{d}x \geqslant \int_a^b f(x)\mathrm{d}x$.

Similarly, $L(P, f)+L(Q, f) = L(R, f) \leqslant \int_a^b f(x)\mathrm{d}x$. Taking the sup over P and Q we get $\int_a^c f(x)\mathrm{d}x+\int_c^b f(x)\mathrm{d}x \leqslant \int_a^b f(x)\mathrm{d}x$.

Definitions 9.13

If $b < a$ we just DEFINE $\int_b^a f(x)\,\mathrm{d}x = -\int_a^b f(x)\,\mathrm{d}x$. We also define $\int_a^a f(x)\,\mathrm{d}x = 0$.

With this convention, Theorem 9.12 gives

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

provided f is continuous on a closed interval containing a, b and c. Now we can prove a key result.

Theorem 9.14

(i) Let h be continuous on $\mathbb{R}$ and let $a \in \mathbb{R}$. Set $H(x) = \int_a^x h(t)\,\mathrm{d}t$. Then $H'(x) = h(x)$ for all $x \in \mathbb{R}$.

(ii) Let f be continuous on $I = [a, b]$. Set $F(x) = \int_a^x f(t)\,\mathrm{d}t$. Then F is continuous on $[a, b]$ and $F'(x) = f(x)$ for $a < x < b$.

Part (ii) is called the (first) fundamental theorem of the calculus. It also explains where the mean value theorem gets its name from. For the mean value theorem tells us that there is some c in (a, b) such that $(F(b)-F(a))/(b-a) = F'(c) = f(c)$. But this tells us that $f(c) = F(b)/(b-a)$ = (average of f on I).

Proof of 9.14

(i) For $x \neq c$ we have

$$H(x) - H(c) = \int_c^x h(t)\,\mathrm{d}t.$$

Let $\varepsilon > 0$ be given. Then we know that there is some $\delta > 0$ such that $|t-c| < \delta$ implies that $|h(t)-h(c)| < \varepsilon/2$. So if $c < x < c+\delta$ we get, using 9.10 and Problem 70,

$$(x-c)(h(c)-\varepsilon/2) = \int_c^x (h(c)-\varepsilon/2)\mathrm{d}t \leqslant \int_c^x h(t)\mathrm{d}t \leqslant \int_c^x (h(c)+\varepsilon/2)\mathrm{d}t = (x-c)\,(h(c)+\varepsilon/2)$$

and so

$$\left| \frac{H(x)-H(c)}{x-c} - h(c) \right| \leqslant \varepsilon/2 < \varepsilon.$$

The same thing works for $c-\delta < x < c$. Because we can do this for any $\varepsilon > 0$ this shows that $\lim_{x \to c} (H(x)-H(c))/(x-c) = h(c)$ and so $H'(c) = h(c)$.

(ii) To see that F is continuous on I, let M be the maximum of $|f(x)|$ on this interval. Then for $y,z \in I$ we have, using (9.10) and (9.12),

$|F(y)-F(z)| = \left| \int_z^y f(t)\,\mathrm{d}t \right| \leqslant M|y-z|$, so that obviously $F(y) \to F(z)$ as $y \to z$.

To see that F has derivative f, we set $h(x) = f(a)$ for $x < a$ and $h(x) = f(x)$ for $a \leqslant x \leqslant b$ and $h(x) = f(b)$ for $x > b$. Then h is continuous on $\mathbb{R}$. Defining H as in part (i) we have $H'(x) = h(x)$ for all x in $\mathbb{R}$. But for $a < x < b$ we have $H(x) = F(x)$ and $h(x) = f(x)$, so that $F'(x) = f(x)$.

Theorem 9.15 The second fundamental theorem of the calculus

Suppose that F, f are real-valued functions on $[a, b]$, that F is continuous and f is Riemann integrable on $[a, b]$, and that $F'(x) = f(x)$ for all x in (a, b). Then

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

Proof .

The proof is based on the mean value theorem. Let $P = \{ x_0, \ldots, x_n \}$ be any partition of $[a, b]$. Then by the mean value theorem there exist points t_k satisfying $x_{k-1} < t_k < x_k$ such that

$$F(b) - F(a) = \sum_{k=1}^{n} (F(x_k) - F(x_{k-1})) = \sum_{k=1}^{n} f(t_k)(x_k - x_{k-1}).$$

But this means that $L(P,f) \leqslant F(b)-F(a) \leqslant U(P,f)$. Hence the lower integral of f is at most $F(b)-F(a)$, and the upper integral of f is at least $F(b)-F(a)$. But the lower and upper integrals of f are, by assumption, the same.

Theorem 9.16 Change of Variables Formula

Let $a < b$ and let g be a real-valued function such that g' is continuous on $[a, b]$. Let f be real-valued and continuous on an interval $[A, B]$ containing $g([a, b])$. Then

$$\int_{g(a)}^{g(b)} f(u)\,\mathrm{d}u = \int_a^b f(g(t))g'(t)\,\mathrm{d}t .$$

Proof

We first extend f to all of $\mathbb{R}$ by setting $f(x) = f(A)$ for $x < A$ and $f(x) = f(B)$ for $x > B$. Now we set $F(y) = \int_{g(a)}^{y} f(u)\,\mathrm{d}u$. Then by 9.14 we have $F'(y) = f(y)$ for all real y. Thus F is differentiable, and so continuous, on $\mathbb{R}$. Hence $H(x) = F(g(x))$ is continuous on $[a, b]$ and for $a < x < b$ the chain rule gives us $H'(x) = f(g(x))g'(x)$. So by 9.15 we have

$$\int_a^b f(g(x))g'(x)\,\mathrm{d}x = H(b)-H(a) = F(g(b))-F(g(a)) = \int_{g(a)}^{g(b)} f(u)\,\mathrm{d}u .$$

9.17 Improper integrals

Suppose that we have a function f continuous and real-valued on the interval $[a, +\infty)$. We cannot integrate f on this interval using the above method, because we would need infinitely many vertices x_k in a partition. However, if $a < L < +\infty$ the integral of f from a to L certainly exists. So we set

$$\int_a^{+\infty} f(x)\,\mathrm{d}x = \lim_{L \to +\infty}\left(\int_a^L f(x)\,\mathrm{d}x\right)$$

if this limit exists AND IS FINITE. In this case, we say that the integral converges, and in all other cases (no limit, or infinite limit) we say that the integral diverges. Such an integral is called an improper integral of the first kind.

It is fairly obvious that if $a < b < +\infty$ then $\int_a^{+\infty} f(x)\,\mathrm{d}x$ converges iff $\int_b^{+\infty} f(x)\,\mathrm{d}x$ converges. This is because $\int_a^L f(x)\,\mathrm{d}x = \int_a^b f(x)\,\mathrm{d}x + \int_b^L f(x)\,\mathrm{d}x$.

Examples 9.18

1. $\int_0^{+\infty} \mathrm{e}^{Ax}\,\mathrm{d}x$ ($A \in \mathbb{R}$). We look at (for $A \neq 0$) $\int_0^L \mathrm{e}^{Ax}\,\mathrm{d}x = (\mathrm{e}^{AL}-1)/A$. If $A > 0$ then this tends to $+\infty$ as $L \to +\infty$, and so the improper integral diverges. If $A < 0$ then $(\mathrm{e}^{AL}-1)/A \to -1/A$ as $L \to +\infty$ and the improper integral converges, its value being $-1/A$. Clearly $A = 0$ gives $\int_0^L 1\,\mathrm{d}x = L$ and the improper integral diverges.

2. $\int_1^{+\infty} 1/x^p\,\mathrm{d}x$ ($p \in \mathbb{R}$). If $p \neq 1$ then $\int_1^L 1/x^p\,\mathrm{d}x = (L^{1-p}-1)/(1-p)$ which, as $L \to +\infty$, tends to

$+\infty$ if if $p < 1$ and to $-1/(1-p)$ if $p > 1$. For $p = 1$ we get $\int_1^L 1/x \, dx = \log L \to +\infty$ as $L \to +\infty$. So this improper integral diverges unless $p > 1$, in which case its value is $1/(p-1)$.

Theorem 9.19

Let f be non-negative and continuous on $[a, +\infty)$. Then $\int_a^{+\infty} f(x) \, dx$ converges if and only if the following is true. There is a positive constant M such that for every L with $a < L < +\infty$ we have $\int_a^L f(x) \, dx \leqslant M$.

This is analogous to the theorem that an infinite series with non-negative terms converges iff the partial sums are bounded above. The **Proof** just consists of noting that $h(L) = \int_a^L f(x) \, dx$ is non-decreasing on $[a, +\infty)$ and so has a limit as $L \to +\infty$. This limit is finite iff h is bounded above.

Theorem 9.20 The Integral Test

Suppose that f is a function continuous, non-negative and non-increasing on $[1, +\infty)$. Then $\int_1^{+\infty} f(x) \, dx$ converges iff the series $\sum_{k=1}^{\infty} f(k)$ converges.

Proof

We just note that, if $N \geqslant 2$ is an integer, then

$$\int_1^N f(x) \, dx = \sum_{k=2}^{N} \left(\int_{k-1}^{k} f(x) \, dx \right).$$

Also

$$f(k) \leqslant \int_{k-1}^{k} f(x) \, dx \leqslant f(k-1).$$

So

$$\sum_{k=2}^{N} f(k) \leqslant \int_1^N f(x) \, dx \leqslant \sum_{k=1}^{N-1} f(k). \tag{1}$$

Suppose then that the improper integral converges. Then there is a constant M such that the integral from 1 to N is always $\leqslant M$, and so the partial sums of the series are bounded above.

Suppose now that the series converges. Then the partial sums of the series are bounded above, by K, say. Take $L > 1$ and choose an integer $N > L$. Then we see from (1) that the integral from 1 to L is bounded above by K, and so the improper integral converges.

As an **example**, we see at once that 9.19 proves that the series $\sum_{k=1}^{\infty} k^{-p}$ converges if $p > 1$ and diverges otherwise.

Theorem 9.21 The comparison test

Suppose that f and g are continuous real-valued functions on $I = [a, +\infty)$ with $|f(t)| \leqslant g(t)$ for all t in I. If $\int_a^{+\infty} g(t)\,dt$ converges, then $\int_a^{+\infty} f(t)\,dt$ converges, and $\left|\int_a^{+\infty} f(t)\,dt\right| \leqslant \int_a^{+\infty} g(t)\,dt$.

Proof

We know from 9.19 that there is a constant M such that $\int_a^L g(t)\,dt \leqslant M$ for all $L > a$. Thus, for all $L > a$, $\int_a^L |f(t)|\,dt \leqslant M$, and so $\int_a^{+\infty} |f(t)|\,dt$ converges. Also $0 \leqslant f(t)+|f(t)| \leqslant 2g(t)$ for all $t \geqslant a$, so that $\int_a^{+\infty} f(t)+|f(t)|\,dt$ converges.

So there are real numbers A and B such that $\int_a^L f(t)+|f(t)|\,dt \to A$ and $\int_a^L |f(t)|\,dt \to B$ as $L \to +\infty$. Thus $\int_a^L f(t)\,dt \to A-B$ as $L \to +\infty$. Here we have used Problem 71.

The last assertion just follows from the fact that $\left|\int_a^L f(t)\,dt\right| \leqslant \int_a^L |f(t)|\,dt \leqslant \int_a^L g(t)\,dt$.

Examples 9.22

1. Consider $\int_0^{+\infty} e^{-t^2}\,dt$. It suffices to consider $\int_1^{+\infty} e^{-t^2}\,dt$. On the range of integration we now have $e^{-t^2} \leqslant e^{-t}$. So using 9.19 we see that the required improper integral does converge.

2. Consider $\int_1^{+\infty} e^{1/t}t^{-1/2}\,dt$. The integrand is here $\geqslant t^{-1/2}$, the corresponding integral of which diverges. Thus this integral diverges.

3. Consider $\int_1^{+\infty} \frac{\log t}{1+t^4}\,dt$. We can do this one two ways. The first is to note that for $t \geqslant 1$ we have $(\log t)/(1+t^4) \leqslant (\log t)/t^4$ and $(\log t)/t$ tends to 0 as $t \to +\infty$. So there is some M such that $(\log t)/(1+t^4) \leqslant t^{-3}$ for $t \geqslant M$. Thus $\int_M^{+\infty} (\log t)/(1+t^4)\,dt$ converges and therefore so does the required integral. The second way is to start as before, and then to CALCULATE $\int_1^L (\log t)/t^4\,dt$ using integration by parts, and finally to look at the limit as $L \to +\infty$ (details omitted).