Analysis I

Prof. T. W. Körner

Lent 2003

Contents

Lecturer's preamble

This course centres around 1 idea, the re-founding of calculus on a rigorous basis.
This is essentially based on the definition of a limit:

$$\forall \epsilon > 0 \; \exists n_0(\epsilon) \text{ such that } |a_n - a| < \epsilon \; \forall n \geq n_0(\epsilon)$$

From this we get the FUNDAMENTAL THEOREM OF ANALYSIS
Every strictly increasing sequence bounded above tends to a limit
This does not work in $\mathbb{Q}$, e.g. look at the decimal expansion of $\sqrt{2}$, as you add more and more digits it is clearly increasing, it is bounded above by 2, but its limit ($\sqrt{2}$) does not exist in $\mathbb{Q}$. However, it always works on $\mathbb{R}$
Using this axiom and the laws of algebra (e.g. $a + b = b + a$), we can now refound the calculus.

NOTE: We can now split proofs and results into two groups, those which are mere algebra (i.e. proof also works on $\mathbb{Q}$) and those using analysis.

1 Why do we bother?

It is surprising how many people think that analysis consists in the difficult proofs of obvious theorems. All we need know, they say, is what a limit is, the definition of continuity and the definition of the derivative. All the rest is 'intuitively clear'.
If pressed they will agree that these definitions apply as much to the rationals $\mathbb{Q}$ as to the real numbers $\mathbb{R}$. They then have to explain the following interesting example.

Example 1.1. If $f : \mathbb{Q} \to \mathbb{Q}$ is given by

$$f(x) = -1 \qquad \text{if } x^2 < 2,$$
$$f(x) = 1 \qquad \text{otherwise},$$

then
(i) f is a continuous function with $f(0) = -1$, $f(2) = 1$ yet there does not exist a c with $f(c) = 0$,
(ii) f is a differentiable function with $f'(x) = 0$ for all x yet f is not constant.

What is the difference between $\mathbb{R}$ and $\mathbb{Q}$ which makes calculus work on one even though it fails on the other? Both are 'ordered fields', that is, both support operations of 'addition' and 'multiplication' together with a relation 'greater than' ('order') with the properties that we expect. If the reader is interested she will find a complete list of the appropriate axioms in texts like the altogether excellent book of Spivak [5] and its many rather less excellent competitors, but, interesting as such things may be, they are irrelevant to our purpose which is not to consider the shared properties of $\mathbb{R}$ and $\mathbb{Q}$ but to identify a *difference*

between the two systems which will enable us to exclude the possibility of a function like that of Example 1.1 for functions from $\mathbb{R}$ to $\mathbb{R}$.
To state the difference we need only recall a definition from course C3.

Definition 1.2. If $a_n \in \mathbb{R}$ for each $n \geq 1$ and $a \in \mathbb{R}$ then we say that $a_n \to a$ if given $\epsilon > 0$ we can find an $n_0(\epsilon)$ such that

$$|a_n - a| < \epsilon \text{ for all } n \geq n_0(\epsilon)$$

The key property of the reals, the *fundamental axiom* which makes everything work was also stated in the course C3.

Axiom 1.3 (The fundamental axiom of analysis)**.** If $a_n \in \mathbb{R}$ for each $n \geq 1$, $A \in \mathbb{R}$ and $a_1 \leq a_2 \leq a_3 \leq \ldots$ and $a_n < A$ for each n then there exists an $a \in \mathbb{R}$ such that $a_n \to a$ as $n \to \infty$.

Less ponderously, and just as rigorously, the fundamental axiom for the real numbers says *every increasing sequence bounded above tends to a limit.*
Everything which depends on the fundamental axiom is analysis, everything else is mere algebra.

2 The axiom of Archimedes

We start by proving the following results on limits, some of which you saw proved in course C3.

Lemma 2.1. (i) The limit is unique. That is, if $a_n \to a$ and $a_n \to b$ as $n \to \infty$ then $a = b$.
(ii) If $a_n \to a$ as $n \to \infty$ and $n(1) < n(2) < n(3) \ldots$ then $a_{n(j)} \to a$ as $j \to \infty$.
(iii) If $a_n = c$ for all n then $a_n \to c$ as $n \to \infty$.
(iv) If $a_n \to a$ and $b_n \to b$ as $n \to \infty$ then $a_n + b_n \to a + b$.
(v) If $a_n \to a$ and $b_n \to b$ as $n \to \infty$ then $a_n b_n \to ab$.
(vi) If $a_n \to a$ as $n \to \infty$, $a_n \neq 0$ for each n and $a \neq 0$, then $a_n^{-1} \to a^{-1}$.
(vii) If $a_n \leq A$ for each n and $a_n \to a$ as $n \to \infty$ then $a \leq A$.

Proof. i) If $b \neq a$ then $|b - a| > 0$ so set $\epsilon = |\frac{b-a}{4}|$
There exists an n_0 such that $|a_n - a| < \epsilon \ \forall n > n_0$
There exists an n_1 such that $|a_n - b| < \epsilon \ \forall n > n_1$
Set $n_2 = \max\{n_0, n_1\}$
We have $|a_{n_2} - a| < \epsilon$ and $|a_{n_2} - b| < \epsilon$ so $|a - b| < 2\epsilon = |\frac{a-b}{2}| < |a - b| \#$

ii) Suppose $a_n \to a$
Then given $\epsilon > 0$ $\exists n_0$ such that $|a_n - a| < \epsilon \ \forall n \geq n_0$
If $j \geq n_0$ then $n(j) \geq n_0$, generates a subsequence, so $|a_{n(j)} - a| < \epsilon$, thus $a_{n(j)} \to a$

iii) If $_n = c$ then given $\epsilon > 0$ set $n_0(\epsilon) = 1$.
Then $|a_n - c| = |c - c| = 0 < \epsilon \ \forall n \geq n_0(\epsilon)$ as required

iv) We know that
Given $\epsilon > 0 \ \exists n_0(\epsilon)$ such that $|a_n - a| < \epsilon \ \forall n \geq n_0(\epsilon)$
Given $\epsilon > 0 \ \exists n_1(\epsilon)$ such that $|b_n - b| < \epsilon \ \forall n \geq n_1(\epsilon)$

Thus if $n \geq n_2(\epsilon) = \max\{n_0(\frac{\epsilon}{2}), n_1(\frac{\epsilon}{2})\}$

$$\begin{aligned} |(a_n + b_n) - (a - b)| &= |(a_n - a) + (b_n - b)| \\ &\leq |a_n - a| + |b_n - b| \\ &< \frac{\epsilon}{2} + \frac{\epsilon}{2} \end{aligned}$$

for $n \geq n_2(\epsilon)$

v) As before, we have that
There exists an n_0 such that $|a_n - a| < \epsilon \ \forall n > n_0$
There exists an n_1 such that $|a_n - b| < \epsilon \ \forall n > n_1$

Thus if $n \geq n_2(\frac{\epsilon}{3}) = \max\{n_1(1), n_0(\frac{\epsilon}{2(|b|+1)}), n_1(\frac{\epsilon}{2(|a|+1)})\}$

$$\begin{aligned} \text{Then } |a_n b_n - ab| &\leq |a_n b_n - ab_n + ab_n - ab| \\ &= |(a_n - a)b_n + a(b_n - b)| \\ &\leq |(a_n - a)b_n| + |a(b_n - b)| \\ &= |a_n - a||b_n| + |a||b_n - b| \\ &\leq |a_n - a|(|b| + 1) + |a||b_n - b| \\ &< \frac{\epsilon}{2(|b| + 1)} + \frac{|a|\epsilon}{2(|a| + 1)} \\ &\leq \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon \end{aligned}$$

Note: If $n \geq n_1(1), |b_n - b| < 1$ so $|b_n| < |b| + 1$
We use $\frac{\epsilon}{2(|a|+1)}$ instead of $\frac{\epsilon}{2|a|}$ in case $a = 0$

vi) Again, given $\epsilon > 0 \ \exists n_0(\epsilon)$ such that $|a_n - a| < \epsilon$
Thus if $n \geq n_1(\epsilon) = \max\{n_0(\frac{|a|}{2}), n_0(\frac{\epsilon|a|^2}{2})\}$

$$\left|\frac{1}{a_n}-\frac{1}{a}\right| = \left|\frac{a-a_n}{aa_n}\right| = \frac{|a-a_n|}{|a||a_n|}$$
$$< \frac{2|a_n-a|}{|a|^2}$$
$$< \frac{\epsilon|a|^2}{2}.(\frac{2}{|a|^2}) = \epsilon$$

Note: If $n \geq n_1(\epsilon)$ then $|a_n - a| < \frac{|a|}{2}$ so $|a_n| > \frac{|a|}{2}$

vii) If $a > A$ then set $\epsilon = \frac{a-A}{2}$. Then $a - a_n \geq a - A > \epsilon\ \forall n$, so $a_n \not\to a$ #

[Remark: $a_n < A, a_n \to a \not\Rightarrow a < A$, observe $-\frac{1}{n} < 0$ but $-\frac{1}{a_n} \to 0$] $\square$

We need the following variation on the fundamental axiom.

Exercise 2.2. A decreasing sequence of real numbers bounded below tends to a limit.

[Hint. If $a \leq b$ then $-b \leq -a$.]

Proof. If $b_1 \geq b_2 \geq b_3 \ldots$ and $b_n \geq B$ then $-b_1 \leq -b_2 \leq -b_3 \ldots$ and $-b_n \leq -B$ so by the fundamental axiom $-b_n \to \beta$, say, so $b_n = (-1).(-b_n) \to (-1).\beta = -\beta$ $\square$

Useful as the results of Lemma 2.1 are, they are also true of sequences in $\mathbb{Q}$. They are therefore mere, if important, algebra. Our first truly 'analysis' result may strike the reader as rather odd.

Theorem 2.3 (Axiom of Archimedes)**.**

$$\frac{1}{n} \to 0 \text{ as } n \to \infty$$

Proof. The sequence $1, \frac{1}{2}, \frac{1}{3}, \ldots$ is decreasing, and bounded below by 0, so $\frac{1}{n} \to l$ for some l.
Observe that $\frac{1}{2n}$ is a subsequence of $\frac{1}{n}$ so $\frac{1}{2n} \to l$.
But $\frac{1}{2n} = \frac{1}{2}\frac{1}{n} \to \frac{1}{2}l = \frac{l}{2}$
$\therefore l = \frac{l}{2} \Rightarrow l = 0$ $\square$

Theorem 2.3 shows that there is no 'exotic' real number $\beth$ say (to choose an exotic symbol) with the property that $1/n > \beth$ for all integers $n \geq 1$ yet $\beth > 0$ (that is, $\beth$ is strictly positive and yet smaller than all strictly positive rationals). There exist number systems with such exotic numbers (the famous 'non-standard analysis' of Abraham Robinson and the 'surreal numbers' of Conway constitute two such systems) but, just as the rationals are, in some sense, too small a system for the standard theorems of analysis to hold so these non-Archimedean systems are, in some sense, too big. Archimedes and Eudoxus realised the need for an axiom to show that there is no exotic number $\daleth$ bigger than any integer[1]

[1] Footnote for passing historians, this is a course in mathematics.

(i.e. $\daleth > n$ for all integers $n \geq 1$; to see the connection with our form of the axiom consider $\gimel = 1/\daleth$). However, in spite of its name, what was an axiom for Archimedes is a theorem for us.

Theorem 2.4. Given any real number K we can find an integer n with $n > K$.

Proof. Assume we can find a K that is greater than all numbers. Clearly it is greater than 1.
Thus $0 < \frac{1}{K} \leq \frac{1}{n}$, and thus $\frac{1}{n} \not\to 0 \#$
By reductio ad absurdum, $\nexists K$ such that $K \geq n \ \forall n \in \mathbb{Z}$ $\square$

3 Series and sums

There is no need to restrict the notion of a limit to real numbers.

Definition 3.1. If $a_n \in \mathbb{C}$ for each $n \geq 1$ and $a \in \mathbb{C}$ then we say that $a_n \to a$ if given $\epsilon > 0$ we can find an $n_0(\epsilon)$ such that

$$|a_n - a| < \epsilon \text{ for all } n \geq n_0(\epsilon)$$

Exercise 3.2. We work in $\mathbb{C}$.
(i) The limit is unique. That is, if $a_n \to a$ and $a_n \to b$ as $n \to \infty$ then $a = b$.
(ii) If $a_n \to a$ as $n \to \infty$ and $n(1) < n(2) < n(3) \dots$ then $a_{n(j)} \to a$ as $j \to \infty$.
(iii) If $a_n = c$ for all n then $a_n \to c$ as $n \to \infty$.
(iv) If $a_n \to a$ and $b_n \to b$ as $n \to \infty$ then $a_n + b_n \to a + b$.
(v) If $a_n \to a$ and $b_n \to b$ as $n \to \infty$ then $a_n b_n \to ab$.
(vi) If $a_n \to a$ as $n \to \infty$, $a_n \neq 0$ for each n and $a \neq 0$, then $a_n^{-1} \to a^{-1}$.

Exercise 3.3. Explain why there is no result in Exercise 3.2 corresponding to part (vii) of Lemma 2.1.

We illustrate some of the ideas introduced by studying infinite sums.

Definition 3.4. We work in $\mathbb{F}$ where $\mathbb{F} = \mathbb{R}$ or $\mathbb{F} = \mathbb{C}$. If $a_j \in \mathbb{F}$ we say that $\sum_{j=1}^{\infty} a_j$ converges to s if

$$\sum_{j=1}^{N} a_j \to s$$

as $N \to \infty$. We write $\sum_{j=1}^{\infty} a_j = s$.
If $\sum_{j=1}^{N} a_j$ does not tend to a limit as $N \to \infty$, we say that the sum $\sum_{j=1}^{\infty} a_j$ diverges.

Note: Work as much as possible with $\sum_1^N a_n$ and as little as possible with $\sum_1^\infty a_n$ until you know that $\sum_1^\infty a_n$ exists

Example: Let $u_n = (-1)^{n+1}, S = \sum_1^\infty u_n$
i) $S = u_1 + \sum_{n=2}^\infty u_n = u_1 + \sum_{n=1}^\infty (-1) u_n = u_1 - \sum_{n=1}^\infty u_n = 1 - S$ so $S = \frac{1}{2}$
ii) $S = \sum_{r=1}^\infty (u_{2r-1} + u_{2r}) = \sum_{r=1}^\infty (1-1) = \sum_{r=1}^\infty 0 = 0$

The above logic is faulty because S does not exist
$\sum_1^N u_n = 1 - 1 + \cdots = \begin{cases} 0 & \text{if } N \text{ even} \\ 1 & \text{if } N \text{ odd} \end{cases}$ does not converge

Lemma 3.5. We work in $\mathbb{F}$ where $\mathbb{F} = \mathbb{R}$ or $\mathbb{F} = \mathbb{C}$.
(i) Suppose $a_j, b_j \in \mathbb{F}$ and $\lambda, \mu \in \mathbb{F}$. If $\sum_{j=1}^\infty a_j$ and $\sum_{j=1}^\infty b_j$ converge then so does $\sum_{j=1}^\infty \lambda a_j + \mu b_j$ and

$$\sum_{j=1}^\infty \lambda a_j + \mu b_j = \lambda \sum_{j=1}^\infty a_j + \mu \sum_{j=1}^\infty b_j$$

(ii) Suppose $a_j, b_j \in \mathbb{F}$ and there exists an N such that $a_j = b_j$ for all $j \geq N$. Then, either $\sum_{j=1}^\infty a_j$ and $\sum_{j=1}^\infty b_j$ both converge or they both diverge. (In other words, initial terms do not matter.)

Proof. i) $\sum_{j=1}^N (\lambda a_j + \mu b_j) = \lambda \sum_{j=1}^N a_j + \mu \sum_{j=1}^N b_j \to \lambda \sum_{j=1}^\infty a_j + \mu \sum_{j=1}^\infty b_j$
ii) Suppose $a_j = b_j$ for $j \geq M$, then $a_j = b_j + c_j$ where $c_j = 0$ for $j \geq M$, so $\sum_1^N c_j = \sum_1^M c_j$ for $N \geq M \to \sum_1^M c_j$
So $\sum_1^\infty c_j$ exists, so if $\sum_1^\infty b_j$ exists, then $\sum_1^\infty (b_j + c_j)$ exists, $\therefore \sum_1^\infty a_j$ exists $\square$

Exercise 3.6. Any problem on sums $\sum_{j=1}^\infty a_j$ can be converted into one on sequences by considering the sequence $s_n = \sum_{j=1}^n a_j$. Show conversely that a sequence s_n converges if and only if, when we set $a_1 = s_1$ and $a_n = s_n - s_{n-1}$ $[n \geq 2]$ we have $\sum_{j=1}^\infty a_j$ convergent. What can you say about $\lim_{n\to\infty} \sum_{j=1}^n a_j$ and $\lim_{n\to\infty} s_n$ if both exist?

The following results are fundamental to the study of sums.

Theorem 3.7 (The comparison test)**.** We work in $\mathbb{R}$. Suppose that $0 \leq b_j \leq a_j$ for all j. Then, if $\sum_{j=1}^\infty a_j$ converges, so does $\sum_{j=1}^\infty b_j$.

Proof. Observe that $\sum_{j=1}^N a_j$ is increasing, and recall that if $c_N \leq k\ \forall N$ and $c_N \to c$ then $c \leq k$, It follows that $\sum_{j=1}^N a_j \leq \sum_{j=1}^\infty a_j$
$0 \leq \sum_{j=1}^N b_j \leq \sum_{j=1}^N a_j \leq \sum_{j=1}^\infty a_j$. $\sum_{j=1}^N b_j$ is an increasing sequence, since $b_j \geq 0$
so since it is bounded above the fundamental theorem tells us that $\sum_{j=1}^N b_j$ tends to a limit as $n \to \infty$ i.e. $\sum_{j=1}^\infty b_j$ exists $\square$

Note that was can use this an another proof of summing a geometric progression. Let $|x| < 1$

$$\sum_{n=0}^N x^n = \frac{1-x}{1-x} \sum_{n=0}^N x^n = \frac{1 - x + x - x^2 + x^2 - \ldots - x^{N+1}}{1-x}$$
$$= \frac{1 - x^{N+1}}{1-x} \to \frac{1}{1-x} \text{ as } N \to \infty$$

Now look at $\sum_{n=1}^{N} \frac{x^n}{n^2+3+\sin x}$ for $0 \leq x < 1$
Since $0 \leq \frac{x^n}{n^2+3+\sin x} \leq x^n, \sum_{n=1}^{\infty} x^n$ exists, $\Rightarrow \sum_{n=1}^{\infty} \frac{x^n}{n^2+3+\sin x}$ exists.

Theorem 3.8. We work in $\mathbb{F}$ where $\mathbb{F} = \mathbb{R}$ or $\mathbb{F} = \mathbb{C}$. If $\sum_{j=1}^{\infty} |a_j|$ converges, then so does $\sum_{j=1}^{\infty} a_j$.

This is saying that, if we take a sequence, and plot the points of the sequence and join them up, $\sum_{j=1}^{\infty} |a_j|$ converges means the length of the line produced is bounded, and thus it converges to a point.

Proof. First consider it over $\mathbb{R}$
Set $a_j^+ = \begin{cases} a_j & \text{if } a_j \geq 0 \\ 0 & \text{if } a_j < 0 \end{cases}$
Set $a_j^- = \begin{cases} 0 & \text{if } a_j \geq 0 \\ -a_j & \text{if } a_j < 0 \end{cases}$ $\therefore a_j = a_j^+ - a_j^-$ and $0 \leq a_j^+ |a_j|, 0 \leq a_j^- \leq |a_j|$
$\sum_1^{\infty} a_j^+, \sum_1^{\infty} a_j^-$ converge because $\sum_1^{\infty} |a_j|$ does, using the comparison test. Thus
$\sum_{j=1}^{\infty} a_j = \sum_{j=1}^{\infty} (a_j^+ - a_j^-)$ convergses to $\sum_1^{\infty} a_j^+ - \sum_1^{\infty} a_j^-$
If $\mathbb{F} = \mathbb{C}$ with $a_j = x_j + iy_j, x_j, y_j \in \mathbb{R}$
$|x_j| \leq |a_j|$ so by comparison $\sum_1^{\infty} |x_j|$ converges, and so by the first part $\sum_1^{\infty} x_j$ also converges. Similarly $\sum_1^{\infty} y_j$ converges
So $\sum_{j=1}^{\infty} a_j = \sum_{j=1}^{\infty} (x_j = iy_j)$ converges to $\sum_{j=1}^{\infty} x_j + i \sum_{j=1}^{\infty} y_j$ $\square$

Theorem 3.8 is often stated using the following definition.

Definition 3.9. We work in $\mathbb{F}$ where $\mathbb{F} = \mathbb{R}$ or $\mathbb{F} = \mathbb{C}$. If $\sum_{j=1}^{\infty} |a_j|$ converges we say that the sum $\sum_{j=1}^{\infty} a_j$ is absolutely convergent.

Theorem 3.8 then becomes the statement that absolute convergence implies convergence.
Here is a trivial but useful consequence of Theorems 3.7 and 3.8.

Lemma 3.10 (Ratio test)**.** Suppose $a_j \in \mathbb{C}$ and $|a_{j+1}/a_j| \to l$ as $j \to \infty$. If $l < 1$, then $\sum_{j=1}^{\infty} |a_j|$ converges. If $l > 1$, then $\sum_{j=1}^{\infty} a_j$ diverges.

Of course Lemma 3.10 tells us nothing if $l = 1$ or l does not exist.

Proof. If $|\frac{a_{j+1}}{a_j}| \to l < 1$ then choose $\epsilon = \frac{1-l}{2}$. $\exists j(0)$ such that $\left| |\frac{a_{j+1}}{a_j}| - l \right| < \epsilon$ for $j \geq j(0)$
so $|\frac{a_{j+1}}{a_j} < l + \epsilon = \frac{1+l}{2} = K < 1$ for $j \geq j(0)$, so $|a_j| \leq |a_{j(0)}| K^{j-j(0)}$ for $j \geq j(0)$
We know that $\sum_1^{\infty} |a_{j(0)} K^{j-j(0)}$ converges (as it is a geometric progression with ratio $0 \leq K < 1$). We also know that the first $j(0)$ terms are irrelevant to convergence, and we know that

$$|a_j| \leq |a_{j(0)}| K^{j-j(0)} \quad j \geq j(0)$$

so by comparison $\sum |a_j|$ converges

To prove the second part (if $|\frac{a_{n+1}}{a_n}| \to l$, $l > 1$ then $\sum a_n$) set $\epsilon = \frac{l-1}{2}$. $\exists N$ such that $\left||\frac{a_n+1}{a_n}| - l\right| < \epsilon = \frac{l-1}{2}$ for $n \geq N$, so $|\frac{a_n+1}{a_n}| > \frac{l+1}{2}$ for $n \geq N$
$|a_n| > K(\frac{l+1}{2})^n$ for $n \geq N$, so $|a_n|$ unbounded, so $\sum a_n$ cannot converge $\square$

Sums which are not absolutely convergent are much harder to deal with in general. It is worth keeping in mind the following trivial observation.

Lemma 3.11. We work in $\mathbb{F}$ where $\mathbb{F} = \mathbb{R}$ or $\mathbb{F} = \mathbb{C}$. If $\sum_{j=1}^{\infty} a_j$ converges, then $a_j \to 0$ as $j \to \infty$.

Proof. Set $S_n = \sum_{j=1}^{n} a_j$. $a_{n+1} = s_{n+1} - s_n \to \sum_1^{\infty} a_j - \sum_1^{\infty} a_j = 0$ $\square$

At a deeper level the following result is sometimes useful.

Lemma 3.12 (Alternating series test)**.** We work in $\mathbb{R}$. If we have a decreasing sequence of positive numbers a_n with $a_n \to 0$ as $n \to \infty$, then $\sum_{j=1}^{\infty}(-1)^{j+1}a_j$ converges. Further

$$\left|\sum_{j=1}^{n}(-1)^{j+1}a_j - \sum_{j=1}^{\infty}(-1)^{j+1}a_j\right| \leq |a_{N+1}|$$

for all $N \geq 1$.

Proof. Set

$$\begin{aligned} S_1 &= a_1 \\ S_2 &= a_1 - a_2 \\ S_3 &= a_1 - a_2 + a_3 \\ &\vdots \end{aligned}$$

Obeserve that $S_{2N+1} = S_{2N-1} + (-a_{2N} + a_{2N+1}) \leq S_{2N-1}$ so the sequence S_{2N+1} is decreasing. But also

$$\begin{aligned} S_{2N+1} &= a_1 - a_2 + a_3 - a_4 + \cdots \\ &= (a_1 - a_2) + (a_3 - a_4) + \ldots + (a_{2N-1} - a_{2N}) + a_{2N+1} \\ &\geq a_1 - a_2 \end{aligned}$$

So S_{2N+1} is a decreasing sequence bounded below, so $S_{2N+1} \to l$. Similarly S_{2N} is an increasing sequence bounded above so $S_{2N} \to l'$. Now $a_{2N+1} = S_{2N+1} - S_{2N} \to l - l'$ and $a_{2N+1} \to 0$ so $l - l' = 0$ so $l = l'$ and $S_n \to l$ as $n \to \infty$
To finish observe that S_{2N+1} is decreasing, S_{2N} is increasing, so $S_{2N+1} \geq l \geq S_{2N}$ so $|l - S_{2N}| \leq |S_{2N} - S_{2N+1}| = a_{2N+1}$. Similarly $|l - S_{2N+1}| \leq a_{2N+2}$ $\square$

The last lemma is sometimes expressed by saying 'the error caused by replacing a convergent infinite alternating sum of decreasing terms by the sum of its first few terms is no greater than the absolute value of the first term neglected'.
Later we will give another test for convergence called the integral test (Lemma 14.4) from which we deduce the result known to many of you that $\sum_{n=1}^{\infty} n^{-1}$ diverges. I will give another proof in the next example.

Example 3.13. (i) $\displaystyle\sum_{n=1}^{\infty} \frac{1}{n}$ diverges.
(ii) $\displaystyle\sum_{n=1}^{\infty} \frac{(-1)^n}{n}$ is convergent but not absolutely convergent.
(iii) If $v_{2n} = 1/n$, $v_{2n-1} = -1/(2n)$ then $\sum_{n=1}^{\infty} v_n$ is not convergent.

Proof. i) This shows that $a_n \to 0 \not\Rightarrow \sum a_n$ converges

$$\begin{aligned}
1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \cdots &= 1 + \frac{1}{2} + (\frac{1}{3} + \frac{1}{4}) + (\frac{1}{5} + \frac{1}{6} + \frac{1}{7} + \frac{1}{8}) \\
&\quad + \cdots + (\frac{1}{2^{N-1}+1} + \frac{1}{2^{N-1}+2} + \cdots + \frac{1}{2^N}) + \cdots \\
&\geq 1 + \frac{1}{2} + (\frac{1}{4} + \frac{1}{4}) + (\frac{1}{8} + \frac{1}{8} + \frac{1}{8} + \frac{1}{8}) \\
&\quad + \cdots + (\frac{1}{2^n} + \frac{1}{2^n} + \frac{1}{2^n} + \cdots + \frac{1}{2^n}) + \cdots \\
&\geq 1 + \frac{1}{2} + \frac{1}{2} + \frac{1}{2} + \cdots
\end{aligned}$$

So $\sum_1^N \frac{1}{r}$ is not bounded and diverges
Note: $1 + \frac{1}{2} + \frac{1}{3} + \cdots = 1 + (\frac{1}{2} + \frac{1}{3}) + (\frac{1}{4} + \frac{1}{5} + \frac{1}{6} + \frac{1}{7}) + \cdots \leq 1 + (\frac{1}{2} + \frac{1}{2}) + (\frac{1}{4} + \frac{1}{4} + \frac{1}{4} + \frac{1}{4}) + \cdots + \frac{1}{2^N} \leq N + 1$. Note that $2^20 \approx 10^6$, so $\sum_1^{10^6} \leq 21$ so it is very slow to diverge
ii) $\frac{1}{n}$ is decreasing, $\frac{1}{n} \to 0$ and $\frac{1}{n} \geq 0$ so we can apply the alternating series test. Thus $\sum \frac{(-1)^n}{n}$ is convergent, and i) tells us that is it not absolutely convergent.
iii) This is similar to i)

$$\begin{aligned}
v_1 + v_2 + \ldots + v_{2n} &= (v_1 + v_2) + (v_3 + v_4) + \cdots \\
&= (1 - \frac{1}{2}) + (\frac{1}{2} - \frac{1}{4}) + \cdots + (\frac{1}{n} - \frac{1}{2n}) \\
&= \frac{1}{2}(1 + \frac{1}{2} + \frac{1}{3} + \cdots) \to \infty
\end{aligned}$$

So $\sum v_j$ does not converge $\square$

4 Least upper bounds

A non-empty bounded set in $\mathbb{R}$ need not have a maximum.

Example 4.1. The set $E = \{-1/n : n \geq 1\}$ is non-empty and any $e \in E$ satisfies the inequalities $-1 \leq e \leq 0$ but E has no largest member.

Proof. If $x \leq 0$ then $\exists n$ such that $-x > \frac{1}{n} > 0$ so $-\frac{1}{n} > x$ and x is not a maximum
If $x \geq 0$ then $x \notin E$
Thus E contains no largest member $\square$

However, as we shall see every non-empty bounded set in $\mathbb{R}$ has a least upper bound (or supremum).

Definition 4.2. Let E be a non-empty set in $\mathbb{R}$. We say that α is a least upper bound for E if

> (i) $\alpha \geq e$ for all $e \in E$ [that is, α is an upper bound for E] and (ii) If $\beta \geq e$ for all $e \in E$ then $\beta \geq \alpha$ [that is, α is the least such upper bound]

If E has a supremum α we write $\sup_{e \in E} e = \sup E = \alpha$.

Lemma 4.3. If the least upper bound exists it is unique.

Proof. Suppose E is a non-empty subset of $\mathbb{R}$ and α, α' are least upper bounds for E. Then since α' is an upper bound, and α is a least upper bound $\alpha \leq \alpha'$. Similarly $\alpha' \leq \alpha$ and so $\alpha' = \alpha$ $\square$

The following remark is trivial but sometimes helpful.

Lemma 4.4. Let E be a non-empty set in $\mathbb{R}$. Then α is a least upper bound for E if and only if we can find $e_n \in E$ with $e_n \to \alpha$ and b_n such that $b_n \geq e$ for all $e \in E$ and $b_n \to \alpha$ as $n \to \infty$.

Proof. First show that α is an upper bound.
If $e \in E$ then $b_n \geq e$. But $b_n \to \alpha$ so $\alpha \geq e$.
Then show α is the least upper bound i.e. if β is an upper bound $\beta \geq \alpha$
If β is an upper bound then, in particular, $\beta \geq e_n$, but $e_n \to \alpha$ so $\beta \geq \alpha$ $\square$

Here is the promised result.

Theorem 4.5. Any non-empty set in $\mathbb{R}$ with an upper bound has a least upper bound.

Proof. Let E be such a set. Since E is non-empty, we can find $a_0 \in E$. Observe that $\exists e \in E$ such that $e \geq a_0$ (e.g. a_0 itself). Since E is bounded, we can find b_0 such that $b_0 \geq e \; \forall e \in E$. Set $c_0 = \frac{a_0+b_0}{2}$.
Either $\exists e \in E$ such that $e \geq c_0$ and we put $a_1 = c_0, b_1 = b_0$
Or $\forall e \in E$ $c_0 > e$ and we put $a_1 = a_0, b_1 = c_0$
Observe that

$a_0 \leq a_1 \leq b_1 \leq b_0$
$b_1 - a_1 = \frac{b_0 - a_0}{2^1}$ and
1) $\exists e_1 \in E$ such that $e_1 \geq a_1$
2) $b_1 \geq e \ \forall e \in E$

Repeating we obtain

$a_0 \leq a_1 \leq \ldots \leq a_n \leq b_n \leq \ldots \leq b_1 \leq b_0$
$b_n - a_n = \frac{b_0 - a_0}{2^n}$ and
1) $\exists e_n \in E$ such that $e_n \geq a_1$
2) $b_n \geq e \ \forall e \in E$

The a_ns form an increasing sequence bounded above, so $a_n \to \alpha$ by the fundamental theorem
$b_n = a_n + (b_n + a_n) = a_n + 2^{-n}(b_0 - a_0) \to \alpha + 0 = \alpha$ so $a_n \leq e_n \leq b_n$ with all the $e_n \in E$, so $e_n \to \alpha$ and $b_n \geq e \ \forall e \in E$ so α is the least upper bound of E □

We observe that this result is actually equivalent to the fundamental axiom.

Theorem 4.6. Theorem 4.5 implies the fundamental axiom.

Proof. On the assumption that every bounded non-empty set has a least upper bound, suppose $a_1 \leq a_2 \leq \ldots \leq a_n \leq A_0 \forall n$
Set $A = \{a_1, a_2, \ldots\}$ is non-empty and bounded, so it has a supremum which we call α. Given $\epsilon > 0$ we know $\alpha - \epsilon$ is not an upper bound, so $\exists N$ such that $a_N > \alpha - \epsilon$. Now if $n \geq N, \alpha \geq a_n \geq a_N \geq \alpha - \epsilon$, so $|a_n - \alpha| < \epsilon$ and $a_n \to \alpha$ □

Of course we have the notion of a greatest lower bound or infimum.

Exercise 4.7. Define the greatest lower bound in the manner of Definition 4.2, prove its uniqueness in the manner of Lemma 4.3 and state and prove a result corresponding to Lemma 4.4.

If E has an infimum β we write $\inf_{e \in E} e = \inf E = \beta$. One way of dealing with the infimum is to use the following observation.

Lemma 4.8. Let E be a non-empty set in $\mathbb{R}$ and write $-E = \{-e : e \in E\}$. Then E has an infimum if and only if $-E$ has a supremum. If E has an infimum $\inf E = -\sup(-E)$.

Proof. If $-E$ has a supremum α
i) $\alpha \geq -e \ \forall e \in E$
ii) If $\beta \geq -e \ \forall e \in E$ then $\beta \geq \alpha$
So
i)' $-a \leq e \ \forall e \in E$
ii)' If $-\beta \leq e \ \forall e \in E$ then $-\beta \leq -\alpha$
So $-\alpha$ is the infimum of E □

Exercise 4.9. Use Lemma 4.8 and Theorem 4.5 to show that any non-empty set in $\mathbb{R}$ with a lower bound has a greatest lower bound.

The notion of a supremum will play an important rôle in our proofs of Theorem 5.12 and Theorem 9.2.
The following result is also equivalent to the fundamental axiom (that is, we can deduce it from the fundamental axiom and conversely, if we take it as an axiom, rather than a theorem, then we can deduce the fundamental axiom as a theorem).

Theorem 4.10 (Bolzano-Weierstrass)**.** *If $x_n \in \mathbb{R}$ and there exists a K such that $|x_n| \leq K$ for all n, then we can find $n(1) < n(2) < \ldots$ and $x \in \mathbb{R}$ such that $x_{n(j)} \to x$ as $j \to \infty$.*

Proof. Set $a_0 = -K, b_0 = K$. $\{m : x_m \in [a_0, b_0]\}$ is infinite. Set $c_0 = \frac{a_0+b_0}{2}$
If $\{m : x_m \in [a_0, c_0]\}$ is infinite we set $a_1 = a_0, b_1 = c_0$, otherwise set $a_1 = c_0, b_1 = b_0$
Observe that

$$a_0 \leq a_1 \leq b_1 \leq b_0$$
$$b_1 - a_1 = 2^{-1}(b_0 - a_0)$$
$$\{m : x_m \in [a_1, b_1]\} \text{ is infinite}$$

Repeat. a_n is increasing and bounded above so tends to a limit, call it α
$b_n = a_n + (b_n - a_n) = a_n + (b_0 - a_0)2^{-n} \to \alpha$
Since $[a_n, b_n]$ contains an x_m for infinitely many m I can choose $m(1) < m(2) < m(3) < \ldots$ such that $x_{m(j)} \in [a_j, b_j]$. Thus $a_j \leq x_{m(j)} \leq b_j$ and as $a_j, b_j \to \alpha$, $x_{m(j)} \to \alpha$ □

The Bolzano-Weierstrass theorem says that every bounded sequence of reals has a convergent subsequence. Notice that we say nothing about uniqueness; if $x_n = (-1)^n$ then $x_{2n} \to 1$ but $x_{2n+1} \to -1$ as $n \to \infty$.
We proved the theorem of Bolzano-Weierstrass by 'lion hunting' but your supervisor may well show you another method. We shall use the Bolzano-Weierstrass theorem to prove that every continuous function on a closed bounded interval is bounded and attains its bounds (Theorem 5.12). The Bolzano-Weierstrass theorem will be much used in the next analysis course because it generalises to many dimensions.
Note also that the Bolzano-Weierstrass theorem implies the fundamental axiom

Proof. Suppose $a_1 \leq a_2 \leq \ldots$ and $a_n \leq A$. Then $\{a_j\}$ is a bounded sequence so by Bolzano-Weierstrass it has a convergent subset i.e. $\exists n(1) < n(2) < \ldots$ such that $a_{n(k)} \to \alpha$
Since $a_{n(k)} \leq a_j\ \forall j \geq n(k)$ we have $\alpha \geq a_n(k)\ \forall k$ so $\alpha \geq a_j\ \forall j$
Now given $\epsilon > 0\ \exists n(k)$ such that $|\alpha - a_{n(k)}| < \epsilon$ so $\alpha \geq a_j \geq a_{n(k)}\ \forall j \geq n(k)$, so $|\alpha - a_j| < \epsilon\ \forall j \geq n(k)$ and thus $a_j \to \alpha$ □

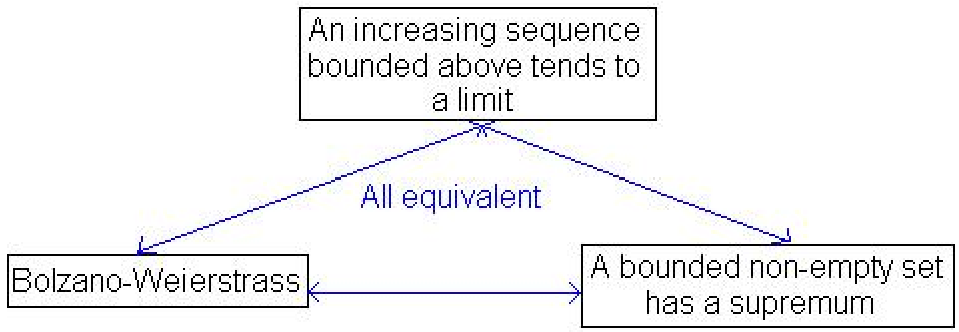

Figure 1: Equivalence of fundamental axioms

5 Continuity

We make the following definition.

Definition 5.1. A function $f : \mathbb{R} \to \mathbb{R}$ is continuous at x if given $\epsilon > 0$ we can find a $\delta(\epsilon, x) > 0$ [read 'a delta depending on epsilon and x'] such that

$$|f(x) - f(y)| < \epsilon$$

for all y with $|x - y| < \delta(\epsilon, x)$.
If f is continuous at each point $x \in \mathbb{R}$ we say that f is a continuous function on $\mathbb{R}$.

I shall do my best to make this seem a reasonable definition but it is important to realise that I am really stating a rule of the game (like a knights move in chess or the definition of offside in football). If you wish to play the game you must accept the rules. Results about continuous functions must be derived from the definition and not stated as 'obvious from the notion of continuity'.
In practice we use a slightly more general definition.

Definition 5.2. Let E be a subset of $\mathbb{R}$. A function $f : E \to \mathbb{R}$ is continuous at $x \in E$ if given $\epsilon > 0$ we can find a $\delta(\epsilon, x) > 0$ [read 'a delta depending on epsilon and x'] such that

$$|f(x) - f(y)| < \epsilon$$

for all $y \in E$ with $|x - y| < \delta(\epsilon, x)$.
If f is continuous at each point $x \in E$, we say that f is a continuous function on E.

However, it will do no harm and may be positively helpful if, whilst you are getting used to the idea of continuity, you concentrate on the case $E = \mathbb{R}$.

Lemma 5.3. Suppose that E is a subset of $\mathbb{R}$, that $x \in E$, and that f and g are functions from E to $\mathbb{R}$.
(i) If $f(x) = c$ for all $x \in E$, then f is continuous on E.
(ii) If f and g are continuous at x, then so is $f + g$.
(iii) Let us define $f \times g : E \to \mathbb{R}$ by $f \times g(t) = f(t)g(t)$ for all $t \in E$. Then if f and g are continuous at x, so is $f \times g$.
(iv) Suppose that $f(t) \neq 0$ for all $t \in E$. If f is continuous at x so is $1/f$.

Proof. i) Trivial
ii) As f, g are continous at x
Given $\epsilon > 0$ $\exists \delta_1(\epsilon) > 0$ such that $|f(x) - f(y)| < \epsilon$ for $|x - y| < \delta_1(\epsilon)$
Given $\epsilon > 0$ $\exists \delta_2(\epsilon) > 0$ such that $|g(x) - g(y)| < \epsilon$ for $|x - y| < \delta_2(\epsilon)$
Now set $\delta_3(\epsilon) = \min\{\delta_1(\frac{\epsilon}{2}), \delta_2(\frac{\epsilon}{2})\}$
Then if $|x - y| < \delta_3(\epsilon)$

$$\begin{aligned} |(f+g)(x) - (f+g)(y)| &= |f(x) + g(x) - f(y) + g(y)| \\ &= |(f(x) - f(y)) + (g(x) - g(y))| \\ &\leq |f(x) - f(y)| + |g(x) - g(y)| \\ &\leq \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon \end{aligned}$$

iii) Is done similarly
iv) f is continuous at x
Thus given $\epsilon > 0$ $\exists \delta(\epsilon) > 0$ such that $|f(x) - f(y)| < \epsilon$ for $|x - y| < \delta(\epsilon)$
Set $\delta_1(\epsilon) = \min\{\delta(\frac{|f(x)|}{2}), \delta(\frac{\epsilon|f(x)|^2}{2})\}$
Then if $|x - y| < \delta_1(\epsilon)$

$$\begin{aligned} |\frac{1}{f(x)} - \frac{1}{f(y)}| &\leq \frac{|f(y) - f(x)|}{|f(x)||f(y)|} \\ &\leq \frac{2|f(x) - f(y)|}{|f(x)|^2|} \\ &< \epsilon \end{aligned}$$

Using $|f(y) - f(x)| < \frac{|f(x)|}{2}$ so $|f(y)| > \frac{|f(x)|}{2}$ □

Lemma 5.4. Let U and V be subsets of $\mathbb{R}$. Suppose $f : U \to \mathbb{R}$ is such that $f(t) \in V$ for all $t \in U$. If f is continuous at $x \in U$ and $g : V \to \mathbb{R}$ is continuous at $f(x)$, then the composition $g \circ f$ is continuous at x.

Proof. Given $\epsilon > 0$ $\exists$ $\delta_1(\epsilon) > 0$ such that $|f(x) - f(y)| < \epsilon$ for $|x - y| < \delta_1(\epsilon)$
Given $\epsilon > 0$ $\exists$ $\delta_2(\epsilon) > 0$ such that $|g(f(x)) - g(u)| < \epsilon$ for $|u - f(x)| < \delta_2(\epsilon)$
Set $\delta_3(\epsilon) = \delta_1(\delta_2(\epsilon))$
Then if $|x - y| < \delta_3(\epsilon)$ we have $|f(x) - f(y)| < \delta_2(\epsilon)$ and so $|g(f(x)) - g(f(y))| < \epsilon$ □

By repeated use of parts (ii) and (iii) of Lemma 5.3 it is easy to show that polynomials $P(t) = \sum_{r=0}^{n} a_r t^r$ are continuous. The details are spelled out in the next exercise.

Exercise 5.5. Prove the following results.
(i) Suppose that E is a subset of $\mathbb{R}$ and that $f : E \to \mathbb{R}$ is continuous at $x \in E$. If $x \in E' \subset E$ then the restriction $f|_{E'}$ of f to E' is also continuous at x.
(ii) If $J : \mathbb{R} \to \mathbb{R}$ is defined by $J(x) = x$ for all $x \in \mathbb{R}$, then J is continuous on $\mathbb{R}$.
(iii) Every polynomial P is continuous on $\mathbb{R}$.
(iv) Suppose that P and Q are polynomials and that Q is never zero on some subset E of $\mathbb{R}$. Then the rational function P/Q is continuous on E (or, more precisely, the restriction of P/Q to E is continuous.)

The following result is little more than an observation but will be very useful.

Lemma 5.6. Suppose that E is a subset of $\mathbb{R}$, that $x \in E$, and that f is continuous at x. If $x_n \in E$ for all n and $x_n \to x$ as $n \to \infty$, then $f(x_n) \to f(x)$ as $n \to \infty$.

Proof. Given $\epsilon > 0\ \exists\ \delta(\epsilon) > 0$ such that $|f(x) - f(y)| < \epsilon$ for $|x - y| < \delta(\epsilon)$
Given $\epsilon > 0\ \exists\ N(\epsilon)$ such that $|x_n - x| < \epsilon$ for $n \geq N(\epsilon)$
If we set $M(\epsilon) = N(\delta(\epsilon))$ then $n \geq M(\epsilon) \Rightarrow |x_n - x| < \delta(\epsilon) \Rightarrow |f(x_n) - f(x)| < \epsilon$ □

So far in this section we have only done algebra but the next result depends on the fundamental axiom. It is one of the key results of analysis and although my recommendation runs contrary to a century of enlightened pedagogy I can see no objections to students learning the proof as a model. Notice that the theorem resolves the problem posed by Example 1.1 (i).

Theorem 5.7 (The intermediate value theorem)**.** If $f : [a, b] \to \mathbb{R}$ is continuous and $f(a) \geq 0 \geq f(b)$ then there exists a $c \in [a, b]$ such that $f(c) = 0$

Proof. This relies on the fact that we are working in $\mathbb{R}$
Set $a_0 = a, b_0 = b, c_0 = \frac{a_0+b_0}{2}$
If $f(c_0) \geq 0$ set $a_1 = c_0, b_1 = b_0$, otherwise set $a_1 = a_0, b_1 = c_0$. Observe that

$$\begin{aligned}
&a_0 \leq a_1 \leq b_1 \leq b_0 \\
&b_1 - a_1 = 2^{-1}(b_0 - a_0) \\
&f(a_1) \geq 0 \geq f(b_1)
\end{aligned}$$

Repeat, obtaining

$$\begin{aligned}
&a_0 \leq a_1 \leq \ldots \leq a_n \leq b_n \leq \ldots \leq b_1 \leq b_0 \\
&b_n - a_n = 2^{-n}(b_0 - a_0) \\
&f(a_n) \geq 0 \geq f(b_n)
\end{aligned}$$

a_n is an increasing sequence bounded above, so $a_n \to \alpha$ with $\alpha \leq b_n \forall n$ as $a_n \leq b_n$
$b_n = a_n + (b_n - a_n) = a_n + 2^{-n}(b_0 - a_0) \to \alpha + 0 = \alpha$, and as $b_n \geq a_n\ \forall n, \alpha \geq a$
Now $a_n \to \alpha$ so $f(a_n) \to f(\alpha)$ and similarly $f(b_n) \to f(\alpha)$. But $f(a_n) \geq 0$ so $f(\alpha) \geq 0$ and $f(b_n) \leq 0$ so $f(\alpha) \leq 0$. Thus $f(\alpha) = 0$ □

The next three exercises are applications of the intermediate value theorem.

Exercise 5.8. Show that any real polynomial of odd degree has at least one root. Is the result true for polynomials of even degree? Give a proof or counterexample.

Proof. Let $p(x) = x^3+ax^2+bx+c = x^3(\frac{a}{x}+\frac{b}{x^2}+\frac{c}{x^3}) \geq x^3(1-\frac{|a|}{x}-\frac{|b|}{x^2}-\frac{|c|}{x^3}) > \frac{x^5}{4}$
If we set $x \geq 4(2+|a|+|b|+|c|), p(x) \geq 0$ and for x sufficiently large and negative $p(x)$ is negative, so by continuity has a root. ☐

Exercise 5.9. Suppose that $g : [0,1] \to [0,1]$ is a continuous function. By considering $f(x) = g(x) - x$, or otherwise, show that there exists a $c \in [0,1]$ with $g(c) = c$. (Thus every continuous map of $[0,1]$ into itself has a fixed point.) Give an example of a bijective (but, necessarily, non-continuous) function $h : [0,1] \to [0,1]$ such that $h(x)\ neqx$ for all $x \in [0,1]$.
[Hint: First find a function $H : [0,1] \setminus \{0,1,1/2\} \to [0,1] \setminus \{0,1,1/2\}$ such that $H(x) \neq x$.]

Proof. Let $f(x) = g(x) - x$
$f(0) \geq 0$ (if $f(0) = 0, 0 = c$ and done)
$f(1) \leq 0$ (if $f(1) = 0, 0 = c$ and done) Therefore using the Intermediate Value Theorem $\exists c$ such that $f(c) = 0$ and thus $g(c) = c$ ☐

Exercise 5.10. Every mid-summer day at six o'clock in the morning, the youngest monk from the monastery of Damt starts to climb the narrow path up Mount Dipmes. At six in the evening he reaches the small temple at the peak where he spends the night in meditation. At six o'clock in the morning on the following day he starts downwards, arriving back at the monastery at six in the evening. Of course, he does not always walk at the same speed. Show that, none the less, there will be some time of day when he will be at the same place on the path on both his upward and downward journeys.

We proved the intermediate value theorem (Theorem 5.7) by lion hunting. We prove the next two theorems by using the Bolzano-Weierstrass Theorem (Theorem 4.10). Again the results are very important and I can see no objection to learning the proofs as a model.

Theorem 5.11. If $f : [a,b] \to \mathbb{R}$ is continuous then we can find an M such that $|f(x)| \leq M$ for all $x \in [a,b]$.

Proof. Note that this needs to be on a closed interval, on $(0,1)$ $f : x \mapsto \frac{1}{x}$ is unbounded
Suppose f is unbounded
Then we can find $x_n \in [a,b]$ such that $|f(x_n)| \geq n$
The x_n lie in a bounded interval so they have a convergent subsequence $x_{n(j)} \to \alpha$ say. Since $x_{n(j)} \leq b, \alpha \leq b$ and since $x_{n(j)} \geq a, \alpha \geq a$ so $\alpha \in [a,b]$
Now f is continuous at α so $\exists\ \delta > 0$ such that $|f(\alpha) - f(x)| < 1\ \forall |x-\alpha| < \delta$
$x \in [a,b]$ so $|f(x)| \leq |f(\alpha)+1|\ \forall |x-\alpha| < \delta, x \in [a,b]$

But $x_{n(j)} \to \alpha$. So $\exists J$ such that if $j \geq J$ $|x_{n(j)} - \alpha| < \delta$
so $n(j) \leq |f(x_{n(j)}| \leq |f(\alpha)| + 1$ $\forall j \geq J$ #
Thus f must be bounded $\square$

In other words a continuous function on a closed bounded interval is bounded. We improve this result in the next theorem.

Theorem 5.12. If $f : [a, b] \to \mathbb{R}$ is continuous then we can find $x_1, x_2 \in [a, b]$ such that

$$f(x_1) \leq f(x) \leq f(x_2)$$

for all $x \in [a, b]$.

Proof. Start by observing that $\{f(x) : x \in [a, b]\}$ is a non-empty bounded set, so it has a supremum M. What we need to show is that $\exists \alpha \in [a, b]$ such that $f(\alpha) = M$
Since M is the supremum of $\{f(x) : x \in [a, b]\}$ we can certainly find $x_n \in [a, b]$ such that $f(x_n) \geq M - \frac{1}{n}$ (if not, then $(M - \frac{1}{n})$ would be an upper bound, but M is the least upper bound). Automatically $M \geq f(x_n)$
By Bolzano-Weierstrass there is a convergenet subsequence $x_{n(j)} \to \alpha$. As before $\alpha \in [a, b]$ and $f(x_{n(j)}) \to f(\alpha)$. But $M \geq f(x_{n(j)}) \geq M - \frac{1}{n(j)}$, so $f(x_{n(j)}) \to M$ so $f(\alpha) = M$ $\square$

In other words a continuous function on a closed bounded interval is bounded and attains its bounds.

6 Differentiation

In this section it will be useful to have another type of limit.

Definition 6.1. Let E be a subset of $\mathbb{R}$, f be some function from E to $\mathbb{R}$, and x some point of E. If $l \in \mathbb{R}$ we say that $f(y) \to l$ as $y \to x$ [or, if we wish to emphasise the restriction to E that $f(y) \to l$ as $y \to x$ through values $y \in E$] if, given $\epsilon > 0$, we can find a $\delta(\epsilon) > 0$ [read 'a delta depending on epsilon'] such that

$$|f(y) - l| < \epsilon$$

for all $y \in E$ with $0 < |x - y| < \delta(\epsilon)$.

As before there is no real loss if the reader initially takes $E = \mathbb{R}$.
The following two exercises are easy but useful.

Exercise 6.2. Let E be a subset of $\mathbb{R}$. Show that a function $f : E \to \mathbb{R}$ is continuous at $x \in E$ if and only if $f(y) \to f(x)$ as $y \to x$.

Exercise 6.3. Let E be a subset of $\mathbb{R}$, f, g be some functions from E to $\mathbb{R}$, and x some point of E.
(i) The limit is unique. That is, if $f(y) \to l$ and $f(y) \to k$ as $y \to x$ then $l = k$.

(ii) If $x \in E' \subseteq E$ and $f(y) \to l$ as $y \to x$ through values $y \in E$, then $f(y) \to l$ as $y \to x$ through values $y \in E'$.
(iii) If $f(t) = c$ for all $t \in E$ then $f(y) \to c$ as $y \to x$.
(iv) If $f(y) \to l$ and $g(y) \to k$ as $y \to x$ then $f(y) + g(y) \to l + k$.
(v) If $f(y) \to l$ and $g(y) \to k$ as $y \to x$ then $f(y)g(y) \to lk$.
(vi) If $f(y) \to l$ as $y \to x$, $f(t) \neq 0$ for each $t \in E$ and $l \neq 0$ then $f(t)^{-1} \to l^{-1}$.
(vii) If $f(t) \leq L$ for each $t \in E$ and $f(y) \to l$ as $y \to x$ then $l \leq L$.

Proof. iii) $f(t) = c \;\forall t$. Given $\epsilon > 0$ set $\delta = 1$. $|f(y) - c| = |c - c| = |0| = 0 < \epsilon$
iv) Since $f(y) \to l, g(y) \to k$ as $y \to x$
Given $\epsilon > 0 \;\exists\; \delta_1(\epsilon) > 0$ such that $|f(y) - l| < \epsilon$ for $|y - x| < \delta_1(\epsilon)$
Given $\epsilon > 0 \;\exists\; \delta_2(\epsilon) > 0$ such that $|g(y) - k| < \epsilon$ for $|y - x| < \delta_2(\epsilon)$
Thus given $\epsilon > 0$ if we set $\delta_3(\epsilon) = \min\{(\delta_1(1), \delta_1(\frac{\epsilon}{2(1+|k|)}), \delta_2(\frac{\epsilon}{2(1+|l|)})\}$ then

$$\begin{aligned}
|f(y)g(y) - lk| &= |f(y)g(y) - f(y)k + f(y)k - lk| \\
&\leq |f(y)g(y) - f(y)k| + |f(y)k - lk| \\
&= |f(y)||g(y) - k| + |f(y) - l||k| \\
&\leq (|l| + 1)|g(y) - k| + |f(y) - l||k|
\end{aligned}$$

□

We can now define the derivative.

Definition 6.4. Let E be a subset of $\mathbb{R}$. A function $f : E \to \mathbb{R}$ is differentiable at $x \in E$ with derivative $f'(x)$ if

$$\frac{f(y) - f(x)}{y - x} \to f'(x)$$

as $y \to x$.
If f is differentiable at each point $x \in E$, we say that f is a differentiable function on E.

As usual, no harm will be done if you replace E by $\mathbb{R}$.
You can think of differentiation as

$$f(y) = \underbrace{f(x) + f'(x)(y - x)}_{\text{linear}} + \text{error term}$$

to be useful the error has to $\to 0$ faster than the linear term.
Note: We can also use this in 3-D with grad

$$\begin{aligned}
\delta f &= \bigtriangledown f.\delta x \\
\delta f &= \frac{\delta f}{\delta x_1}\delta x_1 + \frac{\delta f}{\delta x_2}\delta x_2 + \frac{\delta f}{\delta x_3}\delta x_3 \\
\Rightarrow f(\underline{y}) &= f(\underline{x}) + \bigtriangledown f_0.(\underline{y} - \underline{x}) + \text{error term}
\end{aligned}$$

and as before error $\to 0$ faster than the linear term
Here are some easy consequences of the definition.

Exercise 6.5. Let E be a subset of $\mathbb{R}$, f some function from E to $\mathbb{R}$, and x some point of E. Show that if f is differentiable at x then f is continuous at x.

Proof. Given $\epsilon > 0 \ \exists \ \delta(\epsilon) > 0$ such that $|\frac{f(x)-f(y)}{y-x} - f'(x)| < \epsilon$ for $|y-x| < \delta(\epsilon)$, so

$$\begin{aligned}|f(y) - f(x) - f'(x)(y-x)| &< \epsilon|y-x| \\ \text{thus } |f(y) - f(x)| &< \epsilon|y-x| + |f'(x)(y-x)| \\ &= \epsilon|y-x| + |f'(x)||y-x|\end{aligned}$$

All of the terms in the last expression are small except $|f'(x)|$ which is fixed. So if we set $\delta_1(\epsilon) = \min\{\delta(1), \frac{\epsilon}{2}, \frac{\epsilon}{2(|f'(x)+1|)}\}$ for $\epsilon < 1$. Then if $0 < |x-y| < \delta_1(\epsilon)$

$$\begin{aligned}|f(x) - f(y)| &< 1.|y-x| + |f'(x)||y-x| \\ &< \frac{\epsilon}{2} + \frac{\epsilon}{2} \\ &= \epsilon\end{aligned}$$

$\square$

Exercise 6.6. Let E be a subset of $\mathbb{R}$, f, g be some functions from E to $\mathbb{R}$, and x some point of E. Prove the following results.
(i) If $f(t) = c$ for all $t \in E$ then f is differentiable at x with $f'(x) = 0$.
(ii) If f and g are differentiable at x then so is their sum $f+g$ and

$$(f+g)'(x) = f'(x) + g'(x)$$

(iii) If f and g are differentiable at x then so is their product $f \times g$ and

$$(f \times g)'(x) = f'(x)g(x) + f(x)g'(x)$$

(iv) If f is differentiable at x and $f(t) \neq 0$ for all $t \in E$ then $1/f$ is differentiable at x and

$$(1/f)'(x) = -f'(x)/f(x)^2$$

(v) If $f(t) = \sum_{r=0}^{n} a_r t^r$ on E then f is differentiable at x and

$$f'(x) = \sum_{r=1}^{n} r a_r x^{r-1}$$

Proof. iii)

$$\begin{aligned}\frac{f(y)g(y) - f(x)g(x)}{y-x} &= \frac{f(y)(g(y)-g(x))}{y-x} + \frac{g(x)(f(y)-f(x))}{y-x} \\ &= f(x)g'(x) + g(x)f'(x)\end{aligned}$$

(Note: We have used that, as we pass to the limit, $f(y) \to f(x)$)
iv)

$$
\begin{aligned}
\frac{\frac{1}{f(y)} - \frac{1}{f(x)}}{y - x} &= \frac{1}{f(y)f(x)} \cdot \frac{f(x) - f(y)}{y - x} \\
&= -\frac{1}{f(y)f(x)} \cdot \frac{f(y) - f(x)}{y - x} \\
&\to -\frac{f'(x)}{(f(x))^2}
\end{aligned}
$$

□

The next result is slightly harder to prove than it looks. (we split the proof into two halves depending on whether $f'(x) \neq 0$ or $f'(x) = 0$).

Lemma 6.7. Let U and V be subsets of $\mathbb{R}$. Suppose $f : U \to \mathbb{R}$ is such that $f(t) \in V$ for all $t \in U$. If f is differentiable at $x \in U$ and $g : V \to \mathbb{R}$ is differentiable at $f(x)$, then the composition $g \circ f$ is differentiable at x with

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

Proof. If $f'(x) \neq 0$ then $f(x+h) - f(x) \neq 0$ for h small and thus

$$
\begin{aligned}
\frac{g(f(x+h)) - g(f(x))}{h} &= \frac{g(f(x+h)) - g(f(x))}{f(x+h) - f(x)} \cdot \frac{f(x+h) - f(x)}{h} \\
&= g'(f(x)).f'(x)
\end{aligned}
$$

(since f is continuous at $x \leq 0$, $f(x+h) \to f(x)$)
We have lots of different ways of dealing with $f'(x) = 0$. Here is one
Since g is differentiable at $f(x)$

$$\frac{g(f(x) + h) - g(f(x))}{k} \to g'(f(x))$$

so if k is sufficiently small, say for $0 < |k| < \mu$

$$\left| \frac{g(f(x) + k) - g(f(x))}{k} - g'(f(x)) \right| < 1$$

So $|g(f(x) + k) - g(f(x))| \leq |k|(1 + |g'(f(x))|)$
Now f is continuous at x so $\exists \delta(\mu) > 0$ such that $|f(x+h) - f(x)| < \mu$ for $|h| < \delta(\mu)$
Take $k = f(x+h) - f(x)$. We have

$$|g(f(x+h)) - g(f(x))| \leq |f(x+h) - f(x)|(1 + |g'(f(x))|)$$

so

$$\left| \frac{g(f(x+h)) - g(f(x))}{h} \right| \leq \underbrace{\left| \frac{f(x+h) - f(x)}{h} \right|}_{\to 0} (1 + |g'(f(x))|)$$

Thus $g \circ f$ is differentiable with derivative 0 □

7 The mean value theorem

We have almost finished our project of showing that the horrid situation revealed by Example 1.1 can not occur for the reals.
Our first step is to prove Rolle's theorem.

Theorem 7.1 (Rolle's theorem)**.** If $g : [a, b] \to \mathbb{R}$ is a continuous function with g differentiable on (a, b) and $g(a) = g(b)$, then we can find a $c \in (a, b)$ such that $g'(c) = 0$.

Proof. A continuous function on a closed bounded interval is bounded and attains its bounds (by theorem 5.12)
In other words, we can find $c_1, c_2 \in [a, b]$ such that $g(c_1) \geq g(x) \geq g(c_2)$ $\forall x \in [a, b]$
If both are end points (i.e. $c_1, c_2 \in \{a, b\}$) then $g(c_1) = g(c_2)$ so $g(x) = g(c_1)$ $\forall x \in [a, b]$ so g is constant and we could take $c = \frac{a+b}{2}$. Hence we may assume that at least one of c_1, c_2 is not an end point.
WLOG assume that c_1 is not an end point (otherwise consider $-g$
Let $c_1 = c$, we have $(c - \delta, c + \delta) \subset [a, b]$ for some $\delta > 0$
Observe that $f(c + h) - g(c) \leq 0$ (as c is a maximum point)
So, if $h > 0$ $\frac{g(c+h)-g(c)}{h} \leq 0$ so allowing $h \to 0$ through positive values $g'(c) \leq 0$
Similarly if $h < 0$ $\frac{g(c+h)-g(c)}{h} \geq 0$ so allowing $h \to 0$ through negative values $g'(c) \geq 0$.
Thus $g'(c) = 0$ □

A simple tilt gives the famous mean value theorem.

Theorem 7.2 (The mean value theorem)**.** If $f : [a, b] \to \mathbb{R}$ is a continuous function with f differentiable on (a, b), then we can find a $c \in (a, b)$ such that

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

Proof. Set $g(t) = (f(t) - f(a)) - \frac{t-a}{t-b}(f(b) - f(a))$
$g(t) = Af(t)+B$ where A, B are such that $g(a) = g(b) = 0$. As g is differentiable on (a, b) and continuous on $[a, b]$, by Rolle's Theorem $\exists c \in (a, b)$ such that $g'(c) = 0$ i.e. $0 = f'(c) - \frac{f(b)-f(a)}{b-a}$ □

We now have the results so long desired.

Lemma 7.3. If $f : [a, b] \to \mathbb{R}$ is a continuous function with f differentiable on (a, b), then the following results hold.
(i) If $f'(t) > 0$ for all $t \in (a, b)$ then f is strictly increasing on $[a, b]$. (That is, $f(y) > f(x)$ whenever $b \geq y > x \geq a$.)
(ii) If $f'(t) \geq 0$ for all $t \in (a, b)$ then f is increasing on $[a, b]$. (That is, $f(y) \geq f(x)$ whenever $b \geq y > x \geq a$.)
(iii) **[The constant value theorem]** If $f'(t) = 0$ for all $t \in (a, b)$ then f is constant on $[a, b]$. (That is, $f(y) = f(x)$ whenever $b \geq y > x \geq a$.)

Proof. i) $f(y) - f(x) = (y - x)f'(c)$ with c between x and y by the MVT, and thus as $f'(c)$ is positive and $y - x$ is $f(y) > f(x)$
ii) Same as above, except that $f'(c)$ might be 0, so can get $\geq$ instead of $>$
iii) $f(y) - f(x) = (y - x)f'(c) = 0$
The partical converses are easy:
If f is differentiable, and increasing, then $\frac{f(x+h)-f(x)}{h} \geq 0 \; \forall h \neq 0$ so

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

□

Notice that since we deduce Lemma 7.3 from the mean value theorem we can not use it in the proof of Rolle's theorem.
The mean value theorem has many important consequences, some of which we look at in the remainder of the section.
We start by looking at inverse functions.

Lemma 7.4. (i) Suppose $f : [a, b] \to \mathbb{R}$ is continuous. Then f is injective if and only if it is strictly increasing (that is $f(t) > f(s)$ whenever $a \leq s < t \leq b$) or strictly decreasing.
(ii) Suppose $f : [a, b] \to \mathbb{R}$ is continuous and strictly increasing. Let $f(a) = c$ and $f(b) = d$. Then the map $f : [a, b] \to [c, d]$ is bijective and f^{-1} is continuous on $[c, d]$.

Proof. If f is strictly increasing, $y \neq x$ then either $y > x$ or $x > y$. If $y > x, f(y) > f(x)$ so $f(y) \neq f(x)$, same for $x > y$
Conversely, suppose f is not strictly increasing. Then either

> A) We can find $x_1 < x_2 < x_3$ such that $f(x_1) \leq f(x_2)$ and $f(x_3) \leq f(x_2)$
>
> B) we can find $x_1 < x_2 < x_3$ such that $f(x_1) \geq f(x_2)$ and $f(x_3) \geq f(x_2)$

WLOG assume case A (otherwise look at $-f$)
If $f(x_1) = f(x_2)$ or $f(x_2) = f(x_3)$ then we have a contradiction (f would not be injective). So suppose all are strict inequalities
Choose c with $f(x_2) > c > \max\{f(x_1), f(x_2)\}$. By the Intermediate Value Theorem we can find α_1, α_2 such that $x_1 < \alpha_1 < x_3, x_2 < \alpha_2 < x_3$ and $f(\alpha_1) = x = f(\alpha_2)$. Thus f is not injective
ii) The injectivity follows from i)
Surjectivity follows from the Intermediate Value Theorem: If $x \leq \gamma \leq d$ then since $f(a) = c, f(b) = d \; \exists t$ with $a \leq t \leq b$ such that $f(t) = \gamma$
Thus f is bijective and f^{-1} is defined as a function, $f^{-1} : [c, d] \mapsto [a, b]$
Claim that f^{-1} is continuous
Let $y \in [c, d]$. To simplify matters, suppose $y \neq c, y \neq d$
If $\delta > 0$ consider the interval $[\theta_1, \theta_2] = [a, b] \cap [f^{-1}(y - \frac{\delta}{2}), f^{-1}(y + \frac{\delta}{2})]$
$\theta_1 < f^{-1}(y) < \theta_2$

$f(\theta_1) < y < f(\theta_2)$ so choose $\epsilon > 0$ such that $f(\theta_1) < y - \epsilon < y < y + \epsilon < f(\theta_2)$
If $|y - z| < \epsilon$, $f(\theta_1) < z < f(\theta_2)$ so $\theta_1 < f^{-1}(z) < \theta_2$ and so $|f^{-1}(z) - f^{-1}(y)| < \frac{\delta}{2} < \delta$
The proof if $y = c$ or $y = d$ is similar, but one-sided $\square$

Lemma 7.5 (Inverse rule)**.** Suppose $f : [a, b] \to \mathbb{R}$ is differentiable on $[a, b]$ and $f'(x) > 0$ for all $x \in [a, b]$. Let $f(a) = c$ and $f(b) = d$. Then the map $f : [a, b] \to [c, d]$ is bijective and f^{-1} is differentiable on $[c, d]$ with

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

Proof.
$$\frac{f^{-1}(y+h) - f^{-1}(y)}{h} = \frac{f^{-1}(y+h) - f^{-1}(y)}{f(f^{-1}(y+h)) - f(f^{-1}(y))}$$
But f is continuous, so $f^{-1}(y + h) - f^{-1}(y) \to 0$ and thus

$$\frac{f^{-1}(y+h) - f^{-1}(y)}{h} \to \frac{1}{f'(f^{-1}(y))}$$

$\square$

Note: Suppose we know that f and f^{-1} are differentiable, then $t = f(f^{-1}(t)) = f \circ f^{-1}(t)$. By the chain rule $1 = f'(f^{-1}(t)).(f^{-1})'(t)$ so $(f^{-1})'(t) = \frac{1}{f'(f^{-1}(t))}$
This does not constitute a proof!
In the opinion of the author the true meaning of the inverse rule and the chain rule only becomes clear when we consider higher dimensions in the next analysis course.
We now prove a form of Taylor's theorem.

Theorem 7.6 (nth mean value theorem)**.** Suppose that $b > a$ and $f : [a, b] \to \mathbb{R}$ is $n + 1$ times differentiable. Then

$$f(b) - \sum_{j=0}^{n} \frac{f^{(j)}(a)}{j!}(b-a)^j = \frac{f^{(n+1)}(c')}{(n+1)!}(b-a)^{n+1}$$

for some c' with $a < c' < b$.

Proof. This is complicated, so we try to simplify a little
Step 1: By translation $x \mapsto x - a$ we may suppose $a = 0$, and our statement becomes
$$f(b) - \sum_{n=0}^{n} \frac{f^{(j)}(0)}{j!} b^j = \frac{f^{(n+1)}(c')}{(n+1)!} b^{n+1}$$
with $0 < c' < b$
Step 2: Set $g(t) = f(t) - \sum_{r=0}^{n} \frac{f^{(r)}(0)}{r!} t^r$
Then $g(0) = g'(0) = \cdots = g^{(n)}(0) = 0$ and our theorem becomes

If g is $(n+1)$ times differentiable, and $g(0) = g'(0) = \cdots = g^{(n)}(0) = 0$ then

$$g(b) = \frac{g^{(n+1)}(c)}{(n+1)!} b^{n+1} \qquad (g \text{ theorem})$$

We now need to prove the g theorem
Let $G(x) = g(x) - \frac{x^{n+1}}{b^{n+1}} g(b)$. Then $G(0) = G'(0) = \cdots = G^{(n)}(0) = 0$ and $G(b) = 0$, so by Rolle, $\exists 0 < b_1 < b$ such that $G'(b_1) = 0$
And because $G'(0) = 0$ $\exists 0 < b_2 < b_1$ such that $G''(b_2) = 0$
And because $G''(0) = 0$ $\exists 0 < b_3 < b_2$ such that $G'''(b_3) = 0$ and so on
Until we get to $0 < b_n < b_{n-1}$ with $G^{(n)}(b_n) = 0$. But $G^{(n)}(0) = 0$ so $\exists 0 < c < b_n$ such that $G^{(n+1)}(c) = 0$

$$G^{(n+1)}(x) = g^{(n+1)}(c) - \frac{g(b)(n+1)!}{b^n} \Rightarrow g(b) = \frac{g^{(n+1)}(c) b^{n+1}}{(n+1)!}$$

□

This gives us a global and a local Taylor's theorem.

Theorem 7.7 (Global Taylor theorem)**.** Suppose that $b > a$ and $f : [a, b] \to \mathbb{R}$ is $n+1$ times differentiable. If $x, t \in [a, b]$

$$f(t) = \sum_{j=0}^{n} \frac{f^{(j)}(x)}{j!}(t-x)^j + \frac{f^{(n+1)}(x + \theta(t-x))}{(n+1)!}(t-x)^{n+1}$$

for some $\theta \in (0, 1)$.

Theorem 7.8 (Local Taylor theorem)**.** Suppose that $\delta > 0$ and $f : (x - \delta, x + \delta) \to \mathbb{R}$ is n times differentiable on $(x - \delta, x + \delta)$ and $f^{(n)}$ is continuous at x. Then

$$f(t) = \sum_{j=0}^{n} \frac{f^{(j)}(x)}{j!}(t-x)^j + \epsilon(t)(t-x)^n$$

where $\epsilon(t) \to 0$ as $t \to x$.

Proof. Move to the origin, $f(t) = f(0) + f'(0)t + \cdots + \frac{f^{(n-1)}(0)}{(n-1)!} t^{n-1} + \frac{f^{(n)}(c) t^n}{n!}$
If $f^{(n)}$ is continuous at O

$$\begin{aligned} f^{(n)}(c) - f^{(n)}(0) \to &\text{as } t \to 0 \\ &\text{so } c \to 0 \end{aligned}$$

Write $\epsilon(t) = (f^{(n)}(c) - f^{(n)}(0)) \frac{1}{n!}$

$$f(t) = f(0) + f'(0)t + \cdots + \frac{f^{(n-1)}(0) t^{n-1}}{(n-1)!} + \frac{f^{(n)}(0) t^n}{n!} \epsilon(t) t^n$$

□

Notice that the local Taylor theorem always gives us some information but that the global one is useless unless we can find a useful bound on the $n+1$th derivative.
To reinforce this warning we consider a famous example of Cauchy.

Exercise 7.9. Consider the function $F : \mathbb{R} \to \mathbb{R}$ defined by

$$\begin{aligned} &F(0) = 0 \\ &F(x) = \exp(-1/x^2) \qquad \text{otherwise.} \end{aligned}$$

(i) Prove by induction, using the standard rules of differentiation, that F is infinitely differentiable at all points $x \neq 0$ and that, at these points,

$$F^{(n)}(x) = P_n(1/x)\exp(-1/x^2)$$

where P_n is a polynomial which need not be found explicitly.
(ii) Explain why $x^{-1}P_n(1/x)\exp(-1/x^2) \to 0$ as $x \to 0$.
(iii) Show by induction, using the definition of differentiation, that F is infinitely differentiable at 0 with $F^{(n)}(0) = 0$ for all n. [Be careful to get this part of the argument right.]
(iv) Show that

$$F(x) = \sum_{j=0}^{\infty} \frac{F^{(j)}(0)}{j!} x^j$$

if and only if $x = 0$. (The reader may prefer to say that 'The Taylor expansion of F is only valid at 0'.)
(v) Why does part (iv) not contradict the local Taylor theorem (Theorem 7.8)?

Since examiners are fonder of the global Taylor theorem than it deserves I shall go through the following example.

Example 7.10. Assuming the standard properties of the exponential function show that

$$\exp x = \sum_{j=0}^{\infty} \frac{x^j}{j!}$$

for all x.

Proof. If $f(x) = e^x$ then $f^{(r)}x = e^x$ and using the formula

$$f(x) = f(0) + \frac{f'(0)}{1!}x + \cdots + \frac{f^{(n)}(0)}{n!}x^n + \frac{f^{(n+1)}(c)}{(n+1)!}x^{n+1} \quad |c| < |x|$$

which becomes

$$e^x = 1 + \frac{x}{1!} + \frac{x^2}{2!} + \cdots + \frac{x^n}{n!} + e^c \frac{x^{n+1}}{(n+1)!}$$

Look at $R_n = e^c \frac{x^{n+1}}{(n+1)!}$, the error term
$|R_n| \leq e^{|x|} \frac{|x|^{n+1}}{(n+1)!} \to 0$ (because if $n \geq 2|x| + 1$, $\frac{|x|}{n} \leq \frac{1}{2}$
So $e^x = 1 + x + \frac{x^2}{2!} + \cdots$ □

Please note that in a pure mathematics question many (or even most) of the marks in a question of this type will depend on estimating the remainder term. [In methods questions you may simply be asked to 'find the Taylor's series' without being asked to prove convergence.]

8 Complex variable

The field $\mathbb{C}$ of complex numbers resembles the field $\mathbb{R}$ of real numbers in many ways but not in all.

Lemma 8.1. We can not define an order on $\mathbb{C}$ which will behave in the same way as $>$ for $\mathbb{R}$.

However there is sufficient similarity for us to define limits, continuity and differentiability. (We have already seen some of this in Definition 3.1 and Exercise 3.2.)

Definition 8.2. Let E be a subset of $\mathbb{C}$, f be some function from E to $\mathbb{C}$, and z some point of E. If $l \in \mathbb{C}$ we say that $f(w) \to l$ as $w \to z$ [or, if we wish to emphasise the restriction to E that $f(w) \to l$ as $w \to z$ through values $w \in E$] if, given $\epsilon > 0$, we can find a $\delta(\epsilon) > 0$ [read 'a delta depending on epsilon'] such that

$$|f(w) - l| < \epsilon$$

for all $w \in E$ with $0 < |w - z| < \delta(\epsilon)$.

As usual there is no real loss if the reader initially takes $E = \mathbb{C}$.

Definition 8.3. Let E be a subset of $\mathbb{C}$. We say that a function $f : E \to \mathbb{C}$ is continuous at $z \in E$ if and only if $f(w) \to f(z)$ as $w \to z$.

Exercise 8.4. Let E be a subset of $\mathbb{C}$, f, g be some functions from E to $\mathbb{C}$, and z some point of E.
(i) The limit is unique. That is, if $f(w) \to l$ and $f(w) \to k$ as $w \to z$ then $l = k$.
(ii) If $z \in E' \subseteq E$ and $f(w) \to l$ as $w \to z$ through values $w \in E$, then $f(w) \to l$ as $w \to x$ through values $w \in E'$.
(iii) If $f(u) = c$ for all $u \in E$ then $f(w) \to c$ as $w \to z$.
(iv) If $f(w) \to l$ and $g(w) \to k$ as $w \to z$ then $f(w) + g(w) \to l + k$.
(v) If $f(w) \to l$ and $g(w) \to k$ as $w \to z$ then $f(w)g(w) \to lk$.
(vi) If $f(w) \to l$ as $w \to z$, $f(u) \neq 0$ for each $u \in E$ and $l \neq 0$ then $f(w)^{-1} \to l^{-1}$.

Exercise 8.5. Suppose that E is a subset of $\mathbb{C}$, that $z \in E$, and that f and g are functions from E to $\mathbb{C}$.
(i) If $f(u) = c$ for all $u \in E$, then f is continuous on E.
(ii) If f and g are continuous at z, then so is $f + g$.
(iii) Let us define $f \times g : E \to \mathbb{C}$ by $f \times g(u) = f(u)g(u)$ for all $u \in E$. Then if f and g are continuous at z, so is $f \times g$.

(iv) Suppose that $f(u) \neq 0$ for all $u \in E$. If f is continuous at z so is $1/f$.
(v) If $z \in E' \subset E$ and f is continuous at z then the restriction $f|_{E'}$ of f to E' is also continuous at z.
(vi) If $J : \mathbb{C} \to \mathbb{C}$ is defined by $J(z) = z$ for all $z \in \mathbb{C}$, then J is continuous on $\mathbb{C}$.
(vii) Every polynomial P is continuous on $\mathbb{C}$.
(viii) Suppose that P and Q are polynomials and that Q is never zero on some subset E of $\mathbb{C}$. Then the rational function P/Q is continuous on E (or, more precisely, the restriction of P/Q to E is continuous.)

Exercise 8.6. Let U and V be subsets of $\mathbb{C}$. Suppose $f : U \to \mathbb{C}$ is such that $f(z) \in V$ for all $z \in U$. If f is continuous at $w \in U$ and $g : V \to \mathbb{C}$ is continuous at $f(w)$, then the composition $g \circ f$ is continuous at w.

Definition 8.7. Let E be a subset of $\mathbb{C}$. A function $f : E \to \mathbb{C}$ is differentiable at $z \in E$ with derivative $f'(z)$ if

$$\frac{f(w) - f(z)}{w - z} \to f'(z)$$

as $w \to z$.
If f is differentiable at each point $z \in E$ we say that f is a differentiable function on E.

Exercise 8.8. Let E be a subset of $\mathbb{C}$, f some function from E to $\mathbb{C}$, and z some point of E. Show that if f is differentiable at z then f is continuous at z.

Exercise 8.9. Let E be a subset of $\mathbb{C}$, f, g be some functions from E to $\mathbb{C}$, and z some point of E. Prove the following results.
(i) If $f(u) = c$ for all $u \in E$ then f is differentiable at z with $f'(z) = 0$.
(ii) If f and g are differentiable at z then so is their sum $f + g$ and

$$(f + g)'(z) = f'(z) + g'(z)$$

(iii) If f and g are differentiable at z then so is their product $f \times g$ and

$$(f \times g)'(z) = f'(z)g(z) + f(z)g'(z)$$

(iv) If f is differentiable at z and $f(u) \neq 0$ for all $u \in E$ then $1/f$ is differentiable at z and

$$(1/f)'(z) = -f'(z)/f(z)^2$$

(v) If $f(u) = \sum_{r=0}^{n} a_r u^r$ on E then f is differentiable at z and

$$f'(z) = \sum_{r=1}^{n} r a_r z^{r-1}$$

Proof. v) If $f_n(z) = z^n$ then f is differentiable with $f_n' = n f_{n-1}$
So prove by induction, already proved for 0, 1
Suppose true for n. $f_{n+1} = f_n.f_1$ so by the multiplicative rule
$f_{n+1}' = f_n'.f_1 + f_n.f_1' = n f_{n-1}.f_1 + f_n.1 = (n+1) f_n$ □

Exercise 8.10 (Chain rule)**.** Let U and V be subsets of $\mathbb{C}$. Suppose $f : U \to \mathbb{C}$ is such that $f(z) \in V$ for all $t \in U$. If f is differentiable at $w \in U$ and $g : V \to \mathbb{R}$ is differentiable at $f(w)$, then the composition $g \circ f$ is differentiable at w with

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

In spite of these similarities the subject of complex differentiable functions is very different from that of real differentiable functions. It turns out that 'well behaved' complex functions need not be differentiable.

Example 8.11. Consider the map $\Gamma : \mathbb{C} \to \mathbb{C}$ given by $\Gamma(z) = \overline{z}$. The function Γ is nowhere differentiable.

Proof.

$$\frac{\Gamma(z+h) - \Gamma(z)}{h} = \frac{\overline{(z+h)} - \overline{z}}{h} = \frac{\overline{h}}{h}$$

Claim that $\frac{\overline{h}}{h}$ does not tend to a limit as $h \to 0$
First, let $h \in \mathbb{R}, h \to 0$. $\frac{\overline{h}}{h} = \frac{h}{h} = 1 \to 1$
Now let $h = ik, j \in \mathbb{R}, k \to 0$. $\frac{\overline{h}}{h} = \frac{-ik}{ik} = -1 \to -1$ # $\square$

We can view this another way (this is not a proof)

$$\begin{aligned}\frac{\overline{(z+h)} - \overline{z}}{h} &= \frac{\overline{(z+re^{i\theta})} - \overline{z}}{re^{i\theta}} \\ &= \frac{re^{-i\theta}}{re^{i\theta}} \\ &= e^{-2i\theta}\end{aligned}$$

so the derivative depends on the direction taken
$f(z+iw) = f(z) + Re^{i\theta} + \text{error}$
Also, given that $\overline{z}$ changes sense it is not surprising that it is not differentiable
Because complex differentiability is so much more restrictive than real differentiability we can prove stronger theorems about complex differentiable functions. For example it can be shown that such functions can be written locally as power series[2] (contrast the situation in the real case revealed by Example 7.9). To learn more go to the course P3 on complex methods.

9 Power series

In this section we work in $\mathbb{C}$ unless otherwise stated. We start with a very useful observation.

Lemma 9.1. If $\sum_{n=0}^{\infty} a_n z_0^n$ converges and $|z| < |z_0|$ then $\sum_{n=0}^{\infty} a_n z^n$ converges.

[2] The syllabus says that this fact is part of this course. In this one instance I advise you to ignore the syllabus.

Proof. Since $\sum a_n z_0^n$ converges, $a_n z^n \to 0$. Since the series $a_n z_0^n \to 0$ it is bounded i.e. $\exists K$ such that $|a_n z_0^n| \leq K \ \forall n \geq 1$
(this follows: $a_n z_0^n \to 0 \Rightarrow \exists N : |a_n z_0^n| \leq 1 \ \forall n \geq N$, take $K = \max\{1, |a_0|, |a_1 z|, |a_2 z_0^2|, \ldots, |a_N z_0^N|\}$
Thus $|a_n z^n| = |a_n z_0^n||\frac{z}{z_0}|^n \leq K|\frac{z}{z_0}|^n$
Since $|\frac{z}{z_0}| \leq 1, \sum |\frac{z}{z_0}|^n$ converges over the reals (as it is a geometric progression) (note, we say nothing about $\mathbb{C}$), so by the comparison test $\sum |a_n z^n|$ converges, and since an absolutely convergent series converges, $\sum a_n z^n$ converges $\square$

This gives us the following basic theorem on power series.

Theorem 9.2. Suppose that $a_n \in \mathbb{C}$. Then either $\sum_{n=0}^{\infty} a_n z^n$ converges for all $z \in \mathbb{C}$, or there exists a real number R with $R \geq 0$ such that
(i) $\sum_{n=0}^{\infty} a_n z^n$ converges if $|z| < R$,
(ii) $\sum_{n=0}^{\infty} a_n z^n$ diverges if $|z| > R$.

Proof. If $\sum a_n z^n$ converges $\forall z$, we set $R = \infty$, nothing to prove.
If not, consider $E = \{|z| : \sum a_n z_n \text{diverges}\}$
E is a subset of $\mathbb{R}$, $E \neq \phi$, E bounded below by 0. $R =\in E$ is thus defined. We must now prove that it has the required properties.
If $|z| < R$ then $\sum a_n z^n$ converges by definition
If $|z| > R$ then by definition we can find $|w| \in \mathbb{C}$ such that $|z| > |w|$ and $\sum a_n w^n$ diverges. Now, by the previous lemma, if $\sum a_n z^n$ converged then $\sum a_n w^n$ would converge, but it does not, so $\sum a_n z^n$ diverges $\square$

We call R the radius of convergence of $\sum_{n=0}^{\infty} a_n z^n$. If $\sum_{n=0}^{\infty} a_n z^n$ converges for all z we write $R = \infty$.
Warning: Although the ratio test is a useful first test for trying to find R it is not always successful and it cannot be used in the definition. Look at
$1 + 2z + z^2 + 2z^3 + z^4 + 2z^5 + \cdots$
$|\frac{a_{n+1}}{a_n}|$ is alternately $2|z|$ and $\frac{|z|}{2}$
However, $1+z^2+z^4+\cdots$ converges, and $2z+2z^3+2z^5+\cdots = 2z(1+z^2+z^4+\cdots)$ converges too (if $|z| < 1$, so $1 + 2z + z^2 + 2z^3 + z^4 + 2z^5 + \cdots$ converges for $|z| < 1$
$|z|^n \to \infty$ if $|z| > 1$, so $1 + 2z + z^2 + 2z^3 + z^4 + \cdots$ must diverge if $|z| > 1$, so the radius of convergence if 1
We never say anything about behaviour on the radius of convergence.
The following useful strengthening is left to the reader as an exercise.

Exercise 9.3. Suppose that $\sum_{n=0}^{\infty} a_n z^n$ has radius of convergence R. Then the sequence $|a_n z^n|$ is unbounded if $|z| > R$ and $\sum_{n=0}^{\infty} a_n z^n$ converges absolutely if $|z| < R$.

Note that we say nothing about what happens on the circle of convergence.

Example 9.4. (i) $\sum_{n=1}^{\infty} n^{-2} z^n$ has radius of convergence 1 and converges for all z with $|z| = 1$.
(ii) $\sum_{n=1}^{\infty} z^n$ has radius of convergence 1 and diverges for all z with $|z| = 1$.

Proof. i) If $|z| = 1, |\frac{z^n}{n^2}| = \frac{1}{n^2}$ so $\sum n^{-2} z^n$ is absolutely convergent
Using the ratio test: $|\frac{z^{n+1}}{(n+1)^2} / \frac{z^n}{n^2}| = \frac{|z|}{(1+\frac{1}{n})^2} \to |z|$ so the radius of convergence is 1, and we have convergence on the circle of convergence
ii) $\sum_0^N z^n = \frac{1-z^{N+1}}{1-z}$ for $z \neq 1$ which tends to a limit iff $|z| < 1$
and $\sum_0^N 1^n = N \to \infty$ so $\sum z^n$ has radius of convergence 1 and diverges at every point on the circle of convergence □

A more complicated example is given in question 16 on the 3rd example sheet. It is a remarkable fact that we can operate with power series in the same way as polynomials (within the radius of convergence). In particular we shall show that we can differentiate term by term.

Theorem 9.5. If $\sum_{n=0}^{\infty} a_n z^n$ has radius of convergence R and we write $f(z) = \sum_{n=0}^{\infty} a_n z^n$ then f is differentiable at all points z with $|z| < R$ and

$$f'(z) = \sum_{n=1}^{\infty} n a_n z^{n-1}$$

The proof is starred in the syllabus. We use three simple observations.

Lemma 9.6. If $\sum_{n=0}^{\infty} a_n z^n$ has radius of convergence R then given any $\epsilon > 0$ we can find a $K(\epsilon)$ such that $\sum_{n=0}^{\infty} |a_n z^n| < K(\epsilon)$ for all $|z| \leq R - \epsilon$.

Lemma 9.7. If $\sum_{n=0}^{\infty} a_n z^n$ has radius of convergence R then so do $\sum_{n=1}^{\infty} n a_n z^{n-1}$ and $\sum_{n=2}^{\infty} n(n-1) a_n z^{n-2}$.

Proof. Suppose $\sum a_n z^n$ has radius of convergence R. If $|z| < R$ we can find w_0 with $|z| < |w_0| < R$
Since $|w_0| < R, \sum a_n w_0^n$ converges, so $\exists M$ such that $|a_n w_0^n| \leq M \; \forall n$
Now $|n a_n z^{n-1}| = |a_n w_0^n| |\frac{z}{w_0}|^{n-1} \frac{n}{|w_0|} \leq \frac{Mn}{w_0} |\frac{z}{w_0}|^{n-1} = u_n$
Claim that $\sum u_n$ converges $\because \frac{u_{n+1}}{u_n} = \frac{z}{w_0}(1 + \frac{1}{n}) \to \frac{z}{w_0} < 1$ so ratio test works.
Since $\sum u_n$ converges the comparison test tells us that $\sum |n a_n z^{n-1}|$ converges, so $\sum n a_n z^{n-1}$ converges.
Thus $\sum n a_n z^{n-1}$ has radius of convergence at least R, since $|n a_n z^{n-1}| \geq n|a_n||z|^{n-1} = \frac{1}{|z|} n |a_n z^n|$
$\sum n(n-1) a_n z^{n-2}$ has radius of convergence R by "differentiating" again. □

Lemma 9.8. (i) $\binom{n}{r} \leq n(n-1)\binom{n-2}{r-2}$ for all $2 \leq r \leq n$.
(ii) $|(z+h)^n - z^n - nhz^{n-1}| \leq n(n-1)(|z| + |h|)^{n-2} |h|^2$ for all $z, h \in \mathbb{C}$.

Proof. i)

$$n(n-1)\binom{n-2}{r-2} = \frac{n(n-1)(n-2)}{(r-2)!(n-r)!} = \frac{n!}{(r-2)!(n-r)!} \geq \frac{n!}{r!(n-r)!} = \binom{n}{r}$$

ii)

$$
\begin{aligned}
|(z+h)^n - nz^{n-1}h - z^n| &= \sum_{r=2}^{n} |\binom{n}{r} z^{n-r} h^2| \\
&\leq n(n-1) \sum_{r=2}^{n} \binom{n-2}{r-2} |z^{n-r} h^r| \\
&= n(n-1)h^2 \sum_{r=0}^{n} \binom{n-2}{r} |z^{n-2-r}||h^r| \\
&= n(n-1)|h|^2(|z|+|h|)^{n-2}
\end{aligned}
$$

Thus if $|z|+|h|<R$

$$
\begin{aligned}
|\frac{\sum a_n(z+h)^n - \sum a_n z^n}{h} - \sum n a_n z^{n-1}| &= |\frac{1}{h}(\sum a_n((z+h)^n - z^n - nz^{n-1}h))| \\
&\leq \frac{1}{|h|} \sum |a_n||(z+h)^n - z^n - nz^{n-1}h| \\
&\leq \frac{1}{|h|} \sum |a_n||h|^2(|z|+|h|)^{n-2} n(n-1) \\
&= |h| \sum n(n-1)|a_n|(|z|+|h|)^{n-2}
\end{aligned}
$$

and given that we are within the radius of convergence this sum converges to a bounded value, so $\sum n(n-1)|a_n||w|^n \leq K$ for $|w| < R-\epsilon$ □

In this course we shall mainly work on the real line. Restricting to the real line we obtain the following result.

Theorem 9.9. (i) If $a_j \in \mathbb{R}$ there exists a unique R (the radius of convergence) with $0 \leq R \leq \infty$ such that $\sum_{n=0}^{\infty} a_n x^n$ converges for all real x with $|x| < R$ and diverges for all real x with $x > R$.
(ii) If $\sum_{n=1}^{\infty} a_n x^n$ has radius of convergence R and we write $f(x) = \sum_{n=0}^{\infty} a_n x^n$ then f is differentiable at all points x with $|x| < R$ and

$$f'(x) = \sum_{n=1}^{\infty} n a_n x^{n-1}$$

10 The standard functions

In school you learned all about the functions exp, log, sin and cos and about the behaviour of x^α. Nothing that you learned was wrong (we hope) but you might be hard pressed to prove all the facts you know in a coherent manner,
To get round this problem, we start from scratch making new definitions and assuming nothing about these various functions. One of your tasks is to make sure that the lecturer does not slip in some unproved fact. On the other hand

you must allow your lecturer to choose definitions which allow an easy development of the subject rather than those that follow some 'historic', 'intuitive' or 'pedagogically appropriate' path[3].
Let us start with the exponential function. Throughout this section we shall restrict ourselves to the real line.

Lemma 10.1. The sum $\sum_{n=0}^{\infty} \frac{x^n}{n!}$ has infinite radius of convergence.

Proof.
$$|\frac{x^{n+1}}{(n+1)!}|/|\frac{x^n}{n!}| = \frac{|x|}{n+1} \to 0$$
so by the ratio test $\sum_{n=0}^{\infty} \frac{x^n}{n!}$ converges $\forall x$ so the radius of convergence is ∞ □

We can thus define a function $e : \mathbb{R} \to \mathbb{R}$ by
$$e(x) = \sum_{n=0}^{\infty} \frac{x^n}{n!}$$
(We use $e(x)$ rather than $\exp(x)$ to help us avoid making unjustified assumptions.)

Theorem 10.2. (i) The function $e : \mathbb{R} \to \mathbb{R}$ is everywhere differentiable with $e'(x) = e(x)$.
(ii) $e(x + y) = e(x)e(y)$ for all $x, y \in \mathbb{R}$.
(iii) $e(x) > 0$ for all $x \in \mathbb{R}$.
(iv) e is a strictly increasing function.
(v) $e(x) \to \infty$ as $x \to \infty$, $e(x) \to 0$ as $x \to -\infty$.
(vi) $e : \mathbb{R} \to (0, \infty)$ is a bijection.

Proof. i) Since a power series can be differentiated term by term within its radius of convergence $e(x)$ is differentiable
$$\therefore e'(x) = \sum_{1}^{\infty} n.\frac{x^{n-1}}{n!} = \sum_{1}^{\infty} \frac{x^{n-1}}{(n-1)!} = \sum_{0}^{\infty} \frac{x^n}{n!} = e(x)$$
ii) Let $f(x) = e(a - x)e(x)$ with a a constant
$$\begin{aligned} f'(x) &= -e'(a-x)e(x) + e(a-x)e'(x) \\ &= -e(a-x)e(x) + e(a-x)e(x) \\ &= 0 \end{aligned}$$
so by the mean value theorem, f is constant, $\therefore f(x) = f(0)$ i.e. $e(a - x)e(x) = e(a)e(0)$
Looking at power series $e(0) = 1 + \frac{0}{1!} + \frac{0}{2!} + \cdots = 1$ so $e(a - x)e(x) = e(a)\ \forall a, x$
Set $a = x + y$ to get $e(y)e(x) = e(x + y)$

[3]If you want to see a treatment along these lines see the excellent text of Burn[3].

iii) $e(x) = \sum_0^\infty \frac{x^r}{r!} \geq 1 \ \forall x \geq 0$ but $(e^{-x})e^x = e(-x+x) = e(0) = 1$
Therefore $e^{-x} > 0$ for $x > 0$, and thus $e^x > 0 \ \forall x$
iv) One of the corollaries of the mean value theorem (theorem 7.2) tells us that since $e'(x) = e(x) > 0$ e is strictly increasing
v) $e(x) = \sum_0^\infty \frac{x^r}{r!} \geq 1 + \frac{x}{1!} = 1 + x \to \infty$ so $e(x) \to \infty$ as $x \to \infty$
$e(-x) = \frac{1}{e(x)}$ (by $e(-x)e(x) = e(0) = 1$), and $\frac{1}{e(x)} \to 0$ as $x \to \infty$, so $e(x) \to 0$ as $x \to -\infty$
vi) Since e is strictly increasing it is injective ($x > y \Rightarrow e(x) > e(y)$)
If $0 < c$ since $e(x) \to 0$ as $x \to -\infty$ we can find x_1 with $e(x_1) < c$
Since $e(x) \to \infty$ as $x \to \infty$ we can find x_3 with $e(x_3) > c$
So by the Intermediate Value Theorem $\exists x_2$ such that $e(x_2) = c$
Thus e is surjective ☐

It is worth stating some of our results in the language of group theory.

Lemma 10.3. The mapping e is an isomorphism of the group $(\mathbb{R}, +)$ and the group $((0,\infty), \times)$.

Proof. $e(x+y) = e(x)e(y)$ ☐

Since $e : \mathbb{R} \to (0,\infty)$ is a bijection we can consider the inverse function $l : (0,\infty) \to \mathbb{R}$.

Theorem 10.4. (i) $l : (0,\infty) \to \mathbb{R}$ is a bijection. We have $l(e(x)) = x$ for all $x \in \mathbb{R}$ and $e(l(t)) = t$ for all $t \in (0,\infty)$.
(ii) The function $l : (0,\infty) \to \mathbb{R}$ is everywhere differentiable with $l'(t) = 1/t$.
(iii) $l(uv) = l(u) + l(v)$ for all $u, v \in (0,\infty)$.

Proof. i) Just repeats what it means to be an inverse function
ii) We know that the inverse of a bijective function f is differentiable and $(f^{-1})'(x) = \frac{1}{f'(f^{-1}(x))}$
so $l'(x) = \frac{1}{e'(l(x))} = \frac{1}{e(l(x))} = \frac{1}{x}$
iii) Just isomorphism
e goes from $(\mathbb{R}, +)$ to $((0,\infty), \times)$
l goes from $((0,\infty), \times)$ to $(\mathbb{R}, +)$
(Or, $e(l(u) + l(v)) = e(l(u))e(l(v)) = uv = e(l(uv))$, so as e is bijective, $l(u) + l(v) = l(uv)$) ☐

We now write $e(x) = \exp(x)$, $l(t) = \log t$ (the use of ln is unnecessary). If $\alpha \in \mathbb{R}$ and $x > 0$ we write $r_\alpha(x) = \exp(\alpha \log x)$.

Theorem 10.5. Suppose x, $y > 0$ and $\alpha, \beta \in \mathbb{R}$. Then
(i) $r_\alpha(xy) = r_\alpha(x) r_\alpha(y)$,
(ii) $r_{\alpha+\beta}(x) = r_\alpha(x) r_\beta(x)$,
(iii) $r_\alpha(r_\beta(x)) = r_{\alpha\beta}(x)$
(iv) $r_1(x) = x$.

Proof. i)

$$\begin{aligned}
r_\alpha(xy) &= \exp(\alpha \log(xy)) \\
&= \exp(\alpha \log x + \alpha \log y) \\
&= \exp(\alpha \log x) \exp(\alpha log y) \\
&= r_\alpha(x) r_\alpha(y)
\end{aligned}$$

ii)

$$\begin{aligned}
r_{\alpha+\beta}(x) &= \exp((\alpha + \beta) \log x) \\
&= \exp(\alpha \log x + \beta \log x) \\
&= \exp(\alpha \log x) \exp(\beta \log x) \\
&= r_\alpha(x) r_\beta(x)
\end{aligned}$$

iii)

$$\begin{aligned}
r_\alpha(r_\beta(x)) &= \exp(\alpha \log(\exp(\beta(\log x)))) \\
&= \exp(\alpha\beta \log x) \\
&= r_{\alpha\beta}(x)
\end{aligned}$$

iv)

$$r_1(x) = \exp(\log(x)) = x$$

□

Exercise 10.6. Use the results of Theorem 10.5 to show that if n is a strictly positive integer and $x > 0$ then

$$r_n(x) = \underbrace{xx \ldots x}_{n}$$

Proof. $r_n(x) = r_{\underbrace{1+1+1+\cdots+1}_{n \text{ times}}}(x) = \underbrace{r_1(x) r_1(x) r_1(x) \ldots r_1(x)}_{n \text{ times}} = \underbrace{x.x.x.\ldots.x}_{n \text{ times}}$ □

Thus, if we write $x^\alpha = r_\alpha(x)$, our new notation is *consistent* with our old school notation when α is rational but gives, in addition, a valid definition when α is irrational.

Lemma 10.7. (i) If α is real, $r_\alpha : (0, \infty) \to (0, \infty)$ is everywhere differentiable and $r'_\alpha(x) = \alpha r_{\alpha-1}(x)$.
(ii) If $x > 0$ and we define $f_x(\alpha) = x^\alpha$ then $f_x : \mathbb{R} \to (0, \infty)$ is everywhere differentiable and $f'_x(\alpha) = \log x f_x(\alpha)$.

Proof. i)

$$\begin{aligned}
\frac{d}{dx} r_\alpha(x) &= \frac{d}{dx}(\exp(\alpha \log x)) \\
&= \frac{d}{dx}(e(\alpha l(x))) \\
&= \alpha l'(x) e'(\alpha l(x)) \text{ (chain rule)} \\
&= \frac{\alpha}{x} e(\alpha l(x)) \\
&= \frac{\alpha}{x} r_\alpha(x) \\
&= \alpha_{r-1}(x) r_\alpha(x) \\
&= \alpha r_{\alpha-1}(x)
\end{aligned}$$

ii) $f_x(\alpha) = \exp(x \log \alpha)$
$\therefore (f_x)'(\alpha) = \log \alpha \exp'(x \log \alpha) = \log \alpha f_x(\alpha)$ □

Note: This may appear circular, as we use x^n in the definition of e^x, and then e^x for x^n. However, the logic actually goes as follows:

i) Define x^n for $n \in \mathbb{N}$, find properties

ii) Define $r_\alpha(x)$ for $\alpha \in \mathbb{R}, x > 0$, show that $r_n(x) = x^n$ for $n \in \mathbb{N}$

iii) Decide to write $r_a lpha(x) = x^\alpha$

Finally we look at the trigonometric functions.

Lemma 10.8. The sums $\sum_{n=0}^{\infty} \frac{(-1)^n x^{2n+1}}{(2n+1)!}$ and $\sum_{n=0}^{\infty} \frac{(-1)^n x^{2n}}{(2n)!}$ have infinite radius of convergence.

Proof. Observe that

$$\left|\frac{(-1)^{n+1} x^{2n+3}}{(2n+3)!}\right| / \left|\frac{(-1)^n x^{2n+1}}{(2n+1)!}\right| = \frac{x^2}{(2n+2)(2n+2)} \to 0 \text{ as } n \to 0 \ \forall x$$

$\therefore \sum_{n=0}^{\infty} \frac{(-1)^n x^{2n+1}}{(2n+1)!}$ converges $\forall x$ i.e. infinite radius of convergence
Similarly for the other half □

We can thus define functions s, $c : \mathbb{R} \to \mathbb{R}$ by

$$s(x) = \sum_{n=0}^{\infty} \frac{(-1)^n x^{2n+1}}{(2n+1)!} \text{ and } c(x) = \sum_{n=0}^{\infty} \frac{(-1)^n x^{2n}}{(2n)!}$$

Lemma 10.9. (i) The functions $s, c : \mathbb{R} \to \mathbb{R}$ are everywhere differentiable with $s'(x) = c(x)$ and $c'(x) = -s(x)$.
(ii) $s(x+y) = s(x)c(y) + c(x)s(y)$, $c(x+y) = c(x)c(y) - s(x)s(y)$ and $s(x)^2 + c(x)^2 = 1$ for all $x, y \in \mathbb{R}$.
(iii) $s(-x) = -s(x)$, $c(-x) = c(x)$ for all $x \in \mathbb{R}$.

Proof. i) We can differentiate term by term within the radius of convergence so, for example

$$c'(x) = \sum_{n=1}^{\infty} \frac{d}{dx}(\frac{x^{2n}(-1)^n}{(2n)!}) = \sum_{1}^{\infty}(-1)\frac{x^{2n-1}(-1)^{n-1}}{(2n-1)!} = -\sum_{0}^{\infty}\frac{(-1)^n x^{2n+1}}{(2n+1)!} = -s(x)$$

ii) Fix α, β, let $a = \alpha + \beta$, Consider

$$\begin{aligned} f(x) &= s(x)c(a-x) + s(a-x)c(x) \\ f'(x) &= s'(x)c(a-x) - s(x)c'(a-x) - s'(a-x)c(x) + s(a-x)c'(x) \\ &= c(x)c(a-x) - s(x)s(a-x) - c(a-x)c(x) - s(a-x)s(x) \\ &= 0 \end{aligned}$$

sp f is constant, so $f(x) = f(a)\ \forall x$, i.e. $s(x)c(a-x) + s(a-x)c(x) = s(a)c(a-a) + s(a-a)c(a) = s(a)c(0) + s(0)c(a)$
From the power series $c(0) = 1, s(0) = 0$, so not set $x = \beta$ to get

$$s(\beta)c(\alpha) + s(\alpha)c(\beta) = s(\alpha + \beta)$$

$c(x+y)$ is proved similarly

$$\begin{aligned} \frac{d}{dx}(c^2(x) + s^2(x)) &= 2c'(x)c(x) + 2s'(x)s(x) \\ &= -2s(x)c(x) + 2s(x)c(x) \\ &= 0 \end{aligned}$$

$\therefore c^2(x) + s^2(x)$ is constant, and thus $c^2(x) + s^2(x) = c^2(0) + s^2(0) = 1 + 0 = 1$ and thus $c^2(x) + s^2(x) = 1\ \forall x$
iii) $s(-x) = \sim_{n=0}^{\infty} \frac{(-1)^n(-x)^{2n+1}}{(2n+1)!} = \sum_{n=0}^{\infty} -\frac{(-1)^n x^{2n+1}}{(2n+1)!} = -s(x)$
Similarly $c(-x) = c(x)$ □

Exercise 10.10. Write down what you consider to be the chief properties of sinh and cosh. (They should convey enough information to draw a reasonable graph of the two functions.)
(i) Obtain those properties by starting with the definitions

$$\sinh x = \sum_{n=0}^{\infty} \frac{x^{2n+1}}{(2n+1)!} \text{ and } \cosh x = \sum_{n=0}^{\infty} \frac{x^{2n}}{(2n)!}$$

and proceeding along the lines of Lemma 10.9.
(ii) Obtain those properties by starting with the definitions

$$\sinh x = \frac{\exp x - \exp(-x)}{2} \text{ and } \cosh x = \frac{\exp x + \exp(-x)}{2}$$

We have not yet proved one of the most remarkable properties of the sine and cosine functions, their periodicity.

Theorem 10.11. Let s and c be as in Lemma 10.9.
(i) If $c(a) = 0$ and $c(b) = 0$ then $s(b - a) = 0$.
(ii) We have $s(x) > 0$ for all $0 < x \leq 1$.
(iii) There exists a unique $\omega \in [0, 2]$ such that $c(\omega) = 0$.
(iv) $s(\omega) = 1$.
(v) $s(x + 4\omega) = s(x)$, $c(x + 4\omega) = c(\omega)$.
(vi) The function c is strictly decreasing from 1 to -1 as x runs from 0 to 2ω,

Proof. i)

$$\begin{aligned} s(b-a) &= s(b)c(-a) + s(-a)c(b) \\ &= s(b)c(a) - s(a)c(b) \\ &= 0 - 0 = 0 \end{aligned}$$

ii) $s(x)$ by $s(x) = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \cdots$
and if $0 < x \leq 1$, $\frac{x^{2n+1}}{(2n+1)!} < \frac{x^{2n-1}}{(2n-1)!}$ (proof: $\frac{x^{2n+1}}{(2n+1)!} / \frac{x^{2n-1}}{(2n-1)!} = \frac{x^2}{(2n+1)(2n)} < 1$)
The error in computing an alternating [convergent] series is less than the first term omitted, so $|s(x) - x| \leq \frac{x^3}{3!}$ so $s(x) \geq x - \frac{x^3}{3!} > 0$ for $x \in (0, 1]$
So by i) if a, b are distinct zeroes of $c(x)$ we have $|a - b| > 1$ (†)
iii) If $0 \leq x \leq 1$ then the terms $\frac{x^{2n}}{(2n)!}$ are strictly decreasing, tending to 0
So by alternating series the error in

$$c(x) = \sum (-1)^n \frac{x^{2n}}{(2n)!}$$

is equal to the first term neglected, so $|c(x) - 1| \leq \frac{x^2}{2!}$ so $c(x) \geq 1 - \frac{x^2}{2!} \geq \frac{1}{2}$ $\forall x \in [0, 1]$
Look at $c(2)$: $c(2) = 1 - \frac{2^2}{2!} + \frac{2^4}{4!} - \frac{2^6}{6!} + \cdots$, not decreasing, alternating, but $-\frac{2^2}{2!} + \frac{2^4}{4!} - \frac{2^6}{6!}$ is
So $|c(2) - (1 - \frac{2^2}{2!})| \leq \frac{2^4}{4!}$, sp $|c(2) + 1| \leq \frac{2}{3}$, so $c(2) \leq -\frac{1}{3}$
By the Intermediate Value Theorem since $c(1) > 0 > c(2)$ $\exists \omega \in [1, 2]$ such that $c(\omega) = 0$. By the earlier result (†) it is unique.
The rest is simple algebra, using

$$\begin{aligned} c(2\omega) &= c(\omega)c(\omega) - s(\omega)s(\omega) \\ &= -s^2(\omega) \\ &= -1 + c^2(\omega) \\ &= -1 c(4\omega) \qquad\qquad = c(2\omega)c(2\omega) - s(2\omega)s(2\omega) \\ &= 1 + (1 - c^2(2\omega)) \\ &= 1 \end{aligned}$$

$s'(x) = c(x) \geq 0$ for $0 \leq x \leq \omega$, so since $s(0) = 0$ the Mean Value Theorem says $s(x) \geq 0$ for $0 \leq x \leq \omega$
$s^2(\omega) + c^2(\omega) = 1$ so $s^2(\omega) = 1$ so since $s(\omega) \geq 0, s(\omega) = 1$

$s(2\omega) = 2s(\omega)c(\omega) = 0$ and $s(4\omega) = 2s(2\omega)c(2\omega) = 0$
$s(x+4\omega) = s(x)c(4\omega) + c(x)s(4\omega) = s(x)$ and $c(x+4\omega) = c(x)c(4\omega) - s(x)s(4\omega) = c(x)$ □

We now **define** $\pi = 2\omega$. If $\overline{x}$ and $\overline{y}$ are non-zero vectors in $\mathbb{R}^m$ we know by the Cauchy-Schwarz inequality that $|\overline{x}.\overline{y}| \leq \|\overline{x}\|\|\overline{y}\|$ and we may **define** the angle between the two vectors to be that θ with $0 \leq \theta \leq \pi$ which satisfies

$$\cos\theta = \frac{\overline{x}.\overline{y}}{\|\overline{x}\|\|\overline{y}\|}$$

We can also justify the standard use of polar coordinates.

Lemma 10.12. If $(x,y) \in \mathbb{R}^2$ and $(x,y) \neq (0,0)$ then there exist a unique $r > 0$ and θ with $0 \leq \theta < 2\pi$ such that

$$x = r\cos\theta, y = r\sin\theta$$

Proof. If $\underline{x}, \underline{y}$ are non-zero then Caughy Schwarz tells us

$$-1 \leq \frac{\underline{x}.\underline{y}}{||\underline{x}||\ ||\underline{y}||} < 1$$

We have already seen that $\cos'\theta = -\sin\theta < 0$ for $0 < \theta < \pi$ so cos is strictly decreasing, and continous on $[0,\pi]$ with $\cos 0 = 1, \cos\pi = -1$ so the equation

$$\cos\theta = \frac{\underline{x}.\underline{y}}{||\underline{x}||\ ||\underline{y}||}$$

has exactly only solution θ with $0 \leq \theta \leq \pi$. We call the angle between $\underline{x}$ and $\underline{y}$ θ
Suppose that $(x,y) \neq (0,0)$, then

$$-1 \leq \frac{x}{\sqrt{x^2+y^2}} \leq 1$$

so as before there is $\theta_1 \in [0,\pi]$ such that $\cos\theta_1 = \frac{x}{\sqrt{x^2+y^2}}$ and there is also $\theta_2 \in [\pi, 2\pi]$ such that $\cos\theta_2 = \frac{x}{\sqrt{x^2+y^2}}$
$\cos^2\theta_1 + \sin^2\theta_1 = 1 = \cos^2\theta_2 + \sin^2\theta_2$, $\frac{x}{\sqrt{x^2+y^2}} + \frac{y}{\sqrt{x^2+y^2}} = 1$ and $\sin\theta_1 \geq 0, \sin\theta_2 \leq 0$ so we can choose a unique $\theta \in [0,2\pi]$ such that $\cos\theta = \frac{x}{\sqrt{x^2+y^2}}, \sin\theta = \frac{y}{\sqrt{x^2+y^2}}$
(we must take care if $(x,y) = (1,0)$ or $(-1,0)$)
Writing r for the positive square root of (x^2+y^2)

$$x = r\cos\theta \qquad y = r\sin\theta$$

□

You may be unhappy with a procedure which reduces geometry to analysis. It is possible to produce treatments which soften the blow[4] but what we can not do is to justify analytic results by appealing to geometry and then appeal to analysis to justify the geometry.

11 Onwards to the complex plane

This section contains useful background material which does not really form part of the course. If time is short I shall omit it entirely.
The mean value theorem **fails** for differentiable functions $f : \mathbb{C} \to \mathbb{C}$. (See Example 11.5.) However, the constant value theorem holds.

Theorem 11.1. If $f : \mathbb{C} \to \mathbb{C}$ is differentiable and $f'(z) = 0$ for all $z \in \mathbb{C}$ then f is constant.

Proof. (This is an ad hoc proof, better ones will follow)
Fix y_0, we show that $\Re(f(x + iy_0))$ is constant

Proof. Set $u(x) = \Re(f(x + iy_0))$

$$\begin{aligned}\frac{u(x+\delta x) - u(x)}{\delta x} &= \Re(\frac{f(x+\delta x+iy_0) - f(x+iy_0)}{\delta x}) \\ &= \Re(f'(x+iy_0)) \\ &= 0\end{aligned}$$

So u is differentiable and $u'(x) = 0\ \forall x$, and $u : \mathbb{R} \to \mathbb{R}$ so u is constant

Now set $v(y) = \Re(f(x_0 + iy_0))$ (x_0 fixed)

$$\begin{aligned}\frac{v(y+\delta y) - v(y)}{\delta y} &= \Re(\frac{f(x_0+iy+i\delta y) - f(x_0+iy)}{\delta y}) \\ &= \Re(i.\frac{f(x_0+iy+i\delta y) - f(x_0+iy)}{i\delta y}) \\ &= \Re(if'(x_0+iy)) \\ &= 0\end{aligned}$$

so v is constant, and as $\Re(f(x + iy))$ is constant if we vary x and y $\Re(f(x_1 + iy_1)) = \Re(f(x_1 + iy_2)) = \Re(f(x_2 + iy_2))$ so $\Re(f(x))$ is constant □

Similarly $\Im(f(x))$ is constant (or look at $\Re(if(x))$), so f is constant □

[4] The matter is more subtle than it looks. Classical Euclidean geometry is 'weaker' than the geometry required to justify analysis and if you wished to obtain analysis from geometry you would need to add extra axioms.

The proof of Theorem 11.1 that I shall give is very ad hoc and you will meet better ones later.
Since the constant value theorem holds we can extend the proof of Theorem 10.2 (ii) to this wider context and obtain a version of the exponential function for complex numbers. In this section we work in $\mathbb{C}$ unless otherwise stated.

Lemma 11.2. The sum $\sum_{n=0}^{\infty} \frac{z^n}{n!}$ has infinite radius of convergence.

Proof.
$$|\frac{z^{n+1}}{(n+1)!}/\frac{z^n}{n!}| = \frac{|z|}{n+1} \to 0 \text{ as } n \to \infty$$
□

We can thus define a function $\exp : \mathbb{C} \to \mathbb{C}$ by

$$\exp(z) = \sum_{n=0}^{\infty} \frac{z^n}{n!}$$

Theorem 11.3. (i) The function $e : \mathbb{C} \to \mathbb{C}$ is everywhere differentiable with $e'(z) = e(z)$.
(ii) $e(z + w) = e(z)e(w)$ for all $z, w \in \mathbb{C}$.

Proof. i) Differentiate term by term
ii) Show that $\frac{d}{dz}(e(a - z)e(z)) = 0$ for a constant
Use the constant value theorem to deduce that $e(a - z)e(z) = e(a)\ \forall z$ so $e(z - z)e(z) = e(a)\ \forall z, \forall a$ and so $e(w)e(z) = e(z + w)$ □

Notice that the remaining parts of Theorem 10.2 are either meaningless or false (compare Theorem 10.2 (vi) with Lemma 11.4 (iii) which shows that $\exp : \mathbb{C} \to \mathbb{C}$ is not injective). We must be very careful in making the transition from real to complex.
We obtain a series of famous formulae.

Lemma 11.4. (i) If θ is real

$$\exp i\theta = \cos\theta + i \sin\theta$$

(ii) If x and y are real

$$\exp(x + iy) = (\exp x)(\cos y + i \sin y)$$

(iii) exp is periodic with period $2\pi i$, that is to say

$$\exp(z + 2\pi i) = \exp z$$

for all $z \in \mathbb{C}$.
(iv) $\exp(\mathbb{C}) = \mathbb{C} \setminus \{0\}$.

Proof. i)

$$\begin{aligned}\exp(i\theta) &= \sum_{n=0}^{\infty} \frac{i\theta}{n!} = \sum_{n=0}^{\infty} \frac{(i\theta)^{2n}}{(2n)!} + \sum_{n=0}^{\infty} \frac{(i\theta)^{2n+1}}{(2n+1)!} \\ &= \sum_{n=0}^{\infty} \frac{(-1)^n\theta^{2n}}{(2n)!} + i\sum_{n=0}^{\infty} \frac{(-1)^n\theta^{2n+1}}{(2n+1)!} \\ &= \cos\theta + i\sin\theta \text{ for } \theta \in \mathbb{R}\end{aligned}$$

ii)

$$\begin{aligned}\exp(x+iy) &= \exp(x)\exp(iy) \\ &= \exp(x)(\cos y + i\sin y) \text{ for } x, y \in \mathbb{R}\end{aligned}$$

iii)

$$\begin{aligned}\exp(z+2\pi i) &= \exp(z)\exp(2\pi i) \\ &= \exp(z)(\cos 2\pi + i\sin 2\pi) \\ &= \exp z\end{aligned}$$

so exp is periodic with period $2\pi i$ iv) If $z \in \mathbb{C}$ and $z \neq 0$ then $z = x + iy$ with $(x, y) \neq (0, 0)$, so $z = r\cos\theta + ir\sin\theta$ for some $r > 0, \theta \in \mathbb{R}$, which in turn is equal to $r\exp(i\theta) = \exp(\log r)\exp(i\theta) = \exp(\log r + i\theta)$ so $\exp(\mathbb{C}) \supset \mathbb{C} \setminus \{0\}$
It is impossible to have $\exp(\omega) = 0$ for then $\exp(z) = \exp(z - \omega)\exp(\omega) = 0\ \forall z$, but set $z = 0, \exp(0) = 1\#$
So the range of exp is $\mathbb{C} \setminus \{0\}$, and thus exp takes every value except 0 infinitely often, as it is periodic □

Example 11.5. Observe that $\exp 0 = \exp 2\pi i = 1$ but $\exp'(z) = \exp(z) \neq 0$ for all $z \in \mathbb{C}$. Thus the mean value theorem does not hold for differentiable functions $f : \mathbb{C} \to \mathbb{C}$.

It is also possible to extend the definition of sine and cosine to the complex plane but the reader is warned that the behaviour of the new functions may be somewhat unexpected. Since these extended functions most certainly do not form part of this course (though you will be expected to know them after the Complex Methods course) their study is left as an exercise.

Exercise 11.6. (i) Explain why the infinite sums

$$\sin z = \sum_{n=0}^{\infty} \frac{(-1)^n z^{2n+1}}{(2n+1)!} \text{ and } \cos z = \sum_{n=0}^{\infty} \frac{(-1)^n z^{2n}}{(2n)!}$$

converge everywhere and are differentiable everywhere with $\sin' z = \cos z$, $\cos' z = -\sin z$.
(ii) Show that

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

and

$$\exp iz = \cos z + i \sin z$$

for all $z \in \mathbb{C}$.
(iii) Show that

$$\sin(z+w) = \sin z \cos w + \cos z \sin w, \cos(z+w) = \cos z \cos w - \sin z \sin w$$

and $(\sin z)^2 + (\cos z)^2 = 1$ for all $z, w \in \mathbb{C}$.
(iv) $\sin(-z) = -\sin z$, $\cos(-z) = \cos z$ for all $z \in \mathbb{C}$.
(v) sin and cos are 2π periodic in the sense that

$$\sin(z+2\pi) = \sin z \text{ and } \cos(z+2\pi) = \cos z$$

for all $z \in \mathbb{C}$.
(vi) If x is real then $\sin ix = i \sinh x$ and $\cos ix = \cosh x$.
(vii) Recover the addition formulae for sinh and cosh by setting $z = ix$ and $w = iy$ in part (iii).
(ix) Show that $|\sin z|$ and $|\cos z|$ are bounded if $|\Im z| \le K$ for some K but that $|\sin z|$ and $|\cos z|$ are unbounded on $\mathbb{C}$.

Proof. ii)

$$\begin{aligned}
\exp(iz) - \exp(-iz) &= \sum_0^\infty \frac{(iz)^n}{n!} - \frac{(-iz)^n}{n!} \\
&= \sum_0^\infty \frac{1}{n!}(i^n - (-i)^n)z^n \\
&= \sum_0^\infty \frac{1}{(2r+1)!} 2i^{2r+1} z^{2r+1} \\
&= 2i \sum_0^\infty \frac{1}{(2r+1)!}(-1)^r z^{2r+1}
\end{aligned}$$

so $\sin z = \frac{exp(iz) - exp(-iz)}{2i}$ and similarly $\cos z = \frac{\exp(iz)+\exp(-iz)}{2}$
For the other parts
iii) Use ii) or i) and the CVT
iv) Use ii)
v) By iii) $\sin(z+2\pi) = \sin z \cos 2\pi + \cos z \sin 2\pi = \sin z$
vi) Use ii) ix) Use ii), $|z_1 - z_2| \le \Big||z_1| - |z_2|\Big|$ □

However, as you were already shown in the Algebra and Geometry course and will be shown again in the Complex Methods course

There is no logarithm defined on all of $\mathbb{C} \setminus \{0\}$.

Exercise 11.7. Suppose, if possible, that there exists a continuous $L : \mathbb{C}\setminus\{0\} \to \mathbb{C}$ with $\exp(L(z)) = z$ for all $z \in \mathbb{C} \setminus \{0\}$.
(i) If θ is real, show that $L(\exp(i\theta)) = i(\theta + 2\pi n(\theta))$ for some $n(\theta) \in \mathbb{Z}$.
(ii) Define $f : \mathbb{R} \to \mathbb{R}$ by

$$f(\theta) = \frac{1}{2\pi}\left(\frac{L(\exp i\theta) - L(1)}{i} - \theta\right)$$

Show that f is a well defined continuous function, that $f(\theta) \in \mathbb{Z}$ for all $\theta \in \mathbb{R}$, that $f(0) = 0$ and that $f(2\pi) = -1$.
(iii) Show that the statements made in the last sentence of (ii) are incompatible with the intermediate value theorem and deduce that no function can exist with the supposed properties of L.
(iv) Discuss informally what connection, if any, the discussion above has with the existence of the international date line.

Proof. Let us assume that such a L exists
$\exp(L(\exp(i\theta))) = \exp(i\theta)$ so $L(\exp(i\theta)) = i\theta + 2N\pi i$ for some integer N, call this $L(\exp(i\theta)) = i\theta + 2\pi n(\theta)i$
Let $f(\theta) = \frac{1}{2\pi}(\frac{L(\exp(i\theta))-L(1)}{i} - \theta) = \frac{1}{2\pi}(\frac{i\theta+2\pi n(\theta)i-2\pi n(0)i}{i}) = n(\theta) - n(0) \in \mathbb{Z}$
f is continuous (from chain rule etc), and $f(0) = \frac{1}{2\pi}(\frac{L(1)-L(1)}{i}) = 0, f(1) = \frac{1}{2\pi}(\frac{L(1)-L(1)}{i} - 2\pi) = -1$
Therefore the Intermediate Value Theorem states that $\exists c$ such that $f(c) = -\frac{1}{2} \notin \mathbb{Z}$ # ☐

A similar argument shows that it is not possible to produce a continuous square root on the complex plane.

Exercise 11.8. Show by modifyiing the argument of Exercise 11.7, that there does not exist a continuous $S : \mathbb{C} \to \mathbb{C}$ with $S(z)^2 = z$ for all $z \in \mathbb{C}$.

More generally, it is not possible to define continuous non-integer powers z^α on the complex plane. (Of course, $z \mapsto z^n$ is well behaved if n is an integer.) However, in the special case when x is real and strictly positive we can define

$$x^z = \exp(z \log x)$$

without problems and this enables us to write $\exp z = e^z$ where $e = \exp 1$. Surprisingly, Exercise 11.7 is not an end but a beginning of much important mathematics – but that is another story.
(The next bit was given in lectures, but did not form part of the printed notes)
Take the circle $\{e^{i\theta} : 0 \le \theta < 2\pi\}$. Say that $e^{i\theta_1} \sim e^{i\theta_2}$ iff $\theta_1 - \theta_2 = 2\pi\mathbb{Z}$, so there are equivalence classes $[e^{i\phi}] = \{e^{i\theta} : e^{i\theta} \sim e^{i\phi}\}$
Pick one member from each equivalence class (one representative from each class) and call the result E
There are countably many rationals in $[0, 1)$, call them $x_1, x_2, x_3, \ldots$
Look at $E_n = \{e^{2\pi i x_n + i\theta} : e^{i\theta} \in E\} = e^{2\pi i x_n}E$

$E_n \cap E_m = \phi$ and $\bigcup_{\forall n} E_n =$ whole circle
So what is the "mass" of E_n if the circle "weighs" 1g?
We have a paradox - if it's posiitve then the circle weighs ∞ (made up of a union of countable infinite objects), and if 0 then the circle weighs 0
Therefore we have constructed a set with no "natural" definition of area. We need to carefully define area and also integrals as a result and even still we will find some integrals which cannot be found.

12 The Riemann integral

At school we are taught that an integral is the area under a curve. If pressed the framer of this definition might reply that everybody knows what area is, but then everybody knows what honey tastes like. But does honey taste the same to you as it does to me? Perhaps the question is unanswerable but, for many practical purposes, it is sufficient that we agree on what we call honey.
In order to agree on what an integral is, we need a definition which does not depend on intuition. It is important that, as far as possible, the properties of our formally defined integral shall agree with our intuitive ideas on area **but we have to prove this agreement** and not simply assume it.
In this section we introduce a notion of integral due to Riemann. For the moment we only attempt to define our integral for bounded functions on bounded intervals.
Let $f : [a, b] \to \mathbb{R}$ be a function such that there exists a K with $|f(x)| \leq K$ for all $x \in [a, b]$. [To see the connection with 'the area under the curve' it is helpful to suppose initially that $0 \leq f(x) \leq K$. However, all the definitions and proofs work more generally for $-K \leq f(x) \leq K$.]
A dissection $\mathcal{D}$ of $[a, b]$ is a finite subset of $[a, b]$ containing the end points a and b. By convention, we write

$$\mathcal{D} = \{x_0, x_1, \ldots, x_n\} \text{ with } a = x_0 \leq x_1 \leq x_2 \leq \ldots \leq x_n = b$$

We define the *upper sum* and *lower sum* associated with $\mathcal{D}$ by

$$S(f, \mathcal{D}) = \sum_{j=1}^{n} (x_j - x_{j-1}) \sup_{x \in [x_{j-1}, x_j]} f(x),$$
$$s(f, \mathcal{D}) = \sum_{j=1}^{n} (x_j - x_{j-1}) \inf_{x \in [x_{j-1}, x_j]} f(x)$$

[Observe that, *if the integral $\int_a^b f(t)\,dt$ exists,* then the upper sum ought to provide an upper bound and the lower sum a lower bound for that integral.]
The next lemma is obvious but useful.

Lemma 12.1. If $\mathcal{D}$ and $\mathcal{D}'$ are dissections with $\mathcal{D}' \supseteq \mathcal{D}$ then

$$S(f, \mathcal{D}) \geq S(f, \mathcal{D}') \geq s(f, \mathcal{D}') \geq s(f, \mathcal{D})$$

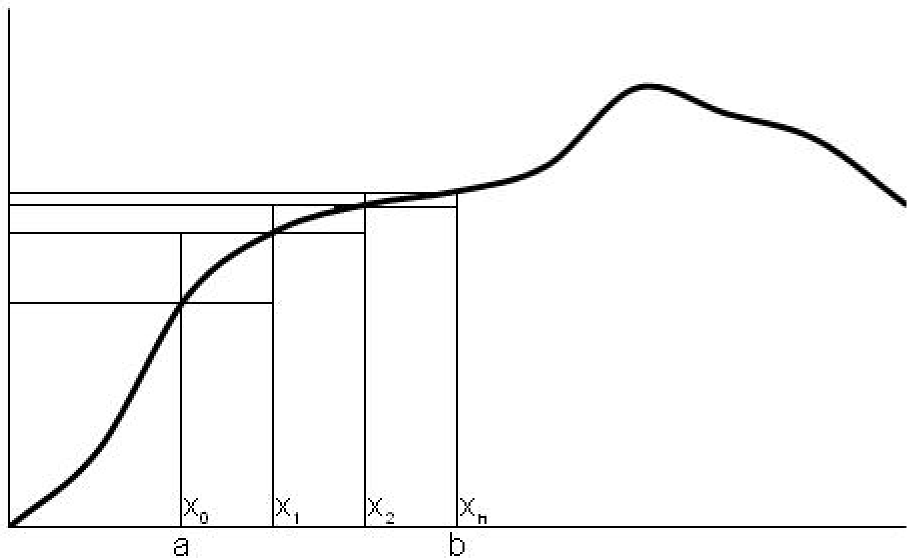

Figure 2: Upper and Lower sums for integrals

Proof. If $x_1 \geq y \geq x_0$ then

$$\begin{aligned}(x_1 - y) \sup_{t\in[y,x_1]} f(t) + (y - x_0) \sup_{t\in[y,x_0]} f(t) &\leq (x_1 - y) \sup_{t\in[x_0,x_1]} f(t) + (y - x) \sup_{t\in[x_0,x_1]} f(t) \\ &= (x_1 - x_0) \sup_{t\in[x_0,x_1]} f(t)\end{aligned}$$

So adding one point does not increase the upper sum, thus by induction

$$S(f, \mathcal{D}) \geq S(f, \mathcal{D}') \geq s(f, \mathcal{D}') \geq s(f, \mathcal{D})$$

□

The next lemma is again hardly more than an observation but it is the key to the proper treatment of the integral.

Lemma 12.2 (Key integration property)**.** If $f : [a, b] \to \mathbb{R}$ is bounded and $\mathcal{D}_1$ and $\mathcal{D}_2$ are two dissections, then

$$S(f, \mathcal{D}_1) \geq S(f, \mathcal{D}_1 \cup \mathcal{D}_2) \geq s(f, \mathcal{D}_1 \cup \mathcal{D}_2) \geq s(f, \mathcal{D}_2). \qquad \bigstar$$

Proof. Immediate from the preceding lemma □

The inequalities ★ tell us that, whatever dissection you pick and whatever dissection I pick, your lower sum cannot exceed my upper sum. There is no way we can put a quart in a pint pot[5].

[5]Or put a litre in a half litre bottle.

Since $S(f, \mathcal{D}) \geq -(b-a)K$ for all dissections $\mathcal{D}$ we can define the *upper integral* as $I^*(f) = \inf_{\mathcal{D}} S(f, \mathcal{D})$. We define the *lower integral* similarly as $I_*(f) = \sup_{\mathcal{D}} s(f, \mathcal{D})$. The inequalities ★ tell us that these concepts behave well.

Lemma 12.3. If $f : [a, b] \to \mathbb{R}$ is bounded, then $I^*(f) \geq I_*(f)$.

Proof. $S(f, D_1) \geq s(D_2, f)$, thence immediate □

[Observe that, *if the integral* $\int_a^b f(t)\,dt$ *exists,* then the upper integral ought to provide an upper bound and the lower integral a lower bound for that integral.]
If $I^*(f) = I_*(f)$, we say that f is Riemann integrable and we write

$$\int_a^b f(x)\,dx = I^*(f)$$

The following lemma provides a convenient criterion for Riemann integrability.

Lemma 12.4. (i) A bounded function $f : [a, b] \to \mathbb{R}$ is Riemann integrable if and only if, given any $\epsilon > 0$, we can find a dissection $\mathcal{D}$ with

$$S(f, \mathcal{D}) - s(f, \mathcal{D}) < \epsilon$$

(ii) A bounded function $f : [a, b] \to \mathbb{R}$ is Riemann integrable with integral I if and only if, given any $\epsilon > 0$, we can find a dissection $\mathcal{D}$ with

$$S(f, \mathcal{D}) - s(f, \mathcal{D}) < \epsilon \text{ and } |S(f, \mathcal{D}) - I| \leq \epsilon$$

Proof. i) Suppose f is Riemann integrable. Then $I^*(f) = I_*(f) = If$ say
Given any $\epsilon > 0$ we can find $\mathcal{D}_1$ such that $I^*(f) \geq S(f, \mathcal{D}_1) - \epsilon$ and $\mathcal{D}_2$ such that $I_*(f) \leq s(f, \mathcal{D}_2) + \epsilon$

$$\begin{aligned}\therefore I*(f) + \epsilon &\geq S(f, \mathcal{D}_1) \\ &\geq S(f, \mathcal{D}_1 \cup \mathcal{D}_2) \\ &\geq s(f, \mathcal{D}_1 \cup \mathcal{D}_2) \\ &\geq s(f, \mathcal{D}_2) = I_*(f) - \epsilon\end{aligned}$$

So $S(f, \mathcal{D}_1 \cup \mathcal{D}_2) - s(f, \mathcal{D}_1 \cup \mathcal{D}_2) < 2\epsilon$
Conversely, suppose given any $\epsilon > 0$ we can find $\mathcal{D}$ with $S(f, \mathcal{D}) - s(f, \mathcal{D}) < \epsilon$. Then

$$\begin{aligned}0 &\leq I^*(f) - I_*(f) \\ &\leq S(f, \mathcal{D}) - s(f, \mathcal{D}) \\ &< \epsilon\end{aligned}$$

So, since ϵ is arbitrary $I^*(f) = I_*(f)$
ii) Observe that i) shows the existence of the integral and note that

$$S(f, \mathcal{D}) \geq I^*(f) \geq I_*(f) \geq s(f, \mathcal{D})$$

so if $S(f, \mathcal{D}) - s(f, \mathcal{D}) < \epsilon$, $|S(f, \mathcal{D}) - I^*(f)| < \epsilon$ and $|s(f, \mathcal{D}) - I_*(f)| < \epsilon$ □

Many students are tempted to use Lemma 12.4 (ii) as the *definition* of the Riemann integral. The reader should reflect that, without the inequality ★, it is not even clear that such a definition gives a unique value for I. (This is only the first of a series of nasty problems that arise if we attempt to develop the theory without first proving ★, so I strongly advise the reader not to take this path.)

We now prove a series of standard results on the integral.

Lemma 12.5. (i) The function $J : [a, b] \to \mathbb{R}$ given by $J(t) = 1$ is integrable and

$$\int_a^b 1\, dx = b - a$$

(ii) If $f, g : [a, b] \to \mathbb{R}$ are Riemann integrable, then so is $f + g$ and

$$\int_a^b f(x) + g(x)\, dx = \int_a^b f(x)\, dx + \int_a^b g(x)\, dx$$

(iii) If $f : [a, b] \to \mathbb{R}$ is Riemann integrable, then $-f$ is and

$$\int_a^b (-f(x))\, dx = -\int_a^b f(x)\, dx$$

(iv) If $\lambda \in \mathbb{R}$ and $f : [a, b] \to \mathbb{R}$ is Riemann integrable, then λf is Riemann integrable and

$$\int_a^b \lambda f(x)\, dx = \lambda \int_a^b f(x)\, dx$$

(v) If $f, g : [a, b] \to \mathbb{R}$ are Riemann integrable functions with $f(t) \geq g(t)$ for all $t \in [a, b]$, then

$$\int_a^b f(x)\, dx \geq \int_a^b g(x)\, dx$$

Proof. i) $S(1, \mathcal{D}) = \sum (x_j - x_{j-1})1 = x_n - x_{n-1} + x_{n-1} - x_{n-2} + \cdots - x_0 = x_n - x_0 = b - a$, and similarly $s(1, \mathcal{D}) = b - a$, so $I^*1 = b - a = I_*1$ so 1 is integrable $\int_a^b 1 = b - a$

ii) Suppose f, g integrable on $[m, l]$. Then given $\epsilon > 0$ we can find a dissection $\mathcal{D}_1$ such that $S(f, \mathcal{D}_1) - s(f, \mathcal{D}_1) < \epsilon$ and a dissection $\mathcal{D}_2$ such that $S(g, \mathcal{D}_2) - s(g, \mathcal{D}_2) < \epsilon$

Take $\mathcal{D} = \mathcal{D}_1 \cup \mathcal{D}_2$

$$S(f, \mathcal{D}_1) \geq S(f, \mathcal{D}) \geq If \geq s(f, \mathcal{D}) \geq s(f, \mathcal{D}_1)$$

so $S(f, \mathcal{D}) - s(f, \mathcal{D}) < \epsilon$, and similarly $S(g, \mathcal{D}) - s(g, \mathcal{D}) < \epsilon$

$$
\begin{aligned}
S(f+g,\mathcal{D}) &= \sum (x_j - x_{j-1}) \sup_{t\in[x_{j-1},x_j]} (f(t)+g(t)) \\
&\geq \sum (x_j - x_{j-1}) \sup_{t\in[x_{j-1},x_j]} f(t) + \sup_{t\in[x_{j-1},x_j]} g(t) \\
&= S(f,\mathcal{D}) + S(g,\mathcal{D}) \\
&\geq s(f,\mathcal{D}) + s(g,\mathcal{D})
\end{aligned}
$$

$\geq s(f+g,\mathcal{D})$ similarly

So

$$
\begin{aligned}
0 &\leq S(f+g,\mathcal{D}) + s(f+g,\mathcal{D}) \\
&\leq (S(f,\mathcal{D}) - s(f,\mathcal{D})) + (S(g,\mathcal{D}) - s(g,\mathcal{D})) \\
&< 2\epsilon
\end{aligned}
$$

so since $\epsilon > 0$ is arbitrary $f+g$ is integrable. Moreover

$$
\begin{aligned}
I(f+g) &\geq s(f,\mathcal{D}) + s(g,\mathcal{D}) \\
&\geq (If - \epsilon) + (Ig - \epsilon) \\
&= If + Ig - 2\epsilon
\end{aligned}
$$

and sum to get $If + Ig + 2\epsilon \geq I(f+g)$, so since this is true $\forall \epsilon > 0$ $If + Ig = I(f+g)$

$$
\int_a^b f + \int_a^b g = \int_a^b (f+g)
$$

iii) Observe that

$$
\begin{aligned}
S(-f,\mathcal{D}) &= -s(f,\mathcal{D}) \qquad \text{so} \qquad I^*(-f) = -I_*(f) \\
s(-f,\mathcal{D}) &= -S(f,\mathcal{D}) \qquad \text{so} \qquad I_*(-f) = -I^*(f)
\end{aligned}
$$

so f integrable $\Rightarrow -f$ integrable
iv) Consider first the case when $\lambda \geq 0$

$$
\begin{aligned}
S(\lambda f,\mathcal{D}) &= \lambda S(f,\mathcal{D}) \\
\text{so } I^*(\lambda f) &= \lambda I^* f \\
\text{similarly } I_*(\lambda f) &= \lambda I_* f
\end{aligned}
$$

so if f is integrable so is λf and $I(\lambda f) = \lambda If$
For the case $\lambda < 0$ consider $|\lambda|(-f) = \lambda f$ and use the case $\lambda \geq 0$
v) If $f \geq g$ then $S(f,\mathcal{D}) \geq S(g,\mathcal{D})$ so $I^* f \geq I^* g$ so if f and g are integrable $If \geq Ig$ □

Lemma 12.6. (i) If $f : [a,b] \to \mathbb{R}$ is Riemann integrable then so is f^2.
(ii) If $f, g : [a,b] \to \mathbb{R}$ are Riemann integrable, then so is the product fg.

Proof. i) Suppose $|f(t)| \leq M \ \forall t$. Then

$$\begin{aligned} |f(t)^2 - f(s)^2| &= |f(t) - f(s)||f(t) + f(s)| \\ &\leq |f(t) - f(s)|(|f(t)| + |f(s)|) \\ &= 2M|f(t) - f(s)| \end{aligned}$$

so $\sup_{t\in J} f(t)^2 - \inf_{s\in J} f(s)^2 \leq 2M(\sup_{t\in J} f(t) - \inf_{s\in J} f(s))$. Thus

$$\begin{aligned} S(f^2, \mathcal{D}) - s(f^2, \mathcal{D}) &= \sum (x_j - x_{j-1}) \sup_{t\in[x_{j-1},x_j]} f(t)^2 - \inf_{s\in[x_{j-1},x_j]} f(s)^2 \\ &\leq 2M \sum (x_j - x_{j-1} (\sup_{t\in[x_{j-1},x_j]} f(t) - \inf_{s\in[x_{j-1},x_j]} f(s)) \\ &= 2M(S(f, \mathcal{D}) - s(f, \mathcal{D})) \end{aligned}$$

If f is integrable then given any $\epsilon > 0$ we can find $\mathcal{D}$ such that $S(f, \mathcal{D}) - s(f, \mathcal{D}) < \frac{\epsilon}{2M+1}$ and so $S(f^2, \mathcal{D}) - s(f^2, \mathcal{D}) < \epsilon$ so f^2 is integrable
iii) If f, g are integrable then $f + g, f - g$ are integrable so $(f + g)^2$ and $(f - g)^2$ are integrable, so $4fg = (f + g)^2 - (f - g)^2$ is integrable, and thus $fg = \frac{1}{4}(4fg)$ is integrable ☐

Lemma 12.7. (i) If $f : [a, b] \to \mathbb{R}$ is Riemann integrable then so is $f_+(t) = \max(f(t), 0)$.
(ii) If $f : [a, b] \to \mathbb{R}$ is Riemann integrable, then $|f|$ is Riemann integrable and

$$\int_a^b |f(x)|\,dx \geq \left| \int_a^b f(x)\,dx \right|$$

Proof. i) $|f(t) - f(s)| \geq |f_+(t) - f_+(s)|$ so $\sup_{t\in J}(f(t)) - \inf_{s\in J}(f(s)) \geq \sup_{t\in J} f_+(t) - \inf_{s\in J} f_+(s)$
so along the same lines as the previous lemma $S(f, \mathcal{D}) - s(f, \mathcal{D}) \geq S(F_+, \mathcal{D}) - s(f_+, \mathcal{D})$, so if f is integrable, given $\epsilon > 0, \exists \mathcal{D}$ such that $S(f, \mathcal{D}) - s(f, \mathcal{D}) < \epsilon$ so $S(f_+, \mathcal{D}) - s(f_+, \mathcal{D}) < \epsilon$. Thus f_+ is integrable
ii) Write

$$\begin{aligned} f_+(t) &= f(t) \quad \text{for } f(t) \geq 0 \\ &= 0 \quad \text{for } f(t) < 0 \\ f_-(t) &= 0 \quad \text{for } f(t) \geq 0 \\ &= -f(t) \quad \text{for } f(t) < 0 \end{aligned}$$

Then $f = f_+ - f_-$ and $|f| = f_+ + f_-$
By i) if f is integrable so are f_+ and f_-, so $|f| = f_+ + f_-$ is integrable
$|f| = f_+ + f_-$ so $\int_a^b |f| = \int f_+ + \int f_-$, $\int_a^b f = \int f_+ - \int f_-$
$\int f_+, \int f_- \geq 0$ so $\int_a^b |f| \geq |\int_a^b f|$ ☐

Notice that we often only need the much weaker inequality

$$\left|\int_a^b f(x)\,dx\right| \le \sup_{t\in[a,b]} |f(t)|(b-a)$$

usually stated as

$$|\text{integral}| \le \text{length} \times \sup$$

(For the simplest proof, write $M = \sup |f|$, then $M \ge f(s) \ge -M\ \forall s \in [a,b]$
so $M(b-a) \ge S(f,\mathcal{D}) \ge s(f,\mathcal{D}) \ge -M(b-a)$
so $M(b-a) \ge I^*f \ge I_*f \ge -M(b-a)$
so if If exists, $|If| \ge M(b-a)$
The next lemma is also routine in its proof but continues our programme of showing that the integral has all the properties we expect.

Lemma 12.8. Suppose that $a \le c \le b$ and that $f : [a,b] \to \mathbb{R}$ is a bounded function. Then, f is Riemann integrable on $[a,b]$ if and only if $f|_{[a,c]}$ is Riemann integrable on $[a,c]$ and $f|_{[c,b]}$ is Riemann integrable on $[c,b]$. Further, if f is Riemann integrable on $[a,b]$, then

$$\int_a^b f(x)\,dx = \int_a^c f|_{[a,c]}(x)\,dx + \int_c^b f|_{[c,b]}(x)\,dx$$

Proof. Suppose $f : [a,b] \to \mathbb{R}$ and $a < c < b$. If $\mathcal{D}_1$ is a dissection of $[a,c]$ with $a = x_0 \le x_1 \le \ldots \le x_n = c$ and $\mathcal{D}_2$ is a dissection of $[c,b]$ with $c = x_{n+1} \le x_{n+1} \le \ldots \le x_m = b$

$$\begin{aligned} S_{[a,b]}(f,\mathcal{D}_1\cup\mathcal{D}_2) &= \sum (x_j - x_{j-1}) \sup_{t\in[a,b]} f(t) \\ &= S_{[a,c]}(f,\mathcal{D}_1) + S_{[c,b]}(f,\mathcal{D}_2) \end{aligned}$$

and so $I^*_{[a,b]}f \le I^*_{[a,c]}f + I^*_{[c,b]}f$
On the other hand if $\mathcal{D}$ is a dissection of $[a,b]$

$$\begin{aligned} S_{[a,b]}(f,\mathcal{D}) &\ge S_{[a,b]}(f,\mathcal{D}\cup\{c\}) \\ &= S_{[a,c]}(f,\mathcal{D}_1) + S_{[c,b]}(f,\mathcal{D}_2) \end{aligned}$$

where $\mathcal{D}_1$ is a dissection of $[a,c]$ and $\mathcal{D}_2$ is a dissection of $[c,b]$, and thus $\mathcal{D}_1\cup\mathcal{D}_2 = \mathcal{D}\cup\{c\}$
so $I^*_{[a,b]}f \ge I^*_{[a,c]}f + I^*_{[c,b]}f$
$\therefore I^*_{[a,b]}f = I^*_{[a,c]}f + I^*_{[c,b]}f$
Similarly $I_{*[a,b]}f = I_{*[a,c]}f + I_{*[c,b]}f$
So if f is integrable over $[a,c]$, $[c,b]$ which gives us

$$I^*_{[a,c]}f = I_{*[a,c]}f = I_{[a,c]}f \qquad I^*_{[c,b]}f = I_{*[c,b]}f = I_{[c,b]}f$$

and thus

$$I^*_{[a,b]}f = I_{*[a,b]}f = I_{[a,b]}f$$

Conversely, if f is integrable over $[a, b]$

$$\begin{aligned} I_{[a,b]}f &= I^*_{[a,b]}f = I^*_{[a,c]}f + I^*_{[c,b]}f \\ &\geq I_{*[a,c]}f + I_{*[c,b]}f = I_{*[a,b]}f \\ &= I_{[a,b]}f \end{aligned}$$

so f is integrable on $[a, c], [c, b]$ with

$$I_{[a,b]}f = I_{[a,c]}f + I_{[c,b]}f$$

□

In a very slightly less precise and very much more usual notation we write

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

There is a standard convention which says that, if $b \geq a$ and f is Riemann integrable on $[a, b]$, we define

$$\int_b^a f(x)\,dx = -\int_a^b f(x)\,dx$$

It is, however, a convention that requires care in use.

Exercise 12.9. Suppose that $b \geq a$, $\lambda, \mu \in \mathbb{R}$, and f and g are Riemann integrable. Which of the following statements are always true and which are not? Give a proof or counter-example. If the statement is not always true, find an appropriate correction and prove it.

(i) $\displaystyle\int_b^a \lambda f(x) + \mu g(x)\,dx = \lambda \int_b^a f(x)\,dx + \mu \int_b^a g(x)\,dx$.

(ii) If $f(x) \geq g(x)$ for all $x \in [a, b]$, then $\displaystyle\int_b^a f(x)\,dx \geq \int_b^a g(x)\,dx$.

13 Some properties of the integral

Not all bounded functions are Riemann integrable

Lemma 13.1. If $f : [0, 1] \to \mathbb{R}$ is given by

$$\begin{aligned} f(x) &= 1 \qquad \text{when } x \text{ is rational,} \\ f(x) &= 0 \qquad \text{when } x \text{ is irrational,} \end{aligned}$$

then f is not Riemann integrable

This does not worry us unduly, but makes it more important to show that the functions we wish to be integrable actually are.
Our first result goes back to Riemann (indeed, essentially, to Newton and Leibniz).

Lemma 13.2. (i) If $f : [a, b] \to \mathbb{R}$ is increasing, then f is Riemann integrable.
(ii)If $f : [a, b] \to \mathbb{R}$ can be written as $f = f_1 - f_2$ with $f_1, f_2 : [a, b] \to \mathbb{R}$ increasing, then f is Riemann integrable.
(iii) If $f : [a, b] \to \mathbb{R}$ is piecewise monotonic, then f is Riemann integrable.

Proof. i) We show that given ϵ $\exists$ a $\mathcal{D}$ such that $S(f, \mathcal{D}) - s(f, \mathcal{D}) < \epsilon$ (the hard bit is to find the dissection)
Choose $\mathcal{D}$ to be $a, a + \frac{b-a}{N}, a + \frac{2(b-a)}{N}, \dots, b$

$$\begin{aligned} S(f, \mathcal{D}) - s(f, \mathcal{D}) &= \sum \Big(\sup_{t \in [a + \frac{(b-a)(r-1)}{N}, a + \frac{(b-a)r}{N}]} f(t) - \inf_{t \in [a + \frac{(b-a)(r-1)}{N}, a + \frac{(b-a)r}{N}]} f(t) \Big) \\ &= \sum (f(a + \frac{(b-a)r}{N}) - f(a + \frac{(b-a)(r-1)}{N})) . \frac{b-a}{N} \\ &= (f(b) - f(a)) \frac{b-a}{N} \end{aligned}$$

By choosing N large enough we have $(f(b) - f(a)) \frac{b-a}{N} < \epsilon$
ii) This means f is the sum of an increasing and decreading function (both Riemann integrable), so thus it is Riemann integrable
iii) Using Lemma 12.8 and induction □

It should be noted that the results of Lemma 13.2 do not require f to be continuous. (For example, the Heaviside function, given by $H(t) = 0$ for $t < 0$, $H(t) = 1$ for $t \geq 0$ is increasing but not continuous.) It is quite hard to find a continuous function which is not the difference of two increasing functions but an example is given in question 18 on the 4th example sheet.
The proof of the next result is starred. Next year you will see a simpler proof based on a different idea (that of uniform continuity).

Theorem 13.3. If $f : [a, b] \to \mathbb{R}$ is continuous then f is Riemann integrable.

Proof. Suppose that $f : [a, b] \to \mathbb{R}$ is not integrable. Then there exists a $\delta > 0$ such that $I^*_{[a,b]} f - I_{*[a,b]} f = \delta(b - a)$
Set $a_0 = a, b_0 = b$, $c_0 = c = \frac{a_0 + b_0}{2}$

$$I^*_{[a,c]} f - I^*_{[c,b]} f = I^*_{[a,b]} f \qquad I_{*[a,c]} f - I_{*[c,b]} f = I_{*[a,b]} f$$

so

$$\begin{aligned} \max\{I^*_{[a,c]} f - I_{*[a,c]} f, I^*_{[c,b]} f - I_{*[c,b]} f\} &\geq \frac{1}{2}(I^*_{[a,b]} f - I_{*[a,b]} f) \\ &= \frac{1}{2}\delta(b - a) \end{aligned}$$

Choose $a_1 = a_0, b_1 = c_0$ if $I^*_{[a,c]} f - I_{*[a,c]} f \geq \frac{1}{2}\delta(b-a)$, otherwise $a_1 = c_0, b_1 = b_0$
Observe

$$
\begin{aligned}
&a_0 \le a_1 \le b_1 \le b \\
&I^*_{[a_1,b_1]} f - I_{*[a_1,b_1]} f \ge (b_1 - a_1)\delta \\
&b_1 - a_1 = 2^{-1}(b_0 - a_0)
\end{aligned}
$$

Extend this by induction to get

$$
\begin{aligned}
&a_0 \le a_1 \le \ldots \le a_n \le b_n \ldots \le b_1 \le b_0 \\
&I^*_{[a_n,b_n]} f - I_{*[a_n,b_n]} f \ge (b_n - a_n)\delta \\
&b_n - a_n = 2^{-n}(b_0 - a_0)
\end{aligned}
$$

By the same arguments that we have made before, the a_n are increasing, bounded above, and so a_n tend to a limit α, and as $b_n - a_n \to 0$ so $b_n \to \alpha$, and further $b_n \le b$ so $\alpha \le b$ and $a_n \le \alpha$ so $a \le \alpha$
Examine what happens at α
f is continuous at α so $\exists \eta > 0$ such that $|f(x) - f(\alpha)| \le \frac{\delta}{4}$ $\forall x \in [a,b], |x - \alpha| \le \eta$
Pick n such that $|a_n \alpha|, |b_n \alpha| < \eta$
Then $f(\alpha) - \frac{\delta}{4} \le f(x) \le f(\alpha) + \frac{\delta}{4}$ for $x \in [a_n, b_n]$, so

$$
\begin{aligned}
I^*_{[a_n,b_n]} f - I_{*[a_n,b_n]} f &\le 2\frac{\delta}{4}(b_n - a_n) \\
&= \frac{\delta}{2}(b_n - a_n) \\
&< \delta(b_n - a_n)\#
\end{aligned}
$$

So $I^* f = I_* f$ and f is integrable $\square$

We complete the discussion of integration and this course with a series of results which apply only to *continuous* functions.
Our first result is an isolated, but useful, one.

Lemma 13.4. If $f : [a,b] \to \mathbb{R}$ is continuous, $f(t) \ge 0$ for all $t \in [a,b]$ and

$$
\int_a^b f(t)\, dt = 0
$$

it follows that $f(t) = 0$ for all $t \in [a,b]$.

Proof. Suppose $f : [a,b] \to \mathbb{R}$ is continuous and $f(t) \ge 0$ $\forall t \in [a,b]$
Suppose $\exists c \in [a,b]$ and $f(c) \ne 0$
Then $f(c) > 0$ and by continuity $\exists \delta > 0$ such that $|f(x) - f(c)| < \frac{f(c)}{2}$ for $|x - c| \le \delta$ and $x \in [a,b]$, so $f(x) > \frac{f(c)}{2}$ is $|x - c| \le \delta$ and $x \in [a,b]$
Write $J = \{a \in [a,b] : |x - c| \le \delta\}$
J is an interval, with the length of $J \ge \delta$
(note: we must allow for the possibility that $c = a$ or $c = b$)

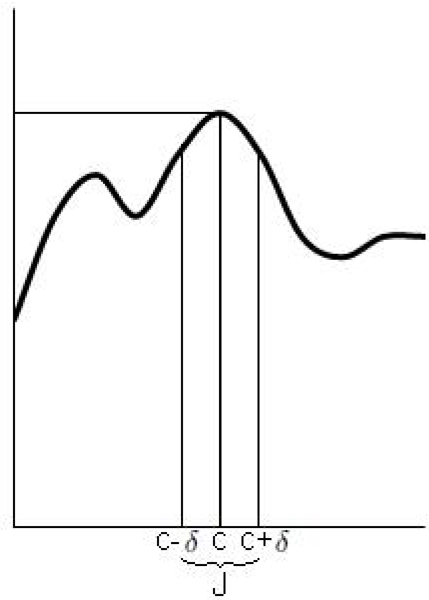

Figure 3: Interval around a point

$$
\begin{aligned}
\int_a^b f(x)\,dx &= \int_J f(x)\,dx + \int_{[a,b]\setminus J} f(x)\,dx \\
&\geq \int_J f(x)\,dx \\
&\geq \text{length } J.\frac{f(x)}{2} \\
&0
\end{aligned}
$$

□

Exercise 13.5. Let $a \leq c \leq b$. Give an example of a Riemann integrable function $f : [a,b] \to \mathbb{R}$ such that $f(t) \geq 0$ for all $t \in [a,b]$ and

$$\int_a^b f(t)\,dt = 0$$

but $f(c) \neq 0$.

Proof. $f(c) = 1, f(t) = 0$ otherwise will do □

We now come to the justly named fundamental theorem of the calculus.

Theorem 13.6 (The fundamental theorem of the calculus)**.** Suppose that $f : (a,b) \to \mathbb{R}$ is a continuous function and that $u \in (a,b)$. If we set

$$F(t) = \int_u^t f(x)\,dx$$

then F is differentiable on (a,b) and $F'(t) = f(t)$ for all $t \in (a,b)$.

Proof.

$$\begin{aligned}
|\frac{F(t+h)-F(t)}{h} - f(t)| &= |\frac{1}{h}(\int_u^{t+h} f(x)\,dx - \int_u^t f(x)\,dx) - f(t)| \\
&= |\frac{1}{h}\int_t^{t+h} f(x)\,dx - f(t)| \\
&= |\frac{1}{h}\int_t^{t+h} (f(u)-f(t))\,du| \quad (\int_t^{t+h} f(t)\,du = hf(t) \because f(t) \text{ constant wrt } u) \\
&\leq \frac{1}{|h|} \text{ length . sup} \\
&= \frac{1}{|h|}|h| \sup_{|u-t|\leq|h|} |f(u)-f(t)| \\
&= \sup_{|u-t|\leq|h|} |f(u)-f(t)| \to 0 \text{ as } h \to 0 \because f \text{ is continuous} \\
\therefore F'(t) &= f(t)
\end{aligned}$$

□

Exercise 13.7. (i) Let H be the Heaviside function $H : \mathbb{R} \to \mathbb{R}$ given by $H(x) = 0$ for $x < 0$, $H(x) = 1$ for $x \geq 0$. Calculate $F(t) = \int_0^t H(x)\,dx$ and show that F is not differentiable at 0. Where does our proof of Theorem 13.6 break down?
(ii) Let $f(0) = 1$, $f(t) = 0$ otherwise. Calculate $F(t) = \int_0^t f(x)\,dx$ and show that F is differentiable at 0 but $F'(0) \neq f(0)$. Where does our proof of Theorem 13.6 break down?

Sometimes we think of the fundamental theorem in a slightly different way.

Theorem 13.8. Suppose that $f : (a,b) \to \mathbb{R}$ is continuous, that $u \in (a,b)$ and $c \in \mathbb{R}$. Then there is a unique solution to the differential equation $g'(t) = f(t)$ $[t \in (a,b)]$ such that $g(u) = c$.

Proof. First, prove existence
Let $g(t) = \int_u^t f(x)\,dx$
By the fundamental theorem of the calculus $g'(t) = f(t)$
Observe $g(u) = 0 + c = c$
Next, show uniqueness

If $h'(t) = g'(t) = f(t)$ with $h(u) = g(u) = c$ set $a(t) = h(t) - g(t)$, and thus $a'(t) = h'(t) - g'(t) = f(t) - f(t) = 0\ \forall t$, so by the Mean Value Theorem $a(t)$ is constant
$\therefore a(u) = h(u) - g(u) = c - c = 0$ so $a(t) = 0\ \forall t$ so $h(t) = g(t)$ $\square$

We call the solutions of $g'(t) = f(t)$ *indefinite integrals* (or, simply, *integrals*) of f.
Yet another version of the fundamental theorem is given by the next theorem.

Theorem 13.9. Suppose that $g : [a, b] \to \mathbb{R}$ has continuous derivative. Then

$$\int_a^b g'(t)\, dt = g(b) - g(a)$$

Proof. Set $G(s) = \int_a^s g'(t)/.dt - g(s)$
$G'(s) = g'(s) - g'(s)$ so by the Mean Value Theorem G is constant, so $G(b) = G(a)$ i.e. $\int_a^b g'(t)\, dt - g(b) = -g(a)$ so $\int_a^b g'(t)\, dt = g(b) - g(a)$ $\square$

Theorems 13.6 and 13.9 show that (under appropriate circumstances) integration and differentiation are inverse operations and the the theories of differentiation and integration are subsumed in the greater theory of the calculus.
We use the fundamental theorem of the calculus to prove the formulae for integration by substitution and integration by parts.

Theorem 13.10 (Change of variables for integrals)**.** Suppose that $f : [a, b] \to \mathbb{R}$ is continuous and $g : [\gamma, \delta] \to \mathbb{R}$ is differentiable with continuous derivative. Suppose further that $g([\gamma, \delta]) \subseteq [a, b]$. Then, if $c, d \in [\gamma, \delta]$, we have

$$\int_{g(c)}^{g(d)} f(s)\, ds = \int_c^d f(g(x))g'(x)\, dx$$

Proof. Rearrange to $\int_c^d f(g(x))g'(x)\, dx - \int_{g(c)}^{g(t)} f(s)\, ds = 0$, and replace by the more general problem

$$\int_c^t f(g(x))g'(x)\, dx - \int_{g(c)}^{g(t)} f(s)\, ds = 0 \quad \forall c \le t \le d$$

Proof begins:

$$\frac{d}{dt}\left(\int_c^t f(g(x))g'(x)\, dx - \int_{g(c)}^{g(t)} f(s)\, ds\right) = f(g(t))g'(t) - g'(t)f(g(t)) = 0$$

(by the chain rule, $\frac{d}{dt} f(g(t)) = g'(t)f'(g(t))$)
so $\int_c^t f(g(x))g'(x)\, dx - \int_{g(c)}^{g(t)} f(s)\, ds$ is constant (MVT)
Taking $t = c$ we see that the constant is zero, so

$$\int_c^t f(g(x))g'(x)\, dx - \int_{g(c)}^{g(t)} f(s)\, ds = 0$$

$\square$

Exercise 13.11. The following exercise is traditional.
(i) Show that integration by substitution, using $x = 1/t$, gives

$$\int_a^b \frac{dx}{1+x^2} = \int_{1/b}^{1/a} \frac{dt}{1+t^2}$$

when $b > a > 0$.
(ii) If we set $a = -1$, $b = 1$ in the formula of (i), we obtain

$$\int_{-1}^1 \frac{dx}{1+x^2} \stackrel{?}{=} -\int_{-1}^1 \frac{dt}{1+t^2}$$

Explain this apparent failure of the method of integration by substitution.
(iii) Write the result of (i) in terms of $\tan^{-1}$ and prove it using standard trigonometric identities.

Lemma 13.12. Suppose that $f : [a,b] \to \mathbb{R}$ has continuous derivative and $g : [a,b] \to \mathbb{R}$ is continuous. Let $G : [a,b] \to \mathbb{R}$ be an indefinite integral of g. Then, we have

$$\int_a^b f(x)g(x)\,dx = [f(x)G(x)]_a^b - \int_a^b f'(x)G(x)\,dx$$

Proof.

$$\begin{aligned}\frac{d}{dt}(&\int_a^t f(x)g(x)\,dx - (f(t)G(t) - f(a)G(a)) + \int_a^t f'(x)G(x))\ dx \\ &= f(t)g(t) - (f'(t)G(t) + f(t)G'(t)) + f'(t)G(t) = 0\end{aligned}$$

So by the MVT $\int_a^t f(x)g(x)\,dx - (f(t)G(t) - f(a)G(a)) + \int_a^t f'(x)G(x)$ is constant. Setting $t = a$ we see that the constant is zero, and the result follows ☐

We obtain the following version of Taylor's theorem by repeated integration by parts.

Theorem 13.13 (A global Taylor's theorem with integral remainder)**.** If $n \geq 1$ and $f : (-a,a) \to \mathbb{R}$ is n times differentiable with continuous nth derivative, then

$$f(t) = \sum_{j=0}^{n-1} \frac{f^{(j)}(0)}{j!} t^j + R_n(f,t)$$

for all $|t| < a$, where

$$R_n(f,t) = \frac{1}{(n-1)!} \int_0^t (t-x)^{n-1} f^{(n)}(x)\,dx$$

Proof. $R_n = \frac{1}{(n-1)!}\int_0^t (t-x)^{n-1} f^{(n)}(x)\,dx$
Integrate by parts:

$$\begin{aligned}
R_n &= \left[\frac{(t-x)^{n-1}}{(n-1)!} f^{(n-1)}(x)\right]_0^t + \frac{1}{(n-2)!}\int_0^t (t-x)^{n-1} f^{(n-1)}(x)\,dx \\
&= \frac{-f^{(n-1)}(0)t^{n-1}}{(n-1)!} + R_{n-1} \\
&= -\frac{f^{(n-1)}(0)t^{n-1}}{(n-1)!} - \frac{f^{(n-2)}(0)t^{n-2}}{(n-2)!} + R_{n-2} \\
&= -\frac{f^{(n-1)}(0)t^{n-1}}{(n-1)!} - \cdots - f'(0)t1! + \underbrace{\int_0^t f'(x)\,dx}_{R_1} \\
&= -\frac{f^{(n-1)}(0)t^{n-1}}{(n-1)!} - \cdots - f'(0)t1! + f(t) - f(0)
\end{aligned}$$

Rearranging:

$$f(t) = f(0) + \frac{f'(0)t}{1!} + \cdots + \frac{f^{(n-1)}(0)t^{n-1}}{(n-1)!} + R_n(f,t)$$

□

In the opinion of the lecturer this form is powerful enough for most purposes and is a form that is easily remembered and proved for examination.

14 Infinite integrals

The reader may already be familiar with definitions of the following type.

Definition 14.1. If $f : [a,b] \to \mathbb{R}$ and $M, P \geq 0$ let us write

$$f_{M,P}(x) = \begin{cases} f(x) & \text{if } -P \leq f(x) \leq M \\ M & \text{if } f(x) > M \\ -P & \text{if } f(x) < -P. \end{cases}$$

If $f_{M,P}$ is Riemann integrable for each $M, P \geq 0$ and

$$\int_a^b f_{M,P}(x)\,dx \to L$$

as $M, P \to \infty$ then we say that f is Riemann integrable and

$$\int_a^b f(x)\,dx = L$$

Definition 14.2. If $f : [a, \infty) \to \mathbb{R}$ is such that $f|_{[a,X]}$ is Riemann integrable for each $X > a$ and $\int_a^X f(x)\,dx \to L$ as $X \to \infty$, then we say that $\int_a^\infty f(x)\,dx$ exists with value L.

It must be said that neither Definition 14.1 nor Definition 14.2 are more than ad hoc.
In the rest of this section we look at Definition 14.2.

Lemma 14.3. Suppose $f : [a, \infty) \to \mathbb{R}$ is such that $f|_{[a,X]}$ is Riemann integrable on $[a, X]$ for each $X > a$. If $f(x) \geq 0$ for all x, then $\int_a^\infty f(x)\,dx$ exists if and only if there exists a K such that $\int_a^X f(x)\,dx \leq K$ for all X.

Proof. Set $u_n = \int_a^n f(x)\,dx$, then $u_{n+1} - u_n = \int_n^{n+1} f(x)\,dx \geq 0$ (as $f \geq 0$) so u_n is an increasing sequence bounded above by the hypothesis, so u_n tends to a limit, L say.
Thus given $\epsilon > 0$ $\exists N(\epsilon)$ such that $L \geq u_n \geq L - \epsilon$ $\forall n \geq N(\epsilon)$ If $X \geq N(\epsilon)$ we can choose m so that $m+1 \geq X \geq m$, with $u_{m+1} \geq u_X \geq u_m$ so $L \geq u_X \geq L - \epsilon$ and $u_X \to L$ as $X \to \infty$ $\square$

We use Lemma 14.3 to prove the integral comparison test.

Lemma 14.4. If $f : [1, \infty) \to \mathbb{R}$ is a decreasing function with $f(x) \to 0$ as $x \to \infty$ then

$$\sum_{n=1}^{\infty} f(n) \text{ and } \int_1^\infty f(x)\,dx$$

either both diverge or both converge.

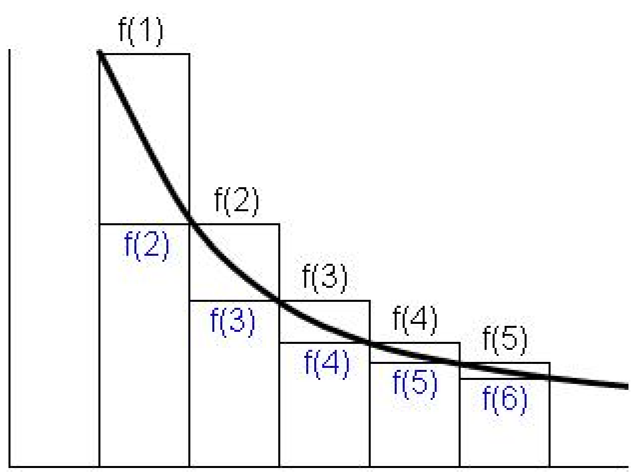

Figure 4: Integration of a decreasing function

Proof. $f(n-1) \geq f(x) \geq f(n)$ for $n \geq x \geq (n-1)$

$$\int_{n-1}^{n} f(n-1)\,dx \geq \int_{n-1}^{n} f(x)\,dx \geq \int_{n-1}^{n} f(n)\,dx$$

i.e. $f(n-1) \geq \int_{n-1}^{n} f(x)\,dx \geq f(n)$. Sum from $n = 2$ to $N+1$

$$\sum_{n=2}^{N+1} f(n-1) \geq \sum_{n=2}^{N+1} \int_{n-1}^{n} f(x)\,dx \geq \sum_{n=2}^{N+1} f(n)$$

$$\therefore \sum_{1}^{N} f(n) \geq \int_{1}^{N+1} f(x)\,dx \geq \sum_{2}^{N+1} f(n)$$

If $\sum f(n)$ converges then $\sum_1^N f(n)$ is bounded

$$\sum_{1}^{N} f(n) \leq A \;\forall N$$

So $A \geq \int_1^{N+1} f(x)\,dx \;\forall N \geq 1$
so since f is positive, $A \geq \int_1^X f(x)\,dx \;\forall X \in \mathbb{R}, X \geq 1$, so $\int_1^\infty f(x)\,dx$ converges.
Conversely, if $\int_1^\infty f(x)\,dx$ converges $\exists B$ such that $B \geq \int_1^X f(x)\,dx \;\forall X$ so $B \geq \sum_2^{N+1} f(n) \;\forall N$ such that $B + f(1) \geq \sum_1^M f(r) \;\forall M$ so $\sum f(n)$ converges □

Example 14.5. If $\alpha > 1$ then $\sum_{n=1}^\infty n^{-\alpha}$ converges. If $\alpha \leq 1$ then $\sum_{n=1}^\infty n^{-\alpha}$ diverges.

Proof.

$$\begin{aligned}\int_1^X x^{-\alpha}\,dx &= \left[\frac{x^{1-\alpha}}{1-\alpha}\right]_1^X \qquad \alpha \neq 1\\ &= \frac{X^{1-\alpha}-1}{1-\alpha}\end{aligned}$$

Which tends to ∞ if $\alpha < 1$ and converges to $\frac{1}{1-\alpha}$ if $\alpha > 1$
Thus $\sum n^{-\alpha}$ diverges for $\alpha < 1$ and converges for $\alpha > 1$
($x^{-1\alpha}$ is decreasing for $\alpha > 0$, $\alpha \leq 0$ is trivial, $x^{-\alpha} \not\to 0$)
In the case $\alpha = 1$, $\frac{1}{x}$ is a decreasing positive function, and $\int_1^X \frac{1}{x}\,dx = \log X \to \infty$ so $\sum \frac{1}{n}$ diverges □

Along the same lines

$$\int_\xi^X \frac{1}{x(\log x)^\alpha}\,dx = \left[\frac{(\log x)^{1-\alpha}}{1-\alpha}\right]_\xi^X$$

which converges if $\alpha > 1$ and diverges if $\alpha < 1$

$$\int_\xi^X \frac{1}{x\log x}\,dx = [\log(\log x)]_\xi^X \to \infty \text{ as } X \to \infty$$

$\sum_{n=\xi}^{\infty} \frac{1}{X(\log X)^\beta}$ converges if $\beta > 1$ and diverges if $\beta \leq 1$
This is really as far as we need to go, but I will just add one further remark.

Lemma 14.6. Suppose $f : [a, \infty) \to \mathbb{R}$ is such that $f|_{[a,X]}$ is Riemann integrable on $[a, X]$ for each $X > a$. If $\int_a^\infty |f(x)|\,dx$ exists, then $\int_a^\infty f(x)\,dx$ exists.

Proof. Let

$$\begin{aligned} f_+(x) &= f(x) \text{ when } f(x) \geq 0 \\ &= 0 \text{ when } f(x) \leq 0 f_-(x) \qquad &= -f(x) \text{ when } f(x) \leq 0 \\ &= 0 \text{ when } f(x) \geq 0 \end{aligned}$$

f_+, f_- are locally integrable with $|f| \geq f_+ \geq 0$ so $\int_a^x |f| \geq \int_a^x f_+ \geq 0$
But $\int_a^\infty |f|$ converges such that $\int_a^X |f| \leq A\ \forall X$ for some A, so $\int_a^X f_+ \leq A\ \forall X$
i.e. $\int_a^\infty f_+(x)\,dx$ tends to a limit $\int_a^\infty f_+(x)\,dx$
$\int_a^\infty f_-$ exists for the same reason, and

$$\int_a^X f = \int_a^X f_+ - f_- \to \int_a^\infty f_+ - \int_a^\infty f_-$$

and thus $\int_a^\infty f$ exists $\square$

It is natural to state Lemma 14.6 in the form 'absolute convergence of the integral implies convergence'.
Speaking broadly, infinite integrals $\int_a^\infty f(x)\,dx$ work well when they are absolutely convergent, that is to say, $\int_a^\infty |f(x)|\,dx < \infty$, but are full of traps for the unwary otherwise. This is not a weakness of the Riemann integral but inherent in any mathematical situation where an object only exists 'by virtue of the cancellation of two infinite objects'. (Question 17 on the 4th example sheet gives an example of an integral which is convergent but not absolutely convergent.)
(The final lemma and example were given in lectures, but did not form part of the printed notes)

Lemma 14.7. If f, g are locally Riemann-integrable and $f \geq g \geq 0$

$$\int_a^\infty f \text{ converges } \to \int_a^\infty g \text{ converges}$$

Proof. Similar to the proof of the preceding lemma. As $f \geq g \geq 0, \int_a^X f(x) \geq \int_a^X g(x) \geq 0$
Thus $\int_a^\infty f$ converges $\Rightarrow \int_a^X f$ is bounded $\Rightarrow \int_a^X g$ is bounded and thus $\int_a^\infty g$ exists $\square$

Example 14.8.

$$\int_{-\infty}^{+\infty} \exp\left(-\frac{x^2}{2}\right) \text{ exists}$$

Proof. $\exp\left(-\frac{x^2}{2}\right) \leq \exp\left(-\frac{x}{2}\right)$ for $x \geq 1$

$$\int_1^X \exp\left(-\frac{x}{2}\right) = \frac{1}{2}(e^{-\frac{1}{2}} - e^{-\frac{x}{2}}) \to \frac{1}{2}e^{-\frac{1}{2}}$$

so $\int_1^\infty \exp\left(-\frac{x}{2}\right)$ exists, so $\int_1^\infty \exp\left(-\frac{x^2}{2}\right)$ exists, so $\int_{-\infty}^{-1} \exp\left(-\frac{x^2}{2}\right)$ exists, and thus $\int_{-\infty}^{+\infty} \exp\left(-\frac{x^2}{2}\right)$ exists □

However, we note that there is no way of writing $\int_{-\infty}^{X} \exp\left(-\frac{x^2}{2}\right) dx$ using $\log, \exp$, polynomials or trig functions.

15 Further reading

The two excellent books Spivak's *Calculus* [5] and J. C. Burkill's *A First Course in Mathematical Analysis* [2] both cover the course completely and should be in your college library[6]. Burkill's book is more condensed and Spivak's more leisurely. A completely unsuitable but interesting version of the standard analysis course is given by Berlinski's *A Tour of the Calculus* [1] — Spivak rewritten by Sterne with additional purple passages by the Ankh-Morpork tourist board. I have written *A Companion to Analysis* [4] which covers this course at a higher level together with the next analysis course. It is available off the web but is unlikely to be as suitable for beginners as Spivak and Burkill. If you do download it, remember that you are under a moral obligation to send me an e-mail about any mistakes you find.

References

[1] D. Berlinski *A Tour of the Calculus* Mandarin Paperbacks 1997.

[2] J. C. Burkill *A First Course in Mathematical Analysis* CUP, 1962.

[3] R. P. Burn *Numbers and Functions* CUP, 1992.

[4] T. W. Körner *A Companion to Analysis* for the moment available via my home page http://www.dpmms.cam.ac.uk/~twk/ .

[5] M. Spivak, *Calculus* Addison-Wesley/Benjamin-Cummings, 1967.

[6] A quieter version of the JCR with shelves of books replacing the bar.