MAS221 Analysis, Semester 2, 2017-18

Sarah Whitehouse

Contents

Welcome to Semester 2 of Analysis! This semester we will continue the study of analysis of functions of one real variable. We will start by studying differentiation and later move on to integration.

Revision

Naturally it is important to have a good grasp of the first semester material. Please have a go at the (not for credit) online revision test as a good way to revise key concepts. In particular, key to everything is the definition of convergence of a sequence (a_n) to a limit l (Definition 2.1 from Semester 1). We revise this here.

Let's run through how we end up with the formal definition of convergence. Informally, we might say that *the terms of the sequence should get closer and closer to l*. This is certainly too vague! For example, the terms of the sequence $(1 + \frac{1}{n})$ get closer and closer to 0, but we can see that the limit is 1 not 0.

So we realize that we need to say something like *the terms of the sequence should get arbitrarily close to l*. The formal definition is designed to capture this idea.

Definition A sequence (a_n) is said to *converge to a limit* $l \in \mathbb{R}$ if given any $\varepsilon > 0$, there exists $N \in \mathbb{N}$ such that for all $n > N$, we have $|a_n - l| < \varepsilon$.

In words: no matter what distance ε is specified, there is an index N beyond which all the terms $a_{N+1}, a_{N+2}, \ldots$ have a distance smaller than ε to l.

We will now go over an example in detail, first guessing the limit by experimenting with particular values and special cases and then giving a detailed proof, using the algebra of limits and the sandwich rule.

Example Let $a, b \in \mathbb{R}$ and consider $\lim_{n\to\infty}(|a|^n + |b|^n)^{1/n}$.

1. Does this limit exist? What do you think it is? If you don't know, try experimenting with some particular values of a and b or try considering special cases (eg $a = 0$ or $a = b$).

2. Prove that your guess is correct.

Solution.

1. We begin by trying to guess what is going on, following the hint.

   If $a = 0$, then we get

   $$\lim_{n\to\infty}(|b|^n)^{1/n} = \lim_{n\to\infty}(|b|) = |b|.$$

   And if $a = b$,

   $$\lim_{n\to\infty}(|a|^n + |b|^n)^{1/n} = \lim_{n\to\infty}(2|b|^n)^{1/n} = \lim_{n\to\infty} 2^{1/n}|b| = \lim_{n\to\infty} 2^{1/n} \lim_{n\to\infty} |b| = |b|,$$

   where we have used the algebra of limits.

For a numerical example, let's take $a = 1066$ and $b = 2018$. Maple tells me that some terms of the sequence are approximately

$$3084, 2282.253273, 2112.645195, 2056.185482, 2034.334311, 2025.242498, \ldots,$$
$$2018.009362, , \ldots, 2018.000289.$$

So it looks like this may be converging to 2018.

Looking at our examples, we are led to guess that the sequence does converge, to $\max\{|a|, |b|\}$.

We can also come up with a rough idea about why this should be right: if $|a| < |b|$, then for large n, $|a|^n$ will be insignificant compared to $|b|^n$, so the terms will behave like $(|b|^n)^{1/n} = |b|$. Now we need to prove that this is actually correct.

2. We will prove that $\lim_{n\to\infty}(|a|^n + |b|^n)^{1/n} = \max\{|a|, |b|\}$.

   Let $M = \max\{|a|, |b|\}$. Then $|a| \leq M$ and $|b| \leq M$, so $|a|^n \leq M^n$ and $|b|^n \leq M^n$. Thus $|a|^n + |b|^n \leq 2M^n$. So

$$(|a|^n + |b|^n)^{1/n} \leq 2^{1/n}M.$$

   Also, $|a|^n + |b|^n \geq M^n$, so $(|a|^n + |b|^n)^{1/n} \geq M$. Thus

$$M \leq (|a|^n + |b|^n)^{1/n} \leq 2^{1/n}M.$$

   Using the algebra of limits,

$$\lim_{n\to\infty} 2^{1/n}M = M \lim_{n\to\infty} 2^{1/n} = M.1 = M.$$

   So it follows using the sandwich rule that $\lim_{n\to\infty}(|a|^n + |b|^n)^{1/n} = M$, as required.

5 Differentiation

5.1 Introduction

The processes of *differentiation* and *integration* constitute the two corner–stones of the *calculus* which revolutionised mathematics (and its applications), starting from the groundbreaking work of Newton and Leibniz in the seventeenth century. In this chapter we will focus on understanding differentiation from a rigorous analytic viewpoint, using the knowledge that we have gained about limits in semester 1. Before we start this process let us remind ourselves what differentiation is for.

The geometric motivation for differentiation is to find the *slope* or *gradient* of the tangent to a curve at a point lying on it. If the curve is given by a formula

$y = f(x)$, then the gradient of the tangent at the point (x, y) appears to be well–approximated by the slope of a chord connecting the very nearby points (x, y) and $(x + \Delta x, y + \Delta y)$. This slope is given by the ratio:

$$\frac{\Delta y}{\Delta x} = \frac{f(x + \Delta x) - f(x)}{\Delta x}.$$

But to get from the slope of the chord to the slope of the tangent, it seems that we must put $\Delta x = 0$. This cannot be taken literally as it gives us the meaningless ratio $0/0$.

The dynamic motivation for differentiation is to find the *instantaneous rate of change* of one quantity with respect to another. Let us assume that we are dealing with a physical quantity $F(t)$ that changes as a function of time t. For example $F(t)$ could be position of a moving particle at time t, in which case the required rate of change is the velocity. Or $F(t)$ could be the charge on a conductor at time t, in which case the rate of change is the current. Then over a very small time interval Δt, the average rate of change is:

$$\frac{\Delta F(t)}{\Delta t} = \frac{F(t + \Delta t) - F(t)}{\Delta t},$$

and we again want to understand what happens when $\Delta t = 0$.

Of course, we now know that we must solve both of these problems by taking a limit as Δx, or Δt, tends to zero.

5.2 Differentiation as a limit

Definition 5.2.1 Let f be a real-valued function with domain D_f. We say that f is *differentiable* at $a \in D_f$ if $\lim_{x \to a} \dfrac{f(x) - f(a)}{x - a}$ exists (and is finite). In this case we write

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

and we call $f'(a) \in \mathbb{R}$, the *derivative of f at a*. We say that f is *differentiable* on $S \subseteq D_f$ if it is differentiable at every point $a \in S$. Then the mapping $a \mapsto f'(a)$ defines a real-valued function which is called the *derivative* of f, and denoted by f'. This has domain

$$D_{f'} = \{x \in D_f \mid f'(x) \text{ exists}\}.$$

In applied mathematics, we may often write $y = f(x)$, and write the function f' as dy/dx. Then $f'(a) = \frac{dy}{dx}\Big|_{x=a}$. When we do analysis, we do not find the dy/dx notation so helpful; it is usually easier to work with f'.

We can of course, iterate the notion of differentiability in the usual way. Suppose that $a \in D_{f'}$ and that f' is differentiable at a, then we define the second derivative $f''(a)$ of f at a by

$$f''(a) = (f')'(a).$$

More generally if $n \in \mathbb{N}$ with $n > 2$ we define the nth derivative of f at a by

$$f^{(n)}(a) = (f^{(n-1)})'(a),$$

whenever the limit on the right-hand side exists. We say that f is *infinitely differentiable* or *smooth* at a if $f^{(n)}(a)$ exists (and is finite) for all $n \in \mathbb{N}$. It is also sometimes useful (especially when considering Taylor series, see section 5.6), to employ the notation $f(a) = f^{(0)}(a)$.

Example 5.2.2 If $f(x) = c$ where c in $\mathbb{R}$ is constant, it is easy to check directly from Definition 5.2.1 that $D_{f'} = D_f = \mathbb{R}$, and $f'(a) = 0$ for all $a \in \mathbb{R}$.

Example 5.2.3 Let $f(x) = x^n$ for $x \in \mathbb{R}$, where $n \in \mathbb{N}$ is fixed. Use the Binomial Theorem to show directly from the definition that f is differentiable at all $x \in \mathbb{R}$, with $f'(x) = nx^{n-1}$.

Solution. Let $a \in \mathbb{R}$. Using the Binomial Theorem, we have

$$\begin{aligned}
&\frac{f(a+h) - f(a)}{h} \\
&\quad = \frac{(a+h)^n - a^n}{h} \\
&\quad = \frac{a^n + na^{n-1}h + \frac{1}{2}n(n-1)a^{n-2}h^2 + \cdots + nah^{n-1} + h^n - a^n}{h} \\
&\quad = na^{n-1} + \frac{1}{2}n(n-1)a^{n-2}h + \cdots + nah^{n-2} + h^{n-1}.
\end{aligned}$$

Thus

$$\lim_{h \to 0} \frac{f(a+h) - f(a)}{h} = na^{n-1},$$

so $f'(x) = nx^{n-1}$ for all $x \in \mathbb{R}$, and $D_{f'} = D_f = \mathbb{R}$.

Next we turn our attention to the relationship between differentiability and continuity.

Theorem 5.2.4 *If a real-valued function f is differentiable at $a \in D_f$ then f is continuous at a.*

Proof. We need to show that $\lim_{x \to a} f(x) = f(a)$. For $x \neq a$, write

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

Since f is differentiable at a, $\lim_{x \to a} \dfrac{f(x) - f(a)}{x - a} = f'(a)$, and of course $\lim_{x \to a}(x - a) = 0$. Hence by algebra of limits (Semester 1, Theorem 3.3.1),

$\lim_{x\to a} \dfrac{f(x)-f(a)}{x-a}.(x-a) = f'(a)\cdot 0 = 0$. So $\lim_{x\to a}(f(x)-f(a))$ exists and equals zero, and the result follows. $\square$

On the other hand, it is not true that every function that is continuous at a is differentiable at a.

Example 5.2.5 Consider the function $f(x) = |x|$ with $D_f = \mathbb{R}$. It is continuous at every point in $\mathbb{R}$. It is also easy to see that it is differentiable at every $x \neq 0$. Show that it is not differentiable at zero, by showing that the relevant left and right limits are different there.

Solution. We have

$$\lim_{h\uparrow 0} \frac{f(0+h)-f(0)}{h} = \lim_{h\uparrow 0} \frac{|h|}{h} = \lim_{h\uparrow 0} \frac{-h}{h} = -1.$$
$$\lim_{h\downarrow 0} \frac{f(0+h)-f(0)}{h} = \lim_{h\downarrow 0} \frac{|h|}{h} = \lim_{h\downarrow 0} \frac{h}{h} = 1.$$

Thus the left and right limits are different and so $\lim_{h\to 0} \dfrac{f(0+h)-f(0)}{h}$ does not exist.

Generalising the last example, we have the following definition.

Definition 5.2.6 We say that the real-valued function f has a *left derivative* at $a \in D_f$ if $f'_-(a) = \lim_{h\uparrow 0} \dfrac{f(a+h)-f(a)}{h}$ exists (and is finite), and that it has a *right derivative* at $a \in D_f$ if $f'_+(a) = \lim_{h\downarrow 0} \dfrac{f(a+h)-f(a)}{h}$ exists (and is finite).

Theorem 5.2.7 *A real-valued function f is differentiable at $a \in D_f$ if and only if both the left and right derivatives at a exist and are equal. In this case*

$$f'(a) = f'_-(a) = f'_+(a).$$

Giving the proof is an exercise on the problem sheets.

5.3 Rules for differentiation

The results in this section should all be familiar from MAS110, but now we can make the proofs rigorous.

Theorem 5.3.1 *Let f and g be real-valued functions that are differentiable at $a \in D_f \cap D_g$. Then the following hold.*

1. *For each $\alpha, \beta \in \mathbb{R}$, the function $\alpha f + \beta g$ is differentiable at a and*
$$(\alpha f + \beta g)'(a) = \alpha f'(a) + \beta g'(a).$$
2. *(The Product Rule) The function fg is differentiable at a and*
$$(fg)'(a) = f'(a)g(a) + f(a)g'(a).$$
3. *(The Quotient Rule) If $g(a) \neq 0$ then f/g is differentiable at a and*
$$\left(\frac{f}{g}\right)'(a) = \frac{g(a)f'(a) - f(a)g'(a)}{g(a)^2}.$$

Proof.

1. This is an easy application of the algebra of limits.
2. For all $h \neq 0$,
$$\begin{aligned} & \frac{(fg)(a+h) - (fg)(a)}{h} \\ = & \frac{f(a+h)g(a+h) - f(a)g(a+h)}{h} + \frac{f(a)g(a+h) - f(a)g(a)}{h} \\ = & \left(\frac{f(a+h) - f(a)}{h}\right) g(a+h) + f(a)\left(\frac{g(a+h) - g(a)}{h}\right). \end{aligned}$$
The result follows by taking limits as $h \to 0$ and using Definition 5.2.1 and the algebra of limits, together with the fact that at a, g is differentiable, hence continuous by Theorem 5.2.4, and so $\lim_{h\to 0} g(a+h) = g(a)$.
3. First observe that by Problem 61, there exists $\delta > 0$ so that $g(x) \neq 0$ for all $x \in (a - \delta, a + \delta)$. Consider $h \in \mathbb{R}$ such that $|h| < \delta$. Then
$$\begin{aligned} & \frac{1}{h}\left(\left(\frac{f}{g}\right)(a+h) - \left(\frac{f}{g}\right)(a)\right) \\ = & \frac{1}{h}\left(\frac{f(a+h)g(a) - f(a)g(a+h)}{g(a)g(a+h)}\right) \\ = & \frac{1}{g(a)g(a+h)}\left(\frac{f(a+h) - f(a)}{h}g(a) - f(a)\frac{g(a+h) - g(a)}{h}\right), \end{aligned}$$
and the result follows by algebra of limits, using the fact that (as above), $\lim_{h\to 0} g(a+h) = g(a)$.

□

Theorem 5.3.2 (Chain rule) *Let f, g be real-valued functions such that the range of g is contained in the domain of f. Suppose that g is differentiable at a and that f is differentiable at $g(a)$. Then $f \circ g$ is differentiable at a and*

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

Proof. For $x \in D_g$, we consider

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

We would like to write this as

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

but that only makes sense when $g(x) \neq g(a)$. To overcome this problem, we introduce the function $Q : D_f \to \mathbb{R}$ defined by

$$Q(y) = \begin{cases} \frac{f(y) - f(g(a))}{y - g(a)} & \text{if } y \neq g(a), \\ f'(g(a)) & \text{if } y = g(a). \end{cases}$$

We claim that

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

for all $x \in D_g$. Indeed, for x such that $g(x) \neq g(a)$, this is clear because the factors of $g(x) - g(a)$ cancel. And for x such that $g(x) = g(a)$, both sides are zero.

So we need to study

$$\lim_{x \to a} Q(g(x)) \frac{g(x) - g(a)}{x - a}.$$

Since g is differentiable at a, by Theorem 5.2.4, g is continuous at a. Since f is differentiable at $g(a)$, Q is continuous at $g(a)$. So $Q \circ g$ is continuous at a and $\lim_{x \to a} Q(g(x))$ exists and is equal to $Q(g(a)) = f'(g(a))$. Using the algebra of limits, we then have

$$\begin{aligned} \lim_{x \to a} Q(g(x)) \frac{g(x) - g(a)}{x - a} &= \lim_{x \to a} Q(g(x)) \lim_{x \to a} \frac{g(x) - g(a)}{x - a} \\ &= f'(g(a))g'(a), \end{aligned}$$

as required. □

5.4 Turning points and Rolle's theorem

Definition 5.4.1 A real-valued function f has a *local minimum* at $a \in D_f$ if there exists $\delta > 0$ such that $(a-\delta, a+\delta) \subset D_f$ and $f(x) \geq f(a)$ for all $x \in (a-\delta, a+\delta)$.

A real-valued function f has a *local maximum* at $a \in D_f$ if there exists $\delta > 0$ such that $(a-\delta, a+\delta) \subset D_f$ and $f(x) \leq f(a)$ for all $x \in (a-\delta, a+\delta)$.

A *turning point* (sometimes called an *extreme point*) for f is a point in its domain that is either a local minimum or a local maximum.

It is important to distinguish carefully between local maxima and minima and global maxima and minima, when the latter exist. For example if $f : [a,b] \to \mathbb{R}$ is continuous, then we know by Theorem 4.3.4 from semester 1 that it attains both its supremum and infimum on $[a,b]$. So these are the global maximum and minimum, respectively. But they are not necessarily turning points, because they might be at the endpoints of the interval $[a,b]$. And, of course, a local minimum/maximum need not be a global one.

Example 5.4.2 Consider the function $f : [-3,2] \to \mathbb{R}$ defined by

$$f(x) = \begin{cases} x+2 & \text{if } x \in [-3,-1), \\ x^2 & \text{if } x \in [-1,2]. \end{cases}$$

Identify any turning points and the global maximum and minimum.

Solution. Note that the function is continuous (by checking what happens at $x=-1$), so as above there is a global maximum and minimum. The global maximum is 4, attained at $x=2$. The global minimum is -1, attained at $x=-3$. Neither of these is a turning point (because they are the endpoints of the interval: we cannot find a δ such that all points within distance δ are contained in the domain). There is a local minimum at $x=0$.

Theorem 5.4.3 *If f is differentiable at $a \in D_f$ and a is a turning point for f, then $f'(a)=0$.*

Proof. We suppose that f has a local minimum at a. So there exists $\delta > 0$ so that $(a-\delta, a+\delta) \subset D_f$ and $f(x) \geq f(a)$ for all $x \in (a-\delta, a+\delta)$. Hence

$$\frac{f(x)-f(a)}{x-a} \leq 0 \qquad \text{for all } x \text{ such that } a-\delta < x < a,$$
$$\frac{f(x)-f(a)}{x-a} \geq 0 \qquad \text{for all } x \text{ such that } a < x < a+\delta.$$

Taking one-sided limits as $x \to a$ and using Problem 31 from semester 1, we deduce that $f'_-(a) \leq 0$ and $f'_+(a) \geq 0$. But f is differentiable at a, so by

Theorem 5.2.7, $f'_-(a) = f'_+(a) = f'(a)$, and so we conclude that $f'(a) = 0$, as required.

The argument in the case of a local maximum is very similar. □

A point where a differentiable function has derivative zero is sometimes called a *stationary* point. So Theorem 5.4.3 says that turning points are stationary points. The converse to Theorem 5.4.3 is false: not all stationary points are turning points. Consider for example the familiar case of $f(x) = x^3$. Then $f'(0) = 0$ but 0 is neither a local maximum nor a local minimum; it is what is called an *inflection point*. We will not pursue the story of classifying stationary points further here. You have seen this before in MAS110, and it is revisited in one of the exercises. Instead we will use Theorem 5.4.3 to explore some new territory.

Theorem 5.4.4 (Rolle's theorem) *Let f be a real-valued function that is continuous on $[a, b]$ and differentiable on (a, b) with $f(a) = f(b)$. Then there exists $c \in (a, b)$ such that $f'(c) = 0$.*

Proof. If f is constant, the result is obvious, so assume that f takes at least two distinct values. By Theorem 4.3.4 from semester 1, f is bounded on $[a, b]$ and attains both its supremum and infimum. It cannot attain both of these at the end-points, as then they would be equal and f would be constant. So there must be a $c \in (a, b)$ where either the supremum or infimum is attained. But then c is a turning point, and so $f'(c) = 0$ by Theorem 5.4.3. □

5.5 Mean value theorems

The next result is a generalisation of Rolle's theorem and a very important result, with lots of interesting consequences.

Theorem 5.5.1 (Mean value theorem) *If a real-valued function f is continuous on $[a, b]$ and differentiable on (a, b), then there exists $c \in (a, b)$ such that*

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

Proof. For all $x \in [a, b]$, define $g(x) = f(x) - \alpha(x - a)$, where $\alpha = \frac{f(b)-f(a)}{b-a}$. Then g is continuous on $[a, b]$, and differentiable on (a, b), because f is. You can check easily that $g(a) = g(b) = f(a)$, and so we may apply Rolle's theorem (Theorem 5.4.4) to deduce that there exists $c \in (a, b)$ such that $g'(c) = 0$. But $g'(c) = f'(c) - \alpha$ and hence $f'(c) = \alpha$, as required. □

Towards the end of last semester, you studied monotonic functions and the inverse function theorem. We can now see some consequences of the mean value theorem in that context.

Corollary 5.5.2 *[Monotonicity revisited] Suppose that a real-valued function f is continuous on $[a, b]$ and differentiable on (a, b). If for all $x \in (a, b)$ we have*

$$\begin{aligned} f'(x) &\geq 0, \quad \textit{then } f \textit{ is monotonic increasing on } [a, b], \\ f'(x) &> 0, \quad \textit{then } f \textit{ is strictly monotonic increasing on } [a, b], \\ f'(x) &\leq 0, \quad \textit{then } f \textit{ is monotonic decreasing on } [a, b], \\ f'(x) &< 0, \quad \textit{then } f \textit{ is strictly monotonic decreasing on } [a, b]. \end{aligned}$$

Proof. We'll just do the first of these, as the others are so similar. Choose arbitrary α, β such that $a \leq \alpha < \beta \leq b$. By the mean value theorem (Theorem 5.5.1), there exists $c \in (\alpha, \beta)$ so that

$$\frac{f(\beta) - f(\alpha)}{\beta - \alpha} = f'(c) \geq 0.$$

Hence $f(\beta) \geq f(\alpha)$ and so f is monotonic increasing, as required. □

Corollary 5.5.2 becomes a powerful tool to study inverses of functions, when used in conjunction with the inverse function theorem (Semester 1, Theorem 4.3.7). Roughly speaking, that result says a strictly monotonic continuous function has an inverse of the same kind. Now we can add information about differentiability.

Theorem 5.5.3 *[Inverses revisited] Suppose that $f : \mathbb{R} \to \mathbb{R}$ is continuous on $[a, b]$ and differentiable on (a, b), and that f' is continuous at $c \in (a, b)$. If $f'(c) \neq 0$ then the following hold.*

1. *There exists $\delta > 0$ so that f is invertible on $[c - \delta, c + \delta]$, and f^{-1} is continuous on $(f(c - \delta), f(c + \delta))$ if $f'(c) > 0$, and on $(f(c + \delta), f(c - \delta))$ if $f'(c) < 0$.*
2. *The mapping f^{-1} is differentiable at $f(c)$ and*

$$(f^{-1})'(f(c)) = \frac{1}{f'(c)}.$$

Proof. We will do the case where $f'(c) < 0$; (the case where $f'(c) > 0$ is similar).

1. Since f' is continuous at c, by Problem 61 there exists $\delta > 0$ so that $f'(x) < 0$ for all $x \in (c - \delta, c + \delta)$, and we can certainly ensure (by choosing a smaller δ, if necessary), that $(c - \delta, c + \delta) \subseteq (a, b)$. By Corollary 5.5.2, f is strictly decreasing on $[c - \delta, c + \delta]$, and the result then follows from Theorem 4.3.7.

2. Let $y = f(c)$, then for $d \in (c-\delta, c) \cup (c, c+\delta)$, we have $f(d) \neq f(c)$ since f is invertible, hence injective. Write $x = f(d)$. Then

$$\begin{aligned} \frac{f^{-1}(y) - f^{-1}(x)}{y - x} &= \frac{c - d}{y - x} \\ &= \frac{1}{\frac{y-x}{c-d}} = \frac{1}{\frac{f(c)-f(d)}{c-d}}. \end{aligned}$$

Now since f^{-1} is continuous on $(f(c+\delta), f(c-\delta))$, as $x \to y$, we have $d \to c$ and so

$$(f^{-1})'(f(c)) = \lim_{x \to y} \frac{f^{-1}(y) - f^{-1}(x)}{y - x} = \lim_{d \to c} \frac{1}{\frac{f(c)-f(d)}{c-d}} = \frac{1}{f'(c)}.$$

□

Note. Theorem 5.5.3 should be familiar to you from calculus as the rule

$$\frac{dx}{dy} = \frac{1}{\frac{dy}{dx}}.$$

We have now established precise conditions under which this holds.

The next result is a useful variation on the mean value theorem theme.

Theorem 5.5.4 (Cauchy's mean value theorem) *Let f and g each be continuous on $[a, b]$ and differentiable on (a, b) with $g'(x) \neq 0$ for all $x \in (a, b)$. Then there exists $c \in (a, b)$ so that*

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

Proof. We note that, as $g'(x) \neq 0$ for all $x \in (a, b)$, we must have $g(a) - g(b) \neq 0$, by Rolle's theorem (Theorem 5.4.4). The rest of the proof follows along similar lines to that of the mean value theorem, and is left to you to do as one of the exercises. □

Corollary 5.5.5 (l'Hôpital's rule) *Suppose that f and g are each differentiable on (a, b), with $g'(x) \neq 0$ for all $x \in (a, b)$.*

1. *If $c \in (a, b)$ with $f(c) = g(c) = 0$, then*

$$\lim_{x \to c} \frac{f(x)}{g(x)} = \lim_{x \to c} \frac{f'(x)}{g'(x)},$$

whenever the limit on the right-hand side exists (and is finite).

2. *If* $\lim_{x\downarrow a} f(x) = \lim_{x\downarrow a} g(x) = 0$*, then*

$$\lim_{x\downarrow a} \frac{f(x)}{g(x)} = \lim_{x\downarrow a} \frac{f'(x)}{g'(x)},$$

whenever the limit on the right-hand side exists (and is finite).

3. *If* $\lim_{x\uparrow b} f(x) = \lim_{x\uparrow b} g(x) = 0$*, then*

$$\lim_{x\uparrow b} \frac{f(x)}{g(x)} = \lim_{x\uparrow b} \frac{f'(x)}{g'(x)},$$

whenever the limit on the right-hand side exists (and is finite).

Proof. We'll only prove (2) as the other proofs are similar. We suppose $\lim_{x\downarrow a} \frac{f'(x)}{g'(x)}$ exists and is finite and denote it by L. Let $\varepsilon > 0$. Then there exists $\delta > 0$ such that $|\frac{f'(t)}{g'(t)} - L| < \varepsilon$ whenever $|a < t < a + \delta|$.

We apply Cauchy's mean value theorem (Theorem 5.5.4) on the interval $[y, x]$ where $a < y < x < a + \delta$ to deduce that there exists $z \in (y, x)$ such that

$$\frac{f(x) - f(y)}{g(x) - g(y)} = \frac{f'(z)}{g'(z)}.$$

Thus

$$\left|\frac{f(x) - f(y)}{g(x) - g(y)} - L\right| = \left|\frac{f'(z)}{g'(z)} - L\right| < \varepsilon$$

Now $f(y) \to 0$ and $g(y) \to 0$ as $y \to c_+$. So letting $y \to c_+$, we get, for $a < x < a + \delta$,

$$\left|\frac{f(x)}{g(x)} - L\right| \leq \epsilon.$$

Thus, $\lim_{x\downarrow a} \frac{f(x)}{g(x)} = L$, as required. □

In the exercises you will see a variant of l'Hôpital's rule, where instead of converging to zero, the functions diverge to infinity at the point of interest. We'll use this in the next example (where we will assume some properties of the exponential function that will be made rigorous later on).

Example 5.5.6 Evaluate $\lim_{x\downarrow 0} x^x$. (Recall, for example from MAS110, that for $a > 0$, the function $f(x) = a^x$ may be defined as $f(x) = e^{x\ln(a)}$ for $x \in \mathbb{R}$. Similarly we may define $x^x = e^{x\ln(x)}$ for $x > 0$.)

Solution. We use the version of l'Hôpital's rule from the exercises and consider

$$\begin{aligned} \lim_{x\downarrow 0} x\ln(x) &= \lim_{x\downarrow 0} \frac{\ln(x)}{1/x} \\ &= -\lim_{x\downarrow 0} \frac{1/x}{1/x^2} = -\lim_{x\downarrow 0} x = 0. \end{aligned}$$

So by continuity of the exponential function,

$$\lim_{x\downarrow 0} x^x = \lim_{x\downarrow 0} e^{x\ln(x)} = e^{\lim_{x\downarrow 0} x\ln(x)} = e^0 = 1.$$

5.6 Taylor's theorem

Let $[a, b]$ be a given interval in $\mathbb{R}$.

Definition 5.6.1 For each $n \in \mathbb{N}$ we define the real vector space $C^n(a, b)$ to consist of functions $f : [a, b] \to \mathbb{R}$ for which

- The nth derivative $f^{(n)}$ of f exists for all points in (a, b).
- $f^{(n)}$ is continuous on (a, b).

We also define the vector space $C^\infty(a, b)$ of functions that are infinitely differentiable on (a, b).

Clearly for all $n \in \mathbb{N}$, we have inclusions

$$C^\infty(a, b) \subseteq C^n(a, b) \subseteq C^{n-1}(a, b) \subseteq \cdots \subseteq C^1(a, b) \subseteq C(a, b),$$

where $C(a, b)$ is the space of continuous functions on (a, b).

Definition 5.6.2 Let $f \in C^n(a, b)$, for some $n \in \mathbb{N}$. Fix $x_0 \in (a, b)$. The real numbers $f^{(k)}(x_0)/k!$, for $k = 0, 1, \ldots, n$ are called the *Taylor coefficients of f at x_0*.

We define a function $T_f^{(n)} \in C^n(a, b)$ by

$$T_f^{(n)}(x) = \sum_{k=0}^{n} \frac{f^{(k)}(x_0)}{k!}(x - x_0)^k.$$

The function $T_f^{(n)}$ is called *the Taylor polynomial of f of degree n around x_0*.

Theorem 5.6.3 (Taylor's theorem) *Let $f \in C^{n+1}(a, b)$ and $x_0 \in (a, b)$. Then for all $x \in (a, b)$,*

$$f(x) = \sum_{k=0}^{n} \frac{f^{(k)}(x_0)}{k!}(x - x_0)^k + R_f^{n+1}(x), \qquad (1)$$

where $R_f^{n+1}(x) = \frac{f^{(n+1)}(c)}{(n+1)!}(x - x_0)^{n+1}$, for some c depending on x, with $c \in (x_0, x)$ if $x > x_0$ and $c \in (x, x_0)$ if $x < x_0$.

Proof. Assume for convenience that $x > x_0$.

Define $M_f : (x_0, x) \to \mathbb{R}$ by

$$M_f(x) = \frac{(n+1)!}{(x-x_0)^{n+1}}[f(x) - T_f^{(n)}(x)], \tag{2}$$

and $g : [x_0, x] \to \mathbb{R}$ by

$$g(t) = -f(x) + f(t) + \sum_{k=1}^{n} \frac{f^{(k)}(t)}{k!}(x-t)^k + \frac{(x-t)^{n+1}}{(n+1)!}M_f(x).$$

Then g is clearly continuous on $[x_0, x]$ and differentiable on (x_0, x). You can easily check that $g(x_0) = g(x) = 0$. Then by Rolle's theorem (Theorem 5.4.4), there exists $c \in (x_0, x)$ with $g'(c) = 0$. Now for $t \in (x_0, x)$

$$\begin{aligned} g'(t) &= f'(t) - \sum_{k=1}^{n} \frac{f^{(k)}(t)}{(k-1)!}(x-t)^{k-1} + \sum_{k=1}^{n} \frac{f^{(k+1)}(t)}{k!}(x-t)^k - \frac{(x-t)^n}{n!}M_f(x) \\ &= \frac{(x-t)^n}{n!}(f^{(n+1)}(t) - M_f(x)). \end{aligned}$$

Then $g'(c) = 0$ tells us that c is such that $f^{(n+1)}(c) = M_f(x)$, and then (1) follows by straightforward algebra from (2). □

Notes.

1. The term $R_f^{n+1}(x) = f(x) - T_f^{(n)}(x)$ measures the error in approximating f by its Taylor polynomial of degree n at f. It is called the *remainder term of degree* $n+1$.
2. If $0 \in (a, b)$, we can take $x_0 = 0$. In this special case, Theorem 5.6.3 is called *Maclaurin's theorem*.

Now suppose that $f \in C^\infty(a, b)$ and that the series $\sum_{k=0}^{\infty} \frac{f^{(k)}(x_0)}{k!}(x-x_0)^k$ converges for all $x \in (a, b)$. If we may write

$$f(x) = \sum_{k=0}^{\infty} \frac{f^{(k)}(x_0)}{k!}(x-x_0)^k = \lim_{n\to\infty} \sum_{k=0}^{n} \frac{f^{(k)}(x_0)}{k!}(x-x_0)^k,$$

we say that f is *represented by its Taylor series* on (a, b). We will learn more about convergence of infinite series of numbers in the next chapter and infinite series of functions later on.

6 Series

Given a sequence (a_n), we can try to make sense of the sum of its terms, $\sum_{n=1}^{\infty} a_n$. We call this a series. We define such an infinite sum via a limit, when it exists. You have already studied series in MAS110, but without knowing the formal definition of limit. So now we revisit this material in a more rigorous way.

6.1 Convergence and absolute convergence

Definition 6.1.1 Given a sequence (a_n) of real numbers, we consider the associated sequence (s_n) of *partial sums*, where

$$s_n = a_1 + a_2 + \cdots + a_n.$$

We say that the sequence (a_n) is *summable* if the sequence (s_n) of partial sums converges.

In this case, we write

$$\lim_{n\to\infty} s_n = \sum_{n=1}^{\infty} a_n$$

or more informally

$$\lim_{n\to\infty} s_n = a_1 + a_2 + a_3 + \cdots$$

If the sequence (a_n) is summable, we also say the *series* $\sum_{n=1}^{\infty} a_n$ converges. This is less formal language than the above, as we are using $\sum_{n=1}^{\infty} a_n$ for both the sequence of sums, and the number it converges to (if any). However, it does not cause problems in practice.

Example 6.1.2 Consider the *geometric series*

$$\sum_{n=0}^{\infty} ar^n$$

where a and r are fixed real numbers. For which r does this converge? And what is the limit in that case?

Solution. This series converges precisely when $|r| < 1$, to the limit

$$\frac{a}{1-r}.$$

(Make sure you know why!)

Note that here the geometric series is indexed so that it begins at $n = 0$ rather than $n = 1$.

The proof of the following is left as an exercise (or see the proof given in MAS110).

Proposition 6.1.3 *Suppose the series* $\sum_{n=1}^{\infty} a_n$ *converges. Then* $a_n \to 0$ *as* $n \to \infty$. □

Note that the converse to the above is *not* true. For example, the *harmonic series* $\sum_{n=1}^{\infty} \frac{1}{n}$ does not converge, yet $\frac{1}{n} \to 0$ as $n \to \infty$.

Example 6.1.4 Show that the series

$$\sum_{n=1}^{\infty} \frac{n}{n+2}$$

does not converge.

Solution. Observe

$$\lim_{n\to\infty} \frac{n}{n+2} = \lim_{n\to\infty} \frac{1}{1+2/n} = 1 \neq 0.$$

Therefore the series does not converge.

Definition 6.1.5 We say a series $\sum_{n=1}^{\infty} a_n$ *converges absolutely* when the series $\sum_{n=1}^{\infty} |a_n|$ converges.

Proposition 6.1.6 *Any series that converges absolutely also converges.*

Proof. Suppose that $\sum_{n=1}^{\infty} |a_n|$ converges. We need to show that $\sum_{n=1}^{\infty} a_n$ converges. Write $\sigma_n = |a_1| + \cdots + |a_n|$ and $s_n = a_1 + \cdots + a_n$. We know that (σ_n) converges and hence it is a Cauchy sequence. Let $\varepsilon > 0$. Then there exists N such that for $m, n > N$, $|\sigma_n - \sigma_m| < \varepsilon$. For $n > m > N$, we have

$$\begin{aligned}
|s_n - s_m| &= |(a_1 + \cdots + a_n) - (a_1 + \cdots + a_m)| = |a_{m+1} + \cdots + a_n| \\
&\leq |a_{m+1}| + \cdots + |a_n| = (|a_1| + \cdots + |a_n|) - (|a_1| + \cdots + |a_m|) \\
&= \sigma_n - \sigma_m = |\sigma_n - \sigma_m| < \varepsilon.
\end{aligned}$$

Thus (s_n) is also Cauchy and hence convergent. □

6.2 Convergence tests

In this section we look at some criteria to tell when a series converges. The first of these is called the *comparison test.*

Theorem 6.2.1 (Comparison test) *Let (a_n) and (b_n) be sequences of real numbers, where $a_n \geq 0$ and $b_n \geq 0$. Suppose that (b_n) is summable and $a_n \leq b_n$ for all sufficiently large n. Then (a_n) is summable.*

Proof. Let

$$s_n = a_1 + \cdots + a_n, \qquad t_n = b_1 + \cdots + b_n.$$

Then the sequence of partial sums (t_n) converges, so it is Cauchy.
Let $\varepsilon > 0$. Pick N such that

- $|t_n - t_m| < \varepsilon$ whenever $m, n \geq N$ and
- $a_n \leq b_n$ whenever $n \geq N$.

Then for $n > m > N$, we have

$$|s_n - s_m| = a_{m+1} + \cdots + a_n \leq b_{m+1} + \cdots + b_n = |t_n - t_m| < \varepsilon.$$

So the sequence of partial sums (s_n) is also Cauchy. Hence it converges, and (a_n) is summable as required. □

We can immediately apply this to absolute convergence.

Corollary 6.2.2 *Let $\sum_{n=1}^{\infty} b_n$ be absolutely convergent. Suppose we have a sequence (a_n) where $|a_n| \leq |b_n|$ for all sufficiently large n. Then the series $\sum_{n=1}^{\infty} a_n$ is also absolutely convergent.* □

Example 6.2.3 Show that the series

$$\sum_{n=1}^{\infty} n3^{-n}$$

converges.

Solution. Observe that for all $n \in \mathbb{N}$, $n \leq 2^n$. Hence

$$n3^{-n} \leq 2^n 3^{-n} = \left(\frac{2}{3}\right)^n$$

and the series

$$\sum_{n=1}^{\infty} \left(\frac{2}{3}\right)^n$$

is a geometric series, with common ratio $\frac{2}{3} < 1$, which converges. Hence the series

$$\sum_{n=1}^{\infty} n3^{-n}.$$

also converges by the comparison test.

Comparing a series to a geometric series, as in the previous example, gives us the *ratio test*, perhaps the most useful of any of the tests for convergence.

Theorem 6.2.4 (Ratio test) *Consider a series of non-zero terms*

$$\sum_{n=1}^{\infty} a_n$$

and suppose

$$\lim_{n\to\infty} \left| \frac{a_{n+1}}{a_n} \right| = r.$$

- *If* $r < 1$*, the series converges absolutely.*
- *If* $r > 1$*, the series does not converge.*

Note that if $r = 1$, we cannot tell by using the ratio test whether or not the series converges.

Proof. Suppose

$$\lim_{n\to\infty} \left| \frac{a_{n+1}}{a_n} \right| = r < 1.$$

Let $r < s < 1$. Then there exists an N such that if $n \geq N$, then

$$\left| \frac{a_{n+1}}{a_n} \right| \leq s,$$

that is, $|a_{n+1}| \leq |a_n|s$. In particular

$$|a_{N+1}| \leq |a_N|s,\ |a_{N+2}| \leq |a_N|s^2,\ |a_{N+3}| \leq |a_N|s^3, \ldots,$$

that is, $|a_{N+k}| \leq |a_N|s^k$ for all k.

Now, since $0 < s < 1$, the geometric series

$$\sum_{k=0}^{\infty} |a_N|s^k$$

converges. So by the comparison test, the series $\sum_{k=0}^{\infty} a_{N+k}$ converges absolutely. Since the series $\sum_{n=0}^{\infty} a_n$ only adds on finitely many terms, it also converges absolutely.

Now suppose

$$\lim_{n\to\infty} \left| \frac{a_{n+1}}{a_n} \right| = r > 1.$$

Then there exists an N such that if $n \geq N$, then

$$\left| \frac{a_{n+1}}{a_n} \right| \geq 1.$$

So,

$$\cdots \geq |a_{N+2}| \geq |a_{N+1}| \geq |a_N|.$$

So $\lim_{k\to\infty} |a_{N+k}| \geq |a_N| > 0$. So the sequence (a_n) does not converge to zero and so the series does not converge. □

Example 6.2.5 Let $x \in \mathbb{R}$. Investigate convergence of the series

$$\sum_{n=1}^{\infty} \frac{x^n}{n}$$

using the ratio test.

Solution. Set $a_n = x^n/n$. Then

$$\frac{a_{n+1}}{a_n} = \frac{x^{n+1}/(n+1)}{x^n/n} = \frac{nx}{n+1} = \frac{x}{1+1/n}$$

So

$$\lim_{n\to\infty} \left| \frac{a_{n+1}}{a_n} \right| = |x|.$$

Thus, by the ratio test, our series converges absolutely if $|x| < 1$, and does not converge if $|x| > 1$.

The ratio test does not tell us about convergence in the above example for $x = \pm 1$. It turns out that the series converges, but not absolutely, when $x = -1$ (to $-\ln(2)$), and does not converge when $x = 1$ (this is the harmonic series).

Our final test for convergence, called the *root test*, is mainly theoretically useful. Before stating it, observe that for any bounded sequence of numbers (b_n), where $b_n \geq 0$, the sequence (c_n), where

$$c_n = \sup\{b_n, b_{n+1}, b_{n+2}, \ldots\}$$

is bounded and monotone decreasing, so the sequence (c_n) converges.

If the sequence (b_n) is not bounded, we adopt the convention that the limit of the sequence (c_n) is ∞.

Theorem 6.2.6 (Root test) *Let $\sum_{n=1}^{\infty} a_n$ be a series, and set*

$$c = \lim_{n\to\infty} \sup\{|a_n|^{\frac{1}{n}}, |a_{n+1}|^{\frac{1}{n+1}}, |a_{n+2}|^{\frac{1}{n+2}}, \ldots\}.$$

- *If $c < 1$, then the series converges absolutely.*
- *If $c > 1$, then the series does not converge.*

Proof. Let
$$c_n = \sup\{|a_n|^{\frac{1}{n}}, |a_{n+1}|^{\frac{1}{n+1}}, |a_{n+2}|^{\frac{1}{n+2}}, \ldots\}.$$
Suppose $c < 1$. Pick $c < d < 1$. Then (taking $\varepsilon = d - c$ in the definition of convergence), we see that there is some N such that $c_n < d$ for all $n \geq N$. Since $d < 1$, the geometric series
$$\sum_{n=1}^{\infty} d^n$$
converges.

Since $|a_n|^{1/n} \leq c_n$, we have $|a_n| \leq c_n^n < d^n$ if $n \geq N$. So by the comparison test, the series $\sum_{n=1}^{\infty} a_n$ converges absolutely.

Now suppose $c > 1$. Then there is some N such that $c_n > 1$ for all $n \geq N$. Hence, for each $n \geq N$, there must be some $m > n$ with $|a_m|^{\frac{1}{m}} > 1$, and so $|a_m| > 1$. Thus we find a subsequence (a_{n_k}) of (a_n) with all terms greater than 1. But this means that the sequence (a_n) does not converge to zero, so the series $\sum_{n=1}^{\infty} a_n$ cannot converge. □

7 Integration

In this chapter we will go over the rigorous definition of the integral. You have seen the idea before in MAS110, but now we can make it precise using limits.

7.1 Integration of step functions

We start by considering functions for which it is easy to capture explicitly the idea of an integral as a (signed) area.

Recall that a *bounded interval*, I, is a set of one of the forms: (c, d), $[c, d]$, $(c, d]$ or $[c, d)$. Here $c, d \in \mathbb{R}$ with $c \leq d$. For each of these we define the *length* of I to be
$$\text{length}(I) = d - c.$$

Definition 7.1.1 Let $a, b \in \mathbb{R}$, with $a < b$. Let $A \subseteq [a, b]$. Then we define the *characteristic function* of A, $\chi_A \colon [a, b] \to \mathbb{R}$, by the formula
$$\chi_A(t) = \begin{cases} 1 & t \in A, \\ 0 & t \notin A. \end{cases}$$

A function $s \colon [a, b] \to \mathbb{R}$ is called a *step function* if we have constants $\alpha_i \in \mathbb{R}$ and bounded intervals $I_i \subseteq [a, b]$ such that
$$s(t) = \sum_{i=1}^{n} \alpha_i \chi_{I_i}(t), \qquad \text{for all } t \in [a, b].$$

Notice that a step function has only finitely many steps.

The idea of the integral of a function is that it finds the (signed) area under its graph. Here signed means that area above the x-axis is counted positively, area under the x-axis is counted negatively. We use this idea to build the definition. Linking this with the "opposite of differentiation" is a profound *theorem*, and not the definition; we will see this theorem soon.

Given a function $s(t) = \alpha\chi_I(t)$, where I is a bounded inteval, the area under the graph is a rectangle, with width length(I) and height $|\alpha|$, and the (signed) area we want is

$$\alpha\text{length}(I).$$

The area under a step function comes from adding areas of rectangles togeher, making the following a reasonable definition.

Definition 7.1.2 Let $s\colon [a, b] \to \mathbb{R}$ be a step function. Write

$$s(t) = \sum_{i=1}^{n} \alpha_i \chi_{I_i}(t).$$

Then we define the *integral* of s to be

$$\int_a^b s(t)\,dt = \sum_{i=1}^{n} \alpha_i \text{length}(I_i).$$

The same step function can be written in different ways.

Example 7.1.3 Define step functions $r, s\colon [0, 5] \to \mathbb{R}$ by

$$r(t) = \chi_{[1,3]}(t) + 2\chi_{[2,4]}(t), \qquad s(t) = \chi_{[1,2)}(t) + 3\chi_{[2,3]}(t) + 2\chi_{(3,4]}(t).$$

Notice that in fact $r(t) = s(t)$ for all t. Calculate $\int_0^5 r(t)\ dt$ and $\int_0^5 s(t)\ dt$.

Solution. We have

$$\int_0^5 r(t)\ dt = 1 \times 2 + 2 \times 2 = 6$$

and

$$\int_0^5 s(t)\ dt = 1 \times 1 + 3 \times 1 + 2 \times 1 = 6.$$

As illustrated in the example, the definition of the integral for a step function does not depend on how the step function is written. We will not give the details of the argument, but it can be done by going via a representation involving *disjoint* intervals and then showing that subdivision of those intervals does not change the integral.

The following is immediate from the definition.

Proposition 7.1.4 *Let $r, s\colon [a,b] \to \mathbb{R}$ be step functions, and $\alpha, \beta \in \mathbb{R}$. Then $\alpha r + \beta s : [a,b] \to \mathbb{R}$ is a step function and*

$$\int_a^b \alpha r(t) + \beta s(t)\, dt = \alpha \int_a^b r(t)\, dt + \beta \int_a^b s(t)\, dt.$$

□

The next result is clear geometrically; the proof is left as an exercise.

Proposition 7.1.5 *Let $r, s\colon [a,b] \to \mathbb{R}$ be step functions with $r(t) \le s(t)$ for all $t \in [a,b]$. Then*

$$\int_a^b r(t)\, dt \le \int_a^b s(t)\, dt.$$

□

Corollary 7.1.6 *Let $s\colon [a,b] \to \mathbb{R}$ be a step function. Then $|s(t)|\colon [a,b] \to \mathbb{R}$ is also a step function and*

$$\left| \int_a^b s(t)\, dt \right| \le \int_a^b |s(t)|\, dt.$$

Proof. If $s(t) = \sum_{i=1}^n \alpha_i \chi_{I_i}(t)$, then $|s(t)| = \sum_{i=1}^n |\alpha_i| \chi_{I_i}(t)$, so $|s(t)|$ is also a step function. For all t,

$$-|s(t)| \le s(t) \le |s(t)|.$$

Hence, by Propositions 7.1.5 and 7.1.4,

$$-\int_a^b |s(t)|\ dt \le \int_a^b s(t)\ dt \le \int_a^b |s(t)|\ dt.$$

The result now follows. □

The final property we need for later work is also clear from the definition.

Proposition 7.1.7 *Let $s\colon [a,b] \to \mathbb{R}$ be a step function. Let $a \le c \le b$. Then the restriction of s to each of $[a,c]$ and $[c,b]$ is again a step function and*

$$\int_a^b s(t)\ dt = \int_a^c s(t)\ dt + \int_c^b s(t)\ dt.$$

□

7.2 The Riemann integral

Let $f\colon [a,b] \to \mathbb{R}$ be a bounded function, that is to say we have constants m and M such that $m \le f(t) \le M$ for all $t \in [a,b]$. Then if $s(t)$ is a step function, and $s(t) \le f(t)$, then $s(t) \le M$ for all t, so

$$\int_a^b s(t)\,dt \le M(b-a).$$

Similarly, if $s(t)$ is a step function, and $s(t) \ge f(t)$ for all t, then

$$\int_a^b s(t)\,dt \ge m(b-a).$$

We want to define the area under the graph of f. We approach this by approximating f by step functions, since we already know how to treat this case, by the previous section.

There are two ways we can go about this approximation: approximating from below, and from above.

Definition 7.2.1 Let $f\colon [a,b] \to \mathbb{R}$ be a bounded function. We define the *lower integral* of f:

$$L\int_a^b f(t)\,dt = \sup\left\{\int_a^b s(t)\,dt \mid s \text{ is a step function, } s(t) \le f(t) \text{ for all } t\right\}.$$

We define the *upper integral* of f:

$$U\int_a^b f(t)\ dt = \inf\left\{\int_a^b s(t)\,dt \mid s \text{ is a step function, } s(t) \ge f(t) \text{ for all } t\right\}.$$

Notice that the set involved in the definition of the lower integral,

$$\left\{\int_a^b s(t)\,dt \mid s \text{ is a step function, } s(t) \le f(t) \text{ for all } t\right\},$$

is a non-empty bounded above subset of $\mathbb{R}$. The integral of a step function is by definition a real number. If $m \le f(t)$ for all $t \in [a,b]$, then $s = m\chi_{[a,b]} : [a,b] \to \mathbb{R}$ is a step function such that $s(t) \le f(t)$ for all t. So $\int_a^b s(t)\,dt = m(b-a)$ is an element of the set. And if $f(t) \le M$ for all $t \in [a,b]$, then the set is bounded above by $M(b-a)$. Thus the set has a supremum.

Similarly, the set involved in the definition of the upper integral is a non-empty bounded below subset of $\mathbb{R}$ and so has an infimum.

Note that if s is a step function, then

$$\int_a^b s(t)\,dt = L\int_a^b s(t)\,dt = U\int_a^b s(t)\,dt.$$

Definition 7.2.2 We call a bounded function $f : [a, b] \to \mathbb{R}$ *Riemann integrable* if

$$L \int_a^b f(t)\, dt = U \int_a^b f(t)\, dt.$$

In this case we define the integral

$$\int_a^b f(t)\, dt = L \int_a^b f(t)\, dt = U \int_a^b f(t)\, dt.$$

Step functions are clearly Riemann integrable, but not every bounded function is.

Example 7.2.3 Define $f\colon [0, 1] \to \mathbb{R}$ by

$$f(t) = \begin{cases} 0 & t \in \mathbb{Q}, \\ 1 & t \notin \mathbb{Q}. \end{cases}$$

Show that f is not Riemann integrable.

Solution. Note that *every* interval, however small, contains both rational and irrational numbers. So if $s\colon [0, 1] \to \mathbb{R}$ is a step function, if $s(t) \geq f(t)$ for all t, then $s(t) \geq 1$. And if $s(t) \leq f(t)$ for all t, then $s(t) \leq 0$. We see that

$$L \int_0^1 f(t)\, dt = 0, \qquad U \int_0^1 f(t)\, dt = 1,$$

and since these are not equal, the function f is not Riemann integrable.

At the end of the next chapter we will prove the major result that if $f\colon [a, b] \to \mathbb{R}$ is continuous, then f is Riemann integrable (Theorem 8.4.4).

Propositions 7.1.4, 7.1.5, 7.1.7, and Corollary 7.1.6 give us the following result.

Proposition 7.2.4 *Let the bounded functions $f, g\colon [a, b] \to \mathbb{R}$ be Riemann integrable.*

- *Let $\alpha, \beta \in \mathbb{R}$. Then $\alpha f + \beta g : [a, b] \to \mathbb{R}$ is Riemann integrable and*

$$\int_a^b \alpha f(t) + \beta g(t)\, dt = \alpha \int_a^b f(t)\, dt + \beta \int_a^b g(t)\, dt.$$

- *Suppose $f(t) \leq g(t)$ for all $t \in [a, b]$. Then*

$$\int_a^b f(t)\, dt \leq \int_a^b g(t)\, dt.$$

- *Let $a \le c \le b$. Then*

$$\int_a^b f(t)\,dt = \int_a^c f(t)\,dt + \int_c^b f(t)\,dt.$$

- *The function $|f(x)| : [a,b] \to \mathbb{R}$ is Riemann integrable, and*

$$\left|\int_a^b f(t)\,dt\right| \le \int_a^b |f(t)|\,dt.$$

□

Corollary 7.2.5 *Let the bounded function $f\colon [a,b] \to \mathbb{R}$ be Riemann integrable. Let*

$$m = \inf\{f(x) \mid x \in [a,b]\}, \qquad M = \sup\{f(x) \mid x \in [a,b]\}.$$

Then

$$m(b-a) \le \int_a^b f(t)\ dt \le M(b-a).$$

Proof. Certainly

$$m \le f(t) \le M$$

for all $t \in [a,b]$, so by Proposition 7.2.4,

$$\int_a^b m\,dt \le \int_a^b f(t)\,dt \le \int_a^b M\,dt.$$

Integrating the lefthand and righthand terms, the result follows. □

7.3 The fundamental theorem of calculus

Remember that integration was defined in terms of areas. In this section we link it with the reverse of differentiation. This is known as the *fundamental theorem of calculus*. You are already familiar with this statement, but now we are in a position to give a full proof.

Theorem 7.3.1 (Fundamental theorem of calculus) *Let $f\colon [a,b] \to \mathbb{R}$ be continuous. Define $F\colon [a,b] \to \mathbb{R}$ by*

$$F(x) = \int_a^x f(t)\,dt.$$

Then F is differentiable, and $F'(x) = f(x)$ for all $x \in [a,b]$.

Proof. Let $h > 0$ and let $x \in [a,b]$. Set

$$m_h = \inf\{f(t)\ t \in [x,x+h]\}, \qquad M_h = \sup\{f(t) \mid t \in [x,x+h]\}.$$

Recall the boundedness theorem (Theorem 4.3.4) from semester one: a continuous function f on a closed bounded interval is bounded and attains its bounds. So $m_n, M_n \in \mathbb{R}$ for all n.

Note that

$$\int_x^{x+h} f(t)\,dt = \int_a^{x+h} f(t)\,dt - \int_a^x f(t)\,dt = F(x+h) - F(x).$$

So by Corollary 7.2.5

$$m_h h \le F(x+h) - F(x) \le M_h h$$

and so

$$m_h \le \frac{F(x+h) - F(x)}{h} \le M_h.$$

The same inequality holds when $h < 0$; the proof is similar.

Since f is continuous at x, we have that

$$\lim_{h\to 0} m_h = \lim_{h\to 0} M_h = f(x).$$

So by the sandwich rule,

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

That is, F is differentiable at x and $F'(x) = f(x)$. □

The version of the theorem that we use in practice to calculate integrals follows as a corollary.

Corollary 7.3.2 *Let $f\colon [a,b] \to \mathbb{R}$ be a continuous function. Suppose we have a differentiable function $F\colon [a,b] \to \mathbb{R}$ such that $F'(x) = f(x)$ for all $x \in [a,b]$. Then*

$$\int_a^b f(t)\,dt = F(b) - F(a).$$

Proof. Let

$$G(x) = \int_a^x f(t)\,dt.$$

Then $G(a) = 0$, and by Theorem 7.3.1, G is differentiable with

$$G'(x) = f(x) = F'(x).$$

So we have a constant C such that $G(x) = F(x) + C$ for all $x \in [a,b]$. Now

$$\int_a^b f(t)\ dt = G(b) = G(b) - G(a) = (F(b) + C) - (F(a) + C) = F(b) - F(a).$$

□

Note that all of the techniques we already know about finding integrals, such as substitution and integration by parts, can be proved for continuous functions with an "antiderivative".

7.4 Improper integrals

This section was not covered in 2017-18 and is not examinable.

8 Sequences and series of functions

In this chapter we will study sequences and series of functions. We start with two different notions of convergence for a sequence of functions. Later we will consider the interaction between continuity and these notions of convergence, as well as looking at integration.

8.1 Pointwise and uniform convergence

Consider a sequence of functions, (f_n), where $f_n\colon [a,b] \to \mathbb{R}$.

Definition 8.1.1 We say the sequence (f_n) *converges pointwise* to a function $f\colon [a,b] \to \mathbb{R}$ if for each $t \in [a,b]$, we have

$$\lim_{n\to\infty} f_n(t) = f(t).$$

In other words, the sequence (f_n) converges pointwise to f if for each $t \in [a,b]$ and $\varepsilon > 0$, there exists $N \in \mathbb{N}$ such that $|f_n(t) - f(t)| < \varepsilon$ whenever $n \geq N$.

Note that in this definition, the N can depend not just on the value of ε, but also on the point t.

Example 8.1.2 Define $f_n\colon [0,2\pi] \to \mathbb{R}$ by $f_n(t) = \cos(t/n)$. Show that the sequence (f_n) converges pointwise and determine the limit function.

Solution. Observe that $t/n \to 0$ as $n \to \infty$, and cos is a continuous function, so for each $t \in [0,2\pi]$, we have

$$\cos\left(\frac{t}{n}\right) \to \cos 0 = 1$$

as $n \to \infty$. Thus the sequence (f_n) has pointwise limit the constant function $f : [0,2\pi] \to \mathbb{R}$, with $f(t) = 1$ for all $t \in [0,2\pi]$.

Example 8.1.3 Define $f_n\colon [0,1] \to \mathbb{R}$ by $f_n(t) = t^n$. Show that the sequence (f_n) converges pointwise and determine the limit function.

Solution. For $t < 1$,

$$f(t) = t^n \to 0$$

as $n \to \infty$.

On the other hand, $1^n = 1$ for all n, so $f(1) \to 1$ as $n \to \infty$. So the sequence (f_n) has pointwise limit f, where

$$f(t) = \begin{cases} 0 & t < 1, \\ 1 & t = 1. \end{cases}$$

Example 8.1.4 Define $f_n \colon [0, 2] \to \mathbb{R}$ by $f_n(t) = t^n$. Show that the sequence (f_n) does not converge pointwise.

Solution. Note that if $t > 1$, then $t^n \to \infty$ as $n \to \infty$, so if $t > 1$ then

$$\lim_{n\to\infty} f_n(t)$$

does not exist, and (f_n) has no pointwise limit.

Instead of pointwise convergence, we could insist that a sequence of functions converges at the same rate at each point. This concept is called *uniform convergence* and it is a much stronger condition. Uniform convergence, as we shall see, preserves such things as continuity and to an extent differentiation. As the second of the above examples show, pointwise convergence does not preserve continuity.

Definition 8.1.5 Let $f_n \colon [a, b] \to \mathbb{R}$. We say the sequence (f_n) *converges uniformly* to a function $f \colon [a, b] \to \mathbb{R}$ when for all $\varepsilon > 0$, we have $N \in \mathbb{N}$ such that $|f_n(t) - f(t)| < \varepsilon$ for all $n \geq N$ and $t \in [a, b]$.

The difference between this definition and pointwise convergence is that the N does not depend on the point t, only on ε. Note that uniform convergence of a sequence (f_n) to a function f *implies* pointwise convergence to the same function f.

Proposition 8.1.6 *Consider a sequence of functions $f_n \colon [a, b] \to \mathbb{R}$. Let $f \colon [a, b] \to \mathbb{R}$. Then the following statements are equivalent.*

- *The sequence (f_n) converges uniformly to f.*
- *Let $M_n = \sup\{|f_n(t) - f(t)| \mid t \in [a, b]\}$. Then $M_n \to 0$ as $n \to \infty$.*

Proof. Let (f_n) converge uniformly to f.

Let $\varepsilon > 0$. Then we have $N \in \mathbb{N}$ such that $|f_n(t) - f(t)| < \frac{\varepsilon}{2}$ whenever $n \geq N$, for all $t \in [a, b]$. Let $n \geq N$. Then

$$M_n = \sup\{|f_n(t) - f(t)| \mid t \in [a, b]\} \leq \frac{\varepsilon}{2} < \varepsilon.$$

Thus $M_n \to 0$ as $n \to \infty$, and the second condition holds.

Conversely, suppose $M_n \to 0$ as $n \to \infty$. Let $\varepsilon > 0$. Then we have $N \in \mathbb{N}$ such that $M_n < \varepsilon$ whenever $n \geq N$. But this means, for $n \geq N$, that

$$\sup_{t\in[a,b]} |f_n(t) - f(t)| < \varepsilon$$

and so $|f_n(t) - f(t)| < \varepsilon$ whenever $n \geq N$, for all $t \in [a, b]$. Thus (f_n) converges to f uniformly, as required. □

The condition on the M_ns in Proposition 8.1.6 is usually the easiest way to actually prove uniform convergence in an example.

Example 8.1.7 For $n \geq 2$, define $f_n \colon [0, 2\pi] \to \mathbb{R}$ by

$$f_n(t) = \frac{1 - 2n}{1 - n} \sin t.$$

First show that the sequence (f_n) converges pointwise and determine the limit function f. Then show that in fact the sequence (f_n) converges uniformly to f.

Solution. Observe that

$$f_n(t) = \frac{\frac{1}{n} - 2}{\frac{1}{n} - 1} \sin t \to 2 \sin t$$

as $n \to \infty$, so (f_n) converges pointwise to f, where $f(t) = 2 \sin t$.

Then

$$|f_n(t) - f(t)| = \left| \frac{1 - 2n}{1 - n} - 2 \right| |\sin t| \leq \left| \frac{1 - 2n}{1 - n} - 2 \right|$$

for all t, and

$$\left| \frac{1 - 2n}{1 - n} - 2 \right| = \frac{1}{1 - n} \to 0$$

as $n \to \infty$.

Thus, if

$$M_n = \sup\{|f_n(t) - f(t)| \mid t \in [0, 2\pi]\}$$

then $0 \leq M_n \leq \frac{1}{1-n}$. And since $\frac{1}{1-n} \to 0$ as $n \to \infty$, $M_n \to 0$ as $n \to \infty$. So (f_n) converges uniformly to f.

8.2 Continuity under uniform limits

Theorem 8.2.1 (Uniform limit theorem) *Let $f_n \colon [a, b] \to \mathbb{R}$ be continuous for each $n \in \mathbb{N}$. Suppose the sequence (f_n) converges uniformly to a function $f \colon [a, b] \to \mathbb{R}$. Then f is continuous.*

The proof is sometimes called the $\varepsilon/3$ proof because of the main trick.

Proof. Let $t_0 \in [a, b]$. We want to prove that f is continuous at t_0. Let $\varepsilon > 0$. Then:

- We have $N \in \mathbb{N}$ such that $|f_n(t) - f(t)| < \frac{\varepsilon}{3}$ whenever $n \geq N$, for all $t \in [a, b]$ (since (f_n) converges uniformly to f).
- We have $\delta > 0$ such that $|f_N(t) - f_N(t_0)| < \frac{\varepsilon}{3}$ whenever $|t - t_0| < \delta$ (since f_N is continuous at t_0).

So, let t be such that $|t - t_0| < \delta$. Using the above two conditions,

$$|f(t)-f(t_0)| \leq |f(t)-f_N(t)|+|f_N(t)-f_N(t_0)|+|f_N(t_0)-f(t_0)| < \frac{\varepsilon}{3}+\frac{\varepsilon}{3}+\frac{\varepsilon}{3} = \varepsilon.$$

Thus f is continuous at t_0. □

The result also holds, with essentially the same proof, for functions $f_n : \mathbb{R} \to \mathbb{R}$.

Example 8.2.2 Let $f_n : [0, 1] \to \mathbb{R}$ be given by $f_n(t) = t^n$. Show that the sequence (f_n) does not converge uniformly.

Solution. We saw in Example 8.1.3 that the sequence (f_n) has pointwise limit f, where

$$f(t) = \begin{cases} 0 & t < 1, \\ 1 & t = 1. \end{cases}$$

If (f_n) did converge uniformly the limit would have to be this function f. But f is clearly not continuous. So (f_n) does not converge uniformly.

8.3 Integration and differentiation

We will look at how uniform convergence interacts with integration. The following result is called the *uniform convergence theorem* for integrals. Many of our results and examples in the rest of this course are applications of this theorem.

Theorem 8.3.1 (Uniform convergence theorem) *Let $f_n \colon [a, b] \to \mathbb{R}$ be a continuous function, for $n \in \mathbb{N}$. Suppose the sequence (f_n) converges uniformly to a function f. Then*

$$\lim_{n\to\infty} \int_a^b f_n(t)\,dt = \int_a^b f(t)\,dt.$$

Proof. By Theorem 8.2.1, the function f is continuous, and therefore integrable. Let $\varepsilon > 0$. Since (f_n) converges uniformly to f, we have $N \in \mathbb{N}$ such that

$$|f_n(t) - f(t)| < \frac{\varepsilon}{2(b-a)}$$

whenever $n \geq N$, for all $t \in [a, b]$.

Hence by Proposition 7.2.4 and Corollary 7.2.5, for $n \geq N$ we have

$$\left| \int_a^b f_n(t)\,dt - \int_a^b f(t)\,dt \right| = \left| \int_a^b f_n(t) - f(t)\,dt \right|$$
$$\leq \int_a^b |f_n(t) - f(t)|\,dt \leq \frac{(b-a)\varepsilon}{2(b-a)} = \frac{\varepsilon}{2} < \varepsilon.$$

The result now follows. □

Thus we can, *under suitable conditions*, swap limits and integral signs. This is frequently useful.

The result actually still holds if each f_n is Riemann integrable rather than continuous, but the proof is much more involved, and the above is enough for our purposes.

The following example shows that we do need *uniform* convergence to be able to swap limits and integrals.

Example 8.3.2 Let $f_n : [0, 1] \to \mathbb{R}$, for $n \in \mathbb{N}$ be defined by

$$f_n(t) = nt(1 - t^2)^n.$$

1. Show that (f_n) converges pointwise to the zero function.
2. Calculate $\lim_{n\to\infty} \int_0^1 f_n(t)\,dt$.
3. Does (f_n) converge uniformly?

Solution. (Outline)

1. This is left as an exercise, with the hint to use the binomial expansion of $(1 + t^2)^n$.
2. Using the substitution $u = t^2$, one can show that $\int_0^1 f_n(t)\,dt = \frac{n}{2(n+1)}$. So

$$\lim_{n\to\infty} \int_0^1 f_n(t)\,dt = \lim_{n\to\infty} \frac{n}{2(n+1)} = \lim_{n\to\infty} \frac{1}{2 + 2/n} = \frac{1}{2}.$$

3. From the previous parts we see that

$$\lim_{n\to\infty} \int_0^1 f_n(t)\,dt = \frac{1}{2} \neq 0 = \int_0^1 \lim_{n\to\infty} f_n\,dt.$$

Thus (f_n) cannot converge uniformly (otherwise this example would contradict the uniform convergence theorem).

Our first immediate application is to differentiation. Indeed, it is natural to ask at this point about swapping limits and differentiation.

Corollary 8.3.3 *Consider differentiable functions $f_n\colon [a,b] \to \mathbb{R}$. Suppose (f_n) converges pointwise to a function f, and the sequence of derivatives (f_n') converges uniformly to a function g. Then f is differentiable, and $f' = g$.*

The proof uses the fundamental theorem of calculus and the uniform convergence theorem, but details are omitted.

Thus we can, *under suitable conditions*, swap limits and differentiation. The following example shows that convergence conditions on the sequence of derivatives are necessary.

Example 8.3.4 Define $f_n\colon [0, 2\pi] \to \mathbb{R}$ by

$$f_n(t) = \frac{1}{n}\sin(n^2 t).$$

Show that (f_n) converges uniformly to the zero function, but (f_n') does not converge pointwise.

Solution. Observe that $|f_n(t)| \leq \frac{1}{n}$ for all $t \in [0, 2\pi]$, and $\frac{1}{n} \to 0$ as $n \to \infty$. Thus the sequence (f_n) converges uniformly to 0, the zero function.

But

$$f_n'(t) = n\cos(n^2 t)$$

and the sequence (f_n') does not converge pointwise (let alone uniformly).

8.4 Uniform continuity

We recall once again what it means for a function to be continuous.

A function $f\colon [a,b] \to \mathbb{R}$ is continuous at a point $x_0 \in [a,b]$ if for all $\varepsilon > 0$ we have $\delta > 0$ such that if $|x - x_0| < \delta$ then $|f(x) - f(x_0)| < \varepsilon$.

The function f is continuous on $[a,b]$ if it is continuous at every point $x_0 \in [a,b]$.

Along similar lines to uniform as opposed to pointwise convergence, we make the following definition.

Definition 8.4.1 Let $f\colon [a,b] \to \mathbb{R}$. We say f is *uniformly continuous* on $[a,b]$ if for all $\varepsilon > 0$ we have $\delta > 0$ such that if $|x - y| < \delta$ then $|f(x) - f(y)| < \varepsilon$ for all $x, y \in [a,b]$.

The difference here is that the δ *can only depend on* ε, not on a chosen point x_0 as well. For a given ε, the same δ has to work across the whole of the interval $[a,b]$.

We similarly define uniform continuity of a function $f\colon \mathbb{R} \to \mathbb{R}$.

It is clear that if f is uniformly continuous then f is continuous. The following example shows that the converse need not hold.

Example 8.4.2 Define $f\colon \mathbb{R} \to \mathbb{R}$ by $f(x) = x^2$. Certainly f is continuous. Show that f is not uniformly continuous.

Solution. Let $x, y \in \mathbb{R}$, and suppose $|x - y| = c > 0$. Then

$$|f(x) - f(y)| = |x^2 - y^2| = |(x - y)(x + y)| = c|x + y|.$$

So if $x, y \geq 1/c$, then $|f(x) - f(y)| > 1$.

Taking $\varepsilon = 1$ in the definition of uniform continuity, we see that there is no $\delta > 0$ such that $|x - y| < \delta$ ensures that $|f(x) - f(y)| < 1$. Thus f is not uniformly continuous.

In view of the above, the following feels surprising: on a *closed bounded interval* there is no difference between continuity and uniform continuity. We will omit the proof.

Theorem 8.4.3 *Let $f\colon [a, b] \to \mathbb{R}$ be continuous. Then f is uniformly continuous.*

As an application, we use this result to supply a sketch of the proof promised earlier that continuous functions on bounded intervals are Riemann integrable.

Theorem 8.4.4 *Let $f\colon [a, b] \to \mathbb{R}$ be continuous. Then f is Riemann integrable.*

Sketch proof.

- By Theorem 8.4.3, the function f is uniformly continuous.
- Let $\varepsilon > 0$. We want to find a step function greater than or equal to f and a step function less than or equal to f with integrals within distance ε of each other.
- Since f is uniformly continuous, we have $\delta > 0$ such that if $|x - y| < \delta$, then
$$|f(x) - f(y)| < \frac{\varepsilon}{2(b - a)}.$$
- Divide the interval $[a, b]$ into subintervals of length smaller than δ.
- Define step functions r, s using the inf and sup of f on each of these subintervals as the steps.
- Then $r(t) \leq f(t)$ and $s(t) \geq f(t)$ for all $t \in [a, b]$ and
$$\left| \int_a^b s(t)\, dt - \int_a^b r(t)\, dt \right| < \varepsilon.$$

- Hence
$$\left|U\int_a^b f(t)\,dt - L\int_a^b f(t)\,dt\right| < \varepsilon.$$
- Since this holds for all $\varepsilon > 0$, the upper and lower integrals must be equal. So by definition, the function f is Riemann integrable. □

8.5 Series of functions

We can treat convergence of series of functions as we did for series of real numbers, via sequences of partial sums.

Definition 8.5.1 Let (f_n) be a sequence of functions, $f_n\colon [a,b] \to \mathbb{R}$. We consider the sequence of partial sums (s_n), where $s_n\colon [a,b] \to \mathbb{R}$ is defined by

$$s_n(t) = f_1(t) + f_2(t) + \cdots + f_n(t).$$

We say the series $\sum_{n=1}^{\infty} f_n$ *converges pointwise* if, for each $t \in [a,b]$, the series $\sum_{n=1}^{\infty} f_n(t)$ converges. That is, for each $t \in [a,b]$, the sequence of partial sums $(s_n(t))$ converges, or equivalently, the sequence (s_n) converges pointwise.

We say that the sequence (f_n) is *uniformly summable*, or that the series $\sum_{n=1}^{\infty} f_n$ *converges uniformly*, on $[a,b]$ if the sequence (s_n) of partial sums converges uniformly.

One also makes the same definition for functions defined on $\mathbb{R}$.

Theorem 8.2.1 and Corollary 8.3.3 immediately give us the following two results.

Theorem 8.5.2 *Let (f_n) be a uniformly summable sequence of continuous functions $f_n\colon [a,b] \to \mathbb{R}$. Then the function $f\colon [a,b] \to \mathbb{R}$ given by*

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

is continuous. □

Theorem 8.5.3 *Let (f_n) be a sequence of differentiable functions, $f_n\colon [a,b] \to \mathbb{R}$, such that the series $\sum_{n=1}^{\infty} f_n$ converges pointwise to a function f, and the series $\sum_{n=1}^{\infty} f_n'$ is uniformly summable. Then f is differentiable and*

$$f'(t) = \sum_{n=1}^{\infty} f_n'(t)$$

for all $t \in [a,b]$. □

In order for the above to be useful, we need a criterion for uniform convergence of a series. The following result, called the *Weierstrass M-test* provides a handy criterion.

Theorem 8.5.4 (Weierstrass M-test) *Let $f_n\colon [a,b] \to \mathbb{R}$ be a sequence of functions. Suppose we have a summable sequence of real numbers (M_n) such that $|f_n(t)| \le M_n$ for all n, and all $t \in [a,b]$. Then the sequence (f_n) is uniformly summable. Further, for each $t \in [a,b]$, the series*

$$\sum_{n=1}^{\infty} f_n(t)$$

converges absolutely.

Proof. For each $t \in [a,b]$, absolute convergence of the series

$$\sum_{n=1}^{\infty} f_n(t)$$

follows immediately from the comparison test. Define $f : [a,b] \to \mathbb{R}$ by $f(t) = \sum_{n=1}^{\infty} f_n(t)$.

Let $\varepsilon > 0$. Since the series $\sum_{n=1}^{\infty} M_n$ converges, we have N such that $\sum_{n=r+1}^{\infty} M_n < \varepsilon$ whenever $r \ge N$.

Let $t \in [a,b]$. Then

$$|f(t) - (f_1(t) + \cdots + f_r(t))| = \left| \sum_{n=r+1}^{\infty} f_n(t) \right| \le \sum_{n=r+1}^{\infty} |f_n(t)| \le \sum_{n=r+1}^{\infty} M_n < \epsilon.$$

Since this holds for all $t \in [a,b]$, it follows that the sequence of partial sums of the series $\sum_{n=1}^{\infty} f_n$ converges uniformly, to f. □

Note that the Weierstrass M-test also works for functions $f_n\colon \mathbb{R} \to \mathbb{R}$.

We will see in the next chapter how to use uniform convergence of series in a class of well-behaved examples. But it can also be used to construct examples with pathological properties.

Proposition 8.5.5 *There is a function $f\colon \mathbb{R} \to \mathbb{R}$ which is continuous everywhere, but differentiable nowhere.*

We just give a quick sketch of the proof.

Sketch proof.

- Define a function $f_n\colon \mathbb{R} \to \mathbb{R}$ by

$$f_n(t) = \frac{1}{10^n}\,(\text{distance from } 10^n t \text{ to the nearest integer})\,.$$

- Let $f = \sum_{n=1}^{\infty} f_n$.
- Use the Weierstrass M-test to show that the series $\sum_{n=1}^{\infty} f_n$ converges uniformly.
- Apply the Uniform Limit Theorem 8.2.1 to show that f is continuous.
- Check directly that f is not differentiable anywhere, by showing that the required limit does not exist. (To give full details of this is quite long. But the idea is straightforward: f_1 has "corners" at $\frac{n}{20}$ for $n \in \mathbb{Z}$, f_2 has "corners" at $\frac{n}{200}$ for $n \in \mathbb{Z}$ and so on - thus f has "corners" everywhere and is differentiable nowhere.) □

9 Applications

9.1 Power series

Definition 9.1.1 A series of the form

$$f(x) = \sum_{n=0}^{\infty} a_n x^n$$

where $a_n \in \mathbb{R}$ are constants is called a *power series.*

Theorem 9.1.2 *For a power series $f(x) = \sum_{n=0}^{\infty} a_n x^n$, one of the following holds:*

- *The series converges only when $x = 0$.*
- *The series converges for all $x \in \mathbb{R}$.*
- *There is a constant $R > 0$, such that the series converges absolutely if $|x| < R$, and does not converge if $|x| > R$.*

Proof. Clearly the series converges for $x = 0$. We apply the root test, Theorem 6.2.6. Let $x \in \mathbb{R}$, and let

$$c = \lim_{n\to\infty} \sup\{|a_n x^n|^{\frac{1}{n}}, |a_{n+1}x^{n+1}|^{\frac{1}{n+1}}, \dots\}.$$

Let

$$\alpha = \lim_{n\to\infty} \sup\{|a_n|^{\frac{1}{n}}, |a_{n+1}|^{\frac{1}{n+1}}, \dots\},$$

so that $c = |x|\alpha$.

If $\alpha = 0$, then $c = 0$, and the series $\sum_{n=0}^{\infty} a_n x^n$ converges absolutely for all x by the root test.

If $\alpha = \infty$, then $c = \infty$ if $x \neq 0$, and the series $\sum_{n=0}^{\infty} a_n x^n$ does not converge for $x \neq 0$ by the root test.

Otherwise, α is a real positive number and we let $R = 1/\alpha$. If $|x| < R$, then $c < 1$ and the series converges absolutely by the root test. If $|x| > R$, then $c > 1$, and the series does not converge, again by the root test. □

Definition 9.1.3 The constant R appearing in the statement of Theorem 9.1.2 is called the *radius of convergence* of the power series. If the series converges only when $x = 0$, we say the radius of convergence is 0. If the series converges for all $x \in \mathbb{R}$, we say the radius of convergence is ∞.

The conventions about 0 and ∞ mean that we don't have to treat special cases all the time.

We leave the proof of the following as an exercise, using the ratio test.

Proposition 9.1.4 *Let*

$$f(x) = \sum_{n=0}^{\infty} a_n x^n$$

be a power series. Then the radius of convergence is

$$R = \lim_{n\to\infty} \left| \frac{a_n}{a_{n+1}} \right|$$

if this limit exists. □

This result is usually the easiest way to find the radius of convergence in examples.

Example 9.1.5 Find the radius of convergence of the series

$$f(x) = \sum_{n=0}^{\infty} n x^n.$$

Solution. We have coefficients $a_n = n$. So

$$\lim_{n\to\infty} \left| \frac{a_n}{a_{n+1}} \right| = \lim_{n\to\infty} \frac{n}{n+1} = \lim_{n\to\infty} \frac{1}{1+1/n} = 1.$$

Thus the radius of convergence is 1.

For computations using power series, the following allows us to use our results on uniform convergence.

Lemma 9.1.6 *Let*

$$\sum_{n=0}^{\infty} a_n x^n$$

be a power series with radius of convergence R*. Let* $0 < S < R$*. Then the series converges uniformly on the interval* $[-S, S]$*.*

Proof. By definition of the radius of convergence, the series

$$\sum_{n=0}^{\infty} |a_n| S^n$$

converges. If $x \in [-S, S]$, then

$$|a_n x^n| \leq |a_n| S^n$$

The result now follows from the Weierstrass M-test (with $M_n = |a_n| S^n$). □

Theorem 9.1.7 (Termwise differentiation of power series) *Let*

$$f(x) = \sum_{n=0}^{\infty} a_n x^n$$

be a power series with radius of convergence $R > 0$*. Then the function* f *is differentiable on* $(-R, R)$*, with*

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

where this series also has radius of convergence R*.*

We just give a sketch of the proof.

Sketch proof

- Consider $\sum_{n=1}^{\infty} n a_n x^{n-1}$ and let the radius of convergence of this power series be R'.
- Use comparison with $\sum_{n=1}^{\infty} a_n x^n$ to show that $R' \leq R$.
- Use Example 9.1.5 and comparison to show that $R \leq R'$. So $R' = R$.
- By Lemma 9.1.6, $\sum_{n=1}^{\infty} n a_n x^{n-1}$ converges uniformly on $[-S, S]$.
- Apply Theorem 8.5.3 to conclude that we can differentiate term by term.

□

9.2 e and the exponential function

This section was not covered in 2017-18 and is not examinable.