MA3211

Mathematical Analysis II

Jia Xiaodong

May 9, 2022

1 Differentiation

1.1 Derivatives

Definition 1.1 (Differentiability at a point)**.** A function f is said to be *differentiable* at a point a if f is defined in some open interval containing a and the limit

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

exists. Then we call $f'(a)$ the *derivative* of f at a. ☐

By letting $h = x - a$ we can also write the limit as

$$\lim_{h \to 0} \frac{f(a + h) - f(a)}{h}.$$

Also geometrically $f'(a)$ is the slope of the tangent line to the curve $y = f(x)$ at $x = a$.

Definition 1.2 (Differentiable functions on open intervals)**.** If a function f is differentiable at every point in (a, b), then we say that f is differentiable on (a, b). ☐

Definition 1.3 (Differentiable functions on closed intervals)**.** If a function $f\colon [a, b] \to \mathbb{R}$ is such that

- f is differentiable on (a, b), and
- the one sided limits

  $$L_1 = \lim_{x \to a^+} \frac{f(x) - f(a)}{x - a} \qquad L_2 = \lim_{x \to b^-} \frac{f(x) - f(b)}{x - b}$$

  exists,

then we say that f is differentiable on $[a, b]$. Also we define $f'(a) = L_1$ and $f'(b) = L_2$. ☐

Similar definitions can be made for half-open intervals, etc.

Definition 1.4 (Derivatives)**.** Let f be a differentiable function on the interval I. Then the *derivative* of f is the function $f': I \to \mathbb{R}$ given by $x \mapsto f'(x)$ for all $x \in I$. We can also write

$$f'(x) = \frac{\mathrm{d}}{\mathrm{d}x} f(x).$$

□

Definition 1.5 (Continuously differentiable functions)**.** A function f is said to be *continuously differentiable* on an interval I if f is differentiable on I and f' is continuous on I. □

Example 1.1. Let $f(x) = c$, a constant. Then

$$\begin{aligned} f'(x) &= \lim_{h \to 0} \frac{f(x+h) - f(x)}{h} \\ &= \lim_{h \to 0} \frac{c - c}{h} \\ &= 0. \end{aligned}$$

◊

Example 1.2. Let $f(x) = x$. Then

$$\begin{aligned} f'(x) &= \lim_{h \to 0} \frac{f(x+h) - f(x)}{h} \\ &= \lim_{h \to 0} \frac{x + h - x}{h} \\ &= 1. \end{aligned}$$

◊

Example 1.3. Let $f(x) = x^n$ where $n \in \mathbb{N}$, Then

$$\begin{aligned} f'(x) &= \lim_{h \to 0} \frac{f(x+h) - f(x)}{h} \\ &= \lim_{h \to 0} \frac{(x+h)^n - x^n}{h} \\ &= \lim_{h \to 0} \frac{(x+h)^n - x^n}{h} \\ &= nx^{n-1}. \end{aligned}$$

◊

Theorem 1.1. *If f is differentiable at a, then it is continuous at a.*

Proof.

$$\begin{aligned} \lim_{x \to a} f(x) &= \lim_{x \to a} (f(x) - f(a) + f(a)) \\ &= \lim_{x \to a} \left[\frac{f(x) - f(a)}{x - a} (x - a) \right] + f(a) \\ &= \lim_{x \to a} \left(\frac{f(x) - f(a)}{x - a} \right) \lim_{x \to a} (x - a) + f(a) \\ &= f'(a) \cdot 0 + f(a) \\ &= f(a). \end{aligned}$$

■

Theorem 1.2 (Differentiation rules). *Let f and g be functions differentiable at a. Then*

i. $\frac{\mathrm{d}}{\mathrm{d}x} f(x) \pm g(x)\Big|_{x=a} = f'(a) \pm g'(a)$.

ii. Product rule. $\frac{\mathrm{d}}{\mathrm{d}x} f(x)g(x)\Big|_{x=a} = f'(a)g(a) + f(a)g'(a)$.

iii. Quotient rule. $\frac{\mathrm{d}}{\mathrm{d}x} \frac{f(x)}{g(x)}\Big|_{x=a} = \frac{f'(a)g(a) - f(a)g'(a)}{g(a)^2}$.

Proof.

i.

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}x} f(x) \pm g(x)\Big|_{x=a} &= \lim_{x\to a} \frac{f(x) \pm g(x) - (f(a) \pm g(a))}{x-a} \\ &= \lim_{x\to a} \frac{f(x) - f(a)}{x-a} \pm \lim_{x\to a} \frac{g(x)-g(a)}{x-a} \\ &= f'(a) \pm g'(a).\end{aligned}$$

ii.

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}x} f(x)g(x)\Big|_{x=a} &= \lim_{x\to a} \frac{f(x)g(x) - f(a)g(a)}{x-a} \\ &= \lim_{x\to a} \frac{[f(x)g(x) - f(x)g(a)] + [f(x)g(a) - f(a)g(a)]}{x-a} \\ &= \lim_{x\to a} \left[f(x)\frac{g(x)-g(a)}{x-a} + \frac{f(x)-f(a)}{x-a} g(a) \right] \\ &= f(a)g'(a) + f'(a)g(a).\end{aligned}$$

iii.

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}x} \frac{f(x)}{g(x)}\Big|_{x=a} &= \lim_{x\to a} \frac{\frac{f(x)}{g(x)} - \frac{f(a)}{g(a)}}{x-a} \\ &= \left(\lim_{x\to a} \frac{f(x)g(a) - f(a)g(x)}{g(x)g(a)} \right) / (x-a) \\ &= \left(\lim_{x\to a} \frac{f(x)g(a) - f(a)g(a) - f(a)g(x) + f(a)g(a)}{x-a} \right) / g(a)^2 \\ &= \frac{f'(a)g(a) - f(a)g'(a)}{g(a)^2}\end{aligned}$$

∎

We have shown that $\frac{\mathrm{d}}{\mathrm{d}x} x^n = nx^{n-1}$ for all $n \in \mathbb{N}$, and using the quotient rule inductively we can show that $\frac{\mathrm{d}}{\mathrm{d}x} \frac{1}{x^n} = -\frac{n}{x^{n+1}}$ for all $n \in \mathbb{N}$. So in general we can say that

$$\frac{\mathrm{d}}{\mathrm{d}x} x^n = nx^{n-1}$$

for all $n \in \mathbb{Z} \setminus \{0\}$. One question is if this holds for real or rational n as well. We will answer this question later.

We will not be proving all the basic identities. Assume the following to be true:

$$\frac{\mathrm{d}}{\mathrm{d}x}\sin x = \cos x \qquad \frac{\mathrm{d}}{\mathrm{d}x}\cos x = -\sin x \qquad \frac{\mathrm{d}}{\mathrm{d}x}e^x = e^x \qquad \frac{\mathrm{d}}{\mathrm{d}x}\ln x = \frac{1}{x}.$$

Theorem 1.3 (Carathéodory's theorem)**.** *Let I be an interval and $f: I \to \mathbb{R}$ and $c \in I$. Then $f'(c)$ exists iff there exists a function ϕ on I such that ϕ is continuous at c and for all $x \in I$,*

$$f(x) - f(c) = \phi(x)(x - c).$$

Proof.

$(\implies)$: Assume that $f'(c)$ exists. Then

$$\begin{aligned}\lim_{x\to c}\phi(x) &= \lim_{x\to c}\frac{f(x)-f(x)}{x-c}\\ &= f'(c).\end{aligned}$$

So we define $\phi(c) = f'(c)$ and ϕ is continuous at c.

$(\impliedby)$: Assume that ϕ is continuous at $x = c$. Then

$$\phi(c) = \lim_{x\to c}\phi(x) = \lim_{x\to c}\frac{f(x)-f(c)}{x-c}$$

which means that $f'(c) = \phi(c)$ and so it exists. ■

Theorem 1.4 (Chain Rule)**.** *Let I and J be intervals. Let $g: I \to \mathbb{R}$ and $f: J \to \mathbb{R}$ be functions such that $f[J] \subseteq I$. If f is differentiable at $a \in J$ and g is differentiable at $f(a)$, then $h = g \circ f$ is differentiable at a and*

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

Proof. Let $b = f(a)$. Given that $f'(a)$ and $g'(b)$ exists, we want to show that $h'(a) = g'(b)f'(a)$. By Carathéodory's theorem, there exists a function $\phi: J \to \mathbb{R}$ and $\psi: I \to \mathbb{R}$ such that firstly

$$f(x) - f(a) = \phi(x)(x - a)$$

for all $x \in J$, ϕ is continuous at a, and $\phi(a) = f'(a)$. Also,

$$g(y) - g(b) = \psi(y)(x - a)$$

for all $y \in J$, ψ is continuous at b, and $\psi(a) = g'(b)$.

Since $f[J] \subseteq I$, for any $x \in J$ we have $y = f(x) \in I$. Now

$$\begin{aligned} g(f(x)) - g(f(a)) &= \phi(f(x))(f(x) - f(a))\\ &= [\phi(f(x))\phi(x)](x - a)\\ h(x) - h(a) &= \alpha(x)(x - a)\end{aligned}$$

We have cast this into the form of Carathéodory's theorem. What is left is to check if α is continuous at a. Since ψ and f are both continuous at a, $\phi \circ f$ is continuous at a. We also know that ϕ is continuous at a. Therefore, $\alpha(x) = (\phi \circ f)(x)\phi(x)$ is continuous at a. ■

An alternative and the more conventional way of writing the chain rule is obtained if we let $u = f(x)$:

$$(g \circ f)'(x) = g'(f(x))f'(x) = g'(u)u' = \frac{\mathrm{d}y}{\mathrm{d}u}\frac{\mathrm{d}u}{\mathrm{d}x}.$$

Example 1.4. Let $f\colon \mathbb{R} \to \mathbb{R}$ be defined by

$$f(x) = \begin{cases} x^2 \sin\left(\frac{1}{x}\right), & \text{if } x \neq 0 \\ 0, & \text{otherwise} \end{cases}$$

At points $x \neq 0$, $f(x)$ is differentiable. Using the chain rule (skipping some steps),

$$\begin{aligned} f'(x) &= \frac{\mathrm{d}}{\mathrm{d}x}\left[x^2 \sin\left(\frac{1}{x}\right)\right] \\ &= 2x \sin\left(\frac{1}{x}\right) - x^2(\cos\frac{1}{x})(\frac{1}{x^2}) \end{aligned}$$

At 0 we will have to do it the hard way.

$$\begin{aligned} f'(0) &= \lim_{x\to 0} \frac{x^2 \sin\frac{1}{x} - 0}{x - 0} \\ &= \lim_{x\to 0} x \sin\frac{1}{x} \end{aligned}$$

Now

$$0 \leq \left|x \sin\frac{1}{x}\right| \leq |x|$$

so by the Squeeze theorem, the previous limit goes to 0 as $x \to 0$. So f is differentiable on $\mathbb{R}$. Is is continuously differentiable? f' is obviously continuous on $x \neq 0$. We need to show if it is true that $\lim_{x\to 0} f'(x) = 0$.

Create a sequence (x_n) where $x_n = \frac{1}{2n\pi}$. Then $\lim_{n\to\infty} x_n = 0$. Furthermore

$$\begin{aligned} \lim_{n\to\infty} f'(x_n) &= \lim_{n\to\infty}(2x_n \sin 2n\pi - \cos 2n\pi) \\ &= -1 \end{aligned}$$

so it is not continous. ◊

Theorem 1.5. *Suppose that I is an interval, $f\colon I \to \mathbb{R}$ is strictly monotone and continuous on I.*

Then $J = f[I]$ is also an interval, and the inverse function $g = f^{-1}\colon J \to \mathbb{R}$ exists and $g(f(x)) = x$ for all $x \in I$. g is also strictly monotone and continuous on J.

Theorem 1.6 (Inverse function theorem)**.** *Suppose that I is an interval, $f\colon I \to \mathbb{R}$ is strictly monotone and continuous on I, and f is differentiable at $c \in I$ and $f'(c) \neq 0$.*

Then $g = f^{-1}$ is differentiable at $f(c)$ and [1]

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

[1] An equivalent and perhaps more convenient form is $(f^{-1})'(b) = \frac{1}{f'(f^{-1}(b))}$.

Proof. Since $f'(c)$ exists, by Carathéodory's theorem, there exists a function $\phi\colon I \to \mathbb{R}$ such that

$$f(x) - f(c) = \phi(x)(x - c)$$

for all $x \in I$, ϕ is continuous at c, and $\phi(c) = f'(c) \neq 0$.

From theorem 5.1, there exists an interval $V = (c - \delta, c + \delta)$ such that

$$\forall x \in V\,[\phi(x) \neq 0].$$

Let $U = f[V]$. U is an interval and $d = f(c) \in U$. Let $x = g(y)$. Then, firstly

$$y - d = f(g(y)) - f(g(d)) = \phi(g(y))(g(y) - g(d)).$$

Secondly, $\phi(g(y)) \neq 0$ since $g(y) \in V$. Rearranging,

$$g(y) - g(d) = \frac{1}{\phi(g(y))}(y - d)$$

g is continuous at d by theorem 1.5. ϕ is continuous at c by supposition. Therefore $\frac{1}{\phi \circ g}$ is continuous at d, and by Carathéodory's theorem, g is differentiable at d and

$$g'(d) = \frac{1}{\phi(g(d))} = \frac{1}{\phi(c)} = \frac{1}{f'(c)}.$$

■

What happens when $f'(c) = 0$ instead?

Theorem 1.7. *Suppose that I is an interval, $f\colon I \to \mathbb{R}$ be strictly monotone and continuous on I, and f is differentiable at $c \in I$ and $f'(c) = 0$. Then $g = f^{-1}$ is not differentiable at $f(c)$.*

Proof. Let $d = f(c)$. Suppose that g was differentiable at d. Then, by the chain rule,

$$(g \circ f)'(c) = g'(f(c))f'(c) = 0$$

However this is clearly not true since $(g \circ f)(x) = x$ and we know the derivative of x should be 1, not 0. ■

Example 1.5. Let $r \in \mathbb{Q}^+$, and let $f(x) = x^r$ for $x > 0$. Now we can show that $f'(x) = rx^{r-1}$.

Write $r = \frac{m}{n}$, where $m, n \in \mathbb{N}$. Then $f = g \circ h$ where $g(x) = x^m$ and $h(x) = x^{\frac{1}{n}}$. We have already established earlier that $g'(x) = mx^{m-1}$. Now $H = h^{-1}(x) = x^n$ and thus $H' = nx^{n-1}$. By the inverse function theorem,

$$h'(x) = \frac{1}{H'(h(x))} = \frac{1}{nx^{\frac{n-1}{n}}} = \frac{1}{n}x^{\frac{1-n}{n}}$$

Using the chain rule,

$$f'(x) = g'(h(x))h'(x) = mx^{\frac{m-1}{n}}\frac{1}{n}x^{\frac{1-n}{n}} = \frac{m}{n}x^{\frac{m-n}{n}} = rx^{r-1}.$$

◊

Theorem 1.8. *Suppose I is an open interval and $f\colon I \to \mathbb{R}$ a continuous function. Define $g(x) = |f(x)|$ for all $x \in I$. If g is differentiable at $c \in I$, then f is also differentiable at c.*

Proof. We will consider the three cases $f(c) > 0$, $f(c) < 0$, and $f(c) = 0$.

For $f(c) > 0$, since f is continuous at c so by theorem 5.1 there is an interval $J = (c-\delta, c+\delta)$ such that $f(x) > 0$ for all $x \in J$. Then

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

It can thus be seen that for $f(c) < 0$ we get $f'(c) = -g'(c)$ using the same method.

For $f(c) = 0$, it is the same as above but let us work out explicitly the value of the derivative.

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

However, we can see that since $g(x) \geq 0$ for all x,

$$\lim_{x\to c^+} \frac{g(x)}{x-c} > 0 \qquad\qquad \lim_{x\to c^-} \frac{g(x)}{x-c} < 0$$

therefore $g'(c) = f'(c) = 0$. ∎

Theorem 1.9 (Straddle lemma)**.** *Let I be an interval and $f\colon I \to \mathbb{R}$ be differentiable at $c \in I$. Then $\forall\epsilon > 0, \exists\delta > 0, \forall u, v \in I$:*

$$c-\delta < u \leq c \leq v < c+\delta \implies |f(v) - f(u) - (v-u)f'(c)| \leq \epsilon(v-u)$$

Proof. Let $\epsilon > 0$. Since f is differentiable at c, the limit $\lim_{x\to c} \frac{f(x)-f(c)}{x-c}$ exists and so

$$\exists\delta > 0 \left[|x-c| < \delta \implies \left| \frac{f(x)-f(c)}{x-c} - f'(c) \right| < \epsilon \right]$$

Pick u and v such that $c-\delta < u \leq c \leq v < c+\delta$. Then

$$\begin{aligned} |f(c) - f(u) - (c-u)f'(c)| &< \epsilon(c-u) \\ |f(v) - f(c) - (v-c)f'(c)| &< \epsilon(v-c) \end{aligned}$$

Therefore

$$\begin{aligned} |f(v) - f(u) - (v-u)f'(c)| &= |f(v) - f(c) - (v-c)f'(c) + f(c) - f(u) - (c-u)f'(c)| \\ &< \epsilon(v-c) + \epsilon(c-u) \\ &= \epsilon(v-u) \end{aligned}$$

∎

1.2 Mean value theorem

After we have established some facts about the derivative, we want to know what it is good for. We know some applications of derivatives in basic calculus. We will establish some of these facts in the following sections.

Definition 1.6 (Absolute and relative maximums)**.** Let I be an interval, and $f\colon I \to \mathbb{R}$ be a function and $x_0 \in I$. Then we say that $f(x_0)$ is an *absolute maximum* of f on I if $\forall x \in I\, [f(x_0) \geq f(x)]$. We say that $f(x_0)$ is a *relative maximum* if $\exists\delta > 0, \forall x \in (x_0-\delta, x_0+\delta)\ [f(x) \leq f(x_0)]$. □

Absolute and relative minimums can be defined similarly. If $f(x_0)$ is either a relative (absolute) maximum or minimum, then we call it a relative (absolute) *extremum.* Note that relative extrema can only occur at an interior point but absolute extrema can occur at the end points of the interval. Therefore an absolute extremum may not necessarily be relative extremum.

Theorem 1.10. *Let $f\colon (a,b) \to \mathbb{R}$ and suppose $f'(c)$ exists for some $c \in (a,b)$.*

i. If $f'(c) > 0$, then $\exists \delta > 0, \forall x \in (c-\delta, c), \forall y \in (c, c+\delta)\ [f(x) < f(c) < f(y)]$.

ii. If $f'(c) < 0$, then $\exists \delta > 0, \forall x \in (c-\delta, c), \forall y \in (c, c+\delta)\ [f(x) > f(c) > f(y)]$.

Proof. The proof for (ii) is similar to (i) so we will only prove (i).

Since $f'(c) = \lim_{x\to c} \frac{f(x)-f(c)}{x-c} > 0$, using theorem 1.5,

$$\exists \delta > 0 \left[0 < |x-c| < \delta \implies \frac{f(x)-f(c)}{x-c} > 0 \right].$$

Taking $x \in (c-\delta, c)$, since $x - c > 0$, we conclude that $f(x) - f(c) < 0$ and $f(x) < f(c)$. Taking $x \in (c, c+\delta)$, $x - c < 0$, and following the same reasoning we arrive at $f(x) > f(c)$. ■

Note that the above conditions does not allow us to conclude that f is an increasing function in a neighbourhood of c. It only compares $f(c)$ to the surrounding points, but says nothing about the nature of those points. The following example illustrates this,

Example 1.6. Let $f\colon \mathbb{R} \to \mathbb{R}$ be defined by

$$f(x) = \begin{cases} x + 2x^2 \cos\left(\frac{1}{x}\right), & \text{if } x \neq 0 \\ 0, & \text{otherwise} \end{cases}$$

We have at $x = 0$

$$\begin{aligned} f'(0) &= \lim_{h\to 0} \frac{f(h) - f(0)}{h} \\ &= \lim_{h\to 0} \frac{h + 2h^2 \cos\left(\frac{1}{h}\right)}{h} \\ &= \lim_{h\to 0} 1 + 2h \cos\left(\frac{1}{h}\right) \\ &= 1 > 0. \end{aligned}$$

However, there is no neighbourhood around $x = 0$ such that f is increasing. Suppose instead that there was such a neighbourhood and let it be $(-\delta, \delta)$. Consider the sequence $(s_n) = \left(\frac{1}{(2n-\frac{1}{2})\pi}\right)$. Then,

$$\begin{aligned} f'(s_n) &= 1 + \frac{4}{(2n-\frac{1}{2})\pi} \cos\left(\left(2n - \frac{1}{2}\right)\pi\right) + 2\sin\left(\left(2n - \frac{1}{2}\right)\pi\right) \\ &= 1 - 2 = -1. \end{aligned}$$

Now since $\lim_{n\to\infty}(s_n) = 0$, there exists some element $s_m \in (-\delta, \delta)$. However, the previous calculation also shows us that since $f'(s_m) < 0$, there exists some α such that for all $y \in (s_m, s_m + \alpha) \subset (-\delta, \delta)$, we have $f(s_m) < f(y)$, or in other words, f is not increasing there. ◊

Theorem 1.11 (Interior extremum theorem)**.** *Let $f\colon (a,b) \to \mathbb{R}$ and suppose $f'(c)$ exists for some $c \in (a,b)$. If $f'(c) \neq 0$, then $f(c)$ cannot be a relative extremum. Equivalently, if $f(c)$ is a relative extremum, then $f'(c) = 0$.*

Proof. Follows directly. ■

Note that the converse is not true, i.e. $f'(c) = 0$ does not always give us a relative extremum. Inflection points serve as a counterexample. Furthermore, $f(c)$ can be a relative extremum but $f'(c)$ might not exist.

Using what we have we can also prove the following fun fact. This also serves as a counterexample for the converse of the intermediate value theorem, since it clearly shows that just satisfying the intermediate value theorem alone is not enough to make something continuous.

Theorem 1.12 (Intermediate value theorem for derivatives)**.** *Let $f\colon [a,b] \to \mathbb{R}$ be differentiable on $[a,b]$ and that $f'(a) < f'(b)$. Suppose there is a $k \in \mathbb{R}$ such that $f'(a) < k < f'(b)$. Then there exists a $c \in (a,b)$ such that $f'(c) = k$.*

Proof. Consider the function $g(x) = f(x) - kx$. Since $f(x)$ and kx are differentiable, $g(x)$ is as well and we have $g'(x) = f'(x) - k$.

Since $f'(a) < k$, so $g'(a) < 0$ and theorem 1.10 tells us that $g(a)$ cannot be an absolute minimum. Similarly, since $k < f'(b)$, we have $g'(b) > 0$, and theorem 1.10 tells us that $g(b)$ cannot be an absolute minimum either. But the extreme value theorem says that g has to have an absolute minimum at some point $z \in (a,b)$. This means $g'(z) = 0$, which then means that $f'(z) = k$. ■

Theorem 1.13 (Rolle's theorem)**.** *If f is continuous on $[a,b]$ and differentiable on (a,b) and $f(a) = f(b)$, then there exists $c \in (a,b)$ such that $f'(c) = 0$.*

Proof. The case when f is a constant function is trivially true.

The second case is when f is not a constant function. Then by the extreme value theorem, there exists $x_1, x_2 \in [a,b]$ such that $\forall x \in [a,b]\ f(x_1) \leq f(x) \leq f(x_2)$. Since f is not constant, $f(x_1) \neq f(x_2)$. However since $f(a) = f(b)$, either x_1 or x_2 is in (a,b), call it c. So f has a relative extremum at c. By the interior extremum theorem, $f'(c) = 0$. ■

We can generalize this theorem to relax the condition of $f(a) = f(b)$. This is done by transforming the function in question until we can satisfy Rolle's theorem.

Theorem 1.14 (Mean value theorem)**.** *If f is continuous on $[a,b]$ and differentiable on (a,b), then*

$$\exists c \in (a,b) \left[f'(c) = \frac{f(b) - f(a)}{b - a} \right].$$

Proof. Define $g\colon [a,b] \to \mathbb{R}$ by $g(x) = f(x) - f(a) - \frac{f(b)-f(a)}{b-a}(x - a)$. It is continuous on $[a,b]$ and differentiable on (a,b). Also, $g(a) = 0 = g(b)$. By Rolle's theorem, $\exists c \in (a,b)\ g'(c) = 0$. This means

$$g'(c) = f'(c) - \frac{f(b) - f(a)}{b - a} = 0.$$

■

Example 1.7. We will use the mean value theorem to show that $e^x \geq 1 + x$.

This is clearly true for $x = 0$.

Consider the case when $x > 0$. Apply the mean value theorem to $f(x) = e^x$ on the interval $[0, x]$. So $\exists c \in (0, x)$ $f'(c) = \frac{f(x)-f(0)}{x}$. Evaluating, we get $e^x = 1 + e^c x$. Since $e^c > 1$, we get $e^x > 1 + x$. The case where $x < 0$ is similar. ◊

Example 1.8. We will use the mean value theorem to show that $\sqrt{1+x} < 1 + \frac{x}{2}$ for all $x > 0$.

Apply the mean value theorem to $f(x) = \sqrt{1+x}$ on the interval $[0, x]$. So $\exists c \in (0, x)$ such that

$$\begin{aligned} f'(c) &= \frac{f(x) - f(0)}{x} \\ \frac{1}{2\sqrt{1+c}} &= \frac{\sqrt{1+x} - 1}{x} \\ \frac{x}{2} &= \sqrt{1+c}\left(\sqrt{1+x} - 1\right) \\ &< \sqrt{1+x}\left(\sqrt{1+x} - 1\right) \\ 1 + \frac{x}{2} &> \sqrt{1+x}. \end{aligned}$$

◊

Theorem 1.15. *If f is continuous on $[a, b]$ and differentiable on (a, b), and $f'(x) = 0$ for all $x \in (a, b)$, then f is constant on $[a, b]$.*

Proof. If $a < x \leq b$ then by the mean value theorem $\exists c \in (a, x)$ $f'(c) = \frac{f(x)-f(a)}{x-a} = 0$. This means $f(x) = f(a)$. ■

Theorem 1.16. *If f is differentiable and its derivative is bounded on $[a, b]$ then f is Lipschitz on $[a, b]$.*

Proof. Let $K \geq |f'(x)|$ for all $x \in [a, b]$. For any $x, y \in [a, b]$ the mean value theorem applied to f on $[x, y]$ states that $\exists c \in (x, y)$

$$\begin{aligned} f'(c) &= \frac{f(x) - f(y)}{x - y} \\ f(x) - f(y) = f'(c)(x - y) &\leq K(x - y) \end{aligned}$$

■

Theorem 1.17. *Let f be differentiable on (a, b). For all $x \in (a, b)$,*

i. $f'(x) \geq 0$ *iff f is increasing on (a, b).*

ii. $f'(x) \leq 0$ *iff f is decreasing on (a, b).*

Proof. We only prove (i) since the proof for (ii) is similar.

($\Longrightarrow$): Let $a < x_1 < x_2 < b$. Apply the mean value theorem to f on $[x_1, x_2]$. Then $\exists c \in (x_1, x_2)$ $f'(x) = \frac{f(x_2)-f(x_1)}{x_2-x_1} \geq 0$. Since $x_2 - x_1 \geq 0$, we get that $f(x_2) \geq f(x_1)$.

($\Longleftarrow$): We have

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

For $h < 0$, we have $x + h < x$ and so $f(x+h) - f(x) \leq 0$. Therefore the term in the limit is positive. For $h > 0$, we have $x + h > x$ and so $f(x+h) - f(x) \geq 0$. Therefore the term in the limit is also positive. Hence, the limit is always positive and thus $f'(x) \geq 0$. ■

There is also a strict version but its converse is not true:

Theorem 1.18. *Let f be differentiable on (a, b). If, for all $x \in (a, b)$,*

i. If $f'(x) > 0$, then f is strictly increasing on (a, b).

ii. If $f'(x) < 0$, then f is strictly decreasing on (a, b).

Theorem 1.19 (First derivative test)**.** *Let f be continuous on $[a, b]$ and $c \in (a, b)$. Suppose that f is differentiable on (a, b) except possibly at c.*

i. If there is a neighbourhood $(c - \delta, c + \delta) \subseteq I$ of c such that $f'(x) \geq 0$ for all $x \in (c - \delta, c)$ and $f'(x) \leq 0$ for $x \in (c, c + \delta)$, then f has a relative maximum at c.

ii. A similar statement can be made for relative minimums.

Proof. We will only prove (i), the proof for (ii) is similar.

Let $x \in (c - \delta, c)$. Applying the mean value theorem to f on the interval $[x, c]$, $\exists x_0 \in (x, c)\ f'(x_0) = \frac{f(c)-f(x)}{c-x}$. By assumption $f'(x_0) \geq 0$ and $c - x > 0$, so $f(c) - f(x) \geq 0$ and hence $f(c) \geq f(x)$.

Let $x \in (c, c + \delta)$. Applying the mean value theorem again now on the interval $[c, x]$, $\exists x_1 \in (c, x)\ f'(x_1) = \frac{f(x)-f(c)}{x-c}$. A similar reasoning allows us to conclude that $f(x) \leq f(c)$. ■

Theorem 1.20 (Cauchy's mean value theorem)**.** *Let f and g be continuous on $[a, b]$ and differentiable on (a, b), and assume that $g'(x) \neq 0$ for all $x \in (a, b)$. Then*

$$\exists c \in (a, b) \left[\frac{f(b) - f(a)}{g(b) - g(a)} = \frac{f'(c)}{g'(c)} \right].$$

Proof. First, $g(a) \neq g(b)$ because Rolle's theorem tells us that $g'(x_0) = 0$ at some point $x_0 \in (a, b)$.

Next, let $h(x) = \frac{f(b)-f(a)}{g(b)-g(a)}(g(x) - g(a)) - (f(x) - f(a))$. Notice that $h(a) = h(b) = 0$, and that $h'(x) = \frac{f(b)-f(a)}{g(b)-g(a)} g'(x) - f'(x)$. Applying Rolle's theorem to $h(x)$, $\exists c \in (a, b)\ h'(c) = 0$, and we get the result we are after. ■

Theorem 1.21 (L'Hôspital's rule)**.** *Let f and g be differentiable on (a, b) and assume that $g(x) \neq 0$ and $g'(x) \neq 0$ for all $x \in (a, b)$. Suppose that $\lim_{x\to a^+} f(x) = \lim_{x\to a^+} g(x) = 0$. Then, if $\lim_{x\to a^+} \frac{f'(x)}{g'(x)} = L$ for $L \in \mathbb{R} \cup \{\pm\infty\}$, then $\lim_{x\to a^+} \frac{f(x)}{g(x)} = L$.*

Proof. First consider the case $L \in \mathbb{R}$. Let $\epsilon > 0$. Then $\exists \delta > 0 \left[|x - a| < \delta \implies \left| \frac{f'(x)}{g'(x)} - L \right| < \frac{\epsilon}{2} \right]$. Now let $a < x < y < a + \delta$. We can apply Cauchy's mean value theorem to the interval (x, y). So $\exists z \in (x, y)$ $\frac{f(y)-f(x)}{g(y)-g(x)} = \frac{f'(z)}{g'(z)}$. This means that $\left| \frac{f(y)-f(x)}{g(y)-g(x)} - L \right| < \frac{\epsilon}{2}$. Now, consider $M = \lim_{y \to a^+} \frac{f(y)-f(x)}{g(y)-g(x)}$. We have from the above $L - \epsilon < L - \frac{\epsilon}{2} \leq M \leq L + \frac{\epsilon}{2} < L + \epsilon$. We arrive at the conclusion that $x - a < \delta \implies \left| \frac{f(x)}{g(x)} - L \right| < \epsilon$.

Next, suppose that $L = \infty$. So

$$\exists \delta > 0, \forall k \left[|x - a| < \delta \implies k < \frac{f'(x)}{g'(x)} \right].$$

Similar as above, Cauchy's mean value theorem gives us $z \in (x, y)$ such that $k < \frac{f(y)-f(x)}{g(y)-f(x)}$. Therefore, we have $\lim_{y \to a^+} \frac{f(y)-f(x)}{g(y)-g(x)} > k$ and we conclude that $x - a < \delta \implies k < \frac{f(x)}{g(x)}$. ■

There is a similar case for infinite limits. We note that we can convert a limit of infinity to one of zero by taking reciprocals.

Theorem 1.22 (L'Hôpital's rule)**.** *Let f and g be differentiable on (a, b) and assume that $g'(x) \neq 0$ for all $x \in (a, b)$. Suppose that* $\lim_{x \to a^+} f(x) = \lim_{x \to a^+} g(x) = \infty$. *Then, if* $\lim_{x \to a^+} \frac{f'(x)}{g'(x)} = L$ *for* $L \in \mathbb{R} \cup \{\pm\infty\}$, *then* $\lim_{x \to a^+} \frac{f(x)}{g(x)} = L$.

Though we only state the Rule for $x \to a^+$, we can also write similar versions for $x \to a^-$, $x \to a$, and $x \to \pm\infty$.

1.3 Higher order derivatives

If f is differentiable, then its derivative f' is a function. So we can consider the differentiability of f' as well.

Definition 1.7 (Higher derivatives.)**.** We denote the derivative of f' as the second derivative of f, and write f'' or $f^{(2)}$. We can keep going with $f''' = f^{(3)}$, $f'''' = f^{(4)}$ and so on.

Furthermore, let the set $C^n(I) = \{f \mid f^{(n)}$ exists and is continuous on $I\}$. Also let $C^\infty(I) = \bigcap_{n=1}^{\infty} C^n(I)$, and we call it the set of all *infinitely differentiable* or *smooth* functions. □

$C^0(I) = C(I)$ is the set of all continuous functions on I. We have a natural hierarchy

$$C^\infty(I) \subset C^n(I) \subset C(I).$$

Theorem 1.23 (Second derivative test)**.** *Let f and its derivative be defined on an interval I. Suppose that c is an interior point of I such that $f'(c) = 0$ and $f''(c)$ exists.*

i. If $f''(c) > 0$, then $f(c)$ is a relative maximum.

ii. If $f''(c) < 0$, then $f(c)$ is a relative minimum.

Proof. We will only prove (i).

Applying theorem 1.10 to f', we can make the conclusion that $\exists\delta > 0, \forall x \in (c-\delta, c), \forall y \in (c, c+\delta)\ f'(x) < f'(c) = 0 < f'(y)$. By the first derivative test this tells us that $f(c)$ is a relative maximum. ■

The theorem does not say anything about $f''(c) = 0$. Again, $f(x) = x^3$ serves as an illustration. The second derivative test is easier to use but less powerful than the first derivative test. For example for $f(x) = x^4$ at $x = 0$ it is also inconclusive, yet the first derivative test gives us an answer since f' changes sign around $x = 0$.

Theorem 1.24. *Suppose that the function $f\colon [a,b] \to \mathbb{R}$ is continuous on $[a,b]$ and f'' exists on (a,b). The graph of f and the line segment joining point a to point b intersects at at least one point. Then $\exists c \in (a,b)\ f''(c) = 0$.*

Proof. Let the point of intersection be $(x, f(x))$, and let the gradient of the line segment be $m = \frac{f(b)-f(a)}{b-a}$. Apply the mean value theorem twice on the intervals $[a,x]$ and $[x,b]$. This tells us that there exists two points, $x_1 \in (a,x)$ and $x_2 \in (x,b)$ such that $f'(x_1) = m = f'(x_2)$. Then, using Rolle's theorem on f' on the interval $[x_1, x_2]$ we obtain our conclusion. ■

Theorem 1.25 (Taylor's theorem)**.** *Let f be a function such that $f \in C^n([a,b])$ and $f^{(n+1)}$ exists on (a,b). If $x_0 \in [a,b]$, then for any $x \in [a,b]$ there exists a point c between x and x_0 such that*

$$f(x) = f(x_0) + f'(x_0)(x-x_0) + \frac{f''(x_0)}{2!}(x-x_0)^2 + \cdots + \frac{f^{(n)}(x_0)}{n!}(x-x_0)^n + \frac{f^{(n+1)}(c)}{(n+1)!}(x-x_0)^{n+1}$$

Proof. Take $x \in [a,b]$, and assume $x_0 \neq x$. Let M be the unique number satisfying the following equation.

$$f(x) = f(x_0) + f'(x_0)(x-x_0) + \cdots + \frac{f^{(n)}(x_0)}{n!}(x-x_0)^n + M(x-x_0)^{n+1}$$

Define $F\colon [a,b] \to \mathbb{R}$ given by $F(t) = f(t) + f'(t)(x-t) + \cdots + \frac{f^{(n)}(x_0)}{n!}(x-x_0)^n + M(x-t)^{n+1}$. F is continuous on $[a,b]$ and differentiable on (a,b). Furthermore, $F(x) = f(x) = F(x_0)$. So by Rolle's theorem, $\exists c \in (x, x_0)\ F'(c) = 0$. Now we compute the derivative of F:

$$\begin{aligned}
F'(t) &= f'(t) + [f''(t)(x-t) - f'(t)] + \left[\frac{f'''(t)}{2!}(x-t)^2 - \frac{f''(t)}{2!}2(x-t)\right] + \cdots + \\
&\quad \left[\frac{f^{(n+1)}(t)}{n!}(x-t)^n - \frac{f^{(n)}(t)}{n!}n(x-t)^{n-1}\right] - M(n+1)(x-t)^n \\
0 &= \frac{f^{(n+1)}(t)}{n!}(x-t)^n - M(n+1)(x-t)^n
\end{aligned}$$

Simplifying the final result gives us the desired form for M. ■

The polynomial

$$P_{n(x)} = f(x_0) + f'(x_0)(x-x_0) + \frac{f''(x_0)}{2!}(x-x_0)^2 + \cdots + \frac{f^{(n)}(x_0)}{n!}(x-x_0)^n$$

is called the n-th *Taylor polynomial* for f at x_0. It has the property that $P_n^{(j)}(x_0) = f^{(i)}(x_0)$. We can use it to estimate the value of f at points near x_0. The error of this estimation is given by

$$R_n(x) = f(x) - P_n(x) = \frac{f^{(n+1)}(c)}{(n+1)!}(x - x_0)^{n+1}$$

where $c \in (x, x_0)$. This is called the *Lagrange form* of the remainder.

Example 1.9. We want to show that $\cos x \geq 1 - \frac{1}{2}x^2$ for all $x \in \mathbb{R}$. Let $f(x) = \cos x$ and set $x_0 = 0$. Applying Taylor's theorem to f with $n = 2$, we get $f(x) = f(0) + f'(0)(x) + \frac{f''(0)}{2!}x^2 + R_2(x)$ where $R_2(x) = \frac{f'''(c)}{3!}x^3$. Now,

$$f(0) = 1 \qquad f'(0) - \sin 0 = 0 \qquad f''(0) = -\cos 0 = -1$$

so $f(x) = 1 - \frac{x^2}{2} + R_2(x)$. However we can see that R_2 is always positive if $|x| < \pi$. If $|x| \geq \pi$, then since $\pi < 3$ and the range of cos is only $[-1, 1]$, we have $f(x) > 1 - \frac{3^2}{2}$ ◊

Example 1.10. Consider $f(x) = e^x$. Since the derivative of f is just f itself, we can write $R_n(x) = \frac{f(c_n)}{(n+1)!}(x - x_0)^{n+1}$. Let I be the closed interval with end points x_0 and x. Since f is continuous on I and continuous functions on closed intervals are bounded, so $\exists M > 0 \ |f(u)| \leq M$. Thus $|R_n(x)| \leq \frac{M}{(n+1)!}|x - x_0|^{n+1}$. Call the term on the right y_n.

We claim that $(y_n) \to 0$. Performing the ratio test,

$$\lim_{n\to\infty}\left|\frac{y_{n+1}}{y_n}\right| = \lim_{n\to\infty}|x - x_0|(n+1) = 0$$

so by the squeeze theorem we find that $(R_n(x)) \to 0$.

Using this fact, we arrive at the conclusion that

$$e^x = \lim_{n\to\infty}(P_n(x) + R_x(x)) = e^{x_0}\sum_{n=0}^{\infty}\frac{(x - x_0)^n}{n!}$$

and more famously with $x_0 = 0$ we get $f(x) = \sum_{n=0}^{\infty}\frac{x^n}{n!}$. ◊

Example 1.11. It should be noted that a function being smooth does not automatically imply that it is equal to its own Taylor series. In fact, it is not true in general. Consider the following function $h \in C^\infty$:

$$h(x) = \begin{cases} e^{-\frac{1}{x^2}}, & \text{if } x \neq 0 \\ 0, & \text{otherwise} \end{cases}.$$

For the case $x \neq 0$, we have

$$h'(x) = \frac{2}{x^3}e^{-\frac{1}{x^2}}$$

For $x = 0$, we have using L'Hôspital's rule and the Squeeze theorem,

$$\begin{aligned} h'(0) &= \lim_{x\to 0}\frac{e^{-\frac{1}{x^2}} - 0}{x - 0} \\ &= \lim_{x\to 0}\frac{e^{-\frac{1}{x^2}}}{x} \\ &= 0. \end{aligned}$$

In fact, $h^{(j)}(0) = 0$ for all j. Then by Taylor's theorem,

$$P_n(x) = h(0) + h'(0)x + \frac{h''(0)}{2!}x^2 + \cdots = 0.$$

Therefore we have this odd situation where $R_n(x) = f(x)$! Therefore the function does not converge to its own Taylor series. ♢

Theorem 1.26 (n-th derivative test). *Let I be an interval. Suppose $f: I \to \mathbb{R}$ has continuous derivatives $f, f', \ldots f^{(n)}$ such that for some $x_0 \in I$, $f'(x_0) = \cdots = f^{(n-1)}(x_0) = 0$. Then*

i. *If n is even and $f^{(n)}(x_0) > 0$, then f has a relative minimum at x_0.*

ii. *If n is even and $f^{(n)}(x_0) < 0$, then f has a relative maximum at x_0.*

iii. *If n is odd, then f has neither a relative minimum nor a relative maximum at x_0.*

Proof.

i. ii. These are similar so we shall just prove (i). Since $f^{(n)}$ is continuous, there exists a neighbourhood $U = (x_0 - \delta, x_0 + \delta)$ such that $\forall x \in U (f^{(n)}(x) > 0)$. Using Taylor's theorem for any $x \in U$ we have

$$f(x) = f(x_0) + \frac{f^{(n)}(c)}{n!}(x - x_0)^n$$

for some $c \in (x, x_0)$. Since $c \in U$ and n is even, the right most term is positive and that gives us $f(x) \geq f(x_0)$ for all $x \in U$. Hence x_0 is a relative minimum.

iii. From the above, we can see that if n is odd then $(x - x_0)$ would have different signs for $x < x_0$ and $x > x_0$. Therefore x cannot be any form of extrema.

■

2 Integration

The first motivation we have for defining integration is to find the area under curves. The second is to be able to use it to construct nice functions.

2.1 Riemann integrals

We deal with the first motivation first. The standard approach is to approximate the area under the curve with increasingly smaller rectangles. The precise area is then the limit of this process.

Definition 2.1 (Partitions of intervals). Let $I = [a, b]$. A finite set $P = \{x_0, x_1, \ldots, x_n\}$ where $a = x_0 < x_1 < \cdots < x_n = b$ is called a *partition* of I. It divides the interval I into subintervals

$$I = [x_0, x_1] \cup [x_i, x_2] \cup \cdots \cup [x_{n-1}, x_n]$$

If P and R are partitions of $[a, b]$, and if $P \subseteq R$, then we say that R is a refinement of P. □

We first define some notation.

Definition 2.2. Let $f: [a, b] \to \mathbb{R}$ be a bounded function, and let $P = \{x_0, \dots .x_n\}$ be a partition of $[a, b]$. For $1 \le i \le n$ let

$$\begin{aligned} M_i &= \sup\{f(x) \mid x \in [x_{i-1}, x_i]\} \\ m_i &= \inf\{f(x) \mid x \in [x_{i-1}, x_i]\} \end{aligned}$$

Also let $\Delta x_i = x_i - x_{i-1}$. □

Definition 2.3 (Upper and lower sums)**.** Let $f: [a, b] \to \mathbb{R}$ be a bounded function, and let $P = \{x_0, \dots .x_n\}$ be a partition of $[a, b]$. Define the *upper sum* of f with respect to the partition P to be

$$U(f, P) = \sum_{i=1}^{n} M_i \Delta x_i.$$

Similarly, define the *lower sum* to be

$$L(f, P) = \sum_{i=1}^{n} m_i \Delta x_i.$$

□

Theorem 2.1. *Let $f: [a, b] \to \mathbb{R}$ be a bounded function and P be a partition of $[a, b]$. Also let $m = \inf f(x)$ and $M = \sup f(x)$. Then*

$$m(b - a) \le L(f, P) \le U(f, P) \le M(b - a).$$

Proof. Let $P = \{x_0, \dots, x_n\}$. First we note that $M_i \le M$ since M_i is only the supremum in a subinterval whereas M is the supremum in the entire interval $[a, b]$. Similarly $m_i \ge m$. Thus,

$$\begin{aligned} U(f, P) &= \sum_{i=1}^{n} M_i \Delta x_i \\ &\le \sum_{i=1}^{n} M(x_i - x_{i-1}) \\ &= M(b - a). \end{aligned}$$

The last step may be realised by writing out a few terms of the sum. Similarly we have

$$\begin{aligned} U(f, P) &\ge L(f, P) \\ &= \sum_{i=1}^{n} m_i \Delta x_i \\ &\ge \sum_{i=1}^{n} m(x_i - x_{i-1}) \\ &= m(b - a). \end{aligned}$$

■

More importantly, the above theorem tells us that $\sup U$ and $\inf L$ exists. Therefore, we are able to get the best over and under-estimates.

Definition 2.4 (Upper and lower intergrals)**.** Let $f\colon [a,b] \to \mathbb{R}$ be a bounded function. The *upper integral* of f on $[a,b]$ is defined to be

$${}_a^b f = U(f) = \inf\{U(f,P) \mid P \text{ partitions } [a,b]\}.$$

and the *lower integral* of f on $[a,b]$ is defined to be

$${}_a^b f = L(f) = \sup\{L(f,P) \mid P \text{ partitions } [a,b]\}.$$

□

Now, we have a theorem that allows us to improve our over and under-estimates by adding in more points to our partition.

Theorem 2.2. *Let $f\colon [a,b] \to \mathbb{R}$ be a bounded function and let P and Q be partitions of $[a,b]$. Then*

$$L(f,P) \leq L(f,P \cup Q) \leq U(f,P \cup Q) \leq U(f,Q).$$

Consequently,

$${}_a^b f \leq {}_a^b f.$$

Proof. Let $Q = \{q_0, \dots, q_n\}$. Now consider the new partition $R = Q \cup \{c\}$ for some $c \in P$, and suppose $q_{j-1} < c < q_j$. Let $y = \sup\{f(x) \mid x \in [q_{j-1}, c]\}$ and let $z = \sup\{f(x) \mid x \in [c, q_j\}$. Then, $y, z \leq M_j = \sup\{f(x) \mid x \in [q_{j-1}, q_j]\}$. Now if we calculate the new upper sum with respect to R, we get

$$\begin{aligned} U(f,R) &= \sum_{i \neq j} M_i \Delta x_i + y(c - q_{j-1}) + z(q_j - c) \\ &\leq \sum_{i \neq j} M_i \Delta x_i + M_j(q_j - q_{j-1}) \\ &= U(f,Q). \end{aligned}$$

We can then inductively show this for all the other points in P. We can also do the same for $L(f,P)$ and $L(f, P \cup Q)$ to get the final result. Furthermore, we also have

$$L(f,P) \leq U(f,Q)$$

for any partitions P and Q. In words, this means that any lower sum is smaller than any upper sum. Therefore, we have

$$L(f,P) \leq \inf\{U(f,Q)\} = {}_a^b f.$$

Since this forms an upper bound on all L, we have

$${}_a^b f = \sup\{L(f,P)\} \leq {}_a^b f.$$

■

Definition 2.5 (Integrability)**.** A bounded function $f\colon [a,b] \to \mathbb{R}$ is said to be *(Riemann) integrable* on $[a,b]$ if

$${}_a^b f = {}_a^b f.$$

Then, we define the *integral* of f on $[a,b]$ as

$$\int_a^b f = {}_a^b f = {}_a^b f.$$

□

Note that the Riemann integral is only defined for bounded functions.

Example 2.1 (Dirichlet function)**.** This is a famous non-integrable function. Let $g\colon [0,1] \to \mathbb{R}$ be defined by

$$g(x) = \begin{cases} 1, & \text{if } x \in \mathbb{Q} \\ 0, & \text{otherwise} \end{cases}.$$

Let $P = \{x_1, x_2, \dots, x_n\}$ be a partition of $[0,1]$. Since $\mathbb{Q}$ is dense in $\mathbb{R}$, we have $m_i(g,P) = 1$ and $M_i(g,P) = 0$. For the same reason we also have $L(g,P) = 0$ and $U(g,P) = 1$. This then means that ${}_0^1 g = 0 \neq {}_0^1 g = 1$. ◊

Example 2.2 (Identity function)**.** Let $h\colon [0,1] \to [0,1]$ be defined by $h(x) = x$. Let P_n be the partition $\{x_0, x_1, \dots, x_n\}$ where $n \in \mathbb{N}$ and $x_k = \frac{k}{n}$, and $\Delta x_k = \frac{1}{n}$. For a fixed n, since h is increasing, we have $M_k = h(x_k)$ and $m_k = h(x_{k-1})$ for all $1 \le k \le n$. Therefore,

$$U(h,P_n) = \sum_{k=1}^{n} \frac{k}{n}\frac{1}{n} = \frac{n+1}{2n}. \qquad L(h,P_n) = \sum_{k=1}^{n} \frac{k-1}{n}\frac{1}{n} = \frac{n-1}{2n}.$$

So we have $\frac{n-1}{2n} \le {}_0^1 h \le {}_0^1 h \le \frac{n+1}{2n}$ for all n. Taking the limit $n \to \infty$, we have $\int_0^1 h = \frac{1}{2}$. ◊

Theorem 2.3. *Let $f\colon [a,b] \to \mathbb{R}$ be bounded and integrable on $[a,b]$ and let $m = \inf\{f(x) \mid x \in [a,b]\}$ and $M = \sup\{f(x) \mid x \in [a,b]\}$. Then we have*

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

ii. If $f(x) > 0$ for all $x \in [a,b]$ then $\int_a^b f \ge 0$.

Proof.

i. Theorem 2.1 already tells us that $m(b-a) \le L(f,P) \le U(f,P) \le M(b-a)$. Since f is integrable, we also have $\int_a^b f = {}_a^b$. But since ${}_a^b = \inf U(f,P)$, we have $\int_a^b f \le U(f,P)$. Similarly for $L(f,P)$.

ii. If $f(x) \ge 0$ for all $x \in [a,b]$, then $m \ge 0$. From the above, we immediately get that $\int_a^b f \ge m(b-a) \ge 0$. ■

The previous examples show that checking integrability from the definition is quite tedious. Following is an easier way to check for integrability.

Theorem 2.4 (Riemann integrability criterion)**.** *Let $f\colon [a,b] \to \mathbb{R}$ be a bounded function. Then f is integrable on $[a,b]$ iff for any $\epsilon > 0$, there exists a partition P of $[a,b]$ such that*

$$U(f,P) - L(f,P) < \epsilon.$$

Proof.

($\Longrightarrow$): Assume that f is integrable on $[a,b]$. Let $\epsilon > 0$. Let $S = \{U(f,T) \mid T \text{ partitions } [a,b]\}$. Since ${}_a^b f = \inf S$, there is a partition Q such that $I(f,Q) < {}_a^b f + \frac{\epsilon}{2}$. Similarly, there exists a partition R such that $L(f,R) > {}_a^b - \frac{\epsilon}{2}$. Now let $P = Q \cup R$. From theorem 2.2, we have

$${}_a^b f - \frac{\epsilon}{2} < L(f,R) \le L(f,P) \le \int_a^b f \le U(f,P) \le U(f,Q) < {}_a^b f + \frac{\epsilon}{2}.$$

($\Longleftarrow$): Let $\epsilon > 0$. Then there exists a partition P such that $U(f, P) - L(f, P) < \epsilon$. We have ${}^{b}_{a} f - {}^{b}_{a} f \leq U(f, P) - L(f, P) < \epsilon$. Taking the limit $\epsilon \to 0$, we obtain ${}^{b}_{a} f - {}^{b}_{a} f = 0$. ■

Theorem 2.5. *If $f\colon [a, b] \to \mathbb{R}$ is monotone on $[a, b]$, then f is integrable on $[a, b]$.*

Proof. First assume that f is increasing. The proof for the case where f is decreasing is similar. Choose $n \in \mathbb{N}$ such that $n > \frac{(b-a)(f(b)-f(a))}{\epsilon}$. Let the partition $P_n = \{x_0, x_1, \dots, x_n\}$ be the partition that subdivides $[a, b]$ evenly, i.e. $\Delta x_i = \frac{b-a}{n}$. Similar to the case for the identity function, we have $M_i = f(x_i)$ and $m_i = f(x_{i-1})$. This gives us

$$U(f, P_n) = \sum_{i=1}^{n} f(x_i)\Delta x_i \qquad\qquad L(f, P_n) = \sum_{i=1}^{n} f(x_{i-1})\Delta x_i$$

And so

$$\begin{aligned} U(f, P_n) - L(f, P_n) &= \sum_{i=1}^{n} (f(x_i) - f(x_{i-1}) \frac{b-a}{n} \\ &= (f(b) - f(a)) \frac{b-a}{n} \\ &< \epsilon. \end{aligned}$$

■

Theorem 2.6. *If $f\colon [a, b] \to \mathbb{R}$ is continuous on $[a, b]$, then f is integrable on $[a, b]$.*

Proof. Let $\epsilon > 0$. Since f is continuous on the compact interval $[a, b]$, it is uniformly continuous on $[a, b]$. So $\exists \delta > 0, \forall x, y \in [a, b] \left[|x - y| < \delta \implies |f(x) - f(y)| < \frac{\epsilon}{b-a} \right]$. Now choose $n \in \mathbb{N}$ such that $n > \frac{b-a}{\delta}$, and let the partition $P_n = \{x_0, x_1, \dots, x_n\}$ be the partition that subdivides $[a, b]$ evenly, i.e. $\Delta x_i = \frac{b-a}{n} i < \delta$.

Consider the subinterval $[x_{i-1}, x_i]$. By the extreme value theorem, there exists some $u_i, v_i \in [x_{i-1}, x_i]$ such that $f(u_i) \leq f(x) \leq f(v_i)$. Since $|u_i - v_i| \leq |x_i - x_{i-1}| < \delta$. Therefore, due to uniform continuity, we have $|f(u_i) - f(v_i)| < \frac{\epsilon}{b-a}$. Thus

$$\begin{aligned} U(f, P_n) - L(f, P_n) &= \sum_{i=1}^{n} (M_i - m_i)\Delta x_i \\ &= \sum_{i=1}^{n} (f(v_i) - f(u_i))\Delta x_i \\ &= \sum_{i=1}^{n} \frac{\epsilon}{b-a} \frac{b-a}{n} \\ &= \epsilon. \end{aligned}$$

■

Theorem 2.7. *Let $f\colon [a, b] \to \mathbb{R}$ be a bounded function. Suppose there exists a sequence of partitions $(P_n)_n$ of $[a, b]$ such that the sequences $(L(f, P_n))_n$ and $U((f, P_n))_n$ both converge to the same value A. Then f is integrable and $\int_a^b f = A$.*

Proof. Since we have $L(f, P_n) \leq {}^{b}_{a} f \leq {}^{b}_{a} f \leq U(f, P_n)$, taking the limit $n \to \infty$ gives us $A \leq {}^{b}_{a} f \leq {}^{b}_{a} f \leq A$. ■

Lemma 2.8. *Let $f, g\colon [a,b] \to \mathbb{R}$ be bounded functions, P a partition of $[a,b]$ and $c \in \mathbb{R}$. Then*

i. $L(cf, P) = \begin{cases} cL(f,P), & \text{if } c > 0 \\ cU(f,P), & \text{if } c < 0 \end{cases}$

ii. $U(cf, P) = \begin{cases} cU(f,P), & \text{if } c > 0 \\ cL(f,P), & \text{if } c < 0 \end{cases}$

iii. $L(f,P) + L(g,P) \leq L(f+g,P) \leq U(f+g,P) \leq U(f,P) + U(g,P).$

Proof.

i. ii. We recall that for any non-empty $S \subseteq \mathbb{R}$,

$$\inf(cS) = \begin{cases} c\inf S, & \text{if } c > 0 \\ c\sup S, & \text{if } c < 0 \end{cases} \qquad \sup(cS) = \begin{cases} c\sup S, & \text{if } c > 0 \\ c\inf S, & \text{if } c < 0 \end{cases}$$

iii. Let $P = \{x_1, x_2, \ldots, x_n\}$. For $1 \leq i \leq n$ and $x \in [x_{i-1}, x_i]$, we have

$$m_i(f,P) \leq f(x) \leq M_i(f,P) \qquad m_i(g,P) \leq f(x) \leq M_i(g,P)$$

and so

$$m_i(f,P) + m_i(g,P) \leq (f+g)(x) \leq M_i(f,P) + M_i(g,P).$$

This means

$$m_i(f,P) + m_i(g,P) \leq m_i(f+g,P) \qquad M_i(f+g) \leq M_i(f,P) + M_i(g,P).$$

Therefore

$$\begin{aligned} L(f,P) + L(g,P) &= \sum_{i=1}^{n} (m_i(f,P) + m_i(g,P))(x_i - x_{i-1}) \\ &\leq \sum_{i=1}^{n} m_i(f+g,P)(x_i - x_{i-1}) \\ &= L(f+g,P) \\ &\leq U(f+g,P) \\ &\leq \sum_{i=1}^{n} (M_i(f,P) + M_i(g,P))(x_i - x_{i-1}) \\ &= U(f,P) + U(g,P). \end{aligned}$$

■

The following are basic facts that we have taken for granted all along. Parts (i) and (ii) show that the set of integrable functions form a real vector space. Part (v) shows that this set also possesses some ring structure. We will also use the following fact:

Lemma 2.9. *For a non-empty bounded $S \subseteq \mathbb{R}$, if there is a $k > 0$ such that $\forall s, t \in S\ |s - t| \leq k$ then we have $\sup S - \inf S \leq k$.*

Theorem 2.10 (Integration rules)**.** *Let $f, g\colon [a,b] \to \mathbb{R}$ be integrable on $[a,b]$ and $c \in \mathbb{R}$. Then* *i.* $\int_a^b cf = c\int_a^b f$.

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

iii. *If* $f(x) \le g(x)$ *for all* $x \in [a,b]$, *then* $\int_a^b f \le \int_a^b g$.

iv. $\left|\int_a^b f\right| \le \int_a^b |f|$.

v. fg *is integrable on* $[a,b]$.

Proof.

i. Assume $c > 0$. The proof for $c < 0$ is similar. By lemma 2.8, we have $U(cf,P) = U(f,P)$. Therefore

$$ {}^b_a cf = \inf U(cf,P) = \inf cU(f,P) = c \inf U(f,P) = c \, {}^b_a f = c\int_a^b f. $$

Similarly, ${}^b_a cf = \cdots = c\int_a^b f$. Therefore cf is integrable and $\int_a^b cf = c\int_a^b f$.

ii. First we show that ${}^b_a(f+g) \le {}^b_a f + {}^b_a g$. Let $\epsilon > 0$. Then there exists partitions P, Q such that $U(f,P) < {}^b_a f + \frac{\epsilon}{2}$, and $U(g,Q) < {}^b_a g + \frac{\epsilon}{2}$, since the upper integral is an infimum. Let $R = P \cup Q$, and by lemma 2.8,

$$ U(f+g,R) \le U(f,R) + U(g,R) \le U(f,P) + U(g,Q) < \int_a^b f + \int_a^b g + \epsilon. $$

Taking the limit $\epsilon \to 0$, we get ${}^b_a(f+g) \le \int_a^b f + \int_a^b g$.

We can also show that ${}^b_a(f+g) \ge {}^b_a f + {}^b_a g$ through similar means. Putting it all together, we have

$$ {}^b_a(f+g) \ge {}^b_a f + {}^b_a g = \int_a^b f + \int_a^b g = {}^b_a f + {}^b_a g \ge {}^b_a(f+g) $$

but since ${}^b_a(f+g) \le {}^b_a(f+g)$ we have equality.

iii. Let $h = g - f$. It is integrable based on the previous two points. We have $h \ge 0$ and by theorem 2.3 we have our result.

iv. Let $\epsilon > 0$. Since f is integrable on $[a,b]$, there exists a partition P such that $U(f,P) - L(f,P) < \epsilon$. We shall prove that $U(|f|,P) - L(|f|,P) \le U(f,P) - L(f,P) < \epsilon$.

Now let there be a partition $P = \{x_1, x_2, \ldots, x_n\}$. Consider $[x_{i-1}, x_i]$. Let $S = \{|f(x)| \mid x \in [x_{i-1}, x_i]\}$. Now S is a bounded set since f is bounded (we can only integrate bounded functions), and so we have $M_i(|f|,P) = \sup S$ and $m_i(|f|,P) = \inf S$. For any $u, v \in [x_{i-1}, x_i]$, we have, by the triangle inequality

$$ ||f(u)| - |f(v)|| \le |f(u) - f(v)| \le M_i(f,P) - m_i(f,P) $$

From the lemma, we have $M_i(|f|, P) - m_i(|f|, P) = \sup S - \inf S \leq M_i(f, P) - m_i(f, P)$. Therefore,

$$\begin{aligned}
U(|f|, P) - L(|f|, P) &= \sum_{i=1}^{n} (M_i(|f|, P) - m_i(|f|, P))\Delta x_i \\
&\leq \sum_{i=1}^{n} (M_i(f, P) - m_i(f, P))\Delta x_i \\
&= U(f, P) - L(f, P) \\
&< \epsilon
\end{aligned}$$

Therefore we have established that $|f|$ is integrable on $[a, b]$. Now since $-|f(x)| \leq f(x) \leq |f(x)|$, by part (iii),

$$-\int_a^b |f| = \int_a^b -|f| \leq \int_a^b f \leq \int_a^b |f|$$

and so $\left|\int_a^b f\right| \leq \int_a^b f$.

v. Since f and g are integrable, they are bounded. Let $K > 0$ be such that $|f(x)|, |g(x)| \leq K$ for all $x \in [a, b]$.

   Let $\epsilon > 0$. Then there exist partitions P and Q such that

$$U(f, P) - L(f, P) < \frac{\epsilon}{2K} \qquad U(g, Q) - L(g, Q) < \frac{\epsilon}{2K}.$$

   Now let $R = P \cup Q = \{x_1, x_2, \dots, x_n\}$. We also have

$$U(f, R) - L(f, R) < \frac{\epsilon}{2K} \qquad U(g, R) - L(g, R) < \frac{\epsilon}{2K}.$$

   For $u, v \in [x_{i-1}, x_i]$,

$$\begin{aligned}
|(fg)(u) - (fg)(v)| &\leq |f(u)g(u) - f(v)g(u)| + |f(v)g(u) - f(v)g(v)| \\
&\leq K|f(u) - f(v)| + K|g(u) - g(v)| \\
&\leq K(M_i(f, R) - m_i(f, R)) + K(M_i(g, R) - m_i(g, R)).
\end{aligned}$$

   From the lemma, we have

$$M_i(fg, R) - m_i(fg, R) \leq K(M_i(f, R) - m_i(f, R)) + K(M_i(g, R) - m_i(g, R)).$$

   Therefore

$$\begin{aligned}
U(fg, R) - L(fg, R) &= \sum_{i=1}^{n} (M_i(fg, R) - m_i(fg, R))\Delta x_i \\
&\leq K\sum_{i=1}^{n} [(M_i(f, R) - m_i(f, R)) + (M_i(g, R) - m_i(g, R))]\Delta x_i \\
&\leq K(U(f, R) - L(f, R)) + K(U(g, R) - L(g, R)) \\
&< K\left(\frac{\epsilon}{2K} + \frac{\epsilon}{2K}\right) \\
&= \epsilon.
\end{aligned}$$

■

The following two theorems are familiar results that tell us we can combine and split intervals.

Theorem 2.11. *Let the function $f\colon[a,b]\to\mathbb{R}$ be integrable on both $[a,c]$ and $[c,b]$. Then f is integrable on $[a,b]$ and*

$$\int_a^c f + \int_c^b f = \int_a^b f.$$

Proof. Let $\epsilon > 0$. Choose partitions P and Q such that $U(f,P) < {}_a^c f + \frac{\epsilon}{2}$, and $U(f,Q) < {}_c^b f + \frac{\epsilon}{2}$. Let $R = P \cup Q$. Then R is a partition of $[a,b]$. If we write out the sum fully we would also see that $U(f,R) = U(f,P) + U(f,Q)$. Therefore, we have

$${}_a^b f \le U(f,R) = U(f,P) + U(f,Q) < \int_a^c f + \int_c^b f + \epsilon.$$

Taking the limit $\epsilon \to 0$, we get ${}_a^b f \le \int_a^c f + \int_c^b f$. We may go through a similar argument to arrive at ${}_a^b f \ge \int_a^c f + \int_c^b f$. Since ${}_a^b f \le {}_a^b$, the result follows. ■

Theorem 2.12. *Let the function f be integrable on $[a,b]$. Then for any $c \in (a,b)$, f is integrable on $[a,c]$ and on $[c,b]$.*

Proof. We only present the proof for $[a,c]$, the proof for $[c,b]$ is similar. Let $\epsilon > 0$, then there exists a partition $P = \{x_1, x_2, \ldots, x_n\}$ of $[a,b]$, such that $U(f,P) - L(f,P) < \epsilon$. Consider the partition $Q = (P \cup \{c\}) \cap [a,c]$. Then

$$U(f,Q) - L(f,Q) \le U(f, P\cup\{c\}) - L(f, P\cup\{c\}) < U(f,P) - L(f,P) < \epsilon.$$

■

Using the established results, we can give some examples of integrable functions.

Example 2.3 (Polynomials)**.** Polynomials on any interval $[a,b]$ are integrable. This is simply because they are continuous, so theorem 2.6 says they are integrable. ♢

Example 2.4. The ceiling $f(x) = \lceil x \rceil$ and floor functions $g(x) = \lfloor x \rfloor$ are integrable on any interval $[a,b]$ by theorem 2.5 since they are monotone. Their absolute values are also integrable (the absolute value of any integrable function is integrable). ♢

Theorem 2.13. *Let the function f be integrable on $[a,b]$. Suppose that there is a positive constant C such that $0 < C \le h(x)$ for all $x \in [a,b]$. Then the function $\frac{1}{h}$ is integrable on $[a,b]$.*

Proof. For any $x, y \in [a,b]$, we have

$$\begin{aligned}\left|\frac{1}{h(x)} - \frac{1}{h(y)}\right| &= \left|\frac{h(y)-h(x)}{h(x)h(y)}\right| \\ &\le \frac{1}{C^2}|h(x)-h(y)| \\ &\le \frac{1}{C^2}a(M_i(h,P) - m_i(h,P)).\end{aligned}$$

By the previous lemma this means

$$M_i\left(\frac{1}{h}, P\right) - m_i\left(\frac{1}{h}, P\right) \le \frac{1}{C^2}(M_i(h,P) - m_i(h,P)).$$

Now let $\epsilon > 0$. Since h is integrable on $[a, b]$, by the Riemann integrability criterion, there exists a partition P such that $U(h, P) - L(h, P) < C^2\epsilon$. Then

$$\begin{aligned} U\Big(\frac{1}{h}, P\Big) - L\Big(\frac{1}{h}, P\Big) &= \sum_{i=1}^{n}\Big(M_i\Big(\frac{1}{h}, P\Big) - m_i\Big(\frac{1}{h}, P\Big)\Big)\Delta x_i \\ &\leq \frac{1}{C^2}\sum_{i=1}^{n}(M_i(h, P) - m_i(h, P))\Delta x_i \\ &= \frac{1}{C^2}(U(h, P) - L(h, P)) \\ &< \frac{1}{C^2}C^2\epsilon \\ &= \epsilon. \end{aligned}$$

■

Definition 2.6 (Length of a curve)**.** Let f be a continuous function on $[a, b]$ and let $P = \{x_0, x_1, \dots, x_n\}$ be a partition of $[a, b]$. We define

$$l(f, P) = \sum_{i=1}^{n}\sqrt{(x_i - x_{i-1})^2 + (f(x_i) - f(x_{i-1}))^2}.$$

Intuitively this is the length of a polygonal line plotted along the curve. We then define the length of f on $[a, b]$ by

$$l(f) = \sup\{l(f, P) \mid P \text{ partitions } [a, b]\}.$$

□

Of course the above definition only makes sense if the set of all $l(f, P)$ is bounded. These are the functions with sensible lengths that we can talk about.

Theorem 2.14. *If $f \in C^1[a, b]$, then its length is defined and is given by*

$$l(f) = \int_a^b \sqrt{1 + (f')^2}.$$

2.2 The fundamental theorem of calculus

We defined integration primarily to help us find the area under curves. It can also be used to define new functions which have nice properties.

Definition 2.7 (Indefinite integrals)**.** Let $f\colon [a, b] \to \mathbb{R}$ be integrable on $[a, b]$. For each $x \in (a, b]$, f is integrable on $[a, x]$. Define

$$F(x) = \int_a^x f.$$

The function $F\colon [a, b] \to \mathbb{R}$ is called the *indefinite integral* of f. We also define $\int_a^a f = 0$. □

Theorem 2.15. *The indefinite integral F of f is uniformly continuous on $[a, b]$.*

Proof. For f to be integrable it has to be bounded. Let $M > 0$ be the upper bound of f. For $x, y \in [a, b]$ where $x < y$, we have

$$\begin{aligned}
|F(x) - F(y)| &= \left|\int_a^x f - \int_a^y f\right| \\
&= \left|\int_a^x f - \int_a^x f + \int_x^y f\right| \\
&= \left|\int_x^y f\right| \\
&\leq \int_x^y |f| \\
&\leq \int_x^y M \\
&= M|x - y|
\end{aligned}$$

which shows that F is Lipschitz. For the intermediary steps we have used properties shown in theorem 2.10. ■

We can also easily let $G(x) = \int_{x_0}^x f$ for any $x_0 \in [a, b]$. This means that in fact $G(x) = F(x) - F(x_0)$. Since $F(x_0)$ is just a constant, this means that most of the properties of $F(x)$ carry on to $G(x)$, which means that the lower limit of the indefinite integral does not really matter.

The next sensible question to ask is if F is differentiable. It turns out that this is not always true. However we can still have some criteria under which it is true.

Theorem 2.16 (Fundamental theorem of calculus part 1)**.** *Let f be integrable on $[a, b]$ and let F be the indefinite integral of f. If f is continuous at a point $c \in [a, b]$, then F is differentiable at c with derivative $F'(c) = f(c)$.*

Proof. Let $\epsilon > 0$. Since f is continuous at c, there exists $\delta > 0$ such that $t \in (c - \delta, c + \delta)$ means that $|f(t) - f(c)| < \frac{\epsilon}{2}$. This also means $f(c) - \frac{\epsilon}{2} < f(t) < f(c) + \frac{\epsilon}{2}$.

Now fix $x \in (c - \delta, c)$. For any $y \in [x, c]$, we have $f(c) - \frac{\epsilon}{2} < f(y) < f(c) + \frac{\epsilon}{2}$. Taking the integral on both sides, we have $\int_x^c f(c) - \frac{\epsilon}{2} \leq \int_x^c f \leq \int_x^c f(c) + \frac{\epsilon}{2}$. But this just gives us

$$\begin{gathered}
\left(f(c) - \frac{\epsilon}{2}\right)(c - x) \leq F(c) - F(x) \leq \left(f(c) + \frac{\epsilon}{2}\right)(c - x) \\
f(c) - \frac{\epsilon}{2} \leq \frac{F(c) - F(x)}{c - x} \leq f(c) + \frac{\epsilon}{2} \\
\left|\frac{F(x) - F(c)}{x - c} - f(c)\right| \leq \frac{\epsilon}{2} < \epsilon
\end{gathered}$$

Therefore, $\lim_{x \to c^-} \frac{F(x) - F(c)}{x - c} = f(c)$. We can repeat the proof for the right hand limit to obtain $\lim_{x \to c^+} \frac{F(x) - F(c)}{x - c} = f(c)$. Therefore the derivative of F at c is $f(c)$. ■

This theorem also means the if f is continuous on some interval then the indefinite integral F is differentiable on the same interval and it is what is known as an *anti-derivative* of f. This also means that every continuous function is a derivative of something — its indefinite integral. In fact, the fundamental theorem tells us that the map from $C^1[a, b]$ to $C[a, b]$ is a surjective linear transformation.

In a way, integration is the reverse process of differentiation. However it should be noted that in general this is not the case. This is highlighted in the following example.

Example 2.5. Let $f\colon[-1,1]\to\mathbb{R}$ be defined by

$$f(x)=\begin{cases}-1, & \text{if } -1\le x<0\\ 1, & \text{if } 0\le x\le 1\end{cases}$$

Let us find the indefinite integral. Let $-1\le x<0$. Then $F(x)=\int_1^x -1=-(x+1)$. For $0\le x\le 1$, we get $F(x)=\int_{-1}^0 -1+\int_0^x 1=-1+x$. Actually we have $F(x)=|x|$. Therefore, $F'(0)$ does not exist. ♢

Example 2.6. Let f be a continuous function. Evaluate

$$\lim_{x\to 0}\frac{1}{x}\int_x^{x^3} f.$$

Let $F(x)=\int_0^x f$. The fundamental theorem of calculus tells us that $F'(x)=f(x)$. Then

$$\begin{aligned}\lim_{x\to 0}\frac{1}{x}\int_x^{x^3} f &= \lim_{x\to 0}\frac{1}{x}\left(\int_0^{x^3} f-\int_0^x f\right)\\ &= \lim_{x\to 0}\frac{1}{x}\big(F(x^3)-F(x)\big)\end{aligned}$$

Since F is continuous, the $\lim_{x\to 0}F(x^3)=\lim_{x\to 0}F(x)=F(0)$. Then we use L'Hôpital's rule, and the limit becomes

$$\begin{aligned}\lim_{x\to 0}\frac{\mathrm{d}}{\mathrm{d}x}\big(F(x^3)-F(x)\big) &= \lim_{x\to 0}3x^2f(x^3)-f(x)\\ &= f(0).\end{aligned}$$

♢

Corollary 2.16.1. *If f is continuous on $[a,b]$ and g is differentiable on $[c,d]$, such that* ran $g\subseteq[a,b]$, *then the function*

$$G(x)=\int_a^{g(x)} f$$

where $x\in[c,d]$, is differentiable on $[c,d]$ and

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

Proof. The indefinite integral of f is given by $F(x)=\int_a^x f$ for $x\in[a,b]$. From the fundamental theorem of calculus we have $F'(x)=f(x)$. Now consider the function $F\circ g\colon[c,d]\to\mathbb{R}$. Then we can see that $G=F\circ g$. Furthermore, using the chain rule $(F\circ g)'(x)=g'(x)F'(g(x))=g'(x)f(g(x))$. ■

Corollary 2.16.2 (Mean value theorem for integrals)**.** *Suppose f is continuous on $[a,b]$, then there exists $c\in(a,b)$ such that $\int_a^b f=f(c)(b-a)$.*

Proof. Let $F(x)=\int_a^x f$ for $x\in[a,b]$. Then F is differentiable and $F'(x)=f(x)$. By the mean value theorem for derivatives, there exists $c\in(a,b)$ such that $F(b)-F(a)=F'(c)(b-a)$, or in other words $\int_a^b f=f(c)(b-a)$. ■

Theorem 2.17 (Fundamental theorem of calculus part 2)**.** *Let g be differentiable and continuous on $[a, b]$. Then*

$$\int_a^b g' = g(b) - g(a).$$

Proof. Since g' is continuous on $[a, b]$ it is also integrable on $[a, b]$. Then define $F(x) = \int_a^x g'$ for $x \in [a, b]$. By the fundamental theorem of calculus (part 1), we have $F' = g'$. Then $(F - g)' = 0$ and theorem 1.15 tells us that $F - g$ is a constant function. Let $F = g + c$. So now we have

$$\int_a^b g' = F(b) - F(a) = g(b) + c - g(a) - c = g(b) - g(a).$$

■

Is the continuity of g' necessary for the theorem to hold? It turns out that we do not need it.

Theorem 2.18 (Cauchy's fundamental theorem of calculus)**.** *Let g be differentiable on $[a, b]$ and let g' be integrable on $[a, b]$. Then*

$$\int_a^b g' = g(b) - g(a).$$

Proof. Let $P = \{x_0, x_1, \ldots, x_n\}$ be a partition of $[a, b]$. For $1 \leq i \leq n$, since g is differentiable, by the mean value theorem there is a point $c_i \in (x_{i-1}, x_i)$ such that

$$g'(c_i) = \frac{g(x_i) - g(x_{i-1})}{x_i - x_{i-1}}$$

Hence $g(x_i) - g(x_{i-1}) = g'(c_i)\Delta x_i$. Then

$$\begin{aligned} g(b) - g(a) &= \sum_{i=1}^{n} (g(x_i) - g(x_{i+1}) \\ &= \sum_{i=1}^{n} g'(c_i)\Delta x_i \end{aligned}$$

Recall that we also have $m_i(g', P) \leq g'(c_i) \leq M_i(g', P)$. Thus

$$L(g', P) = \sum_{i=1}^{n} m_i(g', P)\Delta x_i \leq \sum_{i=1}^{n} g'(c_i)\Delta x_i \leq \sum_{i=1}^{n} M_i(g', P)\Delta x_i = U(g', P).$$

This means, for any partition P of $[a, b]$, we also have

$$L(g', p) \leq {}_a^b\, g' \leq g(b) - g(a) \leq {}_a^b\, g' \leq U(g', P).$$

Since g' is integrable, this give us $\int_a^b g' = g(b) - g(a)$. ■

It is possible for a function to have a non-integrable derivative, which is the reason why we need to specify this condition in the theorem.

Example 2.7. Let

$$g(x) = \begin{cases} x^2 \sin\left(\frac{1}{x^2}\right), & \text{if } x \neq 0 \\ 0, & \text{otherwise} \end{cases}$$

Then

$$g'(x) = \begin{cases} 2x \sin \frac{1}{x^2} - \frac{2}{x} \cos \frac{1}{x^2}, & \text{if } x \neq 0 \\ 0, & \text{otherwise} \end{cases}$$

We claim g' is not integrable on $[-1, 1]$ since it is not bounded there. Consider the sequence $s_n = \frac{1}{\sqrt{2n\pi}}$. Then

$$\begin{aligned} \lim_{n\to\infty} g'(s_n) &= 2s_n \sin 2n\pi - \frac{2}{s_n} \cos 2n\pi \\ &= -2\sqrt{2n\pi} \\ &= \infty \end{aligned}$$

◊

Theorem 2.19 (Integration by parts)**.** *Suppose that the functions $u, v\colon [a, b] \to \mathbb{R}$ are differentiable on $[a, b]$ and their derivatives u' and v' are integrable on $[a, b]$. Then*

$$\int_a^b uv' = u(b)v(b) - u(a)v(a) - \int_a^b vu'$$

Proof. Let $g(x) = u(x)v(x)$. Then $g' = u'v + uv'$ by the product rule. It is clearly integrable. Therefore

$$\int_a^b g' = \int_a^b u'v + \int_a^b uv' = u(b)v(b) - u(a)v(a)$$

and rearrangement gives us our desired result. ■

Theorem 2.20 (Integration by substitution)**.** *Suppose that the function $f\colon [a, b] \to \mathbb{R}$ is differentiable on $[a, b]$ and f' is integrable on $[a, b]$. If $g\colon I \to \mathbb{R}$ is continuous on an interval I containing $f[[a, b]]$, then*

$$\int_a^b (g \circ f) \cdot f' = \int_{f(a)}^{f(b)} g.$$

Proof. Since $g(a) \in I$ we can consider the indefinite integral $G(x) = \int_{f(a)}^{x} g$ for $x \in I$. Then $G'(x) = g(x)$. Let $h = G \circ f$. By the chain rule, $h' = (g \circ f) \cdot f'$. Since f' is given to be integrable, and $g \circ f$ is continuous, h' is integrable, and

$$\int_a^b (g \circ f) \cdot f' = \int_a^b (G \circ f)' = G(f(b)) - G(f(a)) = \int_{f(a)}^{f(b)} g.$$

■

Following is another statement of Taylor's theorem where we express the remainder term in an integral instead of a derivative.

Theorem 2.21 (Taylor's theorem in integral form). *Let f be a function such that $f, f', \dots f^{(n+1)}$ exists on $[a, x]$ and $f^{(n+1)}$ is integrable on $[a, x]$. Then*

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

Proof. As usual let us call the remainder term $R_n(x)$. We can try doing integration by parts, letting $u = f^{(n)}$ and $v = \frac{(x-t)^n}{n!}$, so

$$\begin{aligned}
\frac{1}{n!}\int_a^x (x-t)^n f^{(n+1)} &= \int_a^x u'v \\
&= uv\Big|_a^x - \int_a^x uv' \\
&= f^{(n)}(t)\frac{(x-t)^n}{n!}\Big|_a^x - \int_a^x \frac{(-1)(x-n)^{n-1}}{(n-1)!} f^{(n)} \\
&= -\frac{f^{(n)}(a)}{n!}(x-a)^n + \underbrace{\frac{1}{(n-1)!}\int_a^x (x-t)^{n-1} f^{(n)}}_{R_{n-1}}.
\end{aligned}$$

We may perform induction to obtain

$$R_n(x) = -\sum_{k=1}^{n} \frac{f^{(k)}(a)}{k!}(x-a)^k + R_0(x)$$

It can be seen that $R_0(x) = \int_a^x f' = f(x) - f(a)$, so rearranging the expression above gives us the result we want. ■

2.3 Riemann sums

Recall our original definition of the integrals

$$_a^b = \inf_P U(f, P) \qquad\qquad _a^b = \sup_P L(f, P)$$

This is due to Darboux, and is known as the Riemann-Darboux integral. The original Riemann integral was defined in terms of limits of Riemann sums and we shall show that they are equivalent.

Definition 2.8 (Norm of partitions). Let $P = \{x_1, x_2, \dots, x_n\}$ be a partition of $[a, b]$. Then the *norm* of P is $\|P\| = \max\{\Delta x_i\}$ □

It is fairly easy to see that if partitions P and Q are such that $Q \subseteq P$, then $\|P\| \leq \|Q\|$.

Definition 2.9 (Riemann sum). Let $f : [a, b] \to \mathbb{R}$ be a bounded function. Let $P = \{x_0, x_1, \dots, x_n\}$ be a partition of $[a, b]$ and let $\xi_i \in [x_{i-1}, x_i]$ for $1 \leq i \leq n$. The sum

$$S(f, P)(\xi) = \sum_{i=1}^{n} f(\xi_i)\Delta x_i$$

is called the *Riemann sum* of f with respect to P and ξ. □

The difference between the Riemann, upper and lower sums is that there is no preference for the largest or smallest value in the Riemann sum, we can pick any point we want in the subinterval. It is therefore fairly obvious that

$$L(f,P) \leq S(f,P)(\xi) \leq U(f,P).$$

Definition 2.10 (Limit of riemann sums)**.** Suppose that there exists A such that $\forall \epsilon > 0, \exists \delta > 0$, for some partition P of $[a,b]$ and $\xi = \{\xi_1, \xi_2, \dots, \xi_n\}$,

$$\|P\| < \delta \implies |S(f,P)(\xi) - A| < \epsilon$$

then we say that A is the *limit* of these Riemann sums as $\|P\| \to 0$, and we write

$$\lim_{\|P\| \to 0} S(f,P)(\xi) = A.$$

□

Let us recall that since ${}_a^b f$ is an infimum, for every $\epsilon > 0$, there exists a partition P such that $U(f,P) < {}_a^b f + \epsilon$. What does this say about P?

Lemma 2.22. *Let $f\colon [a,b] \to \mathbb{R}$ be a bounded function and P be a partition of $[a,b]$. Then $\forall \epsilon > 0, \exists \delta > 0$*

$$\|P\| < \delta \implies \begin{cases} U(f,P) & < {}_a^b f + \epsilon \\ L(f,P) & > {}_a^b f - \epsilon \end{cases}$$

or equivalently,

$$\|P\| < \delta \implies \begin{cases} \lim\limits_{\|P\| \to 0} U(f,P) & = {}_a^b f \\ \lim\limits_{\|P\| \to 0} L(f,P) & = {}_a^b f \end{cases}$$

Proof. Let $\epsilon > 0$. Then there is a partition $Q = \{y_1, y_2, \dots, y_n\}$ of $[a,b]$ such that $U(f,Q) < {}_a^b f + \frac{\epsilon}{2}$. Let $\eta = \min_{1 \leq i \leq n} \Delta y_i$ and $M = \sup_{x \in [a,b]} f(x)$. Then let $\delta < \min\left(\eta, \frac{\epsilon}{6(N-1)M}\right)$.

Let the partition $P = \{x_1, x_2, \dots, x_n\}$ be such that $\|P\| < \delta$. In each subinterval $[x_{i-1}, x_i]$, there can either be zero or one $y_j \in [x_{i-1}, x_i]$ due to $\|P\| < \eta$. Divide the subintervals into two sets. Let P_1 contain the subintervals such that $(x_{i-1}, x_i) \cap Q = \{y_j\}$ for some $1 \leq j \leq n-1$. There are at most $n-1$ subintervals in P_1. Let P_2 contain the remaining subintervals where $(x_{i-1}, x_i) \cap Q = \emptyset$.

Let $R = P \cup Q = \{z_1, z_2, \dots, z_r\}$. Divide R into two sets again. Let R_1 contain the subintervals obtained by dividing the subintervals from P_1, i.e. they take the form of $[x_{i-1}, y_j]$ or $[y_j, x_i]$. Let $R_2 = P_2$ contain the other subintervals.

Then

$$\begin{aligned} U(f,P) &= \sum_{x_i \in P_1} M_i(f,P)\Delta x_i + \sum_{x_i \in P_2} M_i(f,P)\Delta x_i \\ U(f,R) &= \sum_{x_i \in R_1} M_i(f,R)\Delta x_i + \sum_{x_i \in R_2} M_i(f,R)\Delta x_i \end{aligned}$$

Taking their difference, we have $0 \le U(f,P) - U(f,R) = \sum_{P_1} - \sum_{R_1} \le \left|\sum_{P_1}\right| + \left|\sum_{R_1}\right|$. We also have the following bounds on the sums:

$$|M_i(f,P)\Delta x_i| \le M\|P\| < M\delta \implies \left|\sum_{P_1}\right| \le (N-1)M\delta$$
$$\left|M_j(f,P)\Delta z_i\right| \le M\|R\| < M\delta \implies \left|\sum_{R_1}\right| \le 2(N-1)M\delta$$

Thus $U(f,P) = (U(f,P) - U(f,R)) + U(f,R) \le \frac{\epsilon}{2} + \overline{\int}_a^b f + \frac{\epsilon}{2}$. The proof for the lower sum is similar and omitted. ■

Theorem 2.23. *Let $f\colon [a,b] \to \mathbb{R}$ be a bounded function. Then f is integrable on $[a,b]$ and $\int_a^b f = A$ iff*

$$\lim_{\|P\|\to 0} S(f,P)(\xi) = A.$$

Proof.

($\implies$): Assume f is integrable on $[a,b]$ and $\int_a^b f = A$. Let $\epsilon > 0$, then by lemma 2.22, $\exists\delta > 0$ such that

$$\begin{aligned}\|P\| < \delta &\implies A - \epsilon < L(f,P) \le S(f,P)(\xi) \le U(f,P) < A + \epsilon\\ &\implies |S(f,P)(\xi) - A| < \epsilon.\end{aligned}$$

for any ξ.

($\impliedby$): Assume $\lim_{\|P\|\to 0} S(f,P)(\xi) = A$. Let $\epsilon > 0$. Then $\exists\delta > 0$ such that $\|P\| < \delta \implies |S(f,P)(\xi) - A| < \epsilon$. Recall $M_i = \sup\{f(x) \mid x \in [x_{i-1}, x_i]\}$ and $U(f,P) = \sum_{i=1}^n M_i\Delta x_i$. There exists $\xi_i \in [x_{i-1}, x_i]$ such that $f(\xi_i) > M_i - \frac{\epsilon}{b-a}$. Then

$$S(f,P)(\xi) = \sum_{i=1}^n f(\xi_i)\Delta x_i > \sum_{i=1}^n \left(M_i - \frac{\epsilon}{b-a}\right)\Delta x_i = U(f,P) - \epsilon.$$

Thus we can choose ξ such that $A + \epsilon > S(f,P)(\xi) > U(f,P) - \epsilon \ge \overline{\int}_a^b f - \epsilon$. This means $\overline{\int}_a^b f < A + 2\epsilon$, and so as $\epsilon \to 0$, we conclude that $\overline{\int}_a^b f \ge A$. There is a similar proof for the lower integral, using which we arrive at the desired conclusion. ■

Corollary 2.23.1. *Let $f\colon [a,b] \to \mathbb{R}$ be integrable on $[a,b]$. For each $n \in \mathbb{N}$, let $P_n = \{x_0^{(n)}, x_1^{(n)}, \dots, x_{m_n}^{(n)}\}$ be a partition of $[a,b]$ and let $\xi^{(n)} = \{\xi_1^{(n)}, \xi_2^{(n)}, \dots, \xi_{m_n}^{(n)}\}$ be such that $\xi_i^{(n)} \in [x_{i-1}^{(n)}, x_i^{(n)}]$ for all $1 \le i \le m_n$. If $\lim_{n\to\infty} \|P_n\| = 0$ then $\lim_{n\to\infty} S(f,P_n)(\xi^{(n)}) = \int_a^b f$.*

Proof. Let $A = \int_a^b f$. Let $\epsilon > 0$. By theorem 2.23, $\lim_{\|P\|\to 0} S(f,P)(\xi) = A$, i.e. $\exists\delta > 0$ such that $\|P\| < \delta \implies |S(f,P)(\xi) - A| < \epsilon$ for any ξ. Since $\|P_n\| \to 0$,

$$\exists k \in \mathbb{N}\left[n > k \implies \|P_n\| < \delta \implies |S(f,P_n)(\xi^{(n)}) - A| < \epsilon\right]$$

■

Example 2.8. We want to find $\lim_{n\to\infty}\sum_{i=1}^{n}\frac{1}{n+i}$. From first principles this is not easy to evaluate, since the number of terms in the limit changes. However let us write this as a Riemann sum.

Let $f(x) = \frac{1}{x}$. Then we can rewrite the terms in the limit as $\sum_{i=1}^{n}\frac{1}{1+\frac{i}{n}}\frac{1}{n} = \sum_{i=1}^{n} f(1+\frac{i}{n})\frac{1}{n}$. Let $P_n = \left\{1, 1+\frac{1}{n}, \dots, 1+\frac{n}{n}\right\}$ and $\xi^{(n)} = \left\{1+\frac{1}{n}, \dots, 1+\frac{n}{n}\right\}$. We have $\Delta x_i^{(n)} = \frac{1}{n}$ and so $\|P_n\| = \frac{1}{n}$, so by corollary 2.23.1

$$\lim_{n\to\infty}\sum_{i=1}^{n}\frac{1}{n+i} = \lim_{n\to\infty}\sum_{i=1}^{n} f(\xi_i^{(n)})\Delta x_i^{(n)} = \lim_{n\to\infty} S(f, P_n)(\xi^{(n)}) = \int_1^2 f = \ln 2$$

Another way is to let $g(x) = \frac{1}{1+x}$. Then we can rewrite the terms as $\sum_{i=1}^{n} g(\frac{i}{n})\frac{1}{n}$. Let $P_n' = \left\{0, \frac{1}{n}, \dots, 1\right\}$, and $\xi'^{(n)} = \left\{\frac{1}{n}, \frac{2}{n}, \dots \frac{n}{n}\right\}$. Similarly we have

$$\lim_{n\to\infty}\sum_{i=1}^{n}\frac{1}{n+i} = \lim_{n\to\infty}\sum_{i=1}^{n} g(\xi_i'^{(n)})\Delta x_i^{(n)} = \lim_{n\to\infty} S(f, P_n')(\xi'^{(n)}) = \int_0^1 g = \ln 2$$

◊

2.4 Improper integrals

In defining $\int_a^b f$ we have assume that the interval of integration is a compact interval and the function f is bounded in this interval. An improper integral is on where either of these assumptions fail.

Definition 2.11 (Improper integral for non-bounded functions)**.** Suppose that f is defined on $[a, b)$ and f is integrable on $[a, c]$ for every $c \in (a, b)$, then we define

$$\int_a^b = \lim_{c\to b^-}\int_a^c f$$

provided that the limit exists[2].

Similarly, if f is defined on $(a, b]$ and f is integrable on $[c, b]$ for every $c \in (a, b)$, then we define

$$\int_a^b f = \lim_{c\to a^+}\int_c^b f$$

provided that the limit exists. □

In the case when f is integrable on $[a, b]$, the indefinite integral is consistent with the definite integral we know. This is because the improper integral $\int_a^x f$ is continuous on $[a, b]$.

Example 2.9. Let $f(x) = x^{-\frac{1}{3}}$ for $x \in (0, 1]$. This is not integrable on $[0, 1]$ (regardless of how we extend it to 0 since it is unbounded). But f is integrable on $[c, 1]$ for any $c \in (0, 1)$ since it is continuous, and we have

$$\int_c^1 x^{-\frac{1}{3}} = \frac{3}{2}\left(1 - c^{\frac{2}{3}}\right).$$

[2] In this section we only consider finite (not $\pm\infty$) limits when we say "exist"

Let us evaluate the limit

$$\lim_{c\to 0^+}\int_c^1 f = \lim_{c\to 0^+}\frac{3}{2}\left(1-c^{\frac{2}{3}}\right) = \frac{3}{2}$$

Hence the improper integral $\int_0^1 f = \frac{3}{2}$.

As another example, let $g(x) = \frac{1}{x^2}$ for $x \in (0,1]$. We have

$$\int_c^1 \frac{1}{x^2} = -1 + \frac{1}{c}.$$

The limit $\lim_{c\to 0^+}\int_c^1 f$ does not exists, and so the indefinite integral diverges. ◊

Definition 2.12 (Improper integral for non-bounded intervals)**.** Suppose that f is defined on $[a,\infty)$ and f is integrable on $[a,c]$ for every $c > a$, then we define

$$\int_a^\infty = \lim_{c\to\infty}\int_a^c f$$

provided that the limit exists.

Similarly, if f is defined on $(-\infty, b]$ and f is integrable on $[c,b]$ for every $c < b$, then we define

$$\int_{-\infty}^b f = \lim_{c\to-\infty}\int_c^b f$$

provided that the limit exists.

Furthermore, if both $\int_a^\infty f$ and $\int_{-\infty}^0 f$ converges, then we define

$$\int_{-\infty}^\infty f = \int_0^\infty f + \int_{-\infty}^0 f.$$

□

Example 2.10. Let $g(x) = \frac{1}{x^2}$ for $x \in (0,\infty)$. We have

$$\int_1^c \frac{1}{x^2} = 1 - \frac{1}{c}.$$

Evaluate

$$\lim_{c\to\infty}\int_1^c \frac{1}{x^2} = \lim_{c\to\infty}(1-\frac{1}{c}) = 1$$

so the improper integral converges to 1. ◊

Theorem 2.24. *Let* $f\colon [c,\infty) \to \mathbb{R}$ *be integrable on* $[c,d]$ *for every* $d > c$*. The improper integral* $\int_c^\infty f$ *converges iff* $\forall\epsilon > 0, \exists M > 0\left[M < a < b \implies \left|\int_a^b f\right| < \epsilon\right]$.

Proof. The limit $\lim_{x\to\infty}\int_c^x f$ exists iff it satisfies the Cauchy criterion, i.e.

$$\forall\epsilon > 0, \exists M > 0, \forall a,b > M\left[\left|\int_a^b f\right| < \epsilon\right]$$

In this case we can let $M < a < b$ since the sign does not matter. This leads directly to the result. ■

Theorem 2.25. *Let $f\colon [a,\infty) \to \mathbb{R}$ be integrable on $[a,b]$ for every $b > a$, and that $f(x) \geq 0$ for all $x \in [a,\infty)$. The improper integral $\int_a^\infty f$ converges iff $\int_a^x f$ is bounded for all $x \in [a,\infty)$.*

Proof. Let $F(x) = \int_a^x f$. Let $a \leq x_1 < x_2$. We have $F(x_2) - F(x_1) = \int_{x_1}^{x_2} f \geq 0$. Thus F is an increasing function. Therefore the limit $\lim_{x\to\infty} F$ exists iff F is bounded on $[a,\infty)$. ■

Corollary 2.25.1. *Let $f\colon [a,\infty) \to \mathbb{R}$ and $g\colon [a,\infty) \to \mathbb{R}$ be integrable on $[a,b]$ for every $b > a$. Furthermore assume that $0 \leq f(x) \leq g(x)$ for every $x \in [a,\infty)$. Then if $\int_a^\infty g$ converges, so does $\int_a^\infty f$.*

Proof. For every $x \in [a,\infty)$, $\int_a^x g$ is an upper bound for $\int_a^x f$. ■

Theorem 2.26 (Integral test)**.** *Let f be a positive decreasing function on $[1,\infty)$ and let $a_n = f(n)$. Then $\int_1^\infty f$ converges iff $\sum_{n=1}^\infty a_n$ converges.*

Proof. Since f is positive and decreasing on the interval $[k-1,k]$ where $k = 2,3,\dots$, we have

$$f(k-1) \geq \int_{k-1}^{k} f \geq f(k).$$

By adding this inequality for $k = 2,3,\dots,n$, we obtain

$$\sum_{k=1}^{n} f(k) - f(1) \leq \int_1^n f \leq \sum_{k=1}^{n-1} f(k).$$

Letting $n \to \infty$, we get our result. ■

3 Sequences of functions

Considering sequences can sometimes lead us to new things. For example there are many famous sequences in $\mathbb{Q}$ that converges to $\mathbb{R}$.

Definition 3.1 (Sequences of functions)**.** Let $E \subseteq \mathbb{R}$. Suppose for each $n \in \mathbb{N}$, we have a function $f_n\colon E \to \mathbb{R}$. Then

$$(f_n) = (f_1, f_2, \dots)$$

is a *sequence of functions on E.* □

A sequence of functions can be seen as a family of sequences enumerated by the set E. For each $x \in E$, $(f_n(x))$ is a sequence of real numbers. This sequence may converge or diverge.

3.1 Pointwise convergence

Definition 3.2 (Pointwise convergence)**.** Suppose that for every $x \in E$ the sequence $(f_n(x))$ converges. Define the function $f\colon E \to \mathbb{R}$ by

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

for $x \in E$. We say that (f_n) converges to f *pointwise* on E. □

In other words, f_n converges to f pointwise on E iff $\forall x \in E\,[f_n(x) \to f(x)]$ iff $\forall x \in E, \forall \epsilon > 0, \exists k \in N,\ n \geq K \implies |f_n(x) - f(x)| < \epsilon$.

How well does pointwise convergence preserve the properties of the functions in the sequence? It does not preserve continuity, consider the following counter-example.

Example 3.1. Consider the sequence of functions with $f_n(x) = x^n$ for $x \in [0, 1]$. Then (f_n) converges to the following step function

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

◊

Neither does it preserve integrability. Even if the limit function was integrable, the integral of the limit function is different from the limit of the integrals of the functions.

Example 3.2. Consider the sequence of functions with each $g_n\colon [0, 1] \to \mathbb{R}$ defined as

$$g_n(x) = \begin{cases} 2n^2 x, & \text{if } 0 \leq x \leq \frac{1}{2n} \\ 2n - 2n^2 x, & \text{if } \frac{1}{2n} \leq x \leq \frac{1}{n} \\ 0, & \text{otherwise} \end{cases}$$

The function g_n is simpler than it seems, its graph is a triangle with base $\frac{1}{n}$ and height n. Does (g_n) converge? Take $x \in (0, 1]$. By the Archimedean property, $\exists k \in \mathbb{N}\ [k > \frac{1}{x}]$. Thus $n \geq k \implies \frac{1}{n} \leq \frac{1}{k} < x \implies g_n(x) = 0$. Hence g_n converges to the zero function pointwise on $[0, 1]$. Every function in the sequence has the same area under them, but somehow their limit function has an area of 0. ◊

Finally, since it does not even preserve continuity, it will not preserve differentiability. However if the limit function was differentiable, would the derivative of the limit function converge to the limit of the derivatives? The answer is no.

Example 3.3. Consider the sequence $h_n(x) = \frac{1}{\sqrt{n}} \sin nx$. Then

$$|h_n(x) - 0| \leq \frac{1}{\sqrt{n}} \to 0$$

and so (h_n) converges to the zero function for all x. Now

$$h_n'(x) = \sqrt{n} \cos nx \to \infty$$

◊

We need a stronger form of convergence.

3.2 Uniform convergence

Definition 3.3 (Uniform convergence)**.** A sequence (f_n) of functions *converges uniformly* to f on $E \subseteq \mathbb{R}$ if $\forall \epsilon > 0, \exists K \in \mathbb{N}, \forall x \in E$,

$$n \geq K \implies |f_n(x) - f(x)| < \epsilon$$

□

Example 3.4. Consider again the sequence $h_n(x) = \frac{1}{\sqrt{n}} \sin nx$. Let $\epsilon > 0$. Choose $K \in \mathbb{N}$ such that $K > \frac{1}{\epsilon^2}$. Then $\forall x \in \mathbb{R}$,

$$n \geq K \implies |h_n(x) - 0| \leq \frac{1}{\sqrt{n}} \leq \frac{1}{\sqrt{K}} < \epsilon.$$

so (h_n) converges uniformly to the zero function. ♢

We will introduce a quantity that makes checking for uniform convergence more convenient.

Definition 3.4 (Uniform norm). Let $E \subseteq \mathbb{R}$ and let $\phi\colon E \to \mathbb{R}$ be a bounded function. The *uniform norm* of ϕ on E is defined as

$$\|\phi\|_E = \sup\{|\phi(x)| \mid x \in E\}$$

□

This is a norm function in the space of all bounded functions on E. An important property of norm functions is the triangle inequality. We have an upper bound

$$|(\phi_1 + \phi_2)(x)| \leq |\phi_1(x)| + |\phi_2(x)| \leq \|\phi_1\|_E + \|\phi_2\|_E$$

so

$$\|\phi_1 + \phi_2\| = \sup\{|(\phi_1 + \phi_2)(x)|\} \leq \|\phi_1\|_E + \|\phi_2\|_E$$

Theorem 3.1. *A sequence of functions (f_n) converges uniformly on E iff* $\lim_{n\to\infty} \|f_n - f\|_E = 0$.

Proof.

$(\implies)$: Assume f_n converges to f uniformly on E. Let $\epsilon > 0$. Then $\exists K \in \mathbb{N}, \forall x \in E$,

$$n \geq K \implies |f_n(x) - f(x)| < \frac{\epsilon}{2}$$

This form an upper bound, thus

$$\|f_n - f\|_E \leq \frac{\epsilon}{2} < \epsilon.$$

$(\impliedby)$: Assume $\|f_n - f\|_E \to 0$. Let $\epsilon > 0$. Then $\exists K \in \mathbb{N}$ such that

$$n \geq K \implies \|f_n - f\|_E < \epsilon \implies |f_n(x) - f(x)| \leq \|f_n - f\|_E < \epsilon$$

■

Theorem 3.2 (Cauchy criterion for uniform convergence). *A sequence of functions (f_n) converges uniformly on E iff*

$$\forall \epsilon > 0, \exists K \in N, \forall n, m \geq K \left[\|f_n - f_m\|_E < \epsilon\right]$$

Proof.

$(\implies)$: Assume (f_n) converges to f uniformly on E. Let $\epsilon > 0$. Then $\exists K \in \mathbb{N}$,

$$n \geq K \implies \|f_n - f\|_E < \frac{\epsilon}{2}$$

Then for $n, m \geq K$,

$$\|f_n - f_m\|_E \leq \|f_n - f\| + \|f - f_m\| < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon$$

($\Longleftarrow$): Let $\epsilon > 0$. This means $\exists K \in \mathbb{N}, \forall a, b > K$,

$$|f_a(x) - f_b(x)| \leq \|f_a - f_b\|_E < \frac{\epsilon}{2}.$$

This means that $(f_n(x))$ is a Cauchy sequence of real numbers, and so converges. Define $f(x) = \lim_{n\to\infty} f_n(x)$ for $x \in E$. Taking the limit,

$$|f_n(x) - f(x)| = \lim_{m\to\infty} |f_n(x) - f_m(x)| \leq \frac{\epsilon}{2} < \epsilon.$$

Thus (f_n) converges to f uniformly on E. ■

Theorem 3.3. *A sequence (f_n) does not converge uniformly to f on E iff for some $\epsilon > 0$, there is a subsequence $\left(f_{n_k}\right)$ and a sequence (x_k) in E such that for all $k \in \mathbb{N}$*

$$\left|f_{n_k}(x_k) - f(x_k)\right| \geq \epsilon$$

Proof.

($\Longrightarrow$): Assume (f_n) does not converge to f uniformly on E. Then $\exists \epsilon > 0, \forall K \in \mathbb{N}, \exists n \geq K, \exists x \in E\left[\|f_n(x) - f(x)\|_E \geq \epsilon\right]$.

Specifically, let $K_1 = 1$. Then $\exists n_1 \geq 1, \exists x_1 \in E\left[\left\|f_{n_1}(x_1) - f(x_1)\right\|_E \geq \epsilon\right]$. Next let $K_2 = n_1 + 1$. Then $\exists n_2 > n_1, \exists x_2 \in E\left[\left\|f_{n_2}(x_2) - f(x_2)\right\|_E \geq \epsilon\right]$. Therefore this way we can inductively define the subsequence (f_{n_k}) and the sequence (x_k) that satisfies the claim.

($\Longleftarrow$): This direction is quite clear by comparing with the definition for uniform convergence. ■

Example 3.5. Consider again the sequence of functions with $f_n(x) = x^n$ for $x \in [0, 1]$. We already know what (f_n) pointwise converges to, call it f. We will show that it does not converge to f uniformly using the previous theorem. Define a sequence $x_k = \left(\frac{1}{2}\right)^{\frac{1}{k}}$. Then $|f_k(x_k) - f(x_k)| = \frac{1}{2}$.

Furthermore we see that removing the problematic end point does not work, since the same argument applies to $[0, 1)$ as well. We have to do more than that. Let $0 < r < 1$ and consider the interval $[0, r]$. Then

$$\|f_n - f\|_{[0,r]} \leq r^n \to 0$$

and so it converges uniformly on this restricted interval. ◊

Example 3.6. Consider the sequence of functions $f_n(x) = x^n(1 - x^n)$ for $x \in [0, 1]$. It converges to the zero function pointwise. However it does not converge uniformly,

$$\|f_n - 0\| = \sup\{x^n(1 - x^n)\} = \sup\left\{\frac{1}{4} - (x^n - \frac{1}{2})^2\right\} = \frac{1}{4}$$

which does not go to 0. We can also define the sequence $x_k = \left(\frac{1}{2}\right)^{\frac{1}{k}}$, and

$$|f_k(x_k) - f(x_k)| = \frac{1}{4}.$$

We can try restricting the interval, let $0 < r < 1$ and consider $[0, r]$. Then

$$|f_n(x) - 0| = \left|x^n - x^{2n}\right| \leq |x^n| + \left|x^{2n}\right| \leq r^n + r^{2n} \to 0.$$

◊

Theorem 3.4. *If (f_n) converges uniformly on A and B, then (f_n) converges uniformly on $A \cup B$.*

Proof. Let $\epsilon > 0$. There exists $K_1, K_2 \in \mathbb{N}$ such that

$$\forall x \in A\, [n \geq K_1 \implies |f_n(x) - f(x)| < \epsilon]$$
$$\forall x \in B\, [n \geq K_2 \implies |f_n(x) - f(x)| < \epsilon]$$

Take $K = \max(K_1, K_2)$ and we have

$$\forall x \in A \cup B\, [n \geq K \implies |f_n(x) - f(x)| < \epsilon]$$

■

However it is not true for infinite unions as the following example shows.

Example 3.7. Let $f_n(x) = \frac{x}{x+n}$. For each $a > 0$, (f_n) converges uniformly to the zero function on $[0, a]$. This is because $\|f_n - 0\|_{[0,a]} \leq \frac{x}{n} \leq \frac{a}{n}$. Thus $\lim_{n\to\infty}\|f_n - 0\| = 0$. However (f_n) does not converge uniformly on $[0, \infty)$, since $|f_n(n) - 0| = \frac{1}{2}$. ♢

Theorem 3.5. *Let (f_n) and (g_n) be sequences of functions on E that converge uniformly on E to f and g respectively. Then $(f_n + g_n)$ converges uniformly on E to $f + g$.*

Proof. Let $\epsilon > 0$. Then $\exists M_1, M_2$ such that $\forall x \in E$,

$$n \geq M_1 \implies |f_n(x) - f(x)| < \frac{\epsilon}{2}$$
$$m \geq M_2 \implies |g_m(x) - g(x)| < \frac{\epsilon}{2}$$

Then let $N = \max(M_1, M_2)$.

$$i \geq N \implies |(f_n + g_n) - (f + g)| \leq |f_n - f| + |g_n - g| \leq \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon$$

■

The same is generally not true for multiplication. Consider the following counterexample.

Example 3.8. Let $f_n(x) = x + \frac{1}{n}$ and $f(x) = x$. Then for all x

$$|f_n(x) - f(x)| = \left|\frac{1}{n}\right|$$

which means $\lim_{n\to\infty}\|f_n - f\|_{\mathbb{R}} = 0$, hence (f_n) converges to f uniformly on $\mathbb{R}$. However consider $(f_n)(f_n) = (f_n^2)$. For any $n \in \mathbb{N}$,

$$\left|f_n^2(n) - f^2(n)\right| = \left|n^2 + \frac{2n}{n} + \frac{1}{n^2} - n^2\right| \geq 2$$

♢

This is fixed by an additional requirement of boundedness. We show this after a lemma.

Lemma 3.6. *Let (f_n) be a sequence of bounded functions that converges uniformly to f on E. Then* *i.* *f is bounded on E.*

ii. *(f_n) is uniformly bounded, i.e.* $\exists M > 0\ [|f_n(x)| \leq M]$.

Proof. Let $\epsilon > 0$. Then

$$\exists K, \forall x \in E\ [n \geq K \implies |f_n(x) - f(x)| < \epsilon]$$

i. Let $|f_n(x)| \leq M_n$ for all $x \in E$. Then $|f(x)| \leq |f(x) - f_K(x)| + |f_K(x)| = \epsilon + M_K$.

ii. We have $|f_n(x)| \leq |f_n(x) - f(x)| + |f(x)| < \epsilon + (\epsilon + M_k)$. Now let $M = \max(M_1, M_2, \dots, M_{k-1}, 2\epsilon + M_k)$.

■

Theorem 3.7. *Let (f_n) and (g_n) be sequences of bounded functions on E that converge uniformly on E to f and g respectively. Then $(f_n g_n)$ converges uniformly on E to fg.*

Proof. Let $\epsilon > 0$. Then $\exists M_1, M_2$ such that $\forall x \in E$,

$$n \geq M_1 \implies |f_n(x) - f(x)| < \frac{\epsilon}{2F}$$
$$m \geq M_2 \implies |g_m(x) - g(x)| < \frac{\epsilon}{2G}$$

Now let $N = \max(M_1, M_2)$. By lemma 3.6, let g be bounded by G and let (f_n) be uniformly bounded by F.

$$i \geq N \implies |f_i g_i - fg| \leq |f_i g_i - f_i g| + |f_i g - fg| \leq \epsilon^2.$$

■

3.3 Properties preserved by uniform convergence

Uniform convergence being a stronger form of convergence preserves more properties than pointwise convergence. We can also think of uniform convergence as a condition upon which we may commute the limit taking operations under certain circumstances.

Theorem 3.8. *If (f_n) converges uniformly to f on an interval I and each f_n is continuous at $x_0 \in I$, then f is continuous at x_0.*

Proof. Let $\epsilon > 0$. Then uniform convergence means

$$\exists K \in \mathbb{N}, \forall x \in I \left[n \geq K \implies |f_n(x) - f(x)| < \frac{\epsilon}{3} \right].$$

More specifically, $|f_K(x) - f(x)| < \frac{\epsilon}{3}$. Continuity of f_K means

$$\exists \delta > 0, \forall x \in I \left[|x - x_0| < \delta \implies |f_K(x) - f(x)| < \frac{\epsilon}{3} \right].$$

Thus

$$|f(x) - f(x_0)| \leq |f(x) - f_K(x)| + |f_K(x) - f_K(x_0)| + |f_K(x_0) - f(x_0)| \leq \frac{\epsilon}{3} + \frac{\epsilon}{3} + \frac{\epsilon}{3}$$

■

It directly follows that if all f_n are continuous on I then f is also continuous on I.

Example 3.9. Consider again the sequence $f_n(x) = x^n$ for $x \in [0, 1]$. Every f_n is continuous on $[0, 1]$. The function f that (f_n) converges to pointwise is not continuous on $[0, 1]$. Thus it follows that (f_n) does not converge uniformly to f. ♢

Theorem 3.9. *Suppose that* (f_n) *converges uniformly to* f *on* $[a, b]$ *and each* f_n *is integrable on* $[a, b]$. *Then* f *is integrable on* $[a, b]$ *and* $\forall x_0 \in [a, b]$, *the sequence* $F_n(x) = \int_{x_0}^{a} f_n$ *converges uniformly to* $F(x) = \int_{x_0}^{x} f$ *on* $[a, b]$.

Proof. Let $\epsilon_n = \|f_n - f\|_{[a,b]}$. Since (f_n) converges to f uniformly on $[a, b]$, we know $\lim_{n\to\infty} \epsilon_n = 0$. Now for $x \in [a, b]$

$$|f_n(x) - f(x)| \leq \|f_n - f\|_{[a,b]} = \epsilon_n$$

and rearranging we have

$$f_n(x) - \epsilon_n \leq f(x) \leq f_n(x) + \epsilon_n.$$

Considering the upper and lower integrals

$$\int_a^b f_n - \epsilon_n(b-a) = {}^b_a(f_n - \epsilon_n) \leq {}^b_a f \leq {}^b_a f \leq {}^b_a(f_n + \epsilon_n) = \int_a^b f_n + \epsilon_n(b-a).$$

As $n \to \infty$ we see that we have ${}^b_a f = {}^b_a f$ and so f is integrable.

Next, assume $x_0 < x$. The other case is similar.

$$|F_n(x) - F(x)| = \left|\int_{x_0}^{x} (f_n - f)\right| \leq \int_{x_0}^{x} |f_n - f| \leq \int_{x_0}^{x} \epsilon_n = \epsilon_n|x - x_0| \leq \epsilon_n(b-a)$$

Thus this forms an upper bound, and we conclude

$$\|F_n - F\|_{[a,b]} \leq \epsilon_n(b-a) \to 0.$$

■

Example 3.10. Let $f(x) = x^n \sin nx$ and consider the limit $\lim_{n\to\infty} \int_0^{\frac{\pi}{4}} f$. For $x \in [0, \frac{\pi}{4}]$, $x < 1$ and we have $\|f_n - f\|_{[0,\frac{\pi}{4}]} \to 0$ so we have uniform convergence. Then

$$\lim_{n\to\infty} \int_0^{\frac{\pi}{4}} x^n \sin nx = \int_0^{\frac{\pi}{4}} \lim_{n\to\infty} f_n(x) = 0$$

♢

The natural question to ask now is with regards to differentiability. Unfortunately, even if all functions in a sequence are differentiable, the limit function might not be differentiable.

Example 3.11. Let $f_n(x) = \sqrt{x^2 + \frac{1}{n}}$. Then (f_n) converges uniformly to $\sqrt{x^2}$, since we have

$$\begin{aligned}|f_n(x) - x^2| &= \frac{x^2 + \frac{1}{n} - x^2}{\sqrt{x^2 + \frac{1}{n}} + \sqrt{x^2}} \\ &\leq \frac{1/n}{1/\sqrt{n}} \\ &= \frac{1}{\sqrt{n}}\end{aligned}$$

and $\lim_{n\to\infty} \frac{1}{\sqrt{n}} = 0$. Furthermore clearly every f_n is differentiable on $(-1, 1)$. However, $\sqrt{x^2} = |x|$ is not differentiable on $(-1, 1)$. ♢

Even if the limit function were to be differentiable, the derivative of the functions might not be equal to the limit of the derivatives.

Example 3.12. We have seen previously that $f_n = \frac{1}{\sqrt{n}}$ converges uniformly to the zero function. But $h_n'(x) = \sqrt{n} \cos nx$ which definitely does not converge to the zero function, for example consider $h_n'(0) = \sqrt{n}$. ♢

Theorem 3.10. *Let (f_n) be a sequence of functions where each $f_n \in C^1([a, b])$ (has continuous first derivative). Furthermore, assume that $(f_n(x_0))$ converges for some point $x_0 \in [a, b]$ and (f_n') converges uniformly on $[a, b]$. Then (f_n) converges uniformly to a differentiable function f on $[a, b]$ and for all $x \in [a, b]$,*

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

Proof. Let $L = \lim_{n\to\infty} f_n(x_0)$ and suppose (f_n') converges uniformly to g. Since f_n' is continuous, by the second fundamental theorem of calculus $\int_{x_0}^{x} f_n' = f_n(x) - f_n(x_0)$. Then, since uniform continuity preserves integrability, $\int_{x_0}^{x} f_n'$ converges uniformly to $\int_{x_0}^{x} g$ on $[a, b]$. Thus, rearrangement shows us that f_n will converges uniformly to $f(x) = L + \int_{x_0}^{x} g$ for $x \in [a, b]$. By the fundamental theorem of calculus, $f'(x) = 0 + g(x) = \lim_{n\to\infty} f_n'(x)$. ■

In fact, we can relax the condition on continuity of f_n'.

Theorem 3.11. *Let (f_n) be a sequence of differentiable functions on $[a, b]$, and $(f_n(x_0))$ converges for some point $x_0 \in [a, b]$, and (f_n') converges uniformly on $[a, b]$. Then (f_n) converges uniformly to a differentiable function f on $[a, b]$ and $\lim_{n\to\infty} f_n' = f'$.*

Proof. Bartle, Introduction to real analysis theorem 8.2.3 ■

4 Series of functions

4.1 Infinite series of functions

Definition 4.1 (Infinite series of functions)**.** If (f_n) is a sequence of functions on E, then $(S_n) = \sum_{n=1}^{\infty} f_n$ is an *infinite series* of functions. For each $n \in \mathbb{N}$, the n-th *partial sum* is the function $S_n(x) = \sum_{i=1}^{n} f_i(x)$ for $x \in E$. □

Definition 4.2 (Convergence of series)**.** The series $\sum_{n=1}^{\infty} f_n$ is said to *converge pointwise (uniformly)* to a function S on E if the sequence (S_n) converges pointwise (uniformly) to S on E.

The series $\sum_{n=1}^{\infty} f_n$ is said to *converge absolutely* to a function S on E if the series $\sum_{n=1}^{\infty} |f_n|$ converges pointwise to S on E. □

Theorem 4.1 (Cauchy criterion for uniform convergence)**.** *Let (f_n) be a sequence of functions on E. Then the series $\sum_{n=1}^{\infty} f_n$ converges uniformly on E iff*

$$\forall \epsilon > 0, \exists K \in \mathbb{N} \left[n > m \geq K \implies \|f_{m+1} + \cdots + f_n\|_E < \epsilon \right].$$

Proof. $\sum_{n=1}^{\infty} f_n$ converges uniformly on E iff (S_n) converges uniformly on E iff the Cauchy criterion holds for the sequence of functions (S_n). Therefore

$$\forall \epsilon > 0, \exists K \in \mathbb{N} \, [n > m \geq K \implies \|S_n - S_m\| - E < \epsilon].$$

■

We can restate the Cauchy criterion without using uniform norm with the following observation

$$\|f_{m+1} + \cdots + f_n\|_E = \sup\{|f_{m+1}(x) + \cdots + f_n(x)| \mid x \in E\} < \epsilon$$
$$\iff$$
$$\forall x \in E \, [|f_{m+1}(x) + \cdots + f_n(x)| < \epsilon]$$

This leads to an easy way of showing non-convergence (after taking the contrapositive).

Corollary 4.1.1. *If $\sum_{n=1}^{\infty} f_n$ converges uniformly on E, then f_n converges to the zero function uniformly on E.*

Proof. Let $\epsilon > 0$. Then $\exists K \in \mathbb{N} \, [n > m \geq K \implies \|f_{m+1} + \cdots + f_n\|_E < \epsilon]$. Take $m = n - 1$. Then $\|f_n\|_E = \|f_n - 0\|_E < \epsilon$. ■

Theorem 4.2 (Weierstrass M-test)**.** *Let (f_n) be a sequence of functions on E and let (M_n) be a sequence of positive real numbers such that $\forall n \in \mathbb{N} \, [\|f_n\|_E \leq M_n]$ (or equivalently $\forall x \in E, \forall n \in \mathbb{N} \, [|f_n(x)| \leq M_n]$). If the series $\sum_{n=1}^{\infty} M_n$ converges, then $\sum_{n=1}^{\infty} f_n$ converges uniformly on E.*

Proof. Let $\epsilon > 0$. Since $\sum_{n=1}^{\infty} M_n$ converges, by the Cauchy criterion for series of real numbers,

$$\exists K \in \mathbb{N} \left[n > m \geq K \implies \sum_{k=m+1}^{n} M_k < \epsilon \right]$$

Then,

$$n > m \geq K \implies \|f_{m+1} + \cdots + f_n\|_E \leq \|f_{m+1}\|_E + \cdots + \|f_n\|_E \leq M_{m+1} + \cdots + M_n < \epsilon.$$

■

Example 4.1. Let $f_n(x) = \frac{\sin nx}{n^2}$ and consider the series of functions $\sum_{n=1}^{\infty} f$ on $\mathbb{R}$. Since $|f_n(x)| \leq \frac{1}{n^2}$, and $\sum_{n=1}^{\infty} \frac{1}{n^2}$ converges, thus by the M-test $\sum_{n=1}^{\infty} f_n$ converges uniformly. ◊

Example 4.2 (Geometric series)**.** Consider the series of functions $(S_n(x)) = \sum_{n=0}^{\infty} x^n$ on $(-1, 1)$. We know from calculus that (S_n) converges pointwise to $\frac{1}{1-x}$ pointwise on $(-1, 1)$. It does not however converge uniformly on this interval, since $\|x^n\|_{(-1,1)} = \sup\{|x^n|\} = 1$. However, if we take $0 < r < 1$ and restrict the interval to $[-r, r]$, then $\|x^n\|_{[-r,r]} \le r^n$ and so the series converges uniformly on $[-r, r]$. ♢

In fact, since f_n is continuous, the partial sums will also be continuous, and so the limit function will also be continuous by the uniform convergence. So infinite series can produce new functions for us with certain desired properties.

Theorem 4.3. *If $\sum_{n=1}^{\infty} f_n$ converges uniformly to f on an interval I and each f_n is continuous at $x_0 \in I$, then f is continuous at x_0.*

Proof. For each $n \in \mathbb{N}$, since all f_i are continuous on at x_0, the partial sum $S_n = \sum_{k=1}^{n} f_k$ is also continuous. Since (S_n) converges uniformly to f on I, then f is also continuous at x_0. ■

It immediately follows that if the functions are continuous on the entire interval then the limit function is also continuous on the entire interval.

Theorem 4.4. *If $\sum_{n=1}^{\infty} f_n$ converges uniformly to f on $[a, b]$ and each f_n is integrable on $[a, b]$, then f is integrable on $[a, b]$ and for every $x \in [a, b]$ we have the uniform convergence*

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

Proof. For each $n \in \mathbb{N}$, since all f_i are integrable on $[a, b]$, the partial sum $S_n = \sum_{k=1}^{\infty} n f_k$ is also integrable. Since (S_n) converges uniformly to f on $[a, b]$, then f is also integrable on $[a, b]$, and $\int_a^x S_n$ converges uniformly to $\int_a^x f$. ■

Example 4.3. Consider the series $\sum_{n=0}^{\infty} \frac{x^n}{n!}$. For any $r > 0$, we have $\|f_n\|_{[-r,r]} \le \frac{r^n}{n!}$. By the ratio test, the series $\sum_{n=0}^{\infty} \frac{r^n}{n!}$ converges, so by the M-test, the given series converges uniformly on $[-r, r]$.

Now define $F(x) = \int_0^x \sum_{n=0}^{\infty} \frac{t^n}{n!} \mathrm{d}t$. Uniform convergence tells us that F converges uniformly to $\sum_{n=0}^{\infty} \frac{1}{n!} \frac{x^{n+1}}{n+1} = \sum_{n=0}^{\infty} \frac{x^n}{n!} - 1$. ♢

Theorem 4.5. *Suppose that $\sum_{n=1}^{\infty} f_n(x_0)$ converges for some $x_0 \in [a, b]$, and $\sum_{n=1}^{\infty} f_n'$ converges uniformly on $[a, b]$, then $\sum_{n=1}^{\infty} f_n$ converges uniformly on $[a, b]$ to a differentiable function f and for $x \in [a, b]$*

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

Proof. For each $n \in \mathbb{N}$, let $S_n = \sum_{k=1}^{n} f_k$. Since $(S_n(x_0))$ converges, and (S_n') converges uniformly on $[a, b]$, then (S_n) converges uniformly to f on $[a, b]$ and $f' = \lim_{n\to\infty} S_n'$. ■

Example 4.4. Let $f_n(x) = (-1)^n \frac{1}{\sqrt{n}} \cos \frac{x}{n}$. Then $\sum_{n=1}^{\infty} f_n(0) = \sum_{n=1}^{\infty} \frac{(-1)^n}{\sqrt{n}}$ converges by the alternating series test. Furthermore, $\sum_{n=1}^{\infty} f_n'(x) = \sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n^{\frac{3}{2}}} \sin \frac{x}{n}$. This converges uniformly on $[-r, r]$ by the M-test (skipped). Then we know that the series converges uniformly to some differentiable function f on $[-r, r]$. ♢

Theorem 4.6. *There exists a function $f: \mathbb{R} \to \mathbb{R}$ such that f is continuous at $\mathbb{R}$ but is not differentiable at any point of $\mathbb{R}$.*

Proof. Let $f(x) = |x|$ for $x \in [-1, 1]$. Extend f to the entire real line by $f(x+2k) = f(x)$ for all $k \in \mathbb{Z}$. For $x, y \in [-1, 1]$, we have $|f(x) - f(y)| = ||x| - |y|| \le |x - y|$. Due to the periodicity this is actually true for all $x, y \in \mathbb{R}$.

Now for $n \in \mathbb{N}$ let $f_n(x) = \left(\frac{3}{4}\right)^n f(4^n x)$, and consider $\sum_{n=0}^{\infty} f_n$. Since $|f_n(x)| \le \left(\frac{3}{4}\right)^n$, so by the M-test the series converges uniformly on $\mathbb{R}$. Furthermore since all f_n are continuous, the sum function is continuous as well.

Let $a \in \mathbb{R}$. There is an integer in either $(4^n a - \frac{1}{2}, 4^n a)$, or $(4^n a, 4^n a + \frac{1}{2})$. Define h_n such that there is no integer between $4^n a$ and $4^n a + 4^n h_n$ (so h_n takes the form of $\pm\frac{4^{-n}}{2}$). Now let $g_n = \frac{f(a+h_n) - f(a)}{h_n}$. If $f'(a)$ exists, then g_n will converge to $f'(a)$. We claim that (g_n) diverges and so $f'(a)$ does not exist.

Firstly $f(a + h_m) - f(a) = \sum_{n=0}^{\infty}\left(\frac{3}{4}\right)^n(f(4^n a + 4^n h_m) - f(4^n a))$. Since $4^n h_m$ is even for $n > m$, the periodicity means that we can reduce it to a finite sum since the difference goes to 0 for $n > m$. Then using the (reverse) triangle inequality

$$
\begin{aligned}
|f(a + h_m) - f(a)| &= \left|\sum_{n=0}^{m}\left(\frac{3}{4}\right)^n (f(4^n a + 4^n h_m) - f(4^n a))\right| \\
&\ge \left(\frac{3}{4}\right)^m |f(4^m a + 4^m h_m) - f(4^m a)| - \sum_{n=0}^{m-1}\left(\frac{3}{4}\right)^n |f(4^n a + 4^n h_m) - f(4^n a)| \\
&\ge \left(\frac{3}{4}\right)^m 4^m h_m - \sum_{n=0}^{m-1}\left(\frac{3}{4}\right)^n |4^n h_m| \\
&= |h_m|\left(3^m - \sum_{n=0}^{m-1} 3^n\right) \\
&= |h_m|\frac{3^m + 1}{2}.
\end{aligned}
$$

Thus $|g_m| \ge \frac{3^m+1}{2}$ which diverges. ∎

4.2 Power series

Definition 4.3 (Power series). A *power series* is a series of functions of the form $\sum_{n=0}^{\infty} a_n (x - x_0)^n$ where $x_0, a_1, a_2, \ldots$ are constants. Furthermore x_0 is called the *centre* of the power series. □

Every power series converges at at least one point, which is at the centre x_0.

Theorem 4.7. *Let $\sum_{n=0}^{\infty} a_n (x - x_0)^n$ be a power series.*

i. If it converges at x_1, then it is absolutely convergent for all x such that $|x - x_0| < |x_1 - x_0|$.

ii. If it diverges at x_2, then it is divergent for all x such that $|x - x_0| > |x_2 - x_0|$.

Proof.

i. We are given that $\sum_{n=0}^{\infty} a_n(x_1 - x_0)^n$ converges. Then $\lim_{n\to\infty} a_n(x_1 - x_0)^n = 0$, so $(a_n(x_1 - x_0)^n)$ is a bounded sequence. Let it be bounded by M. Now let x such that $|x - x_0| < |x_1 - x_0|$. We want to apply the comparison test. Consider

$$|a_n(x - x_0)^n| = |a_n(x_1 - x_0)^n| \left|\frac{(x - x_0)^n}{(x_1 - x_0)^n}\right|$$

Due to the conditions on x, we see that $r = \left|\frac{x-x_0}{x_1-x_0}\right| < 1$. Therefore we have $|a_n(x - x_0)^n| < Mr^n$. But the series $\sum_{n=0}^{\infty} Mr^n$ converges as it is a geometric series. Therefore by the comparison test, $\sum_{n=0}^{\infty} a_n(x - x_0)^n$ converges uniformly.

ii. Suppose $\sum_{n=0}^{\infty} a_n(x_2 - x_0)^n$ diverges. Let x be such that $|x - x_0| > |x_2 - x_0|$. If $\sum_{n=0}^{\infty} a_n(x - x_0)^n$ converges, then by the earlier part, the series should converge at x_2, which is a contradiction.

■

This means that every power series converges either at only one point or in some interval. This leads us to the following definition:

Definition 4.4 (Radius of convergence)**.** Given a power series, let $S = \{|x - x_0| \mid \text{the series converges at } x\}$. Then the *radius of convergence* for the series is defined as $R = \sup S$. □

Theorem 4.8. *A power series $\sum_{n=0}^{\infty} a_n(x - x_0)^n$ with a radius of convergence of R*

i. converges absolutely for all $x \in (x_0 - R, x_0 + R)$, and

ii. diverges for all x where $|x - x_0| > R$.

Proof.

i. Let $x \in (x_0 - R, x_0 + R)$, so $|x - x_0| < R$. Then there exists x_1 such that $|x_1 - x_0| \in S$. Since $|x - x_0| < |x_1 - x_0|$ and the series converges at x_1, the series converges absolutely at x.

ii. Assume $|x - x_0| > R$. If the series converges at x, then $|x - x_0| \in S$ and $|x - x_0| > R$ which is a contradiction.

■

Note that the theorem makes no statement about the end points. We will need to use other methods to determine the behaviour there.

Theorem 4.9. *Let $\sum_{n=0}^{\infty} a_n(x - x_0)^n$ be a power series. Suppose that $a_n \neq 0$. Consider the limit $L = \lim_{n\to\infty}\left|\frac{a_{n+1}}{a_n}\right|$.*

i. If the L exists, then the radius of convergence of the series is given by $R = \frac{1}{L}$ if $L > 0$ and $R = \infty$ if $L = 0$.

ii. If $L = \infty$ then $R = 0$.

Proof. For $x \neq x_0$, apply the ratio test to the series:

$$\begin{aligned}\lim_{n\to\infty}\left|\frac{a_{n+1}(x-x_0)^{n+1}}{a_n(x-x_0)^n}\right| &= \lim_{n\to\infty}\left|\frac{a_{n+1}}{a_n}\right||x-x_0| \\ &= L|x-x_0|\end{aligned}$$

Therefore when $|x-x_0| < \frac{1}{L}$, the series converges, and when $|x-x_0| > \frac{1}{L}$ the series diverges. ■

Example 4.5. Consider $\sum_{n=1}^{\infty}\frac{(-1)^{n+1}}{n}(x-1)^n$. We have $L = \lim_{n\to\infty}\left|\frac{a_{n+1}}{a_n}\right| = \lim_{n\to\infty}\frac{n}{n+1} = 1$. Therefore the series converges absolutely on $(0, 2)$. ♢

Theorem 4.10 (Cauchy-Hadamard formula)**.** *Let* $\sum_{n=0}^{\infty}a_n(x-x_0)^n$ *be a power series and let* $L = \limsup |a_n|^{\frac{1}{n}}$*. The radius of convergence is given by*

$$R = \begin{cases} 0, & \textit{if } L = \infty \\ \frac{1}{L}, & \textit{if } 0 < L < \infty \\ \infty, & \textit{if } L = 0 \end{cases}$$

Proof. Apply the ratio test to the series.

$$\limsup |a_n(x-x_0)^n|^{\frac{1}{n}} = |x-x_0| \limsup |a_n|^{\frac{1}{n}} = |x-x_0|L$$

The root test tells us that the power series converges everywhere if $L = 0$, and if $L = \infty$ it diverges everywhere. Otherwise, if $0 < L < \infty$ then for $|x-x_0| < \frac{1}{L}$, the series converges absolutely, whereas for $|x-x_0| > \frac{1}{L}$ the series diverges. ■

Example 4.6. Consider the series $\sum_{n=0}^{\infty}x^n \sin\frac{n\pi}{4}$. Some of the coefficients are zero, so $\lim_{n\to\infty}\left|\frac{a_{n+1}}{a_n}\right|$ does not exist. We try the Cauchy-Hadamard formula instead. Notice that there exists a constant subsequence converging to 1, for example $(\sin\frac{\pi}{2}, \sin\frac{5\pi}{2}, \dots)$. But $\left|\sin\frac{n\pi}{4}\right|^{\frac{1}{n}} \leq 1$ for all n so $\limsup |a_n|^{\frac{1}{n}} = 1$. The radius of convergence is therefore 1.

At the two endpoints $x = \pm 1$ the series clearly diverges, since the $(\pm 1)^n \sin\frac{n\pi}{4}$ never tends to 0. ♢

Example 4.7. Consider the series $\sum_{n=0}^{\infty}\frac{1}{2^n}x^{2n} = 1 + \frac{x^2}{2} + \frac{x^4}{2^2} + \cdots$. We have $|a_{2n}|^{\frac{1}{2n}} = \frac{1}{\sqrt{2}}$. Therefore $\limsup |a_n|^{\frac{1}{n}} = \frac{1}{\sqrt{2}}$. Alternatively the ratio test gives $\lim_{n\to\infty}\left|\frac{a_{n+1}}{a_n}\right| = \frac{x^2}{2}$ which gives the same radius of convergence of $\sqrt{2}$.

At the endpoints where $x = \pm\sqrt{2}$ the terms in the series become 1 so it clearly diverges there. ♢

4.3 Properties of power series

Power series can be used to create functions with desirable properties.

Definition 4.5 (Sum function)**.** Given a power series $\sum_{n=0}^{\infty}a_n(x-x_0)^n$ with a radius of convergence $R > 0$, we can define a function $f\colon (x_0 - R, x_0 + R) \to \mathbb{R}$ by $f(x) = \sum_{n=0}^{\infty}a_n(x-x_0)^n$. We call f the *sum function* of the series. □

Theorem 4.11. *Let $\sum_{n=0}^{\infty} a_n(x-x_0)^n$ have a radius of convergence $R > 0$ and let a and b be such that $x_0 - R < a < b < x_0 + R$. Then $\sum_{n=0}^{\infty} a_n(x-x_0)^n$ converges uniformly on $[a, b]$.*

Proof. Consider a larger interval that is symmetric about x_0, given by $[x_0 - r, x_0 + r]$, where $r = \max(|a - x_0|, |b - x_0|)$.

Evaluated at $x = x_0 + r$, the series becomes $\sum_{n=0}^{\infty} a_n r^n$ and we know that it converges absolutely. Now for $x \in [a, b]$, since $|x - x_0| \leq r$, we have $|a_n(x-x_0)^n| \leq |a_n r^n|$. Therefore, $\|a_n(x-x_0)\|_{[a,b]} \leq |a_n r^n|$. By the Weierstrass M-test, the power series converges uniformly on $[a, b]$. ■

We immediately get that the sum function is continuous on $[a, b]$. Continuity is preserved even after infinite unions, so in fact the sum function is continuous on the entire interval $(x_0 - R, x_0 + R)$.

Theorem 4.12. *Suppose that $f(x) = \sum_{n=0}^{\infty} a_n(x-x_0)^n$ has a radius of convergence $R > 0$. Then f is infinitely differentiable on $(x_0 - R, x_0 + R)$, and furthermore (term by term differentiation)*

$$f^{(k)}(x) = \sum_{n=k}^{\infty} n(n-1)\dots(n-k+1)a_n(x-x_0)^{n-k}$$

and it has a radius of convergence R.

Proof. Write $f_n(x) = a_n(x-x_0)^n$. We have $f_0'(x) = 0$ and $f_n'(x) = na_n(x-x_0)^{n-1}$.

We check the radius of convergence of $\sum_{n=1}^{\infty} na_n(x-x_0)^{n-1}$. This has the same radius of convergence as $\sum_{n=0}^{\infty}{}^{\infty} na_n(x-x_0)^n$. Now $\limsup |na_n|^{\frac{1}{n}} = \limsup n^{\frac{1}{n}}|a_n|^{\frac{1}{n}} = \limsup |a_n|^{\frac{1}{n}} = \frac{1}{R}$ so $\sum f_n'$ has the same radius of convergence as the original series. Furthermore it converges uniformly on any closed subinterval of $(x_0 - R, x_0 + R)$.

Hence by theorem 4.5, f is differentiable on any closed subinterval of $(x_0 - R, x_0 + R)$, and $f'(x) = \sum f_n'(x)$. Taking the union of all closed subintervals we see that it is also true for the entire interval $(x_0 - R, x_0 + R)$.

We can then inductively apply this for all higher derivatives. ■

Although a power series and its derivatives may have the same radius of convergence, their behaviours at the endpoints might differ. Consider the following example.

Example 4.8. Let $f(x) = \sum_{n=1}^{\infty} \frac{x^n}{n^2}$. Since $\lim_{n\to\infty} \frac{a_{n+1}}{a_n} = 1$, the radius of convergence $R = 1$. At $x = \pm 1$ it also converges.

Now consider $f'(x) = \sum_{n=1}^{\infty} \frac{x^{n-1}}{n}$. It still has a radius of convergence $R = 1$. Now at $x = 1$ we get the harmonic series which diverges. ♢

Corollary 4.12.1. *If $f(x) = \sum_{n=0}^{\infty} a_n(x-x_0)^n$ for $x \in (x_0 - r, x_0 + r)$ for some $r > 0$, then $a_k = \frac{f^{(k)}(x_0)}{k!}$.*

Proof. Let R be the radius of convergence of $\sum_{n=0}^{\infty} a_n(x-x_0)^n$. Then for $r \leq R$, by the previous theorem

$$\begin{aligned} f^{(k)}(x) &= \sum_{n=k}^{\infty} n(n-1)\dots(n-k+1)a_n(x-x_0)^{n-k} \\ &= k!a_k + \frac{(k+1)!}{2}a_{k+1}(x-x_0) + \cdots . \end{aligned}$$

■

This leads directly to the following corollary.

Corollary 4.12.2 (Uniqueness of power series). *If $\sum_{n=0}^{\infty} a_n(x-x_0)^n = \sum_{n=0}^{\infty} b_n(x-x_0)^n$ on some interval $(x_0 - r, x_0 + r)$ for $r > 0$, then $a_n = b_n$ for all n.*

Example 4.9. Consider the power series $\sum_{n=0}^{\infty} \frac{x^{3n}}{n+1}$. We skip the checking but it converges on $[-1, 1)$. We want to find a closed form for its sum function. We want to relate it to the geometric series, and to do this we need to remove the coefficient. One way of doing this is through differentiating. Here is a neat trick:

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}x} x^3 f(x) &= \frac{\mathrm{d}}{\mathrm{d}x} \sum_{n=0}^{\infty} \frac{x^3(n+1)}{n+1} \\
&= \sum_{n=0}^{\infty} 3x^{3n+2} \\
&= 3x^2 \sum_{x=0}^{\infty} x^{3n} \\
&= \frac{3x^2}{1-x^3}
\end{aligned}$$

Integrating, we find that

$$t^3 f(t)\Big|_0^x = \int_0^x \frac{3t^2}{1-t^3} = -\ln\left(1-x^3\right).$$

So $f(x) = -\frac{\ln\left(1-x^3\right)}{x^3}$ and $f(0) = \sum_{n=0}^{\infty} \frac{0^{3n}}{n+1} = 1$. ◊

Theorem 4.13. *Let $f(x) = \sum_{n=0}^{\infty} a_n(x-x_0)^n$ have a radius of convergence $R > 0$. Then for any a and b such that $x_0 - R < a < b < x_0 + R$,*

$$\int_a^b f(x) = \sum_{n=0}^{\infty} \int_a^b a_n(x-x_0)^n.$$

Proof. Theorem 4.4. ■

Example 4.10. The geometric series gives us $\sum_{n=0}^{\infty} (-x)^n = \frac{1}{1+x}$. Integrating, we get $\ln(1+x) = \sum_{n=0}^{\infty} \frac{(-1)^n}{n+1} x^{n+1}$ for $-1 < x < 1$. ◊

This also holds for indefinite integrals. Define $F(x) = \int_{x_0}^{x} f(t) = \sum_{n=0}^{\infty} \int_{x_0}^{x} a_n(t-x_0)^n = \sum_{n=0}^{\infty} \frac{a_n}{n+1}(x - x_0^{n+1}$. If we differentiate this we get the original series back, so they must have the same radius of convergence.

Theorem 4.14 (Abel's formula). *Let (b_n) and (c_n) be sequences of real numbers, and for integers n and m such that $n \geq m \geq 1$, let*

$$B_{nm} = \sum_{k=m}^{n} b_k.$$

Then

$$\sum_{k=m}^{n} b_k c_k = B_{nm} c_n - \sum_{k=m}^{n-1} B_{km}(c_{k+1} - c_k).$$

Proof. Fix $m \geq 1$. For $k > m$, we have $B_{km} - B_{k-1,m} = b_k$ and $b_m = B_{mm}$. Hence for $n > m \geq 1$, we have

$$\begin{aligned}\sum_{k=m}^{n} b_k c_k &= b_m c_m + \sum_{k=m+1}^{n} (B_{km} - B_{k-1,m}) c_k \\ &= B_{nm} c_m + \sum_{k=m+1}^{n} B_{km} c_k - \sum_{k=m+1}^{n} B_{k-1,m} c_k \\ &= B_{nm} c_n + \sum_{k=m}^{n-1} B_{km} c_k - \sum_{k=m}^{n} B_{km} c_{k+1} \\ &= B_{nm} c_n - \sum_{k=m}^{n-1} B_{km} (c_{k+1} - c_k).\end{aligned}$$

■

Theorem 4.15 (Abel's theorem). *Suppose that the power series $\sum_{n=0}^{\infty} a_n (x - x_0)^n$ has a finite radius of convergence $R > 0$.*

i. If the series converges at $x = x_0 + R$ then it converges uniformly on $[x_0, x_0 + R]$.

ii. If the series converges at $x = x_0 - R$ then it converges uniformly on $[x_0 - R, x_0]$.

Proof. Assume that $\sum_{n=0}^{\infty} a_n (x - x_0)^n$ converges at $x = x_0 + R$. In other words $\sum_{n=0}^{\infty} a_n R^n$ converges.

Let $x_1 \in (x_0, x_0 + R]$.

$$\begin{aligned}\sum_{n=0}^{\infty} a_n (x_1 - x_0)^n &= \sum_{n=0}^{\infty} a_n R^n \left(\frac{x_1 - x_0}{R}\right)^n \\ &= \sum_{n=0}^{\infty} b_n c_n\end{aligned}$$

where $b_n = a_n R^n$ and $c_n = \left(\frac{x_1 - x_0}{R}\right)^n$. Now $x_1 - x_0 < R$ so (c_n) is a decreasing sequence. Furthermore, since $\sum_{n=0}^{\infty} b_n$ converges, by the Cauchy criterion

$$\forall \epsilon > 0, \exists K \in \mathbb{N}, \forall n, m \left[n \geq m \geq K \implies \left| \sum_{j=m}^{n} b_j \right| < \epsilon \right].$$

Now applying Abel's formula, for $n \geq m \geq K$,

$$\begin{aligned}\left| \sum_{k=m}^{n} a_k (x_1 - x_0)^k \right| &= \left| B_{nm} c_n + \sum_{k=m}^{n-1} B_{km} (c_k - c_{k+1}) \right| \\ &\leq |B_{nm} c_n| + \sum_{k=m}^{n-1} |B_{km} (c_k - c_{k+1})| \\ &= c_n \left| \sum_{j=m}^{n} b_j \right| + \sum_{k=m}^{n-1} (c_k - c_{k+1}) \left| \sum_{j=m}^{k} b_j \right| \\ &< c_n \epsilon + \sum_{k=m}^{n-1} (c_k - c_{k+1}) \epsilon \\ &= c_m \epsilon \\ &\leq \epsilon.\end{aligned}$$

■

Corollary 4.15.1. *Suppose that $f(x) = \sum_{n=0}^{\infty} a_n(x - x_0)^n$ has a radius of convergence $R > 0$. Then*

i. If the series converges at $x = x_0 + R$,

$$\lim_{x \to (x_0+R)^-} f(x) = \sum_{n=0}^{\infty} a_n R^n$$

ii. If the series converges at $x = x_0 - R$,

$$\lim_{x \to (x_0-R)^-} f(x) = \sum_{n=0}^{\infty} (-1)^n a_n R^n$$

Proof. By Abel's theorem $f(x)$ converges uniformly on $[x_0, x_0+R]$ so the sum function is continuous at $x = x_0 + R$. Hence $\lim_{x \to (x_0+R)^-} f(x) = f(x_0 + R) = \sum_{n=0}^{\infty} a_n R^n$. ■

Theorem 4.16 (Dirichlet's test)**.** *Let (f_n) and (g_n) be sequences of functions on E and suppose that $\forall n \in \mathbb{N}, \forall x \in E$,*

- $\exists M > 0 \left[\left| \sum_{k=1}^{n} f_k(x) \right| \le M \right]$,
- *(g_n) converges uniformly on E to the zero function,*
- *$(g_n(x))$ is a decreasing sequence.*

Then the series $\sum_{n=1}^{\infty} f_n g_n$ converges uniformly on E.

Proof. Let

$$F_{nm}(x) = \sum k = m^n f_k(x).$$

Then we have

$$|F_{nm}(x)| = \left| \sum_{k=1}^{n} f_k(x) - \sum_{k=1}^{m-1} f_k(x) \right| \le \left| \sum_{k=1}^{n} f_k(x) \right| + \left| \sum_{k=1}^{m-1} f_k(x) \right| \le 2M$$

By Abel's formula,

$$\begin{aligned} \left| \sum_{k=m}^{n} f_k(x) g_k(x) \right| &= \left| F_{nm}(x) g_n(x) - \sum_{k=m}^{n-1} F_{km}(g_{k+1}(x) - g_k(x) \right| \\ &= |F_{nm}(x)| g_n(x) + \sum_{k=m}^{n-1} |F_{km}|(g_k(x) - g_{k+1}(x) \\ &\le 2M g_n(x) + \sum_{k=m}^{n-1} 2M(g_k(x) - g_{k+1}(x)) \\ &= 2M g_m(x). \end{aligned}$$

For the second step we note that $g_n(x)$ is positive since it is decreasing and it converges to 0. Now let $\epsilon > 0$. Since g_n converges uniformly on E,

$$\exists K \in \mathbb{N} \left[n \ge K \implies |g_n(x) - 0| < \frac{\epsilon}{2M} \right].$$

Therefore

$$n \ge m \ge K \implies \left| \sum_{k=m}^{n} f_k(x) g_k(x) \right| \le 2M \frac{\epsilon}{2M} = \epsilon$$

■

4.4 Taylor series and Maclaurin series

Definition 4.6 (Taylor series)**.** Let f be infinitely differentiable on $(x_0 - r, x_0 + r)$ for some $r > 0$. The power series

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

is called the *Taylor series of f about x_0*. If $x_0 = 0$ then we call it the *Maclaurin series.* □

Note that this only defines a Taylor series for an infinitely differentiable function but it does not say anything about convergence or equivalence. However, once we can show that the function has a power series representation, then it must be equivalent to the Taylor series.

Example 4.11. We have seen how $\ln(1 + x) = \sum_{n=0}^{\infty} \frac{(-1)^n}{n+1} x^{n+1}$ for $-1 < x < 1$. We want to find a Taylor series about x_0 for $f(x) = \ln(x)$.

$$\begin{aligned}\ln x &= \ln(x_0 + x - x_0) \\ &= \ln x_0(1 + \frac{x - x_0}{x_0}) \\ &= \ln x_0 + \sum_{n=0}^{\infty} \frac{(-1)^n}{n+1}\left(\frac{x - x_0}{x_0}\right)^{n+1}\end{aligned}$$

This converges if $\left|\frac{x-x_0}{x_0}\right| < 1$, or in other words $0 < x < 2x_0$. ♢

Recall Taylor's theorem. It tells us that we can write a function as a sum of two polynomials, $f(x) = P_n(x) + R_n(x)$. Notice that $P_n(x)$ is just a partial sum of the Taylor series. When f is infinitely differentiable, we can let n go to infinity to get the following theorem.

Theorem 4.17. *Let f be infinitely differentiable on $I = (x_0 - r, x_0 + r)$ and let $x \in I$. Then*

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

iff

$$\lim_{n\to\infty} R_n(x) = \lim_{n\to\infty} \frac{f^{(n+1)}(c_n)}{(n + 1)!}(x - x_0)^{n+1} = 0$$

where the c_n are between x and x_0.

Corollary 4.17.1. *If f is infinitely differentiable on $(-r, r)$ and if there is $B > 0$ such that $\left|f^{(n)}(x)\right| < B$ for all $x \in (-r, r)$, then the Maclaurin series of f converges to $f(x)$.*

Proof. Consider the remainder term:

$$\begin{aligned}|R_n(x)| &= \frac{\left|f^{(n+1)}(c_n)\right|}{(n + 1)!} x^{n+1} \\ &\leq \frac{B}{(n + 1)!} r^{n+1}.\end{aligned}$$

Since $\lim_{n\to\infty} \frac{B}{(n+1)!} r^{n+1} = 0$, by the squeeze theorem $\lim_{n\to\infty} R_n = 0$, and hence the Maclaurin series of f converges to $f(x)$. ■

Definition 4.7 (Analytic functions)**.** A function f is *analytic* on (a, b) if f is infinitely differentiable on (a, b) and for any $x_0 \in (a, b)$ the Taylor series of f about x_0 converges to f in a neighbourhood of x_0. □

Theorem 4.18. *If the Taylor series of a function f centred at $x_0 \in (a, b)$ converges to f on an open interval (a, b), then f is analytic on (a, b).*

4.5 Arithmetic on power series

Addition of two series, and multiplication by a single number is easily defined. However the multiplication of two series is not so obvious. Term by term multiplication for finite sums motivate the Cauchy product.

Definition 4.8 (Cauchy product)**.** The *Cauchy product* of $\sum_{n=0}^{\infty} a_n$ and $\sum_{n=0}^{\infty} b_n$ is the series $\sum_{n=0}^{\infty} c_n$ where for each n

$$c_n = \sum_{k=0}^{\infty} a_k b_{n-k} = \sum_{i+j=n} a_i b_j$$

□

Theorem 4.19 (Mertens' theorem)**.** *If $\sum_{n=0}^{\infty} a_n$ converges absolutely and $\sum_{n=0}^{\infty} b_n$ converges, then their Cauchy product converges and is given by*

$$\left(\sum_{n=0}^{\infty} a_n\right)\left(\sum_{n=0}^{\infty} b_n\right).$$

Proof. Let $\alpha = \sum_{n=0}^{\infty} |a_n|$ and $A = \sum_{n=0}^{\infty} a_n$ and $B = \sum_{n=0}^{\infty} b_n$. Furthermore denote $A_k = \sum_{n=0}^{k} a_n$ and $B_k = \sum_{n=0}^{k} b_n$ and $C_k = \sum_{n=0}^{k} c_n$ where $c_n = \sum_{i+j=n} a_i b_j$. (A_k) converges to A, (B_k) converges to B, and we want to show that (C_k) converges to AB.

Consider the following

$$\begin{aligned}
C_k &= (a_0 b_0) + (a_0 b_1 + a_1 b_0) + (a_0 b_2 + a_1 b_1 + a_2 b_0) + \cdots + (a_0 b_k + a_1 b_{k-1} + \cdots + a_k b_0) \\
&= a_0(b_0 + b_1 + \cdots + b_k) + a_1(b_0 + b_1 + \cdots + b_{k-1}) + \cdots + (a_k b_0) \\
&= a_0 B_k + a_1 B_{k-1} + \cdots + a_n B_0 \\
&= a_0(B + \beta_k) + a_1(B + \beta_{k-1}) + \cdots + a_k(B + \beta_0) \\
&= (a_0 + a_1 + \cdots + a_k)B + a_0 \beta_k + a_1 \beta_{k-1} + \cdots + a_k \gamma_0 \\
&= A_n B + \beta_k,
\end{aligned}$$

where $\beta_k = B_k - B$ and $\gamma_k = a_0 \beta_k + \ldots a_k \beta_0$. Hence now we need to show that $\lim_{k\to\infty} \gamma_k = 0$.

Let $\epsilon > 0$. Since $\lim_{n\to\infty} \beta_n = 0$, $\exists i \in \mathbb{N} \left[n \geq i \implies |\beta_n| < \frac{\epsilon}{2\alpha}\right]$. Furthermore, since it is a convergent sequence, it is bounded, so $\exists M > 0, \forall n \left[|\beta_n| \leq M\right]$. Furthermore, we also note that $\sum_{n=0}^{\infty} |a_j|$ satisfies the Cauchy criterion, so $\exists k \in N\left[\sum_{m=k-i+1}^{k}|a_m| < \frac{\epsilon}{2M}\right]$.

$$\begin{aligned}
|\gamma_k| &= |a_0\beta_k + \cdots + a_k\beta_0| \\
&\le |a_0||\beta_k| + \ldots |a_1||\beta_{k-1}| + \cdots + |a_{k-i}||\beta_i| + \cdots + |a_k|\beta_0 \\
&\le \left(\sum_{n=0}^{k-i}|a_n|\right)\frac{\epsilon}{2\alpha} + \left(\sum_{m=k-i+1}^{k}|a_m|\right)M \\
&\le \alpha\frac{\epsilon}{2\alpha} + \frac{\epsilon}{2M}M \\
&= \epsilon.
\end{aligned}$$

■

Corollary 4.19.1. *If both* $\sum_{n=0}^{\infty}a_n$ *and* $\sum_{n=0}^{\infty}b_n$ *converge absolutely then their Cauchy product converges absolutely.*

Proof. Let

$$c_n = \sum_{k=0}^{\infty}a_k b_{n-k} \qquad\qquad d_n = \sum_{k=0}^{\infty}|a_k||b_{n-k}|.$$

Since $\sum_{n=0}^{\infty}|a_n|$ converges absolutely and $\sum_{n=0}^{\infty}b_n$ converges, by Merten's theorem $\sum_{n=0}^{\infty}d_n$ converges. However notice that

$$|c_n| = \left|\sum_{k=0}^{n}a_k b_{n-k}\right| \le \sum_{k=0}^{n}|a_k||b_{n-k}| = d_n$$

which means by the comparison test, $\sum_{n=0}^{\infty}|c_n|$ converges. Hence the Cauchy product converges absolutely. ■

Theorem 4.20. *Let* $f(x) = \sum_{n=0}^{\infty}a_n(x-x_0)^n$ *for* $|x-x_0| < R_1$ *and* $g(x) = \sum_{n=0}^{\infty}b_n(x-x_0)^n$ *for* $|x-x_0| < R_2$*. Let* α *and* β *be constants. Then*

i. $\alpha f(x) + \beta g(x) = \sum_{n=0}^{\infty}(\alpha a_n + \beta b_n)(x-x_0)^n$ *for* $|x-x_0| < \min(R_1, R_2)$.

ii. $f(x)g(x) = \sum_{n=0}^{\infty}c_n(x-x_0)^n$ *where* $c_n = \sum_{k=0}^{\infty}a_k b_{n-k}$ *for* $|x-x_0| < \min(R_1, R_2)$.

Proof. We will only show part (ii). Let x be such that $|x-x_0| < \min(R_1, R_2)$. Let us rewrite $f(x) = \sum_{n=0}^{\infty}\alpha_n$ and $g(x) = \sum_{n=0}^{\infty}\beta_n$. They converge absolutely, and by Merten's theorem their Cauchy product converges to $\sum_{n=0}^{\infty}\gamma_n = \left(\sum_{n=0}^{\infty}\alpha_n\right)\left(\sum_{n=0}^{\infty}\beta_n\right)$ where

$$\begin{aligned}
\gamma_n &= \sum_{k=0}^{n}\alpha_k\beta_{n-k} \\
&= \sum_{k=0}^{n}a_k(x-x_0)^k b_{n-k}(x-x_0)^{n-k} \\
&= \sum_{k=0}^{n}a_k b_{n-k}(x-x_0)^n \\
&= c_n(x-x_0)^n
\end{aligned}$$

■

5 Appendix

This section is a brief recap of concepts covered in Analysis I.

5.1 Preliminaries

Definition 5.1 (Increasing functions). Let $f: I \to R$. Function f is said to be *increasing* on I if $x_1, x_2 \in I\, x_1 < x_2 \implies f(x_1) \leq f(x_2)$. □

A similar definition follows for *decreasing* functions. Since we allow for equality, constant functions are increasing and decreasing functions. If we are strict on the inequality, then we call it a *strictly* increasing (decreasing) function.

5.2 Limits of functions

A quit refresher on limits. Intuitively a function f has a limit L at a if $f(x) \approx L$ as we approach $x = a$.

Definition 5.2 (Limits of functions). A function f is said to have a limit L at $x = a$ if

$$\forall \epsilon > 0, \exists \delta > 0\ [|x - a| < \delta \implies |f(x) - L| < \epsilon].$$

□

Note that the value of $f(a)$ is not important for the limit, only those points around it.

Recall all the other types of limits we can make by switching out the terms in the implication sign.

Limit taken at	Antecedent	Limit equals	Consequent
$x \to a$	$\|x - a\| < \delta$	$L \in \mathbb{R}$	$\|f(x) - L\| < \epsilon$
$x \to a^+$	$a < x < a + \delta$	∞	$f(x) > M$
$x \to a^-$	$a - \delta < x < a$	$-\infty$	$f(x) < M$
$x \to \infty$	$\exists k\ x > k$		
$x \to -\infty$	$\exists k\ x < k$		

An equivalent definition is the following.

Definition 5.3 (Sequential criterion for limits). For a function f, $\lim_{x \to a} f(x) = L$ iff (x_n) is a sequence in the domain of f such that $x_n \neq a$ for all n and $x_n \to a$, then $f(x_n) \to L$. □

Some quick facts:

Theorem 5.1. *Suppose* $\lim_{x \to a} f(x) = L$ *exists.*

i. If $L > 0$*, then*

$$\exists \delta > 0\ [0 < |x - a| < \delta \implies f(x) > 0]$$

ii. If $L \neq 0$, then

$$\exists \delta > 0\ [0 < |x - a| < \delta \implies f(x) \neq 0]$$

Proof.

i. Since the limit exists,

$$\forall \epsilon > 0, \exists \delta > 0\ [|x - a| < \delta \implies |f(x) - L| < \epsilon].$$

Pick $0 < \epsilon < L$. Then

$$\exists \delta > 0\ [|x - a| < \delta \implies (0 < f(x) - L < \epsilon) \vee (0 < L - f(x) < \epsilon) \implies f(x) > 0].$$

ii. Do the same as above but for $L < 0$, then we get the result we want.

■

Theorem 5.2. *Suppose that function f is defined in a deleted neighbourhood of a point c and $L \in \mathbb{R}$. Then $\lim_{x \to c} f(x) = L$ iff $\lim_{x \to c^+} f(x) = L = \lim_{x \to c^-} f(x)$.*

Proof. ($\implies$) follows directly from the definition.

($\impliedby$): Let δ_+ and δ_- be the witnesses for the limits $x \to c^+$ and $x \to c^-$ respectively, then let $\delta = \min(\delta_+, \delta_-)$, and we have

$$\forall \epsilon > 0, \exists \delta > 0, c + \delta_- \leq c + \delta < x < c + \delta \leq c + \delta_+ \implies |f(x) - L| < \epsilon. \tag{1}$$

■

Theorem 5.3. *Let $f\colon (0,1) \to \mathbb{R}$ and $L \in \mathbb{R}$, then $\lim_{x \to 0^+} f(x) = L$ iff $\lim_{y \to \infty} f(\frac{1}{y}) = L$.*

Proof. Let $y = \frac{1}{x}$, then

$$\begin{aligned} &\forall \epsilon > 0, \exists \delta > 0 \left[0 < \frac{1}{y} < \delta \implies \left|f\left(\frac{1}{y}\right) - L\right| < \epsilon\right] \\ \iff &\forall \epsilon > 0, \exists \delta > 0 \left[y > \frac{1}{\delta} \implies \left|f\left(\frac{1}{y}\right) - L\right| < \epsilon\right] \end{aligned}$$

■

5.3 Continuity

Definition 5.4 (Continuity)**.** A function f is continuous at a if $\lim_{x \to a} f(x) = f(a)$. □

Definition 5.5 (Uniform continuity)**.** A function f is *uniformly continuous* on an interval I if

$$\forall \epsilon > 0, \exists \delta > 0, \forall x, y \in I\ [|x - y| < \delta \implies |f(x) - f(y)| < \epsilon].$$ □

Theorem 5.4 (Extreme value theorem)**.** *If f is continuous on $[a, b]$, then*

$$\exists x_1, x_2 \in [a, b], \forall x \in [a, b]\ [f(x_1) \leq f(x) \leq f(x_2)]$$

Theorem 5.5 (Intermediate value theorem)**.** *If f is continuous on $[a, b]$, and $f(a) < k < f(b)$, then there exists $c \in (a, b)$ such that $f(c) = k$.*

5.4 Series

Definition 5.6 (Infinite series). Given a sequence (a_n), define a new sequence (s_n) where $s_n = a_1 + \cdots + a_n = \sum_{k=1}^{n} a_k$. We call it the *infinite series* generated by (a_n) and denote it as $\sum_{n=1}^{\infty} a_n$.

If (s_n) converges to s, then we say $\sum_{n=1}^{\infty} a_n$ converges and define $\sum_{n=1}^{\infty} a_n = s$. Otherwise we say that it diverges. ☐

Theorem 5.6. *If $\sum_{n=1}^{\infty} a_n$ converges, then $\lim_{n\to\infty} a_n = 0$.*

The converse is false, with the harmonic series being a famous counterexample: $\sum_{n=1}^{\infty} \frac{1}{n}$.

Since series are just sequences, then theorems on sequences apply to series as well.

Theorem 5.7 (Cauchy criterion for series). *$\sum_{n=1}^{\infty} a_n$ converges iff*

$$\forall \epsilon > 0, \exists K \in \mathbb{N} \left[n > m \geq K \implies \left| \sum_{k=1}^{n} a_k - \sum_{k=1}^{m} a_k \right| = \left| \sum_{k=m+1}^{n} a_k \right| < \epsilon \right]$$

Definition 5.7 (Positive series). A *positive series* $\sum_{n=1}^{\infty} a_n$ is a series where $a_n \geq 0$ for all n. ☐

Theorem 5.8 (Comparison test). *Let (a_n) and (b_n) be positive series and suppose $\exists K \in \mathbb{N}, \forall n \geq K\, [a_n \leq b_n]$, then if $\sum_{n=1}^{\infty} b_n$ converges, so does $\sum_{n=1}^{\infty} a_n$.*

Theorem 5.9 (Alternating series test). *If (a_n) is a positive and decreasing sequence, then the series $\sum_{n=1}^{\infty} (-1)^n a_n$ converges.*

Definition 5.8 (Absolute convergence). Let (a_n) be a sequence, if $\sum_{n=1}^{\infty} |a_n|$ converges, then we say that $\sum_{n=1}^{\infty} a_n$ converges *absolutely.* ☐

Theorem 5.10. *Absolute convergence implies convergence.*

Theorem 5.11 (Ratio test). *Suppose that all the terms of the series $\sum_{n=1}^{\infty} a_n$ are non-zero and the limit $L = \lim_{n\to\infty} \left| \frac{a_{n+1}}{a_n} \right|$ exists.*

- *If $L < 1$ then the series converges absolutely.*
- *If $L > 1$ then the series diverges.*
- *There is no conclusion if $L = 1$.*

Theorem 5.12 (Root test). *Consider the series $\sum_{n=1}^{\infty} a_n$. Let $L = \limsup |a_n|^{\frac{1}{n}}$.*

i. If $L < 1$ then the series converge absolutely.

ii. If $L > 1$ then the series diverges.

iii. There is no conclusion if $L = 1$.

There is a weaker version of the root test involving just $\lim n \to \inf |a_n|^{\frac{1}{n}}$. But do note that this limit might not exist, whereas the limit superior always exists (even if though it might be ∞).