Math 140B - Notes

Neil Donaldson

February 26, 2025

1 Continuity

The overarching goal of this course and its prequel is to make elementary calculus rigorous. We begin with a review of some basic concepts and conventions.

Sets & Functions In these notes, essentially all functions have the form $f : U \to V$ where both U, V are subsets of the real numbers $\mathbb{R}$. To f are associated several concepts:

Domain $\text{dom}(f) = U$; the *inputs* to f. Often *implied* to be the largest set on which a formula is defined. In calculus examples, the domain is typically a union of (open) intervals.

Codomain $\text{codom}(f) = V$; the *potential outputs* of f. By convention, $V = \mathbb{R}$ unless necessary.

Range $\text{range}(f) = f(U) = \{f(x) : x \in U\}$; the *realized outputs* of f and a subset of V.

Injectivity f is *injective/one-to-one* if $f(x) = f(y) \implies x = y$: distinct inputs produce distinct outputs.

Surjectivity f is *surjective/onto* if $f(U) = V$: all potential outputs are realized.

Inverses f is *bijective/invertible* if it is injective and surjective. Equivalently, $\exists f^{-1} : V \to U$ satisfying

$$\forall u \in U,\ f^{-1}\big(f(u)\big) = u \quad \text{and} \quad \forall v \in V,\ f\big(f^{-1}(v)\big) = v$$

Example 1.1. The function defined by $f(x) = \frac{1}{x(x-2)}$ has implied

$$\text{dom}(f) = \mathbb{R} \setminus \{0, 2\} = (-\infty, 0) \cup (0, 2) \cup (2, \infty)$$

$$\text{range}(f) = (-\infty, -1] \cup (0, \infty)$$

The function is neither injective nor surjective. By restricting the domain & codomain, we obtain a bijection:

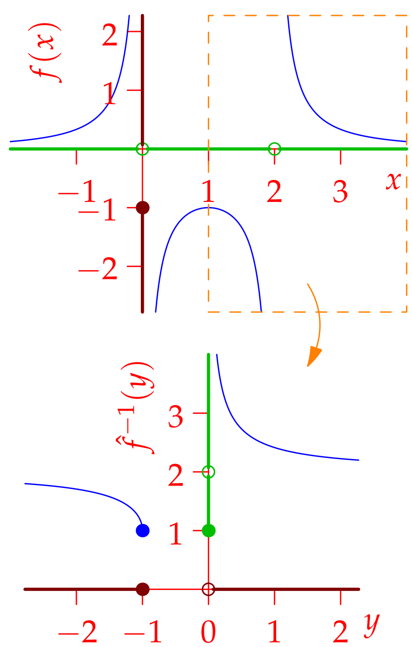
$$\text{dom}(\hat{f}) = [1, 2) \cup (2, \infty)$$

$$\text{codom}(\hat{f}) = (-\infty, -1] \cup (0, \infty)$$

with inverse

$$\hat{f}^{-1}(y) = \begin{cases} 1 + y^{-1}\sqrt{y+1} & \text{if } y > 0 \\ 1 - y^{-1}\sqrt{y+1} & \text{if } y \le -1 \end{cases}$$

Now $\text{dom}(\hat{f}^{-1}) = \text{codom}(\hat{f})$ and $\text{codom}(\hat{f}^{-1}) = \text{dom}(\hat{f})$.



Suprema and Infima A set $U \subseteq \mathbb{R}$ is *bounded above* if it has an *upper bound* M:

$$\exists M \in \mathbb{R} \text{ such that } \forall u \in U,\ u \leq M$$

Axiom 1.2 (Completeness). If $U \subseteq \mathbb{R}$ is non-empty and bounded above, then it has a *least upper bound*, the *supremum* of U

$$\sup U = \min\{M \in \mathbb{R} : \forall u \in U,\ u \leq M\}$$

By convention, $\sup U := \infty$ if U is unbounded above and $\sup \emptyset := -\infty$; now every subset of $\mathbb{R}$ has a supremum. Similarly, the *infimum* of U is its *greatest lower bound*:

$$\inf U = \begin{cases} \max\{m \in \mathbb{R} : \forall u \in U,\ u \geq m\} & \text{if } U \neq \emptyset \text{ is bounded below} \\ -\infty & \text{if } U \neq \emptyset \text{ is unbounded below} \\ \infty & \text{if } U = \emptyset \end{cases}$$

Examples 1.3. Here are four sets with their suprema and infima stated. You should be able to verify these assertions directly from the definitions.

U	$\{1,2,3,4\}$	$(0,5)$	$(-\infty,\pi]$	$\mathbb{R}$	$\{\frac{1}{n} : n \in \mathbb{N}\}$
$\sup U$	4	5	π	∞	1
$\inf U$	1	0	$-\infty$	$-\infty$	0

Note how the supremum/infimum might or might not lie in the set itself.

Interiors, closures, boundaries and neighborhoods These last concepts might not be review, but they will be used repeatedly.

Definition 1.4. Let $U \subseteq \mathbb{R}$. A value $a \in \mathbb{R}$ is *interior* to U if it lies in some open subinterval of U:

$$\exists \delta > 0 \text{ such that } (a-\delta, a+\delta) \subseteq U$$

A *neighborhood* of a is any set to which a is interior: the interval $(a-\delta, a+\delta)$ is an *open δ-neighborhood* of a. A *punctured neighborhood* of a is a neighborhood with a deleted.

The set of points interior to U is denoted U°.

A *limit point* of U is the limit of some sequence $(x_n) \subseteq U$. The *closure* $\overline{U}$ is the set of limit points.

The *boundary* of U is the set $\partial U = \overline{U} \setminus U^\circ$.

Examples 1.5. 1. If $U = [1,3)$, then $U^\circ = (1,3)$, $\overline{U} = [1,3]$ and $\partial U = \{1,3\}$.

2. $\mathbb{Q}^\circ = \emptyset$ and $\partial\mathbb{Q} = \overline{\mathbb{Q}} = \mathbb{R}$.

3. $(-3,5) \cup (5,7]$ is a punctured neighborhood of 5.

1.17 Continuity of Functions

Everything in this section *should* be review.

Definition 1.6. A function $f : U \to \mathbb{R}$ is *continuous at* $u \in U$ if either/both of the following hold:

1. For all sequences $(x_n) \subseteq U$ converging to u, the sequence $(f(x_n))$ converges to $f(u)$.
2. $\forall \epsilon > 0, \exists \delta > 0$ such that $(\forall x \in U)$, $|x - u| < \delta \implies |f(x) - f(u)| < \epsilon$.

A function f is *continuous on* U if it is continuous at every point $u \in U$.

Examples 1.7. 1. We prove that $f(x) = x^3$ is continuous at $u = 2$.

(a) (Limit method) Let $x_n \to 2$. By the *limit laws* (i.e. $\lim(x_n^k) = (\lim x_n)^k$),

$$\lim f(x_n) = \lim x_n^3 = \left(\lim x_n\right)^3 = 2^3 = f(2)$$

(b) (ϵ–δ method) Let $\epsilon > 0$ be given and let $\delta = \min\left(1, \frac{\epsilon}{19}\right)$.

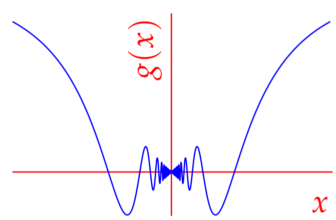
$$|x - 2| < \delta \implies |x - 2| < 1 \implies 1 < x < 3$$

from which

$$\left|x^3 - 2^3\right| = |x - 2| \left|x^2 + 2x + 2^2\right| < 19\,|x - 2| \le \epsilon$$

where we used the triangle inequality.

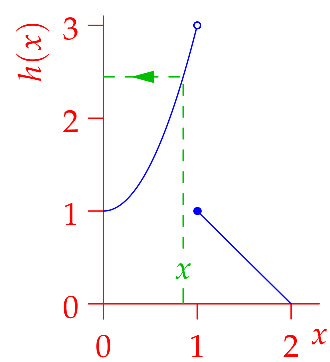
2. Let $g(x) = \begin{cases} x \sin \frac{1}{x} & \text{if } x \neq 0, \\ 0 & \text{if } x = 0 \end{cases}$



Then g is continuous at $x = 0$. Again this can be done with limits or an ϵ–δ argument; both are essentially the *squeeze theorem*.

3. The function defined by

$$h(x) = \begin{cases} 1 + 2x^2 & \text{if } x < 1 \\ 2 - x & \text{if } x \geq 1 \end{cases}$$



is discontinuous at $x = 1$.

(a) The sequence with $x_n = 1 - \frac{1}{n}$ converges to 1, yet

$$\lim h(x_n) = 3 \neq 1 = h(1)$$

(b) Choose $\epsilon = 1$ and suppose $\delta > 0$ is given. Now choose $x = \max\{1 - \frac{\delta}{2}, \frac{1}{\sqrt{2}}\}$ to see that

$$|x - 1| < \delta \quad \text{and} \quad |h(x) - h(1)| \geq 1 = \epsilon$$

Theorem 1.8. *The two parts of Definition 1.6 are equivalent.*

Proof. (1 ⇒ 2) We prove the contrapositive. Suppose condition 2 is *false*; that is,

$$\exists \epsilon > 0, \text{ such that } \forall \delta > 0, \exists x \in U \text{ with } |x-u| < \delta \text{ and } |f(x) - f(u)| \geq \epsilon$$

In particular, for any $n \in \mathbb{N}$ we may let $\delta = \frac{1}{n}$ to obtain

$$\exists \epsilon > 0, \text{ such that } \forall n \in \mathbb{N}, \exists x_n \in U \text{ with } |x_n - u| < \frac{1}{n} \text{ and } |f(x_n) - f(u)| \geq \epsilon$$

The sequence (x_n) shows that condition 1 is *false*:

- $\forall n,\ |x_n - u| < \frac{1}{n}$ whence $x_n \to u$.
- $\forall n,\ |f(x_n) - f(u)| \geq \epsilon > 0$, whence $f(x_n)$ *does not converge to* $f(u)$.

(2 ⇒ 1) Suppose condition 2 is true, that $(x_n) \subseteq U$ converges to u and that $\epsilon > 0$ is given. Then

$$\exists \delta > 0 \text{ such that } |x-u| < \delta \implies |f(x) - f(u)| < \epsilon$$

However, by the definition of convergence ($x_n \to u$),

$$\exists N \in \mathbb{N} \text{ such that } n > N \implies |x_n - u| < \delta \implies |f(x_n) - f(u)| < \epsilon$$

Otherwise said, $f(x_n) \to f(u)$. ∎

Rather than use these definitions every time, it is helpful to have a working dictionary.

Theorem 1.9 (Dictionary of Common Continuous Functions).

1. *Suppose f and g are continuous at u and that k is constant. Then the following are continuous at u (if defined—don't divide by zero!):*

$$f+g, \quad f-g, \quad fg, \quad \frac{f}{g}, \quad |f|, \quad kf, \quad \max(f,g), \quad \min(f,g)$$

2. *If f is continuous at u and h is continuous at $f(u)$, then $h \circ f$ is continuous at u.*
3. *Algebraic functions are continuous: these are functions constructed using finitely many addition/subtraction, multiplication/division and n^{th} root operations.*
4. *The familiar transcendental functions are continuous:* exp, ln, sin, *etc.*

Example 1.10. $f(x) = \sin \frac{\sqrt[3]{x^2+7}}{x-2} + \cos \frac{1}{e^x - 1}$ is continuous on its domain $(-\infty, 0) \cup (0,1) \cup (1, \infty)$.

Theses claims are tedious to prove using elementary definitions. In particular, it is better to defer a proof of the transcendental claim until we can define such functions using power series, after which continuity comes for free.

Exercises 1.17. *Key concepts/results: Suprema/Completeness, Sequential & ϵ-δ continuity*

1. Give examples to show that $g \circ f$ being continuous can happen with:

   (a) f continuous and g discontinuous.

   (b) g continuous and f discontinuous.

   (c) Both f, g discontinuous.

   You may use pictures, but make sure they clearly describe the functions f, g.

2. (a) Prove that the function $f(x) = x^3$ is continuous at $x = -2$ using an ϵ–δ argument.

   (b) Prove that $f(x) = x^3$ is continuous at $x = u$ using an ϵ–δ argument.

3. Prove that the following are discontinuous at $x = 0$: use *both* definitions of continuity.

   (a) $f(x) = 1$ for $x < 0$ and $f(x) = 0$ for $x \geq 0$.

   (b) $g(x) = \sin(1/x)$ for $x \neq 0$ and $g(0) = 0$.

4. If f is continuous at u, prove that it is bounded on some set $(u - \delta, u + \delta) \cap \operatorname{dom}(f)$.

5. Prove the following parts of Theorem 1.9 using ϵ–δ arguments.

   (a) If f, g are continuous at u, then $f - g$ is continuous at u.

   (b) If f, g are continuous at u, then fg is continuous at u.

   (c) If f is continuous at u and h at $f(u)$, then $h \circ f$ is continuous at u.

6. Suppose $f : U \to \mathbb{R}$ is a function whose domain U contains an *isolated point* a: i.e. $\exists r > 0$ such that $(a - r, a + r) \cap U = \{a\}$. Prove that f is continuous at a.

7. Refresh your prerequisites by giving formal proofs:

   (a) (Suprema and sequences) If $M = \sup U$, then $\exists (x_n) \subseteq U$ such that $x_n \to M$.
   (*This has to work even if* $M = \infty$*!*)

   (b) (Limit of a bounded sequence) If $(x_n) \subseteq [a, b]$ and $x_n \to x$, then $x \in [a, b]$.

   (c) (Bolzano–Weierstraß) Every bounded sequence in $\mathbb{R}$ has a convergent subsequence.
   (*Hint: If* $(x_n) \subseteq [a, b]$*, explain why there exist intervals* $I_1 \supseteq I_2 \supseteq I_3 \supseteq \cdots$ *such that* infinitely many (x_n) *lie in each interval* I_k*. Hence obtain a subsequence* (x_{n_k}) *and prove that it is* Cauchy.[1])

8. (Very Hard) Consider the function $f : \mathbb{R} \to \mathbb{R}$ where

$$f(x) = \begin{cases} \frac{1}{q} & \text{whenever } x = \frac{p}{q} \in \mathbb{Q} \text{ with } q > 0 \text{ and } \gcd(p, q) = 1 \\ 0 & \text{if } x \notin \mathbb{Q} \end{cases}$$

   For example, $f(1) = f(2) = f(-7) = 1$, and $f(\frac{1}{2}) = f(-\frac{1}{2}) = f(\frac{3}{2}) = \cdots = \frac{1}{2}$, etc. Prove that f is continuous at each point of $\mathbb{R} \setminus \mathbb{Q}$ and discontinuous at each point of $\mathbb{Q}$.

   (*Hint: for continuity, consider* $A = \{r \in \mathbb{Q} : f(r) \geq \frac{1}{q}\}$ *where* $q \geq \frac{1}{\epsilon} \ldots$)

[1] This is a good moment to review Cauchy completeness: that a sequence is convergent if and only if it is *Cauchy*:

$$\forall \epsilon > 0, \exists N \text{ such that } m, n > N \implies |x_m - x_n| < \epsilon$$

1.18 Properties of Continuous Functions

In this section we describe the behavior of a continuous function on an interval. We first consider the special case when the domain is a closed bounded interval $[a, b]$.

Theorem 1.11 (Extreme Value Theorem). *A continuous function on a closed, bounded interval is bounded and attains its bounds. Otherwise said, if $f : [a, b] \to \mathbb{R}$ is continuous, then*

$$\exists x, y \in [a, b] \text{ such that } f(x) = \sup \operatorname{range}(f) \text{ and } f(y) = \inf \operatorname{range}(f)$$

In particular, the supremum and infimum are finite.

Proof. Suppose f is continuous with domain $[a, b]$ and let $M = \sup\{f(x) : x \in [a, b]\}$. We invoke the three parts of Exercise 1.17.7:

- (Part a) There exists a sequence $(x_n) \subseteq [a, b]$ such that $f(x_n) \to M$.
- (Part c) There exists a convergent subsequence (x_{n_k}) with limit x.
- (Part b) $x \in [a, b]$.

Since f is continuous, we now have $f(x) = \lim_{k\to\infty} f(x_{n_k}) = M$. This shows that M is *finite* and that f attains its least upper bound. For the lower bound, apply this to $-f$. ∎

It is worth considering how the result can fail when one of the hypotheses is weakened. For example:

f discontinuous $f : [0, 1] \to \mathbb{R} : x \mapsto \begin{cases} x & \text{if } x \neq 1 \\ 0 & \text{if } x = 1 \end{cases}$ is bounded but does not attain its bounds.

$\operatorname{dom}(f)$ *not closed* $f : [0, 1) \to \mathbb{R} : x \mapsto x$ is bounded but does not attain its bounds.

$\operatorname{dom}(f)$ *not bounded* $f : [0, \infty) \to \mathbb{R} : x \mapsto x$ is unbounded.

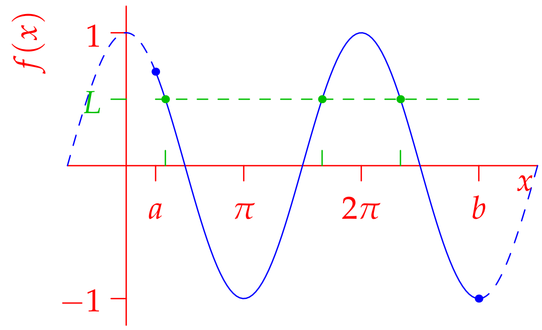
We now consider continuous functions on arbitrary intervals. The next result should be familiar from elementary calculus and is intuitively obvious from the naïve notion of continuity (draw the graph without taking your pen off the page).

Theorem 1.12 (Intermediate Value Theorem). *Let $f : I \to \mathbb{R}$ be continuous on an interval I. Suppose $a, b \in I$ with $a < b$ and that $f(a) \neq f(b)$. If L lies between $f(a)$ and $f(b)$, then $\exists \xi \in (a, b)$ such that $f(\xi) = L$.*

Example 1.13. Let $f(x) = \cos x$ with $a = \frac{\pi}{4}$, $b = 3\pi$ and $L = \frac{1}{2}$; then

$$f(\xi) = L \iff \xi \in \left\{ \frac{\pi}{3}, \frac{5\pi}{3}, \frac{7\pi}{3} \right\}$$

There may therefore be several suitable values of ξ. It is even possible (Exercise 2) for there to be *infinitely many*.

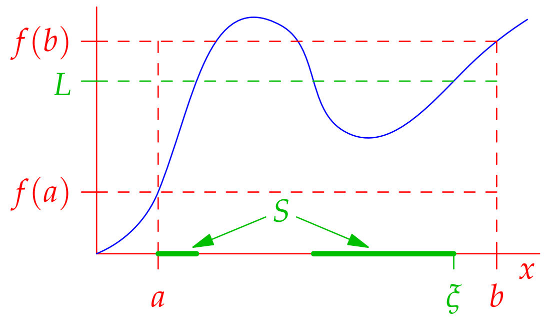


Proof. Suppose WLOG that $f(a) < L < f(b)$ and let

$$S = \{x \in [a, b] : f(x) < L\}$$

Plainly $S \subseteq [a, b)$ is non-empty, hence $\xi := \sup S$ exists and $\xi \in [a, b]$. It remains to show that ξ satisfies the required properties.

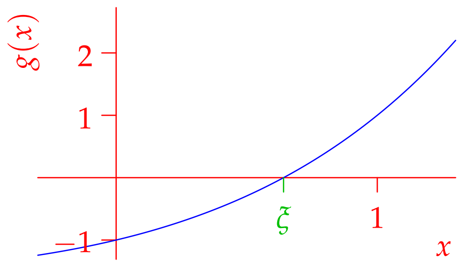
By Exercise 7, $\exists (s_n) \subseteq S$ with $\lim s_n = \xi$. Since f is continuous, $f(\xi) = \lim f(s_n) \leq L$. In particular, $\xi \neq b$.



To finish, play a similar game with the sequence defined by $t_n = \min\{b, \xi + \frac{1}{n}\}$ (see Exercise 4). ∎

Example 1.14. The intermediate value theorem is useful for demonstrating the existence of solutions to equations. For example, we show that the equation $x2^x = 1$ has a solution.

- Observe that $g(x) = x2^x - 1$ is continuous.
- $g(0) = -1 < 0$.
- $g(1) = 1 > 0$.
- By the intermediate value theorem $\exists \xi \in (0, 1)$ such that $g(\xi) = 0$: that is $\xi \cdot 2^\xi = 1$.



It is inefficient, but one can home in on ξ by repeatedly halving the size of the interval: for instance,

$$g(\tfrac{1}{2}) = \tfrac{\sqrt{2}}{2} - 1 < 0, \quad g(\tfrac{3}{4}) = \tfrac{3}{4} \cdot 2^{3/4} - 1 \approx 0.26 > 0 \ldots \implies \tfrac{1}{2} < \xi < \tfrac{3}{4}$$

Corollary 1.15. *Continuous functions map intervals to intervals (or points).*

Proof. An interval I is characterized by the following property

$$\forall x_1, x_2 \in I,\ x \in \mathbb{R},\ x_1 < x < x_2 \implies x \in I$$

Let $f : I \to \mathbb{R}$ be continuous and suppose its range $f(I)$ is not a single point. If $f(a) < L < f(b)$, then $\exists \xi$ between a, b such that $f(\xi) = L$. Otherwise said, $L \in f(I)$ and so $f(I)$ is an interval. ∎

More generally, if $\text{dom}(f) = \bigcup I_n$ is written as a union of disjoint intervals and f is continuous, then

$$\text{range}(f) = \bigcup f(I_n)$$

is also a union of intervals, though these need not be disjoint: a continuous function can bring intervals together, but cannot break them apart.[2]

Example 1.16. The function $f(x) = \sqrt{x^2 - 4}$ has implied domain $(-\infty, -2] \cup [2, \infty)$ and range $[0, \infty)$. Both halves of the domain are mapped onto the same interval range(f) .

Exercises 1.18. *Key concepts: Extreme Value Theorem, Intermediate Value Theorem*
Continuous functions preserve intervals

1. Give examples of the following:
   (a) An unbounded discontinuous function on a closed bounded interval.
   (b) An unbounded continuous function on a non-closed bounded interval.
   (c) A bounded continuous function on a closed unbounded interval which fails to attain its bounds.
2. Consider the function $f(x) = \begin{cases} x \sin \frac{1}{x} & \text{if } x \neq 0 \\ 0 & \text{if } x = 0 \end{cases}$
   (a) Explain why f is continuous on any interval I.
   (b) Suppose $a < 0 < b$ and that $f(a), f(b)$ have opposite signs. If $L = 0$, show that the intermediate value theorem is satisfied by *infinitely many* distinct values ξ.
3. Use the intermediate value theorem to prove that the equation $8x^3 - 12x^2 - 2x + 1 = 0$ has at least 3 real solutions (and thus, by the fundamental theorem of algebra, exactly 3).
4. Complete the proof of the intermediate value theorem by defining $t_n = \min(b, \xi + \frac{1}{n})$.
5. (a) Suppose $f : U \to \mathbb{R}$ is continuous and that $U = \bigcup_{k=1}^{n} I_k$ is the union of a finite sequence (I_k) of closed bounded intervals. Prove that f is bounded and attains its bounds.
   (b) Let $U = \bigcup_{n=1}^{\infty} I_n$, where $I_n = [\frac{1}{2n}, \frac{1}{2n-1}]$ for each $n \in \mathbb{N}$. Give an example of a continuous function $f : U \to \mathbb{R}$ which is either unbounded or does not attain its bounds. Explain.

[2]More generally, if $f : U \to V$ is a continuous function between topological spaces and a, b lie in the same *component* of U, then $f(a), f(b)$ lie in the same component of $f(U)$. In our examples each component is an interval.

1.19 Uniform Continuity

Recall Definition 1.6: $f : U \to \mathbb{R}$ is continuous at all points[3] $y \in U$ provided

$$\forall y \in U,\ \forall \epsilon > 0,\ \exists \delta > 0 \text{ such that } (\forall x \in U)\ |x - y| < \delta \implies |f(x) - f(y)| < \epsilon$$

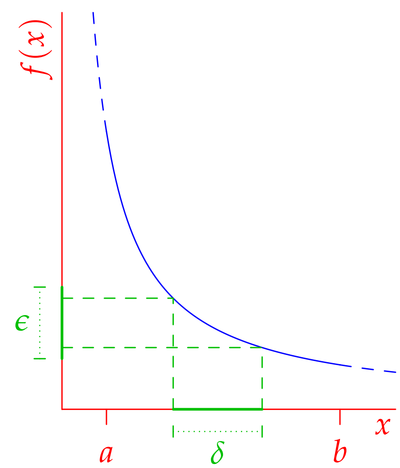
Note the order of the quantifiers: δ is permitted to depend *both* on y and ϵ. In the naïve sense of continuity (x close to $y \implies f(x)$ close to $f(y)$), the meaning of *close* can depend on the *location* y. Uniform continuity is a stronger condition where the **meaning of *close* is independent of location**.

Definition 1.17. $f : U \to \mathbb{R}$ is *uniformly continuous* if

$$\forall \epsilon > 0,\ \exists \delta > 0 \text{ such that } (\forall x, y \in U)\ |x - y| < \delta \implies |f(x) - f(y)| < \epsilon$$

We've included the (typically) hidden quantifiers ($\forall x, y$) to make clear that δ is independent of x, y. Note also that the definition is now symmetric in x, y.

Example 1.18. Consider $f(x) = \frac{1}{x}$.



1. If $0 < a < b \le \infty$, then f is uniformly continuous on $[a, b)$.

   Let $\epsilon > 0$ be given and let $\delta = a^2\epsilon$. Then $\forall x, y \in [a, b)$,

$$|x - y| < \delta \implies \left|\frac{1}{x} - \frac{1}{y}\right| = \left|\frac{y - x}{xy}\right| < \frac{\delta}{xy} \le \frac{\delta}{a^2} = \epsilon$$

2. If $0 < b \le \infty$, then f is *not* uniformly continuous on $(0, b)$.

   Let $\epsilon = 1$ and suppose $\delta > 0$ is given.

   Let $x = \min(\delta, 1, \frac{b}{2})$ and $y = \frac{x}{2}$.

   Certainly $x, y \in (0, b)$ and $|x - y| = \frac{x}{2} \le \frac{\delta}{2} < \delta$. However,

$$|f(x) - f(y)| = \frac{1}{x} \ge 1 = \epsilon$$

Think about how ϵ and δ must relate as one slides the intervals in the picture up/down and left/right.

Some intuition will help make sense of the example.

Bounded/unbounded gradient In part 1 $\epsilon = \delta a^2$, where $\frac{1}{a^2} = |f'(a)|$ bounds the gradient of f. By contrast, the slope of f is *unbounded* in part 2.

Extensibility In part 1 the domain of f may be extended to a (and to b if finite): $g : [a, b] \to \mathbb{R} : x \mapsto \frac{1}{x}$ is continuous. In part 2, this is impossible: there is no continuous function $g : [0, b) \to \mathbb{R}$ such that $g(x) = \frac{1}{x}$ whenever $x > 0$.

Informally, if a continuous function f has bounded gradient, or if you can 'fill in the holes' at the endpoints of $\mathrm{dom}(f)$, then f is uniformly continuous. When uniform continuity is used abstractly in a proof, it is often one of the above properties that is being invoked. The remainder of this section involves making these observations watertight.

[3] To promote symmetry, we use y instead of u for a generic point of $\mathrm{dom}(f)$.

Theorem 1.19. *Let $f : I \to \mathbb{R}$ be continuous on an interval I and differentiable with bounded derivative on the interior I°. Then f is uniformly continuous on I.*

The proof depends on the mean value theorem, which should be familiar from elementary calculus; we'll discuss a proof later on.

Proof. Suppose $|f'(x)| \le M$ on I°. Let $\epsilon > 0$ be given, let $\delta = \frac{\epsilon}{M}$ and suppose $(y, x) \subseteq I$. Then

$$|x - y| < \delta \implies \exists \xi \in I^\circ \text{ such that } f'(\xi) = \frac{f(x) - f(y)}{x - y} \qquad \text{(MVT)}$$
$$\implies |f(x) - f(y)| = |f'(\xi)|\,|x - y| < M\delta = \epsilon$$

∎

Theorem 1.19 isn't a biconditional: for instance, Exercise 1.19.5 shows that $f(x) = \sqrt{x}$ on $[0, \infty)$ and $g(x) = x^{1/3}$ on $\mathbb{R}$ are both uniformly continuous even though both have unbounded slope.

We now discuss extensibility and how uniform continuity relates to continuity on closed sets. First we see that for closed bounded sets, uniform continuity is nothing new.

Theorem 1.20. *If $g : [a, b] \to \mathbb{R}$ is continuous, then it is uniformly continuous.*

Proof. Suppose g is continuous but not uniformly so. Then

$$\exists \epsilon > 0 \text{ such that } \forall \delta > 0,\ \exists x, y \in [a, b] \text{ for which } |x - y| < \delta \text{ and } |g(x) - g(y)| \ge \epsilon \qquad (*)$$

For each $n \in \mathbb{N}$, let $\delta = \frac{1}{n}$ to see that there exists sequences $(x_n), (y_n) \subseteq [a, b]$ satisfying the above. By Bolzano–Weierstraß, the bounded sequence (x_n) has a convergent subsequence $x_{n_k} \to x \in [a, b]$. Clearly

$$|x_{n_k} - y_{n_k}| < \frac{1}{n_k} \to 0 \implies y_{n_k} \to x$$

But then $|g(x_{n_k}) - g(y_{n_k})| \to 0$, which contradicts $(*)$.

∎

Now we build towards a partial converse.

Lemma 1.21. *If $f : U \to \mathbb{R}$ is uniformly continuous and $(x_n) \subseteq U$ is a Cauchy sequence, then $\big(f(x_n)\big)$ is also Cauchy.*

Proof. Let $\epsilon > 0$ be given. Then:

- (Uniform Continuity) $\exists \delta > 0$ such that $|x - y| < \delta \implies |f(x) - f(y)| < \epsilon$.
- (Cauchy) $\exists N \in \mathbb{N}$ such that $m, n > N \implies |x_m - x_n| < \delta$.

Putting these together, we see that

$$\exists N \in \mathbb{N} \text{ such that } m, n > N \implies |f(x_m) - f(x_n)| < \epsilon$$

Otherwise said, $\big(f(x_n)\big)$ is Cauchy.

∎

We now see that a function $f : I \to \mathbb{R}$ is uniformly continuous on a bounded interval if and only if it has a *continuous extension* $g : \overline{I} \to \mathbb{R}$ defined on the closure of its domain.

Theorem 1.22. *Suppose $f : I \to \mathbb{R}$ is continuous where I is a bounded interval with endpoints $a < b$. Define $g : [a, b] \to \mathbb{R}$ via*

$$g(x) = \begin{cases} f(x) & \text{if } x \in I \\ \lim f(x_n) & \text{whenever } (x_n) \subseteq I \text{ and } x_n \to a \\ \lim f(x_n) & \text{whenever } (x_n) \subseteq I \text{ and } x_n \to b \end{cases}$$

Then f is uniformly continuous if and only g is well-defined (g is continuous, if *well-defined).*

Proof. ($\Rightarrow$) Suppose f is uniformly continuous on I and that $a \notin I$. Let $(x_n), (y_n) \subseteq I$ be sequences converging to a. To show that g is well-defined, we must prove that $\big(f(x_n)\big)$ and $\big(f(y_n)\big)$ are convergent, and to the same limit. For this, we define a sequence

$$(u_n) = (x_1, y_1, x_2, y_2, x_3, y_3, \ldots)$$

Since (x_n) and (y_n) have the same limit a, we conclude that $u_n \to a$. But then (u_n) is Cauchy. By Lemma 1.21, $\big(f(u_n)\big)$ is also Cauchy and thus convergent. Since $\big(f(x_n)\big)$ and $\big(f(y_n)\big)$ are subsequences of a convergent sequence, they must also converge to the same (finite) limit.

The argument when $b \notin I$ is identical.

($\Leftarrow$) If g is well-defined then it is continuous (Definition 1.6, part 1); by Theorem 1.20 it is uniformly so. Since $f = g$ on a subset of $\operatorname{dom}(g)$, the same choice of δ will work for f as for g: f is therefore uniformly continuous. ∎

Examples 1.23. 1. Consider $f : x \mapsto x^2$.

(a) If $\operatorname{dom}(f)$ is the open interval $(-3, 10)$, then f is uniformly continuous since its derivative $f'(x) = 2x$ is bounded ($|f'(x)| \le 20$). The continuous extension is $g(x) = x^2$ on $[-3, 10]$.

(b) If $\operatorname{dom}(f)$ is the infinite interval $(-3, \infty)$, then neither Theorem 1.19 nor 1.22 applies: both f' and the domain $(-3, \infty)$ are unbounded.
Instead, note that if $\epsilon = 1$, then for any $\delta > 0$, we can choose $x = \frac{1}{\delta}$ and $y = \frac{1}{\delta} + \frac{\delta}{2}$. Clearly

$$|x - y| = \frac{\delta}{2} < \delta \text{ and } \left|x^2 - y^2\right| = 1 + \frac{\delta^2}{4} > 1 = \epsilon$$

whence f is not uniformly continuous.

2. $f(x) = x \sin \frac{1}{x}$ is continuous on the interval $(0, \infty)$. Strictly, neither Theorem 1.19 nor 1.22 apply since the derivative

$$f'(x) = \sin \frac{1}{x} - \frac{1}{x} \cos \frac{1}{x}$$

is unbounded as is the domain. However, by breaking the domain into two pieces...

- On $[1, \infty)$, the derivative is bounded: $|f'(x)| \le 1 + \frac{1}{|x|} \le 2$ by the triangle inequality. Theorem 1.19 says f is uniformly continuous on $[1, \infty)$.

- f is continuous on $(0, 1]$ and, by the squeeze theorem

$$x_n \to 0^+ \implies \lim f(x_n) = 0$$

Extending f so that $f(0) = 0$ defines a continuous extension. By Theorem 1.22, f is uniformly continuous on $(0, 1]$.

Putting this together (Exercise 6), f is uniformly continuous on $(0, \infty)$. Indeed the function

$$h(x) = \begin{cases} x \sin \frac{1}{x} & \text{if } x \neq 0 \\ 0 & \text{if } x = 0 \end{cases}$$

is uniformly continuous on $\mathbb{R}$.

Exercises 1.19. *Key concepts: Uniform Continuity (same δ for all locations), Bounded gradient, Continuous extensions*

1. Which functions are uniformly continuous? Justify your answers.
   (a) $f(x) = x^4$ on $[-1, 1]$
   (b) $f(x) = x^4$ on $(-1, 1]$
   (c) $f(x) = x^{-4}$ on $(0, 2]$
   (d) $f(x) = x^{-4}$ on $(1, 2]$
   (e) $f(x) = x^2 \sin \frac{1}{x}$ on $(0, 1]$
2. Prove that each function is uniformly continuous by verifying the ϵ–δ property.
   (a) $f(x) = 2x - 14$ on $\mathbb{R}$
   (b) $f(x) = x^3$ on $[1, 5]$
   (c) $f(x) = x^{-1}$ on $(1, \infty)$
   (d) $f(x) = \frac{x+1}{x+2}$ on $[0, 1]$
3. Prove that $f(x) = x^4$ is not uniformly continuous on $\mathbb{R}$.
4. (a) Suppose f is uniformly continuous on a bounded interval I. Prove that f is bounded on I.
   (b) Use part (a) to write down a bounded interval on which the function $f(x) = \tan x$ is defined, but *not* uniformly continuous.
5. Both parts of this question are easy using Exercise 6. Do them explicitly using the ϵ–δ property.
   (a) Let $f(x) = \sqrt{x}$ with domain $[0, \infty)$. Show that $f'(x)$ is unbounded, but that f is still uniformly continuous on $[0, \infty)$.
   (*Hint: let $\delta = (\frac{\epsilon}{2})^3$ and consider the cases $x \geq y \geq 0$, $x \leq y \leq 0$ and $x > 0 > y$ separately*)
   (b) Prove that $g(x) = x^{1/3}$ is uniformly continuous on $\mathbb{R}$.
   (*Hint: let $\delta = \epsilon^2$ and WLOG assume $0 \leq y \leq x$. Now compute $(\sqrt{y} + \epsilon)^2 \ldots$*)
6. Suppose f is uniformly continuous on intervals U_1, U_2 for which $U_1 \cap U_2$ is non-empty. Prove that f is uniformly continuous on $U_1 \cup U_2$.
   (*Hint: if x, y do not lie in the same U_1, U_2, choose some $a \in U_1 \cap U_2$ between x and y*)

1.20 Limits of Functions

In elementary calculus you likely saw many calculations of the following form:

$$\lim_{x\to 3}\frac{x^2-9}{x-3}=\lim_{x\to 3}\frac{(x-3)(x+3)}{x-3}=\lim_{x\to 3}(x+3)=6$$

Loosely speaking, this means that if $(x_n)\subseteq\mathbb{R}\setminus\{3\}$ is a sequence converging to 3, then $\big(f(x_n)\big)$ converges to 6. Our goal in this section is to make this notation precise.

Definition 1.24. Suppose $f:U\to\mathbb{R}$, that $S\subseteq U$, and that a is the limit of some sequence in S. We say that L is the *limit of $f(x)$ as x tends to a along S*, written $\lim\limits_{x\to a^S}f(x)=L$, provided

$$\forall (x_n)\subseteq S,\ \lim x_n=a \implies \lim f(x_n)=L$$

We can now define one-sided and two-sided limits of functions:

Right-hand limit: $\lim\limits_{x\to a^+}f(x)=L$ means $\exists S=(a,b)\subseteq U$ for which $\lim\limits_{x\to a^S}f(x)=L$

Left-hand limit: $\lim\limits_{x\to a^-}f(x)=L$ means $\exists S=(c,a)\subseteq U$ for which $\lim\limits_{x\to a^S}f(x)=L$

Two-sided limit: $\lim\limits_{x\to a}f(x)=L$ means $\exists S=(c,a)\cup(a,b)\subseteq U$ for which $\lim\limits_{x\to a^S}f(x)=L$

- If $U=\operatorname{dom}(f)$ is unbounded, then the one-sided definitions apply when $a=\pm\infty$. We omit the $\pm$ modifiers: for instance,

  $$\lim_{x\to\infty}f(x)=L \iff \lim_{x\to\infty^S}f(x)=L \text{ for some } S=(c,\infty)\subseteq U$$

- Note that f need not be defined at a, though $U=\operatorname{dom}(f)$ must contain at least some *punctured neighborhood* of a (one-sided for a one-sided limit). This will certainly happen if U is a union of intervals of positive length. In such a case, one may simply replace S with $U\setminus\{a\}$ in the definition: this is precisely what we did in the motivating example where $U=\mathbb{R}\setminus\{3\}$.

  Moreover, in such a situation, Definition 1.6 recovers a familiar idea from elementary calculus:

  $$f \text{ is continuous at } a\in U \iff f(a)=\begin{cases}\lim\limits_{x\to a}f(x) & \text{when } a\in U^\circ\\ \lim\limits_{x\to a^\pm}f(x) & \text{when } a\in U\setminus U^\circ\end{cases}\qquad (*)$$

  Warning! When $\operatorname{dom}(f)$ does not contain a punctured neighborhood of a, the right hand side doesn't exist and the assertion is *false*!

- By modifying the proof of Theorem 1.8 when $a,L\in\mathbb{R}$ are finite, we can restate using ϵ-language. For instance, $\lim\limits_{x\to a}f(x)=L$ means

  $$\forall\epsilon>0,\ \exists\delta>0 \text{ such that } (\forall x\in\operatorname{dom}f)\ \ 0<|x-a|<\delta \implies |f(x)-L|<\epsilon$$

  If a and/or L is infinite, use the language of unboundedness: e.g., $\lim\limits_{x\to a}f(x)=\infty$ means

  $$\forall M>0,\ \exists\delta>0 \text{ such that } 0<|x-a|<\delta \implies f(x)>M$$

  There are *fifteen* distinct combinations: *three* two-sided and *six* each of the one-sided limits!

Examples 1.25. 1. Let $f(x) = \frac{2+x}{x}$ where $\text{dom}(f) = U = \mathbb{R} \setminus \{0\} = (-\infty, 0) \cup (0, \infty)$

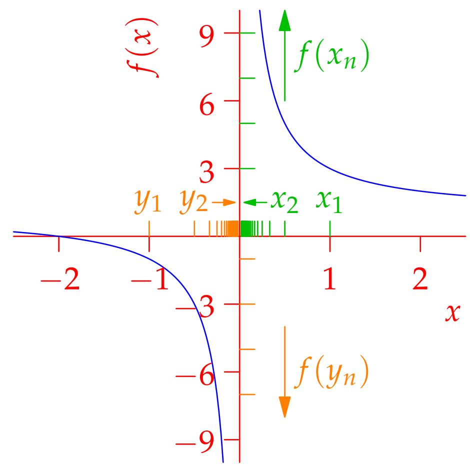
The following should be clear:

$$\lim_{x\to 3} f(x) = \frac{5}{3} \qquad \lim_{x\to\infty} f(x) = 1$$

To compute the first, for instance, we could choose $S = (0,3) \cup (3,\infty)$; if $(x_n) \subseteq S$ and $x_n \to 3$, then the limit laws justify the first claim

$$\lim_{n\to\infty} f(x_n) = \frac{2+3}{3} = \frac{5}{3}$$

as does the fact that f is continuous at $x = 3$. The second claim can be checked similarly.



We can take one-sided limits at $x = 0$:

$$\lim_{x\to 0^+} f(x) = \infty \quad \text{and} \quad \lim_{x\to 0^-} f(x) = -\infty$$

For instance, let $(x_n) \subseteq (0,\infty)$ satisfy $x_n \to 0$. Again, the limit laws show that $\lim_{n\to\infty} f(x_n) = \infty$, which is enough to justify the first claim.

Finally, the sequences defined by $x_n = \frac{1}{n}$ and $y_n = -\frac{1}{n}$ both lie in $S = \mathbb{R} \setminus \{0\}$ and converge to zero, yet

$$\lim_{n\to\infty} f(x_n) = \infty \neq -\infty = \lim_{n\to\infty} f(y_n)$$

It follows that the two-sided limit $\lim_{x\to 0} f(x)$ does not exist.

2. Let $f(x) = \frac{1}{x^2}$ whenever $x \neq 0$ and additionally let $f(0) = 0$. Here the two-sided limit exists

$$\lim_{x\to 0} f(x) = \infty$$

However the value of the function at $x = 0$ does not equal this limit: clearly f is discontinuous at $x = 0$.

3. We revisit our motivating example. Let $f(x) = \frac{x^2-9}{x-3}$ have domain $U = \mathbb{R} \setminus \{3\}$. Whenever $x_n \neq 3$, we see that

$$f(x_n) = \frac{(x_n - 3)(x_n + 3)}{x_n - 3} = x_n + 3$$

By the limit laws, we conclude that $\lim f(x_n) = 3 + 3 = 6$ and so

$$\lim_{x\to 3} \frac{x^2 - 9}{x - 3} = 6$$

Since we referenced the limit laws for sequences so often in the examples, it is appropriate to update them to this new context. We do so without proof.

Corollary 1.26 (Limit Laws for functions). *Suppose $f, g : U \to \mathbb{R}$ satisfy $L = \lim_{x\to a} f(x)$ and $M = \lim_{x\to a} g(x)$ exist. Then,*

1. $\lim_{x\to a}(f+g)(x) = L + M$.
2. $\lim_{x\to a}(fg)(x) = LM$.
3. $\lim_{x\to a}\left(\frac{f}{g}\right)(x) = \frac{L}{M}$ *(requires $M \neq 0$).*
4. *If $L \in \mathbb{R}$ and h is continuous at L, then $\lim_{x\to a}(h\circ f)(x) = h(L)$.*
5. *(Squeeze Theorem) If $L = M$ and $f(x) \le h(x) \le g(x)$ for all $x \in U$, then $\lim_{x\to a} h(x) = L$.*

The corresponding results for one-sided limits also hold.

As with the original limit laws for sequences, parts 1–3 apply provided the limits are not *indeterminate forms* (e.g. $\infty - \infty$, $0\cdot\infty$, $\frac{0}{0}$, $\frac{\infty}{\infty}$). We'll see later how l'Hôpital's rule may be applied to such cases.

Examples 1.27. 1. Since $f(x) = \frac{x^2+5}{3x^2-2}$ is a rational function (continuous at all points of its domain), we quickly conclude that

$$\lim_{x\to 2}\frac{x^2+5}{3x^2-2} = f(2) = \frac{9}{10}$$

Alternatively, we may tediously invoke the other parts of the theorem:

$$\lim_{x\to 2}\frac{x^2+5}{3x^2-2} \overset{(3)}{=} \frac{\lim(x^2+5)}{\lim(3x^2-2)} \overset{(1)}{=} \frac{\lim x^2 + \lim 5}{\lim 3x^2 - \lim 2} \overset{(2)}{=} \frac{(\lim x)^2+5}{(\lim 3)(\lim x)^2 - 2}$$
$$= \frac{2^2+5}{3\cdot 2^2 - 2} = \frac{9}{10}$$

2. As $x \to \infty$, the simplistic approach results in a nonsense indeterminate form:

$$\lim_{x\to\infty}\frac{x^2+5}{3x^2-2} \overset{?}{=} \frac{\lim(x^2+5)}{\lim(3x^2-2)} \overset{?}{=} \frac{\infty}{\infty}$$

However, a little pre-theorem algebra quickly yields[4]

$$\lim_{x\to\infty}\frac{x^2+5}{3x^2-2} = \lim_{x\to\infty}\frac{1+5x^{-2}}{3-2x^{-2}} = \frac{\lim(1+5x^{-2})}{\lim(3-2x^{-2})} = \frac{1}{3}$$

[4]Be careful! The expressions $\frac{x^2+5}{3x^2-2}$ and $\frac{1+5x^{-2}}{3-2x^{-2}}$ do not describe the same function, yet their *limits* at ∞ are equal. The ease of equating these limits is one of the advantages of the '$\exists S$' formulation in Definition 1.24. Think about why; what is a suitable set S in this context?

Classification of Discontinuities

We now consider the ways in which a function can fail to be continuous.

Definition 1.28. Suppose that a function is continuous on an interval except at finitely many values: we call these *isolated discontinuities.*

Examples 1.29. 1. $f(x) = \frac{1}{x}$ has a discontinuity at $x = 0$ since it is continuous on the interval $\mathbb{R}$, except at one point $x = 0$. Note that a function need not be defined at a discontinuity!

2. $f(x) = \frac{1}{\sin \frac{1}{x}}$ has a *non-isolated discontinuity* at $x = 0$: on any interval containing zero, f has infinitely many discontinuities: $x = \frac{1}{\pi n}$ where $|n| \in \mathbb{N}$.

The next result helps us classify isolated discontinuities.

Theorem 1.30. *Let f be defined on a punctured neighborhood of $a \in \mathbb{R}$. Then*

$$\lim_{x\to a} f(x) = L \iff \lim_{x\to a^+} f(x) = L = \lim_{x\to a^-} f(x)$$

Proof. ($\Rightarrow$) Let $S = (c, a) \cup (a, b)$ satisfy the definition for $\lim_{x\to a} f(x) = L$. Since any sequence (say) in S^+ is also in S, plainly $S^+ = (a, b)$ and $S^- = (c, a)$ satisfy the one-sided definitions.

($\Leftarrow$) Suppose $S^- = (c, a)$ and $S^+ = (a, b)$ satisfy the one-sided definitions and denote $S = S^- \cup S^+$. Let $(x_n) \subseteq S$ be such that $x_n \to a$. Clearly (x_n) is the disjoint union of two subsequences $(x_n) \cap S^+$ and $(x_n) \cap S^-$, both of which[5] converge to a. There are three cases:

L finite: Let $\epsilon > 0$ be given. Because of the one-sided limits,

- $\exists N_1$ such that $n > N_1$ and $x_n > a \implies |f(x_n) - L| < \epsilon$
- $\exists N_2$ such that $n > N_2$ and $x_n < a \implies |f(x_n) - L| < \epsilon$

Now let $N = \max(N_1, N_2)$ in the definition of limit to see that $\lim f(x_n) = L$. Since this holds for all sequences $(x_n) \subseteq S$ converging to a, we conclude that $\lim_{x\to a} f(x) = L$.

$L = \pm\infty$**:** This is an exercise.

■

Example 1.31. Recalling elementary calculus, we show that the following is continuous at $x = 1$:

$$f(x) = \begin{cases} x^2 - 3 & \text{if } x \geq 1 \\ 3 - 5x & \text{if } x < 1 \end{cases}$$

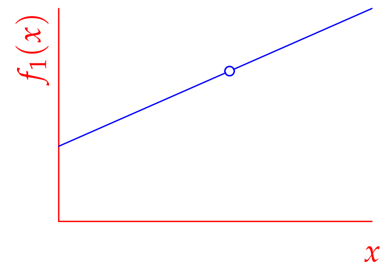
Step 1: Compute the left- and right-handed limits and check that these are equal:

$$\lim_{x\to 1^-} f(x) = \lim_{x\to 1^-} 3 - 5x = -2, \qquad \lim_{x\to 1^+} f(x) = \lim_{x\to 1^+} x^2 - 3 = -2$$

Step 2: Check that the value of the limits equals that of the function: $f(1) = 1^2 - 3 = -2$.

[5] It is possible for *one* of these subsequences to be finite; say if $x_n > a$ for all large n. This is of no concern; one of the ϵ-N conditions would be empty and thus vacuously true.

Recalling $(*)$ on page 13, we describe the different types of isolated discontinuity at some point a.

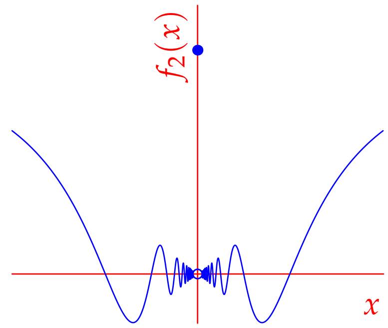


Removable discontinuity The two-sided limit $\lim_{x\to a} f(x) = L$ is finite, and either:

$$f(a) \neq L \text{ or } f(a) \text{ is undefined.}$$

The term comes from the fact that we can remove the discontinuity by changing the behavior of f only at $x = a$:

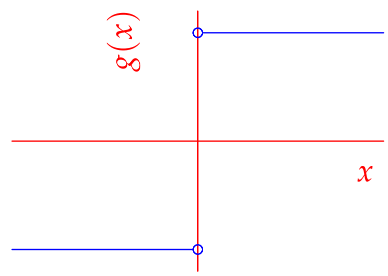
$$\tilde{f}(x) := \begin{cases} f(x) & \text{if } x \neq a \\ \lim_{x\to a} f(x) & \text{if } x = a \end{cases}$$



is now continuous at $x = a$. In the pictures,

$$f_1(x) = \frac{x^2 - 9}{x - 3} \quad \text{and} \quad f_2(x) = \begin{cases} x\sin(\frac{1}{x}) & \text{if } x \neq 0 \\ 1 & \text{if } x = 0 \end{cases}$$

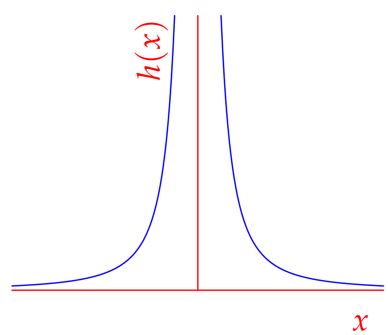
have removable discontinuities at $x = 3$ and 0 respectively.



Jump Discontinuity The one-sided limits are finite but *not equal.* A jump discontinuity cannot be removed by changing or inserting a value at $x = a$. The picture shows

$$g(x) = \frac{|x|}{x} = \begin{cases} 1 & \text{if } x > 0 \\ -1 & \text{if } x < 0 \end{cases}$$

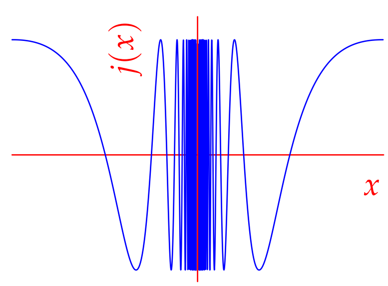
with a jump discontinuity at $x = 0$.



Infinite discontinuity The one-sided limits exist but at least one is infinite. We call the line $x = a$ a *vertical asymptote.* The picture shows

$$h(x) = \frac{1}{x^2}$$

with an infinite discontinuity $x = 0$. The fact that the one-sided limits of h are equal (and infinite) is irrelevant.



Essential discontinuity At least one of the one-sided limits does not exist. The picture shows $j(x) = \sin\frac{1}{x}$ for which neither of the limits $\lim_{x\to 0^\pm} j(x)$ exist.

It is also reasonable to refer to removable, infinite or essential discontinuities at interval endpoints.

Exercises 1.20. *Key concepts:* $\lim_{x\to a} f(x) = L$, ϵ, δ, M, N *versions,* *Limit Laws,* *Discontinuities*

1. Given $f(x) = \frac{x^3}{|x|}$, find $\lim_{x\to\infty} f(x)$, $\lim_{x\to-\infty} f(x)$, $\lim_{x\to 0^-} f(x)$, $\lim_{x\to 0^+} f(x)$ and $\lim_{x\to 0} f(x)$, if they exist.
2. Evaluate the following limits *using the methods of this section*

   (a) $\lim_{x\to a} \frac{\sqrt{x}-\sqrt{a}}{x-a}$ (b) $\lim_{x\to a} \frac{x^{-3/2}-a^{-3/2}}{x-a}$

   (c) $\lim_{x\to 0} \frac{\sqrt{1+3x^2}-1}{x^2}$ (d) $\lim_{x\to -\infty} \frac{\sqrt{4+3x^2}-2}{x}$

3. Suppose that the limits $L = \lim_{x\to a^+} f(x)$ and $M = \lim_{x\to a^+} g(x)$ exist.

   (a) Suppose $f(x) \le g(x)$ for all x in some interval (a, b). Prove that $L \le M$.

   (b) Do we have the same conclusion if we have $f(x) < g(x)$ on (a, b), or can we conclude that $L < M$? Prove your assertion, or give a counter-example.

4. Suppose that $\lim_{x\to\infty} f(x) = \lim_{x\to\infty} g(x) = \infty$. Using *only* this information, which of the following can you *always* evaluate? Prove your assertions in each case.

   (a) $\lim_{x\to\infty}(f+g)(x)$ (b) $\lim_{x\to\infty}(f-g)(x)$ (c) $\lim_{x\to\infty}(fg)(x)$ (d) $\lim_{x\to\infty}(f/g)(x)$

5. Complete the proof of Theorem 1.30 by considering the $L = \pm\infty$ cases.
6. Graph $f : \mathbb{R} \to \mathbb{R}$, find and identify the types of its discontinuities.

$$f(x) = \begin{cases} 0 & x = 0, \pm 1 \\ \frac{x}{|x|} & 0 < |x| < 1 \\ x^2 & |x| > 1 \end{cases}$$

7. Find the discontinuities and identify their types for the following function

$$f(x) = \begin{cases} \frac{1}{x}\sin\frac{1}{x} & \text{if } x < 0 \text{ or } x > 1 \\ \frac{1}{x} & \text{if } 0 < x \le 1 \end{cases}$$

8. Verify the claim following Definition 1.24: $\lim_{x\to a} f(x) = L$ if and only if

$$\forall \epsilon > 0, \exists \delta > 0 \text{ such that } 0 < |x-a| < \delta \implies |f(x) - L| < \epsilon$$

9. Recall Exercise 1.17.6, where we saw that a function $f : U \to \mathbb{R}$ is continuous at any isolated point $a \in U$.

   (a) Any function with domain $\operatorname{dom}(f) = \mathbb{Z}$ is continuous everywhere! Explain why we cannot define any limits $\lim_{x\to a^{(\pm)}} f(x)$ for such a function.

   (*Hint: Being unable to define a limit is different from saying* $\lim f(x) = DNE$*: see page 13.*)

   (b) Suppose $g(x) = x^2 h(x)$ has $\operatorname{dom}(g) = \{0\} \cup \{\frac{1}{n} : n \in \mathbb{Z}\}$, where h is any function taking values in the interval $[-1, 1]$. Explain why g is continuous at every point of its domain.

   (*These awkward examples of continuity can be avoided if we follow our usual approach where a domain is a union of intervals of positive length. This restriction is essentially baked in to the Definition 1.24.*)

2 Sequences and Series of Functions

If (f_n) is a sequence of functions, what should we mean by $\lim f_n$? This question is of huge relevance to the history of calculus: Issac Newton's work in the late 1600's made great use of *power series*, which are naturally constructed as limits of sequences of polynomials.

For instance, for each $n \in \mathbb{N}_0$, we might consider the polynomial function $f_n : \mathbb{R} \to \mathbb{R}$ defined by

$$f_n(x) = \sum_{k=0}^{n} x^k = 1 + x + \cdots + x^n$$

This is easily differentiated and integrated using the power law. What, however, are we to make of the *series*

$$f(x) := \sum_{n=0}^{\infty} x^n = 1 + x + x^2 + \cdots \ ?$$

Does this make sense as a function? What is its domain? Does it equal the limit of the sequence (f_n) in any meaningful way? Is it continuous, differentiable, integrable? If so, can we compute its derivative or integral term-by-term: for instance, is it legitimate to write

$$f'(x) = \sum_{n=1}^{\infty} nx^{n-1} = 1 + 2x + 3x^2 + \cdots \ ?$$

To many in Newton's time, such technical questions were less important than the application of calculus to the natural sciences. For the 18^{th} and 19^{th} century mathematicians who followed, however, the widespread application of calculus only increased the imperative to rigorously address these issues.

2.23 Power Series

First we review some of the important definitions, examples and results concerning infinite series.

Definition 2.1. Let $(b_n)_{n=m}^{\infty}$ be a sequence of real numbers. The *(infinite) series* $\sum b_n$ is the limit of the sequence (s_n) of *partial sums*,

$$s_n = \sum_{k=m}^{n} b_n = b_m + b_{m+1} + \cdots + b_n, \qquad \sum_{n=m}^{\infty} b_n = \lim_{n\to\infty} s_n$$

The series $\sum b_n$ is said to *converge*, *diverge to infinity* or *diverge by oscillation*[6] as does (s_n).

$\sum b_n$ is *absolutely convergent* if $\sum |b_n|$ converges. A convergent series that is not absolutely convergent is *conditionally convergent*.

[6]Recall that every sequence (s_n) has subsequences tending to each of

$$\limsup s_n = \lim_{N\to\infty} \sup\{x_n : n > N\} \quad \text{and} \quad \liminf s_n = \lim_{N\to\infty} \inf\{x_n : n > N\}$$

If (s_n) converges, or diverges to $\pm\infty$, then $\lim s_n = \limsup s_n = \liminf s_n$. The remaining case, divergence by oscillation, is when $\liminf s_n \neq \limsup s_n$: there exist (at least) two subsequences tending to different limits.

Examples 2.2. These examples form the standard reference dictionary for analysis of more complicated series. Make sure they are familiar![7]

1. (Geometric series) If r is constant, then $s_n = \sum_{k=0}^{n} r^k = \frac{1-r^{n+1}}{1-r}$. It follows that

$$\sum_{n=0}^{\infty} r^n \quad \begin{cases} \text{converges (absolutely) to } \frac{1}{1-r} & \text{if } -1 < r < 1 \\ \text{diverges to } \infty & \text{if } r \geq 1 \\ \text{diverges by oscillation} & \text{if } r \leq -1 \end{cases}$$

2. (Telescoping series) If $b_n = \frac{1}{n(n+1)}$, then $s_n = \sum_{k=1}^{n} b_n = 1 - \frac{1}{n+1} \implies \sum_{n=1}^{\infty} \frac{1}{n(n+1)} = 1$.

3. $\sum_{n=1}^{\infty} \frac{1}{n^2}$ is (absolutely) convergent. In fact $\sum_{n=1}^{\infty} \frac{1}{n^2} = \frac{\pi^2}{6}$, though checking this explicitly is tricky.

4. (Harmonic series) $\sum_{n=1}^{\infty} \frac{1}{n}$ is divergent to ∞.

5. (Alternating harmonic series) $\sum_{n=1}^{\infty} \frac{(-1)^n}{n}$ is conditionally convergent.

Theorem 2.3 (Root Test). *Given a series $\sum b_n$, let $\beta = \limsup |b_n|^{1/n}$,*

- *If $\beta < 1$ then the series converges absolutely.*
- *If $\beta > 1$ then the series diverges.*

---

[7] We give sketch proofs or refer to a standard 'test.' Review these if you are unfamiliar.

1. $s_n - rs_n = 1 + r + \cdots + r^n - (r + \cdots + r^n + r^{n+1}) = 1 - r^{n+1} \implies s_n = \frac{1-r^{n+1}}{1-r}$.

2. By partial fractions, $b_n = \frac{1}{n} - \frac{1}{n+1} \implies s_n = \left(1 - \frac{1}{2}\right) + \left(\frac{1}{2} - \frac{1}{3}\right) + \cdots + \left(\frac{1}{n} - \frac{1}{n+1}\right) = 1 - \frac{1}{n+1}$.

3. Use the comparison or integral tests. Alternatively: For each $n \geq 2$, we have $\frac{1}{n^2} < \frac{1}{n(n-1)}$. By part 2,

$$s_n = \sum_{k=1}^{n} \frac{1}{k^2} < 1 + \sum_{k=1}^{n} \frac{1}{k(k-1)} < 1 + \sum_{n=1}^{\infty} \frac{1}{n(n-1)} = 2$$

   Since (s_n) is monotone-up and bounded above by 2, we conclude that $\sum \frac{1}{n^2}$ is convergent.

4. Use the integral test. Alternatively, observe that

$$s_{2^{n+1}} - s_{2^n} = \sum_{k=2^n-1}^{2^{n+1}} \frac{1}{k} \geq \frac{2^n}{2^{n+1}} = \frac{1}{2} \implies s_{2^n} \geq \frac{n}{2} \xrightarrow[n\to\infty]{} \infty$$

   Since $s_n = \sum_{k=1}^{n} \frac{1}{k}$ is monotone-up, we conclude that $s_n \to \infty$.

5. Use the alternating series test, or explicitly check that both the even and odd partial sums (s_{2n}) and (s_{2n+1}) are convergent (monotone and bounded) to the same limit (essentially the proof of the alternating series test).

Root Test: $\beta < 1 \implies \exists \epsilon > 0$ such that $|b_n|^{1/n} \leq 1 - \epsilon$ (for large n) $\implies \sum |b_n|$ converges by comparison with $\sum (1-\epsilon)^n$.
$\beta > 1 \implies$ some subsequence of $(|b_n|^{1/n})$ converges to $\beta > 1 \implies b_n \not\to 0 \implies \sum b_n$ diverges (n^{th}-term test).

The root test is inconclusive if $\beta = 1$. Some simple inequalities[8] yield a test that is often easier to apply.

Corollary 2.4 (Ratio Test). *Given a series $\sum b_n$:*

- *If* $\limsup \left|\frac{b_{n+1}}{b_n}\right| < 1$ *then $\sum b_n$ converges absolutely.*
- *If* $\liminf \left|\frac{b_{n+1}}{b_n}\right| > 1$ *then $\sum b_n$ diverges.*

We are now ready to properly define and analyze our main objects of interest.

Definition 2.5. A *power series centered at* $c \in \mathbb{R}$ with *coefficients* $a_n \in \mathbb{R}$ is a formal expression

$$\sum_{n=m}^{\infty} a_n (x-c)^n$$

where $x \in \mathbb{R}$ is considered a variable. A power series is a *function* whose implied domain is the set of x for which the resulting infinite series converges.

It is common to refer simply to a *series*, and modify by infinite/power only when clarity requires. Almost always $m = 0$ or 1, and it is common for examples to be centered at $c = 0$.

Example 2.6. By the geometric series formula,

$$\sum_{n=0}^{\infty} \frac{(-1)^n}{2^n}(x-4)^n = \frac{1}{1 - \frac{-(x-4)}{2}} = \frac{2}{x-2} \quad \text{whenever} \quad \left|-\frac{x-4}{2}\right| < 1 \iff 2 < x < 6$$

The series is valid (converges) only on the subinterval $(2, 6)$ of the implied domain of the function $x \mapsto \frac{2}{x-2}$.
The behavior as $x \to 2^+$ is unsurprising, since evaluating the power series results in the divergent infinite series

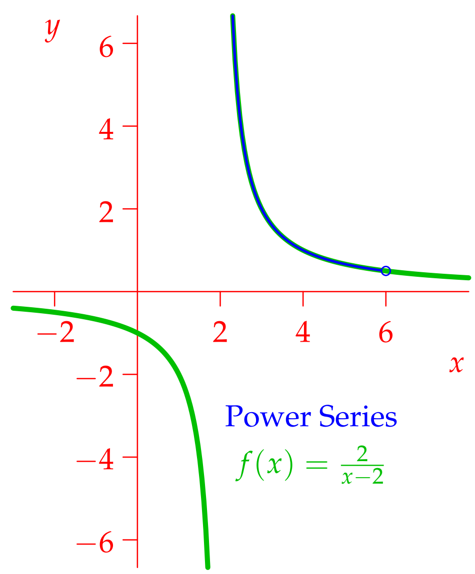
$$\sum 1 = +\infty$$

By contrast, as $x \to 6^-$ we see that limits and infinite series do not interact as we might expect,

$$\lim_{x \to 6^-} \sum_{n=0}^{\infty} \frac{(-1)^n}{2^n}(x-4)^n = \lim_{x \to 6^-} \frac{2}{x-2} = \frac{1}{2}$$

$$\sum_{n=0}^{\infty} \lim_{x \to 6^-} \frac{(-1)^n}{2^n}(x-4)^n = \sum (-1)^n = \text{DNE}$$

with the last series being divergent by oscillation.



The example shows that we cannot blindly take limits inside an infinite sum; understanding precisely when this is possible is one of our primary goals.

[8] You should have encountered these previously: $\liminf \left|\frac{b_{n+1}}{b_n}\right| \le \liminf |b_n|^{1/n} \le \limsup |b_n|^{1/n} \le \limsup \left|\frac{b_{n+1}}{b_n}\right|$

Radius and Interval of Convergence

The implied domain of the series in Example 2.6 turned out to be an *interval* $(2,6)$. Somewhat amazingly, the root test (Theorem 2.3) shows that the same is true for *every* power series!

Theorem 2.7 (Root Test for Power Series). *Given a power series $\sum a_n(x-c)^n$, define*[9]

$$R = \frac{1}{\limsup |a_n|^{1/n}}$$

The precisely one of the following statements holds:

$R \in (0,\infty)$	*the series converges absolutely when* $\lvert x-c\rvert < R$ *and diverges when* $\lvert x-c\rvert > R$
$R = \infty$	*the series converges absolutely for all* $x \in \mathbb{R}$
$R = 0$	*the series converges only at the center* $x = c$

Proof. For each fixed $x \in \mathbb{R}$, let $b_n = a_n(x-c)^n$ and apply the root test to $\sum b_n$, noting that

$$\limsup |b_n|^{1/n} = \begin{cases} \limsup |a_n|^{1/n} |x-c| = \frac{1}{R}|x-c| & \text{if } R \in (0,\infty) \\ 0 & \text{if } R = \infty \text{ or } x = c \\ \infty & \text{if } R = 0 \text{ and } x \neq c \end{cases}$$

In the first situation, $\limsup |b_n|^{1/n} < 1 \iff |x-c| < R$, etc. ∎

Definition 2.8. The *radius of convergence* is the value R defined in Theorem 2.7. The *interval of convergence* is the set of $x \in \mathbb{R}$ for which the series converges; its implied domain.

Radius of convergence	Interval of convergence
$R \neq 0, \infty$	$(c-R, c+R)$, $(c-R, c+R]$, $[c-R, c+R)$ or $[c-R, c+R]$
∞	$\mathbb{R} = (-\infty, \infty)$
0	$\{c\}$

In the first case, convergence/divergence at the endpoints of the interval of convergence must be tested for separately.

The ratio test (Corollary 2.4) provides a more user-friendly version.

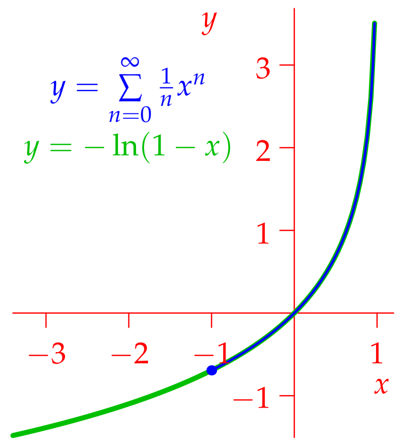
Corollary 2.9 (Ratio Test for Power Series). *If the limit exists,* $R = \lim_{n\to\infty} \left| \frac{a_n}{a_{n+1}} \right|$.

The ratio test is *weaker* than the root test: as Example 2.10.5 shows, there exist series for which the ratio test is be inconclusive.

[9] Since $|a_n| \geq 0$, we here adopt the conventions $\frac{1}{0} = \infty$, $\frac{1}{\infty} = 0$. With similar caveats, one can write $R = \liminf |a_n|^{-1/n}$. Since every sequence has a limit superior, this really is a *definition*. Whether one can easily *compute* R is another matter...

Examples 2.10. 1. The series $\sum_{n=1}^{\infty} \frac{1}{n} x^n$ is centered at 0. The ratio test tells us that

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



Test the endpoints of the interval of convergence separately:

$$x = 1 \qquad \sum \frac{1}{n} = \infty \text{ diverges}$$
$$x = -1 \quad \sum \frac{(-1)^n}{n} \text{ converges (conditionally)}$$

We conclude that the interval of convergence is $[-1, 1)$.

It can be seen (later) that the series converges to $-\ln(1-x)$ on its interval of convergence. As in Example 2.6, this function has a larger domain $(-\infty, -1)$, than that of the series.

2. The series $\sum_{n=1}^{\infty} \frac{1}{n^2} x^n$ similarly has

$$R = \lim_{n\to\infty} \left| \frac{a_n}{a_{n+1}} \right| = \lim_{n\to\infty} \frac{(n+1)^2}{n^2} = 1$$

Since $\sum \frac{1}{n^2}$ is absolutely convergent, we conclude that the power series also converges absolutely at $x = \pm 1$; the interval of convergence is $[-1, 1]$.

3. The series $\sum_{n=0}^{\infty} \frac{1}{n!} x^n$ converges absolutely for all $x \in \mathbb{R}$, since

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

You should recall from elementary calculus that this series converges to the natural exponential function $\exp(x) = e^x$ everywhere on $\mathbb{R}$; indeed this is one of the common *definitions* of the exponential function.

4. The series $\sum_{n=0}^{\infty} n! x^n$ has $R = \lim \frac{n!}{(n+1)!} = 0$. It therefore converges only at its center $x = 0$.

5. Let $a_n = \left(\frac{2}{3}\right)^n$ if n is even and $\left(\frac{3}{2}\right)^n$ if n is odd. If we try to apply the ratio test to the series $\sum_{n=0}^{\infty} a_n x^n$, we see that

$$\left| \frac{a_n}{a_{n+1}} \right| = \begin{cases} \left(\frac{2}{3}\right)^{2n+1} & \text{if } n \text{ even} \\ \left(\frac{3}{2}\right)^{2n+1} & \text{if } n \text{ odd} \end{cases} \implies \limsup \left| \frac{a_n}{a_{n+1}} \right| = \infty \neq 0 = \liminf \left| \frac{a_n}{a_{n+1}} \right|$$

The ratio test is therefore inconclusive. However, by the root test,

$$|a_n|^{1/n} = \begin{cases} \frac{2}{3} & \text{if } n \text{ even} \\ \frac{3}{2} & \text{if } n \text{ odd} \end{cases} \implies R = \frac{1}{\limsup |a_n|^{1/n}} = \frac{1}{3/2} = \frac{2}{3}$$

It is easy to check that the series diverges at $x = \pm\frac{2}{3}$; the interval of convergence is $(-\frac{2}{3}, \frac{2}{3})$.

With the help of the root test the domain of a power series is fully understood. Limits, continuity, differentiability and integrability are more delicate. We will return to these once we've developed some of the ideas around *convergence* for *sequences of functions.*

Exercises 2.23. *Key concepts: Power Series, Radius/interval of convergence* $R = \frac{1}{\limsup |a_n|^{1/n}}$

1. For each power series, find the radius and interval of convergence:

   (a) $\sum \frac{(-1)^n}{n^2 4^n} x^n$ (b) $\sum \frac{(n+1)^2}{n^3}(x-3)^n$ (c) $\sum \sqrt{n} x^n$

   (d) $\sum \frac{1}{n\sqrt{n}}(x+7)^n$ (e) $\sum (x-\pi)^{n!}$ (f) $\sum \frac{3^n}{\sqrt{n}} x^{2n+1}$

2. For each $n \in \mathbb{N}$ let $a_n = \left(\frac{4+2(-1)^n}{5}\right)^n$

   (a) Find $\limsup |a_n|^{1/n}$, $\liminf |a_n|^{1/n}$, $\limsup \left|\frac{a_{n+1}}{a_n}\right|$ and $\liminf \left|\frac{a_{n+1}}{a_n}\right|$.

   (b) Does the series $\sum a_n$ converge? What about $\sum (-1)^n a_n$? Why?

   (c) Find the interval of convergence of the power series $\sum a_n x^n$.

3. Suppose that $\sum a_n x^n$ has radius of convergence R. If $\limsup |a_n| > 0$, prove that $R \le 1$.

4. On the interval $(-\frac{2}{3}, \frac{2}{3})$, express the series in Example 2.10.5 as a simple function.

   (*Hints: Use geometric series formulæ and the fact that the value of an absolutely convergent series is independent of rearrangements*)

5. Consider the power series

$$\sum_{n=1}^{\infty} \frac{1}{3^n n}(x-7)^{5n+1} = \frac{1}{3}(x-7) + \frac{1}{18}(x-7)^6 + \frac{1}{81}(x-7)^{11} + \cdots$$

   Since only one in five of the terms are non-zero, it is a little tricky to analyze using a naïve application of our standard tests.

   (a) Explain why the ratio test for power series (Corollary 2.9) does not apply.

   (b) Writing the series as $\sum a_m (x-7)^m$, observe that

$$a_m = \begin{cases} \frac{5}{3^{\frac{m-1}{5}}(m-1)} & \text{if } m \equiv 1 \mod 5 \\ 0 & \text{otherwise} \end{cases}$$

   Use the root test (Theorem 2.7) and your understanding of elementary limits to directly compute the radius of convergence.

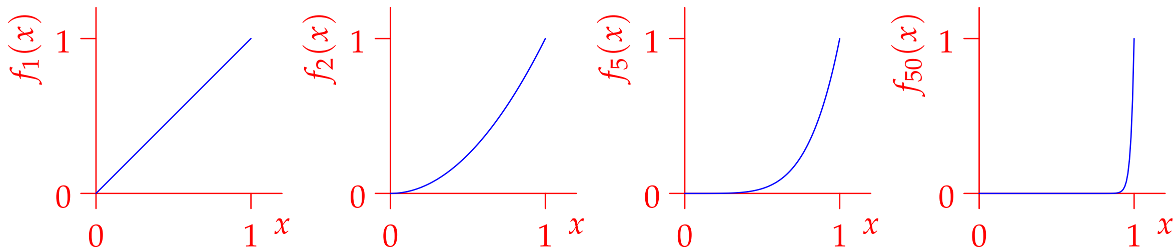
   (c) Alternatively, write $\sum \frac{1}{3^n n}(x-7)^{5n+1} = \sum b_n$. Apply the ratio test for *infinite* series (Corollary 2.4): what do you observe? Use your observation to compute the radius of convergence of the original series in a simpler manner than part (a).

   (d) Finally, check the endpoints to determine the interval of convergence.

2.24 Uniform Convergence

In this section we consider sequences (f_n) of *functions* $f_n : U \to \mathbb{R}$ and their limits.

Example 2.11. For each $n \in \mathbb{N}$, define $f_n : (0,1) \to \mathbb{R} : x \mapsto x^n$. Several examples are graphed.



There are several useful notions of convergence for sequences of functions. The simplest is where, for each input x, $(f_n(x))$ is treated as a distinct sequence of real numbers.

Definition 2.12. Suppose a function f and a sequence of functions (f_n) are given, all with domain U. We say that (f_n) *converges pointwise to f on U* if,

$$\forall x \in U, \lim_{n\to\infty} f_n(x) = f(x)$$

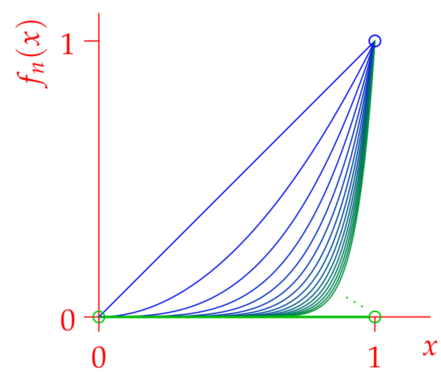
It is common to write '$f_n \to f$ pointwise.' For reference, here are two equivalent rephrasings:

1. $\forall x \in U, \lim_{n\to\infty} |f_n(x) - f(x)| = 0$;
2. $\forall x \in U, \forall \epsilon > 0, \exists N$ such that $n > N \implies |f_n(x) - f(x)| < \epsilon$.

As we'll see shortly, the relative position of the quantifiers ($\forall x$, $\exists N$) is crucial: in this definition, the value of N is permitted to depend on x as well as ϵ.

Example (2.11, mk. II). The sequence (f_n) converges pointwise on the domain $U = (0,1)$ to

$$f : (0,1) \to \mathbb{R} : x \mapsto 0$$



As a sanity check, we prove this explicitly. First observe that

$$|f_n(x) - f(x)| = x^n$$

Suppose $x \in (0,1)$, that $\epsilon > 0$ is given, and let $N = \frac{\ln \epsilon}{\ln x}$. Then

$$n > N \implies n \ln x < \ln \epsilon \implies x^n < \epsilon$$

where the inequality switches sign since $\ln x < 0$.

The example is nice in that a sequence of continuous functions converges pointwise to a continuous function. Unfortunately, this desirable situation is not universal...

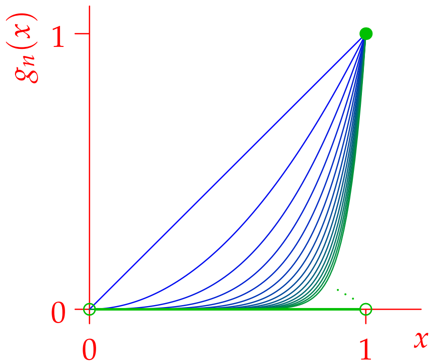
Example (2.11, mk. III). Define

$$g_n : (0,1] \to \mathbb{R} : x \mapsto x^n$$

Each g_n is a continuous function, however its pointwise limit

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

has a *jump discontinuity* at $x = 1$.

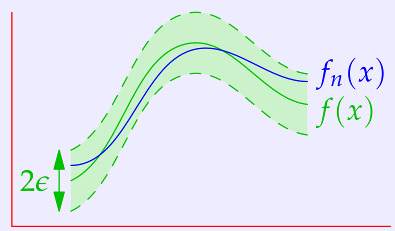


We'd like the limit of a sequence of continuous functions to itself be continuous. With this goal in mind, we make a tighter definition.

Definition 2.13. (f_n) *converges uniformly* to f on U if either

1. $\sup_{x \in U} |f_n(x) - f(x)| \xrightarrow[n\to\infty]{} 0$, or,
2. $\forall \epsilon > 0$, $\exists N$ such that $\forall x \in U, n > N \implies |f_n(x) - f(x)| < \epsilon$

A common notation is $f_n \rightrightarrows f$, though we won't use it.



As pictured, whenever $n > N$, the graph of $f_n(x)$ must lie between $f(x) \pm \epsilon$.
We'll show that statements 1 and 2 are equivalent momentarily. For the present, compare with the corresponding statements for pointwise convergence:

- As with *continuity* versus *uniform continuity*, the distinction comes in the *order of the quantifiers*: in uniform convergence, x is quantified *after* N and so *the same N works for all locations x.*
- Uniform convergence implies pointwise convergence.

Example (2.11, mk. IV). For the final time we revisit our main example. If $f_n(x) = x^n$ and $f(x) = 0$ are defined on $U = (0,1)$, then $f_n \to f$ *non-uniformly.* We show this using both criteria.

1. For *every* n,

$$\sup_{x\in(0,1)} |f_n(x) - f(x)| = \sup\{x^n : 0 < x < 1\} = 1 \not\longrightarrow 0$$

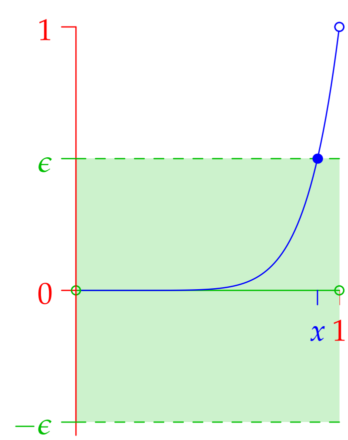
which plainly fails to converge to zero.

2. Suppose the convergence were uniform and let $\epsilon = \frac{1}{2}$. Then

$$\exists N \in \mathbb{N} \text{ such that } \forall x \in (0,1),\ n > N \implies x^n < \frac{1}{2}$$

Since $N \in \mathbb{N}$, a simple choice results in a contradiction;

$$x = \left(\frac{1}{2}\right)^{\frac{1}{N+1}} \in (0,1) \implies x^{N+1} = \frac{1}{2}$$



Theorem 2.14. *The criteria for uniform convergence in Definition 2.13 are equivalent.*

Proof. (1 ⇒ 2) This follows immediately from the fact that

$$\forall x \in U, \ |f_n(x) - f(x)| \le \sup_{x \in U} |f_n(x) - f(x)|$$

(2 ⇒ 1) Suppose $\epsilon > 0$ is given. Then

$$\exists N \in \mathbb{R} \text{ such that } \forall x \in U, \ n > N \implies |f_n(x) - f(x)| < \frac{\epsilon}{2}$$

But then

$$n > N \implies \sup_{x \in U} |f_n(x) - f(x)| \le \frac{\epsilon}{2} < \epsilon$$

■

Amazingly, this subtle change of definition is enough to preserve continuity.

Theorem 2.15. *Suppose (f_n) is a sequence of continuous functions. If $f_n \to f$ uniformly, then f is continuous.*

Proof. We demonstrate the continuity of f at $a \in U$. Let $\epsilon > 0$ be given.

- Since $f_n \to f$ uniformly,

$$\exists N \text{ such that } \forall x \in U, \ n > N \implies |f(x) - f_n(x)| < \frac{\epsilon}{3}$$

- Choose any $n > N$. Since f_n is continuous at a,

$$\exists \delta > 0 \text{ such that } |x - a| < \delta \implies |f_n(x) - f_n(a)| < \frac{\epsilon}{3} \qquad (\dagger)$$

Put these together with the triangle inequality to see that

$$\begin{aligned} |x - a| < \delta \implies |f(x) - f(a)| &\le |f(x) - f_n(x)| + |f_n(x) - f_n(a)| + |f_n(a) - f(a)| \\ &< \frac{\epsilon}{3} + \frac{\epsilon}{3} + \frac{\epsilon}{3} = \epsilon \end{aligned}$$

■

We need not have fixed a at the start of the proof. Rewriting $(\dagger)$ to become

$$\exists \delta > 0 \text{ such that } \forall x, a \in U, \ |x - a| < \delta \implies |f_n(x) - f_n(a)| < \frac{\epsilon}{3}$$

proves a related result.

Corollary 2.16. *Suppose (f_n) is a sequence of **uniformly** continuous functions. If $f_n \to f$ uniformly, then f is also **uniformly** continuous.*

Examples 2.17. 1. Let $f_n(x) = x + \frac{1}{n}x^2$. This is continuous on $\mathbb{R}$ for all x, and converges pointwise to the continuous function $f : x \mapsto x$.

(a) On any bounded interval $[-M, M]$ the convergence $f_n \to f$ is uniform,

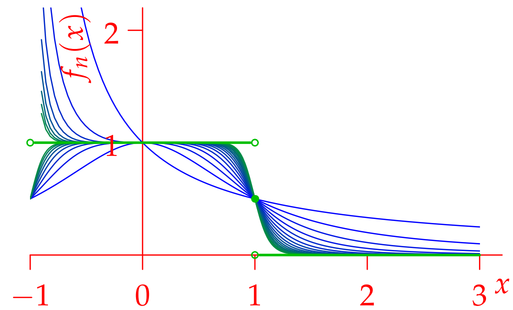
$$\sup_{x\in[-M,M]} |f_n(x) - f(x)| = \sup\left\{\frac{1}{n}x^2 : x \in [-M, M]\right\} = \frac{M^2}{n} \xrightarrow[n\to\infty]{} 0$$

(b) On any unbounded interval, $\mathbb{R}$ say, the convergence is non-uniform,

$$\sup_{x\in\mathbb{R}} |f_n(x) - f(x)| = \sup\left\{\frac{1}{n}x^2 : x \in \mathbb{R}\right\} = \infty$$

2. Consider $f_n(x) = \frac{1}{1+x^n}$; this is continuous on $(-1, \infty)$ and converges pointwise to

$$f(x) = \begin{cases} 0 & \text{if } x > 1 \\ \frac{1}{2} & \text{if } x = 1 \\ 1 & \text{if } -1 < x < 1 \end{cases}$$



We consider the convergence $f_n \to f$ on several intervals.

(a) On $[2, \infty)$, the pointwise limit is continuous. Moreover, $f_n(x)$ is decreasing, whence

$$\sup_{x\in[2,\infty)} |f_n(x) - 0| = \frac{1}{1+2^n} \xrightarrow[n\to\infty]{} 0$$

and the convergence is uniform. Alternatively; if $\epsilon \in (0, 1)$, let $N = \log_2(\epsilon^{-1} - 1)$, then

$$\forall x \geq 2,\ n > N \implies |f_n(x) - 0| = \frac{1}{1+x^n} \leq \frac{1}{1+2^n} < \frac{1}{1+2^N} = \epsilon$$

The same argument shows that $f_n \to f$ uniformly on any interval $[a, \infty)$ where $a > 1$.

(b) On $[1, \infty)$ the convergence is not uniform, since the pointwise limit is discontinuous,

$$f(x) = \begin{cases} 0 & \text{if } x > 1 \\ \frac{1}{2} & \text{if } x = 1 \end{cases}$$

(c) The convergence is not even uniform on the open interval $(1, \infty)$,

$$\sup_{x\in[1,\infty)} |f_n(x) - f(x)| = \sup\left\{\frac{1}{1+x^n} : x > 1\right\} = \frac{1}{2} \not\xrightarrow[n\to\infty]{} 0$$

(d) Similarly, for any $a \in (0, 1)$, the convergence $f_n \to f$ is uniform on $[0, a]$, this time to the (continuous) constant function $f(x) = 1$,

$$\sup_{x\in[0,a]} |f_n(x) - 1| = \left|1 - \frac{1}{1+a^n}\right| = \frac{a^n}{1+a^n} \xrightarrow[n\to\infty]{} 0$$

(e) Finally, on $(-1, 1)$ the convergence is not uniform,

$$\sup_{x\in[0,1)} |f_n(x) - f(x)| = \sup\left\{\frac{x^n}{1+x^n} : x \in [0, 1)\right\} = \frac{1}{2} \not\xrightarrow[n\to\infty]{} 0$$

Exercises 2.24. *Key concepts: Pointwise & Uniform Convergence, Uniform conv preserves continuity*

1. For each sequence of functions defined on $[0,\infty)$:

   (i) Find the pointwise limit $f(x)$ as $n\to\infty$.

   (ii) Determine whether $f_n\to f$ uniformly on $[0,1]$.

   (iii) Determine whether $f_n\to f$ uniformly on $[1,\infty)$.

   (a) $f_n(x)=\dfrac{x}{n}$ (b) $f_n(x)=\dfrac{x^n}{1+x^n}$ (c) $f_n(x)=\dfrac{x^n}{n+x^n}$

   (d) $f_n(x)=\dfrac{x}{1+nx^2}$ (e) $f_n(x)=\dfrac{nx}{1+nx^2}$

2. Let $f_n(x)=\left(x-\frac{1}{n}\right)^2$. If $f(x)=x^2$, we clearly have $f_n\to f$ pointwise on any domain.

   (a) Prove that the convergence is uniform on $[-1,1]$.

   (b) Prove that the convergence is non-uniform on $\mathbb{R}$.

3. For each sequence, find the pointwise limit and decide if the convergence is uniform.

   (a) $f_n(x)=\frac{1+2\cos^2(nx)}{\sqrt{n}}$ for $x\in\mathbb{R}$.

   (b) $f_n(x)=\cos^n(x)$ on $[-\pi/2,\pi/2]$.

4. For each $n\in\mathbb{N}$, consider the continuous function

   $$f_n:[0,1]\to\mathbb{R}:x\mapsto nx^n(1-x)$$

   (a) Given $0\le x<1$, let $a\in(x,1)$. Explain why $\exists N$ such that

   $$n>N\implies |f_{n+1}(x)|\le a\,|f_n(x)|$$

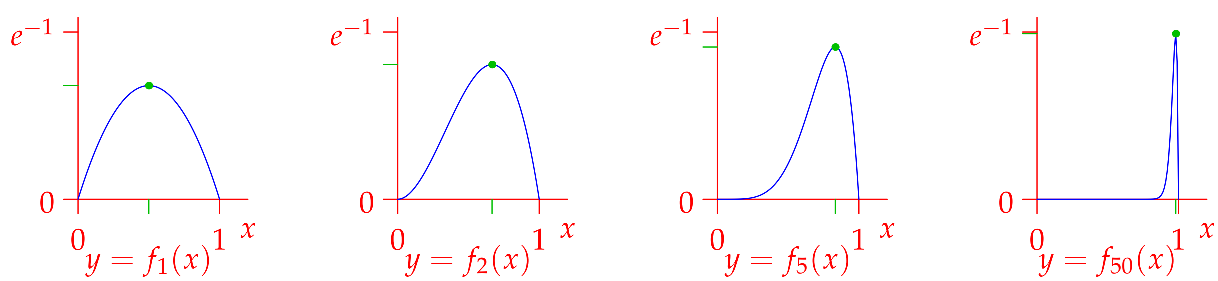
   Hence conclude that the pointwise limit of (f_n) is the zero function.

   (b) Use elementary calculus ($f_n'(x)=0\iff\dots$) to prove that the maximum value of f_n is located at $x_n=\frac{n}{1+n}$. Hence compute

   $$\sup_{x\in[0,1]}|f_n(x)-f(x)|$$

   and use it to show that the convergence $f_n\to 0$ is non-uniform.

   This shows that the converse to Theorem 2.15 is false, even on a bounded interval: the continuous sequence (f_n) converges non-uniformly *to a continuous function. Sketches of several f_n are below.*



5. Explain where the proof of Theorem 2.15 fails if $f_n\to f$ *non-uniformly.*

2.25 More on Uniform Convergence

While we haven't yet developed calculus, our familiarity with basic differentiation and integration makes it natural to pause to consider the interaction of these concepts with sequences of functions.

We also consider a Cauchy-criterion for uniform convergence, which leads to the useful Weierstraß M-test.

Example 2.18. Recall that $f_n(x) = x^n$ converges uniformly to $f(x) = 0$ on any interval $[0, a]$ where $a < 1$. We easily check that

$$\int_0^a f_n(x)\,\mathrm{d}x = \frac{1}{n+1}a^{n+1} \xrightarrow[n\to\infty]{} 0 = \int_0^a f(x)\,\mathrm{d}x$$

In fact the sequence of derivatives converge here also

$$\frac{\mathrm{d}}{\mathrm{d}x} f_n(x) = nx^{n-1} \xrightarrow[n\to\infty]{} 0 = f'(x)$$

It is perhaps surprising that integration interacts more nicely with uniform limits than does differentiation. We therefore consider integration first.

Theorem 2.19. *Let $f_n \to f$ uniformly on $[a, b]$ where the functions f_n are integrable. Then f is integrable on $[a, b]$ and*

$$\lim_{n\to\infty} \int_a^b f_n(x)\mathrm{d}x = \int_a^b f(x)\mathrm{d}x$$

Proof. Given $\epsilon > 0$, note that $\int_a^b \frac{\epsilon}{2(b-a)}\,\mathrm{d}x = \frac{\epsilon}{2}$. Since $f_n \to f$ uniformly, $\exists N$ such that[10]

$$\begin{aligned}
\forall x \in [a, b],\ n > N &\implies |f_n(x) - f(x)| < \frac{\epsilon}{2(b-a)} \\
&\implies f_n(x) - \frac{\epsilon}{2(b-a)} < f(x) < f_n(x) + \frac{\epsilon}{2(b-a)} \\
&\implies \int_a^b f_n(x)\,\mathrm{d}x - \frac{\epsilon}{2} \le \int_a^b f(x)\,\mathrm{d}x \le \int_a^b f_n(x)\,\mathrm{d}x + \frac{\epsilon}{2} \\
&\implies \left| \int_a^b f_n(x)\,\mathrm{d}x - \int_a^b f(x)\,\mathrm{d}x \right| \le \frac{\epsilon}{2} < \epsilon
\end{aligned}$$

■

The appearance of uniform convergence in the proof is subtle. If $N = N(\epsilon)$ were allowed to depend on x, then the integral $\int_a^b f_n(x)\,\mathrm{d}x$ would be meaningless: Which n would we consider? Larger than $N(x, \epsilon)$ for *which* x? Taking n 'larger' than *all* the $N(x, \epsilon)$ might produce the absurdity $n = \infty$!

[10] This assumes f is already integrable. Once we've properly defined (Riemann/Darboux) integrability at the end of the course, we can insert the following

$$\int_a^b f_n(x)\,\mathrm{d}x - \frac{\epsilon}{2} \le L(f) \le U(f) \le \int_a^b f_n(x)\,\mathrm{d}x + \frac{\epsilon}{2} \implies 0 \le U(f) - L(f) \le \epsilon \implies U(f) = L(f)$$

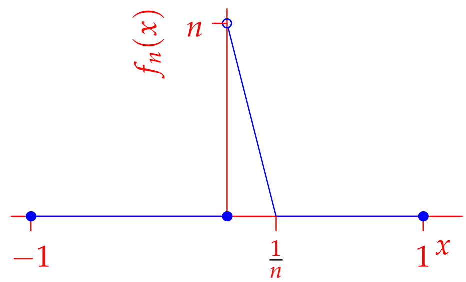
where $U(f)$ and $L(f)$ are the upper and lower Darboux integrals of f; equality shows that f is integrable on $[a, b]$.

Examples 2.20. 1. Uniform convergence is not required for the integrals to converge as we'd like. For instance, recall that extending the previous example to the domain $[0,1]$ results in non-uniform convergence; however, we still have

$$\int_0^1 f_n(x)\,\mathrm{d}x = \frac{1}{n+1} \xrightarrow[n\to\infty]{} 0 = \int_0^1 f(x)\,\mathrm{d}x$$

2. To obtain a sequence of functions $f_n \to f$ for which $\int f_n \not\to \int f$ requires a bit of creativity. Consider the sequence

$$f_n : [-1,1] \to \mathbb{R} : x \mapsto \begin{cases} n - n^2 x & \text{if } 0 < x < \frac{1}{n} \\ 0 & \text{otherwise} \end{cases}$$



If $0 < x < 1$, then for large $n \in \mathbb{N}$ we have

$$x \geq \frac{1}{n} \implies f_n(x) = 0$$

We conclude that $f_n \to 0$ pointwise. Since the area under f_n is a triangle with base $\frac{1}{n}$ and height n, the integral is constant and *non-zero*;

$$\int_{-1}^1 f_n(x)\,\mathrm{d}x = \frac{1}{2} \neq 0 = \int_{-1}^1 f(x)\,\mathrm{d}x$$

It should be obvious that the convergence $f_n \to 0$ is non-uniform; why?

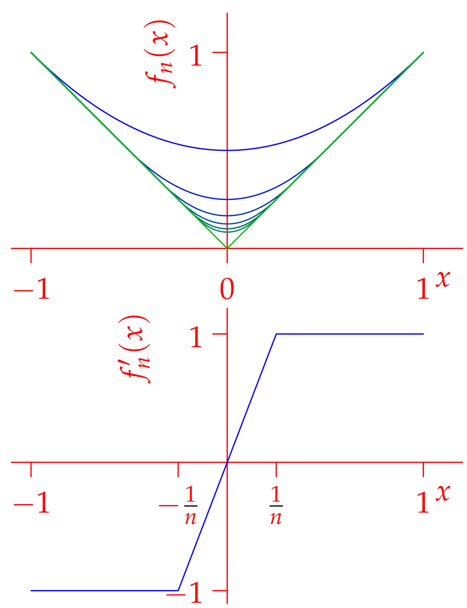
Derivatives and Uniform Limits We've already seen that a uniform limit of differentiable functions *might* be differentiable (Example 2.18). As the next example shows, this should not be expected in general, since even uniform limits of differentiable functions can have corners.

Example 2.21. For each $n \in \mathbb{N}$, consider the function

$$f_n : \mathbb{R} \to \mathbb{R} : x \mapsto \begin{cases} |x| & \text{if } |x| \geq \frac{1}{n} \\ \frac{n}{2}x^2 + \frac{1}{2n} & \text{if } |x| < \frac{1}{n} \end{cases}$$

- Each f_n is differentiable: $f_n'(x) = \begin{cases} 1 & \text{if } x \geq \frac{1}{n} \\ nx & \text{if } |x| < \frac{1}{n} \\ -1 & \text{if } x \leq -\frac{1}{n} \end{cases}$
- f_n converges pointwise to $f(x) = |x|$, which is *non-differentiable* at $x = 0$.
- $f_n \to f$ uniformly since

$$\sup_{x\in[-1,1]} |f_n(x) - f(x)| = f_n(0) = \frac{1}{2n} \to 0$$



If our goal is to transfer differentiability to the limit of a sequence of functions, then we have some work to do.

Theorem 2.22. *Suppose (f_n) is a sequence and f, g functions, all with domain $[a, b]$. Suppose also:*

- *$f_n \to f$ pointwise;*
- *Each f_n is differentiable with continuous derivative;*[11]
- *$f_n' \to g$ uniformly.*

Then $f_n \to f$ uniformly on $[a, b]$ and f is differentiable with derivative g.

The issue in the previous example is that the *pointwise limit* of the derived sequence (f_n') is discontinuous at $x = 0$ and therefore $f_n' \to g$ isn't uniform!

Proof. For any $x \in [a, b]$, the fundamental theorem of calculus (part II) tells us that

$$\int_a^x f_n'(t)\,\mathrm{d}t = f_n(x) - f_n(a)$$

As $n \to \infty$, Theorem 2.19 says the left side converges to $\int_a^x g(t)\,\mathrm{d}t$ and the right to $f(x) - f(a)$ (both pointwise). Since $f_n' \to g$ uniformly, we see that g is continuous and can apply the fundamental theorem (part I): $\int_a^x g(t)\,\mathrm{d}t = f(x) - f(a)$ is differentiable with derivative g.

The uniformity of the convergence $f_n \to f$ follows from Exercise 10. ∎

Uniformly Cauchy Sequences and the Weierstraß M-Test

Recall that one may use Cauchy sequences to demonstrate convergence *without knowing the limit in advance.* An analogous discussion is available for sequences of functions.

Definition 2.23. A sequence of functions (f_n) is *uniformly Cauchy* on U if

$$\forall \epsilon > 0,\ \exists N \in \mathbb{N} \text{ such that } \forall x \in U,\ m, n > N \implies |f_n(x) - f_m(x)| < \epsilon$$

Example 2.24. Let $f_n(x) = \sum\limits_{k=1}^{n} \frac{1}{k^2} \sin k^2 x$ be defined on $\mathbb{R}$. Given $\epsilon > 0$, let $N = \frac{1}{\epsilon}$, then

$$\begin{aligned} m > n > N \implies |f_m(x) - f_n(x)| &= \left| \sum_{k=n+1}^{m} \frac{1}{k^2} \sin k^2 x \right| \le \sum_{k=n+1}^{m} \frac{1}{k^2} \le \sum_{k=n+1}^{m} \frac{1}{k(k-1)} \\ &= \sum_{k=n+1}^{m} \frac{1}{k-1} - \frac{1}{k} = \frac{1}{n} - \frac{1}{m} < \frac{1}{N} = \epsilon \end{aligned}$$

whence (f_n) is uniformly Cauchy.

[11] Without this continuity assumption, the fundamental theorem of calculus doesn't apply and the proof requires an alternative approach. One can also weaken the hypotheses: if $f_n' \to g$ uniformly and $(f_n(x))$ converges for *at least one* $x \in [a, b]$, then there exists f such that $f_n \to f$ is uniform and $f' = g$.

As with sequences of real numbers, uniformly Cauchy sequences converge; in fact uniformly!

Theorem 2.25. *A sequence (f_n) is uniformly Cauchy on U if and only if it converges uniformly to some $f : U \to \mathbb{R}$.*

Proof. ($\Rightarrow$) Let (f_n) be uniformly Cauchy on U. For each $x \in U$, the sequence $(f_n(x)) \subseteq \mathbb{R}$ is Cauchy and thus convergent. Define $f : U \to \mathbb{R}$ to be the pointwise limit:

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

We claim that $f_n \to f$ uniformly. Let $\epsilon > 0$ be given, then $\exists N \in \mathbb{N}$ such that

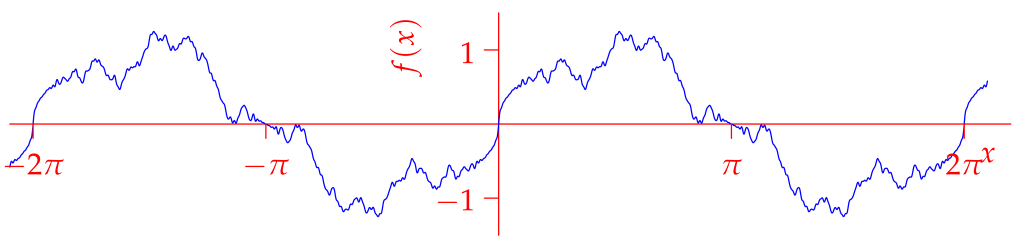
$$\begin{aligned} m > n > N &\implies |f_n(x) - f_m(x)| < \frac{\epsilon}{2} \\ &\implies f_n(x) - \frac{\epsilon}{2} < f_m(x) < f_n(x) + \frac{\epsilon}{2} \\ &\implies f_n(x) - \frac{\epsilon}{2} \le f(x) \le f_n(x) + \frac{\epsilon}{2} && \text{(take limits as } m \to \infty) \\ &\implies |f_n(x) - f(x)| \le \frac{\epsilon}{2} < \epsilon \end{aligned}$$

($\Leftarrow$) This is Exercise 2. ∎

Example (2.24, mk. II). Since (f_n) is uniformly Cauchy on $\mathbb{R}$, it converges uniformly to some $f : \mathbb{R} \to \mathbb{R}$. It seems reasonable to write

$$f(x) = \sum_{n=1}^{\infty} \frac{1}{n^2} \sin n^2 x$$

The graph of this function looks somewhat bizarre:



Since each f_n is (uniformly) continuous, Theorem 2.15 says that f is also (uniformly) continuous. By Theorem 2.19, $f(x)$ is integrable, indeed

$$\int_a^b f(x)\,dx = \lim_{n\to\infty} \sum_{k=1}^{n} -k^{-4} \cos k^2 x \Big|_a^b = \sum_{n=1}^{\infty} \frac{1}{n^4}(\cos n^2 a - \cos n^2 b)$$

which converges (comparison test) for all a, b. By contrast, the derived sequence

$$f_n'(x) = \sum_{k=1}^{n} \cos k^2 x$$

does not converge for *any* x since $\lim_{n\to\infty} \cos n^2 x \neq 0$. We should therefore expect (though we offer no proof) that f is nowhere differentiable.

The example generalizes. Suppose (g_k) is a sequence of functions on U and define the *series* $\sum g_k(x)$ as the pointwise limit of the sequence (f_n) of partial sums

$$\sum_{k=k_0}^{\infty} g_k(x) := \lim_{n\to\infty} f_n(x) \quad \text{where} \quad f_n(x) = \sum_{k=k_0}^{n} g_k(x)$$

whenever the limit exists. The series is said to converge uniformly whenever (f_n) does so. Theorems 2.15, 2.19 and 2.22 immediately translate.

Corollary 2.26. *Let $\sum g_k$ be a series of functions converging uniformly on U. Then:*

1. *If each g_k is (uniformly) continuous then $\sum g_k$ is (uniformly) continuous.*
2. *If each g_k is integrable, then $\int \sum g_k(x)\,\mathrm{d}x = \sum \int g_k(x)\,\mathrm{d}x$.*
3. *If each g_k is continuously differentiable, and the sequence of derived partial sums f_n' converges uniformly, then $\sum g_k$ is differentiable and $\frac{\mathrm{d}}{\mathrm{d}x}\sum g_k(x) = \sum g_k'(x)$.*

As an application of the uniform Cauchy criterion, we obtain an easy test for uniform convergence.

Theorem 2.27 (Weierstraß M-test). *Suppose (g_k) is a sequence of functions on U. Moreover assume:*

1. *(M_k) is a non-negative sequence such that $\sum M_k$ converges.*
2. *Each g_k is bounded by M_k; that is $|g_k(x)| \le M_k$.*

Then $\sum g_k(x)$ converges uniformly on U.

Proof. Let $f_n(x) = \sum_{k=k_0}^{n} g_k(x)$ define the sequence of partial sums. Since $\sum M_k$ converges, its sequence of partial sums is Cauchy (the *Cauchy criterion* for infinite series); given $\epsilon > 0$,

$$\exists N \text{ such that } m > n > N \implies \sum_{k=n+1}^{m} M_k < \epsilon$$

However, by assumption,

$$m > n > N \implies |f_m(x) - f_n(x)| = \left|\sum_{k=n+1}^{m} g_k(x)\right| \le \sum_{k=n+1}^{m} |g_k(x)| \le \sum_{k=n+1}^{m} M_k < \epsilon$$

The sequence of partial sums is uniformly Cauchy and thus uniformly convergent. ∎

Example 2.28. Given the series $\sum_{n=1}^{\infty} \frac{1+\cos^2(nx)}{n^2}\sin(nx)$, we clearly have

$$\left|\frac{1+\cos^2(nx)}{n^2}\sin(nx)\right| \le \frac{2}{n^2} \text{ for all } x \in \mathbb{R}$$

Since $\sum \frac{2}{n^2}$ converges, the M-test shows that the original series converges uniformly on $\mathbb{R}$.

Exercises 2.25. *Key concepts: Uniform convergence preseves integration, Uniform Cauchyness, M-test*

1. For each $n \in \mathbb{N}$, let $f_n(x) = nx^n$ when $x \in [0,1)$ and $f_n(1) = 0$.
   (a) Prove that $f_n \to 0$ pointwise on $[0,1]$.
   (*Hint: recall Exercise 2.24.4 if you're not sure how to prove this*)
   (b) By considering the integrals $\int_0^1 f_n(x)\,\mathrm{d}x$ show that $f_n \to 0$ is not uniform.
2. Prove that if $f_n \to f$ uniformly, then the sequence (f_n) is uniformly Cauchy.
3. (a) Suppose (f_n) is a sequence of bounded functions on U and suppose that $f_n \to f$ converges uniformly on U. Prove that f is bounded on U.
   (b) Give an example of a sequence of bounded functions (f_n) converging pointwise to f on $[0,\infty)$, but for which f is *unbounded.*
4. The sequence defined by $f_n(x) = \frac{nx}{1+nx^2}$ (Exercise 2.24.1) converges uniformly on any closed interval $[a,b]$ where $0 < a < b$.
   (a) Check explicitly that $\int_a^b f_n(x)\,\mathrm{d}x \to \int_a^b f(x)\,\mathrm{d}x$, where $f = \lim f_n$.
   (b) Is the same thing true for derivatives?
5. Let $f_n(x) = n^{-1}\sin n^2x$ be defined on $\mathbb{R}$.
   (a) Prove that f_n converges uniformly on $\mathbb{R}$.
   (b) Check that $\int_0^x f_n(t)\,\mathrm{d}t$ converges for any $x \in \mathbb{R}$.
   (c) Does the derived sequence (f_n') converge? Explain.
6. Use the M-test to prove that $\sum\limits_{n=1}^{\infty} \frac{x^n}{n^2}$ defines a continuous function on $[-1,1]$.
7. Prove that $\sum\limits_{n=1}^{\infty} \frac{x^n \sin x}{(n+1)^3 2^n}$ converges uniformly to a continuous function on the interval $[-2,2]$.
8. Prove that if $\sum g_k$ converges uniformly on a set U and if h is a bounded function on U, then $\sum hg_k$ converges uniformly on U.
   (*Warning: you cannot simply write* $\sum hg_k = h\sum g_k$)
9. Consider Example 2.20.2.
   (a) Check explicitly that the convergence isn't uniform by computing $\sup\limits_{x\in[-1,1]} |f_n(x) - f(x)|$
   (b) Prove that $f_n \to 0$ pointwise on $(0,1]$ using the ϵ–N definition of convergence: that is, given $\epsilon > 0$ and $x \in (0,1]$, find an explicit $N(x,\epsilon)$ such that
   $$n > N \implies |f(x)| < \epsilon$$
   What happens to your choice of $N(x,\epsilon)$ as $x \to 0^+$?
10. Suppose (f_n') converges uniformly on $[a,b]$ and that each f_n' is continuous.
    (a) Use the fact that (f_n') is uniformly Cauchy to prove that (f_n) is uniformly Cauchy and thus converges uniformly to some function f.
    (*Hint:* $|f_n(x) - f_m(x)| = \left|\int_a^x f_n'(t) - f_m'(t)\,\mathrm{d}t\right| \ldots$)
    (b) Explain why we need not have assumed the existence of f in Theorem 2.22.

2.26 Differentiation and Integration of Power Series

We now specialize our recent results to power series. While everything will be stated for series centered at $x = 0$, all are easily translated to arbitrary centers.

Theorem 2.29. *Let $\sum a_n x^n$ be a power series with radius of convergence $R > 0$ and let $T \in (0, R)$. Then:*

1. *The series converges uniformly on $[-T, T]$.*
2. *The series is uniformly continuous on $[-T, T]$ and continuous on $(-R, R)$.*

Proof. This is a straightforward application of the Weierstraß M-test (Theorem 2.27). For each k, define $M_k = |a_k| \, T^k$, and observe that

$$T < R \implies \sum a_n T^n \text{ converges absolutely } \implies \sum M_k \text{ converges}$$

By the M-test and Corollary 2.26, the power series converges uniformly on $[-T, T]$ to a uniformly continuous function.

Finally, every $x \in (-R, R)$ lies in some such interval (take $T = |x|$), whence the power series is continuous on $(-R, R)$. ∎

Example 2.30. On its interval of convergence $(-1, 1)$, the geometric series $\sum\limits_{n=0}^{\infty} x^n$ converges pointwise to $\frac{1}{1-x}$; convergence is uniform on any interval $[-T, T] \subseteq (-1, 1)$.
We needn't use the Theorem for this is simple to verify directly: writing f, f_n for the series and its partial sums,

$$|f_n(x) - f(x)| = \left| \frac{1 - x^{n+1}}{1 - x} - \frac{1}{1 - x} \right| = \left| \frac{x^{n+1}}{1 - x} \right|$$
$$\implies \sup_{x \in [-T, T]} |f_n(x) - f(x)| = \frac{T^{n+1}}{1 - T} \xrightarrow[n \to \infty]{} 0$$

By contrast, the convergence is non-uniform on $(-1, 1)$, since

$$\sup_{x \in (-1,1)} |f_n(x) - f(x)| = \infty$$

Theorem 2.31. *Suppose a power series $\sum a_n x^n$ has radius of convergence $R > 0$. Then the series is integrable and differentiable term-by-term on the interval $(-R, R)$. Indeed for any $x \in (-R, R)$,*

$$\frac{\mathrm{d}}{\mathrm{d}x} \sum_{n=0}^{\infty} a_n x^n = \sum_{n=1}^{\infty} n a_n x^{n-1} \quad \text{and} \quad \int_0^x \sum_{n=0}^{\infty} a_n t^n \, \mathrm{d}t = \sum_{n=0}^{\infty} \frac{a_n}{n+1} x^{n+1}$$

where both series also have radius of convergence R.

Proof. Let $f(x) = \sum a_n x^n$ have radius of convergence R, and observe that

$$\limsup |na_n|^{1/n} = \lim n^{1/n} \limsup |a_n|^{1/n} = \frac{1}{R}$$

whence $\sum na_n x^n$ also has radius of convergence R. At any given non-zero $x \in (-R, R)$, we may write

$$\sum_{n=1}^{\infty} na_n x^{n-1} = x^{-1} \sum_{n=1}^{\infty} na_n x^n$$

to see that the derived series also has radius of convergence R. On any interval $[-T, T] \subseteq (-R, R)$, the derived series converges uniformly (Theorem 2.29). Since each $a_n x^n$ is continuously differentiable, Corollary 2.26 says that f is differentiable on $[-T, T]$ and that

$$f'(x) = \sum_{n=0}^{\infty} \frac{\mathrm{d}}{\mathrm{d}x} a_n x^n = \sum_{n=1}^{\infty} na_n x^{n-1}$$

Since any $x \in (-R, R)$ lies in some such interval $[-T, T]$, we are done.

Exercise 7 discusses the corresponding result for integration. ■

We postpone the canonical examples until after the next result.

Continuity at Endpoints?

There is one small hole in our analysis. A series $\sum a_n x^n$ with radius of convergence R converges and is continuous on $(-R, R)$. But what if it also converges at $x = \pm R$? Is the series continuous at the endpoints? The answer is yes, though demonstrating this small benefit requires a lot of work!

Theorem 2.32 (Abel's Theorem). *Power series are continuous on their full interval of convergence.*

Examples 2.33. 1. Apply our results to the geometric series;

$$\frac{1}{(1-x)^2} = \frac{\mathrm{d}}{\mathrm{d}x} \frac{1}{1-x} = \sum_{n=1}^{\infty} nx^{n-1} = \sum_{n=0}^{\infty} (n+1)x^n = 1 + 2x + 3x^2 + 4x^3 + \cdots$$
$$\ln(1-x) = -\int_0^x \frac{1}{1-t}\,\mathrm{d}t = -\sum_{n=0}^{\infty} \frac{1}{n+1} x^{n+1} = -\sum_{n=1}^{\infty} \frac{1}{n} x^n = -\left(x + \frac{1}{2}x^2 + \frac{1}{3}x^3 + \cdots\right)$$

where both are valid on $(-1, 1)$. In fact the first series has exactly this interval of convergence, whereas the second has $[-1, 1)$. By Abel's Theorem and the fact that logarithms are continuous, we have equality at $x = -1$ and recover the famous identity

$$\ln 2 = \sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n} = 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \cdots$$

This also shows that while integrated and differentiated series have the same radius of convergence as the original, convergence at the endpoints need not be the same.

2. Substitute $x \mapsto -x^2$ in the geometric series and integrate term-by-term: if $|x| < 1$, then

$$\frac{1}{1+x^2} = \sum_{n=0}^{\infty} (-1)^n x^{2n} \implies \arctan x = \sum_{n=0}^{\infty} \frac{(-1)^n}{2n+1} x^{2n+1}$$

In fact the arctangent series also converges at $x = \pm 1$; Abel's Theorem says it is continuous on $[-1, 1]$. Since arctangent is continuous (on $\mathbb{R}$!) we recover another famous identity

$$\frac{\pi}{4} = \arctan 1 = \sum_{n=0}^{\infty} \frac{(-1)^n}{2n+1} = 1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} + \cdots$$

As with the identity for $\ln 2$, this is a very slowly converging alternating series and therefore doesn't provide an efficient method for approximating π.

3. The series $f(x) = \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n)!} x^{2n}$ has radius of convergence ∞. Differentiate to obtain

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

This series is also valid for all $x \in \mathbb{R}$. Differentiating again,

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

Recalling that $f(x) = \cos x$ is the unique solution to the initial value problem

$$\begin{cases} f''(x) = -f(x) \\ f(0) = 1, \ f'(0) = 0 \end{cases}$$

We conclude that, $\forall x \in \mathbb{R}$,

$$\cos x = \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n)!} x^{2n} \qquad \sin x = -f'(x) = \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n+1)!} x^{2n+1}$$

These last expressions can be taken as the *definitions* of sine and cosine. As promised earlier, continuity and differentiability of these functions now come for free! The only real downside of this definition is believing that it has anything to do with right-triangles!

We can similarly define other common transcendental functions using power series: for instance

$$\exp(x) = \sum_{n=0}^{\infty} \frac{1}{n!} x^n$$

Example 2.33.1 could also be taken as a definition of the logarithm on the interval $(0, 2]$,

$$\ln x = \ln(1 - (1-x)) = -\sum_{n=1}^{\infty} \frac{1}{n} (1-x)^n = \sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n} (x-1)^n$$

though this is unnecessary since it is more natural to define ln as the inverse of the exponential.

Proof of Abel's Theorem (non-examinable)

This requires a lot of work, so feel free to omit on a first reading!

First observe that there is nothing to check unless $0 < R < \infty$. By the change of variable $x \mapsto \pm\frac{x}{R}$, it is enough for us to prove the following:

$$\sum_{n=0}^{\infty} a_n \text{ convergent and } f(x) = \sum_{n=0}^{\infty} a_n x^n \text{ on } (-1,1) \implies \lim_{x\to 1^-} f(x) = \sum_{n=0}^{\infty} a_n$$

Proof. Let $s_n = \sum_{k=0}^{n} a_k$ and write $s = \lim s_n = \sum a_n$. It is an easy exercise to check that

$$\sum_{k=0}^{n} a_k x^k = s_n x^n + (1-x)\sum_{k=0}^{n-1} s_k x^k$$

If $|x| < 1$, then (since $s_n \to s$) $\lim s_n x^n = 0$, whence we obtain

$$\forall x \in (-1,1),\ f(x) = (1-x)\sum_{n=0}^{\infty} s_n x^n$$

Let $\epsilon \in (0,1)$ be given and fix $x \in (0,1)$. Then

$$\exists N \in \mathbb{N} \text{ such that } n > N \implies |s_n - s| < \frac{\epsilon}{2} \qquad (*)$$

Use the geometric series formula $\sum_{n=0}^{\infty} x^n = \frac{1}{1-x}$ and write $h(x) = (1-x)\left|\sum_{n=0}^{N}(s_n - s)x^n\right|$ to observe

$$\begin{aligned}
|f(x) - s| &= \left|(1-x)\sum_{n=0}^{\infty} s_n x^n - s\right| = \left|(1-x)\sum_{n=0}^{\infty} s_n x^n - s(1-x)\sum_{n=0}^{\infty} x^n\right| \\
&= \left|(1-x)\sum_{n=0}^{\infty}(s_n - s)x^n\right| = (1-x)\left|\sum_{n=0}^{N}(s_n-s)x^n + \sum_{n=N+1}^{\infty}(s_n-s)x^n\right| \\
&\le (1-x)\left|\sum_{n=0}^{N}(s_n-s)x^n\right| + (1-x)\left|\sum_{n=N+1}^{\infty}(s_n-s)x^n\right| && (\triangle\text{-inequality}) \\
&< h(x) + \frac{\epsilon}{2}(1-x)\left|\sum_{n=N+1}^{\infty} x^n\right| && (\text{by } (*)) \\
&\le h(x) + \frac{\epsilon}{2}
\end{aligned}$$

Since $h > 0$ is continuous and $h(1) = 0$, $\exists\delta > 0$ such that $x \in (1-\delta, 1) \implies h(x) < \frac{\epsilon}{2}$ (the computation of a suitable δ is another exercise).

We conclude that $\lim_{x\to 1^-} f(x) = s$. ∎

Exercises 2.26. *Key concepts:* *Power series continuous on $(-R, R)$, uniformly on any $[-T, T] \subset (-R, R)$, Power series differentiable and integrable term-by-term on $(-R, R)$*

1. (a) Prove that $\sum_{n=1}^{\infty} nx^n = \frac{x}{(1-x)^2}$ for $|x| < 1$.

   (b) Evaluate $\sum_{n=1}^{\infty} \frac{n}{2^n}$, $\sum_{n=1}^{\infty} \frac{n}{4^n}$ and $\sum_{n=1}^{\infty} \frac{(-1)^n n}{4^n}$

2. (a) Starting with a power series centered at $x = 0$, evaluate the integral $\int_0^{1/2} \frac{1}{1+x^4}\,dx$ as an infinite series.

   (b) (Harder) Repeat part (a) but for $\int_0^1 \frac{1}{1+x^4}\,dx$. What extra ingredients do you need?

3. The probability that a standard normally distributed random variable X lies in the interval $[a, b]$ is given by the integral

$$\mathbb{P}(a \le X \le b) = \frac{1}{\sqrt{2\pi}} \int_a^b \exp\left(-\frac{x^2}{2}\right) dx$$

   Find $\mathbb{P}(-1 \le X \le 1)$ as an infinite series.

4. If $f(x) = \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n)!} x^{2n}$ is defined as in Example 2.33.3, prove that $\left(f(x)\right)^2 + \left(f'(x)\right)^2 = 1$. What does (the converse of) Pythagoras' Theorem say about $f(x)$, at least when both it and $f'(x)$ are positive?

   (*Hint: Differentiate and evaluate at zero!*)

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

   (a) Prove that $c'(x) = s(x)$ and that $s'(x) = c(x)$.

   (b) Prove that $c(x)^2 - s(x)^2 = 1$ for all $x \in \mathbb{R}$.

   (*These functions are the hyperbolic sine and cosine:* $s(x) = \sinh x$ *and* $c(x) = \cosh x$)

6. Let $a, b \in (-1, 1)$. Extending Example 2.30, show that the convergence $\sum x^n = \frac{1}{1-x}$ is non-uniform on any interval of the form $(-1, a)$ or $(b, 1)$.

7. Prove the integration part of Theorem 2.31.

8. Prove or disprove: If a series converges absolutely at the *endpoints* of its interval of convergence then its convergence is uniform on the entire interval.

9. Complete the proof of Abel's Theorem:

   (a) Let $s_n = \sum_{k=0}^{n} a_k$ be the partial sum of the series $\sum a_n$. For each n, prove that,

$$\sum_{k=0}^{n} a_k x^k = s_n x^n + (1-x) \sum_{k=0}^{n-1} s_k x^k$$

   (b) Suppose $x > 0$. Let $S = \max\{|s_n - s| : n \le N\}$ and prove that $h(x) \le S(1 - x^{N+1})$. Hence find an explicit δ that completes the final step.

2.27 The Weierstraß Approximation Theorem

A major theme of analysis is *approximation*; for instance power series are an example of (uniform) approximation by polynomials. It is reasonable to ask whether any function can be so approximated. In 1885, Weierstraß answered a specific case in the affirmative.

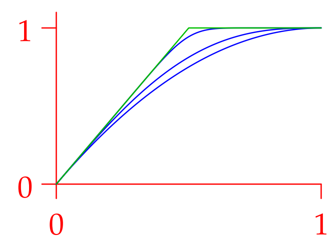
Theorem 2.34 (Weierstraß). *If $f : [a,b] \to \mathbb{R}$ is continuous, then there exists a sequence of polynomials converging uniformly to f on $[a,b]$.*

Suitable polynomials can be defined in various ways. By scaling the domain, it is enough to do this on $[a,b] = [0,1]$ where perhaps the simplest approach is via the *Bernstein Polynomials*,

$$B_n f(x) := \sum_{k=0}^{n} \binom{n}{k} f\left(\frac{k}{n}\right) x^k (1-x)^{n-k} \qquad (\tbinom{n}{k} = \tfrac{n!}{k!(n-k)!} \text{ is the binomial coefficient})$$

We omit the proof due to length; Weierstraß' original argument was completely different. Instead we compute a couple of examples and give an important interpretation/application.

Examples 2.35. 1. Suppose $f(x) = 2x$ if $x < \frac{1}{2}$ and $f(x) = 1$ otherwise.



$$B_1 f(x) = f(0)(1-x)f(0) + f(1)x = x$$

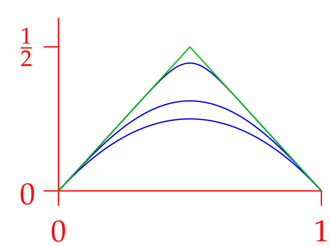
$$\begin{aligned} B_2 f(x) &= f(0)(1-x)^2 + 2f(\tfrac{1}{2})x(1-x) + f(1)x^2 \\ &= 2x(1-x) + x^2 \\ &= x(2-x) \end{aligned}$$

$$\begin{aligned} B_3 f(x) &= f(0)(1-x)^3 + 3f(\tfrac{1}{3})x(1-x)^2 + 3f(\tfrac{2}{3})x^2(1-x) + f(1)x^3 \\ &= 0(1-x)^3 + 2x(1-x)^2 + 3x^2(1-x) + x^3 \\ &= x(2-x) = B_2 f(x) \end{aligned}$$

$$\begin{aligned} B_4 f(x) &= 0(1-x)^4 + 2x(1-x)^3 + 6x^2(1-x)^2 + 4x^3(1-x) + x^4 \\ &= x(x^3 - 2x^2 + 2) \end{aligned}$$

The Bernstein polynomials $B_2 f(x), B_4 f(x)$ and $B_{50} f(x)$ are drawn.

2. Now assume $f(x) = x$ if $x < \frac{1}{2}$ and $f(x) = 1 - x$ otherwise.



$$B_1 f(x) = f(0)(1-x) + f(1)x = 0$$

$$B_2 f(x) = x(1-x)$$

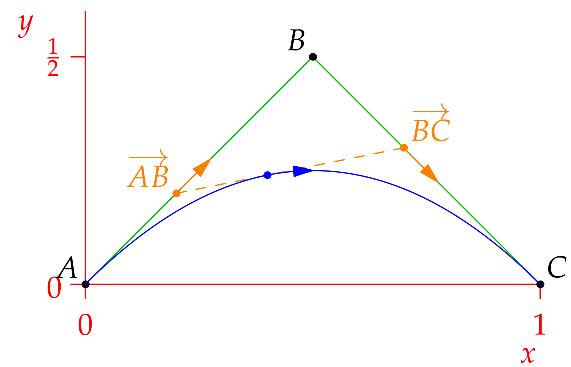
$$\begin{aligned} B_3 f(x) &= 0(1-x)^3 + x(1-x)^2 + x^2(1-x) + 0x^3 \\ &= x(1-x) = B_2 f(x) \end{aligned}$$

$$\begin{aligned} B_4 f(x) &= f(0)(1-x)^4 + f(\tfrac{1}{4})\cdot 4x(1-x)^3 + f(\tfrac{1}{2})\cdot 6x^2(1-x)^2 + f(\tfrac{3}{4})\cdot 4x^3(1-x) + f(1)x^4 \\ &= x(1-x)^3 + 3x^2(1-x)^2 + x^3(1-x) \\ &= x(1-x)(1+x-x^2) \end{aligned}$$

Bézier curves (just for fun!)

The Bernstein polynomials arise naturally when considering *Bézier curves.* These have many applications, particularly in computer graphics. Given three points A, B, C, define points on the line segments $\overrightarrow{AB}$ and $\overrightarrow{BC}$ for each $t \in [0,1]$, via

$$\overrightarrow{AB}(t) = (1-t)A + tB \qquad \overrightarrow{BC}(t) = (1-t)B + tC$$



These points move at a constant speed along the corresponding segments. Now consider a point on the *moving* segment between the points defined above:

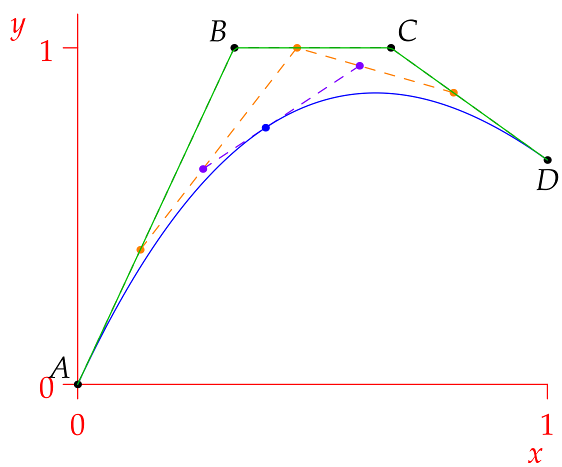
$$R(t) := (1-t)\overrightarrow{AB}(t) + t\overrightarrow{BC}(t) = (1-t)^2 A + 2t(1-t)B + t^2 C$$

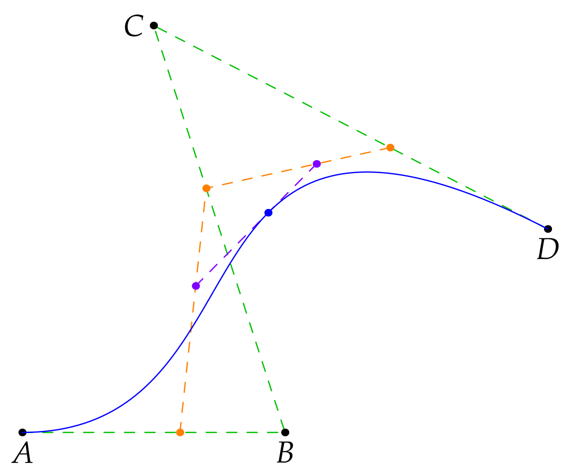
This is the *quadratic Bézier curve with control points A, B, C.* The 2^{nd} Bernstein polynomial for a function f is simply the quadratic Bézier curve with control points $(0, f(0))$, $\left(\frac{1}{2}, f(\frac{1}{2})\right)$ and $(1, f(1))$. The picture[12] above shows $B_2 f(x)$ for the above example.

We can repeat the construction with more control points: with four points A, B, C, D, one constructs $\overrightarrow{AB}(t)$, $\overrightarrow{BC}(t)$, $\overrightarrow{CD}(t)$, then the second-order points between these, and finally the cubic Bézier curve

$$\begin{aligned} R(t) : &= (1-t)\left((1-t)\overrightarrow{AB}(t) + t\overrightarrow{BC}(t)\right) + t\left((1-t)\overrightarrow{BC}(t) + t\overrightarrow{CD}(t)\right) \\ &= (1-t)^3 A + 3t(1-t)^2 B + 3t^2(1-t)C + t^3 D \end{aligned}$$

where we now recognize the relationship to the 3^{rd} Bernstein polynomial.





The pictures show cubic Bézier curves: the first is the graph of the Bernstein polynomial

$$B_3 f(x) = 0(1-x)^3 + 3x(1-x)^2 + 3x^2(1-x) + \tfrac{2}{3}x^3$$

while the second is for the four given control points A, B, C, D.

[12]Click on any of the pictures to see all of them move.

Exercises 2.27. *Key concepts: Every continuous function is the uniform limit of a polynomial sequence*

1. Show that the closed bounded interval assumption in the approximation theorem is required by giving an example of a continuous function $f : (-1, 1) \to \mathbb{R}$ which is *not* the uniform limit of a sequence of polynomials.

2. If $g : [a, b] \to \mathbb{R}$ is continuous, then $f(x) := g\big((b-a)x + a\big)$ is continuous on $[0, 1]$. If $P_n \to f$ uniformly on $[0, 1]$, prove that $Q_n \to g$ uniformly on $[a, b]$, where

$$Q_n(x) = P_n\left(\frac{x-a}{b-a}\right)$$

3. Use the binomial theorem to check that every Bernstein polynomial for $f(x) = x$ is $B_n f(x) = x$ itself!

4. Find a parametrization of the cubic Bézier curve with control points $(1, 0)$, $(0, 1)$, $(-1, 0)$ and $(0, -1)$. Now sketch the curve.

   (*Use a computer algebra package if you like!*)

5. (Hard) Show that the Bernstein polynomials for $f(x) = x^2$ are given by

$$B_n f(x) = \frac{1}{n}x + \frac{n-1}{n}x^2$$

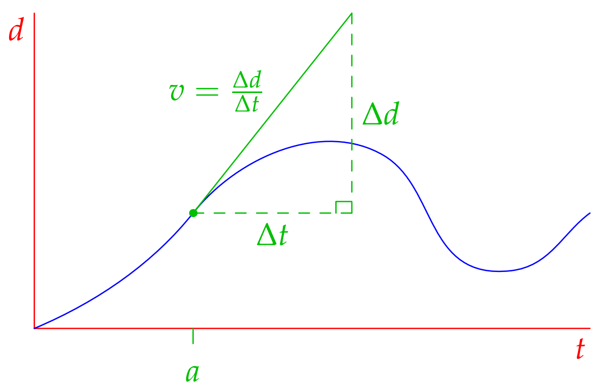
   and thus verify explicitly that $B_n f \to f$ uniformly.

3 Differentiation

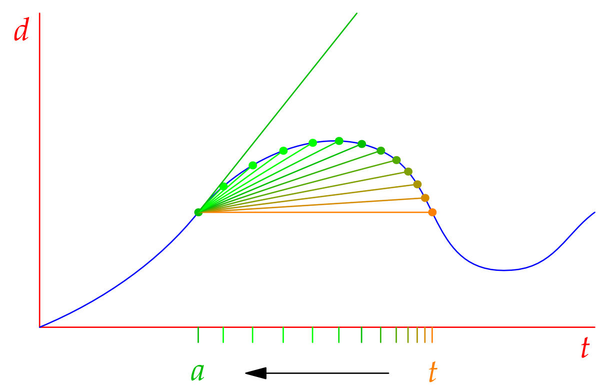
Differentiation grew out of the problem of *instantaneous velocity*. Velocity can only easily be measured as an *average* over a time interval:[13] if an object travels Δd meters in Δt seconds, then its average velocity is $v_{\text{av}} = \frac{\Delta d}{\Delta t}\ \text{ms}^{-1}$. An early 'definition' (dating to the 1300s) makes the instantaneous velocity equal to the constant velocity that would be observed if a body were to stop accelerating: while useless for the purposes of measurement, this is essentially Newton's first law regarding inertial motion (1687). We also see the concept of the *tangent line* beginning to appear. Indeed if one graphs position against time, intuition tells us:

- The graph of inertial (constant speed) motion is a straight line whose slope is the velocity.
- The tangent line to a curve has slope equal to the instantaneous velocity.

The problem of finding, defining and computing instantaneous velocity thus morphed into the consideration of tangent lines to curves. With the advent of analytic geometry in the early 1600s, mathematicians such as Fermat and Descartes pioneered versions of the familiar *secant ('cutting') line* method for computing tangents.



Instantaneous velocity equals constant velocity corresponding to tangent line



Secant lines approximate tangent line as $t \to a$

The average velocity of the particle over the time interval $[a, t]$ is the slope of the secant line, namely

$$v_{\text{av}}(a, t) = \frac{d(t) - d(a)}{t - a}$$

Since the secant lines approximate the tangent line as t approaches a, it seems reasonable that we should compute the instantaneous velocity in this manner:

$$v(a) = \lim_{t \to a} v_{\text{av}}(a, t) = \lim_{t \to a} \frac{d(t) - d(a)}{t - a}$$

This is, of course, the modern definition of the *derivative*.

[13] Even a modern technique such as Doppler-shift compares measurements separated by the extremely small period of a light or soundwave. These are still therefore *average* velocities, albeit taken over very small time intervals.

3.28 Basic Properties of the Derivative

Definition 3.1. Let $f : U \to \mathbb{R}$ and $a \in U^\circ$ an **interior point**. We say that f *is differentiable at a* if the following limit exists (is *finite*!)

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

We call this limit the *derivative of f at a* and denote its value by $\frac{\mathrm{d}f}{\mathrm{d}x}\Big|_{x=a}$ or $f'(a)$.

If $f'(a)$ exists for all $a \in U$ then f is *differentiable (on U)*; the derivative becomes a function $f'(x) = \frac{\mathrm{d}f}{\mathrm{d}x}$.

The two notations are partly attributable to the primary founders of calculus: Issac Newton and Gottfried Leibniz. Each has its pros and cons and you should be comfortable with both.

One-sided derivatives Differentiability only makes sense at interior points of U since the defining limit is two-sided. *Left-* and *right-derivatives* may be defined using one-sided limits; differentiability is then equivalent to these being equal. All results in this section hold for one-sided derivatives with suitable (sometimes tedious) modifications. It is common, though strictly incorrect, to say that f is differentiable on $[a, b)$ if it is differentiable on the interior (a, b) and *right*-differentiable at a. In these notes we will strictly adhere to Definition 3.1: differentiable means *two-sided.*

Examples 3.2. Basic examples should be familiar from elementary calculus.

1. Let $f(x) = x^2 + 4x$. Then, for any $a \in \mathbb{R}$,

$$\begin{aligned}\lim_{x\to a} \frac{f(x) - f(a)}{x - a} &= \lim_{x\to a} \frac{x^2 + 4x - a^2 - 4a}{x - a} = \lim_{x\to a} \frac{(x - a)(x + a + 4)}{x - a} \\ &= \lim_{x\to a}(x + a + 4) = 2a + 4\end{aligned}$$

   Note how the definition of $\lim\limits_{x\to a}$ allows us to cancel the $x - a$ terms from the numerator and denominator. We conclude that f is differentiable (on $\mathbb{R}$) and that $f'(x) = 2x + 4$.

2. Let $g(x) = \frac{x+1}{2x-3}$. Then, for any $a \neq \frac{3}{2}$,

$$\begin{aligned}\lim_{x\to a} \frac{f(x) - f(a)}{x - a} &= \lim_{x\to a} \frac{1}{x - a} \left[\frac{x + 1}{2x - 3} - \frac{a + 1}{2a - 3}\right] = \lim_{x\to a} \frac{5a - 5x}{(x - a)(2x - 3)(2a - 3)} \\ &= \lim_{x\to a} \frac{-5}{(2x - 3)(2a - 3)} = \frac{-5}{(2a - 3)^2}\end{aligned}$$

   f is therefore differentiable on its domain $\mathbb{R} \setminus \{\frac{3}{2}\}$ with derivative $f'(x) = \frac{-5}{(2x-3)^2}$.

The familiar expressions

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

are equivalent to the original definition (Exercise 5). While seemingly simpler, they sometimes lead to nastier calculations: see what happens if you try the previous example in this language…

We now turn to perhaps the most well-known result of elementary calculus.

Theorem 3.3 (Power Law). *Let $r \in \mathbb{R}$. Then $f(x) = x^r$ is differentiable with $f'(x) = rx^{r-1}$.*

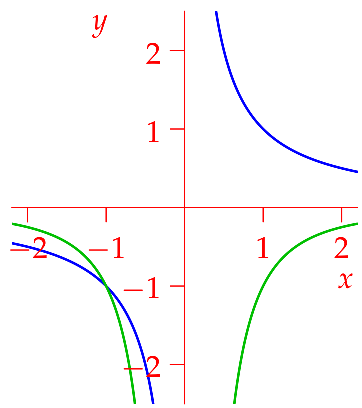
The domains of f and f' depend messily on r, but the formula holds at least on the interval $(0, \infty)$. We leave a complete proof to the exercises and instead consider a few generalizable examples.

Examples 3.4. 1. If $n \in \mathbb{N}$ and $a \in \mathbb{R}$, a simple factorization yields

$$\begin{aligned} \lim_{x \to a} \frac{x^n - a^n}{x - a} &= \lim_{x \to a} \frac{(x-a)(x^{n-1} + ax^{n-2} + \cdots + a^{n-2}x + a^{n-1})}{x - a} \qquad (*) \\ &= \lim_{x \to a} (x^{n-1} + ax^{n-2} + \cdots + a^{n-2}x + a^{n-1}) = na^{n-1} \end{aligned}$$

We conclude that $\frac{\mathrm{d}}{\mathrm{d}x} x^n = nx^{n-1}$.

2. If $f(x) = x^{-1}$ and $a \neq 0$, then



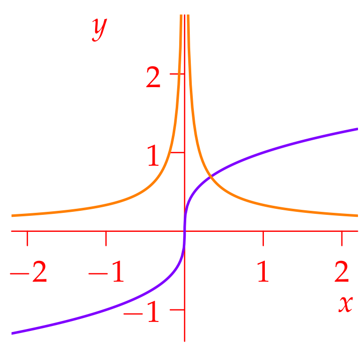
$$\lim_{x \to a} \frac{x^{-1} - a^{-1}}{x - a} = \lim_{x \to a} \frac{a - x}{ax(x-a)} = \lim_{x \to a} \frac{-1}{ax} = -\frac{1}{a^2}$$

from which we conclude that $f'(x) = -x^{-2}$.

A similar approach followed by the factorization $(*)$ proves the power law for all negative integer exponents:

$$\frac{x^{-n} - a^{-n}}{x - a} = \frac{a^n - x^n}{a^n x^n (x - a)} = \cdots$$

3. To differentiate $x^{1/n}$, substitute $x = y^n$ and observe case 1. For instance, if $g(x) = x^{1/3}$ and $a \neq 0$, then $y = x^{1/3}$ and $b = a^{1/3}$ yield



$$\begin{aligned} \lim_{x \to a} \frac{x^{1/3} - a^{1/3}}{x - a} &= \lim_{y \to b} \frac{y - b}{y^3 - b^3} = \frac{1}{3b^2} = \frac{1}{3} a^{-2/3} \\ &\implies g'(x) = \frac{1}{3} x^{-2/3} \end{aligned}$$

Note that g is *not* differentiable at $x = 0$!

We could similarly compute the derivative for all rational exponents, though it is much easier to wait for the chain rule. The power law for irrational exponents is somewhat more ticklish.

Corollary 3.5 (Basic Transcendental Functions). *Recalling our development of power series in Chapter 2, the power law (for positive integers!) is all we need to see that*

$$\frac{\mathrm{d}}{\mathrm{d}x} \exp(x) = \exp(x), \qquad \frac{\mathrm{d}}{\mathrm{d}x} \sin x = \cos x, \qquad \frac{\mathrm{d}}{\mathrm{d}x} \cos x = -\sin x$$

It is also possible to develop these results independently of power series (see e.g. Exercise 12).

Failure of differentiability

It is instructive to consider how a function might fail to be differentiable. Firstly, a familiar fact shows that functions are not differentiable at discontinuities.

Lemma 3.6. *If f is differentiable at a then f is continuous at a.*

Proof. Just take the limit (think carefully why this works!):

$$\lim_{x\to a} f(x) = \lim_{x\to a}\left[\frac{f(x)-f(a)}{x-a}(x-a)+f(a)\right] = f'(a)(0-0)+f(a) = f(a)$$ ∎

It remains to consider situations when a function is continuous but not differentiable.

Examples 3.7. The following exemplify all situations where a function is continuous on an interval and differentiable everywhere except at a single interior point. As with isolated discontinuities, these are classified by considering the three ways in which the derivative limit might not converge.

1. A *vertical tangent line* occurs when the limit is infinite. For instance, $g(x) = x^{1/3}$ at $x = 0$.

2. *Corners* occur when the one-sided limits are unequal (could be infinite). For instance, $f(x) = |x|$ is not differentiable at zero, with one-sided limits

$$\lim_{x\to 0^+}\frac{|x|-|0|}{x-0} = \lim_{x\to 0^+}\frac{x}{x} = 1 \neq \lim_{x\to 0^-}\frac{|x|-|0|}{x-0} = \lim_{x\to 0^-}\frac{-x}{x} = -1$$

   Indeed f is differentiable everywhere except at zero, with

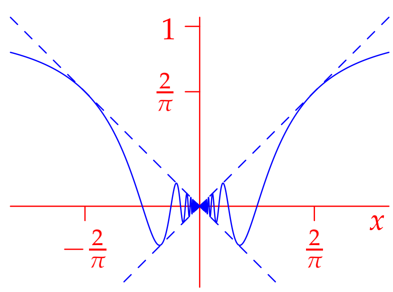
$$f'(x) = \begin{cases} 1 & \text{if } x > 0 \\ -1 & \text{if } x < 0 \end{cases}$$

   A *cusp* describes the special case where the one-sided limits are $\infty \neq -\infty$.

3. A *singularity* is where left- and/or right-limits do not exist. The standard example is

$$f(x) = \begin{cases} x\sin\frac{1}{x} & \text{if } x \neq 0 \\ 0 & \text{if } x = 0 \end{cases}$$

   which is continuous on $\mathbb{R}$ and differentiable everywhere except at zero: the details are in Exercise 10.



Singularities and vertical tangent lines can also prevent one-sided differentiability.

More esoteric examples of non-differentiability are possible:

- Utilizing series, we can create functions which are continuous on an interval but *nowhere differentiable*! For an example, see Exercise 15.
- It is also possible to construct a function which differentiable (and thus continuous) at precisely one point; can you think of an example?

The Basic Rules of Differentiation

Theorem 3.8. *Let f, g be differentiable and k, l be constants.*

1. *(Linearity) The function $kf + lg$ is differentiable with $(kf + lg)' = kf' + lg'$.*
2. *(Product rule) The function fg is differentiable with $(fg)' = f'g + fg'$.*
3. *(Inverse functions) Suppose f is bijective, $b = f^{-1}(a)$ is an interior point of $\operatorname{dom} f^{-1}$, and $f'(a) \neq 0$, then f^{-1} is differentiable at b and*

$$\left.\frac{\mathrm{d}}{\mathrm{d}y}\right|_{y=b} f^{-1}(y) = \frac{1}{f'(a)} = \frac{1}{f'\big(f^{-1}(b)\big)}$$

Proof. Parts 1 and 2 follow from the limit laws:

$$\lim_{x\to a} \frac{(kf+lg)(x) - (kf+lg)(a)}{x-a} = \lim_{x\to a} \left[k\frac{f(x)-f(a)}{x-a} + l\frac{g(x)-g(a)}{x-a} \right] = kf'(a) + lg'(a)$$

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

Note where we used the continuity of g in the second line ($\lim g(x) = g(a)$). Part 3 is an exercise. ∎

The inverse function rule should be intuitive: since the graphs of f and f^{-1} are related by reflection in the diagonal $y = x$, gradients at corresponding points are reciprocals. The result feels even more natural in Leibniz's notation: $\frac{\mathrm{d}x}{\mathrm{d}y} = \frac{1}{\mathrm{d}y/\mathrm{d}x}$.

Examples 3.9. 1. Linearity permits the differentiation of any polynomial: e.g.,

$$\frac{\mathrm{d}}{\mathrm{d}x}(7x^2 + 13x^4) = 7\frac{\mathrm{d}}{\mathrm{d}x}x^2 + 13\frac{\mathrm{d}}{\mathrm{d}x}x^4 = 14x + 52x^3$$

2. The product rule extends the reach of differentiation to include simple combinations: e.g.,

$$\frac{\mathrm{d}}{\mathrm{d}x}(x^4 \sin x) = \left(\frac{\mathrm{d}}{\mathrm{d}x}x^4\right)\sin x + x^4\frac{\mathrm{d}}{\mathrm{d}x}\sin x = 4x^3 \sin x - x^4 \cos x$$

3. Inverse trigonometric functions can now be differentiated: e.g.,

$$y = \sin^{-1} x \implies \frac{\mathrm{d}}{\mathrm{d}x}\sin^{-1} x = \frac{\mathrm{d}y}{\mathrm{d}x} = \left(\frac{\mathrm{d}x}{\mathrm{d}y}\right)^{-1} = \frac{1}{\cos y} = \frac{1}{\sqrt{1-\sin^2 y}} = \frac{1}{\sqrt{1-x^2}}$$

4. Define the natural logarithm to be the inverse of the (bijective!) exponential function $\exp(x)$:

$$y = \ln x \iff x = \exp y$$

It follows that

$$\frac{\mathrm{d}}{\mathrm{d}x}\ln x = \left(\frac{\mathrm{d}x}{\mathrm{d}y}\right)^{-1} = \frac{1}{\exp y} = \frac{1}{x}$$

The full details, and the justification that $\exp x = e^x$, are in Exercise 14.

Theorem 3.10 (Chain Rule). *If g is differentiable at a, and f is differentiable at $g(a)$, then $f \circ g$ is differentiable at a with derivative*

$$(f \circ g)'(a) = f'\big(g(a)\big)\, g'(a)$$

In Leibniz's notation, $\frac{\mathrm{d}(f \circ g)}{\mathrm{d}x} = \frac{\mathrm{d}f}{\mathrm{d}g}\frac{\mathrm{d}g}{\mathrm{d}x}$: this *looks* like a simple cancellation of the $\mathrm{d}g$ terms...[14]

Proof. Since f and g are differentiable, a is interior to $\operatorname{dom}(g)$ and $g(a)$ is interior to $\operatorname{dom}(f)$. Since g is continuous at a, there must exist some open interval $U \ni a$ for which $x \in U \implies g(x) \in \operatorname{dom}(f)$. Define $\gamma : \operatorname{dom}(f) \to \mathbb{R}$ via

$$\gamma(v) = \begin{cases} \frac{f(v) - f(g(a))}{v - g(a)} & \text{if } v \neq g(a) \\ f'\big(g(a)\big) & \text{if } v = g(a) \end{cases} \tag{$*$}$$

Since f is differentiable at $g(a)$,we see that γ is continuous there: indeed $\lim\limits_{v \to g(a)} \gamma(v) = f'\big(g(a)\big)$. For any $x \in U \setminus \{a\}$, let $v = g(x)$ in $(*)$. Then

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

Take limits as $x \to a$ for the result. ∎

Corollary 3.11 (Quotient Rule). *Suppose f and g are differentiable. Then $\frac{f}{g}$ is differentiable whenever $g(x) \neq 0$. Moreover*

$$\left(\frac{f}{g}\right)' = \frac{f'g - fg'}{g^2}$$

The proof is an exercise.

Examples 3.12. 1. By the quotient rule,

$$\frac{\mathrm{d}}{\mathrm{d}x}\tan x = \frac{\mathrm{d}}{\mathrm{d}x}\frac{\sin x}{\cos x} = \frac{\cos^2 x + \sin^2 x}{\cos^2 x} = \sec^2 x$$

2. We can now differentiate highly involved combinations of elementary functions:

$$\frac{\mathrm{d}}{\mathrm{d}x}\left[\tan(e^{4x^2}) - \frac{7x}{\sin x}\right] = 8xe^{4x^2}\sec^2(e^{4x^2}) - \frac{7\sin x - 7x\cos x}{\sin^2 x}$$

[14]This is completely unjustified since $\mathrm{d}g$ does not (for us) have independent meaning. The same problem appears in a famously flawed one-line 'proof' of the chain rule:

$$\lim_{x \to a}\frac{f\big(g(x)\big) - f\big(g(a)\big)}{x - a} \overset{?}{=} \lim_{x \to a}\frac{f\big(g(x)\big) - f\big(g(a)\big)}{g(x) - g(a)}\lim_{x \to a}\frac{g(x) - g(a)}{x - a}$$

The second limit doesn't make sense unless $g(x) \neq g(a)$ for all x on some punctured neighborhood of a: in particular, $g(x)$ cannot be *constant*! The faulty argument may be repaired by replacing this difference quotient with $f'\big(g(a)\big)$ whenever $g(x) = g(a)$, *before* taking the limit. This is precisely what $\gamma\big(g(x)\big)$ does in the correct proof.

Exercises 3.28. *Key concepts: Differentiability, Basic rules: linearity, power, product, chain, quotient*

1. Use Definition 3.1 to calculate the derivatives.

   (a) $f(x) = x^3$ at $x = 2$ (b) $g(x) = x + 2$ at $x = a$

   (c) $f(x) = x^2 \cos x$ at $x = 0$ (d) $r(x) = \frac{3x+4}{2x-1}$ at $x = 1$

2. Differentiate the function $f(x) = \cos(e^{x^5 - 3x})$ using the chain and product rules.

3. (a) Prove the quotient rule (Corollary 3.11) by combining the chain and product rules.

   (b) Prove the inverse derivative rule (Theorem 3.8, part 3).
   (*Hint: You can't simply differentiate* $1 = \frac{dx}{dx} = \frac{d}{dx} f(f^{-1}(x))$ *using the chain rule; why not?*)

4. (a) Find the derivatives of secant, cosecant and cotangent using the quotient rule.

   (b) Why did we choose the positive square-root when computing $\frac{d}{dx} \sin^{-1} x$? What is the standard domain of arcsine, and what happens at $x = \pm 1$?

   (c) Find the derivatives of the inverse trigonometric functions using the inverse function rule.

5. Using the definition of the derivative, and supposing that f is differentiable at a, prove that

$$f'(a) = \lim_{h \to 0} \frac{f(a+h) - f(a)}{h} = \lim_{h \to 0} \frac{f(a+h) - f(a-h)}{2h}$$

6. Use induction to prove the power law $\frac{d}{dx} x^n = nx^{n-1}$ when $n \in \mathbb{N}$ using *only* the product rule and the fact that $\frac{d}{dx} x = 1$.

7. Prove that $f(x) = x\,|x|$ is differentiable everywhere and compute its derivative.

8. Show that $f(x) = x^{2/3}$ has a *cusp* (see Example 3.7.2) at $x = 0$.

9. Show that following function is differentiable everywhere and compute its derivative:

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

   Moreover, prove that the derivative f' is *discontinuous* at $x = 0$.

10. Prove that the function in Example 3.7.3 is differentiable everywhere *except* at $x = 0$.

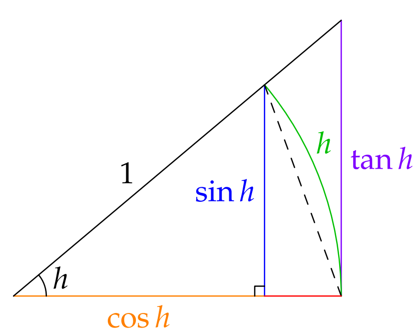
11. Suppose $f(x) = x^2$ whenever $x \in \mathbb{Q}$ and $f(x) = 0$ whenever $x \notin \mathbb{Q}$. At what values of x is f differentiable? Prove your assertion.

12. (a) Suppose $0 < h < \frac{\pi}{2}$. Use the picture to show that

$$0 < \frac{1 - \cos h}{h} < \sin \frac{h}{2} \quad \text{and} \quad \sin h < h < \tan h$$

   Hence conclude that $\lim_{h \to 0} \frac{\sin h}{h} = 1$ and $\lim_{h \to 0} \frac{1 - \cos h}{h} = 0$.

   (b) Use part (a) to prove that $\frac{d}{dx} \sin x = \cos x$



13. (Hard) Use induction to prove the Leibniz rule (general product rule):

$$(fg)^{(n)} = \sum_{k=0}^{n} \binom{n}{k} f^{(k)} g^{(n-k)}$$

Warning! The last two exercises are much longer and & tougher: have a go if you appreciate a challenge.

14. **The Exponential Function & the Power Law**

The ratio tests shows that the power series $\exp(x) := \sum_{n=0}^{\infty} \frac{x^n}{n!}$ converges for all real x. *Define* $e := \exp(1)$. Certainly e^x makes sense whenever $x \in \mathbb{Q}$. If x is irrational, instead define

$$e^x := \sup\{e^q : q \in \mathbb{Q},\ q < x\}$$

The goal of this question is to *prove* that $\exp(x) = e^x$. As a nice bonus we recover Bernoulli's limit identity $e = \lim\limits_{n\to\infty} \left(1 + \frac{1}{n}\right)^n$ and obtain a complete proof of the power law!

(a) For all $x, y \in \mathbb{R}$, prove that $\exp(x + y) = \exp(x)\exp(y)$
(*Hint: use the binomial theorem and change the order of summation*)

(b) Show that $\exp(x)$ is always positive, *even when* $x < 0$.

(c) Prove that $\exp : \mathbb{R} \to (0, \infty)$ is bijective.
(*Hint:* $x \geq 0 \implies \exp(x) \geq 1 + x$*; take limits then apply part (a)*)

(d) Prove that $e^x = \exp(x)$. Do this in three stages:

- If $x \in \mathbb{N}$, use part (a). Now check for $x \in \mathbb{Z}^-$.
- If $x = \frac{m}{n} \in \mathbb{Q}$, first compute $\left[\exp(\frac{m}{n})\right]^n$.
- If x is irrational, consider a sequence of rational numbers $q_n < x$ with $e^{q_n} \to e^x \ldots$

(e) Let $\ln : (0, \infty) \to \mathbb{R}$ be the inverse function of exp. Prove the logarithm laws:

$$\ln(xy) = \ln x + \ln y \quad \text{and} \quad \ln x^r = r \ln x$$

(*Just do this when* $r \in \mathbb{N}$*; in general, another argument like part (d) is required*)

(f) We've already seen that $\frac{\mathrm{d}}{\mathrm{d}y} \ln y = \frac{1}{y}$. Use the fact that

$$\frac{\mathrm{d}}{\mathrm{d}y} \ln y = \lim_{h\to 0} \frac{\ln(y+h) - \ln y}{h}$$

to prove that $\exp(x) = \lim\limits_{n\to\infty} \left(1 + \frac{x}{n}\right)^n$, thus recovering Bernoulli's definition of e.

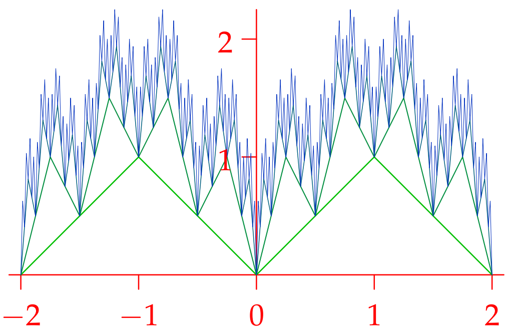
(g) For any $r \in \mathbb{R}$, *define* $x^r := \exp(r \ln x)$. Hence obtain the power law for any exponent.

15. **A Very Strange Function**

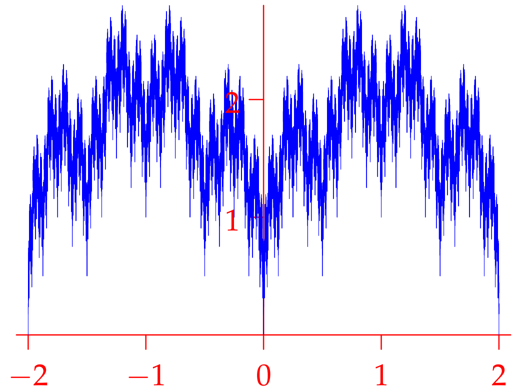
Here is a classic example of a continuous but nowhere-differentiable function!

Let f be the *sawtooth* function defined by $f(x) = |x|$ whenever $x \in [-1,1]$ and extending periodically to $\mathbb{R}$ so that $f(x+2) = f(x)$. Now define $g : \mathbb{R} \to \mathbb{R}$ via

$$g(x) = \sum_{n=0}^{\infty} \left(\frac{3}{4}\right)^n f(4^n x)$$



$f(x)$ and iterations to $n = 3$



$g(x)$ (really $n = 6$, but can you tell?!)

(a) Prove that g is well-defined and continuous on $\mathbb{R}$.

(b) Let $x \in \mathbb{R}$ and $m \in \mathbb{N}$ be fixed. Define $h_m = \pm\frac{1}{2} \cdot 4^{-m}$ where the sign is chosen so that no integers lie strictly between $4^m x$ and $4^m(x + h_m) = 4^m x \pm \frac{1}{2}$.

For each $n \in \mathbb{N}_0$, define

$$k_n = \frac{f\big(4^n(x + h_m)\big) - f(4^n x)}{h_m}$$

Prove the following

i. $|k_n| \leq 4^n$ with equality when $n = m$.

ii. $n > m \implies k_n = 0$.

(*Hint:* $|f(y) - f(z)| \leq |y - z|$*: when is this an equality?*)

(c) Use part (b) to prove that

$$\left| \frac{g(x + h_m) - g(x)}{h_m} \right| \geq \frac{1}{2}(3^m + 1)$$

Hence conclude that g is *nowhere differentiable.*

3.29 The Mean Value Theorem

A key result in elementary calculus, this should be very familiar from your previous studies.

Theorem 3.13 (Mean Value Theorem/MVT). *Let f be continuous on $[a,b]$ and differentiable on (a,b). Then there exists $\xi \in (a,b)$ such that $f'(\xi) = \frac{f(b)-f(a)}{b-a}$.*

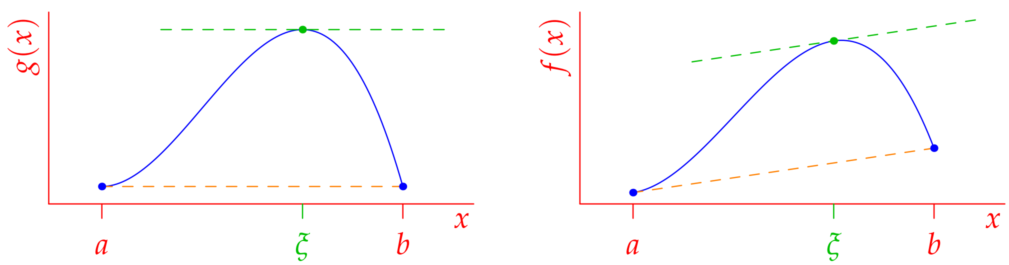
This follows easily from two lemmas.

Lemma 3.14.
1. *(Critical Points) Suppose g is bounded on (a,b) and attains its maximum or minimum at $\xi \in (a,b)$. If g is differentiable at ξ then $g'(\xi) = 0$.*
2. *(Rolle's Theorem) Suppose g is continuous on $[a,b]$, differentiable on (a,b), and $g(a) = g(b)$. Then there exists $\xi \in (a,b)$ such that $g'(\xi) = 0$.*

The main result is obtained by subtracting a straight line and applying Rolle's theorem to

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

and observing that $g(a) = f(a) = g(b)$ and $g'(x) = f'(x) - \frac{f(b)-f(a)}{b-a}$.



Critical Points/Rolle's Theorem — Mean Value Theorem

In the pictures, the orange and green lines are *parallel*: the average slope over the interval $[a,b]$ equals the gradient/derivative $f'(\xi)$.

Proof of Lemma.
1. Suppose $\xi \in (a,b)$ is a maximum: that is, $g(x) \le g(\xi)$ for all $x \ne \xi$. Then

$$\frac{g(x)-g(\xi)}{x-\xi} \begin{cases} \le 0 & \text{whenever } x > \xi \\ \ge 0 & \text{whenever } x < \xi \end{cases}$$

Now take the one-sided limits: since g is differentiable at ξ, we see that

$$0 \le \lim_{x\to\xi^+} \frac{g(x)-g(\xi)}{x-\xi} = g'(\xi) = \lim_{x\to\xi^-} \frac{g(x)-g(\xi)}{x-\xi} \le 0$$

Otherwise said $g'(\xi) = 0$. The case when ξ is a minimum is similar.

2. By the Extreme Value Theorem (1.11), g is bounded and attains its bounds. If the extrema *both* occur at the endpoints a, b, then g is constant: any $\xi \in (a,b)$ satisfies the result. Otherwise, at least one extreme occurs at some $\xi \in (a,b)$: part 1 says that $g'(\xi) = 0$.

■

Examples 3.15. 1. Let $f(x) = (x-1)^2(4-x) + x$ on $[a,b] = [1,4]$: this is roughly the above picture illustrating the mean value theorem. Compute the average slope and the derivative,

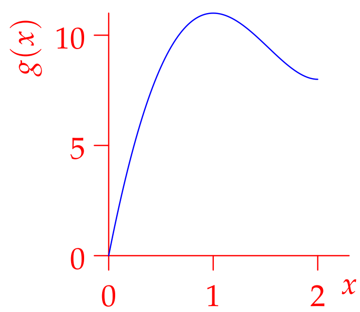
$$\frac{f(b)-f(a)}{b-a} = 1, \qquad f'(x) = 2(x-1)(4-x) - (x-1)^2 + 1 = -3x^2 + 12x - 8$$

and observe that

$$f'(\xi) = \frac{f(b)-f(a)}{b-a} \iff 3\xi^2 - 12\xi + 9 = 0 \iff \xi = 1 \text{ or } 3$$

Since only 3 lies in the interval $(1,4)$, this is the value ξ satisfying the mean value theorem.

2. We find the maximum and minimum values of $g(x) = x^4 - 14x^2 + 24x$ on the interval $[0,2]$. The function is differentiable, with



$$g'(x) = 4x^3 - 28x + 24 = 4(x-2)(x-1)(x+3)$$

By the Lemma, the locations of the extrema are either the end-points $x = 0, 2$ or locations with zero derivative ($x = 1$). Since

$$f(0) = 0, \quad f(1) = 11, \quad f(2) = 8$$

we conclude that $\max(f) = f(1) = 11$ and $\min(f) = f(0) = 0$.

Consequences of the Mean Value Theorem Several simple corollaries relate to monotonicity.

Definition 3.16. Suppose $f : I \to \mathbb{R}$ is defined on an interval I. We say that f is:

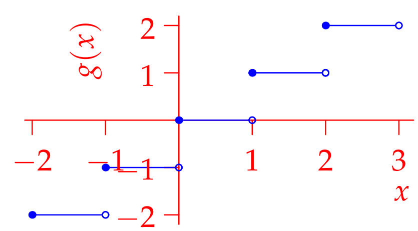
Increasing (monotone-up) on I if $x < y \implies f(x) \le f(y)$

Decreasing (monotone-down) on I if $x < y \implies f(x) \ge f(y)$

We say *strictly* increasing/decreasing if the inequalities are strict.

Examples 3.17. 1. $f : x \mapsto x^2$ is strictly increasing on $[0, \infty)$ and strictly decreasing on $(-\infty, 0]$.

2. The floor function $f : x \mapsto \lfloor x \rfloor$ (the greatest integer less than or equal to x) is increasing, but not strictly, on $\mathbb{R}$.



Corollary 3.18. *Suppose f is differentiable on an interval I. Then*

1. *$f' \ge 0$ on $I \iff f$ is increasing on I*

2. *$f' \le 0$ on $I \iff f$ is decreasing on I*

3. *$f' = 0$ on $I \iff f$ is constant on I*

Proof. (Part 1, $\Rightarrow$) Let $x < y$ where $x, y \in I$. By the mean value theorem, $\exists \xi \in (x, y)$ such that

$$\frac{f(y) - f(x)}{y - x} = f'(\xi) \quad \text{whence} \quad f'(\xi) \geq 0 \implies f(y) \geq f(x)$$

($\Leftarrow$) For the converse, use the definition of derivative: $f'(\xi) = \lim_{x \to \xi} \frac{f(x) - f(\xi)}{x - \xi}$. If f is increasing, then

$$x > \xi \implies f(x) \geq f(\xi) \implies f'(\xi) \geq 0$$

Parts 2 and 3 are similar. ∎

More care is required when relating $f' > 0$ to f being *strictly* increasing (see Exercise 5). The corollary also yields a couple of (hopefully familiar) flashbacks to elementary calculus.

Corollary 3.19. *Let I be an open interval.*

1. *(Anti-derivatives on an interval) If $f'(x) = g'(x)$ on I, then $\exists c$ such that $g(x) = f(x) + c$ on I.*
2. *(First derivative test) Suppose f is continuous on I and differentiable except perhaps at ξ. If*

$$\begin{cases} f'(x) < 0 & \text{whenever } x < \xi, \text{ and} \\ f'(x) > 0 & \text{whenever } x > \xi \end{cases} \qquad \text{then } f \text{ has its minimum value at } x = \xi$$

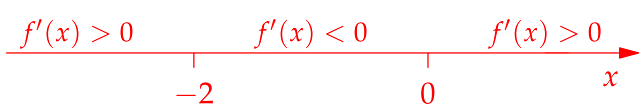
The statement for a maximum is similar.

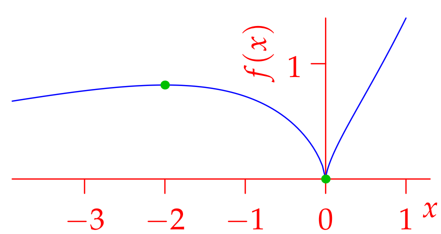
Examples 3.20. 1. Since $\frac{d}{dx} \sin(3x^2 + x) = (6x + 1)\cos(3x^2 + x)$ on (the interval) $\mathbb{R}$, whence all anti-derivatives of $f(x) = (6x + 1)\cos(3x^2 + x)$ are given by

$$\int f(x)\,dx = \int (6x + 1)\cos(3x^2 + x)\,dx = \sin(3x^2 + x) + c$$

As is typical in calculus, we use the *indefinite integral* notation $\int f(x)\,dx$ for anti-derivatives.

2. If $f(x) = x^{2/3}e^{x/3}$, then $f'(x) = \frac{1}{3}x^{-1/3}(2 + x)e^{x/3}$.
By Lemma 3.14, the only possible critical points are at $x = 0$ or -2. The sign of the derivative is also clear:





By the 1st derivative test, f has a maximum at $x = -2$ and a minimum at $x = 0$.

We finish this section by tying together the mean and intermediate value theorems.

Theorem 3.21 (IVT for Derivatives). *Suppose f is differentiable on an interval I containing $a < b$, and that L lies between $f'(a)$ and $f'(b)$. Then $\exists \xi \in (a, b)$ such that $f'(\xi) = L$.*

If $f'(x)$ is *continuous,* this is just the intermediate value theorem applied to f'; surprisingly, continuity of f' is *not* required. A full proof is in Exercise 7.

Exercises 3.29. *Key concepts: Differentiability, Basic rules: linearity, power, product, chain, quotient*

1. Determine whether the conclusion of the mean value theorem holds for each function on the given interval. If so, find a suitable point ξ. If not, state which hypothesis fails.

   (a) x^2 on $[-1,2]$
   (b) $\sin x$ on $[0,\pi]$
   (c) $|x|$ on $[-1,2]$
   (d) $1/x$ on $[-1,1]$
   (e) $1/x$ on $[1,3]$

2. Suppose f and g are differentiable on an interval I containing $a<b$ and that $f(a)=f(b)=0$. By considering $h(x)=f(x)e^{g(x)}$, prove that $f'(\xi)+f(\xi)g'(\xi)=0$ for some $\xi\in(a,b)$.

3. (a) Use the Mean Value Theorem to prove that $x<\tan x$ for all $x\in(0,\frac{\pi}{2})$.
   (b) Prove that $\frac{x}{\sin x}$ is *strictly* increasing on $(0,\frac{\pi}{2})$.
   (c) Prove that $x\le\frac{\pi}{2}\sin x$ for all $x\in[0,\frac{\pi}{2}]$.

4. Suppose that $|f(x)-f(y)|\le(x-y)^2$ for all $x,y\in\mathbb{R}$. Prove that f is a constant function.

5. (a) Prove that $f'>0$ on an interval $I\implies f$ is *strictly* increasing on I.
   (b) Show that the converse of part (a) is *false.*
   (c) Carefully prove the first derivative test (Corollary 3.19).

6. If f is differentiable on an interval I such that $f'(x)\neq 0$ for all $x\in I$, use the intermediate value theorem for derivatives to prove that f is either strictly increasing or strictly decreasing.

7. (Intermediate value theorem for derivatives) Let f,a,b and L be as in Theorem 3.21, define $g:I\to\mathbb{R}$ by $g(x)=f(x)-Lx$, and let $\xi\in[a,b]$ be such that

   $$g(\xi)=\min\{g(x):x\in[a,b]\}$$

   (a) Why can we be sure that ξ exists? If $\xi\in(a,b)$, explain why $f'(\xi)=L$.
   (b) Assume WLOG that $f'(a)<f'(b)$. Prove that $g'(a)<0<g'(b)$. By considering $\lim\limits_{x\to a^+}\frac{g(x)-g(a)}{x-a}$, show that $\exists x>a$ for which $g(x)<g(a)$. Hence complete the proof.

8. Suppose f' exists on (a,b), and is continuous except for a discontinuity at $c\in(a,b)$.

   (a) Suppose $\lim_{x\to c^+}f'(x)=L<f'(c)$. By taking $\epsilon=\frac{f'(c)-L}{2}$ in the definition of this limit and applying IVT for derivatives, obtain a contradiction.
   Hence argue that c cannot be a *removable* or a *jump* discontinuity.
   (b) Similarly, show that f' cannot have an *infinite* discontinuity by considering $\lim\limits_{x\to c^+}f'(x)=\infty$.
   (c) By parts (a) and (b), It remains to see that f' can have an essential discontinuity. Recall (Exercise 3.28.9) that

   $$f:\mathbb{R}\to\mathbb{R}:x\mapsto\begin{cases}x^2\sin\frac{1}{x} & x\neq 0\\ 0 & x=0\end{cases}$$

   is differentiable on $\mathbb{R}$, but has discontinuous derivative at $x=0$.

   i. Use $x_n=\frac{1}{2n\pi}$ and $y_n=\frac{1}{(2n+1)\pi}$ to show that f' has an essential discontinuity at $x=0$.
   ii. Prove that if $\lim s_n=0$ and $\lim f'(s_n)=M$, then $M\in[-1,1]$.
   iii. Prove that for any $L\in[-1,1]$, there is a sequence (t_n) for which $\lim f'(t_n)=L$.
   (*Hint: Use IVT for derivatives*)

3.30 L'Hôpital's Rule

We are often required to consider *indeterminate forms*: limits which do not yield easily to the standard limits laws. For instance, while it is tempting to write

$$\lim_{x\to 0}\frac{\sin 2x}{e^{3x}-1}=\frac{\lim \sin 2x}{\lim e^{3x}-1}=\frac{0}{0} \tag{$*$}$$

this is an incorrect application of the limit laws since the resulting quotient has no meaning.

Definition 3.22. An *indeterminate form* is any limit where a naïve application of the limit laws results in a meaningless expression: the primary types are $\frac{0}{0}$, $\frac{\infty}{\infty}$, $\infty-\infty$, $0\cdot\infty$, 0^0, 0^∞, and 1^∞.

Examples 3.23. 1. $\lim\limits_{x\to 7^+}(x-7)^{\frac{1}{x-7}}$ is an indeterminate form of type 0^∞.

2. Our motivating example $(*)$ may correctly be evaluated using the definition of the derivative:

$$\lim_{x\to 0}\frac{\sin 2x}{e^{3x}-1}=\lim_{x\to 0}\frac{\sin 2x-0}{x-0}\,\frac{x-0}{e^{3x}-1}=\left(\frac{\mathrm{d}}{\mathrm{d}x}\Big|_{x=0}\sin 2x\right)\left(\frac{\mathrm{d}}{\mathrm{d}x}\Big|_{x=0}e^{3x}\right)^{-1}=\frac{2}{3}$$

By considering $\lim\limits_{x\to 0}\frac{3a\sin 2x}{2(e^{3x}-1)}$, we see that an indeterminate form of type $\frac{0}{0}$ can take *any value* a!

The approach generalizes, if non-rigorously: if f,g are differentiable at a and $f(a)=0=g(a)$, then

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

Our goal is to fully justify this result and extend to several situations:

- One-sided limits, including when $a=\pm\infty$.
- When $\lim f(x)=0$ exists, but $f(a)$ does not ($g(x), g(a)$ similarly).
- Indeterminate forms of type $\frac{\infty}{\infty}$ ($\lim f(x)=\infty$, etc.).
- When the RHS cannot be cleanly evaluated: for instance $g'(a)=0$ or if the original limit is $\pm\infty$.

Here is the full result.

Theorem 3.24 (L'Hôpital's Rule). *Let $a\in\mathbb{R}\cup\{\pm\infty\}$ and suppose functions f and g satisfy:*

1. $\lim\limits_{x\to a}\frac{f'(x)}{g'(x)}=L$ *for some* $L\in\mathbb{R}\cup\{\pm\infty\}$*, and,*
2. *(a)* $\lim\limits_{x\to a}f(x)=\lim\limits_{x\to a}g(x)=0$*, or (b)* $\lim\limits_{x\to a}g(x)=\infty$ *(no condition on f)*

Then $\lim\limits_{x\to a}\frac{f(x)}{g(x)}=L$*. The same result holds for one-sided limits.*

The full proof is a behemoth—we postpone this until after several examples. In part because of this, and because examples can often be evaluated more instructively using elementary methods (as in the above example), l'Hôpital's rule is often discouraged in elementary calculus.

Examples 3.25. 1. If $f(x) = e^{4x}$ and $g(x) = 21x - 17$, then $\lim_{x\to\infty} \frac{f(x)}{g(x)}$ has type $\frac{\infty}{\infty}$. By l'Hôpital's rule,

$$\lim_{x\to\infty} \frac{f'(x)}{g'(x)} = \lim_{x\to\infty} \frac{4e^{4x}}{21} = \infty \implies \lim_{x\to\infty} \frac{e^{4x}}{21x - 17} = \infty$$

2. For an example of type $\frac{0}{0}$, consider $f(x) = x^2 - 9$ and $g(x) = \ln(4 - x)$:

$$\lim_{x\to 3^-} \frac{f'(x)}{g'(x)} = \lim_{x\to 3^-} \frac{2x}{-1/(4-x)} = \lim_{x\to 3^-} 2x(x-4) = -6 \implies \lim_{x\to 3^-} \frac{x^2 - 9}{\ln(4-x)} = -6$$

3. One can apply the rule repeatedly: for example

$$\lim_{x\to 0} \frac{e^{4x} - 1 - 4x}{x^2} = \lim_{x\to 0} \frac{4e^{4x} - 4}{2x} = \lim_{x\to 0} \frac{16e^{4x}}{2} = 8$$

This is a generally accepted abuse of protocol: one shouldn't really state the first limit until one knows the last limit exists! As long as everything works, you are fine. However...

4. It is crucially important that the limit $\lim \frac{f'}{g'}$ exists *before* applying l'Hôpital's rule! Consider $f(x) = x + \cos x$ and $g(x) = x$: certainly $\lim_{x\to\infty} \frac{f(x)}{g(x)}$ has type $\frac{\infty}{\infty}$, however

$$\lim_{x\to\infty} \frac{f'(x)}{g'(x)} = \lim_{x\to\infty} 1 - \sin x$$

does not exist! In this case the rule is unnecessary: appealing to the squeeze theorem,

$$\frac{f(x)}{g(x)} = 1 + \frac{\cos x}{x} \xrightarrow[x\to\infty]{} 1$$

5. For another reason for why l'Hôpital's rule is often prohibited in Freshman calculus, consider

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

This appears legitimate. However, recall (Exercise 3.28.12) that this limit is used to demonstrate $\frac{\mathrm{d}}{\mathrm{d}x} \sin x = \cos x$; to use this to calculate the limit on which it depends is circular logic!

The remaining indeterminate forms (Definition 3.22) may be modified so that l'Hôpital's rule applies.

Examples 3.26. 1. An indeterminate form of type $\infty - \infty$ may be transformed to one of type $\frac{0}{0}$ before applying the rule (twice):

$$\begin{aligned} \lim_{x\to 0^+} \frac{1}{e^x - 1} - \frac{1}{x} &= \lim_{x\to 0^+} \frac{x + 1 - e^x}{x(e^x - 1)} && \text{(type } \tfrac{0}{0}\text{)} \\ &= \lim_{x\to 0^+} \frac{1 - e^x}{e^x - 1 + xe^x} && \text{(still type } \tfrac{0}{0}\text{)} \\ &= \lim_{x\to 0^+} \frac{-e^x}{2e^x + xe^x} = -\frac{1}{2} \end{aligned}$$

2. For an indeterminate form of type 1^∞, we use the log laws & continuity of the exponential:

$$\lim_{x\to 0^+}(1+\sin x)^{1/x} = \exp\left(\lim_{x\to 0^+}\frac{1}{x}\ln(1+\sin x)\right) \qquad (\text{type } \tfrac{0}{0})$$
$$= \exp\left(\lim_{x\to 0^+}\frac{\cos x}{1+\sin x}\right) = e^1 = e$$

Proving l'Hôpital's Rule

The complete argument is very lengthy. It starts with an extension of the Mean Value Theorem.

Lemma 3.27 (Extended Mean Value Theorem). *Fix $a<b$, suppose f,g are continuous on $[a,b]$ and differentiable on (a,b). Then there exists $\xi \in (a,b)$ such that*

$$\big(f(b)-f(a)\big)g'(\xi) = \big(g(b)-g(a)\big)f'(\xi)$$

Proof. Apply the standard mean value theorem (really Rolle's theorem) to

$$h(t) = \big(f(b)-f(a)\big)g(t) - \big(g(b)-g(a)\big)f(t)$$

which satisfies $h(a)=h(b)$. ∎

Now for the main event. If you do nothing else, read the following proof of the simplest case. Everything else is a modification.

Proof (Case (a)/type $\frac{0}{0}$, with right limits). Suppose we have a form of type $\frac{0}{0} = \lim_{x\to a^+}\frac{f(x)}{g(x)}$ taking right-limits at a finite location a, and that the resulting limit L is finite.

First observe that condition 1 forces the existence of an interval (a,b) on which f,g are differentiable and $g'(x)\neq 0$. Everything follows from the definition the limit in condition 1, and Lemma 3.27:

$$\text{Given } \epsilon>0,\ \exists\delta\in(0,b-a) \text{ such that } a<\xi<a+\delta \implies \left|\frac{f'(\xi)}{g'(\xi)}-L\right| < \frac{\epsilon}{2} \qquad (*)$$

$$a<y<x<a+\delta \implies \exists\xi\in(y,x) \text{ such that } \frac{f(x)-f(y)}{g(x)-g(y)} = \frac{f'(\xi)}{g'(\xi)} \qquad (\dagger)$$

Since $g'\neq 0$, the usual mean value theorem says

$$\exists c\in(y,x) \text{ such that } g(x)-g(y) = g'(c)(x-y)\neq 0$$

whence we never divide by zero in $(\dagger)$. Combining $(*)$ and $(\dagger)$, observe that

$$a<x<a+\delta \implies \left|\frac{f(x)}{g(x)}-L\right| \overset{2(a)}{=} \lim_{y\to a^+}\left|\frac{f(x)-f(y)}{g(x)-g(y)}-L\right| \overset{(\dagger)}{=} \lim_{y\to a^+}\left|\frac{f'(\xi)}{g'(\xi)}-L\right| \overset{(*)}{\le} \frac{\epsilon}{2} < \epsilon$$

Note that $a<y<\xi(x,y)<x$ is a function of x,y here! Since $\epsilon>0$ is arbitrary, this is the required result. ∎

A complete proof for all indeterminate forms of type $\frac{0}{0}$ follows from some simple modifications.

If $a=-\infty$: Replace the blue part of $(*)$ as follows:

$$\text{Given } \epsilon>0,\ \exists m\le b \text{ such that } \xi<m \implies \left|\frac{f'(\xi)}{g'(\xi)}-L\right|<\frac{\epsilon}{2}$$

The rest of the proof goes through after replacing a with $-\infty$ and $a+\delta$ with m.

If $L=\infty$: Replace the green parts of $(*)$ with Given $M>0$ and $\frac{f'(\xi)}{g'(\xi)}>2M$. Fixing the rest of the proof is again straightforward.

If $L=-\infty$: Replace the green parts of $(*)$ with Given $M>0$ and $\frac{f'(\xi)}{g'(\xi)}<-2M$.

Left-limits: If f,g are differentiable on (c,a), then the blue part may be replaced with either:

- (a finite) $\exists\delta\in(0,a-c)$ such that $a-\delta<\xi<a$
- ($a=\infty$) $\exists m\ge c$ such that $\xi>m$

The blue and green parts of $(*)$ may be replaced independently.

Proof (Case (b), $\lim g(x)=\infty$*).* This requires a little more care.[15] Since $g'\neq 0$, and $\lim_{x\to a^+} g(x)=\infty$, Exercise 3.29.6 says that g is *strictly decreasing* on (a,b). By replacing b by some $\tilde b\in(a,b)$, if necessary, we may assume that

$$a<y<x<b \implies 0<g(x)<g(y) \tag{‡}$$

Assume a and L are finite and obtain $(*)$ and $(\dagger)$ as before. Let $x\in(a,a+\delta)$ be fixed and multiply $(\dagger)$ by $\frac{g(y)-g(x)}{g(y)}$ (this is *positive* by (‡)): a little algebra and the triangle inequality tell us that

$$\begin{aligned} a<y<x &\implies \frac{f(y)}{g(y)}=\frac{f'(\xi)}{g'(\xi)}+\frac{f(x)}{g(y)}-\frac{g(x)}{g(y)}\cdot\frac{f'(\xi)}{g'(\xi)} \\ &\implies \left|\frac{f(y)}{g(y)}-L\right|\le\left|\frac{f'(\xi)}{g'(\xi)}-L\right|+\frac{1}{g(y)}\left(|f(x)|+|g(x)|\left(L+\frac{\epsilon}{2}\right)\right) \end{aligned}$$

Since $\lim_{y\to a^+} g(y)=\infty$ and x is fixed, we see that there exists $\eta\le x-a<\delta$ such that

$$y\in(a,a+\eta) \implies \frac{1}{g(y)}\left(|f(x)|+|g(x)|\left(L+\frac{\epsilon}{2}\right)\right)<\frac{\epsilon}{2}$$

Finally combine with $(*)$: given $\epsilon>0, \exists\eta>0$ such that $y\in(a,a+\eta) \implies \left|\frac{f(y)}{g(y)}-L\right|<\epsilon$.
The same modifications listed above complete the proof. ∎

[15] *Forms of type* $\frac{\infty}{\infty}$*?* Instead of assumption 2. (b), why not simply assume $\lim f=\lim g=\infty$ and write $\frac{f}{g}=\frac{1/g}{1/f}$ to obtain a form of type $\frac{0}{0}$? The problem is that the derivative of the 'new' denominator $\frac{d}{dx}\frac{1}{f}=\frac{-f'}{f^2}$ need not be non-zero on any interval (a,b) and so condition 1. need not hold. We could modify this, but it would make for a weaker theorem. Example 3.25.4 illustrates the issue: $f'(x)=1+\sin x$ has zeros on any unbounded interval.

After the 2. (b) case is proved and we know that $\lim\frac{f}{g}=L$, it is then clear that $\lim f$ must also be infinite (unless $L=0$ in which case $\lim f$ could be anything and need not exist). This situation therefore really does deal with forms of type $\frac{\infty}{\infty}$.

Exercises 3.30. *Key concepts: Types of indeterminate forms, Formal statement of l'Hôpital's rule*

1. Evaluate the limits, if they exist:

   (a) $\lim_{x\to 0} \frac{x^3}{\sin x - x}$

   (b) $\lim_{x\to \frac{\pi}{2}^-} \tan x - \frac{2}{\pi - 2x}$

   (c) $\lim_{x\to 0} (\cos x)^{1/x^2}$

   (d) $\lim_{x\to 0} (1+2x)^{1/x}$

   (e) $\lim_{x\to\infty} (e^x + x)^{1/x}$

2. Suppose f is differentiable on (c, ∞) and that $\lim_{x\to\infty} [f(x) + f'(x)] = L$ is finite.

   (a) Prove that $\lim_{x\to\infty} f(x) = L$ and that $\lim_{x\to\infty} f'(x) = 0$.

   (*Hint: write* $f(x) = \frac{f(x)e^x}{e^x}$)

   (b) Does anything change if L exists and is *infinite*?

3. If $p_n(x)$ is a polynomial of degree n, use induction to prove that $\lim_{x\to\infty} p_n(x)e^{-x} = 0$

4. Let $f(x) = x + \sin x \cos x$, $g(x) = e^{\sin x} f(x)$ and $h(x) = \dfrac{2\cos x}{e^{\sin x}(f(x) + 2\cos x)}$

   (a) Prove that $\lim_{x\to\infty} f(x) = \infty = \lim_{x\to\infty} g(x)$ but that $\lim_{x\to\infty} \frac{f(x)}{g(x)}$ does not exist.

   (b) If $\cos x \neq 0$, and x is large, show that $\frac{f'(x)}{g'(x)} = h(x)$.

   (c) Prove that $\lim_{x\to\infty} h(x) = 0$. Explain why this does not contradict part (a)!

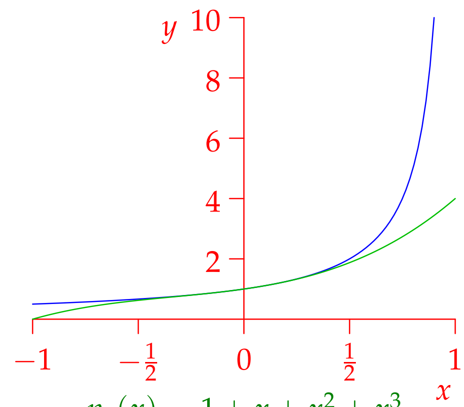
3.31 Taylor's Theorem

A primary goal of power series is the approximation of functions. With this in mind, there are two natural questions to ask of a function f:

1. Given $c \in \text{dom}(f)$, is there a series $\sum a_n(x-c)^n$ which equals $f(x)$ on an interval containing c?
2. If we take the first n terms of such a series, how accurate is this polynomial approximation?

Example 3.28. Recall the geometric series

$$f(x) = \frac{1}{1-x} = \sum_{n=0}^{\infty} x^n \text{ whenever } -1 < x < 1$$



The polynomial approximation

$$p_n(x) = \sum_{k=0}^{n} x^k = 1 + x + \cdots + x^n = \frac{1-x^{n+1}}{1-x}$$

has error

$$R_n(x) = f(x) - p_n(x) = \frac{x^{n+1}}{1-x}$$

If x is close to 0, this is likely very small; for instance if $x \in \left[-\frac{1}{2}, \frac{1}{2}\right]$, then

$$|R_n(x)| \le \frac{1}{1-\frac{1}{2}} \left(\frac{1}{2}\right)^{n+1} = 2^{-n}$$

However, when x is close to 1 the error is unbounded!

The above behavior occurs in general: the truncated polynomials provide better approximations nearer the center of the series. To see this, we first need to consider higher-order derivatives.

Definition 3.29. We write f'' for the *second derivative* of f, namely the derivative of its derivative

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

The existence of $f''(a)$ presupposes that f' exists on an (open) interval containing a. We can similarly consider third, fourth, and higher-order derivatives. As a function, the n^{th} derivative is written

$$f^{(n)}(x) = \frac{\mathrm{d}^n f}{\mathrm{d}x^n}$$

By convention, the *zeroth derivative* is the function itself $f^{(0)}(x) = f(x)$. We say that f is *n times differentiable* at a if $f^{(n)}(a)$ exists, and *infinitely differentiable* (or *smooth*) if derivatives of all orders exist.

Example 3.30. $f(x) = x^2|x|$ is twice differentiable, with $f''(x) = 6|x|$. It is smooth everywhere except at $x = 0$, where third (and higher-order) derivatives do not exist.

Definition 3.31. Suppose f is n times differentiable at $x = c$. The n^{th} *Taylor polynomial* p_n of f centered at c is

$$p_n(x) := \sum_{k=0}^{n} \frac{f^{(k)}(c)}{k!}(x-c)^k = f(c) + f'(c)(x-c) + \frac{f''(c)}{2}(x-c)^2 + \cdots + \frac{f^{(n)}(c)}{n!}(x-c)^n$$

The *remainder* $R_n(x)$ is the error in the polynomial approximation

$$R_n(x) = f(x) - p_n(x) = f(x) - \sum_{j=0}^{n} \frac{f^{(k)}(c)}{k!}(x-c)^k$$

If f is infinitely differentiable at $x = c$, then its *Taylor series* centered at $x = c$ is the power series

$$\mathrm{T}_c f(x) = \sum_{n=0}^{\infty} \frac{f^{(n)}(c)}{n!}(x-c)^n$$

When $c = 0$ this is known as a *Maclaurin series*.[16]

For simplicity we'll mostly work with Maclaurin series, with general situation hopefully being clear.

Examples 3.32. 1. If $f(x) = e^{3x}$, then $f^{(n)}(x) = 3^n e^x$, from which the Maclaurin series is

$$\mathrm{T}_0 f(x) = \sum_{n=0}^{\infty} \frac{3^n}{n!} x^n$$

2. If $g(x) = \sin 7x$, then the sequence of derivatives is

$$7\cos 7x, \quad -7^2 \sin 7x, \quad -7^3 \cos 7x, \quad 7^4 \sin 7x, \quad 7^5 \cos 7x, \quad -7^6 \sin 7x, \quad \ldots$$

At $x = 0$, every even derivative is zero whereas the odd derivatives alternate in sign. The Maclaurin series is easily seen to be

$$\mathrm{T}_0 g(x) = \sum_{n=0}^{\infty} \frac{(-1)^n 7^{2n+1}}{(2n+1)!} x^{2n+1}$$

3. If $h(x) = \sqrt{x}$, then $h'(x) = \frac{1}{2}x^{-1/2}$, $h''(x) = \frac{-1}{2^2}x^{-3/2}$, and $h'''(x) = \frac{3}{2^3}x^{-5/2}$, from which the third Taylor polynomial centered at $c = 1$ is

$$\begin{aligned} p_2(x) &= h(1) + h'(1)(x-1) + \frac{h''(1)}{2}(x-1)^2 + \frac{h'''(1)}{6}(x-1)^3 \\ &= 1 + \frac{1}{2}(x-1) - \frac{1}{8}(x-1)^2 + \frac{1}{16}(x-1)^3 \end{aligned}$$

Rather than computing further examples, we first develop a little theory that makes verifying Taylor series much easier.

[16]Named for Englishman Brook Taylor (1685–1731) and Scotsman Colin Maclaurin (1698–1746). Taylor's general method expanded on examples discovered by James Gregory and Issac Newton in the mid-to-late 1600s.

Differentiation of Taylor Polynomials and Series

Suppose $P(x) = \sum a_j x^j$ is a power series with radius of convergence $R > 0$. As we saw previously (Theorem 2.31), $P(x)$ is differentiable term-by-term on $(-R, R)$. Indeed,

$$
\begin{aligned}
P'(x) &= \sum_{j=1}^{\infty} a_j j x^{j-1} \implies P'(0) = a_1 \\
P''(x) &= \sum_{j=2}^{\infty} a_j j(j-1) x^{j-2} \implies P''(0) = 2a_2 \\
P'''(x) &= \sum_{j=3}^{\infty} a_j j(j-1)(j-2) x^{j-3} \implies P'''(0) = 3!a_3 \\
&\vdots \\
P^{(k)}(x) &= \sum_{j=k}^{\infty} a_j j(j-1)\cdots(j-k+1) x^{j-k} = \sum_{j=k}^{\infty} \frac{j!a_j}{(j-k)!} x^{j-k} \implies P^{(k)}(0) = k!a_k
\end{aligned}
$$

Otherwise said, P is its own Maclaurin series! The same discussion holds for polynomials. Indeed if $P(x) = a_0 + a_1 x + \cdots + a_n x^n$ is a polynomial and f a function, then

$$
P^{(k)}(0) = f^{(k)}(0) \iff a_k = \frac{f^{(k)}(0)}{k!}
$$

If this holds for *all* $k \le n$, then $P = p_n$ is the n^{th} Taylor polynomial of f! With a little modification, we've proved the following:

Theorem 3.33. 1. *If $f(x) = \sum a_n (x-c)^n$ is a power series defined on a neighborhood of c, then $T_c f(x) = f(x)$: the function is its own Taylor series!*

2. *The n^{th} Taylor polynomial of f centered at $x = c$ is the unique polynomial p_n of degree $\le n$ whose value and first n derivatives agree with those of f at $x = c$: that is*

$$
\forall k \le n, \ \ p_n^{(k)}(c) = f^{(k)}(c)
$$

This answers our first motivating question: a function can equal at most one power series with a given center. The second question requires a careful study of the *remainder*: we'll do this shortly.

Examples 3.34 (Common Maclaurin Series). These should be familiar from elementary calculus. Each function equals the given series form our previous discussions of power series: by the Theorem, each series is immediately the Maclaurin series of the given function.

$$
\begin{aligned}
e^x &= \sum_{n=0}^{\infty} \frac{x^n}{n!} && x \in \mathbb{R} & \frac{1}{1-x} &= \sum_{n=0}^{\infty} x^n && x \in (-1,1) \\
\sin x &= \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n+1)!} x^{2n+1} && x \in \mathbb{R} & \ln(1+x) &= \sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n} x^n && x \in (-1,1] \\
\cos x &= \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n)!} x^{2n} && x \in \mathbb{R} & \tan^{-1} x &= \sum_{n=0}^{\infty} \frac{(-1)^n}{2n+1} x^{2n+1} && x \in [-1,1]
\end{aligned}
$$

Examples 3.35 (Modifying Maclaurin Series). By substituting for x in a common series, we quickly obtain new series.

1. Substitute $x \mapsto 7x$ in the Maclaurin series for $\sin x$, to recover our earlier example

$$\sin 7x = \sum_{n=0}^{\infty} \frac{(-1)^n 7^{2n+1}}{(2n+1)!} x^{2n+1}, \quad x \in \mathbb{R}$$

   Note how this requires almost no calculation: since the function equals a series, the Theorem says we have the Maclaurin series for $\sin 7x$!

2. Substitute $x \mapsto x^2$ in the Maclaurin series for e^x to obtain

$$e^{x^2} = \exp(x^2) = \sum_{n=0}^{\infty} \frac{1}{n!} x^{2n}, \quad x \in \mathbb{R}$$

   This would be disgusting to verify directly, given the difficulty of repeatedly differentiating e^{x^2}.

3. We find the Taylor series for $f(x) = \frac{1}{5-x}$ centered at $x = 2$:

$$f(x) = \frac{1}{3+2-x} = \frac{1}{3(1 - \frac{2-x}{3})} = \frac{1}{3} \sum_{n=0}^{\infty} \left(\frac{2-x}{3} \right)^n$$

   which is valid whenever $-1 < \frac{2-x}{3} < 1 \iff -1 < x < 5$.

4. Fix $c \in \mathbb{R}$ and observe that, for all $x \in \mathbb{R}$,

$$e^x = e^{c+x-c} = e^c e^{x-c} = \sum_{n=0}^{\infty} \frac{e^c}{n!} (x-c)^n$$

   We conclude that the series is the Taylor series of e^x centered at $x = c$. Of course this is easily verified using the definition, since $\frac{\mathrm{d}^n}{\mathrm{d}x^n}\Big|_{x=c} e^x = e^c$.

5. Combining the Theorem with the multiple-angle formula, we obtain the Taylor series for $\sin x$ centered at $x = c$:

$$\begin{aligned} \sin x &= \sin(c + x - c) = \sin c \cos(x-c) + \cos x \sin(x-c) \\ &= \sum_{n=0}^{\infty} \frac{(-1)^n \sin c}{(2n)!} (x-c)^{2n} + \sum_{n=0}^{\infty} \frac{(-1)^n \cos c}{(2n+1)!} (x-c)^{2n+1} \end{aligned}$$

Definition 3.36. A function f is *analytic* on a its domain if every $c \in \operatorname{dom} f$ has a neighborhood on which $f(x)$ equals its Taylor series centered at c.

All the examples we've thus far seen are analytic on their domains; indeed the last two of Examples 3.35 *prove* this for the exponential and sine functions. Every analytic function is automatically smooth (infinitely differentiable), however the converse is *false* (Exercise 10). Analyticity is of greater importance in complex analysis where (amazingly!) it is equivalent to complex-differentiability.

Accuracy of Taylor Approximations

Our final goal is to estimate the accuracy of a Taylor polynomial as an approximation to its generating function. Otherwise said, we want to estimate the size of the remainder $R_n(x) = f(x) - p_n(x)$.

Theorem 3.37 (Taylor's Theorem: Lagrange's form). *Suppose f is $n+1$ times differentiable on an open interval I containing c and let $x \in I \setminus \{c\}$. Then there exists ξ between c and x for which the remainder centered at c satisfies*

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

Proof. For simplicity let $c = 0$. Fix $x \neq 0$, define a constant M_x and a function $g : I \to \mathbb{R}$ by

$$R_n(x) = \frac{M_x}{(n+1)!}x^{n+1} \quad \text{and} \quad g(t) = \frac{M_x}{(n+1)!}t^{n+1} + p_n(t) - f(t) = \frac{M_x}{(n+1)!}t^{n+1} - R_n(t)$$

Observe that

$$\begin{aligned} k \leq n+1 \implies g^{(k)}(x) &= \frac{M_x}{(n+1-k)!}t^{n+1-k} + p_n^{(k)}(t) - f^{(k)}(t) \qquad (*) \\ \implies g^{(k)}(0) &= p_n^{(k)}(0) - f^{(k)}(0) = 0 \quad \text{if } k \leq n \end{aligned}$$

where we invoked Theorem 3.33.
Now apply Rolle's Theorem repeatedly (WLOG assume $x > 0$):

- $\exists \xi_1$ between 0 and x such that $g'(\xi_1) = 0$.
- $\exists \xi_2$ between 0 and ξ_1 such that $g''(\xi_2) = 0$, etc.
- Iterate to obtain a sequence (ξ_k) such that

$$0 < \xi_{n+1} < \xi_n < \cdots < \xi_1 < x \quad \text{and} \quad g^{(k)}(\xi_k) = 0$$

Take $\xi = \xi_{n+1}$ and consider $(*)$: since $\deg p_n \leq n$, we see that

$$0 = g^{(n+1)}(\xi) = M_x - f^{(n+1)}(\xi) \implies R_n(x) = f(x) - p_n(x) = \frac{f^{(n+1)}(\xi)}{(n+1)!}x^{n+1}$$

■

Corollary 3.38. *Suppose f is smooth on an open interval I containing c and that all derivatives $f^{(n)}$ of all orders are bounded on I. Then f equals its Taylor series (centered at c) on I.*

Proof. For simplicity, let $c = 0$. Suppose $\left|f^{(n+1)}(\xi)\right| \leq K$ for all $\xi \in I$. Choose any $N > |x|$ and observe that

$$n > N \implies |R_n(x)| \leq \frac{K|x|^{n+1}}{(n+1)!} = \frac{K|x|^{n+1}}{N!(N+1)\cdots(n+1)} \leq \frac{K|x|^N}{N!}\left(\frac{|x|}{N}\right)^{n+1-N} \xrightarrow[n\to\infty]{} 0$$

■

Examples 3.39. 1. The functions sine and cosine have derivatives bounded by 1 on $\mathbb{R}$, and thus both functions equal their Maclaurin series on $\mathbb{R}$. This removes the need to have previously justified these facts using the theory of differential equations.

2. The exponential function does not have bounded derivatives, however we can still apply Taylor's Theorem. For any fixed x, $\exists\xi$ between 0 and x such that

$$|R_n(x)| = \left|\frac{e^{\xi}}{(n+1)!}x^{n+1}\right| \xrightarrow[n\to\infty]{} 0$$

by the same argument in the Corollary. Thus e^x equals its Maclaurin series on the real line (we knew this already from Exercise 3.28.14).

3. Extending Example 3.32.3, we see that $h(x) = \sqrt{x}$ has linear approximation (1st Taylor polynomial) centered at $c = 9$

$$p_1(x) = h(9) + h'(9)(x-9) = 3 + \frac{1}{6}(x-9)$$

This yields the simple approximation

$$\sqrt{10} \approx p_1(10) = 3 + \frac{1}{6} = \frac{19}{6}$$

Taylor's Theorem can be used to estimate its accuracy (remember to shift the center to 9!):

$$R_1(10) = \frac{h''(\xi)}{2!}(10-9)^2 = -\frac{1}{2^2\cdot 2!}\xi^{-3/2} = -\frac{1}{8\xi^{3/2}} \quad \text{for some} \quad \xi \in (9,10)$$

Certainly $\xi^{-3/2} < 9^{-3/2} = \frac{1}{27}$, whence

$$-\frac{1}{216} < R_1(10) < 0 \implies \frac{19}{6} - \frac{1}{216} = \frac{683}{216} < \sqrt{10} < \frac{684}{216} = \frac{19}{6}$$

$\frac{19}{6}$ is therefore an overestimate for $\sqrt{10}$, but is accurate to within $\frac{1}{216} < 0.005$.

Alternative Versions of Taylor's Theorem

There are two further common expressions for the remainder in Taylor's Theorem. These are typically less easy to use than Lagrange's form but can sometimes provide sharper estimates for the remainder, particularly when x is far from the center of the series.

Corollary 3.40. *Suppose $f^{(n+1)}$ is continuous on an open interval I containing c, let $x \in I \setminus \{c\}$, and let $R_n(x) = f(x) - p_n(x)$ be the remainder for the Taylor polynomial centered at c. Then:*

1. *(Integral Remainder)* $R_n(x) = \int_c^x \frac{(x-t)^n}{n!} f^{(n+1)}(t)\,\mathrm{d}t$

2. *(Cauchy's Form)* $\exists\xi$ *between c and x such that* $R_n(x) = \frac{(x-\xi)^n}{n!}(x-c)f^{(n+1)}(\xi)$

Using these expressions it is possible to explicitly prove Newton's binomial series formula.

Corollary 3.41. *If $\alpha \in \mathbb{R}$ and $|x| < 1$, then*

$$\begin{aligned}(1+x)^\alpha &= 1 + \sum_{n=1}^{\infty} \frac{\alpha(\alpha-1)\cdots(\alpha-n+1)}{n!}x^n \\ &= 1 + \alpha x + \frac{\alpha(\alpha-1)}{2!}x^2 + \frac{\alpha(\alpha-1)(\alpha-2)}{3!}x^3 + \frac{\alpha(\alpha-1)(\alpha-2)(\alpha-3)}{4!}x^4 + \cdots\end{aligned}$$

If $\alpha \in \mathbb{N}_0$, this is the usual binomial theorem. Otherwise it is more interesting: for instance,

$$\sqrt{1+x} = (1+x)^{1/2} = 1 + \frac{1}{2}x - \frac{1}{8}x^2 + \frac{1}{16}x^3 - \frac{5}{128}x^4 + \cdots$$
$$\frac{1}{(1+x)^3} = 1 - 3x + 6x^2 - 10x^3 + 15x^4 - \cdots$$

Of course this last could easily be obtained from $\frac{1}{1+x} = \sum(-1)^n x^n$ by differentiating twice!

Exercises 3.31. *Key concepts: Taylor Series/Polynomials, Lagrange's form for Remainder*

1. Compute the Maclaurin series for $\cos x$ directly from the definition and use Taylor's Theorem to indicate why it converges to $\cos x$ for all $x \in \mathbb{R}$.
2. Repeat the previous exercise for $\sinh x = \frac{1}{2}(e^x - e^{-x})$ and $\cosh x = \frac{1}{2}(e^x + e^{-x})$.
3. Find the Maclaurin series for the function $\sin(3x^2)$. How do you know you are correct?
4. Find the Taylor series of $f(x) = x^4 - 3x^2 + 2x - 5$ at $x = 2$ and show that $T_2 f(x) = f(x)$.
5. Find a rational approximation to $\sqrt[3]{9}$ using the first Taylor polynomial for $f(x) = \sqrt[3]{x}$. Now use Taylor's Theorem to estimate its accuracy.
6. If $c \neq 1$, use the fact that $1 - x = (1-c)\left(1 - \frac{x-c}{1-c}\right)$ to obtain the Taylor series of $\frac{1}{1-x}$ centered at c. Hence conclude that $\frac{1}{1-x}$ is analytic on its domain $\mathbb{R} \setminus \{1\}$.
7. We prove that the Maclaurin series $\sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n}x^n$ converges to $\ln(1+x)$ whenever $0 < x \leq 1$.
   (a) Explicitly compute $\frac{d^{n+1}}{dx^{n+1}} \ln(1+x)$.
   (b) Suppose $0 < x \leq 1$. Using Taylor's Theorem, prove that $\lim_{n\to\infty} R_n(x) = 0$.

   (*If $-1 < x < 0$, the argument is tougher, being similar to Exercise 11*)
8. Why can't we use Taylor's Theorem to approximate the error in $\frac{1}{1-x} = 1 + x + R_1(x)$ when $x \geq 1$? Try it when $x = 2$, what happens? What about when $x = -2$?
9. Prove Taylor's Theorem with integral remainder when $c = 0$ by using the following as an induction step: for each $n \in \mathbb{N}$, define

   $$A_n(x) = \int_0^x \frac{(x-t)^n}{n!} f^{(n+1)}(t)\,dt$$

   and use integration by parts to prove that $A_{n+1} = A_n - \frac{x^{n+1}}{(n+1)!}f^{(n+1)}(0)$.

   (*The Cauchy form follows from the intermediate value theorem for integrals which we'll see later*)

10. Consider the function

$$f(x) = \begin{cases} e^{-1/x} & \text{if } x > 0 \\ 0 & \text{otherwise} \end{cases}$$

(a) Prove by induction that there exists a degree $2n$ polynomial q_n for which

$$f^{(n)}(x) = q_n\left(\tfrac{1}{x}\right) e^{-1/x} \text{ whenever } x > 0$$

(b) Prove that f is infinitely differentiable at $x = 0$ with $f^{(n)}(0) = 0$ (*use Exercise 3.30.3*).

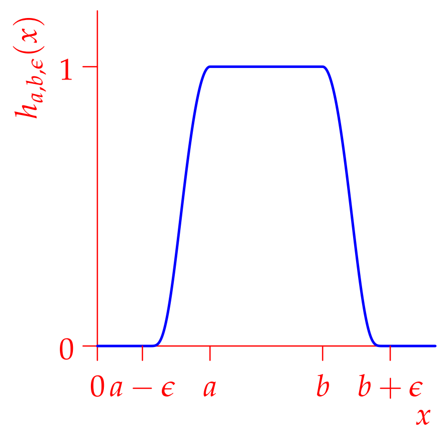
The Maclaurin series of f is identically zero! Moreover, f is smooth (infinitely differentiable) on $\mathbb{R}$ but non-analytic at zero since it does not equal its Taylor series on any open interval containing zero.
A modification allows us to create bump functions, *which find wide use in analysis. If $a < b$, define*

$$g_{a,b} : x \mapsto f(x-a)f(b-x)$$

This is smooth on $\mathbb{R}$ but non-zero only on the interval (a,b). A further modification involving two such functions $g_{a,b}$ creates a smooth function on $\mathbb{R}$ which satisfies

$$h_{a,b,\epsilon}(x) = \begin{cases} 0 & \text{if } x \le a-\epsilon \text{ or } x \ge b+\epsilon \\ 1 & \text{if } a \le x \le b \end{cases}$$

This 'switches on' rapidly from 0 to 1 near a and switches off similarly near b. By letting ϵ be small, we smoothly (but not uniformly) approximate the indicator function on $[a,b]$.



11. (Hard) We prove the binomial series formula (Corollary 3.41).

Let $f(x) = (1+x)^\alpha$ and $g(x) = 1 + \sum_{n=1}^\infty a_n x^n$ where $a_n = \frac{\alpha(\alpha-1)\cdots(\alpha-n+1)}{n!}$. Our goal is to prove that $f = g$ on the interval $(-1,1)$.

(a) Check that $f^{(n)}(0) = n!a_n$ so that g really is the Maclaurin series of f.

(b) i. Prove that the radius of convergence of g is 1.

ii. Prove that $\lim_{n\to\infty} na_n x^n = 0$ whenever $|x| < 1$.

iii. If $|x| < 1$ and ξ lies between 0 and x, prove that $\left|\frac{x-\xi}{1+\xi}\right| \le |x|$.
(*Hint: write $\xi = tx$ for some $t \in (0,1)\ldots$*)

(c) Use Taylor's Theorem with Cauchy remainder to prove that

$$|R_n(x)| < (n+1)\,|a_{n+1}|\,|x|^{n+1}\,(1+\xi)^{\alpha-1}$$

Hence conclude that $g = f$ whenever $|x| < 1$.

(d) Here is an alternative argument for the full result:

i. Show that $(n+1)a_{n+1} + na_n = \alpha a_n$.

ii. Differentiate term-by-term to prove directly that g satisfies the differential equation $(1+x)g'(x) = \alpha g(x)$. Solve this to show that $g = f$ whenever $|x| < 1$.

4 Integration

The theory of infinite series addresses how to sum infinitely many *finite* quantities. Integration, by contrast, is the business of summing infinitely many *infinitesimal* quantities. Attempts to do both have been part of mathematics for well over 2000 years, and the philosophical objections are just as old.[17] The development and increased application of calculus from the late 1600s onward spurred mathematicians to put the theory on a firmer footing, though from Newton and Leibniz it took another 150 years before Bernhard Riemann (1856) provided a thorough development of the integral.

4.32 The Riemann Integral

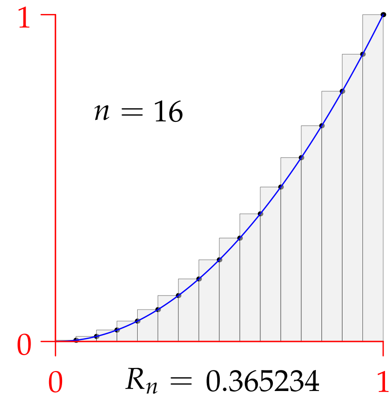
The basic idea behind Riemann integration is to approximate area using a sequence of rectangles whose *width* tends to zero. The following discussion illustrates the essential idea, which should be familiar from elementary calculus.

Example 4.1. Suppose $f(x) = x^2$ is defined on $[0,1]$.

For each $n \in \mathbb{N}$, let $\Delta x = \frac{1}{n}$ and define $x_i = i\Delta x$.

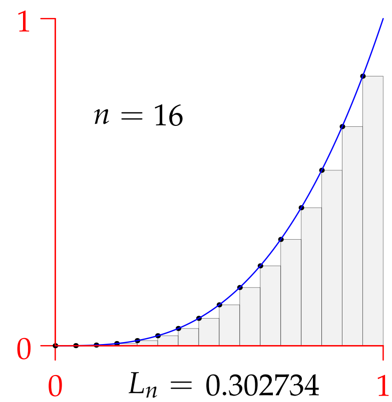
Above each *subinterval* $[x_{i-1}, x_i]$, raise a rectangle of height $f(x_i) = x_i^2$. The sum of the areas of these rectangles is the *Riemann sum with right-endpoints*[18]

$$R_n = \sum_{i=1}^{n} f(x_i)\Delta x = \sum_{i=1}^{n} \frac{i^2}{n^3} = \frac{n(n+1)(2n+1)}{6n^3} = \frac{1}{3} + \frac{3n+1}{6n^2}$$



The *Riemann sum with left-endpoints* is defined similarly:

$$L_n = \sum_{i=1}^{n} f(x_{i-1})\Delta x = \sum_{i=1}^{n} \frac{(i-1)^2}{n^3} = \frac{1}{3} - \frac{3n-1}{6n^2}$$



Since f is an increasing function, the area A under the curve plainly satisfies

$$L_n \le A \le R_n$$

By the squeeze theorem, we conclude that $A = \frac{1}{3}$.

The example should feel convincing, though perhaps this is due to the simplicity of the function. To apply this approach to more general functions, we need to be significantly more rigorous.

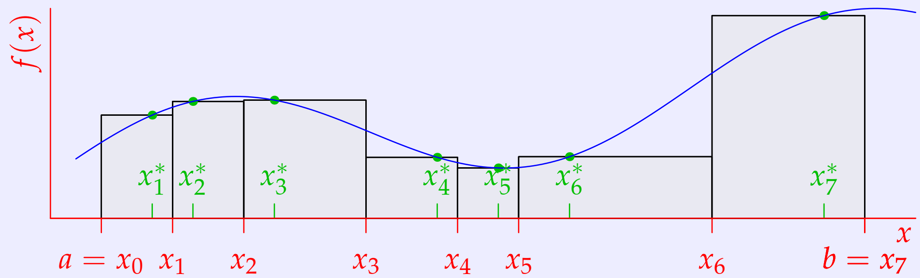
[17]Two of Zeno's ancient paradoxes are relevant here: Achilles and the Tortoise concerns a convergent infinite series, while the Arrow Paradox toys with integration by questioning whether time can be viewed as a sum of instants. Perhaps the most famous contemporary criticism comes from Bishop George Berkeley, who gave his name to the city and first UC campus: in 1734's *The Analyst*, Berkeley savaged the foundations of calculus, describing the infinitesimal increments required in Newton's theory of *fluxions* (derivatives) as merely the "ghosts of departed quantities."

[18]Recall some basic identities: $\sum_{i=1}^{n} i = \frac{1}{2}n(n+1)$, $\sum_{i=1}^{n} i^2 = \frac{1}{6}n(n+1)(2n+1)$, $\sum_{i=1}^{n} i^3 = \frac{1}{4}n^2(n+1)^2$

Definition 4.2. A *partition* $P = \{x_0, \ldots, x_n\}$ of an interval $[a, b]$ is a finite sequence for which

$$a = x_0 < x_1 < \cdots < x_{n-1} < x_n = b$$

Choosing a *sample point* x_i^* in each *subinterval* $[x_{i-1}, x_i]$ results in a *tagged partition*.
The *mesh* of the partition is $\text{mesh}(P) := \max \Delta x_i$, the width $\Delta x_i = x_i - x_{i-1}$ of the largest subinterval.
If $f : [a, b] \to \mathbb{R}$, the *Riemann sum* $\sum_{i=1}^{n} f(x_i^*)\, \Delta x_i$ evaluates the area of a family of n rectangles, as pictured. The heights $f(x_i^*)$ and thus areas can be negative or zero.



In elementary calculus, one typically computes Riemann sums for *equally-spaced* partitions with *left*, *right* or *middle* sample points. The flexibility of tagged partitions makes applying Riemann's definition a challenge, so we instead consider two special families of rectangles.

Definition 4.3. Given a partition P of $[a, b]$ and a bounded function f on $[a, b]$, define

$$M_i = \sup_{x \in [x_{i-1}, x_i]} f(x) \qquad U(f, P) = \sum_{i=1}^{n} M_i\, \Delta x_i$$

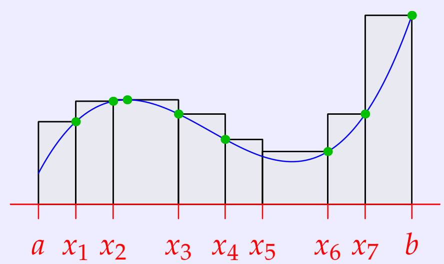
$$m_i = \inf_{x \in [x_{i-1}, x_i]} f(x) \qquad L(f, P) = \sum_{i=1}^{n} m_i\, \Delta x_i$$

$U(f, P)$ and $L(f, P)$ are the *upper* and *lower Darboux sums* for f with respect to P. The *upper* and *lower Darboux integrals* are

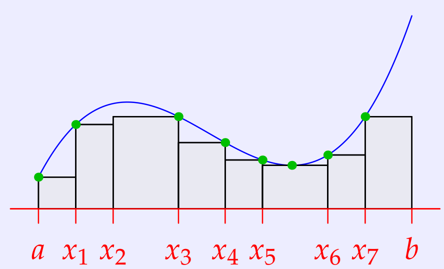
$$U(f) = \inf U(f, P) \qquad L(f) = \sup L(f, P)$$

where the supremum/infimum are taken over all partitions. Necessarily both integrals are *finite.*
We say that f is *(Riemann) integrable* on $[a, b]$ if $U(f) = L(f)$. We denote this value by

$$\int_a^b f \quad \text{or} \quad \int_a^b f(x)\, \mathrm{d}x$$



Upper Darboux sum $U(f, P)$



Lower Darboux sum $L(f, P)$

If the interval is understood or irrelevant, one often simply says that f is integrable and writes $\int f$.

Intuitively, $L(f, P)$ is the sum of the areas of rectangles built on P which just fit under the graph of f. It is also the infimum of all Riemann sums on P. If f is discontinuous, then $L(f, P)$ need not itself be a Riemann sum, as there might not exist suitable sample points!

Examples 4.4. 1. We revisit Example 4.1 in this language.

Given a partition $Q = \{x_0, \ldots, x_n\}$ of $[0,1]$ and sample points $x_i^* \in [x_{i-1}, x_i]$, we compute the Riemann sum for $f(x) = x^2$

$$\sum_{i=1}^{n} f(x_i^*)\,\Delta x_i = \sum_{i=1}^{n} (x_i^*)^2 (x_i - x_{i-1})$$

Since f is increasing, we have $x_{i-1}^2 \le (x_i^*)^2 \le x_i^2$ on each interval, whence

$$L(f,Q) = \sum_{i=1}^{n} (x_{i-1})^2 (x_i - x_{i-1}) \le \sum_{i=1}^{n} (x_i^*)^2 (x_i - x_{i-1}) \le \sum_{i=1}^{n} (x_i)^2 (x_i - x_{i-1}) = U(f,Q)$$

The Darboux sums are therefore the Riemann sums for left- and right-endpoints.

If we take Q_n to be the partition with subintervals of equal width $\Delta x = \frac{1}{n}$, then

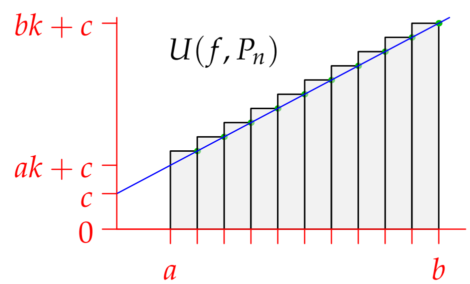
$$U(f) = \inf_P U(f,P) \le U(f,Q_n) = \sum_{i=1}^{n} \left(\frac{i}{n}\right)^2 \Delta x = R_n$$

is the right Riemann sum discussed originally. Similarly $L(f) \ge L_n$. Since L_n and R_n both converge to $\frac{1}{3}$ as $n \to \infty$, the squeeze theorem forces

$$L_n \le L(f) \le U(f) \le R_n \implies L(f) = U(f) = \frac{1}{3}$$

Otherwise said, f is integrable on $[0,1]$ with $\int_0^1 x^2\,dx = \frac{1}{3}$.

2. Suppose $f(x) = kx + c$ on $[a,b]$, and that $k > 0$. Take the evenly spaced partition P_n where $x_i = a + \frac{b-a}{n} i$. Since f is increasing, the upper Darboux sum is again the Riemann sum with right-endpoints:



$$\begin{aligned} U(f,P_n) = R_n &= \sum_{i=1}^{n} f(x_i)\Delta x \\ &= \frac{b-a}{n} \sum_{i=1}^{n} \frac{k(b-a)}{n} i + ak + c \\ &= \frac{b-a}{n} \left[\frac{k(b-a)}{n} \cdot \frac{1}{2} n(n+1) + (ak+c)n\right] \\ &\xrightarrow[n\to\infty]{} \frac{1}{2} k(b-a)^2 + (b-a)(ak+c) = \frac{k}{2}(b^2 - a^2) + c(b-a) \end{aligned}$$

Similarly, the lower Darboux sum is the Riemann sum with left-endpoints:

$$L(f,P_n) = L_n = \frac{b-a}{n} \left[\frac{k(b-a)}{n} \cdot \frac{1}{2} n(n-1) + (ak+c)n\right] \xrightarrow[n\to\infty]{} \frac{k}{2}(b^2 - a^2) + c(b-a)$$

As above, $L_n \le L(f) \le U(f) \le R_n$ and the squeeze theorem prove that f is integrable on $[a,b]$ with $\int_a^b f = \frac{k}{2}(b^2 - a^2) + c(b-a)$.

Now we have some examples, a few remarks are in order.

Riemann versus Darboux Definition 4.3 is really that of the *Darboux integral.* Here is Riemann's definition: $f : [a, b] \to \mathbb{R}$ being integrable with integral $\int_a^b f$ means

$$\forall \epsilon > 0,\ \exists \delta \text{ such that } (\forall P, x_i^*)\ \operatorname{mesh}(P) < \delta \implies \left| \sum_{i=1}^{n} f(x_i^*)\Delta x_i - \int_a^b f \right| < \epsilon$$

This is significantly more difficult to work with, though it can be shown to be equivalent to the Darboux integral. We won't pursue Riemann's formulation further, except to observe that *if* a function is integrable and $\operatorname{mesh}(P_n) \to 0$, then $\int_a^b f = \lim_{n\to\infty} \sum_{i=1}^n f(x_i^*)\Delta x_i$: this allows us to approximate integrals using any sample points we choose, hence why *right*-endpoints ($x_i^* = x_i$) are so common in Freshman calculus.

Monotone Functions Darboux sums are easy to compute for monotone functions. As in the examples, if f is increasing, then each $M_i = f(x_i)$, from which $U(f, P)$ is the Riemann sum with *right-endpoints.* Similarly, $L(f, P)$ is the Riemann sum with *left-endpoints.*

Area If f is positive and continuous,[19] the Riemann integral $\int_a^b f$ serves as a *definition* for the area under the curve $y = f(x)$. This should make intuitive sense:

1. In the second example where we have a straight line, we obtain the same value for the area by computing directly as the sum of a rectangle and a triangle!
2. For any partition P, the area under the curve should satisfy the inequalities

   $$L(f, P) \le \text{Area} \le U(f, P)$$

   But these are precisely the same inequalities satisfied by the integral itself!

   $$L(f, P) \le L(f) = \int_a^b f = U(f) \le U(f, P)$$

In the examples we exhibited a sequence of partitions (P_n) where $U(f, P_n)$ and $L(f, P_n)$ converged to the same limit. The remaining results in this section develop some basic properties of partitions and make this limiting process rigorous.

Definition 4.5. If $P \subseteq Q$ are both partitions of $[a, b]$, we call Q a *refinement* of P.

To refine a partition, we simply throw some more points in!

Lemma 4.6. *Suppose $f : [a, b] \to \mathbb{R}$ is bounded.*

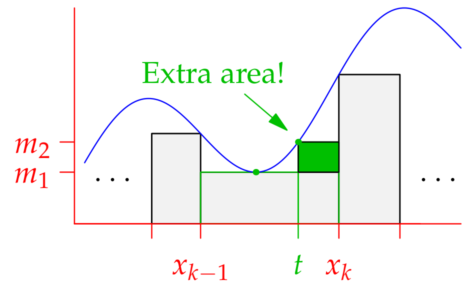
1. *If Q is a refinement of P (on $[a, b]$), then*

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

2. *For any partitions P, Q of $[a, b]$, we have $L(f, P) \le U(f, Q)$.*
3. $L(f) \le U(f)$

[19] We'll see in Theorem 4.17 that every continuous function is integrable.

Proof. 1. We prove inductively. Suppose first that $Q = P \cup \{t\}$ contains exactly one additional point $t \in (x_{k-1}, x_k)$. Write



$$m_1 = \inf\{f(x) : x \in [x_{k-1}, t]\}$$
$$m_2 = \inf\{f(x) : x \in [t, x_{k-1}]\}$$
$$m = \inf\{f(x) : x \in [x_{k-1}, x_k]\} = \min\{m_1, m_2\}$$

The Darboux sums $L(f, P)$ and $L(f, Q)$ are identical except for the terms involving t. This results in extra area:

$$L(f, Q) - L(f, P) = m_1(t - x_{k-1}) + m_2(x_k - t) - m(x_k - x_{k-1})$$
$$= (m_1 - m)(t - x_{k-1}) + (m_2 - m)(x_k - t) \geq 0$$

More generally, since a refinement Q is obtained by adding *finitely many* new points, induction tells us that $P \subseteq Q \implies L(f, P) \leq L(f, Q)$.

The argument for $U(f, Q) \leq U(f, P)$ is similar, and the middle inequality is trivial.

2. If P and Q are partitions, then $P \cup Q$ is a refinement of both P and Q. By part 1,

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

3. This is an exercise.

■

Theorem 4.7. *Suppose $f : [a, b] \to \mathbb{R}$ is bounded.*

1. *(Cauchy criterion) f is integrable $\iff \forall \epsilon > 0, \exists P$ such that $U(f, P) - L(f, P) < \epsilon$.*
2. *f is integrable $\iff \exists (P_n)_{n \in \mathbb{N}}$ such that $U(f, P_n) - L(f, P_n) \to 0$. In such a situation, both sequences $U(f, P_n)$ and $L(f, P_n)$ converge to $\int_a^b f$.*

Part 1 is termed a 'Cauchy' criterion since it doesn't mention the integral (limit).

Proof. We prove the Cauchy criterion, leaving part 2 as an exercise.

($\Rightarrow$) Suppose f is integrable and that $\epsilon > 0$ is given. Since $\inf U(f, Q) = \int f = \sup L(f, R)$, there exist partitions Q, R such that

$$U(f, Q) < \int f + \frac{\epsilon}{2} \quad \text{and} \quad L(f, R) > \int f - \frac{\epsilon}{2}$$

Let $P = Q \cup R$ and apply $(*)$: $L(f, R) \leq L(f, P) \leq U(f, P) \leq U(f, Q)$. But then

$$U(f, P) - L(f, P) \leq U(f, Q) - L(f, R) = U(f, Q) - \int f + \int f - L(f, R) < \epsilon$$

($\Leftarrow$) Assume the right hand side. For every partition, $L(f, P) \leq L(f) \leq U(f) \leq U(f, P)$. Thus

$$0 \leq U(f) - L(f) \leq U(f, P) - L(f, P) < \epsilon$$

Since this holds for all $\epsilon > 0$, we see that $U(f) = L(f)$: that is, f is integrable.

■

Examples 4.8. 1. Consider $f(x) = \sqrt{x}$ on the interval $[0, b]$. We choose a sequence of partitions (P_n) that evaluate nicely when fed to this function:

$$P_n = \{x_0, \ldots, x_n\} \quad \text{where} \quad x_i = \left(\frac{i}{n}\right)^2 b$$
$$\implies \Delta x_i = x_i - x_{i-1} = \frac{b}{n^2}\left(i^2 - (i-1)^2\right) = \frac{(2i-1)b}{n^2}$$

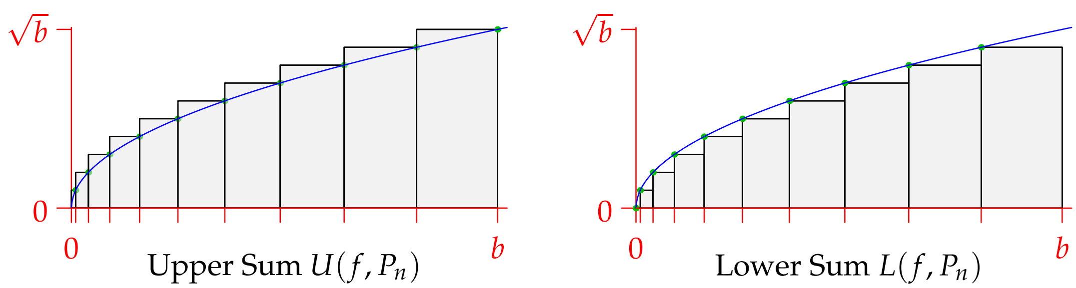
Since f is increasing on $[0, b]$, we see that

$$U(f, P_n) = \sum_{i=1}^{n} f(x_i)\Delta x_i = \sum_{i=1}^{n} \frac{i\sqrt{b}}{n} \cdot \frac{(2i-1)b}{n^2} = \frac{b^{3/2}}{n^3}\sum_{i=1}^{n} 2i^2 - i$$
$$= \frac{b^{3/2}}{n^3}\left[\frac{1}{3}n(n+1)(2n+1) - \frac{1}{2}n(n+1)\right] \xrightarrow[n\to\infty]{} \frac{2}{3}b^{3/2}$$

Similarly

$$L(f, P_n) = \sum_{i=1}^{n} f(x_{i-1})\Delta x_i = \sum_{i=1}^{n} \frac{(i-1)\sqrt{b}}{n} \cdot \frac{(2i-1)b}{n^2} = \frac{b^{3/2}}{n^3}\sum_{i=1}^{n} 2i^2 - 3i + 1$$
$$= \frac{b^{3/2}}{n^3}\left[\frac{1}{3}n(n+1)(2n+1) - \frac{3}{2}n(n+1) + n\right] \xrightarrow[n\to\infty]{} \frac{2}{3}b^{3/2}$$

Since the limits are equal, we conclude that f is integrable and $\int_0^b \sqrt{x}\, dx = \frac{2}{3}b^{3/2}$.



2. Here is the classic example of a *non-integrable function*. Let $f : [a, b] \to \mathbb{R}$ to be the indicator function of the irrational numbers,

$$f(x) = \begin{cases} 1 & \text{if } x \notin \mathbb{Q} \\ 0 & \text{if } x \in \mathbb{Q} \end{cases}$$

Suppose $P = \{x_0, \ldots, x_n\}$ is *any* partition of $[a, b]$. Since any interval of positive length contains both rational and irrational numbers, we see that

$$\sup\{f(x) : x \in [x_{i-1}, x_i]\} = 1 \implies U(f, P) = \sum_{i=1}^{n}(x_i - x_{i-1}) = b - a \implies U(f) = b - a$$
$$\inf\{f(x) : x \in [x_{i-1}, x_i]\} = 0 \implies L(f, P) = 0 \implies L(f) = 0$$

Since the upper and lower Darboux integrals differ, f is not (Riemann) integrable.

As any freshman calculus student can attest, if you can find an anti-derivative, then the fundamental theorem of calculus (Section 4.34) makes evaluating integrals far easier. For instance, you are probably desperate to write

$$\frac{\mathrm{d}}{\mathrm{d}x}\frac{2}{3}x^{3/2} = x^{1/2} \implies \int_0^b \sqrt{x}\,\mathrm{d}x = \frac{2}{3}x^{3/2}\Big|_0^b = \frac{2}{3}b^{3/2}$$

rather than computing Riemann/Darboux sums as in the previous example! However, in most practical situations, no easy-to-compute anti-derivative exists; the best we can do is to approximate using Riemann sums for progressively finer partitions. Thankfully computers excel at such tedious work!

Exercises 4.32. *Key concepts:* *Darboux sums/integrals,* *Partitions, sample points & refinements,*

Cauchy & sequential criteria for integrability

1. Use partitions to find the upper and lower Darboux integrals on the interval $[0, b]$. Hence prove that the function is integrable and compute its integral.

   (a) $f(x) = x^3$ (b) $g(x) = \sqrt[3]{x}$

2. Repeat question 1 for the following two functions. You cannot simply compute Riemann sums for left and right endpoints and take limits: why not?

   (a) $h(x) = x(2-x)$ on $[0, 2]$

   (*Hint: choose a partition with* $2n$ *points such that* $x_n = 1$ *and observe that* $h(2-x) = h(x)$)

   (b) On the interval $[0, 3]$, let $k(x) = \begin{cases} 2x & \text{if } x \le 1 \\ 5-x & \text{if } x > 1 \end{cases}$

   (*Hint: this time try a partition with* $3n$ *points*)

3. Let $f(x) = x$ for rational x and $f(x) = 0$ for irrational x. Calculate the upper and lower Darboux integrals for f on the interval $[0, b]$. Is f integrable on $[0, b]$?

4. Prove part 3 of Lemma 4.6: $L(f) \le U(f)$.

5. Prove part 2 of Theorem 4.7.

   $$f \text{ is integrable} \iff \exists (P_n)_{n\in\mathbb{N}} \text{ such that } \lim_{n\to\infty}\big(U(f, P_n) - L(f, P_n)\big) = 0$$

   Moreover, prove that both $U(f, P_n)$ and $L(f, P_n)$ converge to $\int f$.

6. (a) Reread Definition 4.3. What happens if we allow $f : [a, b] \to \mathbb{R}$ to be *unbounded*?

   (b) (Hard) Read "*Riemann versus Darboux*" on page 73. Explain why being *Riemann* integrable also forces f to be bounded.

   (c) (Hard) Explain the observation that $L(f, P)$ is the infimum of the set of all Riemann sums on P.

7. (If you like coding) Write a short program to estimate $\int_a^b f(x)\,\mathrm{d}x$ using Riemann sums. This can be very simple (equal partitions with right endpoints), or more complex (random partition and sample points given a mesh). Apply your program to estimate $\int_0^5 \sin(x^2 e^{-\sqrt{x}})\,\mathrm{d}x$.

4.33 Properties of the Riemann Integral

The rough take-away of this long section is that everything you think is integrable probably is! Examples will be few, since we have not established many explicit values for integrals.

Theorem 4.9 (Linearity). *If f, g are integrable and k, l are constant, then $kf + lg$ is integrable and*

$$\int kf + lg = k\int f + l\int g$$

Example 4.10. Thanks to examples in the previous section, we can now calculate, e.g.,

$$\int_0^2 5x^3 - 3\sqrt{x}\,\mathrm{d}x = 5\cdot\frac{1}{4}\cdot 2^4 - 3\cdot\frac{2}{3}\cdot 2^{3/2} = 20 - 4\sqrt{2}$$

Proof. Suppose $\epsilon > 0$ is given. By the Cauchy criterion (Theorem 4.7, part 1), there exist partitions R, S such that

$$U(f,R) - L(f,R) < \frac{\epsilon}{2} \quad \text{and} \quad U(g,S) - L(g,S) < \frac{\epsilon}{2}$$

If $P = R \cup S$, then both inequalities are satisfied by P (Lemma 4.6). On each subinterval,

$$\inf f(x) + \inf g(x) \le \inf\big(f(x) + g(x)\big) \quad \text{and} \quad \sup\big(f(x)+g(x)\big) \le \sup f(x) + \sup g(x)$$

since the individual suprema/infima could be 'evaluated' at different places. Thus

$$L(f,P) + L(g,P) \le L(f+g,P) \le U(f+g,P) \le U(f,P) + U(g,P)$$

whence $U(f+g,P) - L(f+g,P) < \epsilon$ and $f+g$ is integrable. Moreover,

$$\int (f+g) - \int f - \int g \le \Big(U(f,P) - \int f\Big) + \Big(U(g,P) - \int g\Big) < \epsilon$$

Using lower Darboux integrals similarly obtains the other half of the inequality

$$-\epsilon < \int (f+g) - \int f - \int g < \epsilon$$

Since this holds for all $\epsilon > 0$, we conclude that $\int (f+g) = \int f + \int g$.
That kf is integrable with $\int kf = k\int f$ is an exercise. Put these together for the result. ∎

Corollary 4.11 (Changing endvalues). *Suppose f is integrable on $[a,b]$ and $g : [a,b] \to \mathbb{R}$ satisfies $f(x) = g(x)$ on (a,b). Then g is also integrable on $[a,b]$ and $\int_a^b g = \int_a^b f$.*

Definition 4.12 (Integration on an open interval). A *bounded* function $g : (a,b) \to \mathbb{R}$ is *integrable* if it has an integrable extension $f : [a,b] \to \mathbb{R}$ where $f(x) = g(x)$ on (a,b). In such a case, we define $\int_a^b g := \int_a^b f$.

The Corollary (its proof is an exercise) shows that the choice of extension is irrelevant.

Theorem 4.13 (Basic integral comparisons). *Suppose f and g are integrable on $[a,b]$. Then:*

1. $f(x) \le g(x) \implies \int f \le \int g$
2. $m \le f(x) \le M \implies m(b-a) \le \int_a^b f \le M(b-a)$
3. fg *is integrable.*
4. $|f|$ *is integrable and* $\left|\int f\right| \le \int |f|$
5. $\max(f,g)$ *and* $\min(f,g)$ *are both integrable.*

Part 3 is *not* integration by parts since it doesn't tell us how $\int fg$ relates to $\int f$ and $\int g$!

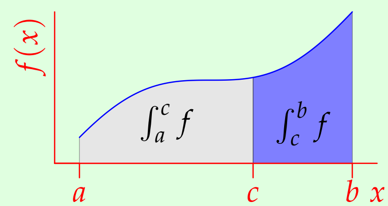
Proof.

1. Since $g-f$ is positive and integrable, $L(g-f,P) \ge 0$ for all partitions P. But then
$$0 \le \inf L(g-f,P) = L(g-f) = \int g-f = \int g - \int f$$
2. Apply part 1 twice.
3. This is an exercise.
4. The integrability is an exercise. For the comparison, apply part 1 to $-|f| \le f \le |f|$.
5. Use $\max(f,g) = \frac{1}{2}(f+g) + \frac{1}{2}|f-g|$, etc., together with the previous parts.

∎

Theorem 4.14 (Domain splitting). *Suppose $f : [a,b] \to \mathbb{R}$ and let $c \in (a,b)$. If f is integrable on both $[a,c]$ and $[c,b]$, then it is integrable on $[a,b]$ and*

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



In light of this result, it is conventional to allow integral limits to be reversed: if $a < b$, then

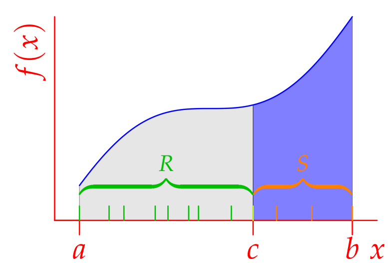
$$\int_b^a f := -\int_a^b f \quad \text{is consistent with} \quad \int_a^a f = 0$$

Proof. Let $\epsilon > 0$ be given, then $\exists R, S$ partitions of $[a,c]$, $[c,b]$ such that

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

Choose $P = R \cup S$ to partition $[a,b]$, then

$$U(f,P) - L(f,P) = U(f,R) + U(f,S) - L(f,R) - L(f,S) < \epsilon$$



Moreover

$$\int_a^b f - \int_a^c f - \int_c^b f \le U(f,P) - L(f,R) - L(f,S) = U(f,P) - L(f,P) < \epsilon$$

Showing that this expression is greater than $-\epsilon$ is similar.

∎

Example 4.15. If $f(x) = \sqrt{x}$ on $[0,1]$ and $f(x) = 1$ on $[1,2]$, then

$$\int_0^2 f = \int_0^1 \sqrt{x}\,dx + \int_1^2 1\,dx = \frac{2}{3} + 1 = \frac{5}{3}$$

Monotonic & Continuous Functions We establish the integrability of two large classes of functions.

Definition 4.16. A function $f : [a,b] \to \mathbb{R}$ is:

Monotonic if it is either *increasing* ($x < y \implies f(x) \le f(y)$) or *decreasing*.

Piecewise monotonic if there is a partition $P = \{x_0, \ldots, x_n\}$ (finite!) of $[a,b]$ such that f is monotonic on each open subinterval (x_{k-1}, x_k).

Piecewise continuous if there is a partition such that f is *uniformly continuous* on each (x_{k-1}, x_k).

Theorem 4.17. *If f is* monotonic *or* continuous *on $[a,b]$, then it is integrable.*

Examples 4.18. 1. Since sine is continuous, we can approximate via a sequence of Riemann sums

$$\int_0^\pi \sin x\,dx = \frac{\pi}{n} \lim_{n\to\infty} \sum_{i=1}^n \sin\frac{\pi i}{n}$$

Evaluating this limit is another matter entirely, one best handled in the next section...

2. Similarly, $e^{\sqrt{x}}$ is integrable and therefore may be approximated via Riemann sums:

$$\int_0^1 e^{\sqrt{x}}\,dx = \frac{1}{n} \lim_{n\to\infty} \sum_{i=1}^n \exp\sqrt{\frac{i}{n}} = \lim_{n\to\infty} \sum_{j=1}^n \frac{2j-1}{n} \exp\frac{j}{n}$$

Both sums use right endpoints: the first has equal subintervals, while the second is analogous to Example 4.8.1. These limits would typically be estimated using a computer.

Proof. Since $[a,b]$ is closed and bounded, a continuous function f is *uniformly* so. Let $\epsilon > 0$ be given:

$$\exists \delta > 0 \text{ such that } \forall x, y \in [a,b],\ |x-y| < \delta \implies |f(x) - f(y)| < \frac{\epsilon}{b-a}$$

Let P be a partition with mesh $P < \delta$. Since f attains its bounds on each $[x_{i-1}, x_i]$,

$$\exists x_i^*, y_i^* \in [x_{i-1}, x_i] \quad \text{such that} \quad M_i - m_i = f(x_i^*) - f(y_i^*) < \frac{\epsilon}{b-a}$$

from which

$$U(f,P) - L(f,P) < \sum_{i=1}^n \frac{\epsilon}{b-a}(x_i - x_{i-1}) = \epsilon$$

The monotonicity argument is an exercise. ∎

Combining the proof with Definition 4.12: every *uniformly continuous* $f : (a,b) \to \mathbb{R}$ is integrable.

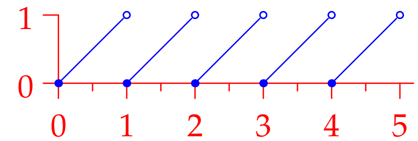
Corollary 4.19. *Piecewise continuous and **bounded** piecewise monotonic functions are integrable.*

Proof. If f is piecewise continuous, then the restriction of f to (x_{k-1}, x_k) has a continuous extension $g_k : [x_{k-1}, x_k] \to \mathbb{R}$; this is integrable by Theorem 4.17. By Corollary 4.11, f is integrable on $[x_{k-1}, x_k]$ with $\int_{x_{k-1}}^{x_k} f = \int_{x_{k-1}}^{x_k} g_k$. Theorem 4.14 ($n-1$ times!) finishes things off:

$$\int_a^b f = \sum_{k=1}^{n} \int_{x_{k-1}}^{x_k} f$$

The argument for piecewise monotonicity is similar. ∎

Example 4.20. The 'fractional part' function $f(x) = x - \lfloor x \rfloor$ is both piecewise continuous and piecewise monotone on any bounded interval. It is therefore integrable on any such interval.



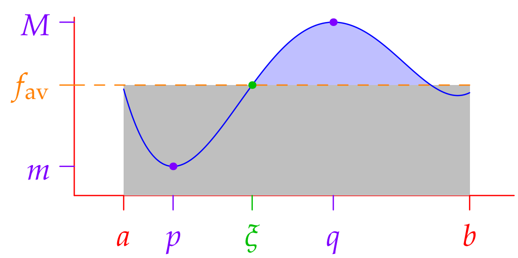
For a final corollary, here is one more incarnation of the intermediate value theorem.

Corollary 4.21 (IVT for integrals). *If f is continuous on $[a, b]$, then $\exists \xi \in (a, b)$ for which*

$$f(\xi) = \frac{1}{b-a} \int_a^b f$$

Proof. Since f is continuous, it is integrable on $[a, b]$. By the extreme value theorem it is also bounded and attains its bounds: $\exists p, q \in [a, b]$ such that

$$f(p) := \inf_{x \in [a,b]} f(x), \qquad f(q) = \sup_{x \in [a,b]} f(x)$$



Applying Theorem 4.13, part 2, with $m = f(p)$ and $M = f(q)$, we see that

$$(b-a)f(p) \le \int_a^b f \le (b-a)f(q)$$

Divide by $b - a$ and apply the usual intermediate value theorem for f to see that the required ξ exists between p and q. ∎

In the picture, when f is positive and continuous, the grey area equals that under the curve; imagine levelling off the blue hill with a bulldozer… The notation $f_{\text{av}} = \frac{1}{b-a}\int_a^b f$ indicates the average value of f on $[a, b]$: to see why this interpretation is sensible, take a sequence of Riemann sums on equally-spaced partitions P_n to see that

$$\frac{1}{b-a} \int_a^b f = \lim_{n\to\infty} \sum_{i=1}^{n} f(x_i^*)\Delta x = \lim_{n\to\infty} \frac{f(x_1^*) + \cdots + f(x_n^*)}{n}$$

is the limit of a sequence of *averages* of equally-spaced samples $f(x_i^*)$.

What can/cannot be integrated?

We now know a great many examples of integrable functions:

- Piecewise continuous & monotonic functions are integrable.
- Linear combinations, products, absolute values, maximums and minimums of (already) integrable functions.

By contrast, we've only seen one non-integrable function (Example 4.8.2). After so many positive integrability conditions, it is reasonable to ask precisely which functions are Riemann integrable. Here is the answer, though it is quite tricky to understand.

Theorem 4.22 (Lebesgue). *Suppose $f : [a, b] \to \mathbb{R}$ is bounded. Then*

$$f \text{ is Riemann integrable} \iff \text{it is continuous except on a set of measure zero}$$

Naïvely, the *measure* of a set is the sum of the lengths of its maximal subintervals, though unfortunately this doesn't make for a very useful definition.[20] Any countable subset has measure zero, so Lebesgue's result is almost as if we can extend Corollary 4.19 to allow for infinite sums. For instance, Exercise 1.17.8 describes a function which is continuous only on the irrationals: it is thus Riemann integrable (indeed $\int_a^b f = 0$ for any $a < b$). There are also uncountable sets with measure zero such as Cantor's middle-third set $\mathcal{C}$: the function

$$f(x) = \begin{cases} 1 & \text{if } x \in \mathcal{C} \\ 0 & \text{otherwise} \end{cases}$$

is continuous except on $\mathcal{C}$ and therefore Riemann integrable; again $\int_0^1 f(x)\,dx = 0$.

Exercises 4.33. *Key concepts: Linear combinations, products, etc., of integrable functions are integrable, Continuous and monotone functions are integrable, Integrability on open intervals*

1. Explain why $\int_0^{2\pi} x^2 \sin^8(e^x)\,dx \le \frac{8}{3}\pi^3$
2. If f is integrable on $[a, b]$ prove that it is integrable on any interval $[c, d] \subseteq [a, b]$.
3. We complete the proof of Theorem 4.9 (linearity of integration).
   (a) Suppose $k > 0$, let $A \subseteq \mathbb{R}$ and define $kA := \{kx : x \in A\}$. Prove that $\sup kA = k \sup A$ and $\inf kA = k \inf A$.
   (b) If $k > 0$ prove that kf is integrable on any interval and that $\int kf = k \int f$.
   (c) How should you modify your argument if $k < 0$?

[20] Formally, the *length* of an open interval (a, b) is $b - a$ and a set $A \subseteq \mathbb{R}$ has *measure zero* if

$$\forall \epsilon > 0,\ \exists \text{ open intervals } I_n \text{ such that } A \subseteq \bigcup_{n=1}^{\infty} I_n \text{ and } \sum_{i=1}^{\infty} \text{length}(I_n) < \epsilon$$

More generally, the *Lebesgue measure* of a set (subject to a technical condition) is the infimum of the sum of the lengths of any countable collection of open covering intervals. *Measure theory* is properly a matter for graduate study. Surprisingly, there exist sets with positive measure that contain no subintervals, and even sets which are non-measurable!

4. Give an example of an integrable but *discontinuous* function on a closed bounded interval $[a, b]$ for which the conclusion of the Intermediate Value Theorem for Integrals is *false*.

5. Use Darboux sums to compute the value of the integral $\int_{1/2}^{15/2} x - \lfloor x \rfloor \, dx$ (Example 4.20).

6. We prove and extend Corollary 4.11. Suppose f is integrable on $[a, b]$.

   (a) If $g : [a, b] \to \mathbb{R}$ satisfies $f(x) = g(x)$ for all $x \in (a, b)$, prove that g is integrable and $\int_a^b g = \int_a^b f$.
   (*Hint: consider $h = f - g$ and show that $\int h = 0$*)

   (b) Now suppose $g : [a, b] \to \mathbb{R}$ satisfies $f(x) = g(x)$ for all $x \in [a, b]$ except at finitely many points. Prove that g is integrable and $\int_a^b g = \int_a^b f$.

7. Show that an increasing function on $[a, b]$ is integrable and thus complete Theorem 4.17.
   (*Hint: Choose a partition with* mesh $P < \frac{\epsilon}{f(b) - f(a)}$)

8. Suppose f and g are integrable on $[a, b]$.

   (a) Define $h(x) = \big(f(x)\big)^2$. We know:
   - f is bounded: $\exists K$ such that $|f(x)| \leq K$ on $[a, b]$.
   - Given $\epsilon > 0$, $\exists P$ such that $U(f, P) - L(f, P) < \frac{\epsilon}{2K}$. For each subinterval $[x_{i-1}, x_i]$, let

   $$M_i = \sup f(x), \qquad m_i = \inf f(x), \qquad \overline{M}_i = \sup h(x), \qquad \overline{m}_i = \inf h(x)$$

   Prove that $\overline{M}_i - \overline{m}_i \leq 2(M_i - m_i)K$. Hence conclude that h is integrable.

   (b) Prove that fg is integrable.
   (*Hint: $fg = \frac{1}{4}(f+g)^2 - \frac{1}{4}(f-g)^2$*)

   (c) Prove that $U(|f|, P) - L(|f|, P) \leq U(f, P) - L(f, P)$ for any partition P. Hence conclude that $|f|$ is integrable.

   (*One can extend these arguments to show that if j is continuous, then $j \circ f$ is integrable. Parts (a) and (c) correspond, respectively, to $j(x) = x^2$ and $j(x) = |x|$.*)

9. (Hard) Let $f(x) = \begin{cases} x & \text{if } x \neq 0 \text{ and } \sin \frac{1}{x} > 0 \\ -x & \text{if } x \neq 0 \text{ and } \sin \frac{1}{x} < 0 \\ 0 & \text{if } x = 0 \end{cases}$

   (a) Show that f is not piecewise continuous on $[0, 1]$.

   (b) Show that f is not piecewise monotonic on $[0, 1]$.

   (c) Show that f is integrable on $[0, 1]$.
   (*Hint: given ϵ, hunt for a suitable partition to make $U(f, P) - L(f, P) < \epsilon$ by considering $[0, x_1]$ differently to the other subintervals*)

   (d) Make a similar argument which proves that $g = \sin \frac{1}{x}$ is integrable on $(0, 1]$.
   (*Hint: Show that g has an integrable extension on $[0, 1]$*)

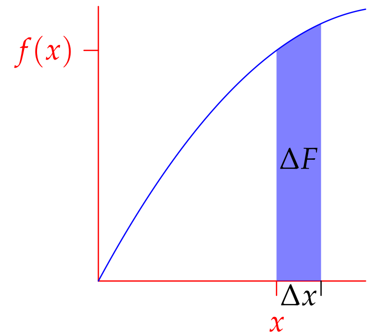
4.34 The Fundamental Theorem of Calculus

The key result linking integration and differentiation is usually presented in two parts. While there are significant subtleties, the rough statements are as follows (we follow the traditional numbering):

Part I Differentiation reverses integration: $\frac{\mathrm{d}}{\mathrm{d}x}\int_a^x f(t)\,\mathrm{d}t = f(x)$

Part II Integration reverses differentiation: $\int_a^b F'(x)\,\mathrm{d}x = F(b) - F(a)$

These facts seemed intuitively obvious to early practitioners of calculus. Given a continuous positive function f:



- Let $F(x)$ denote the area under $y = f(x)$ between 0 and x.
- A small increase Δx results in the area increasing by ΔF.
- $\Delta F \approx f(x)\Delta x$ is approximately the area of a rectangle, whence $\frac{\Delta F}{\Delta x} \approx f(x)$. This is part I.
- $F(b) - F(a) \approx \sum \Delta F_i \approx \sum f(x_i)\Delta x_i$. Since $F' = f$, this is part II.

When Leibniz introduced the symbols $\int$ and d in the late 1600s, it was partly to reflect the fundamental theorem.[21] If you're happy with non-rigorous notions of limit, rate of change, area, and (infinite) sums, the above is all you need!

Of course we are very much concerned with the details: What must we assume about f and F, and how are these properties used in the proof?

Theorem 4.23 (FTC, part I). *Suppose f is integrable on $[a, b]$. For any $x \in [a, b]$, define*

$$F(x) := \int_a^x f(t)\,\mathrm{d}t$$

Then:

1. *F is uniformly continuous on $[a, b]$;*
2. *If f is continuous at $c \in [a, b]$, then F is differentiable[22] at c with $F'(c) = f(c)$.*

Compare this with the naïve version above where we assumed f was continuous. We now require only the *integrability* of f, and its continuity at *one point* for the full result.

[21] $\int$ is a stylized S for *sum*, while d stands for *difference*. Given a sequence $F = (F_0, F_1, F_2, \ldots, F_n)$, construct a new sequence of *differences*

$$\mathrm{d}F = (F_1 - F_0, F_2 - F_1, \ldots, F_n - F_{n-1})$$

which can then be summed:

$$\int \mathrm{d}F = (F_1 - F_0) + (F_2 - F_1) + \cdots (F_n - F_{n-1}) = F_n - F_0 \qquad (*)$$

Viewing a function as an 'infinite sequence' of values spaced along an interval, $\mathrm{d}F$ becomes a sequence of *infinitesimals* and $(*)$ is essentially the fundamental theorem: $\int \mathrm{d}F = F(b) - F(a)$. It is the concept of **function** that is suspect here, not the essential relationship between sums and differences.

[22] Strictly: if $c = a$, then F is *right*-differentiable, etc.

Examples 4.24. Examples in every elementary calculus course.

1. Since $f(x) = \sin^2(x^3 - 7)$ is continuous on any bounded interval, we conclude that

$$\frac{\mathrm{d}}{\mathrm{d}x}\int_4^x \sin^2(t^3-7)\,\mathrm{d}t = \sin^2(x^3-7)$$

If one follows Theorem 4.14 and its conventions, then this is valid for all $x \in \mathbb{R}$.

2. The chain rule permits more complicated examples. For instance: $f(t) = \sin\sqrt{t}$ is continuous on its domain $[0, \infty)$ and $y(x) = x^2 + 3$ has range $[3, \infty) \subseteq \operatorname{dom}(f)$, whence

$$\frac{\mathrm{d}}{\mathrm{d}x}\int_0^{x^2+3} \sin\sqrt{t}\,\mathrm{d}t = \frac{\mathrm{d}y}{\mathrm{d}x}\frac{\mathrm{d}}{\mathrm{d}y}\int_0^y \sin\sqrt{t}\,\mathrm{d}t = 2x\sin\sqrt{x^2+3}$$

3. For a final positive example, we consider when

$$\frac{\mathrm{d}}{\mathrm{d}x}\int_{\sin x}^{e^x} \tan(t^2)\,\mathrm{d}t = e^x\tan(e^{2x}) - \cos x\tan(\sin^2 x)$$

Makes sense. To evaluate this, first choose any constant a and write

$$\int_{\sin x}^{e^x} = \int_a^{e^x} + \int_{\sin x}^a = \int_a^{e^x} - \int_a^{\sin x}$$

before differentiating. This is valid provided $\sin x$, e^x and a all lie in the same subinterval of

$$\operatorname{dom}\tan(t^2) = \mathbb{R} \setminus \{\pm\sqrt{\tfrac{\pi}{2}}, \pm\sqrt{\tfrac{3\pi}{2}}, \pm\sqrt{\tfrac{5\pi}{2}}, \ldots\}$$

Since $|\sin x| \le 1 < \sqrt{\frac{\pi}{2}}$, this requires

$$\left|e^{2x}\right| < \frac{\pi}{2} \iff x < \frac{1}{2}\ln\frac{\pi}{2}$$

Choosing $a = 1$ would certainly suffice.

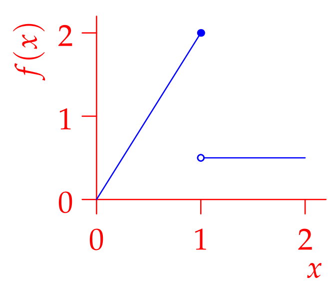
4. Now consider why the theorem requires continuity. The piecewise continuous function

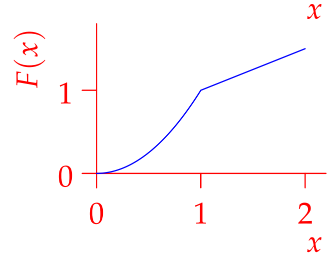
$$f : [0,2] \to \mathbb{R} : x \mapsto \begin{cases} 2x & \text{if } x \le 1 \\ \frac{1}{2} & \text{if } x > 1 \end{cases}$$

has a jump discontinuity at $x = 1$. We can still compute

$$F(x) = \begin{cases} \int_0^x 2t\,\mathrm{d}t = x^2 & \text{if } x \le 1 \\ \int_0^1 2t\,\mathrm{d}t + \int_1^x \frac{1}{2}\mathrm{d}t = \frac{1}{2}(x+1) & \text{if } x > 1 \end{cases}$$

This is continuous, indeed uniformly so! However the discontinuity of f results in F having a *corner* and thus being *non-differentiable* at $x = 1$. Indeed $F'(x) = f(x)$ whenever $x \neq 1$: that is, at all values of x where f is continuous.





Proving FTC I Neither half of the theorem is particularly difficult once you write down what you know and what you need to prove. Here are the key ingredients:

1. Uniform continuity for F means we must control the size of

$$|F(y) - F(x)| = \left| \int_a^y f(t)\,\mathrm{d}t - \int_a^x f(t)\,\mathrm{d}t \right| = \left| \int_x^y f(t)\,\mathrm{d}t \right| \le \int_x^y |f(t)|\,\mathrm{d}t$$

But the boundedness of f allows us to control this last integral...

2. $F'(c) = f(c)$ means showing that $\lim_{x \to c} \frac{F(x)-F(c)}{x-c} = f(c)$, which means controlling the size of

$$\left| \frac{F(x) - F(c)}{x - c} - f(c) \right| = \left| \frac{1}{x-c} \int_c^x f(t)\,\mathrm{d}t - f(c) \right|$$

The trick here will is to bring the *constant* $f(c)$ inside the integral as $\frac{1}{x-c} \int_c^x f(c)\,\mathrm{d}t$ so that the above becomes $\frac{1}{|x-c|} \int_c^x |f(t) - f(c)|\,\mathrm{d}t$. This may now be controlled via the continuity of f...

Proof. 1. Since f is integrable, it is bounded: $\exists M > 0$ such that $|f(x)| \le M$ for all x.

Let $\epsilon > 0$ be given and define $\delta = \frac{\epsilon}{M}$. Then, for any $x, y \in [a, b]$,

$$\begin{aligned} 0 < y - x < \delta \implies |F(y) - F(x)| &= \left| \int_x^y f(t)\,\mathrm{d}t \right| \le \int_x^y |f(t)|\,\mathrm{d}t && \text{(Theorem 4.13, part 4)} \\ &\le M(y - x) && \text{(Theorem 4.13, part 2)} \\ &< M\delta = \epsilon \end{aligned}$$

We conclude that F is uniformly continuous on $[a, b]$.

2. Let $\epsilon > 0$ be given. Since f is continuous at c, $\exists \delta > 0$ such that, for all $t \in [a, b]$,

$$|t - c| < \delta \implies |f(t) - f(c)| < \frac{\epsilon}{2}$$

Now for all $x \in [a, b]$ (except c),

$$\begin{aligned} 0 < |x - c| < \delta \implies \left| \frac{F(x) - F(c)}{x - c} - f(c) \right| &= \left| \frac{1}{x-c} \int_c^x f(t) - f(c)\,\mathrm{d}t \right| && \text{(Theorem 4.9)} \\ &\le \frac{1}{|x-c|} \int_c^x |f(t) - f(c)|\,\mathrm{d}t && \text{(Theorem 4.13)} \\ &\le \frac{1}{|x-c|} \frac{\epsilon}{2} |x - c| = \frac{\epsilon}{2} < \epsilon \end{aligned}$$

Clearly $\lim_{x \to c} \frac{F(x)-F(c)}{x-c} = f(c)$. Otherwise said, F is differentiable at c with $F'(c) = f(c)$. ∎

The Fundamental Theorem, part II As with part I, the *formulaic* part of the result should be familiar, though we are more interested in the assumptions and where they are needed.

Theorem 4.25 (FTC, part II). *Suppose g is continuous on $[a, b]$, differentiable on (a, b), and moreover that g' is integrable on (a, b) (recall Definition 4.12). Then,*

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

Part II is often expressed in terms of *anti-derivatives*: F being an anti-derivative of f if $F' = f$. Combined with FTC, part I, we recover the familiar '$+c$' result and a simpler version of the fundamental theorem often seen in elementary calculus.

Corollary 4.26. *Let f be continuous on $[a, b]$.*

- *If F is an anti-derivative of f, then $\int_a^b f = F(b) - F(a)$.*
- *Every anti-derivative of f has the form $F(x) = \int_a^x f(t)\,dt + c$ for some constant c.*

Examples 4.27. Again, basic examples should be familiar.

1. Plainly $g(x) = x^2 + 2x^{3/2}$ is continuous on $[1, 4]$ and differentiable on $(1, 4)$ with derivative $g'(x) = 2x + 3\sqrt{x}$; this last is continuous (and thus integrable) on $(1, 4)$. We conclude that

$$\int_1^4 2x + 3\sqrt{x}\,dx = x^2 + 2x^{3/2}\Big|_1^4 = (16 + 16) - (1 + 2) = 29$$

2. If $g(x) = \sin(3x^2)$, then $g'(x) = 6x\cos(3x^2)$. Certainly g satisfies the hypotheses of the theorem on any bounded interval $[a, b]$. We conclude

$$\int_a^b 6x\cos(3x^2)\,dx = \sin(3b^2) - \sin(3a^2)$$

   Moreover, every anti-derivative of $f(x) = 6x\cos(3x^2)$ has the form $F(x) = \sin(3x^2) + c$.

3. Recall Example 4.24.4 where the discontinuity of f at $x = 1$ led to the *non-differentiability* of $F(x) = \int_0^x f(t)\,dt$. The function F therefore fails the *hypotheses* of FTC II on the interval $[0, 2]$.

   It almost, however, satisfies the *conclusions* of FTC II, though this is somewhat tautological given the definition of F: except at $x = 1$, F is certainly an anti-derivative of f, and moreover $\int_0^2 f(x)\,dx = F(2) - F(0)$.

   In case you're worried that this makes the theorem trivial, note that other anti-derivatives $\hat{F}$ of f exist (except at $x = 1$) which fail to satisfy the conclusion. For instance

$$\hat{F}(x) = \begin{cases} x^2 & \text{if } x < 1 \\ \frac{1}{2}x & \text{if } x > 1 \end{cases} \implies \hat{F}(2) - \hat{F}(0) = 1 \neq \frac{3}{2} = \int_0^2 f(x)\,dx$$

Proving FTC II Exercise 10 offers a relatively easy proof when $g' = f$ is continuous. For the real McCoy, we can only rely on the *integrability* of g': the trick is to use the mean value theorem to write $g(b) - g(a)$ as a Riemann sum over a suitable partition.

Proof. Suppose $\epsilon > 0$ is given. Since g' is integrable, we may choose some partition P satisfying $U(g', P) - L(g', P) < \epsilon$. Since g satisfies the mean value theorem on each subinterval,

$$\exists \xi_i \in (x_{i-1}, x_i) \quad \text{such that} \quad g'(\xi_i) = \frac{g(x_i) - g(x_{i-1})}{x_i - x_{i-1}}$$

from which

$$g(b) - g(a) = \sum_{i=1}^{n} g(x_i) - g(x_{i-1}) = \sum_{i=1}^{n} g'(\xi_i)(x_i - x_{i-1})$$

This is a Riemann sum for g' associated to the partition P. Since the upper and lower Darboux sums are the supremum and infimum of these, we see that

$$L(g', P) \leq g(b) - g(a) \leq U(g', P)$$

However $\int_a^b g'$ satisfies the same inequality: $L(g', P) \leq \int_a^b g' \leq U(g', P)$. Since these inequalities hold for all $\epsilon > 0$, we conclude that $\int_a^b g' = g(b) - g(a)$. ∎

While we certainly used the integrability of g' in the proof, it might seem strange that we assumed it at all: shouldn't every derivative be integrable? Perhaps surprisingly, the answer is no! If you want a challenge, look up the *Volterra function*, which is differentiable everywhere but whose derivative is non-integrable!

The Rules of Integration

If one wants to *evaluate* an integral, rather than merely show it exists, there are really only two options:

1. Evaluate Riemann sums and take limits. This is often difficult if not impossible to do explicitly.
2. Use FTC II. The problem now becomes the finding of *anti-derivatives*, for which the core method is essentially *guess and differentiate.* To obtain general rules, we can attempt to reverse the rules of differentiation.

Integration by Parts Recall the *product rule*: the product $g = uv$ of two differentiable functions is differentiable with $g' = u'v + uv'$. Now apply Theorems 4.9, 4.13 and FTC II.

Corollary 4.28 (Integration by Parts). *Suppose u, v are continuous on $[a, b]$, differentiable on (a, b), and that u', v' are integrable on (a, b). Then*

$$\int_a^b u'(x)v(x)\,\mathrm{d}x = u(b)v(b) - u(a)v(a) - \int_a^b u(x)v'(x)\,\mathrm{d}x$$

This is significantly less useful than the product rule since it merely transforms the integral of one product into the integral of another.

Examples 4.29. With practice, there is no need to explicitly state u and v.

1. Let $u(x) = x$ and $v'(x) = \cos x$. Then $u'(x) = 1$ and $v(x) = \sin x$. These certainly satisfy the hypotheses. We conclude

$$\begin{aligned}\int_0^{\pi/2} x\cos x\,\mathrm{d}x &= [x\sin x]_0^{\pi/2} - \int_0^{\pi/2} \sin x\,\mathrm{d}x = \frac{\pi}{2}\sin\frac{\pi}{2} - 0 - [-\cos x]_0^{\pi/2} \\ &= \frac{\pi}{2} + \cos\frac{\pi}{2} - \cos 0 = \frac{\pi}{2} - 1\end{aligned}$$

2. Let $u(x) = \ln x$ and $v'(x) = 1$. Then $u'(x) = \frac{1}{x}$ and $v(x) = x$, whence

$$\begin{aligned}\int_e^{e^2} \ln x\,\mathrm{d}x &= [x\ln x]_e^{e^2} - \int_e^{e^2} \frac{x}{x}\,\mathrm{d}x = e^2\ln e^2 - e\ln e - [x]_e^{e^2} \\ &= 2e^2 - e - e^2 + e = e^2\end{aligned}$$

Change of Variables/Substitution We now turn our attention to the *chain rule*. If $g(x) = F(u(x))$, where F and u are differentiable, then g is differentiable with

$$g'(x) = \frac{\mathrm{d}g}{\mathrm{d}x} = \frac{\mathrm{d}F}{\mathrm{d}u}\frac{\mathrm{d}u}{\mathrm{d}x} = F'(u(x))u'(x)$$

Now integrate both sides; the only issue is what assumptions are needed to invoke FTC II.

Theorem 4.30 (Substitution Rule). *Suppose $u : [a,b] \to \mathbb{R}$ and $f : \text{range}(u) \to \mathbb{R}$ are continuous. Suppose also that u is differentiable on (a,b) with integrable derivative u'. Then*

$$\int_a^b f(u(x))\,u'(x)\,\mathrm{d}x = \int_{u(a)}^{u(b)} f(u)\,\mathrm{d}u$$

This is the famous 'u-sub'/change-of-variables formula from elementary calculus.

Proof. We leave as an exercise the verification that both integrals exist. By the intermediate and extreme value theorems, range(u) is a closed bounded interval. Assume range(u) has positive length for otherwise both integrals are trivially zero.

Choose any $c \in \text{range}(u)$ and define

$$F : \text{range}(u) \to \mathbb{R} \text{ by } F(v) := \int_c^v f(t)\,\mathrm{d}t$$

Since f is continuous, by FTC I says that F is differentiable with $F'(u) = f(u)$. But now

$$\begin{aligned}\int_a^b f(u(x))u'(x)\,\mathrm{d}x &= \int_a^b \left[\frac{\mathrm{d}}{\mathrm{d}x}F(u(x))\right]\mathrm{d}x && \text{(chain rule)} \\ &= F(u(b)) - F(u(a)) && \text{(FTC II)} \\ &= \int_{u(a)}^{u(b)} f(u)\,\mathrm{d}u\end{aligned}$$

■

Examples 4.31. Successfully applying the substitution rule can require significant creativity.[23]

1. To evaluate $\int_0^{\sqrt{\pi}} 2x\sin x^2\,dx$, we consider the substitution $u(x) = x^2$ defined on $[0, \sqrt{\pi}]$. Certainly u is continuous; moreover its derivative $u'(x) = 2x$ is integrable on $(0, \sqrt{\pi})$. Finally $f(u) = \sin u$ is continuous on $\text{range}(u) = [0, \pi]$. The hypotheses are satisfied, whence
$$\begin{aligned}\int_0^{\sqrt{\pi}} 2x\sin x^2\,dx &= \int_0^{\sqrt{\pi}} f\big(u(x)\big)u'(x)\,dx = \int_{u(0)}^{u(\pi)} f(u)\,du = \int_0^{\pi} \sin u\,du \\ &= -\cos u\Big|_0^{\pi} = 2\end{aligned}$$

2. For the following integral, a simple factorization suggests the substitution $u(x) = x^2 - 2$. Plainly $u : [\sqrt{2}, \sqrt{3}] \to [0,1]$ and $u'(x) = 2x$ is integrable. Moreover, $f(u) = \frac{1}{u^2+1}$ is continuous on $\text{range}(u) = [0,1]$. We conclude
$$\int_{\sqrt{2}}^{\sqrt{3}} \frac{2x}{x^4-4x^2+5}\,dx = \int_{\sqrt{2}}^{\sqrt{3}} \frac{2x}{(x^2-2)^2+1}\,dx = \int_0^1 \frac{1}{u^2+1}\,du = \arctan u\Big|_0^1 = \frac{\pi}{4}$$

3. The hypotheses on u really are all that's necessary. In particular, u need not be left-/right-differentiable at the endpoints of $[a,b]$. For instance, with $f(u) = u^2$ and $u(x) = \sqrt{x}$ on $[0,4]$, we easily verify
$$\frac{8}{3} = \int_0^4 \frac{1}{2}\sqrt{x}\,dx = \int_0^4 \frac{x}{2\sqrt{x}}\,dx = \int_0^4 f\big(u(x)\big)u'(x)\,dx = \int_0^2 f(u)\,du = \int_0^2 u^2\,du = \frac{8}{3}$$

4. Sloppy 'substitutions' might lead to utter nonsense. For instance, $u(x) = x^2$ suggests
$$\int_{-1}^2 \frac{1}{x}\,dx = \int_{-1}^2 \frac{1}{2x^2}2x\,dx = \int_1^4 \frac{1}{2u}\,du = \frac{1}{2}(\ln 4 - \ln 1) = \ln 2$$
This is total gibberish: the first integral does not exist since $\frac{1}{x}$ is undefined at $0 \in (-1,2)$. Thankfully, the hypotheses of the substitution rule prevent this: $f(u) = \frac{1}{2u}$ is not continuous on $\text{range}(u) = [0,4]$.

   While you are very unlikely to make precisely this mistake, the risk is real in more complicated or abstract situations…

---

[23]Hence the old adage, "Differentiation is a science, whereas integration is an art." To illustrate by example, consider $f(x) = \tan(e^x\cos(3x^2) + 4x^3)$. The derivative is easily found using the product and chain rules:
$$\frac{df}{dx} = \frac{1}{1+(e^x\cos(3x^2)+4x^3)^2}\left(e^x\cos(3x^2) - 6xe^x\sin(3x^2) + 12x^2\right)$$
By contrast, if you want to find an *explicit* anti-derivative of $f(x)$, the integration analogues (parts/substitution) are essentially useless. Similarly, the integral
$$\int_0^1 \tan(e^x\cos(3x^2)+4x^3)\,dx$$
is likely impossible to evaluate explicitly and can only be approximated, say by using Riemann sums.

Exercises 4.34. *Key concepts: Complete statements of FTC parts I & II, Integration by Parts/Substitution*

1. Calculate the following limits:

   (a) $\lim_{x\to 0} \frac{1}{x}\int_0^x e^{t^2}\,\mathrm{d}t$ (b) $\lim_{h\to 0} \frac{1}{h}\int_3^{3+h} e^{t^2}\,\mathrm{d}t$

2. Let $f(t) = \begin{cases} 0 & \text{if } t < 0 \\ t & \text{if } 0 \le t \le 1 \\ 4 & \text{if } t > 1 \end{cases}$

   (a) Determine the function $F(x) = \int_0^x f(t)\,\mathrm{d}t$ and sketch it. Where is F continuous?
   (b) Where is F *differentiable*? Calculate F' at the points of differentiability.

3. Let f be continuous on $\mathbb{R}$.

   (a) Define $F(x) = \int_{x-1}^{x+1} f(t)\,\mathrm{d}t$. Carefully show that F is differentiable on $\mathbb{R}$ and compute F'.
   (b) Repeat for $G(x) = \int_0^{\sin x} f(t)\,\mathrm{d}t$.

4. Recall Examples 4.24.4 and 4.27.3. Describe *all* anti-derivatives F of f on $[0,1)\cup(1,2]$. Which satisfy $\int_0^2 f(x)\,\mathrm{d}x = F(2) - F(0)$?

5. Suppose u, v satisfy the hypotheses of integration by parts. By FTC I, $\int_a^x u'(t)v(t)\,\mathrm{d}t$ is an anti-derivative of $u'(x)v(x)$: what does integration by parts say is another?

6. Use a substitution to integrate $\int_0^1 x\sqrt{1-x^2}\,\mathrm{d}x$

7. Use integration by parts and the substitution rule to evaluate $\int_0^b \arcsin x\,\mathrm{d}x$ for any $b < 1$.

8. Use integration by parts to evaluate $\int_0^b x\arctan x\,\mathrm{d}x$ for any $b > 0$

9. If f and u satisfy the hypotheses of the substitution rule, explain why both $(f\circ u)u'$ and f are integrable on the required intervals.

10. We prove a simpler version of the fundamental theorem when $f : [a,b] \to \mathbb{R}$ is *continuous*.

    Part I Define $F(x) = \int_a^x f(t)\,\mathrm{d}t$. If $c, x \in [a,b]$ where $c \neq x$, prove that

    $$m \le \frac{F(x) - F(c)}{x - c} \le M$$

    where m, M are the maximum and minimum values of $f(t)$ on the closed interval with endpoints c, x; why do m, M exist? Now deduce that $F'(c) = f(c)$.

    Part II Now suppose F is *any* anti-derivative of f on $[a,b]$. Use part (a) and the mean value theorem to prove that $\int_a^b f(t)\,\mathrm{d}t = F(b) - F(a)$.

4.36 Improper Integrals

The Riemann integral has several limitations. Even allowing for functions to be integrable on open intervals (Definition 4.12), the existence of $\int_a^b f(x)\,dx$ requires both:

- That (a, b) be a *bounded* interval.
- That f be *bounded* on (a, b).

Limits provide a natural way to extend the Riemann integral to unbounded intervals and functions.

Definition 4.32. Suppose $f : [a, b) \to \mathbb{R}$ satisfies the following properties:

- f is integrable on every closed bounded subinterval $[a, t] \subseteq [a, b)$.
- If b is finite, then f is unbounded at b (b can be ∞!)

The *improper integral* of f on $[a, b)$ is

$$\int_a^b f(x)\,dx := \lim_{t\to b^-} \int_a^t f(x)\,dx$$

This is *convergent* or *divergent* as is the limit.

If an integral is improper at its lower limit ($f : (a, b] \to \mathbb{R}$, etc.), then $\int_a^b f(x)\,dx := \lim_{s\to a^+} \int_s^b f(x)\,dx$.

If an integral is improper at both ends, choose any $c \in (a, b)$ and define

$$\int_a^b f(x)\,dx = \lim_{s\to a^+} \int_s^c f(x)\,dx + \lim_{t\to b^-} \int_c^t f(x)\,dx$$

provided *both* one-sided improper integrals exist and the limit sum makes sense.

Theorem 4.14 says that the choice of c for a doubly-improper integral is irrelevant.

Many properties of the Riemann integral transfer naturally to improper integrals, though not everything... For example, part 1 of Theorem 4.13 extends:

Theorem 4.33. *If $0 \le f(x) \le g(x)$ on $[a, b)$, then $\int_a^b f \le \int_a^b g$ whenever the integrals exist (standard or improper). In particular:*

- $\int_a^b f = \infty \implies \int_a^b g = \infty$
- $\int_a^b g$ *convergent* $\implies \int_a^b f$ *converges to some value* $\le \int_a^b g$

We leave some of the detail to Exercise 7.

Examples 4.34. 1. $\int_0^t x^2\,\mathrm{d}x = \frac{1}{3}t^3$ for any $t > 0$. Clearly

$$\int_0^\infty x^2\,\mathrm{d}x = \lim_{t\to\infty} \frac{1}{3}t^3 = \infty$$

More formally, the improper integral $\int_0^\infty x^2\,\mathrm{d}x$ diverges to infinity.

2. With $f(x) = x^{-4/3}$ defined on $[1, \infty)$,

$$\int_1^\infty x^{-4/3}\,\mathrm{d}x = \lim_{t\to\infty} \int_1^t x^{-4/3}\,\mathrm{d}x = \lim_{t\to\infty} \left[-3x^{-1/3}\right]_1^t = \lim_{t\to\infty} 3 - 3t^{-1/3} = 3$$

3. Consider $f(x) = |x|\, e^{-x^2/2}$ on $(-\infty, \infty)$. On any bounded interval $[0, t]$,

$$\int_0^t f(x)\,\mathrm{d}x = \int_0^t xe^{-x^2/2}\,\mathrm{d}x = \left[-e^{-x^2/2}\right]_0^t = 1 - e^{-t^2/2} \xrightarrow[t\to\infty]{} 1$$

By symmetry,

$$\int_{-\infty}^\infty |x|\, e^{-x^2/2}\,\mathrm{d}x = 1 + 1 = 2$$

This example arises naturally in probability: multiplying by $\frac{1}{\sqrt{2\pi}}$ computes the expectation of $|X|$ when X is a standard normally-distributed random variable

$$\mathbb{E}(|X|) = \int_{-\infty}^\infty \frac{1}{\sqrt{2\pi}}\, |x|\, e^{-x^2/2}\,\mathrm{d}x = \sqrt{\frac{2}{\pi}}$$

4. Our knowledge of derivatives $\frac{\mathrm{d}}{\mathrm{d}x}\sin^{-1} x = \frac{1}{\sqrt{1-x^2}}$ (or the substitution rule) allows us to evaluate

$$\int_0^1 \frac{1}{\sqrt{1-x^2}}\,\mathrm{d}x = \lim_{t\to 1^-} \int_0^t \frac{1}{\sqrt{1-x^2}}\,\mathrm{d}x = \lim_{t\to 1^-} \sin^{-1} t = \frac{\pi}{2}$$

By symmetry, $\int_{-1}^1 \frac{1}{\sqrt{1-x^2}}\,\mathrm{d}x = \pi$. By comparison, we obtain bounds on another improper integral:

$$\frac{1}{\sqrt{1-x^4}} \le \frac{1}{\sqrt{1-x^2}} \implies \int_{-1}^1 \frac{1}{\sqrt{1-x^4}}\,\mathrm{d}x \le \int_{-1}^1 \frac{1}{\sqrt{1-x^2}}\,\mathrm{d}x = \pi$$

5. Improper integrals need not exist. For instance,

$$\lim_{t\to\infty} \int_0^t \sin x\,\mathrm{d}x = \lim_{t\to\infty} 1 - \cos t$$

diverges by oscillation.

Exercises 4.36. *Key concepts: Formal definition and careful calculation of Improper Integrals*

1. Use your answers from Section 4.34 to decide whether the improper integrals $\int_0^1 \arcsin x \, \mathrm{d}x$ and $\int_0^\infty x \arctan x \, \mathrm{d}x$ exist. If so, what are their values?

2. Let p be a positive constant. Prove:

$$\int_0^1 \frac{1}{x^p} \, \mathrm{d}x = \begin{cases} \frac{1}{1-p} & \text{if } p < 1 \\ \infty & \text{if } p \geq 1 \end{cases} \qquad \int_1^\infty \frac{1}{x^p} \, \mathrm{d}x = \begin{cases} \frac{1}{p-1} & \text{if } p > 1 \\ \infty & \text{if } p \leq 1 \end{cases}$$

   (*The first of these justifies the convergence/divergence properties of p-series via the integral test*)

3. Suppose f is integrable on $[a, b]$. Explain why $\int_a^b f(x) \, \mathrm{d}x = \lim_{t \to b^-} \int_a^t f(x) \, \mathrm{d}x$ is still true, even though the integral is not improper.

4. State a version of integration by parts modified for when $\int_a^b u'(x)v(x) \, \mathrm{d}x$ is improper at b. Now evaluate $\int_0^\infty xe^{-4x} \, \mathrm{d}x$.

5. What is wrong with the following calculation?

$$\int_{-\infty}^{\infty} x \, \mathrm{d}x = \lim_{t \to \infty} \frac{1}{2} x^2 \Big|_{-t}^{t} = \lim_{t \to \infty} \frac{1}{2}(t^2 - t^2) = \lim_{t \to \infty} 0 = 0$$

6. Prove or disprove: if $\int f$ and $\int g$ are convergent improper integrals, so is $\int fg$.

7. Prove part of Theorem 4.33. Suppose $0 \leq f(x) \leq g(x)$ for all $x \in [a, b)$, and that $\int_a^b g$ is a convergent improper integral. Prove that $\int_a^b f$ converges and that $\int_a^b f \leq \int_a^b g$.

Extensions of the Riemann Integral (just for fun)

In the 1890s, Thomas Stieltjes [24] offered a generalization of the Riemann integral.

Definition 4.35. Let $f : [a, b] \to \mathbb{R}$ be bounded and $\alpha : [a, b] \to \mathbb{R}$ monotonically increasing. Given a partition $P = \{x_0, \ldots, x_n\}$ of $[a, b]$, define the sequence of differences

$$\Delta\alpha_i = \alpha(x_i) - \alpha(x_{i-1})$$

The *upper/lower Darboux–Stieltjes sums/integrals* are defined analogously to the pure Riemann case:

$$U(f, P, \alpha) = \sum_{i=1}^{n} \sup_{[x_{i-1}, x_i]} f(x)\,\Delta\alpha_i \qquad L(f, P, \alpha) = \sum_{i=1}^{n} \inf_{[x_{i-1}, x_i]} f(x)\,\Delta\alpha_i$$

$$U(f, \alpha) = \inf_P U(f, P, \alpha) \qquad L(f, \alpha) = \sup_P L(f, P, \alpha)$$

If $U(f, \alpha) = L(f, \alpha)$, we say that f is *Riemann–Stieltjes integrable* of class $\mathcal{R}(\alpha)$ and denote its value $\int_a^b f(x)\,\mathrm{d}\alpha$.

The standard Riemann integral corresponds to $\alpha(x) = x$. It is the ability to choose other functions α that makes the Riemann–Stieltjes integral both powerful and applicable.

Standard Properties Most results in sections 4.32 and 4.33 hold with suitable modifications, as does the discussion of improper integrals. For instance,

$$f \in \mathcal{R}(\alpha) \iff \exists P \text{ such that } U(f, P, \alpha) - L(f, P, \alpha) < \epsilon$$

The result regarding the piecewise continuity of f is a notable exception: depending on α, a piecewise continuous f might not lie in $\mathcal{R}(\alpha)$.

Weighted integrals If α is differentiable, we obtain a standard Riemann integral

$$\int_a^b f(x)\,\mathrm{d}\alpha = \int_a^b f(x)\alpha'(x)\,\mathrm{d}x$$

weighted so that $f(x)$ contributes more when α is increasing rapidly.

Probability If $\alpha(a) = 0$ and $\alpha(b) = 1$, then α may be viewed as a *probability distribution function* and its derivative α' as the corresponding *probability density function*. For example:

1. The *uniform distribution* on $[a, b]$ has $\alpha = \frac{1}{b-a}(x - a)$ so that

   $$\int_a^b f(x)\,\mathrm{d}\alpha = \frac{1}{b-a}\int_a^b f(x)\,\mathrm{d}x$$

   Since α' is constant, the integrals weigh all values of x *uniformly*.
2. The standard *normal distribution* has $\alpha(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-t^2/2}\,\mathrm{d}t$. The fact that $\alpha' = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$ is maximal when $x = 0$ reflects the fact that a normally distributed variable is clustered near its mean.

In all cases, $\int f(x)\,\mathrm{d}\alpha = \mathbb{E}(f(X))$ computes an expectation (see, e.g., Example 4.34.3).

[24] Stieltjes was Dutch; the pronunciation is roughly ‘steelchez.’

Non-differentiable or continuous α This provides major flexibility! For example, if $Q = \{s_0, \ldots, s_n\}$ partitions $[a, b]$, and $(c_k)_{k=1}^n$ is a positive sequence, then

$$\alpha(x) = \begin{cases} 0 & \text{if } x = a \\ \sum_{i=1}^{k} c_i & \text{if } x \in (s_{k-1}, s_k] \end{cases}$$

defines an increasing step function, and the Riemann–Stieltjes integral a weighted *sum*

$$\int_a^b f(x)\,\mathrm{d}\alpha = \sum_{i=1}^{n} c_i f(s_i)$$

Taking an infinite increasing sequence $(s_n) \subseteq [a, b]$ results in an *infinite series*, which helps explain why so many results for series and integrals look similar!

This also touches on probability. For example, let $p \in [0, 1]$, $n \in \mathbb{N}$, and $s_k = k$ on the interval $[0, n]$. If $c_k = \binom{n}{k} p^k (1-p)^{n-k}$, then

$$\int f(x)\,\mathrm{d}\alpha = \sum_{k=0}^{n} \binom{n}{k} p^k (1-p)^{n-k} f(x) = \mathbb{E}(f(X))$$

is the expectation of $f(X)$ when $X \sim B(n, p)$ is binomially distributed.

Lebesgue Integration: Integrals and Convergence

Lebesgue's extension essentially uses rectangles whose *heights* tend to zero: cutting up the area under a curve using *horizontal* instead of *vertical* strips. One of its major purposes is to permit a more general interchange of limits and integration in many cases of *pointwise* (non-uniform) convergence. To see the problem, consider the sequence of piecewise continuous functions

$$f_n : [0, 1] \to \mathbb{R} : x \mapsto \begin{cases} 1 & \text{if } x = \frac{p}{q} \in \mathbb{Q} \text{ with } q \le n \\ 0 & \text{otherwise} \end{cases}$$

Each f_n is Riemann integrable with $\int_0^1 f_n(x)\,\mathrm{d}x = 0$. However, the pointwise limit

$$f(x) = \begin{cases} 1 & \text{if } x \in \mathbb{Q} \\ 0 & \text{if } x \notin \mathbb{Q} \end{cases}$$

is *not* Riemann integrable (compare Example 4.8.2). In the Lebesgue theory, the limit f turns out to be integrable with integral 0, so that

$$\lim_{n\to\infty} \int_0^1 f_n(x)\,\mathrm{d}x = \int_0^1 \lim_{n\to\infty} f_n(x)\,\mathrm{d}x$$

Recall (Theorem 2.19) that the interchange of limits and integrals would be automatic *if* the convergence $f_n \to f$ were *uniform*: of course the convergence isn't uniform here.

Like *measure theory* (recall Theorem 4.22), *Lebesgue integration* is a central topic in graduate analysis.