Calculus or Analysis II: Topology, Functions, Differentiation and Integration

Gudfit

Contents

Chapter 1

Ideas & Motivations

Welcome to Calculus or Analysis II (with *some*[1] theory) by me (Gudfit). The point of these notes is to cover everything I think is important as I build up to my current knowledge, while keeping it free and accessible for everyone from kids to adults.

I aim for each set of notes to be max 200 pages, as rigorous as possible, and far-reaching too. (This will probably be the longest notes as i try to fit applied calculus with theoretical analysis into one single pdf.) That means I'll cover the axioms and proofs of the most interesting stuff, plus I'll pull in other subjects we've already touched on to show how math builds on itself like funky Lego. These notes build on my existing informal logic, algebra I notes and geometry notes, and they're aimed at keeping the proofs, ideas, and build-up of calculus as informal as possible.

It'll be a mix of quick ideas and concepts, but in the appendix for each section, I'll go rigorous with the key axioms pulled from a bunch of books.

As you browse the contents, you may notice that these notes are somewhat dense in terms of theory. This is by design. Many standard Calculus concepts are simply extrapolations of Real Analysis; when we gloss over these roots, it becomes difficult to see the true beauty and underlying simplicity that Calculus offers.

Consequently, we are taking the rigorous "Analysis route." We will build the machinery from the ground up across a series of notes:

Part I: Limits and Convergence.
Part II (This Document): Topology, Functions, Differentiation, and Integration.
Part III: Multivariable Calculus.
Part IV: Calculus on Manifolds.
Part V?: Metrics Spaces.

[1] LOL

Part I

Topology and Limit of Functions

Chapter 2

Introduction to Topology

In the preceding notes, we established the machinery of limits for sequences ($x_n \to l$). We now turn to continuity on $\mathbb{R}$.

Historically, the concepts of calculus were utilised with great efficacy by Newton and Leibniz long before they were placed on a firm logical foundation. For centuries, arguments relied on an intuitive understanding of "approaching" a value or "infinitesimal" changes. While sufficient for the physics of planetary motion, this intuition eventually faltered when confronted with the pathological entities of pure mathematics. To resolve these ambiguities, we must formalise the study of space, proximity, and deformation. This is the domain of *Topology*.

Derived from the Greek *topos* (place) and *mathema* (knowledge), topology is often described as "rubber-sheet geometry" — the study of properties preserved under continuous deformation. However, for the analyst, topology is fundamentally the study of *closeness*. It provides the vocabulary required to define limits and continuity without explicit recourse to metric distances in every instance, allowing us to generalise the familiar properties of the real line to more abstract spaces.

2.1 The Intuition of Continuity

Before providing a definition, we must examine what we intuitively mean when we claim a function $f : \mathbb{R} \to \mathbb{R}$ is "continuous". Several heuristic definitions have been proposed throughout history, notably by Euler and Cauchy.

1. **Geometric Connectivity:** A function is continuous if its graph has no gaps, jumps, or breaks.
2. **Mechanical Tracing:** A function is continuous if its graph can be drawn without lifting the pen from the paper.
3. **Stability:** A function is continuous if a small perturbation in the input produces only a small perturbation in the output.

While the third intuition proves most fruitful for analysis (leading to the $\epsilon - \delta$ definition), the first two suggest a relationship between continuity and the "connectedness" of the domain. To verify these intuitions, we examine functions that test their limits.

Example 2.1.1. The Heaviside Step Function. Consider the function $H : \mathbb{R} \to \mathbb{R}$ defined by:

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

(Note: In engineering contexts, $H(0)$ is often defined as $1/2$ or 1, but the jump discontinuity remains). As illustrated in Figure 2.1, this function exhibits a clear "jump" at $x = 0$. It fails the stability intuition: a movement from $x = -10^{-9}$ to $x = 0$ results in a variation of output of magnitude 2, regardless of how small the input step is. Thus, H is discontinuous at 0.

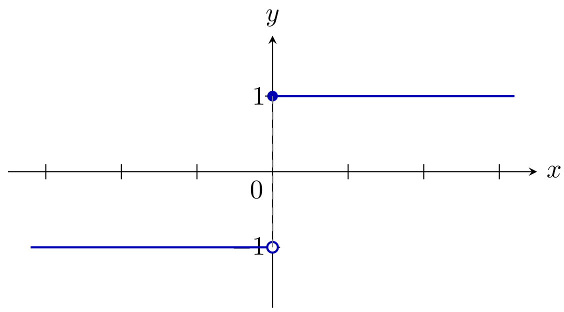


Figure 2.1: The Heaviside step function $H(x)$ with a jump discontinuity at $x = 0$.

The Intermediate Value Property

The intuition that a continuous function "cannot skip values" is formalised by the Intermediate Value Theorem (IVT). It is tempting to define continuity solely by this property.

Definition 2.1.1. ***Intermediate Value Property (IVP)***. A function $f : [a, b] \to \mathbb{R}$ has the Intermediate Value Property if, for any y strictly between $f(a)$ and $f(b)$, there exists some $c \in (a, b)$ such that $f(c) = y$.

If a function allows one to draw its graph without lifting the pen, it must necessarily traverse every vertical position between its endpoints. However, is the converse true? Is a function that satisfies the IVP necessarily continuous in the sense of "stability"?

Example 2.1.2. The Topologist's Sine Curve. Define function $f : \mathbb{R} \to \mathbb{R}$:

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

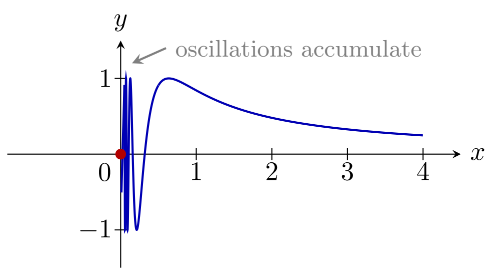


Figure 2.2: The topologist's sine curve $f(x) = \sin(1/x)$ for $x \neq 0$, with $f(0) = 0$. The function oscillates infinitely often as $x \to 0$.

As $x \to 0$, the term $1/x$ grows without bound, causing the sine function to oscillate with infinite frequency between -1 and 1.

- **IVP Compliance:** This function satisfies IVP. On any interval containing 0, the function takes all values in $[-1, 1]$.
- **Failure of Stability:** Near 0, the value of $f(x)$ is unpredictable. An infinitesimal move from 0 could land on 1, -1, or 0. The graph, while not having a simple "jump" like the Heaviside function, is not a simple curve we can trace physically.

This function satisfies definition 2.1.1 but fails the stability criterion; thus, IVP is a necessary but not sufficient condition for continuity.

Pathological Functions We consider functions that defy geometric intuition to demonstrate the necessity of a robust topological framework.

Example 2.1.3. The Dirichlet Function. Let $D : \mathbb{R} \to \mathbb{R}$:

$$D(x) = \begin{cases} 1 & \text{if } x \in \mathbb{Q} \\ 0 & \text{if } x \in \mathbb{R} \setminus \mathbb{Q} \end{cases}$$

Recall that both $\mathbb{Q}$ and $\mathbb{R} \setminus \mathbb{Q}$ are dense in $\mathbb{R}$. In any neighbourhood of any point x, the function D oscillates between 0 and 1 infinitely many times. Visualising this is difficult (see Figure 2.3); the graph appears as two parallel horizontal lines of "dust" rather than solid strokes. Intuitively, this function is continuous *nowhere*.

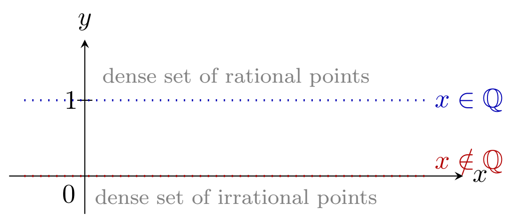


Figure 2.3: Visualisation of the Dirichlet function $D(x)$. The dashed lines indicate that the function is defined on sets that are dense in the real numbers, appearing as "point-dust" rather than continuous segments.

Sequential Characterisation We use theorem 2.1.1. Apply it to the Dirichlet function at an arbitrary point $r \in \mathbb{R}$.

***Theorem* 2.1.1. Sequential Criterion for Continuity**. A function $f : A \to \mathbb{R}$ is continuous at a point $c \in A$ if and only if for every sequence (x_n) in A converging to c, the sequence $(f(x_n))$ converges to $f(c)$.

$$\lim_{n\to\infty} x_n = c \implies \lim_{n\to\infty} f(x_n) = f(c)$$

Let us apply this to the Dirichlet function at an arbitrary point $r \in \mathbb{R}$.

- Due to the density of rationals, there exists a sequence $(q_n) \subseteq \mathbb{Q}$ such that $q_n \to r$. For this sequence, $D(q_n) = 1$ for all n, so $\lim D(q_n) = 1$.
- Due to the density of irrationals, there exists a sequence $(i_n) \subseteq \mathbb{R} \setminus \mathbb{Q}$ such that $i_n \to r$. For this sequence, $D(i_n) = 0$ for all n, so $\lim D(i_n) = 0$.

Since the limits of the image sequences differ $(1 \neq 0)$, the function cannot be continuous at r, regardless of the value of $D(r)$. This sequential approach confirms our intuition that the Dirichlet function is nowhere continuous.

Why Topology?

While sequences allow the detection of discontinuity, the sequential definition is cumbersome for global properties. We seek a definition relying on set structure. Topology abstracts "closeness" via *open sets*. Rather than relying on a metric $d(x, y) = |x - y|$, we specify which subsets are "open neighbourhoods". In the subsequent chapters, we will define the topology of the real line, explore the concepts of open and closed sets, and derive the properties of continuous functions.

Chapter 3

Adherence and Closed Sets

Refining continuity requires formalising proximity. A function f is continuous at x if, whenever a set A is close to x, the image set $f(A)$ is close to $f(x)$. We must quantify "closeness" not as small finite distance, but as infinitesimal proximity.

We begin with the concept of adherence and closed sets, aligning with the sequential definitions established in the previous chapter.

3.1 Adherent and Limit Points

We capture the idea of infinitesimal proximity via sequences.

Definition 3.1.1. ***Adherent Point.*** Let $A \subseteq \mathbb{R}$. A point $x \in \mathbb{R}$ is called an adherent point of A if there exists a sequence (a_n) such that $a_n \in A$ for all n, and $\lim_{n\to\infty} a_n = x$. We denote the set of all adherent points of A as $\text{adh}(A)$ (also written $\bar{A}$).

Essentially, x is adherent to A if x can be approximated arbitrarily well by elements of A. For $x \in A$, the constant sequence $a_n = x$ implies $A \subseteq \text{adh}(A)$. Analysis often requires a stricter condition distinct from trivial adherence.

Remark. (Sequential vs Topological Adherence). A point x is adherent to A if and only if every neighbourhood of x intersects A.

- If $x \in \text{adh}(A)$, there is a sequence $a_n \to x$. For any $r > 0$, eventually $a_n \in B_r(x)$, so $B_r(x) \cap A \neq \emptyset$.
- Conversely, if every $B_{1/n}(x)$ intersects A, choose $a_n \in B_{1/n}(x) \cap A$. Then $a_n \to x$, so $x \in \text{adh}(A)$.

This equivalence is frequently used to switch between perspectives.

Definition 3.1.2. ***Limit Point.*** Let $A \subseteq \mathbb{R}$. A point $x \in \mathbb{R}$ is called a limit point (or accumulation point) of A if there exists a sequence (a_n) of *distinct* points in A such that $\lim_{n\to\infty} a_n = x$.

Example 3.1.1. Intervals. Let $A = (0, 1)$ and $B = [0, 1]$.

- Any $x \in (0, 1)$ is adherent to A (trivial).
- The endpoint 0 is adherent to A because the sequence $(1/n)_{n\geq 2}$ lies in A and converges to 0. Similarly, 1 is adherent.
- Consequently, $\text{adh}(A) = [0, 1]$.
- Similarly, $\text{adh}(B) = [0, 1]$.

In this case, the set of limit points for both A and B is also $[0, 1]$.

The distinction between adherent points and limit points becomes apparent when considering discrete sets.

Example 3.1.2. The Hyperharmonic Set. Define $S = \{1/n : n \in \mathbb{N}\}$.

- **Adherent Points:** Every element $1/n \in S$ is adherent. Furthermore, 0 is adherent because $1/n \to 0$. Thus, $\text{adh}(S) = S \cup \{0\}$.
- **Limit Points:** Take an element $x = 1/k \in S$. Any sequence in S converging to $1/k$ must eventually be constant. Thus, no point in S is a limit point.
- However, the sequence $(1/n)$ consists of distinct points in S and converges to 0. Thus, 0 is the *unique* limit point of S (see Figure 3.1).

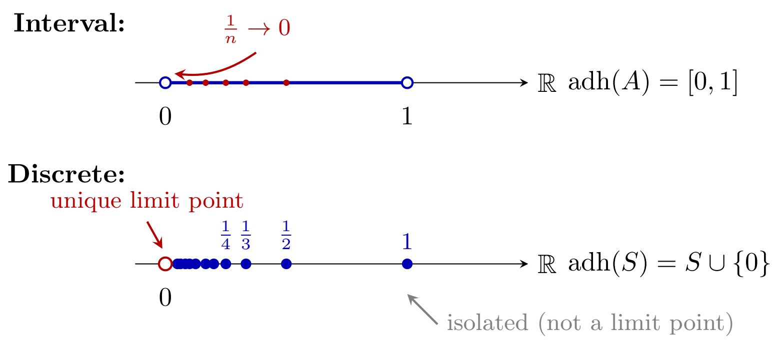


Figure 3.1: Top: The interval $A = (0, 1)$ with sequences converging to the boundary points. Bottom: The hyperharmonic set $S = \{1/n : n \in \mathbb{N}\}$ with 0 as its unique limit point; all other points are isolated.

Remark. Finite sets possess no limit points. If A is finite, any convergent sequence in A must be eventually constant. Thus, the only adherent points of a finite set are the elements of the set itself.

Closed Sets

A set is "closed" if it contains its boundary, or more precisely, if it contains all points that are infinitesimally close to it.

Definition 3.1.3. ***Closed Set***. A set $A \subseteq \mathbb{R}$ is said to be closed if it contains all its adherent points. Equivalently, A is closed if $A = \text{adh}(A)$.

Using our previous examples:

1. The interval $[0, 1]$ is closed because its adherent set is $[0, 1]$.
2. The interval $(0, 1)$ is *not* closed because 0 and 1 are adherent but not contained in the set.
3. The set $S = \{1/n : n \in \mathbb{N}\}$ is not closed because 0 is adherent but $0 \notin S$.
4. Any finite set is closed.

3.2 Open Sets and Interiors

While adherence captures proximity *to* a set, we define the dual notion of being strictly *inside* a set.

Definition 3.2.1. ***Neighbourhood***. For any $x \in \mathbb{R}$ and $r > 0$, the r-neighbourhood (or open ball) of x is the set:

$$B_r(x) = (x - r, x + r) = \{y \in \mathbb{R} : |x - y| < r\}$$

Definition 3.2.2. ***Interior Point***. Let $A \subseteq \mathbb{R}$. A point $x \in A$ is an interior point of A if there exists some $r > 0$ such that the entire neighbourhood $B_r(x)$ is contained within A:

$$B_r(x) \subseteq A$$

The set of all interior points of A is denoted A° or $\text{int}(A)$.

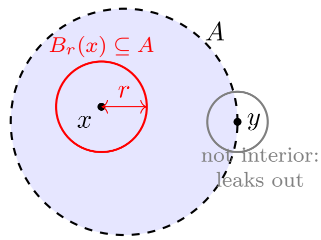


Figure 3.2: Visual representation of an interior point x with its neighborhood $B_r(x)$ entirely contained in A, compared to a boundary point y.

This definition formalises the idea of being "fully inside" a set, insulated from the complement by a buffer of radius r.

Example 3.2.1. Interiors.

- If $A = [0, 1]$, then $x = 1/2$ is an interior point (take $r = 0.1$). However, $x = 0$ is not an interior point, as any interval $(-\epsilon, \epsilon)$ contains negative numbers, which are not in A. Thus, $\text{int}([0, 1]) = (0, 1)$.
- If $A = \mathbb{Q}$, the set of rationals, then $\text{int}(A) = \emptyset$. Any interval $(x - r, x + r)$ contains irrational numbers, so no neighbourhood is ever fully contained in $\mathbb{Q}$.

Definition 3.2.3. ***Open Set***. A set $A \subseteq \mathbb{R}$ is said to be open if every point in A is an interior point. Equivalently, A is open if $A = \text{int}(A)$.

Example 3.2.2. Open Intervals. Any open interval (a, b) is an open set.

Proof. Let $x \in (a, b)$. We must find a radius r such that $(x - r, x + r) \subseteq (a, b)$. Let $r = \min(|x - a|, |x - b|)$. Since $x \in (a, b)$, $r > 0$. For any $y \in (x - r, x + r)$, we have $a \leq x - r < y < x + r \leq b$. Thus $B_r(x) \subseteq (a, b)$, so the set is open. ■

The collection of open sets satisfies specific algebraic properties that form the axioms of general topology.

***Proposition* 3.2.1.** *Union of Open Sets.* The union of any collection of open sets is open.

Proof. Let $\{U_\alpha\}_{\alpha \in \Lambda}$ be an arbitrary collection of open sets, and let $U = \bigcup_{\alpha \in \Lambda} U_\alpha$. Let $x \in U$. By the definition of union, there exists some index $\lambda \in \Lambda$ such that $x \in U_\lambda$. Since U_λ is an open set, x is an interior point of U_λ. Thus, there exists $r > 0$ such that $B_r(x) \subseteq U_\lambda$. Since $U_\lambda \subseteq U$, it follows that $B_r(x) \subseteq U$. Therefore, x is an interior point of U. Since x was arbitrary, U is an open set. ■

Remark. (Finite versus infinite intersections of open sets). The intersection of *finite* open sets is open, but the intersection of *infinite* open sets need not be. Consider $U_n = (-1/n, 1/n)$. Each is open, but $\bigcap_{n=1}^{\infty} U_n = \{0\}$, which is not open.

3.3 Isolated Points and Accumulation

In the previous section, we noted that adherent points are those which can be "touched" by the set A through a sequence. However, adherence encompasses two distinct behaviours: approaching a point via distinct elements (accumulation), and simply being a point that is already "there" but isolated from the rest. To formalise this, first, we establish a geometric characterisation of limit points.

***Proposition* 3.3.1.** *Topological Characterisation of Limit Points.* Let $A \subseteq \mathbb{R}$. A point x is a limit point of A if and only if every open neighbourhood of x contains a point of A distinct from x.

$$x \text{ is a limit point } \iff \forall r > 0, \exists y \in A \cap B_r(x) \text{ such that } y \neq x$$

Proof.

(⇒) Suppose x is a limit point. There exists a sequence (a_n) of distinct points in A converging to x. Since the points are distinct, x can appear in the sequence at most once. By discarding that term (if it exists) and re-indexing, we may assume $a_n \neq x$ for all n. Let $r > 0$. By the definition of convergence, there exists N such that $a_n \in B_r(x)$ for all $n > N$. Thus, $B_r(x)$ contains points of A (specifically $a_{N+1}, a_{N+2}, \dots$) which are distinct from x.

(⇐) Conversely, suppose every neighbourhood contains a point distinct from x. We construct a sequence of distinct points inductively. Let $r_1 = 1$. Choose $a_1 \in A \cap B_{r_1}(x)$ with $a_1 \neq x$. For $n > 1$, define $r_n = \min\left(\frac{1}{n}, \frac{|x-a_1|}{2}, \dots, \frac{|x-a_{n-1}|}{2}\right)$. Since $a_k \neq x$ for all $k < n$, each distance $|x - a_k| > 0$, so $r_n > 0$. Choose $a_n \in A \cap B_{r_n}(x)$ with $a_n \neq x$. By construction, $|x - a_n| < r_n \leq \frac{|x-a_k|}{2} < |x - a_k|$ for all $k < n$. Thus a_n is distinct from all preceding terms. Since $r_n \leq 1/n$, we have $|a_n - x| \to 0$, so $a_n \to x$.

■

This proposition motivates the classification of adherent points into two mutually exclusive categories.

Definition 3.3.1. ***Isolated Point***. Let $x \in A$. We say x is an isolated point of A if x is not a limit point of A. Equivalently, x is isolated if there exists a neighbourhood $B_r(x)$ such that $B_r(x) \cap A = \{x\}$.

***Proposition* 3.3.2.** *Structure of Adherence.* Every adherent point of a set A is either a limit point or an isolated point.

Proof. Let $x \in \text{adh}(A)$.

- If x is a limit point, we are done.
- If x is not a limit point, then by the negation of proposition 3.3.1, there exists some $r > 0$ such that $B_r(x) \cap A \subseteq \{x\}$. Since x is adherent, there must be *some* sequence in A converging to x. For sufficiently large n, elements of this sequence must lie in $B_r(x)$. Since the only available point in this neighbourhood is x itself, the sequence must eventually be constant ($a_n = x$). Thus $x \in A$. Consequently, $B_r(x) \cap A = \{x\}$, satisfying the definition of an isolated point.

■

Example 3.3.1. Decomposition of Adherent Points. Consider $A = (0, 1) \cup \{2\}$.

- The limit points are $[0, 1]$. Any $x \in [0, 1]$ can be approached by distinct points in $(0, 1)$. The point 2 is not a limit point because the neighbourhood $B_{0.5}(2) = (1.5, 2.5)$ contains only 2 from the set A.
- The point 2 is adherent (trivially, as $2 \in A$). Since it is not a limit point, it is an isolated point.

3.4 Basic Properties of Open and Closed Sets

Having defined open and closed sets, we now explore their algebraic structure. The properties of union and intersection differ between these two classes in a perfectly dual way, governed by De Morgan's Laws.

Algebra of Open Sets

***Proposition* 3.4.1.** *Open Set Topology.*

(i) The union of any collection of open sets is open.
(ii) The intersection of any *finite* collection of open sets is open.

Proof.

(i) This follows from proposition 3.2.1.

(ii) Let $U_1, \ldots, U_n$ be open sets and let $U = \bigcap_{i=1}^{n} U_i$. If $U = \emptyset$, it is open (vacuously, every point in the empty set is an interior point). If $U \neq \emptyset$, let $x \in U$. Then $x \in U_i$ for all $i = 1, \ldots, n$. Since each U_i is open, there exist radii $r_i > 0$ such that $B_{r_i}(x) \subseteq U_i$. Let $r = \min(r_1, \ldots, r_n)$. Since the collection is finite, $r > 0$. Then $B_r(x) \subseteq B_{r_i}(x) \subseteq U_i$ for all i. Consequently, $B_r(x) \subseteq \bigcap U_i = U$. Thus x is an interior point.

■

Remark. The finiteness condition in (ii) is crucial. As noted earlier, $\bigcap_{n=1}^{\infty}(-1/n, 1/n) = \{0\}$, which is not open.

Duality of Open and Closed Sets The complement operation captures the relationship between open and closed sets.

***Proposition* 3.4.2.** *Complement Characterisation.* A set $U \subseteq \mathbb{R}$ is open if and only if its complement $U^c = \mathbb{R} \setminus U$ is closed.

Proof.

($\Rightarrow$) Let U be open. We show U^c is closed in the sense of definition 3.1.3. Let $x \in \text{adh}(U^c)$. If x is isolated in U^c, then $x \in U^c$ by definition 3.3.1. Otherwise x is a limit point of U^c by proposition 3.3.2. Suppose, for contradiction, that $x \notin U^c$. Then $x \in U$. Since U is open, there exists $r > 0$ such that $B_r(x) \subseteq U$. This implies $B_r(x) \cap U^c = \emptyset$. However, if x is a limit point of U^c, every neighbourhood of x must contain a point of U^c. This is a contradiction. Thus $x \in U^c$. Since U^c contains its limit points, it is closed.

($\Leftarrow$) Let $F = U^c$ be closed. Let $x \in U$, so $x \notin F$. If every neighbourhood of x met F, then $x \in \text{adh}(F)$ (by the sequential characterisation), hence $x \in F$ by definition 3.1.3, a contradiction. Therefore some $B_r(x)$ is disjoint from F, so $B_r(x) \subseteq U$ and U is open by definition 3.2.3.

■

Algebra of Closed Sets By applying De Morgan's Laws to the properties of open sets, we deduce the properties of closed sets. Recall: $(\bigcup A_\alpha)^c = \bigcap A_\alpha^c$ and $(\bigcap A_\alpha)^c = \bigcup A_\alpha^c$.

***Proposition* 3.4.3.** *Closed Set Topology.*

(i) The intersection of any collection of closed sets is closed.
(ii) The union of any *finite* collection of closed sets is closed.

Proof.

(i) Let $\{F_\alpha\}$ be a collection of closed sets. By proposition 3.4.2, $U_\alpha = F_\alpha^c$ are open. $\bigcap F_\alpha = (\bigcup U_\alpha)^c$. Since $\bigcup U_\alpha$ is open by proposition 3.4.1, its complement is closed by proposition 3.4.2.

(ii) Let $F_1, \ldots, F_n$ be closed. Then $U_i = F_i^c$ are open. $\bigcup_{i=1}^{n} F_i = (\bigcap_{i=1}^{n} U_i)^c$. Since the finite intersection of open sets is open, its complement is closed.

■

Remark. Infinite unions of closed sets need not be closed. Consider $F_n = [1/n, 1 - 1/n]$. $\bigcup_{n=1}^{\infty} F_n = (0, 1)$, which is open.

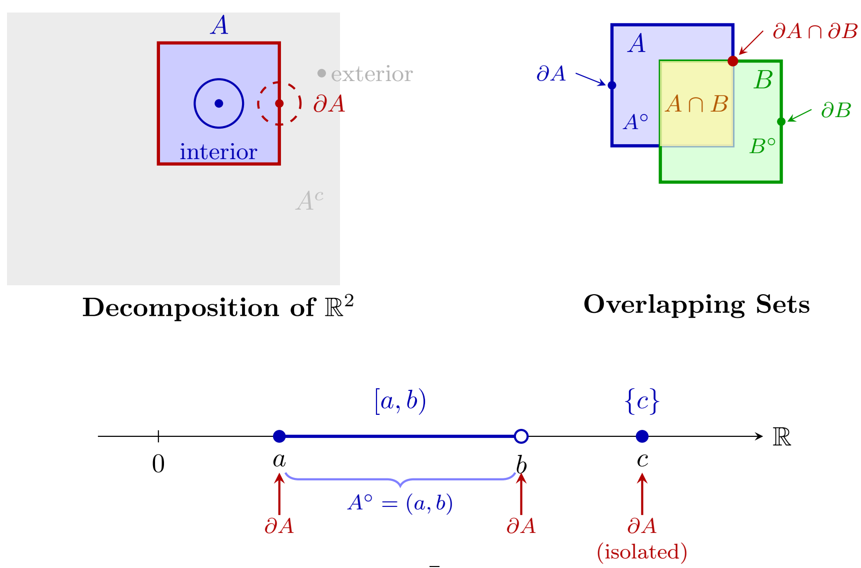
3.5 Closure, Interior, and Boundary

We conclude by defining operators that generate open and closed sets from arbitrary sets.

Definition 3.5.1. ***Closure and Interior***. Let $A \subseteq \mathbb{R}$.

1. The **closure** of A, denoted $\bar{A}$ or $\mathrm{cl}(A)$, is the set of all adherent points of A. It satisfies
$$\bar{A} = A \cup \{x : x \text{ is a limit point of } A\}.$$
Moreover, $\bar{A}$ is the smallest closed set containing A. (Proof: If F is closed and $A \subseteq F$, then any adherent point of A is adherent to F, hence in F. Thus $\bar{A} \subseteq F$).
2. The **interior** of A, denoted A° or $\mathrm{int}(A)$, is the set of all interior points of A. It is the largest open set contained in A.

Definition 3.5.2. ***Boundary***. The boundary of a set A, denoted ∂A, is the set of points that are adherent to both A and its complement A^c.
$$\partial A = \bar{A} \cap \overline{A^c}$$
Equivalently, $x \in \partial A$ if every neighbourhood of x meets both A and A^c; compare proposition 3.3.1.



Example: $A = [a, b) \cup \{c\}, \quad \bar{A} = [a, b] \cup \{c\}, \quad \partial A = \{a, b, c\}$

Figure 3.3: The decomposition of space into the interior, the boundary, and the exterior of a set A.

Example 3.5.1. Topological Operators on $\mathbb{Q}$. Let $A = \mathbb{Q}$.

- **Interior:** $\mathrm{int}(\mathbb{Q}) = \emptyset$. (No interval is purely rational).
- **Closure:** $\bar{\mathbb{Q}} = \mathbb{R}$. (Every real number is a limit of rationals).
- **Boundary:** $\partial \mathbb{Q} = \bar{\mathbb{Q}} \cap \overline{\mathbb{Q}^c} = \mathbb{R} \cap \mathbb{R} = \mathbb{R}$.

This extreme example illustrates why the topology of the real line is non-trivial; a set can have an empty interior yet be its own boundary.

3.6 Exercises

1. **Topological Anatomy.** Consider the set $E \subseteq \mathbb{R}$ defined by:
$$E = (-3, -2] \cup \left\{ 1 + \frac{1}{n} \;\middle|\; n \in \mathbb{N} \right\} \cup \{5\}.$$
Determine the following sets:

(a) The set of limit points of E, denoted E'.
(b) The closure $\bar{E}$.
(c) The interior $\text{int}(E)$.
(d) The boundary ∂E.
(e) The set of isolated points of E.

Remark. Be particularly careful with the point 1. Is it in E? Is it a limit point?

2. Algebra of Closures and Interiors. Let A and B be subsets of $\mathbb{R}$. Determine whether the following equalities hold in general. If true, provide a proof; if false, provide a counter-example.

(a) $\overline{A \cup B} = \bar{A} \cup \bar{B}$.
(b) $\overline{A \cap B} = \bar{A} \cap \bar{B}$.
(c) $\text{int}(A \cap B) = \text{int}(A) \cap \text{int}(B)$.
(d) $\text{int}(A \cup B) = \text{int}(A) \cup \text{int}(B)$.

Remark. Consider intervals sharing a boundary or dense sets.

3. Boundaries and Complements. Compute the boundary ∂A for the following sets:

(a) $A = \mathbb{Z}$.
(b) $A = \{1/n \mid n \in \mathbb{N}\}$.
(c) $A = \mathbb{Q}$.

Based on your calculations, prove or disprove: $\partial A = \partial(A^c)$ for any set A.

4. Supremum and Closure. We defined the supremum of a set bounded above via the Order Axioms. Here we link it to topology. Let $S \subseteq \mathbb{R}$ be a non-empty set bounded above. Prove that $\sup S \in \bar{S}$. Consequently, deduce that if S is closed, then $\sup S \in S$.

Remark. Use the characterisation of the supremum: for any $\epsilon > 0$, there exists $x \in S$ such that $\sup S - \epsilon < x \leq \sup S$.

5. The Boundary Theorem. Prove that for any set $A \subseteq \mathbb{R}$, the boundary ∂A is a closed set.

Remark. Recall that $\partial A = \bar{A} \cap \overline{A^c}$. What do we know about the intersection of closed sets?

6. The Distance Function I. Let $A \subseteq \mathbb{R}$ be a non-empty set. Define the distance from $x \in \mathbb{R}$ to A as $d(x, A) = \inf\{|x-a| : a \in A\}$. Prove that $d(x, A) = 0$ if and only if there exists a sequence $(a_n)_{n=1}^{\infty} \subseteq A$ such that $\lim_{n\to\infty} a_n = x$.

7. ⋆ The Cantor Set. The Cantor ternary set $\mathcal{C}$ is constructed by iteratively removing the middle third of closed intervals. Start with $C_0 = [0, 1]$. Remove $(1/3, 2/3)$ to get $C_1 = [0, 1/3] \cup [2/3, 1]$. Remove the middle thirds of these to get $C_2 = [0, 1/9] \cup [2/9, 1/3] \cup [2/3, 7/9] \cup [8/9, 1]$, and so on. Let $\mathcal{C} = \bigcap_{n=0}^{\infty} C_n$.

(a) Prove that $\mathcal{C}$ is a closed set.
(b) Prove that $\text{int}(\mathcal{C}) = \emptyset$.

Remark. For (b), consider the total length of the intervals removed. Alternatively, show that $\mathcal{C}$ contains no interval of length $\epsilon > 0$ because the maximum length of intervals in C_n is $1/3^n$.

Chapter 4

Continuity

The intuition of a continuous function as one that preserves proximity is formalised by the equivalence of topological, sequential, and metric definitions. This equivalence allows continuity to be expressed interchangeably via adherence, convergence of sequences, or metric inequalities.

4.1 Notions of Continuity

Let us begin by stating the fundamental theorem that links the topological structure of the domain to the metric properties of the codomain. This theorem links the topological structure of the domain to the metric properties of the codomain, allowing translation between the language of sets, sequences, and inequalities.

***Theorem* 4.1.1. The Three Faces of Continuity**. Let $f : A \to \mathbb{R}$ be a function and fix a point $x \in A$. The following statements are equivalent:

(i) **Adherence Preservation:** For every subset $B \subseteq A$, if x is adherent to B, then $f(x)$ is adherent to $f(B)$.

$$x \in \operatorname{adh}(B) \implies f(x) \in \operatorname{adh}(f(B))$$

This requires that any point close to B is mapped to a point close to $f(B)$ (definition 3.1.1).

(ii) **Sequential Continuity:** For every sequence (x_n) in A that converges to x, the image sequence $(f(x_n))$ converges to $f(x)$.

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

(iii) **The $\epsilon - \delta$ Criterion:** For every $\epsilon > 0$, there exists a $\delta > 0$ such that for all $y \in A$:

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

Proof. We proceed by showcasing the cyclic implications $(i) \implies (ii) \implies (iii) \implies (i)$.

$(i) \implies (ii)$ Assume that f preserves adherence. Let (x_n) be a sequence in A such that $x_n \to x$. We aim to show that $f(x_n) \to f(x)$. By the Subsubsequence Criterion, it suffices to show that every subsequence of $(f(x_n))$ admits a further subsequence converging to $f(x)$.

Let $(f(x_{n_k}))_k$ be an arbitrary subsequence of the image sequence. Consider the set of pre-images corresponding to this subsequence, $B = \{x_{n_k} : k \in \mathbb{N}\}$. Since the original sequence (x_n) converges to x, the subsequence (x_{n_k}) also converges to x. Consequently, $x \in \operatorname{adh}(B)$. By hypothesis (i), this implies $f(x) \in \operatorname{adh}(f(B))$.

The set $f(B)$ is precisely the set of values $\{f(x_{n_k}) : k \in \mathbb{N}\}$. Since $f(x)$ is adherent to this set, there exists a sequence (y_m) in $f(B)$ such that $y_m \to f(x)$. If $f(B)$ is finite, then the convergent sequence (y_m) must be eventually constant at $f(x)$. Thus, infinitely many terms of $(f(x_{n_k}))$ are equal to $f(x)$, forming a constant sub-subsequence converging to $f(x)$. If $f(B)$ is infinite, we construct a sub-subsequence by indices. Since $y_m \in f(B)$, for each m, $y_m = f(x_{n_{k_m}})$ for some index k_m. If the indices (k_m) are not strictly increasing, we pass to a subsequence where they are (possible since $\mathbb{N}$ is well-ordered). This yields a sub-subsequence of $(f(x_{n_k}))$ converging to $f(x)$. Thus, by the Subsubsequence Criterion, $f(x_n) \to f(x)$.

$(ii) \implies (iii)$ We proceed by contrapositive. Assume that statement (iii) is false. The negation of (iii) is:

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

Let us fix this specific "bad" ϵ. Since the condition holds for *all* δ, we may choose a sequence of $\delta_n = 1/n$. For each $n \in \mathbb{N}$, there exists a counter-example point $y_n \in A$ such that:

$$|x - y_n| < \frac{1}{n} \quad \text{and} \quad |f(x) - f(y_n)| \geq \epsilon$$

Consider the sequence (y_n). By the Squeeze theorem, $|x - y_n| \to 0$, so $y_n \to x$. However, the distance $|f(x) - f(y_n)|$ remains strictly bounded away from 0 (at least ϵ). Thus, $f(y_n)$ cannot converge to $f(x)$. This contradicts statement (ii). Therefore, (ii) must imply (iii).

$(iii) \implies (i)$ Assume the $\epsilon - \delta$ condition holds. Let $B \subseteq A$ such that $x \in \text{adh}(B)$. We must show $f(x) \in \text{adh}(f(B))$. Recall that $z \in \text{adh}(S)$ if and only if every neighbourhood of z intersects S (section 3.1).

Let $\epsilon > 0$ be given. We need to show that the neighbourhood $(f(x) - \epsilon, f(x) + \epsilon)$ contains a point from $f(B)$. By hypothesis (iii), there exists $\delta > 0$ such that mapping the δ-neighbourhood of x lands inside the ϵ-neighbourhood of $f(x)$:

$$f(A \cap B_\delta(x)) \subseteq B_\epsilon(f(x))$$

Since x is adherent to B, the intersection $B \cap B_\delta(x)$ is non-empty. Let y be a point in this intersection. Then $y \in B$, so $f(y) \in f(B)$. Also $y \in B_\delta(x)$, so by the implication above, $f(y) \in B_\epsilon(f(x))$. Thus, $f(y) \in f(B) \cap B_\epsilon(f(x))$. Since ϵ was arbitrary, $f(x)$ is adherent to $f(B)$.

■

While topological and sequential definitions offer theoretical utility, the $\epsilon - \delta$ criterion provides the standard analytic definition, quantifying stability.

Definition 4.1.1. ***Continuity at a Point***. Let $f : A \to \mathbb{R}$ and let $c \in A$. We say that f is continuous at c if it satisfies the condition:

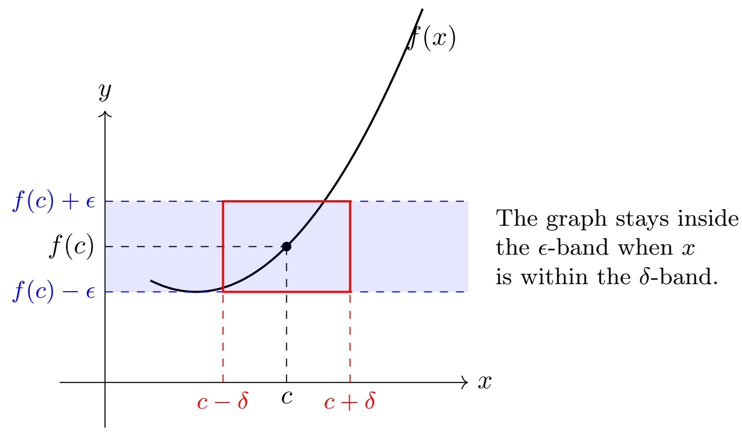
$$\forall \epsilon > 0, \exists \delta > 0 \text{ such that } \forall x \in A, (|x - c| < \delta \implies |f(x) - f(c)| < \epsilon)$$

If f is continuous at every point $c \in A$, we say f is continuous on A.

Note. δ generally depends on both the tolerance ϵ and the point c. We denote this dependency as $\delta(\epsilon, c)$. If δ can be chosen independently of c, the function possesses a stronger property known as *uniform continuity*, which we shall explore in later chapters.

Visualising Continuity

The definition can be interpreted geometrically using "boxes". To claim continuity at c, for any horizontal band of width 2ϵ centred at $f(c)$, we must be able to define a vertical strip of width 2δ centred at c such that the graph of the function restricted to this strip lies entirely within the horizontal band.



Continuity on Discrete Sets A surprising consequence of the definition is the behaviour of functions on isolated points.

Example 4.1.1. Isolated Continuity. Let $A = \mathbb{N}$ and define $f : \mathbb{N} \to \mathbb{R}$ by $f(n) = n^2$. Is f continuous at $c = 1$?

Analysis: We apply the definition. Let $\epsilon > 0$ be arbitrary (e.g., $\epsilon = 0.001$). We need to find $\delta > 0$ such that for all $n \in \mathbb{N}$, $|n - 1| < \delta \implies |n^2 - 1| < 0.001$. If we choose $\delta = 0.5$, the only natural number n satisfying $|n - 1| < 0.5$ is $n = 1$ itself. For $n = 1$, the condition becomes $|1^2 - 1| = 0 < 0.001$, which is trivially true. Thus, the condition holds. In fact, any function defined on a set consisting solely of isolated points is continuous everywhere. This aligns with definition 3.2.3: $\{c\}$ contains a neighbourhood around itself relative to A.

Algebra of Continuous Functions

Just as with limits of sequences, continuity is preserved under standard algebraic operations. This allows us to build complex continuous functions from simple building blocks (like polynomials).

***Theorem* 4.1.2. Algebraic Continuity**. Let $f, g : A \to \mathbb{R}$ be continuous at $c \in A$, and let $\lambda \in \mathbb{R}$. Then the following functions are continuous at c:

1. $f + g$
2. λf
3. fg
4. f/g (provided $g(c) \neq 0$)

Proof. These follow immediately from the sequential characterisation (theorem 4.1.1) and the Algebraic Limit Laws for sequences. For instance, to prove (3): Let (x_n) be a sequence in A with $x_n \to c$. Since f is continuous, $f(x_n) \to f(c)$. Since g is continuous, $g(x_n) \to g(c)$. By the Product Law for sequences, $f(x_n)g(x_n) \to f(c)g(c)$. Thus, the product function maps convergent sequences to convergent sequences, satisfying condition (ii). ■

***Theorem* 4.1.3. Composition of Continuous Functions**. Let $f : A \to B$ and $g : B \to \mathbb{R}$. If f is continuous at $c \in A$ and g is continuous at $f(c) \in B$, then the composition $g \circ f : A \to \mathbb{R}$ is continuous at c.

Proof. Let $x_n \to c$ in A. By the continuity of f, we have $f(x_n) \to f(c)$. Let $y_n = f(x_n)$ and $L = f(c)$. Then $y_n \to L$ is a sequence in B. By the continuity of g at L, we have $g(y_n) \to g(L)$. Substituting back, $g(f(x_n)) \to g(f(c))$. Thus, $g \circ f$ is continuous at c. ■

These theorems imply that all polynomials, rational functions (on their domains), and compositions thereof are continuous. The "nightmare" of $\epsilon - \delta$ need only be faced when proving the continuity of fundamental transcendental functions or pathological examples; for most of analysis, we rely on these algebraic properties.

4.2 The $\epsilon - \delta$ Formalism

While the sequential definition aligns naturally with our intuition of "approaching" a point, it is the $\epsilon - \delta$ formulation (Weierstrass's definition) that constitutes the standard analytic definition. This definition often presents a conceptual hurdle because it reverses the natural flow of a function. A function maps domain to codomain ($x \mapsto f(x)$), yet the definition prescribes a constraint on the codomain (ϵ) to determine a constraint on the domain (δ).

The Challenge of Approximation Recall the definition: f is continuous at c if:

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

The structural difference between this and the sequential definition is the primacy of ϵ. We do not ask "what happens when x is close to c?"; instead, we ask "how close must x be to c to ensure $f(x)$ is within a specific tolerance of $f(c)$?"

The Game Theoretic Interpretation Continuity can be modelled as a game played between a Challenger and a Defender.

1. **The Challenge:** The Challenger chooses a tolerance $\epsilon > 0$.
2. **The Response:** The Defender must calculate a precision $\delta > 0$. This δ acts as a shield; the Defender claims, "Yes, provided the input x deviates from c by less than δ."
3. **The Verification:** If for every x in the δ-neighbourhood, the condition $|f(x) - f(c)| < \epsilon$ holds, the Defender wins that round.

For a function to be continuous, the Defender must possess a winning strategy for *any* ϵ the Challenger proposes, no matter how infinitesimal.

Controlling Gradient via Domain Restriction Consider a function that increases rapidly near c, such as $f(x) = 10^9 x$ near $c = 0$. If the Challenger sets $\epsilon = 1$, the requirement $|f(x) - f(c)| < 1$ implies $|10^9 x| < 1$. Qualitative proximity is insufficient. The steeper the gradient of the function, the smaller the δ required to satisfy a fixed ϵ. Discontinuity arises only when no such δ exists — when the function jumps or oscillates so wildly that no amount of zooming in on the domain stabilises the output.

4.3 Limits of Functions

The concept of a limit $\lim_{x \to c} f(x) = L$ is subtler than continuity because it describes the behaviour of f *near* c, deliberately ignoring the value (or existence) of $f(c)$.

Definition 4.3.1. ***Limit of a Function***. Let $f : A \to \mathbb{R}$ and let c be a limit point of A. We say that the limit of f as x approaches c is L, denoted $\lim_{x \to c} f(x) = L$, if:

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

There are two crucial distinctions between this definition and the definition of continuity:

1. **The Punctured Neighbourhood:** The condition $0 < |x - c|$ explicitly excludes the case $x = c$. We do not care if $f(c)$ exists or if $f(c) = L$. The limit solely characterises the trend of the function as it approaches c.
2. **Restriction to Limit Points:** Unlike continuity, where c may be any point in A, limits require c to be a definition 3.1.2.

Remark. (Why Limit Points?). Suppose we allowed c to be an isolated point of A. By definition of isolation, there exists a $\delta_0 > 0$ such that the punctured neighbourhood $(c - \delta_0, c + \delta_0) \setminus \{c\}$ contains no points of A. Consequently, if we choose any $\delta < \delta_0$, the premise $x \in A \wedge 0 < |x - c| < \delta$ is false for all x. In classical logic, a conditional statement with a false antecedent is vacuously true. Thus, if c were an isolated point, $\lim_{x \to c} f(x) = L$ would be true for *every* real number L. To ensure the uniqueness of limits, we must restrict our attention to limit points, ensuring that the punctured neighbourhood is never empty.

Relationship between Limits and Continuity With the definition of a limit established, we can reframe continuity. As we shall see in theorem 4.6.2, continuity at a limit point is equivalent to the limit equalling the function value.

Remark. If c is an isolated point, the limit is undefined. However, the function is automatically continuous, aligning with the topological triviality of isolated points observed in proposition 3.3.2.

Discontinuity: Negating the Definition

To prove that a function is discontinuous, one must demonstrate that the $\epsilon - \delta$ condition fails. Logical negation of quantifiers is a frequent source of error; thus, we explicitly formulate the negation.

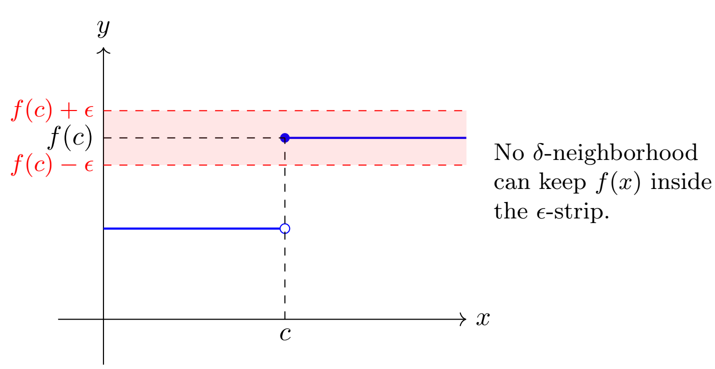
The definition of continuity at c is:

$$\forall \epsilon > 0, \exists \delta > 0, \forall x \in A, (|x - c| < \delta \implies |f(x) - f(c)| < \epsilon)$$

To negate this, we swap universal ($\forall$) and existential ($\exists$) quantifiers and negate the implication. Recall that $\neg(P \implies Q)$ is logically equivalent to $P \wedge \neg Q$.

Definition 4.3.2. ***Discontinuity at a Point***. A function $f : A \to \mathbb{R}$ is discontinuous at $c \in A$ if:

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



The Adversarial Interpretation In the language of our game:

1. The Challenger finds a specific "bad" tolerance ϵ_0.
2. The Defender tries to find a δ to satisfy this tolerance.
3. However, for *every* δ the Defender suggests, the Challenger can find a point x_δ that lies within the δ-shield ($|x_\delta - c| < \delta$) yet violates the tolerance condition ($|f(x_\delta) - f(c)| \geq \epsilon_0$).

This formulation is beneficial for proving the discontinuity of functions like the Dirichlet function or the signum function at 0. It asserts that there is a fixed barrier ϵ_0 such that points arbitrarily close to c map to values at least ϵ_0 away from $f(c)$.

4.4 Examples and Strategies for $\epsilon - \delta$ Proofs

We now turn to the practical application of the $\epsilon - \delta$ criterion. The transition from intuitive limits to inequalities can be jarring; but the key difficulty often lies not in the logic, but in the algebraic manipulation required to find the dependency $\delta(\epsilon)$. We illustrate the technique with two fundamental examples.

Example 4.4.1. Continuity of the Identity. Let $f : \mathbb{R} \to \mathbb{R}$ be defined by $f(x) = x$. We show f is continuous at every point $c \in \mathbb{R}$.

Proof. Fix $c \in \mathbb{R}$ and let $\epsilon > 0$. We seek a $\delta > 0$ such that $|x - c| < \delta \implies |f(x) - f(c)| < \epsilon$. Substituting the function definition, we require:

$$|x - c| < \epsilon$$

This suggests the choice $\delta = \epsilon$. If $|x - c| < \delta = \epsilon$, then trivially $|f(x) - f(c)| = |x - c| < \epsilon$. Thus, the identity function is continuous. ■

The previous example was deceptively simple because the relationship between input error and output error was linear with a slope of 1. Take a non-linear function.

Example 4.4.2. Continuity of the Square. Define $f : \mathbb{R} \to \mathbb{R}$ by $f(x) = x^2$. We show f is continuous at an arbitrary point $c \in \mathbb{R}$.

Proof.

Heuristic: We wish to bound $|x^2 - c^2| < \epsilon$. Factorising the difference:

$$|x^2 - c^2| = |x - c||x + c|$$

We can control the term $|x - c|$ directly with δ. However, the term $|x + c|$ depends on x. To formulate a valid proof, we must bound $|x + c|$ by a constant independent of the specific choice of x (within the neighbourhood).

Bounding Strategy: We may assume *a priori* that $\delta \leq 1$. If $|x - c| < 1$, then by the Triangle Inequality:

$$|x| = |(x - c) + c| \leq |x - c| + |c| < 1 + |c|$$

Consequently, we can bound the problematic term:

$$|x + c| \leq |x| + |c| < (1 + |c|) + |c| = 1 + 2|c|$$

This bound $K = 1 + 2|c|$ depends only on the fixed point c, not on the variable x.

Formal Argument: Let $\epsilon > 0$ be given. Choose $\delta = \min\left(1, \frac{\epsilon}{1+2|c|}\right)$. Let $x \in \mathbb{R}$ satisfy $|x - c| < \delta$.

1. Since $\delta \leq 1$, we have $|x + c| < 1 + 2|c|$.
2. Since $\delta \leq \frac{\epsilon}{1+2|c|}$, we have $|x - c| < \frac{\epsilon}{1+2|c|}$.

Multiplying these inequalities:

$$|x^2 - c^2| = |x - c||x + c| < \left(\frac{\epsilon}{1 + 2|c|}\right)(1 + 2|c|) = \epsilon$$

Thus, f is continuous at c.

■

Remark. The strategy employed above is a standard technique in analysis. We formally summarise this approach below.

A Template for Continuity Proofs

To prove $\lim_{x \to c} f(x) = L$ or continuity at c, follow these steps:

1. **Setup:** "Let $\epsilon > 0$ be given."
2. **Algebraic Simplification:** Manipulate $|f(x) - L|$ to factor out the term $|x - c|$.

$$|f(x) - L| = |x - c| \cdot G(x)$$

3. **Preliminary Bound:** Assume $|x - c| < 1$ (or some other fixed constant) to bound the spurious factor $|G(x)|$ by some constant K depending only on c.
4. **Selection of δ:** Define $\delta = \min(1, \epsilon/K)$.
5. **Verification:** demonstrate that this choice satisfies the condition.

4.5 Limit Laws for Functions

While the $\epsilon - \delta$ definition is the bedrock of proof, computing limits from first principles for every function is tedious and impractical. Just as we did for sequences, we build a set of Limit Laws that allow us to compute limits of complex expressions by breaking them down into simpler components.

Sequential Characterisation of Limits

We begin by explicitly linking functional limits to sequential limits. This connection allows us to import the entire machinery of Analysis I (Algebraic Limit Laws for sequences) into the functional setting without repeating the arduous $\epsilon - \delta$ arithmetic.

***Proposition* 4.5.1.** *Sequential Criterion for Limits.* Let $f : A \to \mathbb{R}$ and let c be a limit point of A. Then:

$$\lim_{x \to c} f(x) = L$$

if and only if for every sequence (x_n) in $A \setminus \{c\}$ such that $\lim_{n \to \infty} x_n = c$, we have:

$$\lim_{n \to \infty} f(x_n) = L$$

Remark. The requirement $x_n \in A \setminus \{c\}$ corresponds to the punctured neighbourhood $0 < |x - c|$ in the definition of a limit. We are strictly forbidden from sampling the function at c itself.

The proof of this proposition is structurally identical to the proof of The Three Faces of Continuity. We leave the details as a crucial exercise in adapting proofs.

Algebraic Limit Laws

Using the sequential criterion, we immediately deduce the arithmetic properties of functional limits.

***Theorem* 4.5.1. Limit Laws.** Let $f, g : A \to \mathbb{R}$ be functions, and let c be a limit point of A. Suppose that limits exist:

$$\lim_{x \to c} f(x) = L \quad \text{and} \quad \lim_{x \to c} g(x) = M$$

Then:

(i) **Sum Law:**

$$\lim_{x \to c}(f(x) + g(x)) = L + M$$

.

(ii) **Scalar Product:** For any $k \in \mathbb{R}$,

$$\lim_{x \to c}(kf(x)) = kL$$

.

(iii) **Product Law:**

$$\lim_{x \to c}(f(x)g(x)) = LM$$

.

(iv) **Quotient Law:** If $M \neq 0$, then

$$\lim_{x \to c}\frac{f(x)}{g(x)} = \frac{L}{M}$$

.

Proof. We illustrate the proof for the Product Law (iii). Let (x_n) be any sequence in $A \setminus \{c\}$ converging to c. By the sequential criterion:

$$\lim_{n \to \infty} f(x_n) = L \quad \text{and} \quad \lim_{n \to \infty} g(x_n) = M$$

By the Product Law for sequences:

$$\lim_{n \to \infty}(f(x_n)g(x_n)) = \left(\lim_{n \to \infty} f(x_n)\right)\left(\lim_{n \to \infty} g(x_n)\right) = LM$$

Since this holds for *any* such sequence, the functional limit is LM. The other parts follow identically. ■

The Quotient Condition The Quotient Law requires a subtle verification. For the function f/g to even be defined near c, we must ensure $g(x) \neq 0$ in some neighbourhood of c.

***Lemma* 4.5.1.** *Locally Non-Vanishing.* If $\lim_{x \to c} g(x) = M$ and $M \neq 0$, then there exists a radius $r > 0$ such that $g(x) \neq 0$ for all $x \in A \cap B_r(c)$ (with $x \neq c$).

Proof. Let $\epsilon = |M|/2$. Since $M \neq 0$, $\epsilon > 0$. By the definition of the limit, there exists $\delta > 0$ such that for $0 < |x - c| < \delta$:

$$|g(x) - M| < \frac{|M|}{2}$$

By the Reverse Triangle Inequality, $|M| - |g(x)| < |M|/2$, which implies $|g(x)| > |M|/2 > 0$. Thus, $g(x)$ is non-zero in this neighbourhood, and the quotient function is well-defined. ■

Continuity of Polynomials and Rational Functions

The Limit Laws allow us to construct a vast library of continuous functions without reverting to $\epsilon - \delta$ definitions.

***Corollary* 4.5.1.** *Polynomial Continuity.* Every polynomial function $P(x) = a_n x^n + \cdots + a_1 x + a_0$ is continuous on $\mathbb{R}$.

Proof. We proceed by induction on the complexity of the function.

1. **Constants:** The constant function $f(x) = k$ is trivially continuous (take any δ).
2. **Identity:** We proved in the previous section that $f(x) = x$ is continuous.
3. **Powers:** By the Product Law, since x is continuous, $x \cdot x = x^2$ is continuous. By induction, x^n is continuous for all $n \in \mathbb{N}$.
4. **Linear Combinations:** By the Scalar Product and Sum Laws, any linear combination $a_n x^n + \cdots + a_0$ is continuous.

■

***Corollary* 4.5.2.** *Rational Continuity.* Any rational function $R(x) = \frac{P(x)}{Q(x)}$ (where P, Q are polynomials) is continuous at every point c in its domain (i.e., where $Q(c) \neq 0$).

Proof. Since P and Q are continuous, $\lim_{x\to c} P(x) = P(c)$ and $\lim_{x\to c} Q(x) = Q(c)$. If c is in the domain, $Q(c) \neq 0$. By the Quotient Law:

$$\lim_{x\to c} \frac{P(x)}{Q(x)} = \frac{P(c)}{Q(c)} = R(c)$$

Thus, R is continuous at c. ■

4.6 The Squeeze Theorem and Trigonometric Limits

In the analysis of sequences, the Squeeze Theorem provided a robust method for determining the limit of a sequence by bounding it between two other sequences converging to the same value. This principle extends naturally to functions.

***Theorem* 4.6.1. The Squeeze Theorem for Functions.** Let $f, g, h : A \to \mathbb{R}$ be functions and let c be a limit point of A. Suppose that there exists a neighbourhood $B_r(c)$ such that for all $x \in A \cap B_r(c)$ (with $x \neq c$):

$$f(x) \leq g(x) \leq h(x)$$

If $\lim_{x\to c} f(x) = L$ and $\lim_{x\to c} h(x) = L$, then $\lim_{x\to c} g(x)$ exists and is equal to L.

Proof. Let $\epsilon > 0$. We must find a $\delta > 0$ such that $0 < |x - c| < \delta$ implies $|g(x) - L| < \epsilon$.

1. Since $\lim_{x\to c} f(x) = L$, there exists $\delta_1 > 0$ such that for $0 < |x - c| < \delta_1$, $|f(x) - L| < \epsilon$, which implies $L - \epsilon < f(x) < L + \epsilon$.
2. Since $\lim_{x\to c} h(x) = L$, there exists $\delta_2 > 0$ such that for $0 < |x - c| < \delta_2$, $|h(x) - L| < \epsilon$, which implies $L - \epsilon < h(x) < L + \epsilon$.

Let $\delta = \min(\delta_1, \delta_2, r)$. For any $x \in A$ satisfying $0 < |x - c| < \delta$, we have:

$$L - \epsilon < f(x) \leq g(x) \leq h(x) < L + \epsilon$$

Thus, $L - \epsilon < g(x) < L + \epsilon$, or $|g(x) - L| < \epsilon$. ■

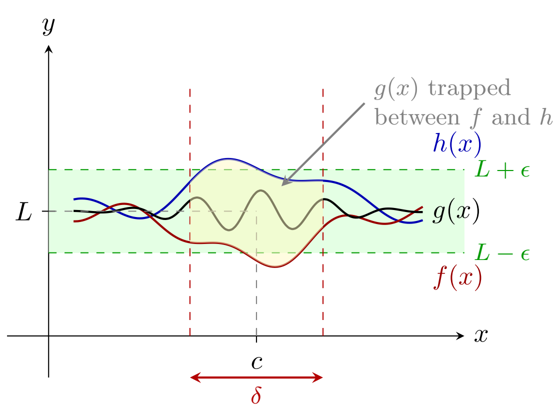


Figure 4.1: The Squeeze Theorem.

Remark. One may also prove this theorem by invoking the Sequential Criterion for limits and applying the Squeeze Theorem for sequences. While unification of these concepts via filter convergence is possible (as explored by A.F. Beardon and others), the direct $\epsilon - \delta$ proof suffices for our purposes.

The Fundamental Trigonometric Limit

The most celebrated application of the Squeeze Theorem is the evaluation of $\lim_{x\to 0} \frac{\sin x}{x}$. This limit is pivotal in describing the derivatives of trigonometric functions. To define the sine and cosine functions requires power series, a topic we shall treat in subsequent chapters. For the present, we accept standard trigonometric properties — specifically the bounding inequality derived from the geometry of the unit circle.

***Proposition* 4.6.1.** *Fundamental Trigonometric Limit.*

$$\lim_{x\to 0} \frac{\sin x}{x} = 1$$

Proof. We rely on the following geometric inequality, which holds for all $x \in (-\pi/2, \pi/2) \setminus \{0\}$:

$$\cos x < \frac{\sin x}{x} < 1$$

Let $f(x) = \cos x$, $g(x) = \frac{\sin x}{x}$, and $h(x) = 1$. As $x \to 0$, we know that $\lim_{x\to 0} \cos x = \cos(0) = 1$ (by the continuity of cosine proved later via power series) and clearly $\lim_{x\to 0} 1 = 1$. Since the limit of the bounding functions is 1, by the Squeeze Theorem,

$$\lim_{x\to 0} \frac{\sin x}{x} = 1$$

. ∎

Continuity via Limits

Throughout this chapter, we have noted the structural similarity between the definition of the limit, $\lim_{x\to c} f(x) = L$, and the definition of continuity at c. The definition of a limit explicitly excludes the value at c (requiring $0 < |x - c|$), while continuity relies on the value at c. We now formalise the relationship between these two concepts.

***Theorem* 4.6.2. Limit Characterisation of Continuity.** Let $f : A \to \mathbb{R}$ be a function and let $c \in A$ be a limit point of A. Then f is continuous at c if and only if:

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

Proof. This follows directly from comparing the logical structures of the definitions definition 4.3.1 and definition 4.1.1.

($\Rightarrow$) Suppose f is continuous at c. By definition, for every $\epsilon > 0$, there exists $\delta > 0$ such that for all $x \in A$:

$$|x - c| < \delta \implies |f(x) - f(c)| < \epsilon$$

This implication holds for all x in the δ-neighbourhood. Specifically, it holds for all x in the *punctured* neighbourhood where $0 < |x - c| < \delta$. Thus, the condition for $\lim_{x\to c} f(x) = f(c)$ is satisfied.

($\Leftarrow$) Suppose $\lim_{x\to c} f(x) = f(c)$. By definition, for every $\epsilon > 0$, there exists $\delta > 0$ such that for all $x \in A$:

$$0 < |x - c| < \delta \implies |f(x) - f(c)| < \epsilon$$

To establish continuity, we must show the condition holds for *all* x with $|x - c| < \delta$. The only point excluded by the limit condition is $x = c$. However, at $x = c$, we have $|f(c) - f(c)| = 0 < \epsilon$, which is trivially true. Thus, the condition holds for the entire neighbourhood, and f is continuous at c.

∎

Note. This theorem allows us to calculate limits by simple substitution, provided we know the function is continuous. For example, $\lim_{x\to 2}(x^2 + 3) = 2^2 + 3 = 7$ is justified precisely because polynomials are continuous functions.

4.7 Exercises

In the following exercises, unless otherwise stated, functions are defined on subsets of $\mathbb{R}$. You may rely on the algebraic limit laws and the sequential criterion where appropriate, but strictly adhere to the $\epsilon - \delta$ definition when explicitly requested.

1. **Basic $\epsilon - \delta$ Construction.** For each of the following functions, propose a candidate limit L as $x \to c$, and prove the limit using the $\epsilon - \delta$ definition. Explicitly define δ in terms of ϵ.
   (a) $f(x) = 3x - 7$ at $c = 2$.
   (b) $g(x) = x^2 + x$ at $c = 3$.
   (c) $h(x) = \frac{1}{x}$ at $c = 1$.

2. **Piecewise Parameters.** Determine the values of the constants a and b that make the function $f : \mathbb{R} \to \mathbb{R}$ continuous everywhere.
$$f(x) = \begin{cases} \frac{x^2-4}{x-2} & \text{if } x < 2 \\ ax^2 - bx + 3 & \text{if } 2 \le x < 3 \\ 2x - a + b & \text{if } x \ge 3 \end{cases}$$

3. **Algebraic Limit Calculations.** Evaluate the following limits. If a limit does not exist, explain why.
   (a) $\lim\limits_{x\to 1} \frac{x^2-3x+2}{x^3-3x^2+2}$.
   (b) $\lim\limits_{x\to 0} \frac{\sqrt{1+\tan x}-\sqrt{1+\sin x}}{x^3}$.
   (c) $\lim\limits_{x\to 0} \frac{1-\cos x\cos(2x)\cos(3x)}{x^2}$.
   (d) $\lim\limits_{x\to 1} \frac{\frac{1}{\sqrt{x}}+\sqrt{x}-2}{(x-1)^2}$.

4. **The Damped Oscillator.** Let $f : \mathbb{R} \to \mathbb{R}$ be defined by:
$$f(x) = \begin{cases} x\sin(1/x) & \text{if } x \neq 0 \\ 0 & \text{if } x = 0 \end{cases}$$
   (a) Prove that f is continuous for all $x \in \mathbb{R}$.
   Remark. For $x = 0$, use the Squeeze Theorem.
   (b) Sketch the graph of f, paying particular attention to the envelope lines $y = x$ and $y = -x$.

5. **Rational Mapping Properties.** Consider the function $f(x) = \frac{x^2-6x+5}{x^2-9x+18}$.
   (a) Determine the domain of continuity of f.
   (b) Prove that if $k \neq 1$, there are exactly two distinct values of x for which $f(x) = k$.
   (c) Illustrate this result by sketching the graph of f and the line $y = k$.

6. **Bounds of Rational Functions.** Let $f(x) = \frac{1}{x^2+1}$.
   (a) Prove that f is bounded on $\mathbb{R}$.
   (b) Find, with proof, $\sup\{f(x) : x \in \mathbb{R}\}$ and $\inf\{f(x) : x \in \mathbb{R}\}$.
   (c) Answer the same questions for the function $g(x) = \frac{4x^2+3}{x^4+1}$.

7. **Sequential Discontinuity.** Consider the signum function:
$$\operatorname{sgn}(x) = \begin{cases} 1 & \text{if } x > 0 \\ 0 & \text{if } x = 0 \\ -1 & \text{if } x < 0 \end{cases}$$
   Use the Sequential Criterion (construct a specific sequence $x_n \to 0$) to prove that sgn is discontinuous at 0. Does the limit $\lim\limits_{x\to 0} \operatorname{sgn}(x)$ exist?

8. **Roots and Rationalisation.** Evaluate $\lim_{x\to 4} \frac{\sqrt{x}-2}{x-4}$. Prove your assertion using the $\epsilon - \delta$ definition.

9. **Local Positivity.** Let f be continuous at c and suppose $f(c) > 0$. Prove that there exists a neighbourhood $B_\delta(c)$ such that $f(x) > 0$ for all $x \in B_\delta(c)$.

   Remark. Apply the definition of continuity with a specific choice of ϵ related to $f(c)$.

10. **The Inverse Image Characterisation.** In the text, we characterised continuity via adherence preservation. A more standard topological definition involves open sets. Prove that a function $f : \mathbb{R} \to \mathbb{R}$ is continuous on $\mathbb{R}$ if and only if for every open set $V \subseteq \mathbb{R}$, the pre-image $f^{-1}(V) = \{x \in \mathbb{R} \mid f(x) \in V\}$ is an open set.

11. **Density of the Rationals.** Let $f, g : \mathbb{R} \to \mathbb{R}$ be continuous functions.

    (a) Prove that if $f(q) = 0$ for all rational numbers $q \in \mathbb{Q}$, then $f(x) = 0$ for all $x \in \mathbb{R}$.
    (b) Conclude that if $f(q) = g(q)$ for all $q \in \mathbb{Q}$, then $f(x) = g(x)$ everywhere.

    Remark. This result asserts that a continuous function is entirely determined by its values on the rational numbers.

12. **Limits of Compositions.** In the text, we proved that the composition of continuous functions is continuous. However, limits are more fragile. Suppose $\lim_{x\to c} f(x) = L$ and $\lim_{y\to L} g(y) = M$.

    (a) Construct a counter-example to show that $\lim_{x\to c} g(f(x)) = M$ is not necessarily true.

        Remark. What if g is discontinuous at L, and f assumes the value L infinitely often near c?

    (b) State and prove a condition on f or g that suffices to make the statement true.

13. **The Dirichlet Function.** Consider the function $D : \mathbb{R} \to \mathbb{R}$ defined by:

$$D(x) = \begin{cases} 1 & \text{if } x \in \mathbb{Q} \\ 0 & \text{if } x \notin \mathbb{Q} \end{cases}$$

    Prove that D is discontinuous at every point $c \in \mathbb{R}$.

    Remark. Use the density of rationals and irrationals. For any δ-neighbourhood, find points mapping to 1 and points mapping to 0.

14. **Thomae's Function (The Popcorn Function).** Consider the function $T : (0, 1) \to \mathbb{R}$ defined by:

$$T(x) = \begin{cases} 0 & \text{if } x \notin \mathbb{Q} \\ 1/q & \text{if } x = p/q \text{ in lowest terms } (\gcd(p, q) = 1) \end{cases}$$

    (a) Prove that $\lim_{x\to c} T(x) = 0$ for every $c \in (0, 1)$.

        Remark. For a given ϵ, how many rational numbers in $(0, 1)$ have denominators q small enough such that $1/q \geq \epsilon$? Is this set finite?

    (b) Conclude that T is continuous at every irrational number but discontinuous at every rational number.

Chapter 5

Continuity Continued

We extend the local definition of continuity (definition 4.1.1) to the entire domain, considering functions that exhibit stability globally.

5.1 Global Continuity and Topological Characterisation

Definition 5.1.1. ***Continuous Function***. Let $f : A \to \mathbb{R}$ be a function. We say that f is a continuous function (or continuous on A) if f is continuous at every point $x \in A$.

To study global continuity, we recast the $\epsilon - \delta$ definition in the language of neighbourhoods. Recall that $B_r(p) = (p - r, p + r)$.

***Proposition* 5.1.1.** *Metric Characterisation of Continuity.* Let $f : A \to \mathbb{R}$ and let $c \in A$. Then f is continuous at c if and only if for every ϵ-neighbourhood $B_\epsilon(f(c))$, there exists a δ-neighbourhood $B_\delta(c)$ such that:

$$f(A \cap B_\delta(c)) \subseteq B_\epsilon(f(c))$$

Proof. This is a set-theoretic restatement of definition 4.1.1. The condition $|x - c| < \delta \implies |f(x) - f(c)| < \epsilon$ is equivalent to $x \in B_\delta(c) \implies f(x) \in B_\epsilon(f(c))$, or $f(A \cap B_\delta(c)) \subseteq B_\epsilon(f(c))$. ■

The Topological Characterisation

The following theorem expresses continuity solely in terms of open sets, justifying the study of general topology.

***Theorem* 5.1.1. Topological Continuity**. Let $A \subseteq \mathbb{R}$ be an open set and let $f : A \to \mathbb{R}$. Then f is continuous on A if and only if for every open set $V \subseteq \mathbb{R}$, the pre-image $f^{-1}(V)$ is an open subset of A.

$$f \text{ is continuous} \iff \forall V \text{ open in } \mathbb{R}, f^{-1}(V) \text{ is open in } A$$

Note. The pre-image is defined as $f^{-1}(V) = \{x \in A : f(x) \in V\}$. This relies solely on the function f, not on the existence of an inverse function.

Proof. We prove both directions of the equivalence.

(⇒) Assume f is continuous. Let $V \subseteq \mathbb{R}$ be an open set. We must show that $U = f^{-1}(V)$ is open. If $U = \emptyset$, it is trivially open. Suppose $U \neq \emptyset$. Let $x \in U$. By definition, $f(x) \in V$. Since V is open, there exists $\epsilon > 0$ such that $B_\epsilon(f(x)) \subseteq V$.

By proposition 5.1.1, there exists $\delta > 0$ such that:

$$f(A \cap B_\delta(x)) \subseteq B_\epsilon(f(x)) \subseteq V$$

Since A is open, we may shrink δ if necessary to ensure $B_\delta(x) \subseteq A$. Thus, $f(B_\delta(x)) \subseteq V$. This implies that for all $y \in B_\delta(x)$, $f(y) \in V$, so $y \in f^{-1}(V) = U$. Consequently, $B_\delta(x) \subseteq U$. Since x was arbitrary, every point in U is an interior point. Thus U is open.

($\Leftarrow$) Assume pre-images of open sets are open. We show f is continuous at an arbitrary point $c \in A$. Let $\epsilon > 0$. Let $V = B_\epsilon(f(c))$. By hypothesis, the set $U = f^{-1}(V)$ is open in A. This means $U = W \cap A$ for some open set $W \subseteq \mathbb{R}$. Since $c \in U$, we have $c \in W$ and $c \in A$. Since W and A are both open in $\mathbb{R}$, their intersection U is also open in $\mathbb{R}$. Thus, c is an interior point of U relative to $\mathbb{R}$; there exists $\delta > 0$ such that $B_\delta(c) \subseteq U$. Tracing the definitions backward:

$$B_\delta(c) \subseteq f^{-1}(B_\epsilon(f(c))) \implies f(B_\delta(c)) \subseteq B_\epsilon(f(c))$$

This satisfies proposition 5.1.1 at c.

■

Remark. If the domain A is not an open set (e.g., $A = [a, b]$), the theorem holds if we use the definition of open sets *relative* to A (subspace topology). Specifically, f is continuous iff $f^{-1}(V) = U \cap A$ for some open set $U \subseteq \mathbb{R}$.

We conclude this section by demonstrating the continuity of the square root function using the $\epsilon - \delta$ definition.

Example 5.1.1. Continuity of $\sqrt{x}$. Let $f : [0, \infty) \to \mathbb{R}$ be defined by $f(x) = \sqrt{x}$. We prove f is continuous on its domain. We divide the proof into two cases: boundary ($c = 0$) and interior ($c > 0$).

Case 1: $c = 0$. Let $\epsilon > 0$. We seek $\delta > 0$ such that for $x \geq 0$:

$$|x - 0| < \delta \implies |\sqrt{x} - 0| < \epsilon$$

This simplifies to $x < \delta \implies \sqrt{x} < \epsilon$. Choosing $\delta = \epsilon^2$, we see that if $0 \leq x < \epsilon^2$, then $\sqrt{x} < \epsilon$. Thus, f is continuous at 0.

Case 2: $c > 0$. Let $\epsilon > 0$. We wish to bound $|\sqrt{x} - \sqrt{c}|$. Using the algebraic identity $(\sqrt{x} - \sqrt{c})(\sqrt{x} + \sqrt{c}) = x - c$, we write:

$$|\sqrt{x} - \sqrt{c}| = \frac{|x - c|}{\sqrt{x} + \sqrt{c}}$$

Since $x \geq 0$, we have $\sqrt{x} \geq 0$. Therefore, the denominator satisfies $\sqrt{x} + \sqrt{c} \geq \sqrt{c}$. Taking the reciprocal reverses the inequality:

$$\frac{1}{\sqrt{x} + \sqrt{c}} \leq \frac{1}{\sqrt{c}}$$

Consequently:

$$|\sqrt{x} - \sqrt{c}| \leq \frac{|x - c|}{\sqrt{c}}$$

We define $\delta = \epsilon\sqrt{c}$. If $|x - c| < \delta$, then:

$$|\sqrt{x} - \sqrt{c}| < \frac{\epsilon\sqrt{c}}{\sqrt{c}} = \epsilon$$

Thus, f is continuous at c.

5.2 Topological Perspectives on Composition

We previously established the continuity of composite functions using sequences (theorem 4.1.3). We now provide a topological proof using neighbourhoods.

Let $h(x) = \sqrt{x^2+1}$. The domain of h is $\mathbb{R}$ since $x^2+1 \geq 1$. We have previously established that polynomials are continuous everywhere, and that the square root function is continuous on $[0, \infty)$. While theorem 4.1.1 allows us to deduce the continuity of sums and products directly, it does not immediately apply to the nested structure $g(f(x))$. We require a characterisation of composition operating directly on the topology of the domain and codomain.

***Theorem* 5.2.1. Metric Continuity of Composition**. Let $f : A \to B$ and $g : B \to \mathbb{R}$ be functions. If f is continuous at $c \in A$ and g is continuous at $f(c) \in B$, then the composite function $g \circ f$ is continuous at c.

Proof. We employ the method of open balls. Let $\epsilon > 0$ be given. We define a target neighbourhood around the image point $(g \circ f)(c)$:

$$V_\epsilon = B_\epsilon(g(f(c)))$$

We must find a neighbourhood U_δ around c such that $(g \circ f)(A \cap U_\delta) \subseteq V_\epsilon$.

Step 1: Continuity of g. Since g is continuous at the point $y = f(c)$, there exists a radius $\eta > 0$ such that:

$$g(B \cap B_\eta(y)) \subseteq V_\epsilon$$

That is, if a point is within η of $f(c)$, its image under g is within ϵ of $g(f(c))$.

Step 2: Continuity of f. We now treat η as the tolerance for the inner function f. Since f is continuous at c, there exists a $\delta > 0$ such that:

$$f(A \cap B_\delta(c)) \subseteq B_\eta(f(c))$$

In other words, if x is within δ of c, then $f(x)$ is within η of $f(c)$.

Combining these inclusions:

$$(g \circ f)(A \cap B_\delta(c)) = g(f(A \cap B_\delta(c))) \subseteq g(B \cap B_\eta(f(c))) \subseteq V_\epsilon$$

Thus, inputs δ-close to c yield outputs ϵ-close to $g(f(c))$, proving continuity. ■

Example 5.2.1. The Absolute Value of a Continuous Function. Let $f : A \to \mathbb{R}$ be a function continuous at $c \in A$. Consider $h(x) = |f(x)|$, which is the composition $g \circ f$ where $g(y) = |y|$. Since the absolute value function is continuous on $\mathbb{R}$, and f is continuous at c, the composition h is continuous at c.

Example 5.2.2. Continuity of Composed Functions. Let $h(x) = \sin(x^2)$. Here, $f(x) = x^2$ is a continuous power function and $g(y) = \sin(y)$ is a continuous trigonometric function. Theorem 5.1.1 implies that for any open set $V \subseteq \mathbb{R}$, the set $g^{-1}(V)$ is open, and subsequently the pre-image of that set under f, namely $f^{-1}(g^{-1}(V))$, is also open. Since the pre-image of any open set under h is open, h is a continuous function on $\mathbb{R}$.

5.3 The Language of Limits

As we progress to more complex arguments, it becomes convenient to adopt a linguistic shorthand that captures the essence of convergence. We formalise the notions of "sufficiently close" and "arbitrarily small".

Definition 5.3.1. *Sufficiently Close*. Let $S \subseteq \mathbb{R}$ and let c be a limit point of S. Let $P(x)$ be a logical property defined for elements of S. We say that $P(x)$ **holds for x sufficiently close to c** if there exists a $\delta > 0$ such that $P(x)$ is true for all $x \in S$ satisfying:

$$0 < |x - c| < \delta$$

Note. This definition excludes the point c, aligning with the definition of a limit. If we wish to include c (as in continuity), we would simply require $|x - c| < \delta$.

Definition 5.3.2. ***Arbitrarily Small***. Let $f : S \to \mathbb{R}$ be a function and let c be a limit point of S. We say that $f(x)$ **becomes arbitrarily small as** x **approaches** c if for every $\epsilon > 0$, the condition $|f(x)| < \epsilon$ holds for x sufficiently close to c. Equivalently:

$$\forall \epsilon > 0, \exists \delta > 0, \forall x \in S, (0 < |x - c| < \delta \implies |f(x)| < \epsilon)$$

By combining these definitions, we can restate the definition of a limit $\lim_{x \to c} f(x) = L$ concisely.

Remark. (The Linguistic Characterisation). The statement $\lim_{x \to c} f(x) = L$ is equivalent to asserting that:

The error $|f(x) - L|$ becomes arbitrarily small for x sufficiently close to c.

This phrasing allows us to reason about limits intuitively. For example, the Product Law can be summarised: if f is close to L and g is close to M, then fg is close to LM. Specifically, we can write $f(x) = L + \alpha(x)$ and $g(x) = M + \beta(x)$, where the error terms α and β become arbitrarily small. The product is:

$$f(x)g(x) = (L + \alpha(x))(M + \beta(x)) = LM + L\beta(x) + M\alpha(x) + \alpha(x)\beta(x)$$

Since α and β are arbitrarily small, their linear combination and product also become arbitrarily small (for x sufficiently close to c), proving the limit is LM.

***Proposition* 5.3.1.** *Caveat on Usage*. While this language is ubiquitous in mathematical literature, it is a tool for *description*, not *demonstration*. When constructing a proof, one must invariably "unpack" the terms "sufficiently close" and "arbitrarily small" back into their constituent δ and ϵ inequalities.

5.4 Infinite Limits

We extend the language to cases where values grow without bound. Instead of quantities becoming "arbitrarily small", we consider quantities becoming "arbitrarily large".

Definition 5.4.1. ***Unbounded Intervals***. Let $a \in \mathbb{R}$. We define the following sets:

$$[a, \infty) = \{x \in \mathbb{R} : x \geq a\}$$

$$(-\infty, a] = \{x \in \mathbb{R} : x \leq a\}$$

$$(-\infty, \infty) = \mathbb{R}$$

The symbols ∞ and $-\infty$ represent unboundedness and are not real numbers.

To discuss limits as $x \to \infty$, the domain of the function must allow x to become arbitrarily large.

Definition 5.4.2. ***Arbitrarily Large Sets***. A non-empty subset $S \subseteq \mathbb{R}$ is said to contain **arbitrarily large elements** if for every $N \in \mathbb{N}$, there exists $x \in S$ such that $x > N$. Similarly, S contains **arbitrarily large negative elements** if for every $N \in \mathbb{N}$, there exists $x \in S$ such that $x < -N$.

Limits at Infinity

The definition of a limit as $x \to \infty$ parallels the definition of a sequence limit $n \to \infty$.

Definition 5.4.3. ***Limit at Infinity***. Let $f : S \to \mathbb{R}$ be a function where S contains arbitrarily large elements. We say that the limit of f as x approaches infinity is L, denoted $\lim_{x \to \infty} f(x) = L$, if:

$$\forall \epsilon > 0, \exists N \in \mathbb{R} \text{ such that } \forall x \in S, (x > N \implies |f(x) - L| < \epsilon)$$

The definition for $\lim_{x \to -\infty} f(x) = L$ is analogous, requiring $x < -N$.

Remark. If the domain $S = \mathbb{N}$, this definition coincides exactly with the definition of convergence for a sequence (a_n). Thus, functional limits at infinity are a generalisation of sequential limits.

Example 5.4.1. Rational Decay. Let $f : (0, \infty) \to \mathbb{R}$ be defined by $f(x) = 1/x$. We show $\lim_{x\to\infty} 1/x = 0$. Let $\epsilon > 0$. By the Archimedean Property, choose $N > 1/\epsilon$. If $x > N$, then $x > 1/\epsilon$, which implies $0 < 1/x < \epsilon$. Thus $|f(x) - 0| < \epsilon$. It follows immediately that $\lim_{x\to-\infty} 1/x = 0$ as well.

Infinite Limits of Functions

We now consider the case where the function values themselves become unbounded. This corresponds to the notion of a vertical asymptote.

Definition 5.4.4. ***Divergence to Infinity***. Let $f : A \to \mathbb{R}$ and let c be a limit point of A.

1. We say f diverges to infinity at c, denoted $\lim_{x\to c} f(x) = \infty$, if for every $M > 0$, there exists $\delta > 0$ such that for all $x \in A$:
$$0 < |x - c| < \delta \implies f(x) > M$$
2. We say f diverges to minus infinity at c, denoted $\lim_{x\to c} f(x) = -\infty$, if for every $M > 0$, there exists $\delta > 0$ such that for all $x \in A$:
$$0 < |x - c| < \delta \implies f(x) < -M$$

In the language of the previous section: $\lim_{x\to c} f(x) = \infty$ means that $f(x)$ can be made *arbitrarily large* by keeping x *sufficiently close* to (but distinct from) c.

Example 5.4.2. Reciprocal Squared. Let $f(x) = 1/x^2$ defined on $\mathbb{R} \setminus \{0\}$. We claim $\lim_{x\to 0} 1/x^2 = \infty$. Let $M > 0$ be given (arbitrarily large). We seek δ such that $0 < |x| < \delta \implies 1/x^2 > M$. The inequality $1/x^2 > M$ is equivalent to $x^2 < 1/M$, or $|x| < 1/\sqrt{M}$. Choosing $\delta = 1/\sqrt{M}$, the implication holds. Note that for $g(x) = 1/x$, the limit at 0 does not exist in this sense because the values approach ∞ from the right and $-\infty$ from the left. This motivates the study of one-sided limits.

Remark. The algebraic limit laws can be extended to infinite limits, provided one avoids indeterminate forms (e.g., $\infty - \infty$, $0 \cdot \infty$, ∞/∞).

***Theorem* 5.4.1. Infinite Product Law**. Let $f, g : A \to \mathbb{R}$ and let c be a limit point of A. Suppose:

$$\lim_{x\to c} f(x) = L > 0 \quad \text{and} \quad \lim_{x\to c} g(x) = \infty$$

Then $\lim_{x\to c}(f(x)g(x)) = \infty$.

Proof. Let $K > 0$ be an arbitrary threshold for the product. We need to find δ such that the product exceeds K.

1. **Control f:** Since $f(x) \to L > 0$, the function eventually stays positive and bounded away from zero. Let $\epsilon = L/2$. There exists $\delta_1 > 0$ such that for $0 < |x - c| < \delta_1$, $|f(x) - L| < L/2$. This implies $f(x) > L/2$.
2. **Control g:** We need $g(x)$ to be large enough so that $(L/2)g(x) > K$. Thus we need $g(x) > 2K/L$. Since $g(x) \to \infty$, there exists $\delta_2 > 0$ such that for $0 < |x - c| < \delta_2$, $g(x) > 2K/L$.

Let $\delta = \min(\delta_1, \delta_2)$. For $0 < |x - c| < \delta$:

$$f(x)g(x) > \left(\frac{L}{2}\right)\left(\frac{2K}{L}\right) = K$$

Thus, the product diverges to ∞. ■

5.5 One-Sided Limits

The definition of a limit $\lim_{x \to c} f(x)$ considers behaviour as x approaches c from both sides simultaneously. In many contexts, it is necessary to restrict the approach to a single direction.

Definition 5.5.1. ***Right-Hand Limit*.** Let $f : A \to \mathbb{R}$ and let c be a limit point of the set $A \cap (c, \infty)$. The right-hand limit of f at c is the limit of the restriction of f to the domain $A \cap (c, \infty)$. We denote this by:

$$\lim_{x \to c^+} f(x) = L \quad \text{or} \quad f(c^+) = L$$

Formally: $\forall \epsilon > 0, \exists \delta > 0$ such that $\forall x \in A$, $(c < x < c + \delta \implies |f(x) - L| < \epsilon)$.

The **left-hand limit**, denoted $\lim_{x \to c^-} f(x)$, is defined analogously by restricting the domain to $A \cap (-\infty, c)$.

Remark. One-sided limits allow us to analyse "infinite limits" via a change of variable. For example, $\lim_{x \to \infty} f(x)$ is equivalent to $\lim_{t \to 0^+} f(1/t)$. This observation unifies the theory of limits at infinity with the theory of limits at a finite point.

The relationship between the global limit and the one-sided limits provides a powerful criterion for continuity.

***Theorem* 5.5.1. One-Sided Continuity Criterion.** Let $f : A \to \mathbb{R}$ and let c be a limit point of both $A \cap (c, \infty)$ and $A \cap (-\infty, c)$. Then f is continuous at c if and only if:

$$\lim_{x \to c^+} f(x) = \lim_{x \to c^-} f(x) = f(c)$$

Proof. We prove this both ways:

($\Rightarrow$) Suppose f is continuous at c. Then for any $\epsilon > 0$, there exists $\delta > 0$ such that $|x - c| < \delta \implies |f(x) - f(c)| < \epsilon$. The inequality holds for $x \in (c, c + \delta)$ (implies the Right-Hand Limit is $f(c)$) and for $x \in (c - \delta, c)$ (implies the Left-Hand Limit is $f(c)$).

($\Leftarrow$) Suppose both one-sided limits exist and equal $f(c)$. Let $\epsilon > 0$. There exists $\delta_1 > 0$ such that $c < x < c + \delta_1 \implies |f(x) - f(c)| < \epsilon$. There exists $\delta_2 > 0$ such that $c - \delta_2 < x < c \implies |f(x) - f(c)| < \epsilon$. Let $\delta = \min(\delta_1, \delta_2)$. Then for any x with $|x - c| < \delta$:
- If $x > c$, the first condition applies.
- If $x < c$, the second condition applies.
- If $x = c$, $|f(c) - f(c)| = 0 < \epsilon$.

Thus f is continuous at c.

∎

This theorem categorises discontinuities at a point c:

1. **Removable Discontinuity:** $\lim_{x \to c^+} f(x) = \lim_{x \to c^-} f(x) = L$, but $L \neq f(c)$ (or $f(c)$ is undefined). We can repair the function by defining $f(c) = L$.
2. **Jump Discontinuity:** The limits exist but $\lim_{x \to c^+} f(x) \neq \lim_{x \to c^-} f(x)$. No redefinition of $f(c)$ can make the function continuous (e.g., the Heaviside step function).
3. **Essential Discontinuity:** At least one of the one-sided limits does not exist (including divergence to $\pm\infty$).

Asymptotic Behaviour of Functions

In many applications, the exact value of a function is less important than its asymptotic behaviour—how it scales as the input grows arbitrarily large.

Polynomial Growth Polynomials form the backbone of approximation theory. Their behaviour at infinity is governed entirely by their "leading term". Intuitively, for very large x, the highest power of x dwarfs all lower powers combined.

***Theorem* 5.5.2. Limits of Polynomials**. Let $P(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_0$ be a polynomial of degree $n \geq 1$ with $a_n \neq 0$. Then:

$$\lim_{x\to\infty} P(x) = \begin{cases} \infty & \text{if } a_n > 0 \\ -\infty & \text{if } a_n < 0 \end{cases}$$

As $x \to -\infty$, the limit depends on the parity of the degree n:

$$\lim_{x\to-\infty} P(x) = \begin{cases} \infty & \text{if } n \text{ is even and } a_n > 0 \\ -\infty & \text{if } n \text{ is even and } a_n < 0 \\ -\infty & \text{if } n \text{ is odd and } a_n > 0 \\ \infty & \text{if } n \text{ is odd and } a_n < 0 \end{cases}$$

Proof. We factor x^n from the expression:

$$P(x) = x^n \left(a_n + \frac{a_{n-1}}{x} + \frac{a_{n-2}}{x^2} + \cdots + \frac{a_0}{x^n} \right)$$

We analyse the behaviour of the term in parentheses as $|x| \to \infty$. Using the algebraic limit laws and the fact that $\lim\limits_{x\to\pm\infty} \frac{1}{x^k} = 0$ for $k \geq 1$:

$$\lim_{x\to\pm\infty} \left(a_n + \sum_{k=1}^{n} \frac{a_{n-k}}{x^k} \right) = a_n + 0 = a_n$$

Consequently, the limit of the product behaves as $\lim\limits_{x\to\pm\infty} P(x) = \lim\limits_{x\to\pm\infty} (x^n \cdot a_n)$. The result then follows from the properties of power functions:

- As $x \to \infty$, $x^n \to \infty$. Thus $P(x) \to (\text{sgn } a_n)\infty$.
- As $x \to -\infty$, if n is even, $x^n \to \infty$. If n is odd, $x^n \to -\infty$. The sign of a_n then determines the final direction.

■

Example 5.5.1. Dominance at Infinity. Let $P(x) = -3x^5 + 1000x^4 + x$. Although the coefficient 1000 is large, as $x \to \infty$, the term $-3x^5$ dominates.

$$\lim_{x\to\infty} (-3x^5 + 1000x^4 + x) = \lim_{x\to\infty} x^5 \left(-3 + \frac{1000}{x} + \frac{1}{x^4} \right) = -\infty$$

Exponential Dominance Recall that factorials dominate exponentials, which in turn dominate polynomials. Specifically, an exponential function with base $a > 1$ grows faster than *any* polynomial, no matter how high the degree.

***Theorem* 5.5.3. Exponential vs. Polynomial Growth**. Let $P(x)$ be any non-constant polynomial and let $a > 1$. Then:

$$\lim_{x\to\infty} \frac{a^x}{|P(x)|} = \infty \quad \text{and} \quad \lim_{x\to\infty} \frac{P(x)}{a^x} = 0$$

Note. Since we established $\lim\limits_{n\to\infty} n^k/a^n = 0$ for integers, the monotonicity of a^x allows us to extend this to the continuum.

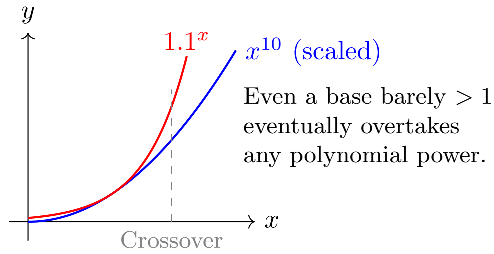


Figure 5.1: Competition between polynomial and exponential growth.

This result implies that in the "battle of infinities", exponentials are a higher order of infinity than polynomials. This has profound implications for the convergence of improper integrals and the complexity of algorithms.

Example 5.5.2. A Limit at Infinity. Compute $\lim_{x\to\infty} \frac{x^{100}}{e^{0.01x}}$. Let $t = 0.01x$. Then $x = 100t$. As $x \to \infty$, $t \to \infty$.

$$\lim_{x\to\infty} \frac{x^{100}}{e^{0.01x}} = \lim_{t\to\infty} \frac{(100t)^{100}}{e^t} = 100^{100} \lim_{t\to\infty} \frac{t^{100}}{e^t}$$

By the growth hierarchy, $t^{100}/e^t \to 0$. Thus, the limit is 0.

5.6 Applications: Curve Sketching

The analytical tools of infinite limits and continuity provide a systematic framework for determining the shape of a curve $y = f(x)$ without resorting to laborious point-plotting. By identifying the domain, symmetries, and asymptotic behaviour, we can construct an accurate qualitative graph.

- **Domain:** Identify values where f is undefined (e.g., division by zero, negative square roots).
- **Symmetry:**
  - If $f(-x) = f(x)$, the graph is symmetric about the y-axis (Even).
  - If $f(-x) = -f(x)$, the graph is symmetric about the origin (Odd).
- **Intercepts:** Points where the curve crosses the axes ($x = 0$ or $y = 0$).

We define the linear behaviours of a function at infinity or near singularities.

Definition 5.6.1. ***Asymptotes***.

1. A line $x = c$ is a **vertical asymptote** if $\lim_{x\to c^+} f(x) = \pm\infty$ or $\lim_{x\to c^-} f(x) = \pm\infty$.
2. A line $y = L$ is a **horizontal asymptote** if $\lim_{x\to\infty} f(x) = L$ or $\lim_{x\to-\infty} f(x) = L$.
3. A line $y = mx + c$ is an **oblique (or inclined) asymptote** at $+\infty$ if:

$$\lim_{x\to\infty} [f(x) - (mx + c)] = 0$$

The parameters m and c may be found via limits:

$$m = \lim_{x\to\infty} \frac{f(x)}{x}, \quad c = \lim_{x\to\infty} (f(x) - mx)$$

Case Study: Rational Functions

Rational functions $f(x) = P(x)/Q(x)$ exhibit a rich variety of asymptotic behaviours dependent on the degrees of the numerator and denominator.

Example 5.6.1. A Function with Three Asymptotes. Consider $f(x) = \dfrac{x^2 + 1}{x^2 - 3x + 2}$.

1. **Domain:** The denominator factors as $(x-1)(x-2)$. Thus $D = \mathbb{R} \setminus \{1, 2\}$.
2. **Vertical Asymptotes:** At $x = 1$ and $x = 2$, the denominator vanishes while the numerator is non-zero (2 and 5 respectively). Checking signs near $x = 1$: as $x \to 1^-$, numerator > 0, denominator $(1^- - 1)(1^- - 2) \approx (-)(-)$; ratio is positive. $\lim_{x\to 1^-} f(x) = +\infty$. As $x \to 1^+$, denominator $\approx (+)(-)$; ratio is negative. $\lim_{x\to 1^+} f(x) = -\infty$. Similar analysis at $x = 2$ yields $\lim_{x\to 2^-} f(x) = -\infty$ and $\lim_{x\to 2^+} f(x) = +\infty$.
3. **Horizontal Asymptote:** Since the degrees of P and Q are equal:
$$\lim_{x\to\pm\infty} \frac{x^2(1 + 1/x^2)}{x^2(1 - 3/x + 2/x^2)} = 1$$
Thus $y = 1$ is the horizontal asymptote.
4. **Critical Points:**
$$f'(x) = \frac{2x(x^2 - 3x + 2) - (x^2 + 1)(2x - 3)}{(x^2 - 3x + 2)^2} = \frac{-3x^2 + 2x + 3}{(x^2 - 3x + 2)^2}$$
Setting $f'(x) = 0$ gives $3x^2 - 2x - 3 = 0$. Roots are $x = \frac{1\pm\sqrt{10}}{3}$. $c_1 \approx -0.72$ and $c_2 \approx 1.39$. Since $1 < c_2 < 2$, there is a local extremum between the asymptotes.

Example 5.6.2. Rational Function. Sketch the curve $y = \frac{(x-1)(x-3)}{(x-2)(x-4)}$.

- **Domain:** The function is undefined where the denominator vanishes, so $D = \mathbb{R} \setminus \{2, 4\}$.
- **Horizontal Asymptotes:** As $x \to \pm\infty$, the highest degree terms dominate:
$$y \approx \frac{x^2}{x^2} \to 1$$
Thus, $y = 1$ is the horizontal asymptote. To find intersections with the asymptote, we set $y = 1$:
$$1 = \frac{x^2 - 4x + 3}{x^2 - 6x + 8} \implies x^2 - 6x + 8 = x^2 - 4x + 3 \implies -2x = -5 \implies x = 2.5$$
The curve crosses the asymptote $y = 1$ at the point $(2.5, 1)$.
- **Vertical Asymptotes:** There are singularities at $x = 2$ and $x = 4$. We analyse the one-sided limits:
  - **At $x = 2$:** The numerator approaches $(1)(-1) = -1$ (negative).
    - As $x \to 2^-$: Denom $\approx (0^-)(-2) = 0^+$. Thus $y \to -\infty$.
    - As $x \to 2^+$: Denom $\approx (0^+)(-2) = 0^-$. Thus $y \to +\infty$.
  - **At $x = 4$:** The numerator approaches $(3)(1) = 3$ (positive).
    - As $x \to 4^-$: Denom $\approx (2)(0^-) = 0^-$. Thus $y \to -\infty$.
    - As $x \to 4^+$: Denom $\approx (2)(0^+) = 0^+$. Thus $y \to +\infty$.
- **Intercepts:**
  - **x-intercepts:** $y = 0 \implies (x-1)(x-3) = 0 \implies x = 1, 3$.
  - **y-intercept:** $x = 0 \implies y = \frac{(-1)(-3)}{(-2)(-4)} = \frac{3}{8}$.

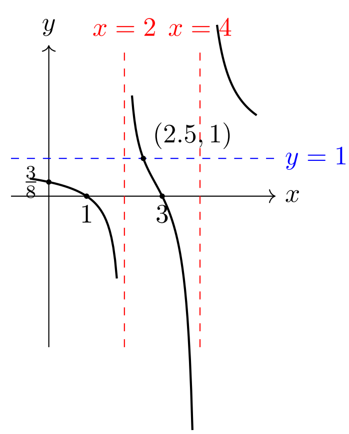


Figure 5.2: Sketch of $y = \frac{(x-1)(x-3)}{(x-2)(x-4)}$.

Example 5.6.3. Even Function. Sketch $y = x^2 + \frac{1}{x^2} - 2$.

- **Symmetry:** replacing x with $-x$ leaves y unchanged. Symmetric about y-axis.
- **Asymptotes:** $x = 0$ is a vertical asymptote ($y \to \infty$). As $x \to \infty$, $y \approx x^2$ (parabolic growth).
- **Roots:** $x^2 + 1/x^2 - 2 = 0 \implies x^4 - 2x^2 + 1 = 0 \implies (x^2 - 1)^2 = 0$. Roots at $x = \pm 1$.

This function behaves like a parabola for large x and shoots to infinity at the origin, touching the axis at ± 1.

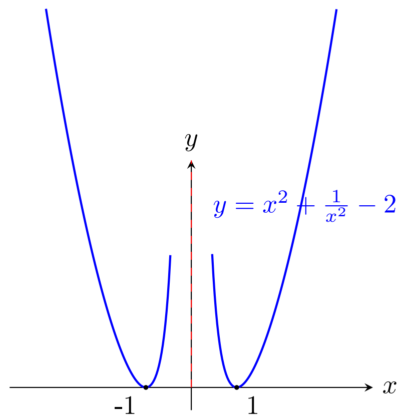


Figure 5.3: Sketch of $y = x^2 + \frac{1}{x^2} - 2$.

Example 5.6.4. Odd Rational Function. Sketch $y = \frac{x}{x^2-1}$.

1. **Symmetry:** $f(-x) = \frac{-x}{(-x)^2-1} = -\frac{x}{x^2-1} = -f(x)$. The function is odd (rotational symmetry about the origin).
2. **Vertical Asymptotes:** $x^2 - 1 = 0 \implies x = \pm 1$.
   - As $x \to 1^+$: $y \approx \frac{1}{0^+} \to +\infty$.
   - As $x \to 1^-$: $y \approx \frac{1}{0^-} \to -\infty$.
   - By odd symmetry, as $x \to -1^-$: $y \to -\infty$, and as $x \to -1^+$: $y \to +\infty$.
3. **Horizontal Asymptote:** As $x \to \pm\infty$, $y \approx \frac{x}{x^2} = \frac{1}{x} \to 0$. $y = 0$ is the horizontal asymptote.
4. **Intercepts:** Only at $(0, 0)$.

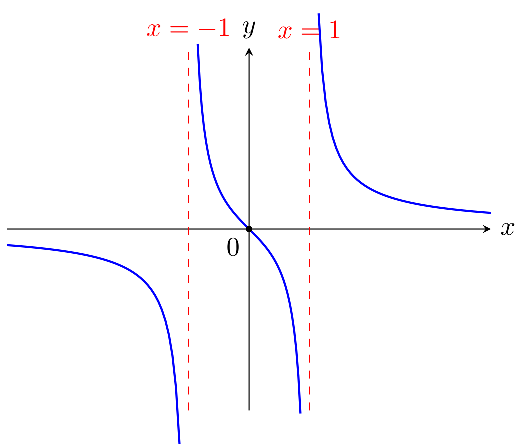


Figure 5.4: Sketch of $y = \frac{x}{x^2-1}$.

5.7 Exercises

1. **Calculations at Infinity.** Evaluate the following limits, justifying your steps using the Limit Laws or the Growth Hierarchy.

   (a)
   $$\lim_{x\to\infty} \frac{3x^3 - 5x + 1}{4x^3 + 2x^2 - 7}$$

   (b)
   $$\lim_{x\to-\infty} \frac{\sqrt{4x^2+1}}{x+2}$$

   (c)
   $$\lim_{x\to\infty} \frac{x^{100}}{2^x}$$

   (d)
   $$\lim_{x\to\infty} (\sqrt{x^2+x} - x)$$

2. **Curve Sketching Practice.** Using the techniques established in the previous section, sketch the general shapes of the curves given by the following equations. Explicitly label all asymptotes and intercepts.

   (a)
   $$y = x^n + x^{-n}$$
   for $n = 2$ (even) and $n = 1$ (odd).

   (b)
   $$y = \frac{x^2}{x+1}.$$

   (c)
   $$y = \frac{x}{x^2+1}.$$

   (d)
   $$y = \frac{x^2+1}{(x-2)(x-4)}.$$

3. **Sequential Proof.** Prove the continuity of $f(x) = |x|$ using the sequential criterion.

4. **The Variable Change Principle.** Prove that $\lim_{x\to\infty} f(x) = L$ if and only if $\lim_{t\to 0^+} f(1/t) = L$. Use this result to evaluate:
   $$\lim_{x\to\infty} x \sin\left(\frac{1}{x}\right).$$

5. **Closed Sets and Continuity.** We established that a function is continuous if and only if the pre-image of every open set is open.

   (a) Prove the dual statement: $f : \mathbb{R} \to \mathbb{R}$ is continuous if and only if for every closed set $C \subseteq \mathbb{R}$, the pre-image $f^{-1}(C)$ is closed.

   (b) **Open Maps.** A function is called an *open map* if the image of every open set is open.

      (i) Show that $f(x) = x^2$ is *not* an open map on $\mathbb{R}$.

      (ii) Show that $f(x) = \sin x$ is *not* an open map on $\mathbb{R}$.

      (iii) Construct a continuous function that is an open map but is not monotonic.

6. **The Pasting Lemma (Closed Case).** Let A and B be closed subsets of $\mathbb{R}$ such that $A \cup B = \mathbb{R}$. Let $f : A \to \mathbb{R}$ and $g : B \to \mathbb{R}$ be continuous functions such that $f(x) = g(x)$ for all $x \in A \cap B$. Define $h : \mathbb{R} \to \mathbb{R}$ by $h(x) = f(x)$ for $x \in A$ and $h(x) = g(x)$ for $x \in B$.

   (a) Use the closed set characterization of continuity (the preimage of a closed set is closed) to prove that h is continuous.

(b) Provide a specific counter-example to show that h may fail to be continuous if A is closed and B is open.

7. Sequential vs. Functional Limits at Infinity.

(a) Prove that if $\lim_{x\to\infty} f(x) = L$, then for every sequence of integers (n_k) diverging to infinity, $\lim_{k\to\infty} f(n_k) = L$.

(b) Show that the converse is false by constructing a function f such that $f(n) = 0$ for all $n \in \mathbb{N}$, yet $\lim_{x\to\infty} f(x)$ does not exist.

(c) $\star$ Prove that if f is non-decreasing on $(0, \infty)$ and the sequential limit $\lim_{n\to\infty} f(n) = L$ exists, then the functional limit $\lim_{x\to\infty} f(x) = L$ exists.

8. The Topologist's Sine Curve. Consider the function $f(x) = \sin(1/x)$ for $x > 0$.

(a) Find two sequences (a_n) and (b_n) approaching 0 from the right such that $\lim f(a_n) = 1$ and $\lim f(b_n) = -1$.

(b) Conclude that $\lim_{x\to 0^+} \sin(1/x)$ does not exist.

(c) Now consider $g(x) = x \sin(1/x)$. Prove that $\lim_{x\to 0} g(x) = 0$.

(d) $\star$ Let $A = \{(x, \sin(1/x)) : x > 0\} \cup \{(0, y) : y \in [-1, 1]\}$. This set in $\mathbb{R}^2$ is known as the Topologist's Sine Curve. Visualise this set. Is the function describing this curve continuous at $x = 0$?

9. $\star$ **Thomae's Function Revisited: One-Sided Limits.** Recall Thomae's function $f(x) = 1/q$ if $x = p/q$ and 0 if irrational.

(a) Evaluate $\lim_{x\to c} f(x)$ for any rational number c.

(b) Does the limit exist? Does it equal $f(c)$?

(c) Use this to re-verify that f is discontinuous at every rational and continuous at every irrational.

10. Asymptotic Equivalence and "Little o" Notation. In the analysis of algorithms and number theory, we often compare the growth rates of functions. We write $f(x) \sim g(x)$ (read "f is asymptotically equivalent to g") if $\lim_{x\to\infty} \frac{f(x)}{g(x)} = 1$.

(a) Prove that $\sim$ is an equivalence relation on the set of functions continuous on $[a, \infty)$ which do not vanish for sufficiently large x.

(b) **The Exponentiation Trap.** Provide a counter-example to show that $f(x) \sim g(x)$ does **not** imply $e^{f(x)} \sim e^{g(x)}$.

(c) **The Logarithmic Rescue.** Prove that if $f(x) \sim g(x)$ and $f(x) \to \infty$ as $x \to \infty$, then $\ln(f(x)) \sim \ln(g(x))$.

Remark. Write $f(x) = g(x)h(x)$ where $h(x) \to 1$. Expand the logarithm of the product.

11. The Closed Graph Theorem (Metric Version). Let $f : \mathbb{R} \to \mathbb{R}$ be a function. We define the *graph* of f as the subset of the Cartesian plane:

$$\Gamma_f = \{(x, y) \in \mathbb{R}^2 : y = f(x)\}$$

(a) Prove that if f is continuous on $\mathbb{R}$, then Γ_f is a closed subset of $\mathbb{R}^2$ (under the standard Euclidean metric).

Remark. Use the sequential characterisation. Let $((x_n, f(x_n)))$ be a sequence of points in the graph converging to some point $(a, b) \in \mathbb{R}^2$. You must show that $(a, b) \in \Gamma_f$, i.e., $b = f(a)$.

(b) Provide a counter-example to show that the converse is false. That is, construct a discontinuous function f whose graph Γ_f is nonetheless a closed set.

Remark. Consider the function $f(x) = 1/x$ for $x \neq 0$ and define $f(0)$ judiciously, or consider a function with a vertical asymptote.

12. Periodic Functions and Limits. A function $f : \mathbb{R} \to \mathbb{R}$ is said to be *periodic* with period $P > 0$ if $f(x + P) = f(x)$ for all $x \in \mathbb{R}$.

(a) Prove that if f is continuous and periodic, then f is bounded on $\mathbb{R}$.

Remark. You may assume the standard result that a continuous function on a closed interval $[0, P]$ is bounded.

(b) Prove that if f is periodic and $\lim_{x\to\infty} f(x) = L$ exists, then f must be a constant function.

(c) Deduce that the sine and cosine functions do not tend to any limit at infinity.

13. The Distance Function II. Let $S \subseteq \mathbb{R}$ be a non-empty set. Define $d_S(x) = \inf\{|x - s| : s \in S\}$.

(a) Prove that $|d_S(x) - d_S(y)| \leq |x - y|$ for all $x, y \in \mathbb{R}$.

(b) Prove that $d_S(x) = 0$ if and only if $x \in \bar{S}$, where $\bar{S}$ denotes the closure of S.

(c) For any $\epsilon > 0$, prove that the set $U_\epsilon = \{x \in \mathbb{R} : d_S(x) < \epsilon\}$ is an open set.

14. Bernstein Polynomials. For a function $f : [0, 1] \to \mathbb{R}$, the n-th *Bernstein polynomial* $B_n(f; x)$ is defined by:

$$B_n(f; x) = \sum_{k=0}^{n} f\left(\frac{k}{n}\right)\binom{n}{k} x^k (1 - x)^{n-k}.$$

(a) **The Partition of Unity.** Use the Binomial Theorem to prove that for the constant function $f(x) = 1$, $B_n(1; x) = 1$.

(b) **The First Moment.** Let $f(x) = x$. Use the identity $k\binom{n}{k} = n\binom{n-1}{k-1}$ to prove that:

$$\sum_{k=0}^{n} \frac{k}{n}\binom{n}{k} x^k (1 - x)^{n-k} = x.$$

(c) **The Second Moment.** Let $f(x) = x^2$. Prove that:

$$B_n(x^2; x) = \left(1 - \frac{1}{n}\right) x^2 + \frac{1}{n} x.$$

Remark. Write $k^2 = k(k-1) + k$. You will need to extract factors of $n(n-1)$ from the binomial coefficient.

(d) **Uniform Convergence.** Using the results above, calculate the explicit error $B_n(x^2; x) - x^2$. Show that this error tends to 0 as $n \to \infty$ independently of $x \in [0, 1]$.

15. Continuity of Monotone Surjections. While continuity usually implies the Intermediate Value Property, the converse is false, structure rescues us. Let $f : [a, b] \to [c, d]$ be a *strictly increasing* function.

(a) Prove that if f is surjective (i.e., $f([a, b]) = [c, d]$), then f must be continuous on $[a, b]$.

Remark. Proceed by contradiction. If f is discontinuous at x_0, it must have a jump discontinuity (why?). If there is a jump, can the function hit every value in the target interval?

(b) Prove the **Continuous Inverse Theorem**: If $f : [a, b] \to [c, d]$ is a continuous, strictly increasing bijection, then its inverse function $f^{-1} : [c, d] \to [a, b]$ is also continuous.

(c) Does this result hold if the domain is not an interval (e.g., a union of two disjoint intervals)? Provide a counter-example.

Chapter 6

Compactness

While limits and continuity rely on local behaviour, compactness is a global property enforcing well-behavedness on functions. The Extreme Value Theorem (as stated in calculus I) asserts that a continuous function $f : [a, b] \to \mathbb{R}$ is bounded and attains its maximum and minimum values. That is, there exist $c, d \in [a, b]$ such that $f(c) \leq f(x) \leq f(d)$ for all $x \in [a, b]$. This theorem fails for open intervals (take $f(x) = 1/x$ on $(0, 1)$) or unbounded intervals (take $f(x) = x$ on $\mathbb{R}$). The property distinguishing $[a, b]$ from $(0, 1)$ or $\mathbb{R}$ in this context is compactness.

6.1 Sequential Compactness

While compactness allows definition via open covers, in real analysis it is naturally characterised via sequences.

Definition 6.1.1. ***Compact Set***. A subset $K \subseteq \mathbb{R}$ is said to be compact if every sequence in K possesses a subsequence that converges to a limit in K. Formally, for every sequence $(x_n) \subseteq K$, there exists a subsequence (x_{n_k}) and a point $x \in K$ such that $\lim\limits_{k\to\infty} x_{n_k} = x$.

This definition imposes two requirements: the subsequence must converge (stability), and the limit must remain within the set (containment).

Example 6.1.1. Finite Sets. Any finite set $K = \{y_1, \ldots, y_m\}$ is compact.

Proof. Let (x_n) be a sequence in K. By the Pigeonhole Principle, at least one element $y \in K$ must appear infinitely often in the sequence. The indices corresponding to this repetition form a subsequence converging to $y \in K$. ■

Example 6.1.2. Closed Intervals. The interval $[a, b]$ is compact.

Proof. Let (x_n) be a sequence in $[a, b]$. Since $a \leq x_n \leq b$ for all n, the sequence is bounded. By the Bolzano-Weierstrass Theorem (Analysis I), every bounded sequence in $\mathbb{R}$ admits a convergent subsequence (x_{n_k}) with limit L. Since $[a, b]$ is a closed set, it contains all its adherent points. Thus $L \in [a, b]$. ■

It is instructive to examine sets that fail to be compact.

Example 6.1.3. Non-Compact Sets.

1. **Open Intervals:** The set $(0, 1)$ is not compact. Take the sequence $x_n = 1/n$. While the sequence converges to 0, and thus every subsequence converges to 0, the limit $0 \notin (0, 1)$. The compactness condition requires the limit to be *in* the set.

2. **Unbounded Sets:** The set $\mathbb{Z}$ (or $\mathbb{R}$) is not compact. The sequence $x_n = n$ tends to infinity. It possesses no convergent subsequence in $\mathbb{R}$, violating the condition.

The Heine-Borel Theorem

The examples above suggest a strong relationship between compactness, closedness, and boundedness. In the setting of $\mathbb{R}$ (and $\mathbb{R}^n$), these properties completely characterise compact sets.

Recall the definition of boundedness from Analysis I:

Definition 6.1.2. *Bounded Set*. A set $S \subseteq \mathbb{R}$ is bounded if there exists $M > 0$ such that $|x| \leq M$ for all $x \in S$.

***Theorem* 6.1.1. Heine-Borel**. A subset $K \subseteq \mathbb{R}$ is compact if and only if it is both closed and bounded.

Proof. We prove the two implications separately.

($\implies$) Suppose K is compact.

- **Boundedness:** Suppose for contradiction that K is unbounded. Then for every $n \in \mathbb{N}$, there exists $x_n \in K$ such that $|x_n| > n$. The sequence (x_n) tends to infinity; consequently, no subsequence of (x_n) can converge to a real number. This contradicts the definition of compactness. Thus, K must be bounded.
- **Closedness:** Suppose K is not closed. Then there exists a limit point x of K such that $x \notin K$. Since x is a limit point, there exists a sequence $(x_n) \subseteq K$ converging to x. Every subsequence of (x_n) must also converge to x. Since $x \notin K$, the sequence (x_n) has no subsequence converging to a point *in* K. This contradicts compactness. Thus, K must be closed.

($\impliedby$) Suppose K is closed and bounded. Let (x_n) be an arbitrary sequence in K. Since K is bounded, the sequence (x_n) is bounded. By the Bolzano-Weierstrass Theorem, (x_n) has a convergent subsequence (x_{n_k}) with limit $L \in \mathbb{R}$. Since $(x_{n_k}) \subseteq K$ and $x_{n_k} \to L$, L is an adherent point of K. Because K is closed, it contains all its adherent points, so $L \in K$. Thus, every sequence in K has a subsequence converging to a point in K, so K is compact.

■

Remark. Heine-Borel is a powerful tool because checking closedness and boundedness is typically easier than testing every possible sequence. However, caution is required: in infinite-dimensional spaces (a topic for functional analysis), closed and bounded sets are not necessarily compact.

6.2 Cantor's Intersection Theorem

The structural similarity between compact sets and closed intervals extends to the property of nested intersections. Recall the Nested Interval Theorem: if $I_1 \supseteq I_2 \supseteq \dots$ is a sequence of closed bounded intervals, their intersection is non-empty. We generalise this to arbitrary compact sets.

***Theorem* 6.2.1. Nested Compact Sets**. Let $(K_n)_{n=1}^{\infty}$ be a sequence of non-empty compact subsets of $\mathbb{R}$ such that $K_{n+1} \subseteq K_n$ for all n. Then:

$$\bigcap_{n=1}^{\infty} K_n \neq \emptyset$$

Proof. For each $n \in \mathbb{N}$, pick an element $x_n \in K_n$. This is possible since each K_n is non-empty. Consider the sequence (x_n). Since $K_{n+1} \subseteq K_n$, we have $x_n \in K_1$ for all n. Since K_1 is compact, there exists a subsequence (x_{n_k}) converging to some point $x \in K_1$.

We claim that $x \in K_m$ for every $m \in \mathbb{N}$. Fix an integer m. For sufficiently large k, we have $n_k \geq m$. Due to the nested property, the tail of the subsequence $(x_{n_k})_{k \geq j}$ lies entirely within K_m. Since K_m is compact (and thus closed), the limit of any convergent sequence in K_m must belong to K_m. Therefore, $x \in K_m$. Since m was arbitrary, x belongs to the intersection of all sets K_n. ■

Remark. If the sets are not compact, the intersection may be empty.

1. **Open sets:** Let $U_n = (0, 1/n)$. Then $\bigcap U_n = \emptyset$.
2. **Closed but unbounded:** Let $F_n = [n, \infty)$. Then $\bigcap F_n = \emptyset$.

Compactness is the essential ingredient ensuring the limit point is not "lost" at the boundary or at infinity.

The Cantor Set, constructed by iteratively removing the middle third of closed intervals, is a compact set that is uncountable, has measure zero, and contains no intervals. Being closed and bounded, it satisfies all properties derived in this chapter.

Open Covers

Open covers allow us to localise properties. If a property holds locally (in a small neighbourhood of every point), under what conditions does it hold globally? Compactness provides the answer: it allows us to stitch together finitely many local patches.

Definition 6.2.1. ***Open Cover***. Let $K \subseteq \mathbb{R}$. An open cover of K is a collection of open sets $\mathcal{O} = \{U_\alpha : \alpha \in \Lambda\}$ such that the union of these sets contains K:

$$K \subseteq \bigcup_{\alpha \in \Lambda} U_\alpha$$

A subcover is a sub-collection of $\mathcal{O}$ that still covers K. If this sub-collection consists of finitely many sets, we call it a finite subcover.

Example 6.2.1. Finite Sets. Any open cover of a finite set $K = \{x_1, \ldots, x_n\}$ admits a finite subcover. If $\mathcal{O}$ covers K, then for each x_i, there is some set $U_i \in \mathcal{O}$ such that $x_i \in U_i$. The collection $\{U_1, \ldots, U_n\}$ is a finite subcover.

Example 6.2.2. Failure of Compactness (Open Interval). Take the open interval $(0, 1)$. This set is not compact. Define the collection:

$$\mathcal{O} = \left\{ \left(\frac{1}{n}, 1 \right) : n \in \mathbb{N}, n \geq 2 \right\}$$

Clearly, $\bigcup_{n=2}^{\infty} (1/n, 1) = (0, 1)$, so this is an open cover. Suppose there existed a finite subcover. Let N be the largest index in this finite sub-collection. Since the sets are nested, the union of the finite subcover is simply the largest set $(1/N, 1)$. However, $(0, 1) \not\subseteq (1/N, 1)$ because the interval $(0, 1/N]$ is missed. Thus, no finite subcover exists.

Example 6.2.3. Failure of Compactness (Unbounded). Let $K = \mathbb{N}$. Consider the open cover $\mathcal{O} = \{(n - 1, n + 1) : n \in \mathbb{N}\}$. Each integer n is contained in $(n - 1, n + 1)$, but removing any single set $(k - 1, k + 1)$ leaves the integer k uncovered. Thus, no finite subcollection can cover the infinite set $\mathbb{N}$.

These examples reinforce our previous findings: open intervals and unbounded sets fail to be compact. We now formalise the "covering" definition of compactness.

Definition 6.2.2. ***Topological Compactness***. A set $K \subseteq \mathbb{R}$ is compact if every open cover of K admits a finite subcover.

The Lebesgue Number Lemma

To prove that our two definitions of compactness (sequential vs covering) are equivalent, we require the Lebesgue number lemma. This states that if a compact set is covered by open sets, there is a minimum "safety radius" δ such that any ball of radius δ fits entirely inside at least one of the covering sets.

***Lemma* 6.2.1.** *Lebesgue Number Lemma.* Let K be a sequentially compact subset of $\mathbb{R}$ and let $\mathcal{O}$ be an open cover of K. Then there exists a $\delta > 0$ such that for every $x \in K$, the ball $B_\delta(x) = (x - \delta, x + \delta)$ is contained in some element $U \in \mathcal{O}$.

Proof. We proceed by contradiction. Suppose no such δ exists. Then for every $n \in \mathbb{N}$, the value $1/n$ fails to be a Lebesgue number. This implies that for each n, there exists a point $x_n \in K$ such that the ball $B_{1/n}(x_n)$ is *not* contained in any single set $U \in \mathcal{O}$.

This construction yields a sequence (x_n) in K. Since K is sequentially compact, there exists a subsequence (x_{n_k}) converging to some limit $z \in K$. Since $\mathcal{O}$ covers K, there exists some open set $U_z \in \mathcal{O}$ such that $z \in U_z$. Because U_z is open, there exists an $\epsilon > 0$ such that $B_\epsilon(z) \subseteq U_z$.

Now, we examine the tail of the subsequence. Choose k sufficiently large such that two conditions are met:

1. The point x_{n_k} is close to z: $|x_{n_k} - z| < \epsilon/2$.
2. The radius of the ball around x_{n_k} is small: $1/n_k < \epsilon/2$.

We apply the Triangle Inequality to show that the ball around x_{n_k} lies inside $B_\epsilon(z)$. Let $y \in B_{1/n_k}(x_{n_k})$. Then:

$$|y - z| \le |y - x_{n_k}| + |x_{n_k} - z| < \frac{1}{n_k} + \frac{\epsilon}{2} < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon$$

Thus, $B_{1/n_k}(x_{n_k}) \subseteq B_\epsilon(z) \subseteq U_z$. This contradicts the defining property of the sequence (x_n), which stated that $B_{1/n}(x_n)$ is not contained in any element of the cover. Consequently, our assumption was false, and a Lebesgue number δ must exist. ■

Equivalence of Definitions

We now establish the equivalence unifiying the topological and analytic views of compactness.

***Theorem* 6.2.2. The Compactness Theorem**. A subset $K \subseteq \mathbb{R}$ is sequentially compact if and only if it is topologically compact (every open cover has a finite subcover).

Proof. We prove the two implications separately.

($\implies$) Sequential implies Covering. Let K be sequentially compact and let $\mathcal{O}$ be an open cover of K. By lemma 6.2.1, there exists a Lebesgue number $\delta > 0$ for this cover. We construct a finite cover using balls of radius δ. Pick an arbitrary point $x_1 \in K$. If $B_\delta(x_1)$ covers K, we are done. If not, pick $x_2 \in K \setminus B_\delta(x_1)$. Proceeding inductively, if $\{B_\delta(x_1), \dots, B_\delta(x_m)\}$ does not cover K, pick $x_{m+1} \in K \setminus \bigcup_{j=1}^m B_\delta(x_j)$.

Consider the sequence (x_m) generated by this process. By construction, for any distinct indices $i \neq j$, $|x_i - x_j| \ge \delta$. Such a sequence cannot have a convergent subsequence (a Cauchy subsequence would eventually have terms closer than δ). This contradicts the sequential compactness of K. Therefore, the process must terminate after finitely many steps. We obtain points $x_1, \dots, x_N$ such that $K \subseteq \bigcup_{j=1}^N B_\delta(x_j)$. Since δ is a Lebesgue number, each ball $B_\delta(x_j)$ is contained in some set $U_j \in \mathcal{O}$. Thus, $\{U_1, \dots, U_N\}$ is a finite subcover of $\mathcal{O}$.

($\impliedby$) Covering implies Sequential. Let K be topologically compact. Let (x_n) be a sequence in K. Let $S = \{x_n : n \in \mathbb{N}\}$ be the set of values in the sequence. If S is finite, one value repeats infinitely often, yielding a convergent subsequence. Assume S is infinite. We define L to be the set of limit points of S. We aim to show $L \cap K \neq \emptyset$.

Suppose for contradiction that $L \cap K = \emptyset$. This implies that no point in K is a limit point of S. Consequently, for every $y \in K$, there exists a neighbourhood U_y of y that contains at most one point of S (specifically, if $y \in S$, it contains y; if not, it contains nothing from S). The collection $\mathcal{U} = \{U_y : y \in K\}$ is an open cover of K. By compactness, there exists a finite subcover $\{U_{y_1}, \dots, U_{y_m}\}$. Since $S \subseteq K \subseteq \bigcup_{j=1}^{m} U_{y_j}$ and each U_{y_j} contains at most one point of S, the set S must be finite. This contradicts the assumption that S is infinite. Thus, S must have a limit point $x \in K$. By the definition of a limit point, there exists a subsequence of (x_n) converging to x.

■

6.3 Continuity and Compactness

We now explore the interaction between compactness and continuous functions. A topological principle states that "compactness is preserved by continuity".

***Theorem* 6.3.1. Preservation of Compactness**. Let $f : K \to \mathbb{R}$ be a continuous function. If K is compact, then the image set $f(K)$ is compact.

Proof. Let (y_n) be an arbitrary sequence in the image set $f(K)$. We must find a subsequence converging to a point in $f(K)$. Since $y_n \in f(K)$, for each n, there exists some $x_n \in K$ such that $f(x_n) = y_n$. This generates a sequence (x_n) in K. Because K is compact, there exists a subsequence (x_{n_k}) converging to some $x \in K$. By the sequential characterisation of continuity (theorem 4.1.1), since $x_{n_k} \to x$, it follows that $f(x_{n_k}) \to f(x)$. Substituting back, $y_{n_k} \to f(x)$. Since $x \in K$, the limit $f(x)$ belongs to $f(K)$. Thus, (y_n) has a convergent subsequence in $f(K)$, so $f(K)$ is compact. ■

By combining Preservation of Compactness with Heine-Borel, we obtain a simple proof of the Extreme Value Theorem.

***Theorem* 6.3.2. Extreme Value Theorem**. Let K be a compact set and let $f : K \to \mathbb{R}$ be continuous. Then f is bounded and attains its maximum and minimum values on K.

$$\exists c, d \in K \text{ such that } f(c) = \inf_{x \in K} f(x) \text{ and } f(d) = \sup_{x \in K} f(x)$$

Proof. By Theorem 6.3.1, the image set $f(K)$ is compact. By Theorem 6.1.1, a compact subset of $\mathbb{R}$ is closed and bounded.

Boundedness: Since $f(K)$ is bounded, f is a bounded function.

Attainment: Let $M = \sup f(K)$ and $m = \inf f(K)$. By the definition of supremum and infimum, M and m are adherent points of $f(K)$. Since $f(K)$ is closed, it contains its adherent points. Thus $M \in f(K)$ and $m \in f(K)$. Therefore, there exist $c, d \in K$ such that $f(d) = M$ and $f(c) = m$.

■

Remark. This theorem demonstrates that compactness is the "correct" generalisation of a closed interval. The properties making $[a, b]$ nice for optimisation are its compactness, not its connectedness or linearity.

Uniform Continuity

Recall the proof of the continuity of $f(x) = x^2$ in the previous chapter. Given $\epsilon > 0$ and a point c, we bounded $|x^2 - c^2| = |x - c||x + c|$. To control $|x + c|$, we assumed $|x - c| < 1$, leading to $|x + c| < 1 + 2|c|$. This yielded $\delta = \min(1, \epsilon/(1 + 2|c|))$.

Notice that as $|c| \to \infty$, δ shrinks towards zero. For a fixed ϵ, the "sensitivity" of the function varies across the domain. Near the origin, a large δ suffices; near infinity, the function is steep, requiring a microscopic δ.

This motivates a stronger form of continuity where δ depends solely on ϵ.

Definition 6.3.1. ***Uniform Continuity***. Let $A \subseteq \mathbb{R}$. A function $f : A \to \mathbb{R}$ is said to be uniformly continuous on A if for every $\epsilon > 0$, there exists a $\delta > 0$ such that for all $x, y \in A$:

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

In logical notation, the distinction lies in the order of quantifiers:

- **Continuity on** A**:** $\forall x \in A, \forall \epsilon > 0, \exists \delta > 0, \forall y \in A, (|x - y| < \delta \implies |f(x) - f(y)| < \epsilon)$.
- **Uniform Continuity on** A: $\forall \epsilon > 0, \exists \delta > 0, \forall x \in A, \forall y \in A, (|x - y| < \delta \implies |f(x) - f(y)| < \epsilon)$.

Example 6.3.1. Linear Functions. The function $f(x) = x$ is uniformly continuous on $\mathbb{R}$. For any $\epsilon > 0$, we may simply choose $\delta = \epsilon$. Then $|x - y| < \delta$ immediately implies $|x - y| < \epsilon$. This δ works for all pairs x, y, regardless of their location.

Example 6.3.2. The Square Function. The function $f(x) = x^2$ is continuous on $\mathbb{R}$ but not uniformly continuous on $\mathbb{R}$. To prove this, we verify the negation of the definition. We must show:

$$\exists \epsilon > 0 \text{ such that } \forall \delta > 0, \exists x, y \in \mathbb{R} \text{ with } |x - y| < \delta \text{ but } |x^2 - y^2| \geq \epsilon$$

Let $\epsilon = 1$. Let $\delta > 0$ be arbitrary. We construct points x, y that are δ-close but mapped far apart. Choose $x = \frac{1}{\delta}$ and $y = \frac{1}{\delta} + \frac{\delta}{2}$. Then $|x - y| = \frac{\delta}{2} < \delta$. However, the distance between images is:

$$|f(x) - f(y)| = y^2 - x^2 = (y - x)(y + x) = \frac{\delta}{2}\left(\frac{2}{\delta} + \frac{\delta}{2}\right) = 1 + \frac{\delta^2}{4} > 1$$

Thus, the condition fails for $\epsilon = 1$. The function grows too rapidly for a single δ to suffice everywhere.

The strategy used above can be generalised. To demonstrate a lack of uniform continuity, it suffices to find two sequences that approach each other while their images stay apart.

***Proposition* 6.3.1.** *Sequential Failure of Uniform Continuity.* A function $f : A \to \mathbb{R}$ is not uniformly continuous if and only if there exists an $\epsilon_0 > 0$ and two sequences $(x_n), (y_n)$ in A such that:

$$\lim_{n\to\infty} |x_n - y_n| = 0 \quad \text{but} \quad |f(x_n) - f(y_n)| \geq \epsilon_0 \text{ for all } n$$

For $f(x) = x^2$, the sequences $x_n = n$ and $y_n = n + 1/n$ satisfy this criterion, as $|x_n - y_n| = 1/n \to 0$ while $|x_n^2 - y_n^2| = 2 + 1/n^2 \geq 2$.

The Heine-Cantor Theorem

The example in proposition 6.3.1 relied on the domain being unbounded. Intuitively, if the domain is confined (bounded) and includes its endpoints (closed), such behaviour should be impossible.

***Theorem* 6.3.3. Heine-Cantor Theorem.** Let K be a compact subset of $\mathbb{R}$ and let $f : K \to \mathbb{R}$ be a continuous function. Then f is uniformly continuous on K.

Proof. Let $\epsilon > 0$ be given. We aim to find a global $\delta > 0$ depending only on ϵ. Since f is continuous, for every point $x \in K$, there exists a local radius $\delta_x > 0$ such that:

$$y \in K, |y - x| < \delta_x \implies |f(y) - f(x)| < \frac{\epsilon}{2}$$

(Note the use of $\epsilon/2$, a standard preparatory step for the Triangle Inequality).

Consider the collection of open intervals $\mathcal{U} = \{B(x, \frac{\delta_x}{2}) : x \in K\}$. Note that we use half the local radius for the covering sets. Clearly, $\mathcal{U}$ is an open cover of K. Since K is compact, there exists a finite subcover. Let the centres of the finite subcover be $x_1, \ldots, x_n$. Thus, $K \subseteq \bigcup_{i=1}^{n} B(x_i, \frac{\delta_{x_i}}{2})$. We define the global δ to be the minimum of these specific radii:

$$\delta = \min\left\{\frac{\delta_{x_1}}{2}, \ldots, \frac{\delta_{x_n}}{2}\right\}$$

Since the set is finite and each $\delta_{x_i} > 0$, we have $\delta > 0$. Now, let $u, v \in K$ be any points satisfying $|u - v| < \delta$. We must show $|f(u) - f(v)| < \epsilon$. Since the balls cover K, the point u must belong to some specific ball $B(x_k, \frac{\delta_{x_k}}{2})$ from our finite subcover. This implies $|u - x_k| < \frac{\delta_{x_k}}{2}$. We now check the distance of v from the centre x_k:

$$|v - x_k| \leq |v - u| + |u - x_k| < \delta + \frac{\delta_{x_k}}{2}$$

Since $\delta \leq \frac{\delta_{x_k}}{2}$, we have:

$$|v - x_k| < \frac{\delta_{x_k}}{2} + \frac{\delta_{x_k}}{2} = \delta_{x_k}$$

Since both u and v are within distance δ_{x_k} of x_k, the continuity condition at x_k applies to both:

$$|f(u) - f(v)| \leq |f(u) - f(x_k)| + |f(x_k) - f(v)| < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon$$

Thus, f is uniformly continuous. ■

6.4 Connectedness

We turn now to the topological property that captures the idea of a set being "in one piece". For example, $[0, 1] \cup [2, 3]$ is disconnected because it consists of two disjoint components separated by a gap. Similarly, $\mathbb{Q}$ is totally disconnected because between any two rationals lies an irrational gap. First we formalise the notion of a separation. It is not enough for two sets to be disjoint (like $(0, 1)$ and $(1, 2)$); they must be separated in a way that prevents "gluing" at the boundary.

Definition 6.4.1. ***Separated Sets***. Two non-empty sets $A, B \subseteq \mathbb{R}$ are said to be separated if:

$$\bar{A} \cap B = \emptyset \quad \text{and} \quad A \cap \bar{B} = \emptyset$$

That is, neither set contains an adherent point of the other.

Definition 6.4.2. ***Connected Set***. A set $S \subseteq \mathbb{R}$ is disconnected if it can be written as the union of two separated sets. That is, if there exist A, B such that $S = A \cup B$ and A, B are separated. If S is not disconnected, it is said to be connected.

Example 6.4.1. Examples of Connectedness.

1. The set $S = [0, 1] \cup [2, 3]$ is disconnected. Let $A = [0, 1]$ and $B = [2, 3]$. Since the closure $\bar{A} = [0, 1]$ (definition 3.5.1) is disjoint from B, these sets form a separation.
2. The set $S = (0, 1) \cup (1, 2)$ is disconnected. Although the sets approach the same point 1, the point $1 \notin S$. Let $A = (0, 1)$ and $B = (1, 2)$. The closure $\bar{A} = [0, 1]$ does not intersect B, and $\bar{B} = [1, 2]$ does not intersect A.
3. The set $S = [0, 1] \cup (1, 2) = [0, 2)$ is connected. If we try to split it at 1 using $A = [0, 1]$ and $B = (1, 2)$, we find $\bar{B} = [1, 2]$, so $A \cap \bar{B} = \{1\} \neq \emptyset$. The sets are not separated.

Characterisation of Connected Sets

While the definition involving closures is precise, it is often cumbersome. We present equivalent conditions that are often easier to apply.

***Theorem* 6.4.1. Characterisations of Connectedness**. Let $S \subseteq \mathbb{R}$ be a non-empty set. The following are equivalent:

(i) S is connected.
(ii) Every continuous function $f : S \to \{0,1\}$ is constant. (Here $\{0,1\}$ has the discrete metric).
(iii) **The Interval Property:** If $a, b \in S$ and $a < c < b$, then $c \in S$. (i.e., S contains the interval $[a,b]$).

Proof. We proceed by the cyclic implications $(i) \implies (ii) \implies (iii) \implies (i)$.

$(i) \implies (ii)$ Suppose S is connected. Let $f : S \to \{0,1\}$ be continuous. Let $A = f^{-1}(\{0\})$ and $B = f^{-1}(\{1\})$. Since f maps into a discrete set, f is locally constant. The pre-images of closed sets are closed in the subspace topology of S. Suppose f is not constant. Then both A and B are non-empty. Since f A is closed relative to S (theorem 5.1.1), meaning $\bar{A} \cap S = A$. Thus $\bar{A} \cap B = (\bar{A} \cap S) \cap B = A \cap B = \emptyset$. Similarly $A \cap \bar{B} = \emptyset$. This implies A, B form a separation of S, contradicting the connectedness of S. Thus f must be constant.

$(ii) \implies (iii)$ Suppose all continuous maps to $\{0,1\}$ are constant. Assume for contradiction that condition (iii) fails. Then there exist $a, b \in S$ and a point $c \notin S$ such that $a < c < b$. Define a function $f : S \to \{0,1\}$ by:

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

Since $c \notin S$, this function is defined on all of S. The function is continuous because for any $x \in S$, we can choose a neighbourhood of x (e.g., radius $|x - c|$) that does not cross c, on which f is constant. However, $f(a) = 0$ and $f(b) = 1$, so f is not constant. Contradiction.

$(iii) \implies (i)$ Suppose S has the Interval Property. Assume for contradiction that S is disconnected, so $S = A \cup B$ with A, B separated. Pick $a \in A$ and $b \in B$. Without loss of generality, assume $a < b$. By the Interval Property, $[a,b] \subseteq S$. Let $K = A \cap [a,b]$. Since A is closed in S (it equals $\bar{A} \cap S$), K is closed in the subspace $[a,b]$. Let $s = \sup(K)$. Since K is closed and bounded, $s \in K \subseteq A$. Because $b \in B$ and $A \cap B = \emptyset$, we must have $s < b$. By the property of the supremum, every point in the interval $(s, b]$ must belong to B (otherwise there would be a point of A larger than s). This implies that s is a limit point of B from the right. Thus $s \in \bar{B}$. Consequently, $s \in A \cap \bar{B}$, which implies $A \cap \bar{B} \neq \emptyset$. This contradicts the separation condition ($A \cap \bar{B} = \emptyset$). Thus S must be connected.

∎

***Corollary* 6.4.1.** *Connected Subsets of* $\mathbb{R}$. A subset of $\mathbb{R}$ is connected if and only if it is an interval. The connected sets are precisely the singletons, open intervals (a,b), closed intervals $[a,b]$, and half-open intervals (bounded or unbounded).

We conclude this section by re-deriving the Intermediate Value Theorem (IVT) as a topological result.

***Proposition* 6.4.1.** *Preservation of Connectedness.* Let $f : S \to \mathbb{R}$ be a continuous function. If S is connected, then the image $f(S)$ is connected.

Proof. Let $g : f(S) \to \{0,1\}$ be a continuous function. Consider the composition $h = g \circ f$. Since f and g are continuous, $h : S \to \{0,1\}$ is continuous. Since S is connected, h must be constant. This implies that g must be constant on $f(S)$. By Theorem 6.4.1, $f(S)$ is connected. ∎

***Theorem* 6.4.2. The Intermediate Value Theorem.** Let $f : [a,b] \to \mathbb{R}$ be a continuous function such that $f(a) \neq f(b)$. If y is any value strictly between $f(a)$ and $f(b)$, there exists $c \in (a,b)$ such that $f(c) = y$.

Proof. The interval $[a,b]$ is connected. By proposition 6.4.1, the image $f([a,b])$ is connected, and thus must be an interval (corollary 6.4.1). Since $f(a)$ and $f(b)$ are in the image, the entire interval connecting them (denoted I_{val} with endpoints $f(a), f(b)$) must be contained in the image. Since $y \in I_{val}$, it follows that $y \in f([a,b])$. Thus, there exists $c \in [a,b]$ such that $f(c) = y$. ∎

6.5 Darboux Continuity

It is natural to ask whether the converse holds: does the Intermediate Value Property (IVP) imply continuity? To investigate this, we isolate the property itself.

Definition 6.5.1. ***Darboux Continuity***. A function $f : I \to \mathbb{R}$ is said to be Darboux continuous (or has the Intermediate Value Property (IVP)) if for any sub-interval $[a, b] \subseteq I$ and any value γ strictly between $f(a)$ and $f(b)$, there exists $c \in (a, b)$ such that $f(c) = \gamma$.

While The Intermediate Value Theorem asserts that continuity implies Darboux Continuity, the converse is false. The property of traversing all intermediate values does not prevent the function from oscillating wildly.

Example 6.5.1. The Topologist's Sine Curve. Recall the function defined in Chapter 1:

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

As established previously, this function is discontinuous at $x = 0$. However, it possesses the Intermediate Value Property (IVP). On any interval $[0, b]$, the function takes the value 0 at the endpoint. For $x > 0$, as $x \to 0$, the term $1/x \to \infty$, causing $\sin(1/x)$ to oscillate through the entire range $[-1, 1]$ infinitely many times. Thus, f assumes every value between $f(0) = 0$ and $f(b)$ within the interval. Consequently, f is Darboux continuous despite its discontinuity.

Remark. Historically, Gaston Darboux showed that derivatives of differentiable functions satisfy the Intermediate Value Property (IVP), even if the derivative itself is not continuous. Thus, the class of Darboux continuous functions is significant in the study of differentiation.

Monotonicity and Continuity While Darboux Continuity alone is insufficient to guarantee continuity, imposing a structural constraint on the function's growth bridges the gap. If a function is strictly monotone, it cannot oscillate; therefore, if it also skips no values (Darboux), it must be continuous.

Definition 6.5.2. ***Strict Monotonicity***. A function $f : I \to \mathbb{R}$ is strictly increasing if for all $x, y \in I$, $x < y \implies f(x) < f(y)$. It is strictly decreasing if $x < y \implies f(x) > f(y)$.

***Theorem* 6.5.1. Continuity of Monotone Darboux Functions**. Let I be an interval and let $f : I \to \mathbb{R}$ be a strictly increasing function. If f is Darboux continuous, then f is continuous on I.

Proof. Fix an interior point $c \in I$. We analyse the left and right limits of f at c. Define the left limit candidate $L^- = \sup\{f(x) : x \in I, x < c\}$ and the right limit candidate $L^+ = \inf\{f(x) : x \in I, x > c\}$. Since f is strictly increasing, for any $x < c < z$, we have $f(x) < f(c) < f(z)$. This implies:

$$L^- \leq f(c) \leq L^+$$

We claim that $L^- = f(c)$. Suppose for contradiction that $L^- < f(c)$. Since f is Darboux continuous, the image $f(I)$ must be an interval. However, the value $y = \frac{L^- + f(c)}{2}$ satisfies $L^- < y < f(c)$. For any $x < c$, $f(x) \leq L^- < y$. For any $x \geq c$, $f(x) \geq f(c) > y$. Thus, y is never attained by f, contradicting the Darboux property. Therefore, $L^- = f(c)$. By a symmetric argument, $L^+ = f(c)$. Since the left and right limits exist and equal $f(c)$, the function is continuous at c. (The argument for endpoints is identical, considering only the relevant one-sided limit). ∎

Note. Continuity of Monotone Darboux Functions provides a powerful tool for proving the continuity of inverse functions. If a continuous function f is strictly increasing, its inverse f^{-1} is strictly increasing and maps an interval to an interval (surjective). This implies f^{-1} has the Intermediate Value Property (IVP), and thus, by the theorem above, f^{-1} is automatically continuous.

6.6 Exercises

1. **Compactness in Discrete Metrics.** Let X be an infinite set equipped with the discrete metric $d(x, y) = 1$ if $x \neq y$ and 0 if $x = y$.
   (a) Determine which subsets of X are compact using the sequential definition of compactness.
   (b) Does the Heine-Borel theorem (Compact $\iff$ Closed and Bounded) hold in this metric space? Justify your answer.

2. **Intersection of Nested Sets.** Give explicit counter-examples to show that the intersection $\bigcap_{n=1}^{\infty} K_n$ may be empty if the sets K_n are nested ($K_{n+1} \subseteq K_n$) but we relax the conditions:
   (a) K_n are closed but not bounded.
   (b) K_n are bounded but not closed.

3. **Finite Intersection Property.** Let K be a compact set. Prove that a collection of closed subsets $\{F_\alpha\}_{\alpha \in \Lambda}$ of K has a non-empty intersection $\bigcap_{\alpha \in \Lambda} F_\alpha \neq \emptyset$ if and only if every finite sub-collection has a non-empty intersection.

   Remark. Consider the complements $U_\alpha = K \setminus F_\alpha$ and use the definition of compactness via open covers.

4. **The Cantor Set.** The Cantor middle-thirds set $\mathcal{C}$ is constructed by removing the open middle third $(1/3, 2/3)$ from $[0, 1]$, then removing the middle thirds $(1/9, 2/9)$ and $(7/9, 8/9)$ from the remaining segments, and so on ad infinitum.
   (a) Prove that $\mathcal{C}$ is a compact set.
   (b) Prove that $\mathcal{C}$ is uncountable by identifying points in $\mathcal{C}$ with ternary expansions using only the digits 0 and 2.

   Remark. Map each binary sequence (b_n) to $\sum_{n=1}^{\infty} \frac{2b_n}{3^n}$.

5. **Compactness of the Limit Point Set.** Let (x_n) be a bounded sequence in $\mathbb{R}$. Let L be the set of all subsequential limits of (x_n). Prove that L is a compact, non-empty set.

6. **Local vs Global Boundedness.** A function $f : \mathbb{R} \to \mathbb{R}$ is called *locally bounded* if for every $x \in \mathbb{R}$, there exists a neighbourhood U_x on which f is bounded.
   (a) Prove that if f is locally bounded on a compact set K, then f is bounded on K.
   (b) Construct a function $f : \mathbb{R} \to \mathbb{R}$ that is locally bounded everywhere but not bounded on $\mathbb{R}$.

7. **Attainment of Bounds.** Let $f : \mathbb{R} \to \mathbb{R}$ be a continuous function. Suppose that $\lim_{x\to\infty} f(x) = \lim_{x\to-\infty} f(x) = 0$.
   (a) Prove that f is bounded on $\mathbb{R}$.
   (b) Prove that f attains a maximum or a minimum value (or both) on $\mathbb{R}$.
   (c) Give an example where the maximum is attained but the minimum is not.

8. **Proper Maps.** A continuous function $f : \mathbb{R} \to \mathbb{R}$ is called a **proper map** if the pre-image of every compact set is compact.
   (a) Prove that if f is proper, then $\lim_{|x|\to\infty} |f(x)| = \infty$.
   (b) Prove that if f is a non-constant polynomial, then f is a proper map.

9. **Semi-Continuity.** A function $f : \mathbb{R} \to \mathbb{R}$ is called *lower semi-continuous* (LSC) at c if for every $\epsilon > 0$, there exists $\delta > 0$ such that:
$$|x - c| < \delta \implies f(x) > f(c) - \epsilon$$
   (a) Prove that the indicator function of an open set is LSC.
   (b) Prove that if f is LSC on a compact set K, then f attains its minimum on K.
   (c) Prove that a function is continuous if and only if it is both upper and lower semi-continuous.

10. **Fixed Point Theorems.** Let $f : [a, b] \to [a, b]$ be a continuous function.

(a) Prove that f has a fixed point, i.e., there exists $c \in [a, b]$ such that $f(c) = c$.
(b) **Contraction Mapping:** Suppose further that there exists a constant $0 < k < 1$ such that $|f(x) - f(y)| \leq k|x - y|$ for all $x, y \in [a, b]$. Prove that the fixed point is unique.
(c) **Weak Contraction:** Let K be a compact subset of $\mathbb{R}$ and let $f : K \to K$ be a function such that $|f(x) - f(y)| < |x - y|$ for all $x \neq y$. Prove that f has a unique fixed point.

Remark. The function $g(x) = |x - f(x)|$ is continuous. Does it attain a minimum?

11. Inverse Functions on Intervals II. Let $f : [a, b] \to \mathbb{R}$ be a continuous, injective function.

(a) Prove that f must be strictly monotone.
(b) Prove that $f([a, b])$ is a compact interval and that the inverse function $f^{-1} : f([a, b]) \to [a, b]$ is uniformly continuous.

Remark. Use connectedness of $[a, b]$ and the Heine-Cantor theorem.

12. Inverse Functions on Compact Sets. Let K be a compact subset of $\mathbb{R}$ and let $f : K \to Y$ be a continuous bijection onto a subset $Y \subseteq \mathbb{R}$. Prove that the inverse function $f^{-1} : Y \to K$ is continuous.

Remark. Use the topological characterisation of continuity. Show that the pre-image of a closed set under f^{-1} is closed in Y.

13. Uniform Continuity I: Concrete Checks. Determine whether the following functions are uniformly continuous on the specified domains. Provide $\epsilon - \delta$ proofs or counterexamples via sequences.

(a) $f(x) = x \sin(1/x)$ on $(0, 1)$.
(b) $g(x) = \sin(x^2)$ on $[0, \infty)$.
(c) $h(x) = \frac{1}{1+x^2}$ on $\mathbb{R}$.

14. Uniform Continuity II: The Square Root. Prove that $f(x) = \sqrt{x}$ is uniformly continuous on $[0, \infty)$.

Remark. Split the domain into a compact part $[0, 1]$ and an unbounded part $[1, \infty)$. Use the Lipschitz property on the latter.

15. Uniform Continuity III: Periodic Functions. Let $f : \mathbb{R} \to \mathbb{R}$ be a continuous function that is periodic with period $P > 0$ (i.e., $f(x + P) = f(x)$ for all x). Prove that f is uniformly continuous on $\mathbb{R}$.

16. Linear Growth. Let $f : \mathbb{R} \to \mathbb{R}$ be a uniformly continuous function. Prove that there exist constants $A, B > 0$ such that for all $x \in \mathbb{R}$:

$$|f(x)| \leq A|x| + B$$

17. The Extension Theorem. Let $f : (a, b) \to \mathbb{R}$ be a continuous function. Prove that f is uniformly continuous on (a, b) if and only if f can be extended to a continuous function $\tilde{f} : [a, b] \to \mathbb{R}$.

Remark. For the forward direction, use Cauchy sequences to define limits at the endpoints.

18. Extension from Dense Sets. Let $D \subseteq \mathbb{R}$ be a dense subset (e.g., $\mathbb{Q}$) and let $f : D \to \mathbb{R}$ be a uniformly continuous function.

(a) Prove that if (x_n) is a Cauchy sequence in D, then $(f(x_n))$ is a Cauchy sequence in $\mathbb{R}$.
(b) Use this to prove that f has a unique continuous extension to the closure $\bar{D}$.
(c) Show that this fails if f is merely continuous but not uniformly continuous (consider $f(x) = 1/x$ on $(0, 1)$).

19. Hölder Continuity. A function $f : A \to \mathbb{R}$ is said to be **Hölder continuous** of order α if there exists a constant $C > 0$ such that for all $x, y \in A$:

$$|f(x) - f(y)| \leq C|x - y|^\alpha$$

(a) Prove that if f is Hölder continuous of order $\alpha > 0$ on A, then f is uniformly continuous on A.
(b) Prove that if $\alpha > 1$ and A is an interval, then f must be a constant function.

(c) Let $f(x) = \sqrt{x}$ on $[0,1]$. Prove directly from the definition that f is Hölder continuous of order $\alpha = 1/2$.

20. The Distance Function III. Let $K \subseteq \mathbb{R}$ be a non-empty closed set. Define $d_K(x) = \inf\{|x-y| : y \in K\}$.

(a) Prove that for every $x \in \mathbb{R}$, there exists a point $k \in K$ such that $d_K(x) = |x-k|$.
(b) Suppose K is also a convex set. Prove that for each $x \in \mathbb{R}$, the point $k \in K$ that minimizes the distance to x is unique.

21. Distance Between Sets. Let $K \subseteq \mathbb{R}$ be a non-empty compact set and let $F \subseteq \mathbb{R}$ be a closed set disjoint from K.

(a) Using the previous exercise, define $g(x) = d_F(x)$. Prove that $g(x) > 0$ for all $x \in K$.
(b) Using the Extreme Value Theorem on g, prove that there exists a strictly positive minimum distance between the sets:

$$d(K,F) = \inf\{|x-y| : x \in K, y \in F\} > 0$$

(c) Show by counter-example that this fails if K is closed but not compact.

22. $\star$ **Equicontinuity and Arzelà-Ascoli.** Let $\{f_n\}$ be a sequence of continuous functions on $[0,1]$. We say the family is *equicontinuous* if for every $\epsilon > 0$, there exists $\delta > 0$ such that for all n and all $x, y \in [0,1]$:

$$|x-y| < \delta \implies |f_n(x) - f_n(y)| < \epsilon$$

(Crucially, δ does not depend on n).

(a) Suppose $\{f_n\}$ is equicontinuous and *pointwise bounded*. Let $D = \{r_1, r_2, \dots\}$ be the set of rationals in $[0,1]$. Using a diagonalisation argument, show there exists a subsequence $\{f_{n_k}\}$ that converges at every rational point r_i.
(b) Prove that this subsequence actually converges *uniformly* on the entire interval $[0,1]$ to a continuous function f.

23. Connectedness of the Graph. Let $f : \mathbb{R} \to \mathbb{R}$. Let $\Gamma_f = \{(x, f(x)) : x \in \mathbb{R}\} \subseteq \mathbb{R}^2$ be the graph of f.

(a) Prove that if f is continuous, then Γ_f is a connected subset of $\mathbb{R}^2$.
(b) Is the converse true? If the graph Γ_f is connected, must f be continuous?

24. The Pasting Lemma (Open and Local Case). Let $\{U_\alpha\}_{\alpha \in I}$ be a collection of open sets in $\mathbb{R}$ such that $\bigcup_{\alpha \in I} U_\alpha = \mathbb{R}$. Suppose for each $\alpha \in I$, there is a continuous function $f_\alpha : U_\alpha \to \mathbb{R}$ such that $f_\alpha(x) = f_\beta(x)$ for all $x \in U_\alpha \cap U_\beta$.

(a) Prove that the globally defined function $h(x) = f_\alpha(x)$ for $x \in U_\alpha$ is continuous.
(b) Contrast this with the closed case: show that if we have an infinite collection of closed sets K_n, the resulting glued function is not necessarily continuous. (Hint: Consider $K_n = \{1/n\}$ and $K_0 = \{0\}$).

25. $\star$ **Uniform Continuity Preview.** Consider $f(x) = x^3$.

(a) In a continuity proof at a point c, one convenient choice is

$$\delta = \min\left(1, \frac{\epsilon}{3c^2 + 3|c| + 1}\right).$$

Show that as $|c| \to \infty$, the required δ for a fixed ϵ tends to 0.
(b) Prove that it is *impossible* to choose a single $\delta > 0$ that satisfies the continuity condition for $\epsilon = 1$ for all $x \in \mathbb{R}$ simultaneously.
(c) Contrast this with $g(x) = \sin x$ by showing that $|g(x) - g(y)| \le |x-y|$ for all x, y and hence g is uniformly continuous on $\mathbb{R}$.

26. $\star$ **Continuity of Roots.** Let $n \in \mathbb{N}$. Prove that the function $f(x) = x^{1/n}$ is continuous on $[0, \infty)$.

Remark. Split the proof into continuity at $c = 0$ and continuity at $c > 0$. For $c > 0$, use the identity $a^n - b^n = (a-b)(a^{n-1} + a^{n-2}b + \cdots + b^{n-1})$ to bound $|x-c|$.

27. ⋆ **Total Disconnectedness I: Rationals.** Prove that the set $\mathbb{Q}$ of rational numbers is totally disconnected. That is, show that if $S \subseteq \mathbb{Q}$ is a connected set containing more than one element, we arrive at a contradiction.

Chapter 7

Perfect Sets and the Structure of Open Sets

We begin by introducing a class of sets that are, in a topological sense, self-contained and dense within themselves. This leads to the construction of the Cantor set, a pathological counterexample to many intuitive notions in analysis, and finally to a complete characterisation of open sets on the real line.

7.1 Perfect Sets

We have previously defined a limit point of a set S as a point x such that every neighbourhood of x contains a point of S distinct from x. A closed set is one that contains all its limit points. We now consider the converse property: a set where every point is a limit point.

Definition 7.1.1. ***Perfect Set***. A subset $P \subseteq \mathbb{R}$ is said to be a perfect set if it is closed and every point in P is a limit point of P. In the language of derived sets (where P' denotes the set of limit points of P), P is perfect if it is closed and satisfies $P = P'$.

Note. The terminology "perfect" suggests a set that is "complete" (closed) and lacks "isolated defects" (isolated points).

Example 7.1.1. Examples of Perfect Sets.

1. Any closed interval $[a, b]$ with $a < b$ is a perfect set. Every point in the interval can be approached by a sequence of distinct points within the interval.
2. The entire real line $\mathbb{R}$ is a perfect set.
3. Finite sets are never perfect (unless empty). As established in Chapter 2, points in a finite set are isolated.

One might ask whether there exists a perfect set that contains no intervals. The answer, provided by the Cantor set, is affirmative. However, such a set cannot be "small" in terms of cardinality.

***Theorem* 7.1.1. Uncountability of Perfect Sets**. Let P be a non-empty perfect set. Then P is uncountable.

Proof. We proceed by contradiction. Suppose P is countable. We may enumerate its elements as $P = \{x_1, x_2, x_3, \dots\}$. We shall construct a sequence of nested compact intervals whose intersection with P is non-empty but contains no point from the enumeration.

Construction of I_1: Pick an arbitrary $x_1 \in P$. Since x_1 is a limit point, for any $\epsilon > 0$, the neighbourhood $(x_1 - \epsilon, x_1 + \epsilon)$ contains infinitely many points of P. Choose a closed interval $I_1 \subset \mathbb{R}$ centred at x_1

such that $I_1 \cap P \neq \emptyset$. To ensure we can exclude x_1 later, ensure the interior of I_1 contains points of P distinct from x_1.

Inductive Step: Suppose we have constructed a closed interval I_n such that $I_n \cap P \neq \emptyset$. We wish to construct $I_{n+1} \subseteq I_n$ such that $I_{n+1} \cap P \neq \emptyset$ and $x_n \notin I_{n+1}$. Since $I_n \cap P$ is non-empty and P is perfect, the set $I_n \cap P$ must be infinite. If it were finite, then any point $p \in I_n \cap P$ would be isolated in P (we could pick a radius smaller than the distance to any other point in the finite set). This contradicts the fact that every point in a perfect set is a limit point. Thus, I_n contains at least two points of P. We can therefore choose a point $y \in (I_n \cap P)$ such that $y \neq x_n$. Since y is a limit point of P, we can construct a small closed interval I_{n+1} around y such that $I_{n+1} \subseteq I_n, I_{n+1} \cap P \neq \emptyset$, and $x_n \notin I_{n+1}$.

The Intersection: Consider the sequence of sets $K_n = I_n \cap P$. Since each I_n is closed and P is closed, each K_n is closed. Furthermore, $K_n \subseteq I_1$, so the sets are bounded. Thus, each K_n is a compact set. By construction, $K_{n+1} \subseteq K_n$ and each K_n is non-empty. By Theorem 6.2.1, the intersection is non-empty:

$$K = \bigcap_{n=1}^{\infty} K_n = \bigcap_{n=1}^{\infty} (I_n \cap P) \neq \emptyset$$

Let $z \in K$. Then $z \in P$. However, for every n, $z \in I_{n+1}$, which implies $z \neq x_n$ (since we explicitly excluded x_n from I_{n+1}). Thus, z is an element of P that is not in the list $\{x_1, x_2, \dots\}$.

This contradicts the assumption that P was enumerated by (x_n). Therefore, P must be uncountable. ■

7.2 The Cantor Set

We now construct the Cantor Ternary Set, denoted $\mathcal{C}$. This set provides a counterexample to the intuition that "uncountable sets must have non-empty interior" or "any set that occupies no length must be countable". We begin with the closed unit interval $C_0 = [0, 1]$ and define a recursive process of "removing the middle third".

- **Step 1:** Remove the open middle third $(1/3, 2/3)$ from C_0.

$$C_1 = [0, 1/3] \cup [2/3, 1]$$

- **Step 2:** Remove the middle thirds of the two remaining intervals. That is, remove $(1/9, 2/9)$ and $(7/9, 8/9)$.

$$C_2 = [0, 1/9] \cup [2/9, 1/3] \cup [2/3, 7/9] \cup [8/9, 1]$$

- **Step n:** C_n is the union of 2^n disjoint closed intervals, each of length 3^{-n}. C_{n+1} is obtained by removing the open middle third of each interval in C_n.

Definition 7.2.1. ***Cantor Set*.** The Cantor set $\mathcal{C}$ is the intersection of all sets in the construction:

$$\mathcal{C} = \bigcap_{n=0}^{\infty} C_n$$

Figure 7.1: The first few iterations of the Cantor set construction.

Topological Properties

***Proposition* 7.2.1.** *Compactness.* The Cantor set $\mathcal{C}$ is a compact set.

Proof. Each set C_n is a finite union of closed intervals, hence closed. The intersection of any collection of closed sets is closed. Thus $\mathcal{C}$ is closed. Since $\mathcal{C} \subseteq [0,1]$, it is bounded. By Theorem 6.1.1, $\mathcal{C}$ is compact. ■

***Proposition* 7.2.2.** *Non-Empty.* The Cantor set is non-empty.

Proof. The endpoints of the intervals in C_n are never removed. For instance, $0, 1, 1/3, 2/3, 1/9, \dots$ remain in $\mathcal{C}$ forever. Since the set of endpoints is non-empty, $\mathcal{C} \neq \emptyset$. ■

We now address the "size" of the Cantor set in terms of length.

***Proposition* 7.2.3.** *Total Length Removed.* The sum of the lengths of the removed intervals is 1.

Proof. In step 1, we remove one interval of length 1/3. In step 2, we remove two intervals of length 1/9. In step n, we remove 2^{n-1} intervals of length $1/3^n$. The total length removed is the geometric series:

$$L = \sum_{n=1}^{\infty} \frac{2^{n-1}}{3^n} = \frac{1}{3} \sum_{n=0}^{\infty} \left(\frac{2}{3}\right)^n = \frac{1}{3} \cdot \frac{1}{1 - 2/3} = \frac{1}{3} \cdot 3 = 1$$

Since we started with an interval of length 1 and removed total length 1, the remaining set occupies no interval length inside $[0,1]$. In an informal sense, it has total length 0. ■

Despite having zero length, the set is surprisingly robust.

***Theorem* 7.2.1. The Cantor Set is Perfect.** The Cantor set $\mathcal{C}$ is a perfect set.

Proof. We know $\mathcal{C}$ is closed. We must show that every point $x \in \mathcal{C}$ is a limit point of $\mathcal{C}$. Let $x \in \mathcal{C}$ and let $\epsilon > 0$. We must find $y \in \mathcal{C}$ such that $y \neq x$ and $|x - y| < \epsilon$.

Recall that $\mathcal{C} \subseteq C_n$ for all n. The set C_n consists of disjoint intervals of length 3^{-n}. Let J_n be the specific component interval of C_n that contains x. Choose n sufficiently large such that $3^{-n} < \epsilon$. The interval J_n has two endpoints, say a_n and b_n. Since the endpoints of construction intervals are never removed, $a_n, b_n \in \mathcal{C}$. Since $x \in J_n$, it must be that x is equal to one of these endpoints or lies strictly between them. In any case, we can choose an endpoint $y_n \in \{a_n, b_n\}$ such that $y_n \neq x$ (unless J_n has degenerated to a point, which it has not; it has length 3^{-n}). Since $x, y_n \in J_n$, we have $|x - y_n| \leq 3^{-n} < \epsilon$. Thus, we have found a point in $\mathcal{C}$ arbitrarily close to x. Hence, x is a limit point. ■

***Corollary* 7.2.1.** *Cardinality of the Cantor Set.* The Cantor set $\mathcal{C}$ is uncountable.

Proof. $\mathcal{C}$ is a non-empty perfect set. By Theorem 7.1.1, it is uncountable. ■

Finally, we classify the interior of the set.

***Proposition* 7.2.4.** *Empty Interior.* The Cantor set contains no intervals. That is, $\text{int}(\mathcal{C}) = \emptyset$.

Proof. Suppose for contradiction that there exists an open interval $(a, b) \subseteq \mathcal{C}$. Then $(a, b) \subseteq C_n$ for all n. However, C_n is a union of intervals of length 3^{-n}. If (a, b) is contained in C_n, it must be contained in one of these components. Thus $b - a \leq 3^{-n}$. This must hold for all $n \in \mathbb{N}$. Taking the limit as $n \to \infty$, we get $b - a \leq 0$, which contradicts the definition of a non-empty open interval. Thus, $\mathcal{C}$ contains no intervals. ■

The Cantor set is therefore an uncountable, compact, perfect set with empty interior that, in terms of length, occupies no interval length inside $[0,1]$. It serves as a stark warning against relying on geometric intuition when dealing with infinite sets.

The Structure of Open Sets

We conclude this section by resolving the structure of open sets in $\mathbb{R}$. It turns out that *every* open set, no matter how complex, is merely a collection of simple open intervals.

Definition 7.2.2. ***Convex Set***. A set $S \subseteq \mathbb{R}$ is convex if whenever $x, y \in S$ with $x < y$, every point between them is in S; that is,

$$x, y \in S,\ x < y \implies [x, y] \subseteq S.$$

Equivalently, S is convex if for all $x, y \in S$ and $t \in [0, 1]$, we have $tx + (1 - t)y \in S$.

Note. In $\mathbb{R}$, convex sets are precisely the intervals.

***Theorem* 7.2.2. Structure of Open Sets**. Let $G \subseteq \mathbb{R}$ be a non-empty open set. Then G can be written as a countable union of disjoint open intervals:

$$G = \bigsqcup_n I_n$$

where each I_n is an open interval (possibly unbounded). Uniqueness follows because each interval is exactly a connected component of G, and connected components are uniquely determined.

Proof. We construct the intervals using an equivalence relation based on connectivity within G.

Step 1: The Relation. Define a relation $\sim$ on G by:

$$x \sim y \iff [\min(x, y), \max(x, y)] \subseteq G$$

In other words, x and y are related if the entire closed segment connecting them lies within G. This is an equivalence relation.

- **Reflexivity:** $[x, x] = \{x\} \subseteq G$ is trivial.
- **Symmetry:** The definition is symmetric in x and y.
- **Transitivity:** Suppose $x \sim y$ and $y \sim z$. Assume without loss of generality $x < y < z$. Since $[x, y] \subseteq G$ and $[y, z] \subseteq G$, their union $[x, z] \subseteq G$. Thus $x \sim z$.

Step 2: The Equivalence Classes. Let I_x be the equivalence class of a point $x \in G$. Let $I_x = \{y \in G : x \sim y\}$. We claim I_x is an open interval. Since G is open, there exists $\epsilon > 0$ such that $(x-\epsilon, x+\epsilon) \subseteq G$. For any y in this neighbourhood, $[y, x] \subset (x - \epsilon, x + \epsilon) \subseteq G$, so $y \sim x$. Thus I_x contains an interval around x. Furthermore, if $a, b \in I_x$ with $a < b$, then $x \sim a$ and $x \sim b$, implying $[a, x]$ and $[x, b]$ (or appropriate segments) are in G. By convexity of intervals, $[a, b] \subseteq I_x$. It follows that I_x is a convex subset of $\mathbb{R}$ with open neighbourhoods around every point. Thus I_x is an open interval (possibly unbounded).

Step 3: Disjointness. Equivalence classes form a partition. Thus, if x and y are in different classes, $I_x \cap I_y = \emptyset$. If they are in the same class, $I_x = I_y$. So G is the disjoint union of these component intervals $\{I_\alpha\}$.

Step 4: Countability. We must show the collection of intervals is countable. By the density of the rationals, each open interval I_α contains at least one rational number $q_\alpha \in \mathbb{Q}$. Since the intervals are disjoint, these rationals must be distinct. We can define an injective map from the collection of intervals to $\mathbb{Q}$ by choosing one rational per interval. Since $\mathbb{Q}$ is countable, the collection of intervals must be countable.

■

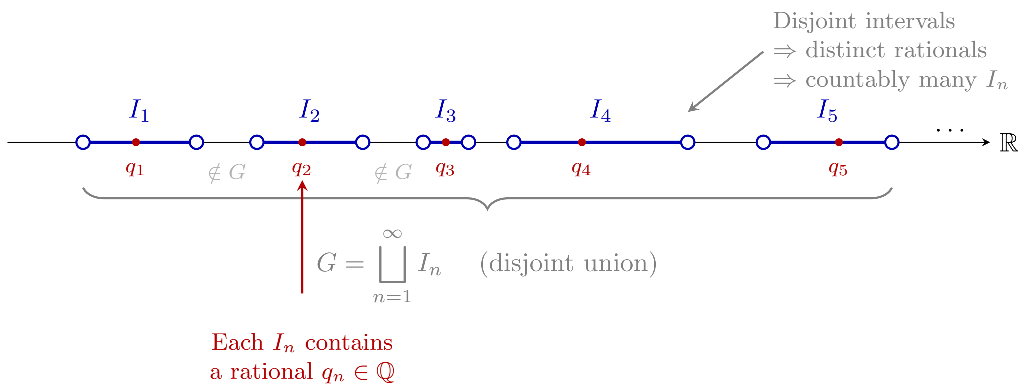


Figure 7.2: Every open set $G \subseteq \mathbb{R}$ decomposes uniquely into countably many disjoint open intervals $\{I_n\}$. By density of $\mathbb{Q}$, each interval contains a rational q_n.

Remark. This theorem allows us to visualise any open set G as a "broken" line, consisting of separated segments. It simplifies the study of measures and integration on open sets, as we can often reduce problems to calculations on single intervals and sum the results.

The Baire Category Theorem

The Baire Category Theorem provides critical insight into the structure of the real line.

Definition 7.2.3. ***Dense Set***. A subset $A \subseteq \mathbb{R}$ is said to be dense in $\mathbb{R}$ if its closure is the entire real line, i.e., $\bar{A} = \mathbb{R}$. Equivalently, A is dense if every non-empty open interval in $\mathbb{R}$ contains an element of A.

We have previously established that $\mathbb{Q}$ and $\mathbb{R} \setminus \mathbb{Q}$ are both dense in $\mathbb{R}$. The Baire Category Theorem asserts that the real line is "too thick" to be constructed from "thin" sets.

***Theorem* 7.2.3. Baire Category Theorem**. Let $(G_n)_{n=1}^{\infty}$ be a sequence of open, dense subsets of $\mathbb{R}$. Then their intersection

$$G = \bigcap_{n=1}^{\infty} G_n$$

is also dense in $\mathbb{R}$.

Proof. We employ a nested interval argument. Let I_0 be an arbitrary non-empty open interval. We aim to show that $G \cap I_0 \neq \emptyset$.

1. Since G_1 is dense and open, the intersection $G_1 \cap I_0$ is a non-empty open set. Thus, we can choose a closed interval J_1 of positive length such that $J_1 \subseteq G_1 \cap I_0$.
2. Now consider G_2. Since G_2 is dense and open, $G_2 \cap \text{int}(J_1)$ is non-empty and open. We can choose a closed interval J_2 of positive length such that $J_2 \subseteq G_2 \cap \text{int}(J_1)$.
3. Proceeding inductively, for each $n \in \mathbb{N}$, we construct a closed interval J_n such that:

$$J_n \subseteq G_n \cap \text{int}(J_{n-1})$$

This yields a nested sequence of non-empty compact intervals $J_1 \supseteq J_2 \supseteq \dots$. By the Nested Interval Property, the intersection is non-empty:

$$\bigcap_{n=1}^{\infty} J_n \neq \emptyset$$

Let x be a point in this intersection. By construction, $x \in J_n \subseteq G_n$ for all n, so $x \in \bigcap G_n = G$. Furthermore, $x \in J_1 \subseteq I_0$. Thus, $G \cap I_0 \neq \emptyset$. Since I_0 was arbitrary, G is dense in $\mathbb{R}$. ■

It is useful to consider the complement perspective.

Definition 7.2.4. ***Nowhere Dense***. A set $S \subseteq \mathbb{R}$ is said to be nowhere dense if the interior of its closure is empty: $\text{int}(\bar{S}) = \emptyset$. Intuitively, a nowhere dense set does not "cluster" sufficiently to fill any interval, no matter how small.

For example, $\mathbb{Z}$ is nowhere dense in $\mathbb{R}$. The Cantor set $\mathcal{C}$, despite being uncountable, is also nowhere dense. theorem 7.2.3 can be reformulated as follows: $\mathbb{R}$ cannot be written as a countable union of nowhere dense sets. In the language of topology, $\mathbb{R}$ is of *Second Category*.

7.3 The Taxonomy of Discontinuities

Earlier, we examined the Dirichlet function, which is discontinuous everywhere. We now reintroduce a function with a more exotic structure: one that is continuous on the irrationals but discontinuous on the rationals.

The Thomae Function

Also known as the Riemann function, this example demonstrates that the set of points of continuity can be dense while the set of discontinuities is also dense.

Definition 7.3.1. ***Thomae Function***. We define $f : \mathbb{R} \to \mathbb{R}$ by:

$$f(x) = \begin{cases} 1 & \text{if } x = 0, \\ 1/n & \text{if } x \in \mathbb{Q} \setminus \{0\},\ x = m/n \text{ in lowest terms}, \\ 0 & \text{if } x \in \mathbb{R} \setminus \mathbb{Q}. \end{cases}$$

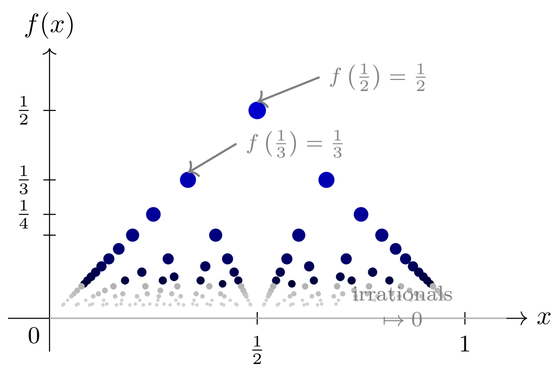


Figure 7.3: The Thomae function on $(0, 1]$. Rationals m/n in lowest terms map to $1/n$; irrationals map to 0. Points accumulate densely near the x-axis as denominators increase.

***Proposition* 7.3.1.** *Discontinuity on Rationals.* The Thomae function is discontinuous at every rational number.

Proof. Let $c \in \mathbb{Q}$. Then $f(c) > 0$. Since the irrationals are dense in $\mathbb{R}$, there exists a sequence (x_k) of irrational numbers converging to c. For each k, $f(x_k) = 0$. Thus, $\lim_{k\to\infty} f(x_k) = 0 \neq f(c)$. By the sequential characterisation of continuity, f is discontinuous at c. ■

***Proposition* 7.3.2.** *Continuity on Irrationals.* The Thomae function is continuous at every irrational number.

Proof. Let c be an irrational number. Then $f(c) = 0$. We wish to show that $\lim_{x\to c} f(x) = 0$. Let $\epsilon > 0$ be given. We must find a $\delta > 0$ such that if $|x - c| < \delta$, then $|f(x)| < \epsilon$.

Consider the value of $f(x)$. If x is irrational, $f(x) = 0 < \epsilon$ trivially. If x is rational, $x = m/n$, then $f(x) = 1/n$. We require $1/n < \epsilon$, or equivalently $n > 1/\epsilon$. Thus, the "bad" rational numbers are those with a denominator n satisfying $1 \leq n \leq 1/\epsilon$.

Let $N = \lceil 1/\epsilon \rceil$. Take the interval bounded by integers $(\lfloor c \rfloor, \lceil c \rceil)$. Within this interval, there are only finitely many rational numbers with denominator $n \leq N$. Let S_ϵ be this finite set of rationals:

$$S_\epsilon = \left\{ \frac{m}{n} : 1 \leq n \leq N, \ \lfloor c \rfloor < \frac{m}{n} < \lceil c \rceil \right\}$$

Since c is irrational, $c \notin S_\epsilon$. Since S_ϵ is finite, the distance between c and the nearest element of S_ϵ is strictly positive. Let $\delta = \min\{|c - r| : r \in S_\epsilon\}$. Now, consider any x such that $|x - c| < \delta$.

- If x is irrational, $|f(x) - 0| = 0 < \epsilon$.
- If x is rational, $x = m/n$. By our choice of δ, $x \notin S_\epsilon$. Consequently, the denominator n must be greater than N. Therefore, $f(x) = 1/n < 1/N \leq \epsilon$.

Thus, f is continuous at c. ■

This result yields a surprising conclusion: the set of discontinuities of the Thomae function is exactly $\mathbb{Q}$.

7.4 Oscillation and the Quantification of Discontinuity

To systematically classify sets of discontinuities, we need a local measure of "how discontinuous" a function is at a point. We use the concept of oscillation.

Definition 7.4.1. ***Diameter***. Let $S \subseteq \mathbb{R}$ be a bounded set. The diameter of S is the supremum of distances between pairs of points in S:

$$\operatorname{diam}(S) = \sup\{|x - y| : x, y \in S\}$$

Definition 7.4.2. ***Oscillation***. Let $f : A \to \mathbb{R}$ be a function and let x be an interior point of A (or, more generally, a limit point of A). The oscillation of f at x, denoted $\operatorname{osc}_x(f)$ (or $\omega_f(x)$), is defined as the limit of the diameter of the image of small closed neighbourhoods around x:

$$\operatorname{osc}_x(f) = \lim_{\delta \to 0^+} \operatorname{diam}\big(f(A \cap [x - \delta, x + \delta])\big)$$

Note. The limit exists because the function $g(\delta) = \operatorname{diam}(f(A \cap B_\delta(x)))$ is monotone decreasing as $\delta \to 0$ and bounded below by 0.

The oscillation provides a precise numerical value representing the "jump" of the function at a point. For the Heaviside step function at 0, the oscillation is 2 (the jump from -1 to 1). For the Thomae function at a rational $x = m/n$, the oscillation is $1/n$.

***Proposition* 7.4.1.** *Oscillation Criterion for Continuity.* A function f is continuous at x if and only if $\operatorname{osc}_x(f) = 0$.

Proof.

($\Longrightarrow$) Suppose f is continuous at x. Let $\epsilon > 0$. There exists $\delta > 0$ such that for all $y \in [x - \delta, x + \delta]$, $|f(y) - f(x)| < \epsilon/2$. By the Triangle Inequality, for any y, z in this neighbourhood:

$$|f(y) - f(z)| \leq |f(y) - f(x)| + |f(x) - f(z)| < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon$$

Thus, $\operatorname{diam}(f(A \cap [x - \delta, x + \delta])) \leq \epsilon$. As $\delta \to 0$, the diameter approaches 0.

($\Longleftarrow$) Suppose $\operatorname{osc}_x(f) = 0$. Let $\epsilon > 0$. By definition, there exists $\delta > 0$ such that $\operatorname{diam}(f(A \cap [x - \delta, x + \delta])) < \epsilon$. Since $x \in [x - \delta, x + \delta]$, for any y in this neighbourhood (intersected with A), we have:

$$|f(y) - f(x)| \leq \operatorname{diam}(f(A \cap [x - \delta, x + \delta])) < \epsilon.$$

Thus, f is continuous at x.

■

Using oscillation, we can characterize exactly which sets can be the set of discontinuities of a function. Let D_f be the set of points where f is discontinuous. We can write:

$$D_f = \{x : \mathrm{osc}_x(f) > 0\} = \bigcup_{n=1}^{\infty} \left\{x : \mathrm{osc}_x(f) \geq \frac{1}{n}\right\}$$

It can be shown that the sets $K_n = \{x : \mathrm{osc}_x(f) \geq 1/n\}$ are closed sets. Consequently, D_f is a countable union of closed sets (known as an F_σ set).

Remark. This characterisation tells us that while $\mathbb{Q}$ can be a set of discontinuities (as it is an F_σ set), the set of irrationals $\mathbb{R} \setminus \mathbb{Q}$ cannot be the set of discontinuities of any function, as it is not an F_σ set. This distinction relies on Theorem 7.2.3.

7.5 Classification of Discontinuities

A function f is continuous at x if the limit $\lim_{t \to x} f(t)$ exists and equals $f(x)$. When this condition fails, we classify the discontinuity based on the behaviour of the limit. For precision, we assume f is defined on an open interval containing x.

Definition 7.5.1. ***Types of Discontinuity***. Let $f : (a, b) \to \mathbb{R}$ and let $x \in (a, b)$.

1. **Removable Discontinuity:** If $\lim_{t \to x} f(t)$ exists but implies $\lim_{t \to x} f(t) \neq f(x)$ (or f is undefined at x).
2. **Jump Discontinuity:** If the one-sided limits $\lim_{t \to x^-} f(t)$ and $\lim_{t \to x^+} f(t)$ both exist but are unequal.
3. **Essential Discontinuity:** If at least one of the one-sided limits does not exist.

The terminology "removable" is justified by the fact that we can modify the function at a single point (defining $f(x) := \lim_{t \to x} f(t)$), to restore continuity. Jump discontinuities represent a finite "break" in the graph, such as in the Heaviside function. Essential discontinuities, such as that of $f(x) = \sin(1/x)$ at $x = 0$, exhibit oscillatory or unbounded behaviour that cannot be tamed by simple redefinition.

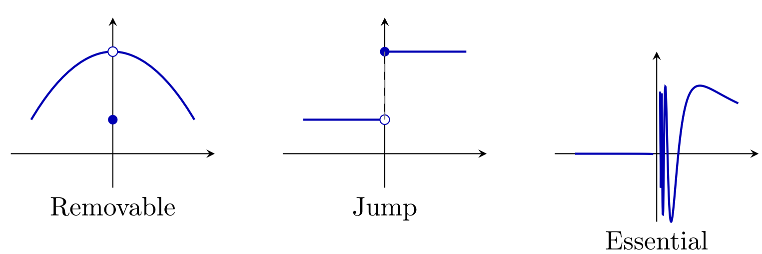


Figure 7.4: Visual classification of discontinuities.

A powerful property of monotone functions is that their behaviour is "tame": they cannot oscillate wildly. Consequently, they cannot possess essential discontinuities.

***Theorem* 7.5.1. Monotone Discontinuities**. Let $f : (a, b) \to \mathbb{R}$ be monotonically increasing. Then:

(i) For any $x \in (a, b)$, the one-sided limits $f(x^-)$ and $f(x^+)$ exist and satisfy:

$$\sup_{a<t<x} f(t) = f(x^-) \leq f(x) \leq f(x^+) = \inf_{x<t<b} f(t)$$

(ii) The only discontinuities of f are jump discontinuities.

(iii) The set of discontinuities of f is countable.

Proof.

Existence of Limits: Let $x \in (a, b)$. Take the set $S_x = \{f(t) : a < t < x\}$. Since f is increasing, this set is bounded above by $f(x)$. By the Least Upper Bound Axiom, $\alpha = \sup S_x$ exists. We claim $\lim_{t \to x^-} f(t) = \alpha$. Let $\epsilon > 0$. By the definition of supremum, there exists $t_0 < x$ such that $\alpha - \epsilon < f(t_0) \leq \alpha$. Since f is increasing, for all $t \in (t_0, x)$, we have $f(t_0) \leq f(t) \leq \alpha$. Thus $\alpha - \epsilon < f(t) \leq \alpha$, which implies $|f(t) - \alpha| < \epsilon$. A strictly analogous argument using the infimum establishes the existence of the right-hand limit $f(x^+)$ and the ordering $f(x^-) \leq f(x) \leq f(x^+)$.

Classification: If f is discontinuous at x, then the equality $f(x^-) = f(x) = f(x^+)$ must fail. Since the limits exist, it is not an essential discontinuity. Since $f(x^-) \leq f(x^+)$, the only failure mode is $f(x^-) < f(x^+)$. Even if $f(x)$ equals one of the limits, the gap between the left and right limits remains. Thus, it is a jump discontinuity.

Countability: Let D be the set of discontinuities. For each $x \in D$, the open interval $J_x = (f(x^-), f(x^+))$ is non-empty. Because f is increasing, these intervals are disjoint. Suppose $x_1 < x_2$ are two points in D. Then $x_1 < t < x_2$ for some t, implying:

$$f(x_1^+) \leq f(t) \leq f(x_2^-)$$

Thus, the interval J_{x_1} lies strictly below J_{x_2}. Since $\mathbb{Q}$ is dense in $\mathbb{R}$, we can choose a distinct rational number $q_x \in J_x$ for each $x \in D$. This defines an injective map from D to $\mathbb{Q}$. Hence, D is countable.

■

7.6 The Structure of the Set of Discontinuities

The Thomae function, which is continuous on the irrationals and discontinuous on the rationals, raises a fundamental structural question: can an arbitrary subset of $\mathbb{R}$ serve as the set of discontinuities for some function? For instance, does there exist a function discontinuous *precisely* on the irrationals?

To answer this, we utilise the concept of oscillation defined in the previous section. Recall that $D_f = \{x : \text{osc}_x(f) > 0\}$.

We first define the class of sets that will characterise discontinuities.

Definition 7.6.1. F_σ ***and*** G_δ ***Sets.***

1. A subset of $\mathbb{R}$ is called an F_σ **set** if it is a countable union of closed sets.
2. A subset of $\mathbb{R}$ is called a G_δ **set** if it is a countable intersection of open sets.

(The notation comes from the French *fermé* (closed) and sum (σ), and *gebiet* (area/open) and *durchschnitt* (intersection)).

Remark. Since the complement of an open set is closed, the complement of a G_δ set is an F_σ set.

Example 7.6.1. Examples of F_σ and G_δ sets.

1. **Open Intervals:** Any open interval (a, b) is an F_σ set.

$$(a, b) = \bigcup_{n=1}^{\infty} \left[a + \frac{1}{n}, b - \frac{1}{n}\right]$$

2. **Closed Intervals:** Any closed interval $[a, b]$ is a G_δ set.

$$[a, b] = \bigcap_{n=1}^{\infty} \left(a - \frac{1}{n}, b + \frac{1}{n}\right)$$

3. **Rationals:** $\mathbb{Q}$ is an F_σ set (it is a countable union of singleton sets, which are closed). It is *not* a G_δ set.

4. **Irrationals:** $\mathbb{R} \setminus \mathbb{Q}$ is a G_δ set (as the complement of an F_σ set). It is *not* an F_σ set.

***Theorem* 7.6.1. Structure of Discontinuities**. Let $f : \mathbb{R} \to \mathbb{R}$ be any function. The set of discontinuities D_f is an F_σ set.

Proof. Recall that f is discontinuous at x if and only if $\text{osc}_x(f) > 0$. We can stratify the set of discontinuities based on the magnitude of the oscillation. Let

$$D_k = \left\{ x \in \mathbb{R} : \text{osc}_x(f) \geq \frac{1}{k} \right\}$$

Then clearly $D_f = \bigcup_{k=1}^{\infty} D_k$. To prove that D_f is F_σ, it suffices to show that each set D_k is closed.

Let (x_n) be a sequence in D_k converging to a limit x. We must show that $x \in D_k$, i.e., $\text{osc}_x(f) \geq 1/k$. Suppose for contradiction that $\text{osc}_x(f) < 1/k$. Then there exists $\delta > 0$ such that the diameter of the image of the neighbourhood $B_\delta(x)$ is strictly less than $1/k$. Since $x_n \to x$, for sufficiently large n, $x_n \in B_{\delta/2}(x)$. Furthermore, since $B_{\delta/2}(x_n) \subseteq B_\delta(x)$, the oscillation at x_n is bounded by the diameter of the image of $B_\delta(x)$. Thus, $\text{osc}_{x_n}(f) \leq \text{diam}(f(B_\delta(x))) < 1/k$. This contradicts the assumption that $x_n \in D_k$. Therefore, $\text{osc}_x(f) \geq 1/k$, so D_k is closed. ■

We now possess the machinery to prove that the sets of discontinuities cannot be arbitrary. Specifically, we can rule out the set of irrational numbers.

***Corollary* 7.6.1.** *Discontinuities on Irrationals.* There is no function $f : \mathbb{R} \to \mathbb{R}$ that is continuous on $\mathbb{Q}$ and discontinuous on $\mathbb{R} \setminus \mathbb{Q}$.

Proof. Suppose such a function exists. Then $D_f = \mathbb{R} \setminus \mathbb{Q}$. By the theorem above, D_f must be an F_σ set. Thus, the irrationals would be a countable union of closed sets: $\mathbb{R} \setminus \mathbb{Q} = \bigcup_{n=1}^{\infty} F_n$. This implies that the real line can be written as:

$$\mathbb{R} = \mathbb{Q} \cup (\mathbb{R} \setminus \mathbb{Q}) = \left(\bigcup_{q \in \mathbb{Q}} \{q\} \right) \cup \left(\bigcup_{n=1}^{\infty} F_n \right)$$

The set $\mathbb{Q}$ is a countable union of singletons (which are closed sets with empty interior). What about the sets F_n? Since $F_n \subseteq \mathbb{R} \setminus \mathbb{Q}$, F_n cannot contain any non-empty open interval (as every open interval contains rationals). Thus, each F_n is a closed set with empty interior; that is, F_n is nowhere dense. Hence $\mathbb{R}$ would be a countable union of nowhere dense sets. Consequently, we have expressed $\mathbb{R}$ as a countable union of nowhere dense sets. This contradicts Theorem 7.2.3, which states that $\mathbb{R}$ is not a countable union of nowhere dense sets (it is of the second category). Therefore, no such function can exist. ■

This result highlights a fundamental asymmetry in the topology of the real line. While we can define a function discontinuous on the "small" set $\mathbb{Q}$ (Thomae's function), we cannot do the inverse, despite both sets being dense. The structural property of being an F_σ set is the decisive obstruction.

7.7 Exercises

1. **Perfect Sets and Compactness.**
   (a) If P is a perfect set and K is a compact set, is the intersection $P \cap K$ always compact?
   (b) Is the intersection $P \cap K$ always perfect? Prove or give a counter-example.
   (c) Does there exist a perfect set consisting entirely of rational numbers?

2. **The Cantor Set Construction.** Let C be the standard Cantor set constructed by removing middle thirds.

(a) Let $x \in C_1$. Prove that there exists $x_1 \in C \cap C_1$ with $x_1 \neq x$ satisfying $|x - x_1| \leq 1/3$.
(b) Generalise this to prove that for each $n \in \mathbb{N}$ and any $x \in C$, there exists $x_n \in C$ different from x such that $|x - x_n| \leq 3^{-n}$.
(c) Show that the set of endpoints of the construction intervals is dense in C.

3. **Cantor-Like Sets.** Repeat the Cantor construction starting with $[0, 1]$, but at each step remove the open middle *fourth* from each component (instead of the third).
   (a) Is the resulting set compact? Is it perfect?
   (b) Compute the total length of this set by summing the lengths of the removed intervals.
   (c) What is the interior of this set?

4. **Total Disconnectedness II: Cantor and Irrationals.** A set E is totally disconnected if for any two distinct points $x, y \in E$, there exist separated sets A and B such that $x \in A$, $y \in B$, and $E = A \cup B$.
   (a) Show that every countable subset of $\mathbb{R}$ is totally disconnected.
   (b) Prove that the Cantor set $\mathcal{C}$ is totally disconnected.

   Remark. Let $x < y$ be in $\mathcal{C}$. Pick a removed interval from the construction that lies between x and y.

   (c) Is the set of irrational numbers totally disconnected?

5. **Separation by Open Sets.** Let A and B be non-empty subsets of $\mathbb{R}$. Prove that if there exist disjoint open sets U and V such that $A \subseteq U$ and $B \subseteq V$, then A and B are separated (i.e., $\bar{A} \cap B = \emptyset$ and $A \cap \bar{B} = \emptyset$).

6. **Connectedness and Closure.**
   (a) Prove that the closure of a connected set is connected.
   (b) Prove that the union of two connected sets with non-empty intersection is connected.

7. **Dense and Nowhere-Dense Sets.**
   (a) Show that a set E is nowhere dense in $\mathbb{R}$ if and only if its complement E^c contains a dense open set.
   (b) Classify the following sets as dense, nowhere dense, or neither:
      (i) $A = \mathbb{Q} \cap [0, 5]$.
      (ii) $B = \{1/n : n \in \mathbb{N}\}$.
      (iii) The set of irrational numbers.
      (iv) The Cantor set.

8. **Constructing Perfect Sets.** Let $\{r_1, r_2, \dots\}$ be an enumeration of the rationals. Let $\epsilon_n = 1/2^n$. Define the open set $O = \bigcup_{n=1}^{\infty}(r_n - \epsilon_n, r_n + \epsilon_n)$. Let $F = O^c$.
   (a) Prove that F is a non-empty closed set consisting only of irrational numbers.
   (b) Does F contain any non-empty open intervals? Is F totally disconnected?
   (c) Is F necessarily perfect? If not, can we modify the construction (e.g., choice of ϵ_n) to ensure F is a non-empty perfect set of irrationals?

9. **Sequential Characterisation of Closure.** Let $E \subseteq \mathbb{R}$.
   (a) Prove that $x \in \bar{E}$ if and only if there exists a sequence (x_n) in E with $x_n \to x$.
   (b) Deduce that E is closed if and only if every convergent sequence in E has its limit in E.

10. **Thomae's Function II: Level Sets.** Let $t(x)$ be the Thomae function.
    (a) For $\epsilon > 0$ and any bounded interval I, show that $E_\epsilon = \{x \in I : t(x) \geq \epsilon\}$ is finite.
    (b) Use this to show that for any $c \in \mathbb{R}$ there exists a sequence of rationals $(x_n) \to c$ with $t(x_n) \to 0$.

11. **Infinite Limits.** We define $\lim_{x \to c} f(x) = \infty$ to mean: $\forall M > 0, \exists \delta > 0$ such that $0 < |x - c| < \delta \implies f(x) > M$.
    (a) Prove that $\lim_{x \to 0} 1/x^2 = \infty$.

(b) Formulate a definition for $\lim_{x\to\infty} f(x) = L$. Using this, show $\lim_{x\to\infty} 1/x = 0$.
(c) Formulate a definition for $\lim_{x\to\infty} f(x) = \infty$ and give an example.

12. Variations on Continuity. Let $f : \mathbb{R} \to \mathbb{R}$.

(a) Define "onetinuous" at c as: $\forall \epsilon > 0$, we can choose $\delta = 1$ such that $|x-c| < \delta \implies |f(x)-f(c)| < \epsilon$. Find a function that is onetinuous everywhere.
(b) Define "equaltinuous" at c as: $\forall \epsilon > 0$, we can choose $\delta = \epsilon$. Find a function that is equaltinuous but not constant.
(c) Define "lesstinuous" at c as: $\forall \epsilon > 0, \exists \delta < \epsilon$. Is every continuous function lesstinuous? Is every lesstinuous function continuous?

13. Composition of Limits. Assume $\lim_{x\to p} f(x) = q$ and $\lim_{x\to q} g(x) = r$.

(a) Prove that the statement holds if g is continuous at q.
(b) Does it hold if only f is continuous?

14. Algebra of Discontinuities. Construct examples or explain why they are impossible:

(a) Two functions f, g discontinuous at 0 such that $f + g$ is continuous at 0.
(b) A function f continuous at 0 and g discontinuous at 0 such that $f + g$ is continuous at 0.
(c) A function f continuous at 0 and g discontinuous at 0 such that fg is continuous at 0.

15. Zero Sets. Let $h : \mathbb{R} \to \mathbb{R}$ be continuous. Prove that the zero set $Z = \{x : h(x) = 0\}$ is a closed set.

16. Cauchy's Functional Equation. Let $f : \mathbb{R} \to \mathbb{R}$ be a function satisfying the additivity property:

$$f(x+y) = f(x) + f(y) \quad \text{for all } x, y \in \mathbb{R}.$$

(a) Prove that $f(0) = 0$ and $f(-x) = -f(x)$.
(b) Prove that $f(qx) = qf(x)$ for all $q \in \mathbb{Q}$ and $x \in \mathbb{R}$.
(c) Prove that if f is continuous at the single point $c = 0$, then f is continuous on all of $\mathbb{R}$.
(d) Conclude that if f is continuous at 0, then $f(x) = cx$ for some constant c.

17. Uniform Continuity IV: More Checks. Decide if the following are uniformly continuous on the given domains.

(a) $f(x) = x^3$ on $\mathbb{R}$. On a bounded interval $[a, b]$.
(b) $f(x) = 1/x$ on $(0, 1)$. On $[1, \infty)$.
(c) $f(x) = \sqrt[3]{x}$ on $\mathbb{R}$.

18. Lipschitz Condition. A function f is M-Lipschitz if $|f(x) - f(y)| \leq M|x - y|$.

(a) Prove that any Lipschitz function is uniformly continuous.
(b) Is the converse true? Consider $f(x) = \sqrt{x}$ on $[0, 1]$.

19. Endpoint Extensions. Let $g : (a, b) \to \mathbb{R}$ be continuous.

(a) Prove that g can be extended to a continuous function on $[a, b]$ if and only if $\lim_{x\to a^+} g(x)$ and $\lim_{x\to b^-} g(x)$ both exist and are finite.
(b) Apply this to $f(x) = \sin(1/x)$ and $h(x) = 1/x$ on $(0, 1)$. Which functions admit continuous extensions to $[0, 1]$?

20. Topological Characterisation II: Basis Test.

(a) Prove that g is continuous on $\mathbb{R}$ if and only if $g^{-1}((a, b))$ is open for every open interval (a, b).
(b) Determine if true (prove or give a counter-example):
   (i) $f(K)$ is compact whenever K is compact.
   (ii) $f(F)$ is closed whenever F is closed.

21. Oscillation Sets in Examples. For a function f, let $D_\alpha = \{x : \text{osc}_x(f) \geq \alpha\}$.

(a) Let t be the Thomae function. Show that for $\alpha > 0$, D_α is contained in the set of rationals with denominator at most $\lfloor 1/\alpha \rfloor$ (in lowest terms). Conclude that D_α is finite on any bounded interval.

(b) Let $f(x) = \sin(1/x)$ for $x \neq 0$ and $f(0) = 0$. Compute D_α for $\alpha > 0$.

22. F_σ and G_δ Operations.

(a) Show that every singleton $\{x\}$ is a G_δ set.
(b) Use this to show that a countable union of G_δ sets need not be a G_δ set.
(c) Show that a countable intersection of F_σ sets need not be an F_σ set.

23. Constructing Discontinuities.

(a) Let F be a closed set. Construct a function f whose set of discontinuities is exactly F.
(b) Let O be an open set. Construct a function g whose set of discontinuities is exactly O.

24. Intermediate Value Property. Let $f : [0,1] \to \mathbb{R}$ be continuous with $f(0) = f(1)$. Prove there exist x, y with $|x - y| = 1/2$ such that $f(x) = f(y)$.

25. Single-Point Continuity. Can a function be continuous for exactly one value of x and discontinuous for all other values?

(a) Construct such a function using a modification of the Dirichlet function (e.g., consider $f(x) = x$ for $x \in \mathbb{Q}$ and $f(x) = 0$ for $x \notin \mathbb{Q}$).
(b) Prove that your example is continuous at exactly one point.
(c) Generalise this: Construct a function that is continuous on exactly the set $\{0, 1, 2\}$ and discontinuous everywhere else.

26. Discontinuous Bijections. The theorem on the continuity of monotone functions implies that a strictly increasing bijection $f : [a, b] \to [c, d]$ must be continuous.

(a) Construct a function $f : (0, 1) \to (0, 1)$ which takes every value $y \in (0, 1)$ exactly once (a bijection), yet is discontinuous for at least one value of x.
(b) Construct a bijection $f : (0, 1) \to (0, 1)$ that is discontinuous *everywhere*.

27. The Stability of the Supremum. Let f be bounded on $[a, b]$. Define the running supremum function $M(x)$ by:

$$M(x) = \sup\{f(t) : a \leq t \leq x\}.$$

(a) Prove that $M(x)$ is a monotonically increasing function.
(b) Prove that if f is continuous at $c \in (a, b)$ and $f(c) < M(c)$, then there is an interval containing c in which $M(x)$ is constant.

Remark. Intuitively, if the current value is strictly below the "high water mark" established earlier, a small movement does not change the record.

28. Midpoint Convexity and Continuity. Let f be defined on an interval (a, b). Suppose f is bounded on (a, b) and satisfies the midpoint convexity inequality for every pair $x_1, x_2 \in (a, b)$:

$$f\left(\frac{x_1 + x_2}{2}\right) \leq \frac{f(x_1) + f(x_2)}{2}.$$

Prove that f is continuous on (a, b).

Remark. This result, due to Jensen, establishes that for convex functions, boundedness implies continuity. Without boundedness, this is false (Hamming functions).

29. The Cantor-Lebesgue Function (The Devil's Staircase). This exercise constructs a continuous, non-decreasing function that maps the Cantor set onto $[0, 1]$ even though the Cantor set occupies no interval length inside $[0, 1]$. Recalling the construction of the Cantor set $\mathcal{C}$, define a function $f : [0, 1] \to [0, 1]$ as follows:

- On the first removed interval $(1/3, 2/3)$, set $f(x) = 1/2$.
- On the next removed intervals $(1/9, 2/9)$ and $(7/9, 8/9)$, set $f(x) = 1/4$ and $f(x) = 3/4$ respectively.
- Continue this process. If (a, b) is one of the 2^{n-1} intervals removed at step n, set $f(x) = k/2^n$ for the appropriate odd integer k.

(a) This definition covers $[0,1] \setminus \mathcal{C}$. Extend f to the whole interval by defining $f(x) = \sup\{f(t) : t \in [0,1] \setminus \mathcal{C}, t < x\}$ for $x \in \mathcal{C}$.
(b) Prove that f is a continuous, non-decreasing function on $[0,1]$ with $f(0) = 0$ and $f(1) = 1$.
(c) Show that f is constant on every interval in the complement of $\mathcal{C}$, yet f is surjective.

30. Sets of Continuity as G_δ Sets. We proved that the set of discontinuities D_f is always an F_σ set (a countable union of closed sets).

(a) Let C_f be the set of points where $f : \mathbb{R} \to \mathbb{R}$ is continuous. Prove that C_f is a G_δ set (a countable intersection of open sets).
(b) Give an example where C_f is dense and G_δ but not open.

Remark. Use the Thomae function.

31. Periodicity II: Value Distribution. Let $f : \mathbb{R} \to \mathbb{R}$ be continuous and periodic with period $T > 0$.

(a) Prove that $f(\mathbb{R})$ is a compact interval $[m, M]$ and that f attains m and M on every interval of length T.
(b) Show that every value $y \in (m, M)$ is attained infinitely many times.

32. ⋆ The Harmonic Comb. Consider the function $f : \mathbb{R} \setminus \{0\} \to \mathbb{R}$ defined by $f(x) = x\lfloor \frac{1}{x} \rfloor$.

(a) **Sketching and Geometry.** Sketch the graph of f on the interval $[-1, 1]$. Clearly identify the behavior of the function in the intervals $(\frac{1}{n+1}, \frac{1}{n}]$. What geometric shapes do the segments of the graph form?
(b) **Limits and Asymptotics.**
   (i) Calculate $\lim_{x\to 0} f(x)$ using the inequality $t - 1 < \lfloor t \rfloor \leq t$.
   (ii) Calculate $\lim_{x\to\infty} f(x)$.
   (iii) Determine the one-sided limits $\lim_{x\to(\frac{1}{n})^+} f(x)$ and $\lim_{x\to(\frac{1}{n})^-} f(x)$ for integer $n \geq 1$.
(c) **Continuity Analysis.**
   (i) Can f be extended to a function continuous at $x = 0$? If so, what value must $f(0)$ take?
   (ii) Classify the discontinuities of f at the points $x = \frac{1}{n}$. Are they removable, jump, or essential?
   (iii) Calculate the oscillation $\operatorname{osc}_c(f)$ at $c = \frac{1}{n}$. Does $\operatorname{osc}_c(f) \to 0$ as $n \to \infty$?
(d) **Uniformity.** Let $\tilde{f}$ be the extension of f to include $\tilde{f}(0) = \lim_{x\to 0} f(x)$. Is $\tilde{f}$ uniformly continuous on the closed interval $[0, 1]$? Why/Why not?

Chapter 8

Further Analytical Properties of Functional Limits

While the fundamental definition of a limit and its algebraic properties have been established, a complete treatment requires a deeper examination of the topological structure of the domain and the behaviour of specific classes of functions, particularly rational powers. Furthermore, the relationship between limits and order (specifically strict inequalities), merits a precise metric justification.

8.1 Cluster Points and Topological Neighbourhoods

We defined limit points (or accumulation points), which can be characterised via sequences of distinct terms. In many texts, these are synonymously referred to as *cluster points*. The existence of a limit $\lim_{x \to c} f(x)$ is predicated on c being a cluster point of the domain; otherwise, the punctured neighbourhood would be empty, rendering the condition vacuously true.

Example 8.1.1. Cluster Points. Consider the set $A = (0, 1)$. The points 0 and 1 are cluster points of A, despite not being elements of A.

- For 0, the sequence $a_n = \frac{1}{n+1}$ consists of elements in $(0, 1)$ and converges to 0.
- For 1, the sequence $b_n = 1 - \frac{1}{n+1}$ consists of elements in $(0, 1)$ and converges to 1.

Indeed, every point $x \in [0, 1]$ is a cluster point of $(0, 1)$. Similarly, every real number $x \in \mathbb{R}$ is a cluster point of the rationals $\mathbb{Q}$, reflecting the density of $\mathbb{Q}$ in $\mathbb{R}$.

To refine the $\epsilon - \delta$ definition, we introduce specific notation for symmetric neighbourhoods.

Notation 8.1.1. Neighbourhood Systems. We denote by $\mathcal{N}_a$ the collection of all open neighbourhoods of a point a. A *symmetric neighbourhood* of a is an interval of the form $(a - \delta, a + \delta)$ for some $\delta > 0$. We denote the collection of such sets by $S\mathcal{N}_a$. A *deleted symmetric neighbourhood*, denoted $S\mathcal{N}_a^*$, excludes the centre point a:

$$V \in S\mathcal{N}_a^* \iff V = (a - \delta, a + \delta) \setminus \{a\}$$

The definition of the limit $\lim_{x \to c} f(x) = L$ can thus be rephrased purely in terms of these sets:

$$\forall U \in S\mathcal{N}_L, \exists V \in S\mathcal{N}_c^* \text{ such that } \forall x \in \text{dom}(f), (x \in V \implies f(x) \in U)$$

This highlights that the limit preserves proximity: establishing a symmetric target U necessitates finding a symmetric source V.

We have previously established the continuity of polynomials. We now extend this to rational powers x^r for $r \in \mathbb{Q}$, specifically for positive bases.

***Proposition* 8.1.1.** *Limits of Rational Powers.* Let $r \in \mathbb{Q}$ and let $f : (0, \infty) \to \mathbb{R}$ be defined by $f(x) = x^r$. Then for any $c > 0$:

$$\lim_{x \to c} x^r = c^r$$

Proof. Let $r = m/k$ where $m, k \in \mathbb{N}$. We apply the Sequential Criterion for limits. Let (x_n) be a sequence in $(0, \infty) \setminus \{c\}$ such that $x_n \to c$. We examine the sequence $x_n^{m/k}$. From the properties of sequences, specifically the limit properties of the k-th root and integer powers:

$$\lim_{n \to \infty} x_n^{1/k} = c^{1/k}$$

Raising this convergent sequence to the power m (using the Product Law for sequences applied m times):

$$\lim_{n \to \infty} (x_n^{1/k})^m = (c^{1/k})^m = c^{m/k}$$

Thus, $\lim_{x \to c} x^{m/k} = c^{m/k}$. For $r \in \mathbb{Q}$ with $r < 0$, we write $x^r = 1/x^{-r}$. Since $-r > 0$, the denominator converges to c^{-r}. Provided $c \neq 0$, the Quotient Law implies the limit is $1/c^{-r} = c^r$. ■

While the Squeeze Theorem addresses non-strict inequalities ($f(x) \leq g(x)$), strict inequalities are generally not preserved in the limit (e.g., $1/n > 0$ yet the limit is 0). However, strict inequality *between the limits* enforces strict inequality between the functions *locally*.

***Proposition* 8.1.2.** *Local Separation.* Let $f, g : X \to \mathbb{R}$ be functions and let c be a cluster point of X. Suppose that:

$$\lim_{x \to c} f(x) = A \quad \text{and} \quad \lim_{x \to c} g(x) = B$$

If $A < B$, then there exists a neighbourhood of c (specifically a $\delta_0 > 0$) such that for all $x \in X$ with $0 < |x - c| < \delta_0$:

$$f(x) < g(x)$$

Proof. We construct disjoint neighbourhoods around A and B. Let ϵ be a positive number such that $3\epsilon < B - A$. For instance, $\epsilon = (B - A)/4$. This ensures that the ϵ-neighbourhoods of A and B are disjoint and separated by a gap of at least ϵ.

1. Since $f(x) \to A$, there exists $\delta_f > 0$ such that for $0 < |x - c| < \delta_f$:
$$|f(x) - A| < \epsilon \implies f(x) < A + \epsilon$$
2. Since $g(x) \to B$, there exists $\delta_g > 0$ such that for $0 < |x - c| < \delta_g$:
$$|g(x) - B| < \epsilon \implies g(x) > B - \epsilon$$

Let $\delta_0 = \min(\delta_f, \delta_g)$. For any $x \in X$ satisfying $0 < |x - c| < \delta_0$:

$$f(x) < A + \epsilon$$

We observe that $B - \epsilon - (A + \epsilon) = B - A - 2\epsilon$. By our choice of ϵ, $2\epsilon < 2(B - A)/3 < B - A$, so $B - A - 2\epsilon > 0$. More simply:

$$f(x) < A + \epsilon < B - \epsilon < g(x)$$

Thus $f(x) < g(x)$. ■

8.2 Exponentials and Logarithms

Let's address the definition of the exponential function a^x for $a > 0$ and $x \in \mathbb{R}$. While integer powers are defined via repeated multiplication and rational powers via roots ($a^{m/n} = \sqrt[n]{a^m}$), the meaning of an expression such as 2^π is not immediately algebraic. We construct these functions by extending the rational powers to the real numbers via continuity, utilising the completeness of $\mathbb{R}$.

Construction of the Exponential Function

Let $a > 0$. The case $a = 1$ is trivial, yielding the constant function $1^x = 1$. We focus on $a > 1$. The case $0 < a < 1$ will follow by reciprocation. Recall from elementary algebra that for rational numbers $r, s \in \mathbb{Q}$, the following identities hold:

$$a^{r+s} = a^r a^s, \quad a^{r-s} = \frac{a^r}{a^s}, \quad (a^r)^s = a^{rs}$$

Our strategy is to define a^x for $x \in \mathbb{R}$ as the limit of a^r as $r \in \mathbb{Q}$ approaches x. To ensure this is well-defined, we must verify the monotonicity and order continuity of the map $r \mapsto a^r$ on $\mathbb{Q}$.

***Lemma* 8.2.1.** *Rational Monotonicity.* Let $a > 1$. For any rational numbers $r_1, r_2 \in \mathbb{Q}$, if $r_1 < r_2$, then $a^{r_1} < a^{r_2}$.

Proof. Let $x, y > 0$ and $n \in \mathbb{N}$. We know that $x < y \iff x^n < y^n$. Since $a > 1$, we have $a^{1/n} > 1$ (as raising 1 to the n-th power yields 1). Consequently, $a^{m/n} > 1$ for all $m, n \in \mathbb{N}$. Thus, $a^q > 1$ for any positive rational q. If $r_1 < r_2$, let $q = r_2 - r_1 \in \mathbb{Q}^+$. Then:

$$\frac{a^{r_2}}{a^{r_1}} = a^{r_2 - r_1} = a^q > 1 \implies a^{r_2} > a^{r_1}$$

■

We now establish that the rational exponential is continuous at 0. This local continuity is sufficient to establish global continuity later via the algebraic properties.

***Lemma* 8.2.2.** *Rational Continuity at 0.* Let $a > 1$. Then $\lim_{\mathbb{Q} \ni r \to 0} a^r = 1$.

Proof. We must show that for any $\epsilon > 0$, there exists $\delta > 0$ such that for all $r \in \mathbb{Q}$ with $0 < |r| < \delta$, we have $|a^r - 1| < \epsilon$. From earlier results, we know that $\lim_{n\to\infty} a^{1/n} = 1$ Thus, there exists an integer n_0 such that for all $n \geq n_0$:

$$1 - \epsilon < a^{-1/n} < a^{1/n} < 1 + \epsilon$$

Set $\delta = 1/n_0$. If $r \in \mathbb{Q}$ satisfies $|r| < \delta$, then $-1/n_0 < r < 1/n_0$. By the monotonicity of rational powers (lemma 8.2.1):

$$1 - \epsilon < a^{-1/n_0} < a^r < a^{1/n_0} < 1 + \epsilon$$

This implies $|a^r - 1| < \epsilon$. ■

This lemma implies that if a sequence of rationals (r_n) converges to a rational r_0, then $a^{r_n} \to a^{r_0}$. We now define a^x for irrational x using the Supremum Principle, effectively filling in the gaps.

Definition 8.2.1. ***The Exponential Function***. Let $a > 1$ and $x \in \mathbb{R}$. We define:

$$a^x = \sup\{a^r : r \in \mathbb{Q}, r < x\}$$

For $0 < b < 1$, we define $b^x = (1/b)^{-x}$.

***Proposition* 8.2.1.** *Equivalence of Definitions.* For $a > 1$ and $x \in \mathbb{R}$, let

$$S_x = \sup\{a^r : r \in \mathbb{Q}, r < x\} \quad \text{and} \quad I_x = \inf\{a^r : r \in \mathbb{Q}, r > x\}$$

Then $S_x = I_x$. Furthermore, if $x \in \mathbb{Q}$, this value coincides with the algebraic definition of a^x.

Proof. The set $\{a^r : r < x\}$ is bounded above by a^R for any rational $R > x$, so S_x exists. Similarly, I_x exists. Clearly $S_x \leq I_x$. For any rationals $r_1 < x < r_2$, we have $a^{r_1} \leq S_x \leq I_x \leq a^{r_2}$. Thus:

$$1 \leq \frac{I_x}{S_x} \leq \frac{a^{r_2}}{a^{r_1}} = a^{r_2 - r_1}$$

We can choose sequences of rationals $r_{1,n} \nearrow x$ and $r_{2,n} \searrow x$. Then $r_{2,n} - r_{1,n} \to 0$. By the continuity at 0, $a^{r_{2,n} - r_{1,n}} \to 1$. By the Squeeze Theorem, $I_x/S_x = 1$, so $S_x = I_x$. If $x \in \mathbb{Q}$, we simply take constant sequences to see the value matches. ■

Properties of the Real Exponential

With the definition established via the supremum of rational powers, we verify that the fundamental properties of exponents extend to the real domain. The continuity of these functions is not merely a property but the bridge that allows algebraic laws to pass from $\mathbb{Q}$ to $\mathbb{R}$.

***Theorem* 8.2.1. Properties of Exponentials**. Let $a > 0$ with $a \neq 1$. The function $f(x) = a^x$ satisfies:

(i) **Continuity:** f is continuous on $\mathbb{R}$.
(ii) **Addition Law:** $a^{x+y} = a^x a^y$ for all $x, y \in \mathbb{R}$.
(iii) **Power Law:** $(a^x)^y = a^{xy}$ for all $x, y \in \mathbb{R}$.
(iv) **Monotonicity:** If $a > 1$, f is strictly increasing. If $0 < a < 1$, f is strictly decreasing.
(v) **Range:** f is a bijection from $\mathbb{R}$ to $(0, \infty)$.

Proof. We assume $a > 1$. The case $0 < a < 1$ follows from the identity $a^x = (1/a)^{-x}$.

(i) Continuity: We first establish continuity at $x = 0$. We know $\lim_{\mathbb{Q}\ni r\to 0} a^r = 1$. For any real x close to 0, we can squeeze it between rationals $p < x < q$ that are arbitrarily close to 0. By monotonicity (proven below independently), $a^p < a^x < a^q$. As $p, q \to 0$, $a^p, a^q \to 1$, so $a^x \to 1$. Thus $\lim_{x\to 0} a^x = 1$. Now consider any $x \in \mathbb{R}$. We use the algebraic identity $a^{x+h} = a^x a^h$ (which holds for rationals and extends to reals by the supremum definition - see below).

$$\lim_{h\to 0} a^{x+h} = \lim_{h\to 0} a^x a^h = a^x \lim_{h\to 0} a^h = a^x \cdot 1 = a^x$$

Thus f is continuous everywhere.

(ii) Addition Law: For $x, y \in \mathbb{R}$, choose sequences of rationals $r_n \to x$ and $s_n \to y$. By continuity (established above), $a^{r_n} \to a^x$ and $a^{s_n} \to a^y$. For rationals, $a^{r_n+s_n} = a^{r_n} a^{s_n}$. Taking limits:

$$a^{x+y} = \lim_{n\to\infty} a^{r_n+s_n} = \lim_{n\to\infty} (a^{r_n} a^{s_n}) = a^x a^y$$

(iii) Power Law: Similarly, using sequences of rationals converging to x and y and the continuity of the exponential, the relation $(a^r)^s = a^{rs}$ extends to reals.

(iv) Monotonicity: Let $x < y$. By the density of rationals, we can choose $r_1, r_2 \in \mathbb{Q}$ such that $x < r_1 < r_2 < y$. By the definition of the exponential:

$$a^x \leq a^{r_1} < a^{r_2} \leq a^y$$

The strict inequality $a^{r_1} < a^{r_2}$ holds for rationals (Lemma 8.2.1). Thus $a^x < a^y$.

(v) Bijectivity: Injectivity follows immediately from strict monotonicity. For Surjectivity: Since $a > 1$, we have $\lim_{n\to\infty} a^n = \infty$ and $\lim_{n\to-\infty} a^n = 0$. Since $f(x) = a^x$ is continuous, by the Intermediate Value Theorem, it attains every value between these limits. Thus, the range is $(0, \infty)$.

■

***Lemma* 8.2.3.** *Different Bases.* If $a, b > 0$, then $a^x b^x = (ab)^x$ for all $x \in \mathbb{R}$.

Proof. This identity holds for rational exponents. Since a^x, b^x and $(ab)^x$ are continuous functions of x, the identity extends to all real x by continuity. ■

Since the exponential function $f : \mathbb{R} \to (0, \infty)$ is bijective (for $a \neq 1$), it possesses a unique inverse.

Definition 8.2.2. ***Logarithm***. Let $a > 0, a \neq 1$. The logarithm to base a, denoted $\log_a : (0, \infty) \to \mathbb{R}$, is the inverse of the exponential function. It is defined by the relation:

$$y = a^x \iff x = \log_a y$$

The natural logarithm, denoted ln or log, corresponds to the base e (Euler's number). The notation lg is often used for base 10.

The properties of the logarithm follow directly from the properties of the exponential function.

Theorem **8.2.2. Properties of Logarithms**. Let $a > 0, a \neq 1$.

(i) **Product to Sum:** $\log_a(xy) = \log_a x + \log_a y$ for all $x, y > 0$.
(ii) **Quotient:** $\log_a(x/y) = \log_a x - \log_a y$.
(iii) **Power:** $\log_a(x^\alpha) = \alpha \log_a x$ for any $\alpha \in \mathbb{R}$.
(iv) **Change of Base:** For any $b > 0, b \neq 1$:

$$\log_b y = \frac{\log_a y}{\log_a b}$$

(v) **Monotonicity:** If $a > 1$, $\log_a$ is strictly increasing. If $0 < a < 1$, it is strictly decreasing.
(vi) **Continuity:** $\lim_{n\to\infty} y_n = y > 0 \implies \lim_{n\to\infty} \log_a y_n = \log_a y$. (This follows because $\log_a$ is the inverse of a continuous strictly monotone function).

Proof. We prove (iv) and (v).

(iv) Let $x = \log_b y$. Then $b^x = y$. Taking $\log_a$ of both sides:

$$\log_a(b^x) = \log_a y$$

Using the power property (derived from $(a^u)^v = a^{uv}$): $x \log_a b = \log_a y$. Thus $x = \frac{\log_a y}{\log_a b}$.

(v) Assume $a > 1$. Let $0 < y_1 < y_2$. Let $x_1 = \log_a y_1$ and $x_2 = \log_a y_2$. If $x_1 \geq x_2$, then since a^x is increasing, $y_1 = a^{x_1} \geq a^{x_2} = y_2$, a contradiction. Thus $x_1 < x_2$.

■

We can now define the power function $f(x) = x^s$ for any real exponent s, where the base x varies.

Note. For irrational s, the domain is restricted to $x > 0$ to ensure real-valued outputs.

Theorem **8.2.3. Continuity of Power Functions**. Fix $s \in \mathbb{R}$. The function $f : (0, \infty) \to \mathbb{R}$ defined by $f(x) = x^s$ is continuous. Furthermore, if $x_n \to c > 0$ and $s_n \to s$, then:

$$\lim_{n\to\infty} x_n^{s_n} = c^s$$

Proof. We express the power function using the natural exponential and logarithm:

$$x^s = (e^{\ln x})^s = e^{s \ln x}$$

Let $y_n = s_n \ln x_n$. Since ln is continuous, $\ln x_n \to \ln c$. Since the product of convergent sequences converges, $y_n = s_n \ln x_n \to s \ln c$. Since the exponential function is continuous, $e^{y_n} \to e^{s \ln c} = c^s$. ■

8.3 Fundamental Transcendental Limits

While the algebraic limit laws suffice for polynomial and rational functions, analysis frequently demands the evaluation of limits involving exponential, logarithmic, and trigonometric functions. These "fundamental limits" often take the indeterminate form 1^∞ or $0/0$ and cannot be evaluated by direct substitution. In this section, we derive these limits, connecting the discrete definition of Euler's number e to its continuous functional counterparts.

The Continuous Exponential Limit

In our study of sequences, we defined the Euler number as the limit of a sequence:

$$e = \lim_{n\to\infty}\left(1+\frac{1}{n}\right)^n$$

We now extend this convergence to the continuous variable $x \to \infty$.

***Theorem* 8.3.1. The Continuous Definition of e.**

$$\lim_{x\to\infty}\left(1+\frac{1}{x}\right)^x = e$$

Proof. We rely on the sequential limit and the properties of the floor function (integer part). Let $\lfloor x\rfloor$ denote the greatest integer less than or equal to x. For $x \geq 1$, we have:

$$\lfloor x\rfloor \leq x < \lfloor x\rfloor + 1$$

Taking reciprocals reverses the inequalities:

$$\frac{1}{\lfloor x\rfloor + 1} < \frac{1}{x} \leq \frac{1}{\lfloor x\rfloor}$$

Adding 1 to each term:

$$1+\frac{1}{\lfloor x\rfloor + 1} < 1+\frac{1}{x} \leq 1+\frac{1}{\lfloor x\rfloor}$$

We raise these terms to powers. Since all bases exceed 1 and the exponent is positive, raising preserves inequalities. However, we must be careful with the exponents to maintain the inequality direction. We define $n = \lfloor x\rfloor$.

$$\left(1+\frac{1}{n+1}\right)^n < \left(1+\frac{1}{x}\right)^x < \left(1+\frac{1}{n}\right)^{n+1}$$

We analyse the bounding sequences as $n \to \infty$ (which occurs as $x \to \infty$):

1. **Upper Bound:**
$$\left(1+\frac{1}{n}\right)^{n+1} = \left(1+\frac{1}{n}\right)^n\left(1+\frac{1}{n}\right) \to e\cdot 1 = e.$$
2. **Lower Bound:**
$$\left(1+\frac{1}{n+1}\right)^n = \frac{(1+\frac{1}{n+1})^{n+1}}{1+\frac{1}{n+1}} \to \frac{e}{1} = e.$$

By the Squeeze Theorem for functions (Theorem 4.6.1), the limit is e. ■

***Corollary* 8.3.1.** *Limit at Negative Infinity.* For sufficiently large negative x, the base $1 + 1/x$ remains positive.

$$\lim_{x\to-\infty}\left(1+\frac{1}{x}\right)^x = e$$

Proof. Let $y = -x$. As $x \to -\infty$, $y \to \infty$.

$$\left(1+\frac{1}{x}\right)^x = \left(1-\frac{1}{y}\right)^{-y} = \left(\frac{y-1}{y}\right)^{-y} = \left(\frac{y}{y-1}\right)^y = \left(1+\frac{1}{y-1}\right)^y$$

Let $z = y - 1$. As $y \to \infty$, $z \to \infty$.

$$\left(1+\frac{1}{z}\right)^{z+1} = \left(1+\frac{1}{z}\right)^z\left(1+\frac{1}{z}\right)$$

The first factor approaches e and the second approaches 1. Thus the limit is e. ■

Limits at Zero

Using the substitution $t = 1/x$, we can translate limits at infinity to limits at zero. This yields the standard forms used in differentiation.

***Theorem* 8.3.2. The Exponential Limit at 0.**

$$\lim_{x\to 0}(1+x)^{1/x} = e$$

Proof. We invoke the One-Sided Continuity Criterion. We must show the limit is e as $x \to 0^+$ and $x \to 0^-$.

- **Right Limit ($x \to 0^+$):** Let $t = 1/x$. As $x \to 0^+$, $t \to \infty$.

$$\lim_{x\to 0^+}(1+x)^{1/x} = \lim_{t\to\infty}\left(1+\frac{1}{t}\right)^t = e$$

- **Left Limit ($x \to 0^-$):** Let $t = 1/x$. As $x \to 0^-$, $t \to -\infty$.

$$\lim_{x\to 0^-}(1+x)^{1/x} = \lim_{t\to-\infty}\left(1+\frac{1}{t}\right)^t = e$$

Since both one-sided limits are e, the limit exists and equals e. ■

The following limits effectively calculate the derivatives of $\ln x$ at $x = 1$ and e^x at $x = 0$, forming the basis of calculus for transcendental functions.

***Theorem* 8.3.3. The Logarithmic Limit.**

$$\lim_{x\to 0}\frac{\ln(1+x)}{x} = 1$$

Proof. We use the continuity of the logarithm function.

$$\frac{\ln(1+x)}{x} = \frac{1}{x}\ln(1+x) = \ln\left((1+x)^{1/x}\right)$$

Let $u = (1+x)^{1/x}$. As $x \to 0$, we established that $u \to e$. Since ln is continuous at e:

$$\lim_{x\to 0}\ln\left((1+x)^{1/x}\right) = \ln\left(\lim_{x\to 0}(1+x)^{1/x}\right) = \ln(e) = 1$$

■

***Theorem* 8.3.4. The Exponential Difference Quotient.**

$$\lim_{x\to 0}\frac{e^x-1}{x} = 1$$

Proof. We employ a change of variable. Let $y = e^x - 1$. As $x \to 0$, $e^x \to 1$, so $y \to 0$. Rearranging for x:

$$y + 1 = e^x \implies x = \ln(1+y)$$

Substituting this into the limit:

$$\lim_{x\to 0}\frac{e^x-1}{x} = \lim_{y\to 0}\frac{y}{\ln(1+y)} = \lim_{y\to 0}\frac{1}{\frac{\ln(1+y)}{y}}$$

Using the Quotient Law and the Logarithmic Limit derived above:

$$\frac{1}{\lim_{y\to 0}\frac{\ln(1+y)}{y}} = \frac{1}{1} = 1$$

■

***Theorem* 8.3.5. General Power Limit**. For any $\alpha \in \mathbb{R}$,

$$\lim_{x \to 0} \frac{(1+x)^\alpha - 1}{x} = \alpha$$

Proof. We write the power as an exponential: $(1+x)^\alpha = e^{\alpha \ln(1+x)}$. Let $u = \alpha \ln(1+x)$. As $x \to 0$, $\ln(1+x) \to 0$, so $u \to 0$. We rewrite the expression as a product of two known limits:

$$\frac{(1+x)^\alpha - 1}{x} = \frac{e^u - 1}{x} = \frac{e^u - 1}{u} \cdot \frac{u}{x}$$

Substituting $u = \alpha \ln(1+x)$:

$$= \frac{e^u - 1}{u} \cdot \frac{\alpha \ln(1+x)}{x} = \alpha \left(\frac{e^u - 1}{u} \right) \left(\frac{\ln(1+x)}{x} \right)$$

As $x \to 0$, $u \to 0$. The first term approaches 1 (Exponential Difference Quotient) and the second term approaches 1 (Logarithmic Limit).

$$\text{Limit} = \alpha \cdot 1 \cdot 1 = \alpha$$

■

We conclude with a sophisticated limit that requires the composition of the results above.

Example 8.3.1. Compound Exponential. Evaluate the limit:

$$L = \lim_{x \to \infty} \left(1 + \frac{x}{x^2 + 1} \right)^{2x}$$

Let us discretise this using a sequence $x_n \to \infty$.

$$a_n = \left(1 + \frac{x_n}{x_n^2 + 1} \right)^{2x_n}$$

We define $y_n = \frac{x_n}{x_n^2+1}$. Note that as $n \to \infty$, $y_n \to 0$. We can rewrite the expression in the form of the standard limit $(1 + y_n)^{1/y_n}$:

$$a_n = \left[(1 + y_n)^{\frac{1}{y_n}} \right]^{s_n}$$

We must determine the exponent s_n.

$$s_n = 2x_n \cdot y_n = 2x_n \cdot \frac{x_n}{x_n^2 + 1} = \frac{2x_n^2}{x_n^2 + 1} = \frac{2}{1 + 1/x_n^2}$$

As $n \to \infty$, $s_n \to 2$. Using the continuity of the power function (Theorem 8.2.3):

$$L = \lim_{n \to \infty} \left[(1 + y_n)^{\frac{1}{y_n}} \right]^{s_n} = \left[\lim_{y \to 0} (1 + y)^{1/y} \right]^{\lim s_n} = e^2$$

8.4 Exercises

1. **Cluster Points of Irrationals.** Prove that every real number $x \in \mathbb{R}$ is a cluster point of the set of irrational numbers $\mathbb{R} \setminus \mathbb{Q}$. Consequently, construct a sequence of irrational numbers converging to x.

   Remark. Recall that between any two real numbers there lies an irrational number. This exercise confirms that $\mathbb{R} \setminus \mathbb{Q}$ is dense in $\mathbb{R}$.

2. **Limit Equivalence.** Let $f : D \to \mathbb{R}$ and let c be a limit point of D. Prove that the following are equivalent:

(i) The functional limit $\lim_{x\to c} f(x)$ exists.

(ii) For every sequence $(x_n) \subseteq D \setminus \{c\}$ converging to c, the sequence $(f(x_n))$ is Cauchy.

Remark. This is known as the Cauchy Criterion for Functional Limits. It allows us to prove existence without knowing the limit value L.

3. The Vanishing Ruler. Consider the function $f : \mathbb{R} \to \mathbb{R}$ defined by:

$$f(x) = \begin{cases} x & \text{if } x \in \mathbb{Q} \\ 0 & \text{if } x \notin \mathbb{Q} \end{cases}$$

(a) Prove that $\lim_{x\to 0} f(x) = 0$.

(b) Prove that for any $c \neq 0$, the limit $\lim_{x\to c} f(x)$ does not exist.

Remark. For (a), use the Squeeze Theorem. For (b), construct two sequences converging to c with different limits.

4. Generalised Decay.

(a) Prove that for any $n \in \mathbb{N}$, $\lim_{x\to\infty} \frac{1}{x^n} = 0$.

(b) Let $P(x)$ be a polynomial of degree $k < n$. Prove that $\lim_{x\to\infty} \frac{P(x)}{x^n} = 0$.

5. Limits of Roots.

(a) Prove that $\lim_{x\to 0^+} \sqrt{x} = 0$ using the $\epsilon - \delta$ definition.

(b) Prove that $\lim_{x\to 1} x^{1/3} = 1$.

(c) **Challenge:** Prove that $\lim_{x\to\infty} (x^{1/n} - 1) = 0$ is false, but $\lim_{n\to\infty} (x^{1/n} - 1) = 0$ for fixed $x > 0$.

6. Sequential Exponentials. Let (x_n) and (y_n) be convergent sequences with limits x and y respectively, where $x > 0$. Prove that:

$$\lim_{n\to\infty} x_n^{y_n} = x^y$$

Remark. Utilise the continuity of the exponential composition $x^y = \exp(y \ln x)$.

7. The Continuous Definition of e. Prove that $\lim_{n\to\infty} \left(1 + \frac{x}{n}\right)^n = e^x$ for any $x \in \mathbb{R}$.

Remark. You may assume the standard limit $\lim_{t\to 0} (1 + t)^{1/t} = e$. Use the substitution $t = x/n$.

Part II

Differentiation

Chapter 9

Differentiation

Finally (I hear you say), we arrive at differentiation. The classical motivation is the need to quantify instantaneous rates of change. Continuity prevents breaks in a function, while differentiability imposes a stronger local linearity.

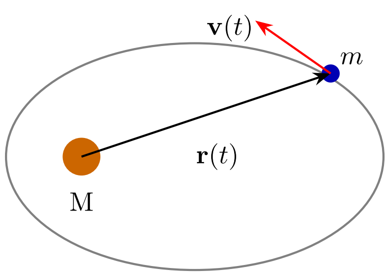
Motivation: The Problem of Instantaneous Velocity

Consider the classical Kepler problem of two celestial bodies, a planet m and a star M. To determine the trajectory of the planet, Newton formulated the law of motion $ma = F$ and the law of universal gravitation. To utilise these laws, one must express acceleration (otherwise known as the rate of change of velocity), in terms of position. This necessitates a precise definition of instantaneous velocity.

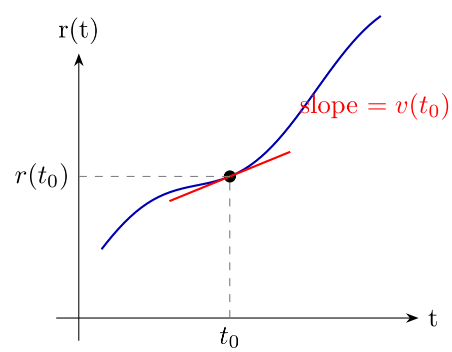
The simplest form of motion is uniform rectilinear motion. If a body moves without external forces, its displacement is linear in time:

$$\mathbf{r}(t) - \mathbf{r}(t_0) = \mathbf{v} \cdot (t - t_0)$$

Here, $\mathbf{v}$ is the constant velocity vector. In general motion, velocity varies. However, physical intuition suggests that over an infinitesimal time interval, the motion is approximately uniform. If we observe the trajectory over a sufficiently small interval around t_0, the curve is indistinguishable from a straight line.



Two-body system



Instantaneous velocity

Mathematically, we seek a vector $\mathbf{v}(t_0)$ such that:

$$\mathbf{r}(t) - \mathbf{r}(t_0) \approx \mathbf{v}(t_0)(t - t_0)$$

as $t \to t_0$. Precisely, the error in this approximation must vanish faster than the time step itself. This reduces the physical problem of velocity to the geometric problem of finding the tangent to a curve, and the algebraic problem of finding the best linear approximation to a function.

9.1 The Derivative

We begin by formalising the notion of the rate of change for a real-valued function.

Definition 9.1.1. ***Derivative***. Let $I \subseteq \mathbb{R}$ be an interval not reducing to a singleton, and let $c \in I$. A function $f : I \to \mathbb{R}$ is said to be differentiable at c if the limit of the Newton quotient exists:

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

If this limit exists, we denote it by $f'(c)$ (or $\frac{df}{dx}(c)$) and call it the derivative of f at c.

Note. If c is an endpoint of the interval I, the limit is understood to be the appropriate one-sided limit.

Geometric Interpretation

Geometrically, the Newton quotient represents the slope of the *secant line* passing through the points $(c, f(c))$ and $(x, f(x))$.

$$m_{\text{sec}} = \frac{f(x) - f(c)}{x - c}$$

As x approaches c, the point $(x, f(x))$ slides along the curve towards $(c, f(c))$. If the function is smooth, the secant lines converge to a limiting position. This limiting line is the *tangent line* to the graph at c, and its slope is $f'(c)$.

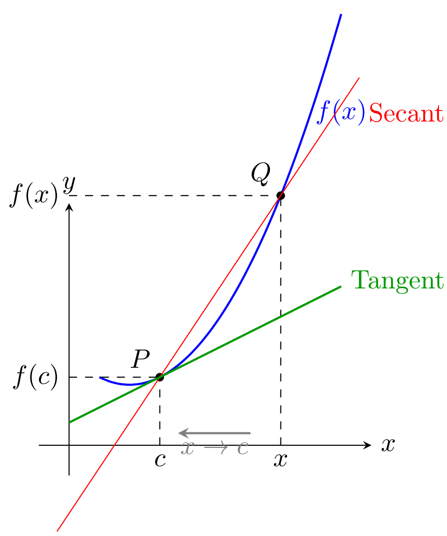


Figure 9.1: The derivative as the limit of the slope of secant lines. As $x \to c$, the red secant line approaches the green tangent line.

9.2 Linear Approximation

While the geometric interpretation of the tangent slope is intuitive, the most powerful perspective for analysis is to view the derivative as the *best linear approximation*.

If a function f is differentiable at c, it behaves locally like a linear function. We can manipulate the definition of the derivative to make this explicit. Assume $f'(c)$ exists. By the properties of limits, we can define an error function $\eta(x)$ such that:

$$\frac{f(x) - f(c)}{x - c} - f'(c) = \eta(x)$$

where $\lim_{x \to c} \eta(x) = 0$. Rearranging this equation yields:

$$f(x) - f(c) = f'(c)(x - c) + \eta(x)(x - c)$$

or equivalently:

$$f(x) = \underbrace{f(c) + f'(c)(x - c)}_{\text{Linear Approximation}} + \underbrace{\eta(x)(x - c)}_{\text{Error Term}}$$

The term $L(x) = f(c) + f'(c)(x - c)$ represents the line passing through $(c, f(c))$ with slope $f'(c)$. The residual term $r(x) = \eta(x)(x - c)$ represents the deviation of the function from this line. Crucially, not only does the error $r(x)$ tend to zero as $x \to c$, but it does so *faster* than the displacement $x - c$.

$$\lim_{x \to c} \frac{r(x)}{x - c} = \lim_{x \to c} \eta(x) = 0$$

Definition 9.2.1. ***Differential***. Let f be differentiable at x. The differential of f, denoted df, is defined by:

$$df = f'(x)\, dx$$

This identifies the linear part of the change in f under a change in x.

Landau's Little-o Notation

To formalise the notion of an error term vanishing "faster" than the variable, we introduce the standard asymptotic notation.

Definition 9.2.2. ***Little-o Notation***. Let $\phi : (-\epsilon, \epsilon) \to \mathbb{R}$. We say that $\phi(h)$ is $o(h)$ (read "little-o of h") as $h \to 0$ if:

$$\lim_{h \to 0} \frac{\phi(h)}{h} = 0$$

This notation captures the scale of the function near zero. If $\phi(h) = o(h)$, then $\phi(h)$ is negligible compared to h for sufficiently small h. Using this notation, with $h = x - c$, we can reformulate differentiability.

***Proposition* 9.2.1.** *Carathéodory's Characterisation*. A function f is differentiable at c if and only if there exists a constant $\lambda \in \mathbb{R}$ and a function $r(h)$ such that for all h in a neighbourhood of 0:

$$f(c + h) = f(c) + \lambda h + r(h)$$

where $r(h) = o(h)$ as $h \to 0$. In this case, $\lambda = f'(c)$.

This characterisation emphasises that $f'(c)$ is the unique coefficient λ that makes the linear map $h \mapsto \lambda h$ a "good" approximation of the increment $f(c + h) - f(c)$. Specifically, differentiability implies:

$$f(c + h) \approx f(c) + f'(c)h$$

This perspective is fundamental because it separates the linear part of the change from the higher-order corrections. In applied mathematics and physics, this justifies the process of *linearisation*, where a non-linear system is approximated by a linear one near an equilibrium point.

Example 9.2.1. Linearisation of the Square. Consider $f(x) = x^2$ at c.

$$f(c + h) = (c + h)^2 = c^2 + 2ch + h^2$$

Here, $f(c) = c^2$. The term linear in h is $2ch$, implying $f'(c) = 2c$. The error term is h^2. We check if $h^2 = o(h)$:

$$\lim_{h \to 0} \frac{h^2}{h} = \lim_{h \to 0} h = 0$$

Thus, the condition is satisfied. The approximation $x^2 \approx c^2 + 2c(x - c)$ is the best linear fit near c.

Example 9.2.2. Non-Differentiability of Absolute Value. Consider $f(x) = |x|$ at $c = 0$.

$$f(0 + h) = |h|$$

Can we write $|h| = 0 + \lambda h + o(h)$? This would require $\lim_{h \to 0} \frac{|h| - \lambda h}{h} = 0$, or $\lim_{h \to 0} \left(\frac{|h|}{h} - \lambda \right) = 0$. However, $\frac{|h|}{h}$ is $\text{sgn}(h)$, which is 1 for $h > 0$ and -1 for $h < 0$. No single constant λ can satisfy the limit for both sides. Thus, $|x|$ does not admit a linear approximation at 0 and is not differentiable there.

Differentiability and Continuity

Example 9.2.3. The Weierstrass Function. There exist continuous functions that are nowhere differentiable. A classical example is

$$f(t) = \sum_{n=0}^{\infty} \frac{\cos(5^n t)}{2^n}.$$

The function is continuous everywhere and differentiable nowhere.

The existence of a linear approximation imposes a strict constraint on the local behaviour of the function. We have already seen continuous functions that are nowhere differentiable; see example 9.2.3. However, the converse relationship is rigid.

***Theorem* 9.2.1. Differentiability implies Continuity**. If f is differentiable at c, then f is continuous at c.

Proof. We use the algebraic limit laws (theorem 4.5.1). We wish to show $\lim_{x \to c} f(x) = f(c)$, or equivalently $\lim_{x \to c} (f(x) - f(c)) = 0$. For $x \neq c$, we can write:

$$f(x) - f(c) = \frac{f(x) - f(c)}{x - c} \cdot (x - c)$$

Taking the limit as $x \to c$:

$$\lim_{x \to c} (f(x) - f(c)) = \lim_{x \to c} \left(\frac{f(x) - f(c)}{x - c} \right) \cdot \lim_{x \to c} (x - c)$$

Since f is differentiable, the first limit is $f'(c)$. The second limit is 0.

$$\lim_{x \to c} (f(x) - f(c)) = f'(c) \cdot 0 = 0$$

Thus f is continuous at c. ■

Remark. The converse is false, as in the example $f(x) = |x|$ at $x = 0$. Continuity is necessary but not sufficient for differentiability.

9.3 Basic Properties of the Derivative

We turn to structural properties of differentiable functions; the linearisation perspective $f(c + h) = f(c) + f'(c)h + o(h)$ governs them.

Just as limits respect algebraic operations, so too does differentiation. Let $f, g : I \to \mathbb{R}$ be functions differentiable at $c \in I$.

***Theorem* 9.3.1. Arithmetic of Derivatives**.

1. **Sum Rule:**

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

.

2. **Product Rule (Leibniz Rule):**

$$(fg)'(c) = f(c)g'(c) + f'(c)g(c)$$

.

3. **Reciprocal Rule:** If $g(c) \neq 0$, then

$$\left(\frac{1}{g}\right)'(c) = -\frac{g'(c)}{(g(c))^2}$$

.

4. **Quotient Rule:** If $g(c) \neq 0$, then

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

.

Proof. We demonstrate the Product Rule and the Reciprocal Rule; the Quotient Rule follows by combining them.

Product Rule: Consider the Newton quotient for the product fg:

$$\frac{f(c+h)g(c+h) - f(c)g(c)}{h}$$

We add and subtract the cross-term $f(c+h)g(c)$ in the numerator:

$$= \frac{f(c+h)g(c+h) - f(c+h)g(c) + f(c+h)g(c) - f(c)g(c)}{h}$$

$$= f(c+h)\frac{g(c+h) - g(c)}{h} + g(c)\frac{f(c+h) - f(c)}{h}$$

As $h \to 0$, $f(c+h) \to f(c)$ by continuity. The difference quotients converge to $g'(c)$ and $f'(c)$ respectively. Thus, the limit is $f(c)g'(c) + g(c)f'(c)$.

Reciprocal Rule: Since g is differentiable at c and $g(c) \neq 0$, continuity ensures there exists a neighbourhood of c where $g(x) \neq 0$, so the function $1/g$ is well-defined locally.

$$\frac{\frac{1}{g(c+h)} - \frac{1}{g(c)}}{h} = \frac{g(c) - g(c+h)}{hg(c+h)g(c)} = -\frac{1}{g(c+h)g(c)} \cdot \frac{g(c+h) - g(c)}{h}$$

Taking the limit as $h \to 0$ yields $-\frac{1}{(g(c))^2} \cdot g'(c)$.

■

The Chain Rule The differentiation of composite functions is handled by the Chain Rule. While often memorised via the notation

$$\frac{dy}{dx} = \frac{dy}{du}\frac{du}{dx},$$

a proof using Newton quotients requires care, as the intermediate difference Δu could be zero. The linear approximation framework bypasses this difficulty elegantly.

***Theorem* 9.3.2. The Chain Rule**. Let $f : I \to \mathbb{R}$ and $g : J \to \mathbb{R}$ be functions such that $f(I) \subseteq J$. If f is differentiable at $c \in I$ and g is differentiable at $d = f(c) \in J$, then the composition $h = g \circ f$ is differentiable at c, and:

$$h'(c) = g'(f(c)) \cdot f'(c)$$

Proof. Since f is differentiable at c, for small k we have:

$$f(c+k) - f(c) = f'(c)k + \phi(k)k \tag{9.1}$$

where $\phi(k) \to 0$ as $k \to 0$, and we define $\phi(0) = 0$. Similarly, since g is differentiable at $d = f(c)$, for small η we have:

$$g(d+\eta) - g(d) = g'(d)\eta + \psi(\eta)\eta \tag{9.2}$$

where $\psi(\eta) \to 0$ as $\eta \to 0$, and $\psi(0) = 0$.

Let k be a small non-zero increment. Let $\eta = f(c+k) - f(c)$. By the continuity of f, $\eta \to 0$ as $k \to 0$. Substituting this η into eq. (9.2):

$$g(f(c+k)) - g(f(c)) = g'(d)[f(c+k) - f(c)] + \psi(\eta)[f(c+k) - f(c)]$$

Using eq. (9.1) to substitute for $[f(c+k) - f(c)]$:

$$(g \circ f)(c+k) - (g \circ f)(c) = g'(d)[f'(c)k + \phi(k)k] + \psi(\eta)[f'(c)k + \phi(k)k]$$

Dividing by k:

$$\frac{(g \circ f)(c+k) - (g \circ f)(c)}{k} = g'(d)f'(c) + g'(d)\phi(k) + \psi(\eta)f'(c) + \psi(\eta)\phi(k)$$

As $k \to 0$, we have $\phi(k) \to 0$. Also, since $\eta \to 0$, $\psi(\eta) \to 0$. Thus, all terms on the right-hand side vanish except the first.

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

■

9.4 Derivatives of Elementary Functions

We apply these rules to standard functions.

The Constant and Linear Functions Consider the constant function $f(x) = c$ for all $x \in \mathbb{R}$. Intuitively, the graph is a horizontal line with zero slope. Formally, for any $x \in \mathbb{R}$:

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

Note the order of operations: the expression simplifies to 0 *before* the limit is taken. The function is strictly defined for $h \neq 0$, avoiding the indeterminate form $0/0$.

Similarly, for the affine function $f(x) = mx + b$, the Newton quotient yields the expected slope:

$$\frac{m(x+h) + b - (mx + b)}{h} = \frac{mh}{h} = m$$

Thus, $f'(x) = m$ for all x.

The Quadratic Function Let $f(x) = x^2$. We compute the derivative at an arbitrary point $x \in \mathbb{R}$.

$$\frac{f(x+h) - f(x)}{h} = \frac{(x+h)^2 - x^2}{h} = \frac{x^2 + 2xh + h^2 - x^2}{h} = \frac{2xh + h^2}{h} = 2x + h$$

As $h \to 0$, the term $2x + h$ converges to $2x$. Thus, $(x^2)' = 2x$. This result, combined with the Product Rule, yields the inductive power rule $(x^n)' = nx^{n-1}$ for $n \in \mathbb{N}$.

The Sine Function Let $f(x) = \sin(x)$. Applying the definition:

$$\frac{\sin(x+h) - \sin(x)}{h}$$

Using the addition formula $\sin(x+h) = \sin x \cos h + \cos x \sin h$, we expand the numerator:

$$\frac{\sin x \cos h + \cos x \sin h - \sin x}{h} = \sin x \left(\frac{\cos h - 1}{h}\right) + \cos x \left(\frac{\sin h}{h}\right)$$

To evaluate the limit as $h \to 0$, we must determine the behaviour of two fundamental limits:

$$L_1 = \lim_{h \to 0} \frac{\sin h}{h} \quad \text{and} \quad L_2 = \lim_{h \to 0} \frac{\cos h - 1}{h}$$

By the Fundamental Trigonometric Limit (proposition 4.6.1), $L_1 = 1$.

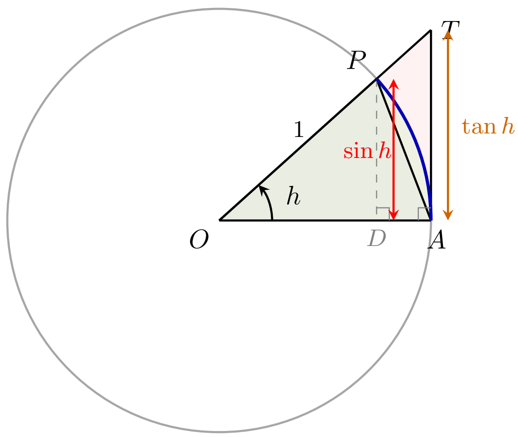


Figure 9.2: Geometric justification for the limit of $(\sin h)/h$. We compare the areas of triangle OAP, sector OAP, and triangle OAT.

For L_2, we manipulate the expression using the Pythagorean identity:

$$\frac{\cos h - 1}{h} = \frac{\cos h - 1}{h} \cdot \frac{\cos h + 1}{\cos h + 1} = \frac{\cos^2 h - 1}{h(\cos h + 1)} = \frac{-\sin^2 h}{h(\cos h + 1)}$$

Separating the limits and using L_1:

$$L_2 = -\left(\lim_{h \to 0} \frac{\sin h}{h}\right) \cdot \left(\lim_{h \to 0} \frac{\sin h}{\cos h + 1}\right) = -(1) \cdot \left(\frac{0}{1+1}\right) = 0$$

Returning to the derivative of the sine function:

$$(\sin x)' = \sin x \cdot (0) + \cos x \cdot (1) = \cos x$$

It follows immediately from the chain rule ($\cos x = \sin(\pi/2 - x)$) that $(\cos x)' = -\sin x$.

The Absolute Value Function Finally, we examine $f(x) = |x|$ at $x = 0$.

$$\frac{f(0+h) - f(0)}{h} = \frac{|h|}{h} = \text{sgn}(h)$$

The limit as $h \to 0^+$ is 1, while the limit as $h \to 0^-$ is -1. Since the one-sided limits disagree, the derivative does not exist at $x = 0$. This confirms our linear approximation intuition: there is no single line that approximates the "corner" at the origin better than all others.

Polynomials and Power Functions

***Proposition* 9.4.1.** *Monomial Derivative via the Binomial Formula.* For $n \in \mathbb{N}$, $\frac{d}{dx}(x^n) = nx^{n-1}$.

Proof. Consider the Newton quotient for $f(x) = x^n$ at x:

$$\frac{(x+h)^n - x^n}{h}$$

By the binomial formula,

$$(x+h)^n = x^n + \binom{n}{1}x^{n-1}h + \binom{n}{2}x^{n-2}h^2 + \cdots + h^n.$$

Substituting and dividing by h gives

$$\frac{(x+h)^n - x^n}{h} = nx^{n-1} + \binom{n}{2}x^{n-2}h + \cdots + h^{n-1}.$$

Letting $h \to 0$ yields nx^{n-1}. ■

Combined with the linearity of the derivative, this implies that for any polynomial $P(x) = \sum a_k x^k$, the derivative is $P'(x) = \sum k a_k x^{k-1}$. This rule extends to rational exponents $x^{p/q}$ via the Chain Rule (implicit differentiation of $y^q = x^p$) and to real exponents x^α for $x > 0$.

Oscillatory Discontinuities We have seen that differentiability implies continuity. However, the derivative function f' need not be continuous. Requiring f' to be continuous is stronger than mere differentiability.

Example 9.4.1. The Differentiable but Discontinuous Derivative. Consider the function tempered by a quadratic factor:

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

1. **At $x = 0$:** We compute the derivative from first principles.

$$f'(0) = \lim_{h \to 0} \frac{h^2 \sin(1/h) - 0}{h} = \lim_{h \to 0} h \sin(1/h) = 0$$

The limit is zero by Theorem 4.6.1. Thus f is differentiable at 0 with $f'(0) = 0$.

2. **For $x \neq 0$:** We use the Chain and Product rules:

$$f'(x) = 2x\sin(1/x) + x^2\cos(1/x) \cdot \left(-\frac{1}{x^2}\right) = 2x\sin(1/x) - \cos(1/x)$$

3. **Continuity of f':** As $x \to 0$, the term $2x\sin(1/x) \to 0$, but $\cos(1/x)$ oscillates between -1 and 1 and has no limit. Thus, $\lim_{x \to 0} f'(x)$ does not exist.

Therefore, $f'(x)$ exists everywhere but is discontinuous at $x = 0$.

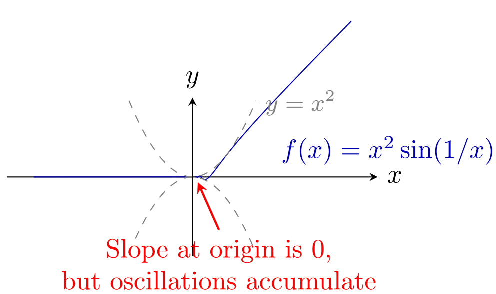


Figure 9.3: The function $f(x) = x^2 \sin(1/x)$ is squeezed between the parabolas $y = \pm x^2$. It is differentiable at 0, but the derivative oscillates wildly near 0.

Derivative of Inverse Functions

The derivatives of trigonometric, exponential, and logarithmic functions are standard. We mention specifically the technique of Implicit Differentiation via the Chain Rule to handle inverse functions, as this relies on the theory developed above.

***Theorem* 9.4.1. Inverse Function Theorem**. Let I, J be intervals in $\mathbb{R}$ and let $f : I \to J$ be a bijective function. Suppose that:

1. f is differentiable at $x_0 \in I$.
2. $f'(x_0) \neq 0$.
3. The inverse function $f^{-1} : J \to I$ is continuous at $y_0 = f(x_0)$.

Then f^{-1} is differentiable at y_0 and its derivative is given by the reciprocal relationship:

$$(f^{-1})'(y_0) = \frac{1}{f'(x_0)} = \frac{1}{f'(f^{-1}(y_0))}$$

Proof. Let $g = f^{-1}$. Since f is bijective, for every $y \in J$ there exists a unique $x \in I$ such that $f(x) = y$, specifically $x = g(y)$. We evaluate the difference quotient for g at y_0. Let $y \in J$ with $y \neq y_0$. Let $x = g(y)$ and $x_0 = g(y_0)$. Since g is injective, $x \neq x_0$.

$$\frac{g(y) - g(y_0)}{y - y_0} = \frac{x - x_0}{f(x) - f(x_0)} = \frac{1}{\frac{f(x)-f(x_0)}{x-x_0}}$$

We now take the limit as $y \to y_0$. Since g is continuous at y_0 (hypothesis 3), we have $\lim_{y \to y_0} g(y) = g(y_0)$, which implies $x \to x_0$.

$$\lim_{y \to y_0} \frac{g(y) - g(y_0)}{y - y_0} = \lim_{x \to x_0} \frac{1}{\frac{f(x)-f(x_0)}{x-x_0}} = \frac{1}{f'(x_0)}$$

The limit exists because $f'(x_0)$ exists and is non-zero. ■

Remark. If f is strictly monotonic and continuous on an interval I, it can be shown that f^{-1} is automatically continuous, satisfying hypothesis (3). Thus, strict monotonicity and non-zero derivatives are sufficient conditions.

Example 9.4.2. Derivative of Arcsine. Consider the function $f : (-\pi/2, \pi/2) \to (-1, 1)$ defined by $f(x) = \sin x$. The function is bijective and differentiable with derivative $f'(x) = \cos x$. Note that $\cos x \neq 0$ on this open interval. Let $g(y) = \arcsin y$ be the inverse function. Applying the theorem at a point $y \in (-1, 1)$:

$$g'(y) = \frac{1}{f'(g(y))} = \frac{1}{\cos(\arcsin y)}$$

Let $\theta = \arcsin y$. Then $\sin\theta = y$. Since $\theta \in (-\pi/2, \pi/2)$, $\cos\theta > 0$. Using the identity $\cos^2\theta + \sin^2\theta = 1$:

$$\cos\theta = \sqrt{1 - \sin^2\theta} = \sqrt{1 - y^2}$$

Substituting this back yields the standard formula:

$$\frac{d}{dy}(\arcsin y) = \frac{1}{\sqrt{1 - y^2}}$$

By symmetry, the derivative of $\arccos y$ is $-1/\sqrt{1 - y^2}$.

Example 9.4.3. Derivative of Arctangent. Consider the bijective function $f : (-\pi/2, \pi/2) \to \mathbb{R}$ defined by $f(x) = \tan x$. Its derivative is $f'(x) = \sec^2 x = 1 + \tan^2 x$, which is never zero. Let $g(y) = \arctan y$. By the Inverse Function Theorem:

$$g'(y) = \frac{1}{f'(g(y))} = \frac{1}{1 + \tan^2(\arctan y)}$$

Since $\tan(\arctan y) = y$, this simplifies directly to:

$$\frac{d}{dy}(\arctan y) = \frac{1}{1 + y^2}$$

9.5 Higher Order Derivatives

Since the derivative f' of a function f is itself a function, we may consider its derivative. This process can be repeated, leading to the concept of higher order derivatives.

Definition 9.5.1. ***Higher Order Derivatives***. Let $f : \text{dom}(f) \subseteq \mathbb{R} \to \mathbb{R}$. We define the derivatives of order n recursively:

$$\begin{aligned} f^{(0)}(x) &= f(x) \\ f^{(1)}(x) &= f'(x) \\ f^{(n+1)}(x) &= \frac{d}{dx}\left[f^{(n)}(x)\right] \end{aligned}$$

for all x where the derivative exists. If f has continuous derivatives up to order k on a set U, we say $f \in C^k(U)$. If f possesses derivatives of all orders on U, we say f is smooth, denoted $f \in C^\infty(U)$.

Physical Interpretation In physics, if $s(t)$ represents the position of a particle at time t, the derivatives correspond to kinematic quantities:

$$\begin{aligned} v(t) &= s'(t) = \dot{s}(t) && \text{(Velocity)} \\ a(t) &= s''(t) = \ddot{s}(t) && \text{(Acceleration)} \\ j(t) &= s'''(t) = \dddot{s}(t) && \text{(Jerk)} \end{aligned}$$

Newton's Second Law relates forces to the second derivative: $F_{\text{net}} = ma(t)$. The third derivative, *jerk*, is the rate of change of acceleration.

Example 9.5.1. Kinematics. Let the position of a particle be $s(t) = 3t^2 + t^3$.

- **Velocity:** $v(t) = \frac{d}{dt}(3t^2 + t^3) = 6t + 3t^2$.
- **Acceleration:** $a(t) = \frac{d}{dt}(6t + 3t^2) = 6 + 6t$.
- **Jerk:** $j(t) = \frac{d}{dt}(6 + 6t) = 6$.

Example 9.5.2. Higher Order Derivative of the Exponential. For $\lambda \in \mathbb{R}$, let $f(x) = e^{\lambda x}$. By the Chain Rule, $f'(x) = \lambda e^{\lambda x}$. By induction,

$$\frac{d^n}{dx^n} e^{\lambda x} = \lambda^n e^{\lambda x}.$$

Reconstructing Functions Knowledge of all derivatives of a function at a single point often allows us to reconstruct the function globally (a concept formally treated in the study of Taylor Series).

Example 9.5.3. Polynomial Reconstruction. Let $f(x) = x^5 + x^4 + x^3 + x^2 + x + 1$. Evaluating the derivatives at $x = 0$:

$$\begin{aligned} f(0) &= 1 \\ f'(0) &= 1 \\ f''(0) &= 2 \\ f'''(0) &= 6 \\ f^{(4)}(0) &= 24 \\ f^{(5)}(0) &= 120 \end{aligned}$$

Notice that $f^{(n)}(0) = n!$ for the coefficient of x^n. We can express $f(x)$ as:

$$f(x) = f(0) + f'(0)x + \frac{f''(0)}{2!}x^2 + \frac{f'''(0)}{3!}x^3 + \cdots + \frac{f^{(5)}(0)}{5!}x^5$$

Differentiability Classes Not all functions possess derivatives of all orders. We examine boundary cases where differentiability fails at a certain order.

Remark. (Point-wise vs. Local Differentiability). Differentiability at a point c is a point-wise property and does not guarantee the existence of the derivative in a neighborhood of c. Because the second derivative $f''(c)$ is the derivative of the function f', its existence requires $f'(x)$ to be defined on an interval containing c. A function can be differentiable at a single point without being differentiable anywhere else in its vicinity, which precludes the existence of higher order derivatives at that point.

Example 9.5.4. Limited Differentiability of $x^{3/2}$. Consider $f(x) = x^{3/2}$ for $x \geq 0$.

$$f'(x) = \frac{3}{2}x^{1/2}, \quad f''(x) = \frac{3}{4}x^{-1/2}$$

At $x = 0$:

- $f'(0) = 0$ (well-defined).
- $f''(0)$ involves $0^{-1/2}$, which is undefined.

Thus, f is differentiable at 0, but not *twice* differentiable there. However, for $x > 0$, f is smooth (C^∞).

Example 9.5.5. The Function $x|x|$. Consider $f(x) = x|x|$. We may write this piecewise:

$$f(x) = \begin{cases} x^2 & x \geq 0 \\ -x^2 & x < 0 \end{cases}$$

First Derivative:

$$f'(x) = \begin{cases} 2x & x > 0 \\ -2x & x < 0 \end{cases}$$

At $x = 0$, the left limit is $\lim_{h \to 0^-} \frac{-h^2}{h} = 0$, and the right limit is $\lim_{h \to 0^+} \frac{h^2}{h} = 0$. Thus $f'(0) = 0$.

Second Derivative:

$$f''(x) = \begin{cases} 2 & x > 0 \\ -2 & x < 0 \end{cases}$$

At $x = 0$, the left limit of the Newton quotient for f' is -2, while the right limit is 2. Since $-2 \neq 2$, $f''(0)$ does not exist.

Conclusion: $f \in C^1(\mathbb{R})$, but $f \notin C^2(\mathbb{R})$.

Remark. (Generalisation). The function $f(x) = x^k|x|$ is k-times differentiable at zero, but fails to be $(k+1)$-times differentiable. This hierarchy illustrates the subtle distinction between merely being differentiable and being *smooth*.

Parametric Differentiation

Often curves are not given as $y = f(x)$ but parametrically as $(x(t), y(t))$. If we view y as a function of x via the parameter t, the Chain Rule yields:

$$\frac{dy}{dt} = \frac{dy}{dx}\frac{dx}{dt} \implies \frac{dy}{dx} = \frac{y'(t)}{x'(t)}$$

provided $x'(t) \neq 0$. This is simply the ratio of the velocities in the vertical and horizontal directions.

Example 9.5.6. Tangent to a Cycloid. Consider the cycloid generated by a circle of radius $r = 1$ rolling on the x-axis, given by the parametric equations:

$$x(t) = t - \sin t, \quad y(t) = 1 - \cos t$$

To find the slope of the tangent line at $t = \pi/2$:

$$\frac{dx}{dt} = 1 - \cos t, \quad \frac{dy}{dt} = \sin t$$

$$\frac{dy}{dx} = \left.\frac{\sin t}{1 - \cos t}\right|_{t=\pi/2} = \frac{1}{1-0} = 1$$

This ratio represents the vertical velocity divided by the horizontal velocity.

Second Derivative of Parametric Curves Caution is required when calculating higher order derivatives for parametric curves. The second derivative $\frac{d^2y}{dx^2}$ is the rate of change of the *slope* with respect to x, not t.

$$\frac{d^2y}{dx^2} = \frac{d}{dx}\left(\frac{dy}{dx}\right) = \frac{\frac{d}{dt}\left(\frac{dy}{dx}\right)}{\frac{dx}{dt}}$$

Example 9.5.7. Concavity of the Astroid. Let $x(t) = \cos^3 t$ and $y(t) = \sin^3 t$ for $t \in (0, \pi/2)$. First derivative:

$$\frac{dx}{dt} = -3\cos^2 t \sin t, \quad \frac{dy}{dt} = 3\sin^2 t \cos t$$

$$\frac{dy}{dx} = \frac{3\sin^2 t \cos t}{-3\cos^2 t \sin t} = -\frac{\sin t}{\cos t} = -\tan t$$

Second derivative:

$$\frac{d^2y}{dx^2} = \frac{\frac{d}{dt}(-\tan t)}{\frac{dx}{dt}} = \frac{-\sec^2 t}{-3\cos^2 t \sin t} = \frac{1}{3\sin t \cos^4 t}$$

For $t \in (0, \pi/2)$, $\frac{d^2y}{dx^2} > 0$, indicating the curve is concave up in the first quadrant.

9.6 Logarithmic Differentiation

Logarithmic differentiation is a powerful technique used to simplify the differentiation of functions that are products, quotients, or powers of multiple functions. The primary strategy is to transform operations of multiplication and exponentiation into addition and multiplication by constants, respectively, using the properties of the natural logarithm.

Definition 9.6.1. ***Procedure***. To differentiate a complicated function $y = f(x)$:

1. Take the natural logarithm of both sides: $\ln(y) = \ln(f(x))$.
2. Use logarithmic identities to expand the right-hand side into a sum or difference of simpler log terms.
3. Differentiate both sides implicitly with respect to x. Recall the key identity:

$$\frac{d}{dx}(\ln u) = \frac{1}{u} \cdot \frac{du}{dx}$$

4. Solve for $\frac{dy}{dx}$ by multiplying both sides by y (and substituting the original expression for y back in).

Example 9.6.1. Complex Product and Quotient. Find $\frac{dy}{dx}$ for $y = (2-x)^{-1}(x+32)^{1/4}(x^2-3)^4$. Taking the natural logarithm of both sides:

$$\ln(y) = \ln\left[(2-x)^{-1}(x+32)^{1/4}(x^2-3)^4\right]$$

Using log properties:

$$\ln(y) = -\ln(2-x) + \frac{1}{4}\ln(x+32) + 4\ln(x^2-3)$$

Differentiating with respect to x:

$$\frac{1}{y}\frac{dy}{dx} = -\frac{1}{2-x}(-1) + \frac{1}{4(x+32)} + 4\cdot\frac{2x}{x^2-3}$$

$$\frac{1}{y}\frac{dy}{dx} = \frac{1}{2-x} + \frac{1}{4(x+32)} + \frac{8x}{x^2-3}$$

Multiplying by y:

$$\frac{dy}{dx} = (2-x)^{-1}(x+32)^{1/4}(x^2-3)^4\left[\frac{1}{2-x} + \frac{1}{4(x+32)} + \frac{8x}{x^2-3}\right]$$

Example 9.6.2. Product with an Exponent. Find the derivative of $y = xe^{x^2+9}$. Taking logarithms:

$$\ln(y) = \ln(x) + \ln(e^{x^2+9}) = \ln(x) + x^2 + 9$$

Differentiating:

$$\frac{1}{y}\frac{dy}{dx} = \frac{1}{x} + 2x$$

Solving for $\frac{dy}{dx}$:

$$\frac{dy}{dx} = y\left(\frac{1}{x} + 2x\right) = xe^{x^2+9}\left(\frac{1}{x} + 2x\right) = e^{x^2+9}(1+2x^2)$$

Example 9.6.3. Variable Exponent. Find the derivative of $y = (x^2+1)^{\sin x}$. Unlike power functions (x^n) or exponential functions (b^x), this function has a variable in both the base and the exponent. We must use logarithmic differentiation. Taking natural logarithms of both sides:

$$\ln(y) = \sin x \cdot \ln(x^2+1)$$

Differentiating:

$$\frac{1}{y}\frac{dy}{dx} = (\cos x)\cdot\ln(x^2+1) + (\sin x)\cdot\frac{1}{x^2+1}(2x)$$

$$\frac{1}{y}\frac{dy}{dx} = \cos x\ln(x^2+1) + \frac{2x\sin x}{x^2+1}$$

Solving for $\frac{dy}{dx}$:

$$\frac{dy}{dx} = (x^2+1)^{\sin x}\left[\cos x\ln(x^2+1) + \frac{2x\sin x}{x^2+1}\right]$$

Example 9.6.4. Derivative of x^x. Let $f(x) = x^x$ for $x > 0$. Set $g(x) = \ln f(x) = x\ln x$. Then $g'(x) = \ln x + 1$. Since $f(x) = e^{g(x)}$, we have $f'(x) = e^{g(x)}g'(x) = x^x(\ln x + 1)$.

Example 9.6.5. Multiple Factors. Let a, b, c be constants. Differentiate

$$y = \left(\frac{1}{x-a}\right)\left(\frac{1}{x-b}\right)^2\left(\frac{1}{x-c}\right)^3$$

Taking logs:

$$\ln y = -\ln(x-a) - 2\ln(x-b) - 3\ln(x-c)$$

Differentiating:

$$\frac{1}{y}y' = -\frac{1}{x-a} - \frac{2}{x-b} - \frac{3}{x-c}$$

$$y' = y\left(-\frac{1}{x-a} - \frac{2}{x-b} - \frac{3}{x-c}\right)$$

Example 9.6.6. Polynomial Products. Differentiate $y = (x^2+1)(x-3)^2(x^3+x)^3(x-1)^4$.

$$\ln y = \ln(x^2+1) + 2\ln(x-3) + 3\ln(x^3+x) + 4\ln(x-1)$$

$$\frac{1}{y}y' = \frac{2x}{x^2+1} + \frac{2}{x-3} + \frac{3(3x^2+1)}{x^3+x} + \frac{4}{x-1}$$

$$y' = y\left[\frac{2x}{x^2+1} + \frac{2}{x-3} + \frac{3(3x^2+1)}{x^3+x} + \frac{4}{x-1}\right]$$

Example 9.6.7. Pre-simplification via Logarithms. Sometimes the function is already a logarithm. In this case, we use log rules to simplify *before* differentiating, rather than taking the log again. Find $\frac{dy}{dx}$ for $y = \ln\left(\frac{\sin x\sqrt{x}}{x^2+3x-2}\right)$. Simplify first:

$$y = \ln(\sin x) + \frac{1}{2}\ln x - \ln(x^2+3x-2)$$

Differentiate directly:

$$\frac{dy}{dx} = \frac{\cos x}{\sin x} + \frac{1}{2x} - \frac{2x+3}{x^2+3x-2} = \cot x + \frac{1}{2x} - \frac{2x+3}{x^2+3x-2}$$

Remark. (Contrast). Consider $y = \ln((x+1)^{30}+2)$. Here, the argument is a sum, not a product. We cannot simplify $\ln(A+B)$. We must use the Chain Rule directly:

$$\frac{dy}{dx} = \frac{1}{(x+1)^{30}+2} \cdot 30(x+1)^{29}$$

***Proposition* 9.6.1.** *Power Function Derivative via Fundamental Limits.* For $f(x) = x^\alpha$ where $\alpha \in \mathbb{R}$ and $x > 0$, $f'(x) = \alpha x^{\alpha-1}$.

Proof. Let $h \neq 0$. Then

$$\frac{(x+h)^\alpha - x^\alpha}{h} = x^{\alpha-1}\frac{(1+h/x)^\alpha - 1}{h/x}.$$

Set $t = h/x$. As $h \to 0$, $t \to 0$. By theorem 8.3.5, the fraction tends to α, hence $f'(x) = \alpha x^{\alpha-1}$. ∎

***Theorem* 9.6.1. General Power Rule**. Let $n \in \mathbb{R}$ and $x > 0$. Then

$$\frac{d}{dx}(x^n) = nx^{n-1}$$

Proof. Let $y = x^n$. For $x > 0$, we take the natural logarithm:

$$\ln y = n\ln x$$

Differentiating with respect to x:

$$\frac{1}{y}\frac{dy}{dx} = n \cdot \frac{1}{x}$$

Solving for y':

$$\frac{dy}{dx} = y \cdot \frac{n}{x} = x^n \cdot \frac{n}{x} = nx^{n-1}$$

∎

Remark. If $n < 0$ is an integer, the domain includes negative numbers. The proof above requires $x > 0$ for $\ln x$ to be defined. However, the result holds for $x < 0$ as well, which can be verified by writing $x^n = (x^{-1})^{-n}$ or by considering $\ln|y|$.

9.7 Applications of the Derivative: Extrema

The derivative provides a local linear model; its sign governs local monotonicity and guides the search for maxima and minima.

Global and Local Extrema

We begin by formalising the notion of "peaks" and "valleys". Let $f : A \to \mathbb{R}$ be a function.

Definition 9.7.1. ***Global Maximum and Minimum***. Let $A \subseteq \mathbb{R}$ and $f : A \to \mathbb{R}$.

1. A point $c \in A$ is a global maximum point for f on A if $f(c) \geq f(x)$ for all $x \in A$. The value $f(c)$ is the maximum value.
2. A point $c \in A$ is a global minimum point for f on A if $f(c) \leq f(x)$ for all $x \in A$. The value $f(c)$ is the minimum value.

Collectively, these are referred to as global extrema.

As discussed in the chapter on Compactness (see theorem 6.3.2), global extrema are not guaranteed to exist unless the domain A is compact and f is continuous. Even when they exist, identifying them by checking every point in A is impossible. To render the problem tractable, we localise the definition.

Definition 9.7.2. ***Local Extrema***. Let $f : A \to \mathbb{R}$ and let $c \in A$.

1. We say c is a local maximum point if there exists a $\delta > 0$ such that $f(c) \geq f(x)$ for all $x \in A \cap (c-\delta, c+\delta)$.
2. We say c is a local minimum point if there exists a $\delta > 0$ such that $f(c) \leq f(x)$ for all $x \in A \cap (c-\delta, c+\delta)$.

Note. Every global extremum is automatically a local extremum (take any δ), but the converse is false. The function $f(x) = x(x-1)(x+1)$ has local hills and valleys, but on $\mathbb{R}$ it is unbounded.

Fermat's Stationary Point Theorem

The fundamental link between extrema and differentiation is due to Fermat. It formalises the intuition that at the peak of a smooth mountain, the ground must be flat. If the tangent line had a non-zero slope, one could move slightly uphill or downhill, contradicting the assumption of a maximum.

***Theorem* 9.7.1. Fermat's Theorem**. Let $f : (a, b) \to \mathbb{R}$ be a function and let $c \in (a, b)$. If c is a local extremum of f and f is differentiable at c, then $f'(c) = 0$.

Proof. Suppose c is a local maximum. By definition, there exists $\delta > 0$ such that for all h with $|h| < \delta$ (and $c + h \in (a, b)$), we have $f(c+h) \leq f(c)$, implying $f(c+h) - f(c) \leq 0$. We examine the sign of the Newton quotient for small h.

1. **Right-hand limit ($h > 0$):** Since $h > 0$ and the numerator is non-positive:
$$\frac{f(c+h) - f(c)}{h} \leq 0$$
Taking the limit as $h \to 0^+$, we have $f'(c) \leq 0$.
2. **Left-hand limit ($h < 0$):** Since $h < 0$, dividing the non-positive numerator by the negative denominator yields a non-negative quotient:
$$\frac{f(c+h) - f(c)}{h} \geq 0$$
Taking the limit as $h \to 0^-$, we have $f'(c) \geq 0$.

Since f is differentiable at c, the left and right limits must coincide. Thus $0 \leq f'(c) \leq 0$, which implies $f'(c) = 0$. The proof for a local minimum is identical, with the inequalities reversed. ■

Remark. (Interior Points Only.). Fermat's Theorem explicitly requires the domain to be an open interval (or for c to be an interior point). If f is defined on a closed interval $[a, b]$, the maximum may occur at an endpoint. For $f(x) = x$ on $[0, 1]$, the maximum is at 1, but $f'(1) = 1 \neq 0$. At an endpoint, we can only form one-sided limits, so the "trapping" argument $0 \leq f'(c) \leq 0$ fails.

Critical Points and the Classification Strategy

Fermat's Theorem provides a necessary condition, but not a sufficient one.

1. **Stationary Points:** Points where $f'(c) = 0$.
2. **Singular Points:** Points where $f'(c)$ does not exist (e.g., $f(x) = |x|$ has a minimum at 0, but is not differentiable there).

Furthermore, $f'(c) = 0$ does not guarantee an extremum. Consider $f(x) = x^3$ at $c = 0$. Here $f'(0) = 0$, but 0 is neither a maximum nor a minimum; it is a point of inflection.

Definition 9.7.3. ***Critical Point***. Let f be defined on a domain A. An interior point $c \in A$ is called a critical point if either $f'(c) = 0$ or $f'(c)$ does not exist.

This leads to a robust algorithm for finding global extrema on closed intervals, grounded in the Extreme Value Theorem.

***Theorem* 9.7.2. Location of Extrema**. Let $f : [a, b] \to \mathbb{R}$ be a continuous function. The global maximum and global minimum of f on $[a, b]$ must occur at one of the following locations:

1. The endpoints a and b.
2. A point $c \in (a, b)$ where $f'(c) = 0$ (Stationary point).
3. A point $c \in (a, b)$ where f is not differentiable (Singular point).

Proof. By the Extreme Value Theorem, extrema exist. Let x_{max} be the location of the maximum. If x_{max} is an endpoint, we are in case (1). If $x_{max} \in (a, b)$, then if f is differentiable at x_{max}, by Fermat's Theorem $f'(x_{max}) = 0$ (case 2). If f is not differentiable there, we are in case (3). ■

9.8 Optimisation Examples

The theorem reduces the search to finitely many candidates.

Example 9.8.1. A Piecewise Function. Consider $f : [-1, 1] \to \mathbb{R}$ defined by:

$$f(x) = \begin{cases} |x| & x \neq 0, x \neq \pm 1 \\ 1/2 & x = 0 \\ 1/3 & x = \pm 1 \end{cases}$$

- **Endpoints:** $f(\pm 1) = 1/3$.
- **Differentiation:** On $(-1, 0) \cup (0, 1)$, $f(x) = |x|$. $f'(x) = \text{sgn}(x)$, which is never 0. There are no stationary points.
- **Singularities:** At $x = 0$, the derivative does not exist. $f(0) = 1/2$.

Comparing candidates $\{1/3, 1/2\}$, one might erroneously conclude the maximum is $1/2$. However, $f(0.9) = 0.9 > 1/2$. The global supremum is 1, but it is never achieved. The failure lies in the premise: f is not continuous at 0 or ± 1, so the Extreme Value Theorem does not apply.

Example 9.8.2. Rational Function on a Compact Set. Find the extrema of $f(x) = x^4 - 2x^2$ on $[-2, 2]$. Since f is a polynomial, it is continuous and differentiable everywhere.

1. **Stationary Points:** $f'(x) = 4x^3 - 4x = 4x(x-1)(x+1)$. The critical points are $x = 0, 1, -1$.
2. **Evaluate Candidates:**
   - Critical points: $f(0) = 0$, $f(1) = -1$, $f(-1) = -1$.
   - Endpoints: $f(2) = 16 - 8 = 8$, $f(-2) = 8$.
3. **Conclusion:** Global Max is 8 (at $x = \pm 2$). Global Min is -1 (at $x = \pm 1$).

Example 9.8.3. Optimisation of $f(x) = x(a-x)^2$. Let $a > 0$ and consider $f(x) = x(a-x)^2$ on $[0, a]$.

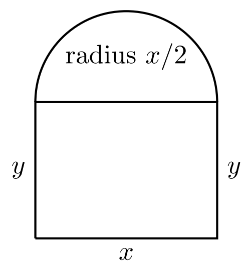
$$f'(x) = (a-x)^2 - 2x(a-x) = 3x^2 - 4ax + a^2.$$

The critical points satisfy $3x^2 - 4ax + a^2 = 0$, hence $x = a$ and $x = a/3$. At $x = a/3$,

$$f''(a/3) = 6(a/3) - 4a = -2a < 0,$$

so the maximum occurs at $x = a/3$ with value $f(a/3) = \frac{4a^3}{27}$.

Example 9.8.4. The Roman Window. A classic optimisation problem involves maximising the area of a "Roman window" (a rectangle surmounted by a semicircle) given a fixed perimeter L.



Let x be the width of the base and y be the height of the rectangular vertical sides. The perimeter consists of the base, two vertical sides, and the semi-circular arc:

$$P = x + 2y + \frac{1}{2}(\pi x) = x\left(1 + \frac{\pi}{2}\right) + 2y = L$$

We express y in terms of x:

$$2y = L - x\left(1 + \frac{\pi}{2}\right) \implies y = \frac{L}{2} - \frac{x}{2}\left(1 + \frac{\pi}{2}\right)$$

The area A is the sum of the rectangle and the semicircle:

$$A(x) = xy + \frac{1}{2}\pi\left(\frac{x}{2}\right)^2 = x\left[\frac{L}{2} - \frac{x}{2}\left(1 + \frac{\pi}{2}\right)\right] + \frac{\pi x^2}{8}$$

Simplifying the function $A(x)$:

$$A(x) = \frac{Lx}{2} - \frac{x^2}{2} - \frac{\pi x^2}{4} + \frac{\pi x^2}{8} = \frac{Lx}{2} - \frac{x^2}{2} - \frac{\pi x^2}{8} = \frac{Lx}{2} - \frac{x^2}{2}\left(1 + \frac{\pi}{4}\right)$$

Domain: The geometric constraints are $x \geq 0$ and $y \geq 0$. $y \geq 0 \implies L \geq x(1 + \pi/2) \implies x \leq \frac{L}{1+\pi/2}$.
Thus we maximise $A(x)$ on the closed interval $[0, \frac{L}{1+\pi/2}]$.

Differentiation:

$$A'(x) = \frac{L}{2} - x\left(1 + \frac{\pi}{4}\right)$$

Setting $A'(x) = 0$:

$$x\left(1 + \frac{\pi}{4}\right) = \frac{L}{2} \implies x\left(\frac{4+\pi}{4}\right) = \frac{L}{2} \implies x = \frac{2L}{4+\pi}$$

We verify this critical point lies in the domain. Since $4 + \pi > 2 + \pi$, $\frac{2L}{4+\pi} < \frac{L}{1+\pi/2}$, so it is a valid interior point. **Comparison:**

- Endpoints: $A(0) = 0$. At the upper bound ($y = 0$), the area is positive but smaller than at the critical point.
- Critical Point: Since $A''(x) = -(1+\pi/4) < 0$, the function is concave down, ensuring a maximum.

Thus, the dimensions maximising area are $x = \frac{2L}{4+\pi}$.

Interpretation of the Derivative's Sign

While Fermat's Theorem identifies candidates, it does not determine the nature of the extremum (max, min, or inflection). To do this, we require the Mean Value Theorem (MVT) to link the sign of f' to monotonicity. However, we can state the expectation that motivates the next chapter:

- If $f'(x) > 0$ on an interval, f is increasing.
- If $f'(x) < 0$ on an interval, f is decreasing.

If f' changes from positive to negative at c, then c is a local maximum. This is the First Derivative Test. The formal proof of these assertions relies on the theorems of Rolle and Lagrange, to which we turn to next.

9.9 The Mean Value Theorems

We have established that local extrema of differentiable functions occur at critical points where the derivative vanishes (Fermat's Theorem). To link local derivatives to global behaviour on an interval, we require the Mean Value Theorem and its generalisations.

Rolle's Theorem

If a smooth curve begins and ends at the same height, it must have a horizontal tangent in between.

***Theorem* 9.9.1. Rolle's Theorem**. Let $f : [a, b] \to \mathbb{R}$ be continuous on $[a, b]$ and differentiable on (a, b). If $f(a) = f(b)$, then there exists a point $\xi \in (a, b)$ such that $f'(\xi) = 0$.

Proof. Since f is continuous on the compact interval $[a, b]$, by the Extreme Value Theorem, f attains a global maximum M and a global minimum m on $[a, b]$.

1. **Case 1:** $M = m$. In this case, $f(x)$ is constant on $[a, b]$. Consequently, $f'(x) = 0$ for all $x \in (a, b)$, and the assertion holds trivially.
2. **Case 2:** $M \neq m$. Since $f(a) = f(b)$, at least one of the extrema must occur at an interior point $\xi \in (a, b)$. Suppose $f(\xi) = M$ (the argument for the minimum is identical). Since ξ is an interior local extremum and f is differentiable at ξ, Fermat's Theorem implies $f'(\xi) = 0$.

∎

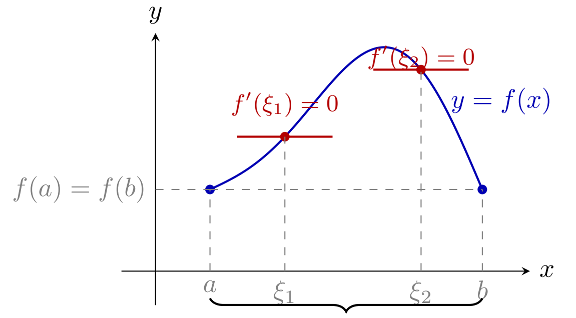


Figure 9.4: Rolle's Theorem ensures at least one point with a horizontal tangent.

By rotating the coordinate system, we generalise Rolle's Theorem to functions where $f(a) \neq f(b)$. The "horizontal" tangent becomes a tangent parallel to the secant line connecting the endpoints.

***Theorem* 9.9.2. Mean Value Theorem (MVT).** Let $f : [a, b] \to \mathbb{R}$ be continuous on $[a, b]$ and differentiable on (a, b). Then there exists a point $\xi \in (a, b)$ such that:

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

Proof. We define the secant line passing through $(a, f(a))$ and $(b, f(b))$ by the linear function:

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

Consider the auxiliary function $g(x) = f(x) - L(x)$, which represents the vertical distance between the graph of f and the secant line. The function g is continuous on $[a, b]$ and differentiable on (a, b) as it is a linear combination of such functions. Furthermore:

$$g(a) = f(a) - L(a) = f(a) - f(a) = 0$$
$$g(b) = f(b) - L(b) = f(b) - f(b) = 0$$

Since $g(a) = g(b)$, Rolle's Theorem applies. There exists $\xi \in (a, b)$ such that $g'(\xi) = 0$. Differentiating g:

$$g'(\xi) = f'(\xi) - L'(\xi) = f'(\xi) - \frac{f(b) - f(a)}{b - a}$$

Setting this to zero yields the result. ■

Consequences of the Mean Value Theorem

The MVT allows us to bound the growth of a function using bounds on its derivative.

***Proposition* 9.9.1.** *Characterisation of Constant Functions.* Let f be differentiable on an interval I. Then $f'(x) = 0$ for all $x \in I$ if and only if f is constant on I.

Proof. The forward implication is trivial. Conversely, suppose $f'(x) = 0$ for all $x \in I$. Pick any two distinct points $x_1, x_2 \in I$ with $x_1 < x_2$. By the MVT applied to $[x_1, x_2]$, there exists ξ such that:

$$f(x_2) - f(x_1) = f'(\xi)(x_2 - x_1) = 0 \cdot (x_2 - x_1) = 0$$

Thus $f(x_2) = f(x_1)$. Since x_1, x_2 were arbitrary, f is constant. ■

***Proposition* 9.9.2.** *The Monotonicity Criterion.* Let f be differentiable on an interval I.

1. If $f'(x) \geq 0$ for all $x \in I$, then f is monotonically increasing.
2. If $f'(x) > 0$ for all $x \in I$, then f is strictly increasing.

(Analogous results hold for decreasing functions with negative derivatives).

Proof. Let $x_1, x_2 \in I$ with $x_1 < x_2$. By the MVT, $f(x_2) - f(x_1) = f'(\xi)(x_2 - x_1)$ for some $\xi \in (x_1, x_2)$. Since $x_2 - x_1 > 0$, the sign of $f(x_2) - f(x_1)$ is determined entirely by the sign of $f'(\xi)$. If $f'(\xi) \geq 0$, then $f(x_2) \geq f(x_1)$. If $f'(\xi) > 0$, then $f(x_2) > f(x_1)$. ■

***Corollary* 9.9.1.** *Injectivity.* If $f'(x) \neq 0$ for all x in an interval I, then f is injective on I.

Applications to Inequalities The MVT provides a rigorous method for proving inequalities. By treating an inequality of the form $A(x) \geq B(x)$ as a monotonicity problem for $f(x) = A(x) - B(x)$, we can establish bounds for transcendental functions.

Example 9.9.1. Exponential Bounds. Show that $e^x \geq 1 + x$ for all $x \in \mathbb{R}$. Let $f(x) = e^x - (1 + x)$. Then $f'(x) = e^x - 1$.

- For $x > 0$, $e^x > 1$, so $f'(x) > 0$. Thus f is strictly increasing on $[0, \infty)$. Since $f(0) = 0$, $f(x) \geq 0$ for $x > 0$.
- For $x < 0$, $e^x < 1$, so $f'(x) < 0$. Thus f is strictly decreasing on $(-\infty, 0]$. Since $f(0) = 0$, $f(x) > f(0) = 0$ for $x < 0$.

Example 9.9.2. Young's Inequality. Let $p, q \in (1, \infty)$ such that $\frac{1}{p} + \frac{1}{q} = 1$. For any $u, v > 0$:

$$uv \leq \frac{u^p}{p} + \frac{v^q}{q}$$

Proof. Fix $v > 0$. We define the function $g(u) = u^p/p + v^q/q - uv$ for $u > 0$. Differentiating with respect to u:

$$g'(u) = u^{p-1} - v$$

The stationary point occurs when $u^{p-1} = v$, or $u_0 = v^{1/(p-1)}$. Note that since $1/p + 1/q = 1$, we have $(p - 1) = p/q$, so $1/(p - 1) = q/p$. Thus the critical point is $u_0 = v^{q/p}$. Checking the second derivative: $g''(u) = (p - 1)u^{p-2}$. Since $p > 1$, $g''(u) > 0$ for all $u > 0$. This implies that the derivative $g'(u)$ is strictly increasing. Since $g'(u_0) = 0$, we have $g'(u) < 0$ for $u < u_0$ and $g'(u) > 0$ for $u > u_0$. By the Monotonicity Criterion, g decreases on $(0, u_0]$ and increases on $[u_0, \infty)$. Thus u_0 is the unique global minimum. The minimum value is:

$$\begin{aligned} g(u_0) &= \frac{(v^{q/p})^p}{p} + \frac{v^q}{q} - v^{q/p}v \\ &= \frac{v^q}{p} + \frac{v^q}{q} - v^{q/p+1} \end{aligned}$$

Using $q/p + 1 = q(1/p + 1/q) = q(1) = q$, we have:

$$g(u_0) = v^q \left(\frac{1}{p} + \frac{1}{q}\right) - v^q = v^q(1) - v^q = 0$$

Since the global minimum is 0, $g(u) \geq 0$ for all u, proving the inequality. ■

The First Derivative Test

With the Monotonicity Criterion established via the Mean Value Theorem, we can now formally justify the First Derivative Test. This theorem allows us to classify the critical points identified by Fermat's Theorem.

***Theorem* 9.9.3. First Derivative Test**. Let f be continuous on an interval containing a critical point c. Suppose f is differentiable on $(c-\delta, c) \cup (c, c+\delta)$ for some $\delta > 0$.

1. If $f'(x) > 0$ on $(c-\delta, c)$ and $f'(x) < 0$ on $(c, c+\delta)$, then f has a local maximum at c.
2. If $f'(x) < 0$ on $(c-\delta, c)$ and $f'(x) > 0$ on $(c, c+\delta)$, then f has a local minimum at c.
3. If $f'(x)$ has the same sign on both sides of c, then f has no extremum at c.

Proof. We prove the first case (local maximum).

- On the interval $(c-\delta, c)$, we have $f'(x) > 0$. By the Monotonicity Criterion, f is strictly increasing on this interval. Since f is continuous at c, for any $x \in (c-\delta, c)$, we must have $f(x) < f(c)$.
- On the interval $(c, c+\delta)$, we have $f'(x) < 0$. By the Monotonicity Criterion, f is strictly decreasing on this interval. Since f is continuous at c, for any $x \in (c, c+\delta)$, we must have $f(x) < f(c)$.

Combining these results, $f(x) < f(c)$ for all $x \in (c-\delta, c+\delta)$ with $x \neq c$. Thus, by definition, c is a local maximum. The proof for the local minimum is analogous. For the third case, suppose $f'(x) > 0$ on both sides. Then f is increasing on $(c-\delta, c)$ (so $f(x) < f(c)$ for $x < c$) and increasing on $(c, c+\delta)$ (so $f(x) > f(c)$ for $x > c$). Since $f(c)$ is not the largest or smallest value in the neighbourhood, it is not an extremum. ■

Remark. This test is often more robust than the Second Derivative Test because it applies even when $f''(c) = 0$ or $f''(c)$ does not exist (provided f is continuous), whereas the Second Derivative Test is inconclusive in those cases.

Cauchy's Mean Value Theorem

A generalisation involving two functions allows us to eliminate the explicit dependence on the interval width $b-a$. This result is the cornerstone of L'Hôpital's Rule.

***Theorem* 9.9.4. Cauchy's Finite Increment Theorem**. Let $f, g : [a, b] \to \mathbb{R}$ be continuous on $[a, b]$ and differentiable on (a, b). Then there exists $\xi \in (a, b)$ such that:

$$f'(\xi)\,[g(b) - g(a)] = g'(\xi)\,[f(b) - f(a)]$$

If $g'(x) \neq 0$ for all $x \in (a, b)$, this can be written as:

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

Proof. We construct a linear combination that vanishes at the endpoints. Let:

$$H(x) = [f(b) - f(a)]\,g(x) - [g(b) - g(a)]\,f(x)$$

Observe that:

$$H(a) = f(b)g(a) - f(a)g(a) - g(b)f(a) + g(a)f(a) = f(b)g(a) - g(b)f(a)$$

$$H(b) = f(b)g(b) - f(a)g(b) - g(b)f(b) + g(a)f(b) = -f(a)g(b) + g(a)f(b)$$

Thus $H(a) = H(b)$. By Rolle's Theorem, there exists $\xi \in (a, b)$ such that $H'(\xi) = 0$. Differentiating H:

$$H'(\xi) = [f(b) - f(a)]g'(\xi) - [g(b) - g(a)]f'(\xi) = 0$$

Rearranging terms yields the result. ■

Although f' need not be continuous (as in $x^2 \sin(1/x)$), derivatives cannot exhibit jump discontinuities. They share the Intermediate Value Property with continuous functions.

***Theorem* 9.9.5. Darboux's Theorem**. Let $f : [a, b] \to \mathbb{R}$ be a differentiable function. If k is a value strictly between $f'(a)$ and $f'(b)$, then there exists a point $c \in (a, b)$ such that $f'(c) = k$.

Proof. Without loss of generality, assume $f'(a) < k < f'(b)$. We consider the auxiliary function $g(x) = f(x) - kx$. Differentiating, $g'(x) = f'(x) - k$. Evaluating at the endpoints:

$$g'(a) = f'(a) - k < 0$$
$$g'(b) = f'(b) - k > 0$$

Since f is differentiable, it is continuous on $[a, b]$, and so is g. By the Extreme Value Theorem, g attains a global minimum on $[a, b]$. We argue that this minimum cannot occur at the endpoints:

- Since $g'(a) < 0$, $\lim_{h \to 0^+} \frac{g(a+h)-g(a)}{h} < 0$. For sufficiently small $h > 0$, $g(a + h) < g(a)$. Thus a is not the minimum.
- Since $g'(b) > 0$, $\lim_{h \to 0^-} \frac{g(b+h)-g(b)}{h} > 0$. The ratio is positive, but h is negative, implying the numerator is negative: $g(b + h) < g(b)$. Thus b is not the minimum.

Therefore, the minimum must occur at some interior point $c \in (a, b)$. By Fermat's Theorem, stationary points of the minimum satisfy $g'(c) = 0$. Consequently, $f'(c) - k = 0 \implies f'(c) = k$. ■

***Corollary* 9.9.2.** *Continuity of Derivatives.* If f is differentiable on an interval I, then f' cannot have any simple jump discontinuities. If f' is discontinuous at c, it must be an essential discontinuity (an oscillatory singularity).

This theorem forces f' to vary without skipping values.

9.10 Summary and More Examples

We conclude this chapter with a summary of the derivatives of standard functions. These can all be derived from first principles (as we did for constants, powers, and sine) or via the chain and inverse function rules.

Function $f(x)$	**Derivative** $f'(x)$	**Function** $f(x)$	**Derivative** $f'(x)$
c (constant)	0	$\sin x$	$\cos x$
x^n	nx^{n-1}	$\cos x$	$-\sin x$
e^x	e^x	$\tan x$	$\sec^2 x$
a^x $(a > 0)$	$a^x \ln a$	$\cot x$	$-\csc^2 x$
$\ln \lvert x \rvert$	$1/x$	$\sec x$	$\sec x \tan x$
$\log_a \lvert x \rvert$	$\frac{1}{x \ln a}$	$\csc x$	$-\csc x \cot x$
$\sin^{-1} x$	$\frac{1}{\sqrt{1-x^2}}$	$\sinh x$	$\cosh x$
$\cos^{-1} x$	$-\frac{1}{\sqrt{1-x^2}}$	$\cosh x$	$\sinh x$
$\tan^{-1} x$	$\frac{1}{1+x^2}$	$\tanh x$	$\text{sech}^2 x$

Example 9.10.1. Chain Rule in Action. Consider $f(x) = \tan(5 - \sin(x^2))$. Let $u = x^2$, $v = \sin u$, $w = 5 - v$, and $y = \tan w$.

$$\frac{dy}{dx} = \frac{dy}{dw} \cdot \frac{dw}{dv} \cdot \frac{dv}{du} \cdot \frac{du}{dx}$$
$$\frac{dy}{dx} = (\sec^2 w) \cdot (-1) \cdot (\cos u) \cdot (2x)$$

Substituting back:

$$f'(x) = -2x \cos(x^2) \sec^2(5 - \sin(x^2))$$

Example 9.10.2. Implicit Differentiation. Find y' for the curve $x^2 + \sin(xy) + y^2 = 0$. Differentiating with respect to x:

$$2x + \cos(xy) \cdot (y + xy') + 2yy' = 0$$

Grouping y' terms:

$$y'(x \cos(xy) + 2y) = -(2x + y \cos(xy))$$

$$y' = -\frac{2x + y\cos(xy)}{x\cos(xy) + 2y}$$

Example 9.10.3. Tangent and Normal to the Folium of Descartes. Consider the curve given by $x^3 + y^3 - 9xy = 0$. We wish to find the equation of the tangent and the normal lines at the point $(2, 4)$. First, verify the point lies on the curve: $2^3 + 4^3 - 9(2)(4) = 8 + 64 - 72 = 0$. Differentiating implicitly with respect to x:

$$3x^2 + 3y^2y' - 9(y + xy') = 0$$

Dividing by 3 and grouping y' terms:

$$x^2 - 3y + y'(y^2 - 3x) = 0 \implies y' = \frac{3y - x^2}{y^2 - 3x}$$

Evaluating at $(2, 4)$:

$$m = y'(2) = \frac{3(4) - 2^2}{4^2 - 3(2)} = \frac{12 - 4}{16 - 6} = \frac{8}{10} = \frac{4}{5}$$

Tangent Line: $y - 4 = \frac{4}{5}(x - 2) \implies 4x - 5y + 12 = 0$.

Normal Line: The slope is $-1/m = -5/4$.

$$y - 4 = -\frac{5}{4}(x - 2) \implies 5x + 4y - 26 = 0$$

Example 9.10.4. Parametric Tangents. Find the tangent to the curve $x = t\cos t$, $y = t\sin t$ at $t = \pi/2$.

$$\frac{dx}{dt} = \cos t - t\sin t \implies x'(\pi/2) = 0 - \pi/2 = -\pi/2$$

$$\frac{dy}{dt} = \sin t + t\cos t \implies y'(\pi/2) = 1 + 0 = 1$$

Slope $m = \frac{dy}{dx} = \frac{1}{-\pi/2} = -\frac{2}{\pi}$. At $t = \pi/2$, $x = 0$ and $y = \pi/2$. Equation: $y - \pi/2 = -\frac{2}{\pi}(x - 0) \implies y = -\frac{2}{\pi}x + \frac{\pi}{2}$.

Example 9.10.5. Higher Order Derivatives. Let $y = x^3/4 - 5x + 3$.

$$y' = \frac{3}{4}x^2 - 5$$

$$y'' = \frac{3}{2}x$$

$$y''' = \frac{3}{2}$$

$$y^{(4)} = 0$$

For polynomials, the $(n + 1)$-th derivative of an n-th degree polynomial is identically zero.

9.11 Exercises

1. **Tangent and Normal Lines.** Find the equation of the tangent and normal lines to the given curves at the specified points.
   (a) The curve $y = x^3 - 4$ at each of its points of intersection with the x-axis and the y-axis.
   (b) The curve $2y = (x + 2)^2$ at each point on it where $y = 2$.
   (c) The curve $y = 4(x - 1)^3 + 3x$ at the origin $(0, 0)$. Prove further that this tangent intersects the curve again at a second distinct point.

2. **High-Order Derivatives.** Calculate the n-th derivative $y^{(n)}$ for the following functions. Look for a pattern.

(a) $y = \frac{x-1}{x+1}$
(b) $y = \frac{1}{x^2-4}$

Remark. Use partial fractions.

(c) $y = \sin(3x)\sin(5x)$

Remark. Use the product-to-sum trigonometric identity.

(d) $y = \frac{a}{a^2-x^2}$

3. Leibniz Rule Application. Let $y = x^2 \sin x$. Using the Leibniz rule for the n-th derivative of a product, show that for $n \geq 2$:

$$y^{(n)} = x^2 \sin(x + \frac{n\pi}{2}) - 2nx\cos(x + \frac{n\pi}{2}) - n(n-1)\sin(x + \frac{n\pi}{2})$$

Use this to deduce the value of the n-th derivative at $x = 0$.

4. Parametric Tangents. The curve known as the Astroid is given by $x = a\cos^3 t$, $y = a\sin^3 t$.

(a) Find the equation of the tangent line at a general parameter t.
(b) Show that the length of the segment of the tangent line intercepted between the coordinate axes is constant and equal to a.

5. Implicit Differentiation. Find $\frac{dy}{dx}$ and $\frac{d^2y}{dx^2}$ at the point $(1, 1)$ for the curve defined by:

$$x^3 + y^3 - 3xy - x + y = -1$$

6. Continuity vs. Differentiability. Construct examples to demonstrate the independence of continuity and differentiability properties:

(a) A function continuous everywhere but differentiable everywhere except at $x = \pm 1$.

(b) A function differentiable at $x = 0$ but discontinuous at every other point.

(c) State for what values of x the function $f(x) = \lfloor x \rfloor$ fails to be continuous and differentiable.

(d) State for what values of x the function $g(x) = |x - 1|$ fails to be differentiable.

7. Odd and Even Derivatives. Let $f : \mathbb{R} \to \mathbb{R}$ be a differentiable function.

(a) Prove that if f is an even function ($f(-x) = f(x)$), then f' is an odd function.
(b) Prove that if f is an odd function ($f(-x) = -f(x)$), then f' is an even function.
(c) Is the converse true? Does f' being odd imply f is even?

8. The Derivative of the Inverse. Let f be strictly increasing and differentiable on an interval I, with $f'(x) \neq 0$.

(a) Using the definition of the derivative and the continuity of the inverse, prove the formula $(f^{-1})'(y) = 1/f'(f^{-1}(y))$.
(b) If $f(x) = x + \sin x$, find $(f^{-1})'(1)$. Note that you cannot find a closed form for $f^{-1}(y)$.

9. Infinite Differentiability. Consider the function $f : \mathbb{R} \to \mathbb{R}$ defined by:

$$f(x) = \begin{cases} e^{-1/x^2} & x \neq 0 \\ 0 & x = 0 \end{cases}$$

(a) Prove that f is continuous at 0.
(b) Prove that $f'(0)$ exists and equals 0.
(c) Prove by induction that for any $n \in \mathbb{N}$, the n-th derivative $f^{(n)}(0)$ exists and equals 0.
(d) Conclude that a function can be smooth (C^∞) and have all derivatives vanish at a point without being the zero function.

10. Searching for c. For each function, find the exact value of c satisfying the Mean Value Theorem on the given interval $[a, b]$. If no such c exists, explain which condition of the theorem is violated.

(a) $f(x) = x(x-2)$ on $[1, 3]$.
(b) $f(x) = |x|$ on $[-1, 2]$.
(c) $f(x) = 1/x$ on $[-1, 1]$.
(d) $f(x) = x^3$ on $[-1, 1]$.

11. **Limits of the Mean Value.** Let $f(x) = x^n$ for $n \geq 2$. By the Mean Value Theorem (MVT) on $[0, h]$, there exists $\theta_h \in (0, 1)$ such that $f(h) - f(0) = hf'(\theta_h h)$. Calculate explicitly the value of θ_h as a function of h and find $\lim_{h \to 0} \theta_h$.

12. **A "Proof" of Continuity.** Critique the following argument: "Let f be differentiable on (a, b). By the MVT, for any $x, x+h \in (a, b)$, $f(x+h) - f(x) = hf'(x + \theta h)$ where $0 < \theta < 1$. Letting $h \to 0$, the LHS tends to 0. The RHS tends to $0 \cdot f'(x) = 0$. Thus $f'(x + \theta h)$ must tend to $f'(x)$. Therefore, the derivative f' is continuous."

    Remark. Consider the function $f(x) = x^2 \sin(1/x)$ discussed in the text.

13. **Stability of the Supremum.** Let f be continuous on $[a, b]$. Define $M(x) = \sup\{f(t) : a \leq t \leq x\}$. Suppose f is differentiable at $x_0 \in (a, b)$ and $f(x_0) < M(x_0)$. Prove that there is a neighbourhood of x_0 where $M(x)$ is constant.

14. **Bounded Derivative Implies Lipschitz.** Prove that if f is differentiable on an interval I and there exists K such that $|f'(x)| \leq K$ for all $x \in I$, then f is Lipschitz continuous on I. That is, $|f(x) - f(y)| \leq K|x - y|$.

15. **Uniqueness of Differential Equations.** Suppose $f, g : \mathbb{R} \to \mathbb{R}$ are differentiable functions satisfying $f'(x) = g'(x)$ for all x and $f(0) = g(0)$. Prove that $f(x) = g(x)$ for all x.

16. **Inequalities via Derivatives.** Prove the following inequalities for $x > 0$:

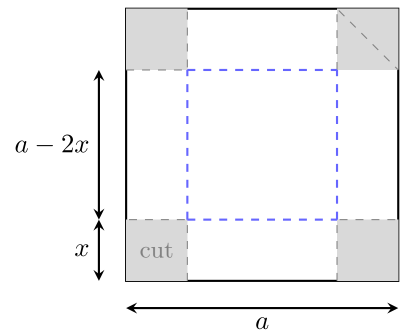
    (a) $\sin x < x$
    (b) $\ln(1 + x) < x$
    (c) $e^x > 1 + x + \frac{x^2}{2}$

17. **Optimisation Geometry.** Find the maximum and minimum values of the following functions:

    (a) $f(x) = \frac{x}{1+x^2}$ on $\mathbb{R}$.
    (b) $g(x) = a \sec x + b \csc x$ on $(0, \pi/2)$ where $a, b > 0$.

18. **Fermat's Reflection Principle.** Two points A and B lie on the same side of a line L. A ray of light travels from A to a point P on L and then to B. The light travels in straight lines. Fermat's Principle states that light takes the path requiring the minimum time (and thus minimum distance in a uniform medium). Use calculus to prove that for the path length $AP + PB$ to be minimised, the angle of incidence must equal the angle of reflection.

19. **The Box Problem.** Find the maximum volume of an open rectangular box that can be made from a square sheet of cardboard of side length a by cutting equal squares from the four corners and folding up the sides. See fig. 9.5.



Flat template

Figure 9.5: The box problem

20. Series Comparison via Limits. Let $\sum a_n$ and $\sum b_n$ be series with positive terms. Prove the Limit Comparison Test: If $\lim_{n\to\infty} \frac{a_n}{b_n} = c$ where $0 < c < \infty$, then $\sum a_n$ converges if and only if $\sum b_n$ converges.

Remark. Use the definition of the limit to bound a_n by $(c-\epsilon)b_n$ and $(c+\epsilon)b_n$ eventually.

21. Generalised Trigonometric Derivatives. Determine the n-th derivative for the following functions:

(a) $y = e^{ax}\sin(bx)$.
(b) $y = \frac{1}{x^2+a^2}$.

Remark. Factorise the denominator over $\mathbb{C}$.

22. Parity of Derivatives. Let $y = \frac{1}{x^2-4}$. Show that the value of the n-th derivative $y^{(n)}$ at $x = 0$ is 0 if n is even, and $-n!(1/2)^{n+1}$ if n is odd.

23. Rational Extrema. Investigate the maxima and minima of the function:

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

where a, b are distinct positive constants.

24. Cauchy Mean Value Theorem. The Generalised (Cauchy) Mean Value Theorem states that for functions f, g continuous on $[a, b]$ and differentiable on (a, b), there exists $c \in (a, b)$ such that $f'(c)(g(b) - g(a)) = g'(c)(f(b) - f(a))$.

(a) Verify this theorem for $f(t) = t^2$ and $g(t) = 4t^3 - 3t$ on the interval $[-1, 1]$.
(b) Interpret the result geometrically in the xy-plane for the curve defined parametrically by $x = f(t), y = g(t)$.

25. Power Dominance. Let $a > 1$ and $k \in \mathbb{N}$. We proved that a^n grows faster than n^k for integer n. Extend this to the continuous variable:

$$\lim_{x\to\infty} \frac{x^k}{a^x} = 0$$

Remark. Use the floor function $\lfloor x \rfloor$ to bound x between integers and apply the Squeeze Theorem.

26. Oscillatory Non-Existence. We examined $\sin(1/x)$. Now consider the function $f(x) = \operatorname{sgn}(\sin(1/x))$.

(a) Describe the behaviour of f near $x = 0$.
(b) Construct sequences (a_n) and (b_n) converging to 0 such that $f(a_n) \to 1$ and $f(b_n) \to -1$.
(c) Conclude that $\lim_{x\to 0} f(x)$ does not exist.

27. One-Sided Limits and Monotonicity. Let $f : (a, b) \to \mathbb{R}$ be a non-decreasing function. Let $c \in (a, b)$.

(a) Prove that the one-sided limits $\lim_{x\to c^-} f(x)$ and $\lim_{x\to c^+} f(x)$ always exist.

(b) Show that $\lim_{x \to c^-} f(x) \le f(c) \le \lim_{x \to c^+} f(x)$.
(c) Deduce that the set of discontinuities of a monotone function is at most countable.

28. Asymptotic Derivative Behaviour. Let $f : (0, \infty) \to \mathbb{R}$ be differentiable. Prove that if $\lim_{x \to \infty} f'(x) = 1$, then:

$$\lim_{x \to \infty} \frac{f(x)}{x} = 1.$$

Remark. Apply the Mean Value Theorem (MVT) on intervals of the form $[x, 2x]$.

Advanced Analytical Methods

29. Legendre Polynomials. The n-th Legendre polynomial is defined by the Rodrigues formula:

$$P_n(x) = \frac{1}{2^n n!} \frac{d^n}{dx^n} (x^2 - 1)^n.$$

(a) Calculate $P_0(x), P_1(x)$, and $P_2(x)$.
(b) Use the Generalised Product Rule (Leibniz Rule) to prove that $P_n(1) = 1$ and $P_n(-1) = (-1)^n$.
(c) Show that $P_n(x)$ satisfies the root property described in algebraic theory: all roots of $P_n(x)$ are real and lie in $[-1, 1]$.

30. Smooth Functions with Compact Support. Earlier in Exercise **9.**, we examined $f(x) = e^{-1/x^2}$. We now construct a "bump function" supported on a finite interval. Let $T(x) = \frac{1}{1-x^2}$ for $|x| < 1$. Define $g : \mathbb{R} \to \mathbb{R}$ by:

$$g(x) = \begin{cases} e^{-T(x)} & \text{if } |x| < 1 \\ 0 & \text{if } |x| \ge 1 \end{cases}$$

(a) Let $F_n(x)$ denote the n-th derivative of $e^{-T(x)}$ for $|x| < 1$. Prove by induction that $F_n(x) = P_n(x)T(x)^{k_n} e^{-T(x)}$ for some polynomial P_n and integer k_n.
(b) Using the limits of rational-exponential products, show that $\lim_{x \to 1^-} g^{(n)}(x) = 0$.
(c) Conclude that g is a C^∞ (smooth) function that is non-zero only on $(-1, 1)$.

31. Bernstein Polynomial Identities. For $n \in \mathbb{N}$ and $t \in \mathbb{R}$, consider the function $h_n(t) = (tp+q)^n$. By computing the first two derivatives of $h_n(t)$ in two different ways (chain rule vs. binomial expansion), prove the following identities used in approximation theory:

(a)

$$\sum_{k=1}^{n} k \binom{n}{k} t^{k-1} p^k q^{n-k} = np(tp + q)^{n-1}.$$

(b)

$$\sum_{k=2}^{n} k(k-1) \binom{n}{k} t^{k-2} p^k q^{n-k} = n(n-1)p^2(tp + q)^{n-2}.$$

32. Differential Equations and Exponentials. Fix $\lambda \in \mathbb{R}$. Suppose $u : \mathbb{R} \to \mathbb{R}$ is differentiable and satisfies $u'(t) = \lambda u(t)$ for all t. Prove that $u(t) = ce^{\lambda t}$ for some constant c.

Remark. Consider the derivative of the auxiliary function $h(t) = u(t)e^{-\lambda t}$.

33. The Wronskian. Let u, v be twice differentiable solutions to the harmonic equation $y'' + by' + cy = 0$. The Wronskian is defined as $W(t) = u(t)v'(t) - u'(t)v(t)$.

(a) Differentiate $W(t)$ and show that $W'(t) + bW(t) = 0$.
(b) Deduce that $W(t) = W(0)e^{-bt}$.
(c) Using this, prove that if u, v satisfy $u'' + u = 0$ (harmonic oscillator with $b = 0, c = 1$) and satisfy the same initial conditions $u(0) = v(0)$ and $u'(0) = v'(0)$, then $u(t) \equiv v(t)$.

34. Characterisation of Sine and Cosine. Using the Wronskian uniqueness result from the previous problem, prove the Pythagorean identity for any solution to $u'' + u = 0$:

$$\left(\frac{u'(t)}{u'(0)}\right)^2 + \left(\frac{u(t)}{u(0)}\right)^2 = 1$$

(Assuming appropriate non-zero initial conditions).

35. The Limit of the Derivative. Suppose $f : [a, b] \to \mathbb{R}$ is continuous on $[a, b]$ and differentiable on (a, b). Prove that if $\lim_{x \to a^+} f'(x) = A$, then the right-hand derivative $f'_+(a)$ exists and equals A.

Remark. This converse to "differentiability implies continuity of the derivative" is false, but this specific limit property holds. Use the Mean Value Theorem (MVT) on $[a, x]$.

36. Taylor-Type Inequalities.

(a) Prove that for $x \geq 0$, $\sin x \geq x - \frac{x^3}{6}$.

(b) Prove that for $x \geq 0$, $1 + x + \frac{x^2}{2} + \cdots + \frac{x^n}{n!} \leq e^x$.

37. Monotonicity of the Power-Exponential. Consider $f : (0, \infty) \to \mathbb{R}$ defined by $f(x) = (1 + 1/x)^x$. Prove that f is strictly increasing.

Remark. It is easier to analyse the derivative of $\ln f(x)$. You may need the inequality $\ln(1 + t) > \frac{t}{1+t}$ for $t > 0$.

Chapter 10

Taylor's Theorem

By the linear approximation formula, if f is differentiable at c, then

$$f(x) = f(c) + f'(c)(x - c) + o(x - c) \quad \text{as } x \to c$$

Linearisation captures the local rate of change but not higher-order geometry, so we seek a polynomial of degree at most n whose derivatives agree with those of f through order n at c.

10.1 Taylor Polynomials

Let I be an interval and let $f : I \to \mathbb{R}$ be a function that is n-times differentiable at a point $c \in I$. We seek a polynomial $P_n(x)$ of degree at most n whose derivatives agree with those of f at c:

$$P_n^{(k)}(c) = f^{(k)}(c) \quad \text{for } k = 0, 1, \dots, n$$

Let $P_n(x) = a_0 + a_1(x - c) + a_2(x - c)^2 + \cdots + a_n(x - c)^n$. Differentiating this polynomial k times and evaluating at $x = c$ eliminates all terms $(x - c)^j$ where $j \neq k$, leaving only the constant term from the k-th differentiation:

$$P_n^{(k)}(c) = k!\, a_k$$

Equating this to $f^{(k)}(c)$ determines the coefficients uniquely: $a_k = \frac{f^{(k)}(c)}{k!}$.

Definition 10.1.1. ***Taylor Polynomial***. Let $f : I \to \mathbb{R}$ be n-times differentiable at $c \in I$. The n-th order Taylor polynomial of f centred at c, denoted $T_n(x)$, is defined by:

$$T_n(x) = \sum_{k=0}^{n} \frac{f^{(k)}(c)}{k!}(x - c)^k$$

If $c = 0$, this is often referred to as the **Maclaurin polynomial**.

Note. The term "order n" refers to the number of derivatives used. The algebraic degree of T_n is at most n, but could be lower if $f^{(n)}(c) = 0$.

Example 10.1.1. The Exponential Function. Let $f(x) = e^x$ and $c = 0$. Since $f^{(k)}(x) = e^x$ for all k, we have $f^{(k)}(0) = 1$. The Taylor polynomial of order n is:

$$T_n(x) = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots + \frac{x^n}{n!}$$

Example 10.1.2. The Sine Function. Let $f(x) = \sin x$ and $c = 0$. The derivatives cycle: $\sin, \cos, -\sin, -\cos$. At 0, these values are $0, 1, 0, -1$. Thus, only odd powers survive:

$$T_{2n+1}(x) = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \cdots + (-1)^n \frac{x^{2n+1}}{(2n+1)!}$$

Remark. $T_{2n+1}(x) = T_{2n+2}(x)$ because the coefficient of x^{2n+2} is zero.

Example 10.1.3. The Cosine Function. Let $f(x) = \cos x$ and $c = 0$. The derivatives cycle: $\cos, -\sin, -\cos, \sin$. At 0, these values are $1, 0, -1, 0$. Thus, only even powers survive:

$$T_{2n}(x) = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \cdots + (-1)^n \frac{x^{2n}}{(2n)!}$$

Remark. $T_{2n}(x) = T_{2n+1}(x)$ because the coefficient of x^{2n+1} is zero.

10.2 Taylor's Theorem

The polynomial $T_n(x)$ is constructed solely from local information at c. To determine how well $T_n(x)$ approximates $f(x)$ for $x \neq c$, we analyse the remainder term:

$$R_n(x) = f(x) - T_n(x)$$

The following theorem, a powerful generalisation of the Mean Value Theorem, provides an explicit form for this error.

***Theorem* 10.2.1. Taylor's Theorem (Lagrange Form).** Let $f : [a, b] \to \mathbb{R}$ be such that $f^{(n)}$ is continuous on $[a, b]$ and $f^{(n+1)}$ exists on (a, b). Let $c, x \in [a, b]$ with $c \neq x$. Then there exists a point ξ strictly between c and x such that:

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

That is, the remainder is given by $R_n(x) = \frac{f^{(n+1)}(\xi)}{(n+1)!}(x - c)^{n+1}$.

Proof. Fix c and x, and define M by

$$f(x) = T_n(x) + M(x - c)^{n+1}$$

We show that $M = \frac{f^{(n+1)}(\xi)}{(n+1)!}$ for some ξ.

To apply Rolle's Theorem, we construct a specific auxiliary function $g(t)$ on the interval between c and x. Let:

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

We verify the boundary values of g:

1. **At $t = x$:** For $k \geq 1$, each term contains $(x - t)^k$ which vanishes at $t = x$. The only surviving term in the sum is for $k = 0$, which is $f(x)$. The remainder term $M(x - t)^{n+1}$ also vanishes.
$$g(x) = f(x) - f(x) - 0 = 0$$
2. **At $t = c$:** By the definition of M:
$$g(c) = f(x) - \left(\sum_{k=0}^{n} \frac{f^{(k)}(c)}{k!}(x - c)^k \right) - M(x - c)^{n+1} = f(x) - T_n(x) - R_n(x) = 0$$

Since f satisfies the differentiability hypotheses, g is continuous on the closed interval between c and x and differentiable on the interior. By Rolle's Theorem, there exists ξ strictly between c and x such that $g'(\xi) = 0$.

We differentiate $g(t)$ with respect to t. Note that for the sum, we use the product rule on terms of the form $\frac{f^{(k)}(t)}{k!}(x - t)^k$:

$$\frac{d}{dt}\left(\sum_{k=0}^{n} \frac{f^{(k)}(t)}{k!}(x - t)^k \right) = f'(t) + \sum_{k=1}^{n} \frac{f^{(k+1)}(t)}{k!}(x - t)^k - \sum_{k=1}^{n} \frac{f^{(k)}(t)}{(k - 1)!}(x - t)^{k-1}$$

Reindexing the second sum with $j = k - 1$:

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

The term $j = 0$ is $f'(t)$, which cancels the separate $f'(t)$ term. The remaining terms cancel with the first sum (shifted), leaving only the $k = n$ term from the first sum.

$$\frac{d}{dt}\left(\sum_{k=0}^{n} \cdots\right) = \frac{f^{(n+1)}(t)}{n!}(x-t)^{n}$$

Now we differentiate the remainder part $-M(x-t)^{n+1}$:

$$\frac{d}{dt}\left(-M(x-t)^{n+1}\right) = -M(n+1)(x-t)^{n}(-1) = M(n+1)(x-t)^{n}$$

Combining these into $g'(t)$ (noting the negative sign in front of the sum in the definition of g):

$$g'(t) = 0 - \frac{f^{(n+1)}(t)}{n!}(x-t)^{n} + M(n+1)(x-t)^{n}$$

$$g'(t) = (x-t)^{n}\left[M(n+1) - \frac{f^{(n+1)}(t)}{n!}\right]$$

Setting $g'(\xi) = 0$, and noting that $\xi \neq x$ implies $(x-\xi)^{n} \neq 0$:

$$M(n+1) - \frac{f^{(n+1)}(\xi)}{n!} = 0 \implies M = \frac{f^{(n+1)}(\xi)}{(n+1)!}$$

Substituting M back into the original equation for $f(x)$ completes the proof. ■

Remark. For $n = 0$, this theorem recovers the Mean Value Theorem: $f(x) = f(c) + f'(\xi)(x-c)$. Thus, Taylor's Theorem can be viewed as an extension of the MVT to higher orders.

Peano Form of the Remainder

While the Lagrange form provides an explicit formula, we sometimes only require the asymptotic behaviour of the error as $x \to c$.

***Proposition* 10.2.1.** *Taylor's Theorem (Peano Form).* Let f be n-times continuously differentiable in a neighbourhood of c. Then

$$f(x) = T_n(x) + o(|x-c|^{n}) \quad \text{as } x \to c.$$

Equivalently,

$$\lim_{x \to c} \frac{f(x) - T_n(x)}{|x-c|^{n}} = 0.$$

Proof. We argue by induction on n.

Base case ($n = 1$). Here $T_1(x) = f(c) + f'(c)(x-c)$. The claim $f(x) = T_1(x) + o(|x-c|)$ is equivalent to $\lim_{x\to c} \frac{f(x)-f(c)-f'(c)(x-c)}{x-c} = 0$, which is exactly the definition of $f'(c)$.

Inductive step. Assume the result holds for order $n-1$, i.e. for any function g that is $(n-1)$-times continuously differentiable near c,

$$g(x) = T_{n-1}(g)(x) + o\big(|x-c|^{n-1}\big) \quad (x \to c).$$

Let f be n-times continuously differentiable near c, and define the remainder $R_n(x) := f(x) - T_n(x)$. Then $R_n(c) = 0$. Differentiating,

$$R_n'(x) = f'(x) - \frac{d}{dx}\left(\sum_{k=0}^{n} \frac{f^{(k)}(c)}{k!}(x-c)^k\right) = f'(x) - \sum_{j=0}^{n-1} \frac{(f')^{(j)}(c)}{j!}(x-c)^j = f'(x) - T_{n-1}(f')(x).$$

By the inductive hypothesis applied to $g = f'$, $R_n'(x) = o\big(|x-c|^{n-1}\big) \quad$ as $x \to c$, equivalently,

$$\lim_{x\to c} \frac{R_n'(x)}{|x-c|^{n-1}} = 0.$$

Now fix $x \neq c$ sufficiently close to c. Apply the Mean Value Theorem to R_n on the interval with endpoints c and x. There exists ξ strictly between c and x such that $R_n(x) - R_n(c) = R_n'(\xi)(x-c)$, so (since $R_n(c) = 0$) $R_n(x) = R_n'(\xi)(x-c)$. Therefore,

$$\frac{|R_n(x)|}{|x-c|^n} = \frac{|R_n'(\xi)|}{|x-c|^{n-1}} = \frac{|R_n'(\xi)|}{|\xi-c|^{n-1}} \left(\frac{|\xi-c|}{|x-c|}\right)^{n-1} \leq \frac{|R_n'(\xi)|}{|\xi-c|^{n-1}},$$

because ξ lies between c and x, hence $|\xi - c| \leq |x-c|$. As $x \to c$, we also have $\xi \to c$, and thus

$$\frac{|R_n'(\xi)|}{|\xi-c|^{n-1}} \longrightarrow 0.$$

By the inequality above, this forces

$$\lim_{x\to c} \frac{|R_n(x)|}{|x-c|^n} = 0,$$

i.e. $R_n(x) = o(|x-c|^n)$ as $x \to c$.

■

The Peano form is asymptotic: it describes the size of the remainder as $x \to c$. Unlike the Lagrange form, it does not identify ξ or give a pointwise formula for $R_n(x)$ for fixed $x \neq c$.

10.3 Applications of Taylor's Theorem

The Lagrange remainder term allows us to approximate functions with rigorous error bounds.

Example 10.3.1. Approximating e. We compute e to a given precision. Let $f(x) = e^x$ expanded at $c = 0$.

$$e = f(1) = 1 + 1 + \frac{1}{2!} + \cdots + \frac{1}{n!} + R_n(1)$$

The remainder is $R_n(1) = \frac{e^\xi}{(n+1)!}$ for some $\xi \in (0,1)$. Since e^x is increasing, $e^\xi < e^1 < 3$. Thus:

$$R_n(1) < \frac{3}{(n+1)!}$$

To get an error less than 10^{-3}, we need $3/(n+1)! < 10^{-3}$, or $(n+1)! > 3000$. Since $6! = 720$ and $7! = 5040$, we choose $n = 6$.

$$e \approx 1 + 1 + 0.5 + 0.1666\cdots + \cdots + \frac{1}{720}$$

Example 10.3.2. Irrationality of e. Taylor's theorem provides a slick proof that e is irrational. Suppose $e = p/q$ for integers $p, q > 0$. Choose $n > q$ and $n > 3$. From the expansion above:

$$e = \sum_{k=0}^{n} \frac{1}{k!} + \frac{e^\xi}{(n+1)!}$$

Multiply both sides by $n!$:

$$n!e = \sum_{k=0}^{n} \frac{n!}{k!} + \frac{e^\xi}{n+1}$$

The LHS is $n!(p/q)$. Since $n > q$, q divides $n!$, so LHS is an integer. The sum term $\sum \frac{n!}{k!}$ is an integer (since $k \leq n$). This forces the remainder term $\frac{e^\xi}{n+1}$ to be an integer. However, $1 < e < 3$, so $0 < e^\xi < 3$. Since $n > 3$, we have $0 < \frac{e^\xi}{n+1} < \frac{3}{4}$. No integer lies strictly between 0 and 3/4. Contradiction.

If f is infinitely differentiable ($f \in C^\infty$), we can define the Taylor polynomial $T_n(x)$ for every n. If the sequence of polynomials converges as $n \to \infty$, we obtain a power series.

Definition 10.3.1. ***Taylor Series***. Let $f \in C^\infty(I)$ and $c \in I$. The Taylor series of f at c is:

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

We say f is analytic at c if this series converges to $f(x)$ on some neighbourhood of c.

Note. Convergence of the series is equivalent to $\lim_{n\to\infty} R_n(x) = 0$.

***Theorem* 10.3.1. Convergence of the Exponential Series**. For any $x \in \mathbb{R}$, $e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!}$.

Proof. Fix x. The remainder is $R_n(x) = \frac{e^\xi}{(n+1)!}x^{n+1}$ where ξ is between 0 and x. Note that $|\xi| \leq |x|$, so $e^\xi \leq e^{|x|}$.

$$|R_n(x)| \leq e^{|x|}\frac{|x|^{n+1}}{(n+1)!}$$

For any fixed real number K, the factorial grows faster than the power K^n (by the Ratio Test on the sequence terms). Thus $\lim_{n\to\infty} \frac{|x|^{n+1}}{(n+1)!} = 0$. Consequently, $R_n(x) \to 0$ for all x. ∎

The Warning: Smooth vs Analytic

It is a common misconception that if a function is smooth (infinitely differentiable), it must equal its Taylor series. This is true for complex differentiable functions, but false for real functions.

Example 10.3.3. The Bump Function. Consider the function $f : \mathbb{R} \to \mathbb{R}$ defined by:

$$f(x) = \begin{cases} e^{-1/x^2} & x \neq 0 \\ 0 & x = 0 \end{cases}$$

1. **Continuity at 0:** As $x \to 0$, $-1/x^2 \to -\infty$, so $e^{-1/x^2} \to 0$. Thus f is continuous.
2. **Differentiability at 0:** We consider the Newton quotient:

$$\lim_{h\to 0} \frac{e^{-1/h^2} - 0}{h}$$

   Let $t = 1/h$. As $h \to 0$, $t \to \infty$. The limit becomes $\lim_{t\to\infty} te^{-t^2}$. By the growth hierarchy (exponentials beat polynomials), this limit is 0. Thus $f'(0) = 0$.
3. **Higher Derivatives:** It can be shown by induction that for $x \neq 0$, $f^{(n)}(x) = P_n(1/x)e^{-1/x^2}$ where P_n is a polynomial. The limit as $x \to 0$ is always 0.

Therefore, $f^{(n)}(0) = 0$ for all n. The Maclaurin series of f is:

$$\sum_{n=0}^{\infty} \frac{0}{n!}x^n = 0 + 0 + 0 + \cdots = 0$$

This series converges to 0 for all x. However, $f(x) > 0$ for all $x \neq 0$. Thus, $f(x) \neq T_\infty(x)$ except at the single point $x = 0$. This function is C^∞ but not analytic.

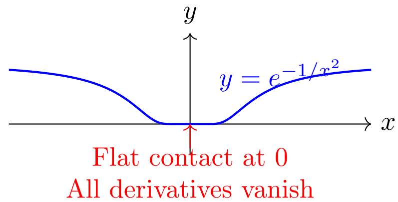


Figure 10.1: The function e^{-1/x^2} is incredibly flat at the origin, flatter than any polynomial.

10.4 The Second Derivative Test

We can now justify the Second Derivative Test for local extrema using Taylor's Theorem.

***Theorem* 10.4.1. Second Derivative Test**. Let f be twice differentiable on (a, b) and let $c \in (a, b)$ be a stationary point ($f'(c) = 0$).

1. If $f''(c) < 0$, then f has a local maximum at c.
2. If $f''(c) > 0$, then f has a local minimum at c.

Proof. We use the Peano form of Taylor's Theorem at c with $n = 2$:

$$f(x) = f(c) + f'(c)(x-c) + \frac{f''(c)}{2}(x-c)^2 + o((x-c)^2)$$

Since $f'(c) = 0$, this simplifies to:

$$f(x) - f(c) = (x-c)^2 \left[\frac{f''(c)}{2} + \frac{o((x-c)^2)}{(x-c)^2} \right]$$

Let $\epsilon(x) = \frac{o((x-c)^2)}{(x-c)^2}$. By definition, $\lim_{x \to c} \epsilon(x) = 0$. Suppose $f''(c) > 0$. Since $\epsilon(x) \to 0$, there exists a neighbourhood $(c - \delta, c + \delta)$ where $|\epsilon(x)| < \frac{f''(c)}{4}$. In this neighbourhood, the term in the square brackets satisfies:

$$\frac{f''(c)}{2} + \epsilon(x) > \frac{f''(c)}{2} - \frac{f''(c)}{4} = \frac{f''(c)}{4} > 0$$

Since $(x-c)^2 > 0$ for $x \neq c$, we have $f(x) - f(c) > 0$, or $f(x) > f(c)$. Thus c is a local minimum. The proof for the maximum is analogous. ■

Remark. If Second Derivative Test gives no information, the point c may be a local minimum, a local maximum, or neither (an inflection point). In such cases, one must examine higher-order derivatives or apply the First Derivative Test.

10.5 L'Hôpital's Rule

The most famous application of the Generalised Mean Value Theorem is the resolution of indeterminate forms. Limits such as $\lim_{x \to 0} \frac{\sin x}{x}$ or $\lim_{x \to \infty} \frac{\ln x}{x}$ cannot be evaluated by the quotient law for limits because they yield the meaningless forms $0/0$ or ∞/∞. L'Hôpital's Rule allows us to replace the ratio of functions with the ratio of their derivatives.

While often applied mechanically, the rigorous justification (particularly for the ∞/∞ case) requires careful use of the limit definitions and Cauchy's Finite Increment Theorem.

***Theorem* 10.5.1. L'Hôpital's Rule**. Let $-\infty \leq a < b \leq \infty$ and let f, g be differentiable functions on (a, b) such that $g'(x) \neq 0$ for all $x \in (a, b)$. Suppose that

$$\lim_{x \to b^-} \frac{f'(x)}{g'(x)} = L \in \mathbb{R} \cup \{\pm\infty\}$$

If either:

1. $\lim_{x \to b^-} f(x) = 0$ and $\lim_{x \to b^-} g(x) = 0$ (The 0/0 case), or
2. $\lim_{x \to b^-} |g(x)| = \infty$ (The ∞/∞ case),

then:

$$\lim_{x \to b^-} \frac{f(x)}{g(x)} = L$$

Note. The theorem holds analogously for limits as $x \to a^+$, or for two-sided limits at a finite point c. The case $b = \infty$ reduces to the finite case via the substitution $t = 1/x$.

Proof. We treat the two cases separately. We assume $L \in \mathbb{R}$ for the exposition; the infinite cases follow similar logic with appropriate modification of inequalities.

Case 1: The 0/0 Form. Assume b is finite and that $f(x) \to 0, g(x) \to 0$ as $x \to b^-$. We define the extensions of f and g to the point b by setting $f(b) = 0$ and $g(b) = 0$. Since the limits are zero, these extended functions are continuous on the interval $(a, b]$. Let $x \in (a, b)$. We apply Theorem 9.9.4 to the interval $[x, b]$. There exists $c \in (x, b)$ such that:

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

Substituting the zeroes:

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

As $x \to b^-$, the constraint $x < c < b$ forces $c \to b^-$. Since $\frac{f'(c)}{g'(c)} \to L$, it follows immediately that $\frac{f(x)}{g(x)} \to L$.

Case 2: The ∞/∞ Form. This case is more subtle because we cannot "extend" the functions to infinity. We must use the general $\epsilon - \delta$ machinery coupled with Cauchy's Finite Increment Theorem on an interior interval.

Let $\epsilon > 0$. Since $\frac{f'(x)}{g'(x)} \to L$, there exists $x_0 \in (a, b)$ such that for all $c \in (x_0, b)$:

$$\left| \frac{f'(c)}{g'(c)} - L \right| < \frac{\epsilon}{2}$$

Fix a point $y \in (x_0, b)$. For any $x \in (y, b)$, we apply Theorem 9.9.4 to the interval $[y, x]$. There exists $c \in (y, x)$ (and thus $c > x_0$) such that:

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

Let $R = \frac{f'(c)}{g'(c)}$. By our choice of x_0, we have $|R - L| < \epsilon/2$. We now manipulate the ratio algebraically to isolate $\frac{f(x)}{g(x)}$.

$$\frac{f(x) - f(y)}{g(x) - g(y)} = \frac{f(x)}{g(x)} \cdot \frac{1 - \frac{f(y)}{f(x)}}{1 - \frac{g(y)}{g(x)}}$$

Therefore:

$$\frac{f(x)}{g(x)} = R \cdot \frac{1 - \frac{g(y)}{g(x)}}{1 - \frac{f(y)}{f(x)}}$$

As $x \to b^-$, we have $|g(x)| \to \infty$. Consequently, $\frac{g(y)}{g(x)} \to 0$ and $\frac{f(y)}{f(x)} \to 0$ (assuming $f \to \infty$; if f remains bounded the result is trivial). Thus, there exists $x_1 > y$ such that for all $x \in (x_1, b)$, the term in the brackets is sufficiently close to 1 that the entire expression stays within ϵ of L.

To be precise, as $x \to b$, the correction factor approaches 1. Since R is $\epsilon/2$-close to L, the product will eventually be ϵ-close to L. Thus $\lim_{x\to b^-} \frac{f(x)}{g(x)} = L$.

■

Example 10.5.1. Basic Application. Evaluate $\lim_{x\to 0} \frac{\sqrt{1+x}-1-x/2}{x^2}$. Direct substitution yields 0/0. Applying L'Hôpital's Rule:

$$\lim_{x\to 0} \frac{\frac{1}{2}(1+x)^{-1/2} - \frac{1}{2}}{2x}$$

This is still 0/0. We apply the rule a second time:

$$\lim_{x\to 0} \frac{-\frac{1}{4}(1+x)^{-3/2}}{2} = \frac{-1/4}{2} = -\frac{1}{8}$$

Example 10.5.2. Circular Reasoning. Consider $\lim_{x\to 0} \frac{\sin x}{x}$. This is a 0/0 form. Applying L'Hôpital's Rule gives $\lim_{x\to 0} \frac{\cos x}{1} = 1$. While the answer is correct, the logic is circular. The derivative of $\sin x$ is established *using* the fact that $\lim_{x\to 0} \frac{\sin x}{x} = 1$. Thus, one should not use L'Hôpital's Rule to prove this specific limit, though it serves as a consistency check.

Example 10.5.3. Standard Applications.

1. **Trigonometric Limit:** Consider $\lim_{x\to 0} \frac{1-\cos x}{x^2}$. This is a 0/0 form. $f'(x) = \sin x$ and $g'(x) = 2x$. The new limit is $\lim_{x\to 0} \frac{\sin x}{2x}$. This is still 0/0. Applying the rule again: $\lim_{x\to 0} \frac{\cos x}{2} = \frac{1}{2}$. Thus the original limit is 1/2.
2. **Exponential Dominance:** Consider $\lim_{x\to\infty} \frac{x^n}{e^x}$ for $n \in \mathbb{N}$. This is an ∞/∞ form. Differentiating n times:
$$\lim_{x\to\infty} \frac{x^n}{e^x} = \lim_{x\to\infty} \frac{nx^{n-1}}{e^x} = \cdots = \lim_{x\to\infty} \frac{n!}{e^x} = 0$$
This confirms the growth hierarchy established in earlier chapters.

Example 10.5.4. Indeterminate Powers. Evaluate $\lim_{x\to 0^+} x^x$. This is of the form 0^0. We use logarithms to convert it to a quotient. Let $y = x^x$, so $\ln y = x \ln x$.

$$\lim_{x\to 0^+} x \ln x = \lim_{x\to 0^+} \frac{\ln x}{1/x}$$

This is now $\frac{-\infty}{\infty}$. Applying L'Hôpital's Rule:

$$\lim_{x\to 0^+} \frac{1/x}{-1/x^2} = \lim_{x\to 0^+} (-x) = 0$$

Since $\ln y \to 0$, by continuity of the exponential, $y \to e^0 = 1$. Thus $\lim_{x\to 0^+} x^x = 1$.

Other Indeterminate Forms L'Hôpital's Rule strictly applies to 0/0 and ∞/∞. Other forms such as $0 \cdot \infty$, $\infty - \infty$, 0^0, 1^∞, and ∞^0 must be algebraically manipulated into a quotient before the rule can be used.

Example 10.5.5. Form ∞/∞ with One-Sided Limits. Evaluate $\lim_{x\to\pi/2} \frac{\sec x}{1+\tan x}$. As $x \to \pi/2$, both $\sec x$ and $\tan x$ approach $\pm\infty$. We evaluate the left and right limits. Consider $x \to \pi/2^-$. The form is ∞/∞.

$$\lim_{x\to\pi/2^-} \frac{\sec x \tan x}{\sec^2 x} = \lim_{x\to\pi/2^-} \frac{\tan x}{\sec x} = \lim_{x\to\pi/2^-} \sin x = 1$$

Similarly, for $x \to \pi/2^+$, the limit is 1. Thus, the limit exists and equals 1. Alternatively, simplifying the original expression yields $\frac{1/\cos x}{(\cos x+\sin x)/\cos x} = \frac{1}{\cos x+\sin x}$, which immediately tends to 1 as $x \to \pi/2$.

Example 10.5.6. Form $\infty - \infty$. Evaluate $\lim_{x \to 0} \left(\frac{1}{\sin x} - \frac{1}{x}\right)$. This yields $\infty - \infty$. We combine the fractions:

$$\lim_{x \to 0} \frac{x - \sin x}{x \sin x}$$

This is now 0/0. Applying L'Hôpital's Rule:

$$\lim_{x \to 0} \frac{1 - \cos x}{\sin x + x \cos x}$$

Still 0/0. Applying the rule again:

$$\lim_{x \to 0} \frac{\sin x}{\cos x + (\cos x - x \sin x)} = \frac{0}{1 + 1 - 0} = 0$$

Example 10.5.7. Form $0 \cdot \infty$. Evaluate $\lim_{x \to \infty} x \sin(1/x)$. Let $t = 1/x$. As $x \to \infty$, $t \to 0^+$.

$$\lim_{t \to 0^+} \frac{\sin t}{t} = 1$$

No L'Hôpital's required if one recognises the standard limit.

Remark. (Warning on Applicability). L'Hôpital's Rule requires the limit of the derivatives to *exist*. Consider $\lim_{x \to \infty} \frac{x + \sin x}{x}$. This is ∞/∞. Differentiation yields $\frac{1 + \cos x}{1}$. As $x \to \infty$, $1 + \cos x$ oscillates and does not converge. Thus L'Hôpital's Rule provides no information. However, standard algebraic manipulation shows $\frac{x + \sin x}{x} = 1 + \frac{\sin x}{x} \to 1$.

10.6 Convexity

We conclude the differential calculus of real-valued functions by examining convexity, a global curvature constraint.

Geometrically, a set in $\mathbb{R}^n$ is convex if the line segment connecting any two points lies entirely within the set. For a function $f : I \to \mathbb{R}$, we apply this to the epigraph, so the chord connecting any two points on the curve lies above the curve.

To formalise the segment between two points, we introduce convex combinations.

***Proposition* 10.6.1.** *Convex Combination of Scalars.* Let $x_1, x_2 \in \mathbb{R}$ with $x_1 < x_2$. A point x lies in the interval $[x_1, x_2]$ if and only if there exist real numbers $t_1, t_2 \geq 0$ such that:

$$t_1 + t_2 = 1 \quad \text{and} \quad x = t_1 x_1 + t_2 x_2$$

Proof. **(⇒)** Suppose $x \in [x_1, x_2]$. Define $t_1 = \frac{x_2 - x}{x_2 - x_1}$ and $t_2 = \frac{x - x_1}{x_2 - x_1}$. Since $x_1 \leq x \leq x_2$, both t_1 and t_2 are non-negative. Their sum is:

$$t_1 + t_2 = \frac{x_2 - x + x - x_1}{x_2 - x_1} = 1$$

The linear combination yields:

$$t_1 x_1 + t_2 x_2 = \frac{x_1(x_2 - x) + x_2(x - x_1)}{x_2 - x_1} = \frac{x_1 x_2 - x_1 x + x_2 x - x_2 x_1}{x_2 - x_1} = \frac{x(x_2 - x_1)}{x_2 - x_1} = x$$

(⇐) Suppose $x = t_1 x_1 + t_2 x_2$ with $t_1, t_2 \geq 0$ and sum 1. Substituting $t_1 = 1 - t_2$:

$$x = (1 - t_2)x_1 + t_2 x_2 = x_1 + t_2(x_2 - x_1)$$

Since $0 \leq t_2 \leq 1$ and $x_2 - x_1 > 0$, we have $0 \leq x - x_1 \leq x_2 - x_1$, implying $x_1 \leq x \leq x_2$.

■

Remark. (Physical Interpretation). Interpreting t_i as normalised masses yields the centre of mass formula $\bar{x} = \sum t_i x_i$ with $\sum t_i = 1$.

The Secant Line Condition

Given a function $f : I \to \mathbb{R}$ and two points $x_1, x_2 \in I$, the secant line $L(x)$ connecting $(x_1, f(x_1))$ and $(x_2, f(x_2))$ is given by the Lagrange interpolation formula:

$$L(x) = f(x_1)\frac{x_2 - x}{x_2 - x_1} + f(x_2)\frac{x - x_1}{x_2 - x_1}$$

Note that the coefficients are precisely the t_1 and t_2 from proposition 10.6.1 for x. Thus, the value of the secant line at a point $x = t_1x_1 + t_2x_2$ is simply $t_1f(x_1) + t_2f(x_2)$. The condition that the graph of f lies below this secant line can be stated succinctly.

Definition 10.6.1. ***Convex Function***. A function $f : I \to \mathbb{R}$ is convex if for all $x_1, x_2 \in I$ and all $t \in [0, 1]$:

$$f((1 - t)x_1 + tx_2) \leq (1 - t)f(x_1) + tf(x_2)$$

Using the notation t_1, t_2, this is:

$$f(t_1x_1 + t_2x_2) \leq t_1f(x_1) + t_2f(x_2)$$

whenever $t_1, t_2 \geq 0$ and $t_1 + t_2 = 1$.

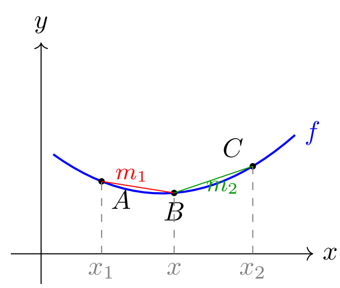
Definition 10.6.2. ***Concave Function***. A function f is concave if the inequality is reversed:

$$f((1 - t)x_1 + tx_2) \geq (1 - t)f(x_1) + tf(x_2)$$

Equivalently, f is concave if and only if $-f$ is convex.

Note. In elementary calculus texts, convex functions are often termed "concave up" and concave functions "concave down". We shall avoid this terminology.

While the definition involving convex combinations is geometrically intuitive, it is often easier to work with the slopes of chords directly. The convexity condition implies that as we move from left to right, the slope of the secant lines must increase.



Slope $m_1 \leq$ Slope m_2

Figure 10.2: The monotonicity of slopes. For a convex function, the slope of the chord connecting x_1 and x is less than or equal to the slope of the chord connecting x and x_2.

***Proposition* 10.6.2.** *Slope Monotonicity Lemma.* Let $f : I \to \mathbb{R}$. The function f is convex if and only if for all $x_1, x, x_2 \in I$ with $x_1 < x < x_2$:

$$\frac{f(x) - f(x_1)}{x - x_1} \leq \frac{f(x_2) - f(x)}{x_2 - x}$$

Proof. By definition 10.6.1, with $t = \frac{x - x_1}{x_2 - x_1}$, we have:

$$f(x) \leq \frac{x_2 - x}{x_2 - x_1}f(x_1) + \frac{x - x_1}{x_2 - x_1}f(x_2)$$

We subtract $f(x_1)$ from both sides:

$$f(x) - f(x_1) \le \left(\frac{x_2 - x}{x_2 - x_1} - 1\right) f(x_1) + \frac{x - x_1}{x_2 - x_1} f(x_2)$$

Simplifying the coefficient of $f(x_1)$:

$$\frac{x_2 - x - (x_2 - x_1)}{x_2 - x_1} = \frac{-(x - x_1)}{x_2 - x_1}$$

Thus:

$$f(x) - f(x_1) \le \frac{x - x_1}{x_2 - x_1}[f(x_2) - f(x_1)]$$

Dividing by $x - x_1$ (which is positive):

$$\frac{f(x) - f(x_1)}{x - x_1} \le \frac{f(x_2) - f(x_1)}{x_2 - x_1}$$

This shows the slope of the first segment is bounded by the slope of the total segment. By a symmetric manipulation (subtracting $f(x_2)$ instead), one shows that the total slope is bounded by the slope of the second segment. Thus:

$$\frac{f(x) - f(x_1)}{x - x_1} \le \frac{f(x_2) - f(x)}{x_2 - x}$$

The converse follows by reversing the algebraic steps. ■

Convexity and Differentiation

The slope monotonicity lemma suggests a strong link between convexity and the increasing nature of the derivative. If f is differentiable, we can take the limits as $x \to x_1$ or $x \to x_2$.

***Theorem* 10.6.1. First Derivative Criterion**. Let $f : I \to \mathbb{R}$ be a differentiable function. The following are equivalent:

1. f is convex on I.
2. f' is monotonically increasing on I.
3. The graph of f lies above its tangents: for all $x, c \in I$,

$$f(x) \ge f(c) + f'(c)(x - c)$$

Proof. $(1 \Rightarrow 2)$ Let $x_1 < x_2$. For any $x \in (x_1, x_2)$, proposition 10.6.2 gives:

$$\frac{f(x) - f(x_1)}{x - x_1} \le \frac{f(x_2) - f(x)}{x_2 - x}$$

Letting $x \to x_1^+$ in the LHS gives $f'(x_1)$. Letting $x \to x_2^-$ in the RHS gives $f'(x_2)$. However, to compare them rigorously, pick u, v such that $x_1 < u < v < x_2$. Applying the lemma repeatedly gives:

$$\frac{f(u) - f(x_1)}{u - x_1} \le \frac{f(v) - f(u)}{v - u} \le \frac{f(x_2) - f(v)}{x_2 - v}$$

Taking limits $u \to x_1$ and $v \to x_2$ implies $f'(x_1) \le f'(x_2)$. Thus f' is increasing.

$(2 \Rightarrow 3)$ By the Mean Value Theorem, there exists ξ between c and x such that $f(x) - f(c) = f'(\xi)(x - c)$. If $x > c$, then $\xi > c$. Since f' is increasing, $f'(\xi) \ge f'(c)$. Multiplying by positive $(x - c)$ preserves the inequality:

$$f(x) - f(c) \ge f'(c)(x - c)$$

If $x < c$, then $\xi < c$. Since f' is increasing, $f'(\xi) \le f'(c)$. Multiplying by negative $(x - c)$ reverses the inequality, yielding again:

$$f'(\xi)(x - c) \ge f'(c)(x - c) \implies f(x) - f(c) \ge f'(c)(x - c)$$

(3 ⇒ 1) Let $x_1, x_2 \in I$ and let $x = tx_1 + (1-t)x_2$. By hypothesis (3) at $c = x$:

$$f(x_1) \geq f(x) + f'(x)(x_1 - x)$$

$$f(x_2) \geq f(x) + f'(x)(x_2 - x)$$

Multiply the first by t and the second by $(1-t)$ and sum:

$$tf(x_1) + (1-t)f(x_2) \geq f(x) + f'(x)[t(x_1 - x) + (1-t)(x_2 - x)]$$

The term in the bracket is $tx_1 + (1-t)x_2 - x = x - x = 0$. Thus $tf(x_1) + (1-t)f(x_2) \geq f(x)$, which is the definition of convexity.

■

***Theorem* 10.6.2. Second Derivative Test for Convexity**. Let f be twice differentiable on I. Then f is convex if and only if $f''(x) \geq 0$ for all $x \in I$.

Proof. This follows from theorem 10.6.1: f is convex $\iff$ f' is increasing. By the Monotonicity Criterion, if $f''(x) \geq 0$, then f' is increasing. Conversely, if f' is increasing, then for $h > 0$, $\frac{f'(x+h)-f'(x)}{h} \geq 0$, and for $h < 0$, $\frac{f'(x+h)-f'(x)}{h} \geq 0$ (negative over negative). Taking the limit as $h \to 0$ gives $f''(x) \geq 0$. ■

Example 10.6.1. Examples of Convex Functions.

1. $f(x) = x^2$ on $\mathbb{R}$. $f''(x) = 2 > 0$. Strictly convex.
2. $f(x) = e^x$ on $\mathbb{R}$. $f''(x) = e^x > 0$. Strictly convex.
3. $f(x) = -\ln x$ on $(0, \infty)$. $f'(x) = -1/x$, $f''(x) = 1/x^2 > 0$. Convex.

The definition of convexity in definition 10.6.1 can be extended by induction to any finite number of points. This result, known as Jensen's Inequality, is the workhorse of analytical inequalities.

***Theorem* 10.6.3. Jensen's Inequality**. Let $f : I \to \mathbb{R}$ be a convex function. For any points $x_1, \ldots, x_n \in I$ and any non-negative weights $t_1, \ldots, t_n$ such that $\sum t_i = 1$:

$$f\left(\sum_{i=1}^{n} t_i x_i\right) \leq \sum_{i=1}^{n} t_i f(x_i)$$

Proof. We proceed by induction on n. The case $n = 1$ is trivial ($t_1 = 1$). The case $n = 2$ is the definition of convexity. Assume the inequality holds for $n = k$. Consider $k+1$ points with weights t_i. If $t_{k+1} = 1$, then other $t_i = 0$ and the result holds. If $t_{k+1} < 1$, let $T = \sum_{i=1}^{k} t_i = 1 - t_{k+1} > 0$. We can write the convex combination as:

$$\sum_{i=1}^{k+1} t_i x_i = t_{k+1} x_{k+1} + T \sum_{i=1}^{k} \frac{t_i}{T} x_i$$

Let $y = \sum_{i=1}^{k} \frac{t_i}{T} x_i$. Note that the weights t_i/T sum to 1, so $y \in I$ by the inductive hypothesis (specifically, the domain is convex). By the two-point case of definition 10.6.1:

$$f(t_{k+1} x_{k+1} + Ty) \leq t_{k+1} f(x_{k+1}) + T f(y)$$

By the inductive hypothesis applied to y:

$$f(y) = f\left(\sum_{i=1}^{k} \frac{t_i}{T} x_i\right) \leq \sum_{i=1}^{k} \frac{t_i}{T} f(x_i)$$

Combining these:

$$f\left(\sum_{i=1}^{k+1} t_i x_i\right) \leq t_{k+1} f(x_{k+1}) + T \sum_{i=1}^{k} \frac{t_i}{T} f(x_i) = \sum_{i=1}^{k+1} t_i f(x_i)$$

■

Example 10.6.2. AM-GM Inequality. Let $x_1, \ldots, x_n > 0$. Consider the convex function $f(x) = -\ln x$. Let weights $t_i = 1/n$. By Jensen's Inequality:

$$-\ln\left(\frac{1}{n}\sum x_i\right) \leq \sum \frac{1}{n}(-\ln x_i) = -\frac{1}{n}\ln\left(\prod x_i\right) = -\ln\left((\prod x_i)^{1/n}\right)$$

Since $-\ln$ is decreasing, removing the function reverses the inequality:

$$\frac{1}{n}\sum_{i=1}^{n} x_i \geq \sqrt[n]{\prod_{i=1}^{n} x_i}$$

This recovers the Arithmetic Mean - Geometric Mean inequality.

Using Jensen's Inequality, we derive Hölder's and Minkowski's inequalities. Hölder's inequality generalises Cauchy-Schwarz to the ℓ^p norms.

***Theorem* 10.6.4. Hölder's Inequality**. Fix a real number $p > 1$ and define $q > 1$ by the conjugacy relation $\frac{1}{p} + \frac{1}{q} = 1$ (equivalently $(p-1)q = p$). For any natural number n and any non-negative real numbers $a_1, \ldots, a_n$ and $b_1, \ldots, b_n$:

$$\sum_{k=1}^{n} a_k b_k \leq \left(\sum_{k=1}^{n} a_k^p\right)^{1/p} \left(\sum_{k=1}^{n} b_k^q\right)^{1/q}$$

Proof. If $\sum b_k^q = 0$, then all $b_k = 0$ and the inequality is trivial ($0 \leq 0$). Assume $B = \sum_{k=1}^{n} b_k^q > 0$. We employ the convex function $f(x) = x^p$ on the domain $[0, \infty)$. (For $p > 1$, $f''(x) = p(p-1)x^{p-2} \geq 0$ for $x > 0$, and convexity extends to 0 by continuity). Define the normalised weights t_k and auxiliary points x_k as follows:

$$t_k = \frac{b_k^q}{B}, \quad x_k = \begin{cases} a_k b_k^{-(q-1)} B & \text{if } b_k \neq 0 \\ 0 & \text{if } b_k = 0 \end{cases}$$

Note that $\sum t_k = \frac{1}{B}\sum b_k^q = 1$, so the weights are valid for Jensen's Inequality. We compute the convex combination $\sum t_k x_k$: If $b_k \neq 0$, $t_k x_k = \frac{b_k^q}{B} a_k b_k^{-(q-1)} B = a_k b_k$. If $b_k = 0$, then $t_k = 0$, so $t_k x_k = 0 = a_k b_k$. Thus $\sum_{k=1}^{n} t_k x_k = \sum_{k=1}^{n} a_k b_k$. Applying theorem 10.6.3 gives $f(\sum t_k x_k) \leq \sum t_k f(x_k)$:

$$\left(\sum_{k=1}^{n} a_k b_k\right)^p \leq \sum_{k=1}^{n} t_k x_k^p$$

We examine the term on the right-hand side. Since $\frac{1}{p} + \frac{1}{q} = 1$, we have $q = \frac{p}{p-1}$, so $q - 1 = \frac{1}{p-1}$, which implies $p(q-1) = q$.

$$x_k^p = \left(a_k b_k^{-(q-1)} B\right)^p = a_k^p b_k^{-p(q-1)} B^p = a_k^p b_k^{-q} B^p$$

Multiplying by the weight t_k:

$$t_k x_k^p = \frac{b_k^q}{B} \cdot a_k^p b_k^{-q} B^p = a_k^p B^{p-1}$$

Summing over k:

$$\sum_{k=1}^{n} t_k x_k^p = B^{p-1} \sum_{k=1}^{n} a_k^p$$

Let $A = \sum a_k^p$. The inequality becomes:

$$\left(\sum_{k=1}^{n} a_k b_k\right)^p \leq A B^{p-1}$$

Taking the p-th root of both sides yields:

$$\sum_{k=1}^{n} a_k b_k \leq A^{1/p} B^{(p-1)/p}$$

Since $\frac{p-1}{p} = 1 - \frac{1}{p} = \frac{1}{q}$, we obtain $A^{1/p} B^{1/q}$, which is the desired result. ∎

***Corollary* 10.6.1.** *Cauchy-Schwarz Inequality.* Setting $p = q = 2$ yields the standard Cauchy-Schwarz inequality:

$$\sum_{k=1}^{n} |x_k y_k| \leq \left(\sum_{k=1}^{n} x_k^2\right)^{1/2} \left(\sum_{k=1}^{n} y_k^2\right)^{1/2}$$

The triangle inequality $|x + y| \leq |x| + |y|$ states that the length of a sum does not exceed the sum of the lengths. Minkowski's inequality establishes this property for the general p-norm, defined for a vector $\mathbf{x} = (x_1, \ldots, x_n) \in \mathbb{R}^n$ as:

$$\|\mathbf{x}\|_p = \left(\sum_{k=1}^{n} |x_k|^p\right)^{1/p}$$

***Theorem* 10.6.5. Minkowski's Inequality**. For any real number $p \geq 1$ and any $x_1, \ldots, x_n, y_1, \ldots, y_n \in \mathbb{R}$:

$$\left(\sum_{k=1}^{n} |x_k + y_k|^p\right)^{1/p} \leq \left(\sum_{k=1}^{n} |x_k|^p\right)^{1/p} + \left(\sum_{k=1}^{n} |y_k|^p\right)^{1/p}$$

Succinctly, $\|\mathbf{x} + \mathbf{y}\|_p \leq \|\mathbf{x}\|_p + \|\mathbf{y}\|_p$.

Proof. The case $p = 1$ is simply the scalar triangle inequality summed over k. Assume $p > 1$. Let $S = \sum |x_k + y_k|^p$. If $S = 0$, the result is trivial. We write $|x_k + y_k|^p = |x_k + y_k| \cdot |x_k + y_k|^{p-1} \leq (|x_k| + |y_k|)|x_k + y_k|^{p-1}$. Summing over k:

$$S \leq \sum_{k=1}^{n} |x_k||x_k + y_k|^{p-1} + \sum_{k=1}^{n} |y_k||x_k + y_k|^{p-1}$$

We apply Hölder's inequality to the first sum with conjugate exponents p and q. Note that $(p - 1)q = p$.

$$\sum |x_k||x_k + y_k|^{p-1} \leq \left(\sum |x_k|^p\right)^{1/p} \left(\sum |x_k + y_k|^{(p-1)q}\right)^{1/q} = \|\mathbf{x}\|_p \cdot S^{1/q}$$

Similarly for the second sum:

$$\sum |y_k||x_k + y_k|^{p-1} \leq \|\mathbf{y}\|_p \cdot S^{1/q}$$

Adding these bounds:

$$S \leq (\|\mathbf{x}\|_p + \|\mathbf{y}\|_p)S^{1/q}$$

Dividing by $S^{1/q}$ (assuming $S > 0$) and using $1 - 1/q = 1/p$:

$$S^{1/p} \leq \|\mathbf{x}\|_p + \|\mathbf{y}\|_p$$

■

10.7 Applications: Curve Sketching II

Differential calculus provides a systematic toolkit for visualising the behaviour of functions. Often, the explicit formula for $f(x)$ is secondary to the information contained in its derivatives.

Example 10.7.1. Reconstructing from the Derivative. Suppose a function f is continuous on $\mathbb{R}$ and its derivative is given by:

$$f'(x) = (x - 1)^2(x - 2)(x - 4)$$

We determine the shape of f without knowing $f(x)$ explicitly.

1. **Critical Points:** $x = 1, 2, 4$.
2. **Sign Analysis:**
    - On $(-\infty, 1)$: $(+)(-)(-) > 0 \implies f$ increasing.
    - On $(1, 2)$: $(+)(-)(-) > 0 \implies f$ increasing.

- On $(2,4)$: $(+)(+)(-) < 0 \implies f$ decreasing.
- On $(4,\infty)$: $(+)(+)(+) > 0 \implies f$ increasing.

3. **Classification:**
   - At $x = 1$: f' does not change sign (positive to positive). This is an inflection point, not an extremum. The tangent is horizontal, but the function continues to rise.
   - At $x = 2$: f' changes from positive to negative. Local Maximum.
   - At $x = 4$: f' changes from negative to positive. Local Minimum.
4. **Concavity:** We compute $f''(x)$ using the product rule on $u = (x-1)^2$ and $v = x^2 - 6x + 8$.

$$f''(x) = 2(x-1)(x^2 - 6x + 8) + (x-1)^2(2x-6)$$

Factor out $2(x-1)$:

$$f''(x) = 2(x-1)[(x^2 - 6x + 8) + (x-1)(x-3)] = 2(x-1)(2x^2 - 10x + 11)$$

Roots are $x = 1$ and the roots of $2x^2 - 10x + 11 = 0$, which are $\frac{10\pm\sqrt{12}}{4} = \frac{5\pm\sqrt{3}}{2}$. Thus we have three inflection points: 1, ≈ 1.63, and ≈ 3.37.

Example 10.7.2. An Asymptotically Flat Function. Consider $f(x) = \frac{(1+x)^2}{1+x^2}$.

1. **Asymptotics:** As $x \to \pm\infty$, $f(x) \approx x^2/x^2 \to 1$. Horizontal asymptote $y = 1$.
2. **Intersections:** $f(x) = 0 \implies x = -1$. $f(0) = 1$.
3. **Derivatives:**

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

Critical points at $x = \pm 1$.

- $x = -1$: $f'(-1) = 0$. Sign changes from $-$ to $+$. Local Minimum (value 0).
- $x = 1$: $f'(1) = 0$. Sign changes from $+$ to $-$. Local Maximum (value 2).

4. **Convexity:**

$$f''(x) = \frac{-4x(1+x^2)^2 - 2(1-x^2)\cdot 2(1+x^2)(2x)}{(1+x^2)^4} = \frac{4x(x^2-3)}{(1+x^2)^3}$$

Inflection points at $x = 0, \pm\sqrt{3}$. The curve starts at $y = 1$ at $-\infty$, dips to 0 at $x = -1$, rises to 2 at $x = 1$, and decays back to 1.

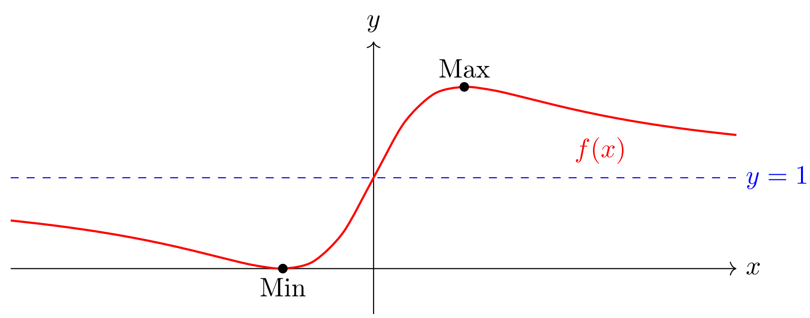


Figure 10.3: Graph of $f(x) = (1+x)^2/(1+x^2)$.

The graph is shown in Figure 10.3.

Systematic Analysis of Functions

By examining the first and second derivatives, domain constraints, and asymptotic limits, we can reconstruct the global shape of a function and identify its key features such as extrema, inflection points, and intervals of monotonicity, without relying on graphical plotting tools.

Example 10.7.3. Monotonicity without Extrema. Consider the function $f : \mathbb{R} \to \mathbb{R}$ defined by $f(x) = x + \sin x$. Differentiation yields:

$$f'(x) = 1 + \cos x$$

Stationary points occur where $\cos x = -1$, which implies $x_n = (2n+1)\pi$ for $n \in \mathbb{Z}$. Since $\cos x \geq -1$ for all real x, it follows that $f'(x) \geq 0$ everywhere. The derivative vanishes only at the discrete set of points $\{x_n\}$. As f' does not change sign (remaining non-negative on either side of x_n), these critical points are not extrema but rather points of inflection. Consequently, f is strictly increasing on $\mathbb{R}$ and possesses no local extrema.

Example 10.7.4. Non-Differentiable Critical Points. Let $f : [-1, 2] \to \mathbb{R}$ be defined by $f(x) = x^{4/3} - 4x^{1/3}$. We determine the intervals of monotonicity and the absolute extrema on this compact domain. Differentiation gives:

$$f'(x) = \frac{4}{3}x^{1/3} - \frac{4}{3}x^{-2/3} = \frac{4}{3}\left(\frac{x-1}{x^{2/3}}\right)$$

The critical points are classified as follows:

1. **Stationary Point:** $f'(x) = 0 \implies x = 1$.
2. **Singular Point:** $f'(x)$ is undefined at $x = 0$ (due to the denominator), although f is continuous there.

We examine the sign of f' on the sub-intervals:

- On $(-1, 0)$: The numerator $(x - 1)$ is negative, and $x^{2/3} > 0$. Thus $f'(x) < 0$ (Decreasing).
- On $(0, 1)$: The numerator is negative, denominator positive. Thus $f'(x) < 0$ (Decreasing).
- On $(1, 2)$: The numerator is positive, denominator positive. Thus $f'(x) > 0$ (Increasing).

At the singular point $x = 0$, the derivative does not change sign; the function decreases through the cusp. At $x = 1$, the sign changes from negative to positive, indicating a local minimum. **Absolute Extrema:**

- Endpoints: $f(-1) = 1 + 4 = 5$; $\quad f(2) = 2^{1/3}(2 - 4) = -2\sqrt[3]{2} \approx -2.52$.
- Critical Points: $f(0) = 0$; $\quad f(1) = 1 - 4 = -3$.

The absolute maximum is 5 (at $x = -1$) and the absolute minimum is -3 (at $x = 1$).

Example 10.7.5. Domain Restrictions. Analyse the function $f(x) = x^2\sqrt{5 - x}$. **Domain:** The radical requires $5 - x \geq 0$, so $\text{dom}(f) = (-\infty, 5]$. **Differentiation:** Using the Product Rule:

$$f'(x) = 2x\sqrt{5-x} - \frac{x^2}{2\sqrt{5-x}} = \frac{4x(5-x) - x^2}{2\sqrt{5-x}} = \frac{x(20-5x)}{2\sqrt{5-x}}$$

Critical points occur at $x = 0$ and $x = 4$. The derivative is undefined at the boundary $x = 5$ (vertical tangent). **Sign Analysis:**

- On $(-\infty, 0)$: $x < 0$ and $(20 - 5x) > 0 \implies f'(x) < 0$ (Decreasing).
- On $(0, 4)$: $x > 0$ and $(20 - 5x) > 0 \implies f'(x) > 0$ (Increasing).
- On $(4, 5)$: $x > 0$ and $(20 - 5x) < 0 \implies f'(x) < 0$ (Decreasing).

Extrema:

- $x = 0$: Local Minimum ($f(0) = 0$).
- $x = 4$: Local Maximum ($f(4) = 16\sqrt{1} = 16$).
- $x = 5$: Endpoint Minimum ($f(5) = 0$).

Example 10.7.6. Trigonometric Extrema. Examine $f(x) = 2\sin x + \cos(2x)$ for local extrema. The function is 2π-periodic. We compute the derivative:

$$f'(x) = 2\cos x - 2\sin(2x) = 2\cos x - 4\sin x\cos x = 2\cos x(1 - 2\sin x)$$

In the principal domain $[0, 2\pi)$, critical points occur where $\cos x = 0$ or $\sin x = 1/2$:

$$x \in \left\{\frac{\pi}{6}, \frac{\pi}{2}, \frac{5\pi}{6}, \frac{3\pi}{2}\right\}$$

We employ the Second Derivative Test: $f''(x) = -2\sin x - 4\cos(2x)$.

1. $x = \pi/6$: $f''(\pi/6) = -1 - 4(1/2) = -3 < 0 \implies$ Local Max ($y = 3/2$).
2. $x = \pi/2$: $f''(\pi/2) = -2 - 4(-1) = 2 > 0 \implies$ Local Min ($y = 1$).
3. $x = 5\pi/6$: $f''(5\pi/6) = -1 - 4(1/2) = -3 < 0 \implies$ Local Max ($y = 3/2$).
4. $x = 3\pi/2$: $f''(3\pi/2) = 2 - 4(-1) = 6 > 0 \implies$ Local Min ($y = -3$).

Thus, the global maxima are $3/2$ and the global minima are -3.

Example 10.7.7. Polynomial Extrema via Second Derivative. Find the local extrema of $f(x) = x^3 - 3x^2 - 24x + 5$.

$$f'(x) = 3x^2 - 6x - 24 = 3(x - 4)(x + 2)$$

Stationary points are $x = 4$ and $x = -2$.

$$f''(x) = 6x - 6$$

Evaluating the concavity:

- $f''(-2) = -18 < 0 \implies x = -2$ is a Local Maximum.
- $f''(4) = 18 > 0 \implies x = 4$ is a Local Minimum.

Example 10.7.8. Reconstructing from the Derivative. Suppose a function f is continuous on $\mathbb{R}$ with derivative $f'(x) = (x - 1)(x + 2)(x - 3)$. We determine the nature of the critical points without an explicit formula for f. The critical points are $\{-2, 1, 3\}$. We construct a sign chart:

Interval	Test Point	$f'(x)$ Sign	Behaviour
$(-\infty, -2)$	-3	$(-)(-)(-) = -$	Decreasing
$(-2, 1)$	0	$(-)(+)(-) = +$	Increasing
$(1, 3)$	2	$(+)(+)(-) = -$	Decreasing
$(3, \infty)$	4	$(+)(+)(+) = +$	Increasing

Classification:

- $x = -2$: Transition $- \to + \implies$ Local Minimum.
- $x = 1$: Transition $+ \to - \implies$ Local Maximum.
- $x = 3$: Transition $- \to + \implies$ Local Minimum.

Example 10.7.9. Optimisation on Open Intervals. Find the extrema of $f(x) = \frac{1}{1-x^2}$ on $(0, 1)$. **Boundary Behaviour:** $\lim_{x\to 0^+} f(x) = 1$ and $\lim_{x\to 1^-} f(x) = \infty$. **Derivative:**

$$f'(x) = \frac{2x}{(1 - x^2)^2}$$

For $x \in (0, 1)$, $f'(x) > 0$. The function is strictly increasing. Since the domain is open and monotonic, there are no local extrema. The value 1 is the infimum but is not attained.

One of the most practical applications of the derivative is the solution of optimisation problems where a physical or geometric quantity must be maximised or minimised subject to constraints.

Example 10.7.10. The Optimal Can. We wish to design a cylindrical can with a fixed volume V_0 such that the surface area (representing the material cost) is minimised. Let r be the base radius and h the height. The constraint is $V_0 = \pi r^2 h$, which implies $h = \frac{V_0}{\pi r^2}$. The surface area A includes the curved side and two circular caps:

$$A = 2\pi r h + 2\pi r^2$$

Substituting for h, we obtain A as a function of r on $(0, \infty)$:

$$A(r) = 2\pi r \left(\frac{V_0}{\pi r^2}\right) + 2\pi r^2 = \frac{2V_0}{r} + 2\pi r^2$$

Differentiating with respect to r:

$$A'(r) = -\frac{2V_0}{r^2} + 4\pi r$$

Setting $A'(r) = 0$ yields the critical dimension:

$$4\pi r = \frac{2V_0}{r^2} \implies r^3 = \frac{V_0}{2\pi} \implies r = \sqrt[3]{\frac{V_0}{2\pi}}$$

To verify minimality, we check the second derivative:

$$A''(r) = \frac{4V_0}{r^3} + 4\pi$$

Since $V_0, r > 0$, $A''(r) > 0$ everywhere. Thus $A(r)$ is strictly convex, and the unique critical point is the global minimum. **Physical Interpretation:** From the critical condition $2\pi r^3 = V_0$, substitute back into the expression for h:

$$h = \frac{V_0}{\pi r^2} = \frac{2\pi r^3}{\pi r^2} = 2r$$

Thus, the most efficient cylindrical can has a height equal to its diameter.

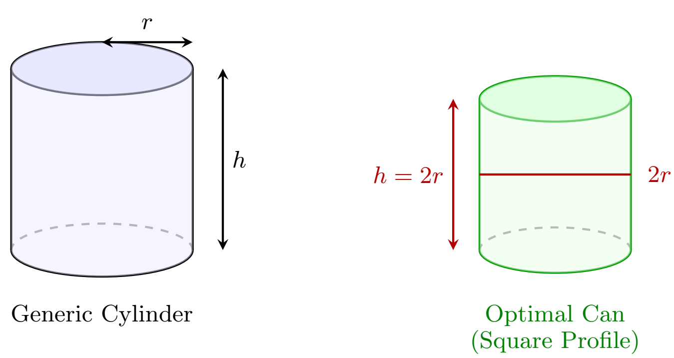


Figure 10.4: Geometric optimisation of a cylinder. The surface area is minimised when the profile is square ($h = 2r$).

10.8 Exercises

1. **Taylor Polynomial Calculation.** Compute the Taylor polynomial $T_n(x)$ centred at c for the following functions:
   (a) $f(x) = \sqrt{1+x}$ at $c = 0$ for $n = 3$.
   (b) $f(x) = \ln x$ at $c = 1$ for $n = 4$.
   (c) $f(x) = e^{x^2}$ at $c = 0$ for $n = 6$.
   (d) $f(x) = \tan x$ at $c = 0$ for $n = 5$.

2. **Limits via Taylor Expansions.** Evaluate the following limits by replacing the functions with their appropriate Taylor polynomials (Peano form). Do not use L'Hôpital's Rule.

(a) $\lim_{x\to 0} \frac{e^x-1-x-x^2/2}{x^3}$
(b) $\lim_{x\to 0} \frac{\ln(1-x)+\sin x+x^2/2}{x^3}$
(c) $\lim_{x\to 0} \frac{\sqrt{1+2x}-1-x+x^2/2}{x^3}$

3. **Numerical Estimation.** Use the Lagrange form of the remainder to estimate $\sqrt{e}$ with an error strictly less than 10^{-4}. How many terms of the Maclaurin expansion of e^x are required?

4. **Convexity Checks.** Determine the intervals on which the following functions are convex or concave.
   (a) $f(x) = x^4 - 4x^3 + 10$
   (b) $f(x) = xe^{-x}$
   (c) $f(x) = \arctan x$

5. **L'Hôpital's Practice.** Evaluate the following limits, verifying the hypotheses of L'Hôpital's Rule carefully.
   (a) $\lim_{x\to\infty} \left(1+\frac{1}{x}\right)^x$
   Remark. Consider the logarithm.
   (b) $\lim_{x\to 0} \left(\csc x - \frac{1}{x}\right)$
   (c) $\lim_{x\to 0^+} x(\ln x)^2$

6. **Uniqueness of the Taylor Polynomial.** Let $n \in \mathbb{N}$ and $x_0 \in \mathbb{R}$.
   (a) Let $P(x) = \sum_{k=0}^{n} c_k(x-x_0)^k$. Prove that $P^{(k)}(x_0) = k!c_k$ for $0 \le k \le n$.
   (b) Suppose $Q(x)$ is a polynomial of degree at most n such that $Q^{(k)}(x_0) = f^{(k)}(x_0)$ for all $0 \le k \le n$. Prove that $Q(x)$ must be the Taylor polynomial $T_n(x)$.
   Remark. Consider the difference $D(x) = T_n(x) - Q(x)$ and show all its derivatives at x_0 vanish.

7. **Contact of Order n.** Let f and g be n-times differentiable at c. We say that f and g have *contact of order n* at c if $f^{(k)}(c) = g^{(k)}(c)$ for all $k = 0, 1, \ldots, n$. Prove that f and g have contact of order n at c if and only if:
$$\lim_{x\to c} \frac{f(x)-g(x)}{(x-c)^n} = 0.$$

8. **Algebra of Convex Functions.** Let I be an interval. Prove the following properties of convex functions:
   (a) If f and g are convex on I and $\lambda > 0$, then $f+g$ and λf are convex.
   (b) If f is convex on I, then $g(x) = \exp(f(x))$ is convex on I.
   (c) If $\{f_\alpha\}_{\alpha\in A}$ is a family of convex functions on I such that $F(x) = \sup_{\alpha\in A} f_\alpha(x)$ is finite for all x, then $F(x)$ is convex.

9. **The Vanishing Derivatives.**
   (a) Prove that for $x \ne 0$, the n-th derivative has the form $f^{(n)}(x) = R_n(x)e^{-1/x^2}$, where $R_n(x)$ is a rational function.
   (b) Show that the Taylor series of f centred at 0 converges to the zero function everywhere, and thus only agrees with f at the origin.

10. **A L'Hôpital Counter-example.** Let $f(x) = x + \sin x\cos x$ and $g(x) = e^{\sin x}(x + \sin x \cos x)$.
   (a) Show that $\lim_{x\to\infty} f(x) = \lim_{x\to\infty} g(x) = \infty$.
   (b) Show that the limit of the ratio of derivatives $\lim_{x\to\infty} f'(x)/g'(x)$ does not exist.
   (c) Compute the actual limit $\lim_{x\to\infty} f(x)/g(x)$ algebraically.
   (d) Explain why L'Hôpital's Rule failed.

11. **Determinants and Wronskians.** Let $f_1, f_2, \ldots, f_n$ be $(n-1)$-times differentiable functions. The Wronskian $W(x)$ is the determinant of the matrix where the entry (i,j) is $f_j^{(i-1)}(x)$.

(a) Calculate the Wronskian of $\{1, x, x^2, \ldots, x^{n-1}\}$.
(b) Calculate the Wronskian of $\{e^{\lambda_1 x}, e^{\lambda_2 x}, \ldots, e^{\lambda_n x}\}$ where λ_i are distinct constants.

Remark. This connects the independence of functions to the non-vanishing of a geometric volume in function space.

12. Algebraic Computation of Taylor Coefficients. Suppose $a, b \in \mathbb{R}$ with $b \geq 0$. Consider $f(x) = \frac{1+ax^2}{1+bx^2}$. Find the Taylor polynomial of degree 4 for f at $x_0 = 0$. For which values of a, b does this polynomial coincide with the degree 4 Taylor polynomial of $\cos x$ at 0?

Remark. Avoid the Quotient Rule. Write $(1+bx^2)f(x) = 1+ax^2$ and differentiate this product using the Leibniz rule n times at $x = 0$.

13. Convergence of Newton's Method. Let f be twice differentiable on $[a, b]$, $f(a) < 0$, $f(b) > 0$, $f' \geq \delta > 0$, and $f'' \geq 0$. Let ξ be the unique root in (a, b). Newton's method generates a sequence $x_{n+1} = x_n - f(x_n)/f'(x_n)$.

(a) By using Taylor's theorem, show that $x_{n+1} - \xi = \frac{f''(\theta_n)}{2f'(x_n)}(x_n - \xi)^2$ for some θ_n between x_n and ξ.
(b) Deduce that if start with $x_0 > \xi$, the sequence decreases monotonically to ξ with quadratic convergence (the error is squared at each step).

14. Approximation of the Sine Inverse.

(a) Find the Taylor series for $(1 - x^2)^{-1/2}$ using the Binomial Theorem.
(b) By integrating term-by-term (assuming this is valid within the radius of convergence), derive the series for $\arcsin x$:

$$\arcsin x = x + \sum_{n=1}^{\infty} \frac{(2n)!}{4^n (n!)^2 (2n+1)} x^{2n+1}$$

(c) Use this to approximate π by evaluating at $x = 1/2$.

15. Young's Inequality via Convexity. Using the fact that $\ln x$ is concave on $(0, \infty)$, prove that if $p, q \in (1, \infty)$ satisfy $\frac{1}{p} + \frac{1}{q} = 1$, then for any $u, v > 0$:

$$uv \leq \frac{u^p}{p} + \frac{v^q}{q}.$$

Remark. Consider the log of both sides or apply Jensen's inequality directly to the function $-\ln x$ with appropriate weights.

16. Hermite Polynomials. Consider the Gaussian function $\gamma(x) = e^{-x^2/2}$.

(a) Prove that for any $n \in \mathbb{N}$, there exists a polynomial $H_n(x)$ of degree n such that $\gamma^{(n)}(x) = (-1)^n H_n(x)\gamma(x)$. These are the Hermite polynomials.
(b) Prove the recurrence relation $H_{n+1}(x) = xH_n(x) - H_n'(x)$.
(c) Compute H_1, H_2, H_3 explicitly.
(d) Determine the intervals of convexity for $\gamma(x)$.

17. Log-Concavity. Consider the function $f : (e, \infty) \to \mathbb{R}$ defined by $f(x) = \ln(\ln x)$.

(a) Calculate $f'(x)$ and $f''(x)$ to prove that f is concave.
(b) Use the definition of concavity to prove that for $x, y > e$:

$$\sqrt{\ln x \ln y} \leq \ln\left(\frac{x+y}{2}\right).$$

18. The Legendre Transform. Let $f : I \to \mathbb{R}$ be a strictly convex, differentiable function. We define its Legendre transform $f^*(t)$ by:

$$f^*(t) = \sup_{x \in I} (tx - f(x)).$$

(a) Show that for a given t, the supremum is attained at the unique x satisfying $f'(x) = t$. Let this point be $x(t)$. Thus $f^*(t) = tx(t) - f(x(t))$.

(b) **Involution.** Prove that $(f^*)^*(x) = f(x)$.
(c) **Example.** Compute the Legendre transform of $f(x) = \frac{1}{p}x^p$ for $p > 1$. Use this to recover Young's Inequality in the form $xy \le f(x) + f^*(y)$.

Remark. Geometric interpretation: f^* encodes the curve f not as a set of points, but as the envelope of its tangent lines.

19. Geometric Curvature. Let a curve be defined parametrically by $x(t), y(t)$.

(a) The curvature κ is defined as the rate of change of the angle of the tangent vector with respect to arc length. Show that for a graph $y = f(x)$, this is given by:

$$\kappa(x) = \frac{|f''(x)|}{(1 + (f'(x))^2)^{3/2}}.$$

(b) Compare the sign of $f''(x)$ with the direction of turning.
(c) Find the point on the parabola $y = x^2$ where the curvature is maximised.

20. Fixed Point Analysis. Let $f : [0, 1] \to [0, \infty)$ be a C^2 function satisfying: (i) $f' \ge 0$, (ii) $f'' > 0$, (iii) $f(0) > 0$, $f(1) = 1$, and $f'(1) > 1$.

(a) Prove that $f(x) \in [0, 1]$ for all $x \in [0, 1]$.
(b) Prove that there exists a unique fixed point $x^* \in (0, 1)$ such that $f(x^*) = x^*$.
(c) Consider the sequence $x_{n+1} = f(x_n)$ starting with $x_0 \in (0, 1)$. Prove that $\lim x_n = x^*$.

Remark. This models the extinction probability in a branching process.

21. Generalized Means. Let $\mathbf{x} = (x_1, \dots, x_n)$ with $x_i > 0$ and weights $\alpha_i > 0$ such that $\sum \alpha_i = 1$. The mean of order $t \ne 0$ is defined as $M_t(\mathbf{x}, \alpha) = \left(\sum \alpha_i x_i^t\right)^{1/t}$.

(a) Using L'Hôpital's Rule (or Taylor expansion), prove that $\lim_{t\to 0} M_t(\mathbf{x}, \alpha) = \prod x_i^{\alpha_i}$ (the geometric mean).
(b) Prove that $\lim_{t\to\infty} M_t(\mathbf{x}, \alpha) = \max\{x_1, \dots, x_n\}$.
(c) Prove that the function $t \mapsto M_t(\mathbf{x}, \alpha)$ is monotonically increasing.

22. The Error in Proportional Parts. Linear interpolation approximates $f(a + k)$ by $f(a) + \frac{k}{h}(f(a + h) - f(a))$ for $0 < k < h$. Using Taylor's Theorem, prove that the error in this approximation is bounded by $\frac{h^2}{8} \sup |f''|$, assuming f is twice differentiable.

23. Polynomial Expansion of Trigonometric Functions. Let m be a positive integer.

(a) Prove that $\sin((2m + 1)\theta)$ can be expressed as a polynomial in $\sin\theta$ of degree $2m + 1$.
(b) Let $y = \sin((2m + 1) \arcsin x)$. Show that $(1 - x^2)y'' - xy' + (2m + 1)^2 y = 0$.
(c) Use Leibniz's Rule to find the coefficients of the polynomial relating y and x.

24. The Asymptotic Limit. Find $\lim_{x\to\infty} x \left[\frac{1}{e} - \left(\frac{x}{x+1}\right)^x\right]$.

Remark. This requires expanding the base $(1 + 1/x)^{-x}$ carefully to second order.

25. Hyperbolic Taylor Series.

(a) Find the Maclaurin series for $\cosh x$ and $\sinh x$.
(b) Prove they converge to their respective functions for all $x \in \mathbb{R}$.
(c) Prove that for $x \ne 0$:

$$\cosh x > 1 + \frac{x^2}{2} \quad \text{and} \quad \frac{\sinh x}{x} > 1 + \frac{x^2}{6}.$$

Part III

Integration

Chapter 11

Introduction to Integration

We have analysed local behaviour via differentiation and linearisation. We now turn to the inverse problem of *quadrature*: assigning area to plane regions. In elementary geometry, the area of a rectangle is the product of its side lengths, and dissection extends this to polygons. Curvilinear regions require a limiting process.

11.1 The Axiomatic Theory of Area

Ideally, we would like to assign a non-negative real number, called area, to every subset of the plane $\mathbb{R}^2$. Let $\mathcal{P}(\mathbb{R}^2)$ denote the power set of $\mathbb{R}^2$. We seek a function $\alpha : \mathcal{P}(\mathbb{R}^2) \to [0, \infty]$ satisfying intuitive geometric properties.

No non-trivial area function exists on all subsets of $\mathbb{R}^2$ that respects rigid motions, a consequence of the Axiom of Choice. For our purposes, we restrict to a collection $\mathcal{M} \subset \mathcal{P}(\mathbb{R}^2)$ of measurable sets containing standard geometric regions.

We characterise the area function $\alpha : \mathcal{M} \to [0, \infty)$ by the following axioms:

***Axiom* 11.1.1. *Non-Negativity*.** For every $S \in \mathcal{M}$, $\alpha(S) \geq 0$.

***Axiom* 11.1.2. *Additivity*.** If $S, T \in \mathcal{M}$, then $S \cup T \in \mathcal{M}$ and $S \cap T \in \mathcal{M}$. Furthermore,

$$\alpha(S \cup T) = \alpha(S) + \alpha(T) - \alpha(S \cap T)$$

If S and T are disjoint (or share only a boundary of zero area), this simplifies to $\alpha(S \cup T) = \alpha(S) + \alpha(T)$.

***Axiom* 11.1.3. *Difference Property*.** If $S, T \in \mathcal{M}$ with $S \subseteq T$, then $T \setminus S \in \mathcal{M}$ and:

$$\alpha(T \setminus S) = \alpha(T) - \alpha(S)$$

This formalises the intuition of "cutting out" a shape.

***Axiom* 11.1.4. *Invariance (Congruence)*.** If a set S can be transformed into T via a rigid motion (translation, rotation, or reflection, formally an isometry), then $\alpha(S) = \alpha(T)$.

***Axiom* 11.1.5. *Normalisation*.** If R is a rectangle with side lengths l and w, then $R \in \mathcal{M}$ and $\alpha(R) = l \times w$.

The first four axioms describe cutting and pasting, but they do not determine the area of a circle. To handle curvilinear figures, we require a limiting process.

The Method of Exhaustion

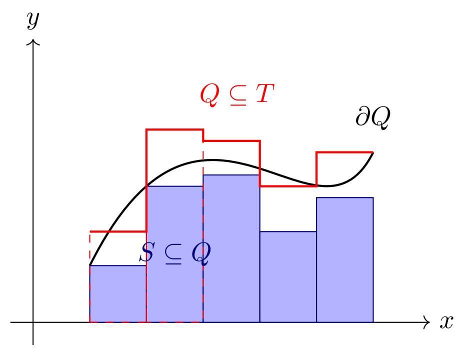
To assign an area to a more complex set Q, we approximate it from the inside and the outside using simple figures. We define a step region as a finite union of adjacent rectangles.

***Axiom* 11.1.6. *Exhaustion Property*.** Let $Q \subset \mathbb{R}^2$. Suppose that there exists a unique real number c such that for all step regions S and T with $S \subseteq Q \subseteq T$, we have:

$$\alpha(S) \leq c \leq \alpha(T)$$

Then Q is measurable and $\alpha(Q) = c$.

This axiom formalises the ancient Greek method of exhaustion. To find the area of a circle, one inscribes regular polygons (which can be decomposed into triangles and thus rectangles) and circumscribes them. As the number of sides increases, the area of the inner polygon increases and the outer decreases, squeezing the area of the circle between them.



The area of Q is squeezed between
the inner region S and outer region T.

Figure 11.1: Approximation of a region Q by step regions. The blue rectangles form the inner approximation S; the red outline suggests the outer approximation T.

The inner and outer approximations are shown in Figure 11.1.

11.2 Axiomatic Characterisation of the Integral

While the general theory of area applies to arbitrary shapes, calculus focuses on a specific type of region: the area under the graph of a function. This specialisation allows us to develop powerful computational tools.

Let $f : [a, b] \to \mathbb{R}$ be a continuous function. We wish to define the integral of f from a to b, denoted by $I_a^b(f)$, representing the signed area between the graph of f and the x-axis. Rather than constructing this via limits of sums (the Riemann sum approach, which we tackle later), we first characterise it by its essential properties derived from the area axioms.

***Theorem* 11.2.1.** Axiomatic Characterisation of the Integral Suppose that for every interval $[a, b] \subset \mathbb{R}$ and every continuous function $f : [a, b] \to \mathbb{R}$, we can assign a real number $I_a^b(f)$ satisfying the following two properties:

1. **Boundedness:** If there exist constants m, M such that $m \leq f(x) \leq M$ for all $x \in [a, b]$, then:

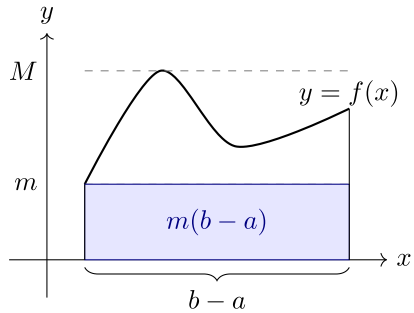
$$m(b - a) \leq I_a^b(f) \leq M(b - a)$$

2. **Additivity of Domain:** For any $c \in [a, b]$,

$$I_a^b(f) = I_a^c(f) + I_c^b(f)$$

Then the function $F(x) = I_a^x(f)$ is differentiable on (a, b) and $F'(x) = f(x)$.

This characterisation links the geometric axioms to antidifferentiation.



Property 1: The area under the curve is bounded
by the rectangles formed by the min and max values.

Figure 11.2: Geometric interpretation of the Boundedness Axiom.

The implication of Theorem 11.2.1 is the Fundamental Theorem of Calculus. See Figure 11.2.

Proof. Proof of Characterisation. Let $x \in (a, b)$ and let $h > 0$ be small enough such that $x + h \in (a, b)$. We wish to compute the derivative of the accumulation function $F(x) = I_a^x(f)$. Consider the Newton quotient:

$$\frac{F(x+h) - F(x)}{h}$$

Using the **Additivity** property, we can decompose the integral:

$$I_a^{x+h}(f) = I_a^x(f) + I_x^{x+h}(f)$$

Thus, the difference is the integral over the small strip $[x, x + h]$:

$$F(x+h) - F(x) = I_x^{x+h}(f)$$

Now we apply the **Boundedness** property. Since f is continuous on the compact interval $[x, x + h]$, by theorem 6.3.2 it attains a minimum value m_h and a maximum value M_h on this interval. That is, there exist $s_h, t_h \in [x, x + h]$ such that $f(s_h) = m_h$ and $f(t_h) = M_h$. The boundedness axiom implies:

$$f(s_h) \cdot h \le I_x^{x+h}(f) \le f(t_h) \cdot h$$

Dividing by h (which is positive):

$$f(s_h) \le \frac{F(x+h) - F(x)}{h} \le f(t_h)$$

We now let $h \to 0$. Since $s_h, t_h \in [x, x+h]$, by the Squeeze Theorem, $s_h \to x$ and $t_h \to x$. Crucially, because f is *continuous*, $\lim_{h\to 0} f(s_h) = f(x)$ and $\lim_{h\to 0} f(t_h) = f(x)$. Therefore, by the Squeeze Theorem again:

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

The argument for $h < 0$ is identical (taking care with signs). Thus $F'(x) = f(x)$. ■

Geometric Intuition

Why should the rate of change of area be the height of the curve? Consider the area $A(x)$ accumulated up to x. If we increase x by a tiny amount h, the area increases by a thin vertical strip. This strip is approximately a rectangle of width h and height $f(x)$.

$$A(x+h) \approx A(x) + f(x) \cdot h$$

$$\frac{A(x+h) - A(x)}{h} \approx f(x)$$

As $h \to 0$, the approximation becomes exact.

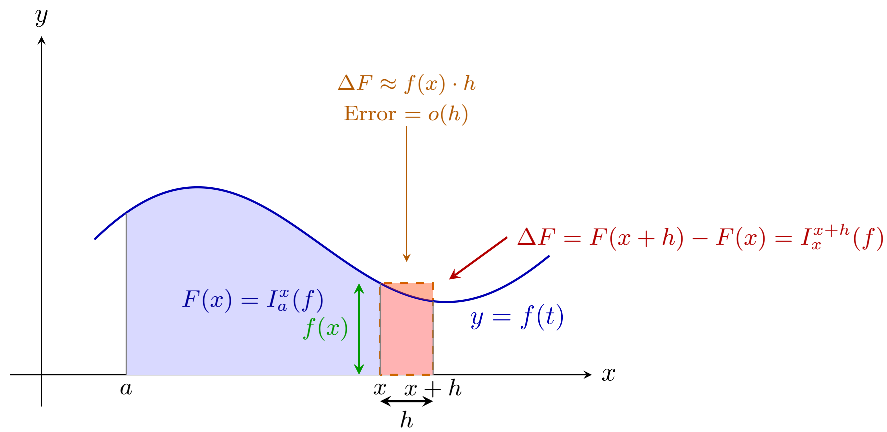


Figure 11.3: The Fundamental Theorem visualised. The change in area ΔF over a width h is approximately the rectangle area $f(x) \cdot h$. The error (the region between the curve and the top of the rectangle) vanishes faster than h as $h \to 0$.

The geometric interpretation is shown in Figure 11.3.

***Corollary* 11.2.1.** *Uniqueness of the Integral.* Suppose an association I exists satisfying the axioms of Theorem 11.2.1. Then I is uniquely determined. Specifically, if F is any antiderivative of f (i.e., $F' = f$), then:

$$I_a^b(f) = F(b) - F(a)$$

Proof. Let $G(x) = I_a^x(f)$. By the theorem, $G'(x) = f(x)$. If F is another primitive such that $F'(x) = f(x)$, then $(G - F)'(x) = 0$ for all x. By theorem 9.9.2, $G(x) - F(x) = C$ for some constant C. To find C, set $x = a$.

$$G(a) = I_a^a(f) = 0$$

(Note: $I_a^a = 0$ follows from $m(a - a) \leq I_a^a \leq M(a - a)$). Thus $0 - F(a) = C$, so $C = -F(a)$. Therefore, $G(x) = F(x) - F(a)$. Setting $x = b$ yields $I_a^b(f) = F(b) - F(a)$. ■

This corollary provides the standard computational method for integrals: find an antiderivative and evaluate the difference at the endpoints. However, we stress that we have not yet proven that such an assignment $I_a^b(f)$ actually *exists* for all continuous functions. We have only shown that *if* it exists, it must behave this way. The construction of this function (proving its existence), requires the machinery of Riemann Sums, to which we turn in the next chapter.

11.3 The Definition of the Riemann Integral

We now construct the integral formally. The approach we adopt is due to Darboux, a refinement of Riemann's original construction using sums. Our objective is to assign a precise meaning to the area under the graph of a bounded function $f : [a, b] \to \mathbb{R}$. Motivated by Exhaustion Property, we approximate the region from the inside and outside using collections of adjacent rectangles.

Definition 11.3.1. ***Partition***. A partition P of the closed interval $[a, b]$ is a finite set of points $P = \{x_0, x_1, \ldots, x_n\} \subset [a, b]$ such that:

$$a = x_0 < x_1 < x_2 < \cdots < x_n = b$$

We denote the length of the k-th subinterval $[x_{k-1}, x_k]$ by $\Delta x_k = x_k - x_{k-1}$. The mesh size of P is

$$\|P\| = \max_{1 \le k \le n} \Delta x_k.$$

Definition 11.3.2. ***Tagged Partition***. A tagged partition of $[a, b]$ is a pair (P, ξ), where $P = \{x_0, \ldots, x_n\}$ is a partition and $\xi = (\xi_1, \ldots, \xi_n)$ satisfies $\xi_k \in [x_{k-1}, x_k]$ for each k.

Definition 11.3.3. ***Riemann Sum***. Let $f : [a, b] \to \mathbb{R}$ be bounded and let (P, ξ) be a tagged partition. The Riemann sum of f associated to (P, ξ) is

$$S(f, P, \xi) = \sum_{k=1}^{n} f(\xi_k)\Delta x_k.$$

Example 11.3.1. Uniform Partition. For $n \in \mathbb{N}$, the uniform partition U_n of $[a, b]$ is given by

$$x_k = a + \frac{k}{n}(b - a), \quad k = 0, 1, \ldots, n.$$

Then $\|U_n\| = \frac{b-a}{n}$.

Given a bounded function f and a partition P, we define the bounds of f on each subinterval. Let:

$$m_k = \inf\{f(x) : x \in [x_{k-1}, x_k]\}$$

$$M_k = \sup\{f(x) : x \in [x_{k-1}, x_k]\}$$

Since f is bounded, these finite real numbers always exist.

Definition 11.3.4. ***Darboux Sums***. The Darboux lower sum $L(f, P)$ and the Darboux upper sum $U(f, P)$ of f with respect to the partition P are defined as:

$$L(f, P) = \sum_{k=1}^{n} m_k \Delta x_k \quad \text{and} \quad U(f, P) = \sum_{k=1}^{n} M_k \Delta x_k$$

Geometrically, $L(f, P)$ represents the area of the union of rectangles inscribed strictly *under* the graph, while $U(f, P)$ represents the area of rectangles circumscribing the graph.

***Lemma* 11.3.1.** *Bounds for Riemann Sums.* Let $f : [a, b] \to \mathbb{R}$ be bounded. For any partition P and any tag ξ,

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

Proof. For each k, we have $m_k \le f(\xi_k) \le M_k$. Multiplying by Δx_k and summing yields the inequalities. ∎

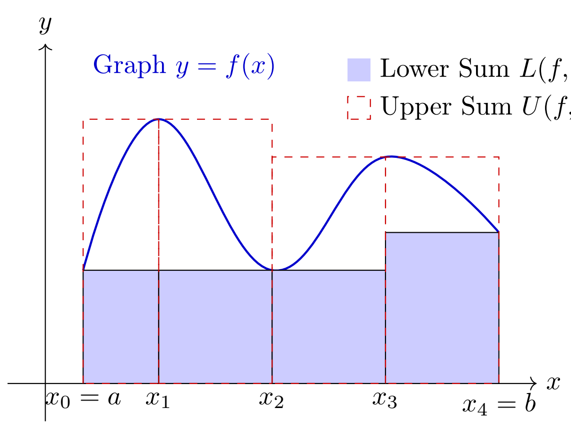


Figure 11.4: Visualisation of Darboux sums. The solid blue rectangles correspond to the lower sum (using m_k), and the dashed red rectangles extend to the upper sum (using M_k). The true area lies trapped between these two values.

The lower and upper rectangles are shown in Figure 11.4.

A First Computation

Let $f(x) = x^2$ on $[0, 1]$. For $N \in \mathbb{N}$, take the uniform partition U_N with $x_k = k/N$ and $\Delta x_k = 1/N$. Let $\xi_k = x_{k-1}$ (left tags) and $\eta_k = x_k$ (right tags). The corresponding Riemann sums are

$$L_N = S(f, U_N, \xi) = \frac{1}{N^3} \sum_{k=0}^{N-1} k^2, \qquad U_N = S(f, U_N, \eta) = \frac{1}{N^3} \sum_{k=1}^{N} k^2.$$

Hence $U_N - L_N = 1/N$. Using

$$\sum_{k=1}^{N} k^2 = \frac{N(N+1)(2N+1)}{6},$$

we obtain

$$\lim_{N\to\infty} L_N = \lim_{N\to\infty} U_N = \frac{1}{3}.$$

The geometry is shown in Figure 11.5.

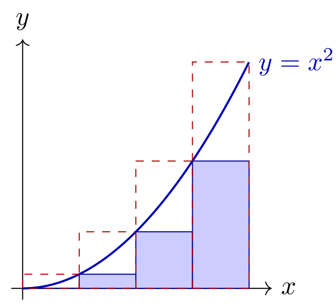


Figure 11.5: Left (filled) and right (dashed) rectangles for $y = x^2$ on $[0, 1]$.

Intuitively, adding more points to the partition should improve the approximation: the lower sums should increase (filling more gaps) and the upper sums should decrease (trimming the excess). We formalise this via the notion of refinement.

Definition 11.3.5. ***Refinement***. A partition Q is called a refinement of a partition P if $P \subseteq Q$. That is, Q contains all points of P and possibly additional points.

***Lemma* 11.3.2.** *Refinement Lemma.* Let P be a partition of $[a, b]$ and let Q be a refinement of P. Then:

$$L(f, P) \le L(f, Q) \quad \text{and} \quad U(f, Q) \le U(f, P)$$

Proof. It suffices to prove this for the case where Q contains exactly one more point than P, say $Q = P \cup \{z\}$. By induction, the result holds for any finite number of additional points. Suppose $P = \{x_0, \ldots, x_n\}$. Let the new point z fall in the k-th subinterval, i.e., $x_{k-1} < z < x_k$. The term $m_k \Delta x_k$ in $L(f, P)$ corresponds to the interval $[x_{k-1}, x_k]$. In $L(f, Q)$, this interval is split into two: $[x_{k-1}, z]$ and $[z, x_k]$. Let $m_k' = \inf\{f(x) : x \in [x_{k-1}, z]\}$ and $m_k'' = \inf\{f(x) : x \in [z, x_k]\}$. Since these are subsets of the original interval, the infimum can only increase (or stay the same):

$$m_k' \ge m_k \quad \text{and} \quad m_k'' \ge m_k$$

We rewrite the contribution to the sum:

$$\begin{aligned} m_k \Delta x_k &= m_k(x_k - x_{k-1}) \\ &= m_k((x_k - z) + (z - x_{k-1})) \\ &= m_k(z - x_{k-1}) + m_k(x_k - z) \end{aligned}$$

Comparing this with the contribution in Q:

$$m'_k(z - x_{k-1}) + m''_k(x_k - z)$$

Since $m'_k \geq m_k$ and $m''_k \geq m_k$, and lengths are positive, the contribution in Q is greater than or equal to the contribution in P. All other intervals remain unchanged. Thus $L(f, P) \leq L(f, Q)$. The proof for the upper sum is strictly analogous, noting that the supremum on a subinterval is less than or equal to the supremum on the whole interval. ■

This lemma yields a fundamental inequality: no matter how coarse or fine the partitions are, any lower sum is bounded by any upper sum.

***Proposition* 11.3.1.** *Fundamental Inequality.* Let P and Q be any two partitions of $[a, b]$. Then:

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

Proof. Consider the common refinement $R = P \cup Q$. Since R is a refinement of P, $L(f, P) \leq L(f, R)$. Since R is a refinement of Q, $U(f, R) \leq U(f, Q)$. Trivially, for any partition, the minimum is less than the maximum, so $L(f, R) \leq U(f, R)$. Combining these:

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

■

Since the set of lower sums is bounded above (by any upper sum), and the set of upper sums is bounded below, their supremum and infimum exist.

Definition 11.3.6. *Upper and Lower Integrals*. The lower Riemann integral of f on $[a, b]$, denoted $L(f)$ or $\underline{\int_a^b} f$, is the supremum of all lower sums:

$$L(f) = \sup_P L(f, P)$$

The upper Riemann integral of f on $[a, b]$, denoted $U(f)$ or $\overline{\int_a^b} f$, is the infimum of all upper sums:

$$U(f) = \inf_P U(f, P)$$

From the fundamental inequality, we always have $L(f) \leq U(f)$.

We are finally in a position to define the Riemann integral.

Definition 11.3.7. *Riemann Integrable*. A bounded function $f : [a, b] \to \mathbb{R}$ is said to be Riemann integrable if its upper and lower integrals coincide:

$$L(f) = U(f)$$

In this case, the common value is called the Riemann integral of f from a to b, denoted by:

$$\int_a^b f$$

Remark. A note on notation. You will frequently encounter $\int_a^b f(x)\,dx$. This introduces a dummy variable and a symbol dx that suggests a sum of $f(x)\Delta x$. Formally, integration acts on the function f itself, so we prefer $\int_a^b f$ for theory and use $\int_a^b f(x)\,dx$ when it aids computation. The symbol dx gains precise meaning in the theory of differential forms.

11.4 Criteria for Integrability

We have defined the Riemann integral via the equality of the lower and upper integrals. However, calculating the supremum over all possible partitions is impractical for checking integrability. We require a workable criterion, similar to how the Cauchy criterion allows us to verify the convergence of a sequence without knowing its limit.

***Theorem* 11.4.1. Riemann's Criterion**. A bounded function $f : [a, b] \to \mathbb{R}$ is Riemann integrable if and only if for every $\epsilon > 0$, there exists a partition P_ϵ of $[a, b]$ such that:

$$U(f, P_\epsilon) - L(f, P_\epsilon) < \epsilon$$

Proof. We prove this both directions.

($\implies$): Suppose f is Riemann integrable. Let $I = \int_a^b f$. By definition, $L(f) = U(f) = I$. Let $\epsilon > 0$. Since the upper integral is the infimum of the upper sums, there exists a partition P_1 such that:

$$U(f, P_1) < I + \frac{\epsilon}{2}$$

Similarly, since the lower integral is the supremum of the lower sums, there exists a partition P_2 such that:

$$L(f, P_2) > I - \frac{\epsilon}{2}$$

Let $P_\epsilon = P_1 \cup P_2$ be the common refinement. By the Refinement Lemma (Refinement Lemma), lower sums increase and upper sums decrease with refinement:

$$U(f, P_\epsilon) \leq U(f, P_1) < I + \frac{\epsilon}{2}$$

$$L(f, P_\epsilon) \geq L(f, P_2) > I - \frac{\epsilon}{2}$$

Subtracting these inequalities gives $U(f, P_\epsilon) - L(f, P_\epsilon) < \left(I + \frac{\epsilon}{2}\right) - \left(I - \frac{\epsilon}{2}\right) = \epsilon$.

($\impliedby$): Suppose the condition holds. For any partition P, we have $L(f, P) \leq L(f) \leq U(f) \leq U(f, P)$. Thus, for the specific partition P_ϵ:

$$0 \leq U(f) - L(f) \leq U(f, P_\epsilon) - L(f, P_\epsilon) < \epsilon$$

Since $U(f) - L(f)$ is a non-negative constant less than every positive ϵ, it must be zero. Thus f is integrable.

■

The utility of this criterion is apparent; it allows us to prove that continuous functions are integrable. We use uniform continuity, which follows from continuity on a compact set by theorem 6.3.3.

***Theorem* 11.4.2. Integrability of Continuous Functions**. If $f : [a, b] \to \mathbb{R}$ is continuous, then f is Riemann integrable.

Proof. By theorem 6.3.2, f is bounded. By theorem 6.3.3, f is uniformly continuous. Fix $\epsilon > 0$. By uniform continuity, there exists a $\delta > 0$ such that for any $x, y \in [a, b]$:

$$|x - y| < \delta \implies |f(x) - f(y)| < \frac{\epsilon}{b - a}$$

Choose a partition $P = \{x_0, \ldots, x_n\}$ such that the width of every subinterval satisfies $\Delta x_k < \delta$. On each subinterval $[x_{k-1}, x_k]$, f attains its maximum M_k and minimum m_k (by the Extreme Value Theorem). Let $t_k, s_k \in [x_{k-1}, x_k]$ such that $f(t_k) = M_k$ and $f(s_k) = m_k$. Since $|t_k - s_k| \leq \Delta x_k < \delta$, we have:

$$M_k - m_k = f(t_k) - f(s_k) < \frac{\epsilon}{b - a}$$

We now compute the difference between the upper and lower sums:

$$U(f,P) - L(f,P) = \sum_{k=1}^{n} (M_k - m_k)\Delta x_k < \sum_{k=1}^{n} \frac{\epsilon}{b-a}\Delta x_k$$

Factoring out the constant:

$$= \frac{\epsilon}{b-a}\sum_{k=1}^{n} \Delta x_k = \frac{\epsilon}{b-a}(b-a) = \epsilon$$

By theorem 11.4.1, f is integrable. ■

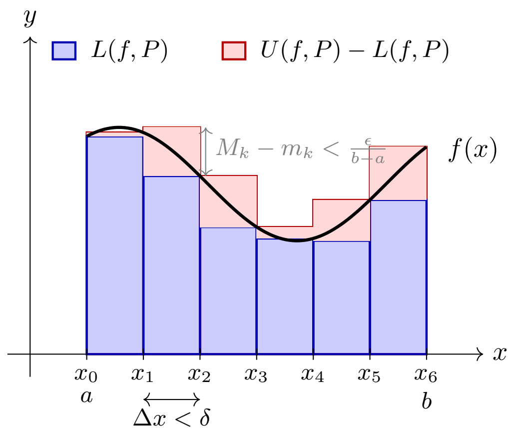


Figure 11.6: Visualising the proof for continuous functions. If the strips are narrow enough ($\Delta x < \delta$), the function cannot fluctuate wildly within a strip.

The partition control is illustrated in Figure 11.6.

Remark. This result can be extended to functions with a finite number of discontinuities, or indeed any function whose set of discontinuities has "measure zero" (Lebesgue's Criterion). However, continuity is sufficient for most elementary applications.

11.5 Properties of the Riemann Integral

Having established existence, we turn to the algebraic structure of the integral. The set of Riemann integrable functions on $[a,b]$, denoted $\mathcal{R}[a,b]$, forms a vector space, and the integral is a linear map on this space.

***Theorem* 11.5.1. Linearity of the Integral**. Let $f, g \in \mathcal{R}[a,b]$ and let $k \in \mathbb{R}$. Then:

1. $kf \in \mathcal{R}[a,b]$ and $\int_a^b kf = k\int_a^b f$.
2. $f+g \in \mathcal{R}[a,b]$ and $\int_a^b (f+g) = \int_a^b f + \int_a^b g$.

Proof. We prove the second property (additivity); the first (homogeneity) is left as an exercise. Let $P = \{x_0, \ldots, x_n\}$ be any partition. Consider a subinterval $[x_{i-1}, x_i]$. Recall the inequality for suprema and infima:

$$\inf_{x\in I} f(x) + \inf_{x\in I} g(x) \leq \inf_{x\in I}(f(x)+g(x))$$

$$\sup_{x\in I}(f(x)+g(x)) \leq \sup_{x\in I} f(x) + \sup_{x\in I} g(x)$$

This occurs because the extrema for f and g need not occur at the same point. Multiplying by Δx_i and summing yields the fundamental inequality for sums:

$$L(f,P) + L(g,P) \leq L(f+g,P) \leq U(f+g,P) \leq U(f,P) + U(g,P) \tag{11.1}$$

Step 1: Integrability. Let $\epsilon > 0$. Since f, g are integrable, there exist partitions P_f, P_g such that $U(f, P_f) - L(f, P_f) < \epsilon/2$ and $U(g, P_g) - L(g, P_g) < \epsilon/2$. Let $P = P_f \cup P_g$ be the common refinement. By the refinement lemma and eq. (11.1):

$$\begin{aligned} U(f+g, P) - L(f+g, P) &\leq [U(f, P) + U(g, P)] - [L(f, P) + L(g, P)] \\ &= [U(f, P) - L(f, P)] + [U(g, P) - L(g, P)] \\ &< \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon \end{aligned}$$

Thus $f + g$ is integrable.

Step 2: Value of the Integral. From eq. (11.1), we have $L(f, P) + L(g, P) \leq L(f + g, P)$. Taking the supremum over all partitions P:

$$\int_a^b f + \int_a^b g = \sup_P L(f, P) + \sup_P L(g, P) \leq \sup_P L(f + g, P) = \int_a^b (f + g)$$

Similarly, from $U(f + g, P) \leq U(f, P) + U(g, P)$, taking the infimum yields:

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

Since $\int_a^b (f + g)$ is sandwiched between the same value, equality holds.

■

***Proposition* 11.5.1.** *Positivity.* If $f \in \mathcal{R}[a, b]$ and $f(x) \geq 0$ for all $x \in [a, b]$, then

$$\int_a^b f \geq 0.$$

Proof. For any partition P, $m_k \geq 0$, so $L(f, P) \geq 0$. Hence $L(f) \geq 0$, and $\int_a^b f = L(f) \geq 0$. ■

***Corollary* 11.5.1.** *Monotonicity.* If $f, g \in \mathcal{R}[a, b]$ and $f(x) \leq g(x)$ for all $x \in [a, b]$, then

$$\int_a^b f \leq \int_a^b g.$$

Proof. The function $g - f$ is integrable by linearity and satisfies $g - f \geq 0$. Apply positivity. ■

***Corollary* 11.5.2.** *Bounds.* Let $f \in \mathcal{R}[a, b]$ and set $m = \inf_{x \in [a,b]} f(x)$ and $M = \sup_{x \in [a,b]} f(x)$. Then

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

Proof. The constant functions m and M are continuous and therefore integrable. The inequalities $m \leq f \leq M$ and monotonicity give the result. ■

Definition 11.5.1. *Average Value*. If $f \in \mathcal{R}[a, b]$, the average value of f on $[a, b]$ is

$$\text{Mean}(f) = \frac{1}{b - a} \int_a^b f.$$

***Theorem* 11.5.2. Integral Mean Value Theorem.** If $f : [a, b] \to \mathbb{R}$ is continuous, there exists $\xi \in [a, b]$ such that

$$f(\xi) = \text{Mean}(f) = \frac{1}{b - a} \int_a^b f.$$

Proof. By the Extreme Value Theorem, f attains a minimum m and a maximum M on $[a, b]$. By the previous corollary,

$$m \leq \text{Mean}(f) \leq M.$$

Since f is continuous, the Intermediate Value Theorem yields ξ with $f(\xi) = \text{Mean}(f)$. ■

11.6 Sets of Measure Zero

We have seen that continuous functions are Riemann integrable. The intuition behind this is that for a continuous function, the local oscillation $M_k - m_k$ on any subinterval can be made arbitrarily small by choosing a sufficiently fine partition. Consequently, the difference $U(f,P) - L(f,P) = \sum (M_k - m_k)\Delta x_k$ becomes negligible.

However, continuity is a sufficient condition, not a necessary one. Consider a function that is zero everywhere except at a finite number of points where it takes the value 1. Intuitively, the "area" under this graph is zero. The few points where the function jumps contribute terms where $M_k - m_k$ is large, but we can make the corresponding widths Δx_k tiny. This suggests that a function is integrable if its discontinuities are "sparse" or "small" in some sense.

To formalise this notion of "smallness", we introduce the concept of a set of measure zero. This concept is central to the Lebesgue theory of integration but provides the complete characterisation of Riemann integrability as well.

Definition 11.6.1. ***Set of Measure Zero***. A subset $S \subseteq \mathbb{R}$ is said to have measure zero if for every $\epsilon > 0$, there exists a countable collection of open intervals $\{(a_n, b_n)\}_{n=1}^{\infty}$ such that:

$$S \subseteq \bigcup_{n=1}^{\infty} (a_n, b_n) \quad \text{and} \quad \sum_{n=1}^{\infty} (b_n - a_n) < \epsilon$$

Example 11.6.1. Finite Sets. Any finite set $S = \{x_1, \dots, x_N\}$ has measure zero. Fix $\epsilon > 0$. For each x_k, consider the interval $I_k = (x_k - \frac{\epsilon}{2N}, x_k + \frac{\epsilon}{2N})$. Clearly $S \subset \bigcup I_k$. The total length is:

$$\sum_{k=1}^{N} \text{length}(I_k) = \sum_{k=1}^{N} \frac{\epsilon}{N} = \epsilon$$

Since ϵ was arbitrary, S has measure zero.

Example 11.6.2. Countable Sets. Any countable set $S = \{x_1, x_2, \dots\}$ has measure zero. Fix $\epsilon > 0$. We cannot use the ϵ/N trick since N is infinite. Instead, we use a geometric series to ensure the sum converges. Cover x_n with the interval $I_n = (x_n - \frac{\epsilon}{2^{n+1}}, x_n + \frac{\epsilon}{2^{n+1}})$. The length of I_n is $\frac{\epsilon}{2^n}$. The total length is:

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

Thus, the set of rational numbers $\mathbb{Q}$ has measure zero, despite being dense in $\mathbb{R}$.

***Proposition* 11.6.1.** *Properties of Measure Zero Sets.*

1. Any subset of a set of measure zero has measure zero.
2. A countable union of sets of measure zero has measure zero.

Proof. (1) is trivial. For (2), let $\{S_k\}_{k=1}^{\infty}$ be sets of measure zero. Fix $\epsilon > 0$. Since S_k has measure zero, it can be covered by a collection of intervals $\mathcal{O}_k$ with total length less than $\epsilon/2^k$. The union of all these collections $\bigcup \mathcal{O}_k$ covers $\bigcup S_k$. The total length is bounded by $\sum_{k=1}^{\infty} \frac{\epsilon}{2^k} = \epsilon$. ∎

Remark. Not all measure zero sets are countable. The Cantor set is an uncountable subset of $[0,1]$ constructed by removing the middle thirds. The total length removed is $1/3 + 2/9 + 4/27 + \cdots = 1$. Thus the remaining set has "measure" $1 - 1 = 0$, yet it has the same cardinality as $\mathbb{R}$.

11.7 The Riemann-Lebesgue Theorem

We now state the criterion for Riemann integrability in terms of oscillation.

Definition 11.7.1. ***Oscillation***. Let $f : [a, b] \to \mathbb{R}$ be bounded. The oscillation of f on an interval $I \subseteq [a, b]$ is $\text{osc}(f, I) = \sup_{x \in I} f(x) - \inf_{x \in I} f(x)$. The oscillation of f at a point x is defined as:

$$\text{osc}_f(x) = \lim_{\delta \to 0^+} \text{osc}(f, (x - \delta, x + \delta) \cap [a, b])$$

Note that f is continuous at x if and only if $\text{osc}_f(x) = 0$.

Theorem **11.7.1. Riemann-Lebesgue Theorem**. Let $f : [a, b] \to \mathbb{R}$ be a bounded function. Then f is Riemann integrable if and only if the set of discontinuities of f has measure zero.

The difference $U(f, P) - L(f, P)$ is a sum of oscillations $(M_k - m_k)\Delta x_k$. Continuity makes these oscillations small, and discontinuities can be confined to intervals of arbitrarily small total length when the discontinuity set has measure zero.

Proof. Let D be the set of discontinuities of f. We define the sets:

$$D_k = \left\{ x \in [a, b] : \text{osc}_f(x) \geq \frac{1}{k} \right\}$$

Then $D = \bigcup_{k=1}^{\infty} D_k$. By proposition 11.6.1, D has measure zero if and only if each D_k has measure zero.

Step 1: Integrable $\implies$ Measure Zero. Suppose f is integrable. Fix $k \in \mathbb{N}$ and $\epsilon > 0$. We wish to show D_k can be covered by intervals of small total length. Since f is integrable, there exists a partition P such that:

$$U(f, P) - L(f, P) < \frac{\epsilon}{k}$$

Let the sum be split into "bad" subintervals (those containing points of D_k in their interior) and "good" ones. Let $\mathcal{I}_{bad}$ be the set of indices j such that the open interval (x_{j-1}, x_j) contains a point of D_k. If $(x_{j-1}, x_j) \cap D_k \neq \emptyset$, then the oscillation over that interval is at least $1/k$. Thus $M_j - m_j \geq 1/k$.

$$\frac{\epsilon}{k} > \sum_{j=1}^{n} (M_j - m_j)\Delta x_j \geq \sum_{j \in \mathcal{I}_{bad}} (M_j - m_j)\Delta x_j \geq \sum_{j \in \mathcal{I}_{bad}} \frac{1}{k} \Delta x_j$$

Cancelling $1/k$, we get $\sum_{j \in \mathcal{I}_{bad}} \Delta x_j < \epsilon$. The intervals indexed by $\mathcal{I}_{bad}$ cover all of D_k except possibly the finite set of endpoints of the partition P. Since a finite set has measure zero, D_k is covered by a union of intervals with total length less than ϵ. Thus D_k has measure zero, implying D has measure zero.

Step 2: Measure Zero $\implies$ Integrable. Suppose D has measure zero. Let $M = \sup |f(x)|$. If $M = 0$, then f is identically zero and thus integrable. Assume $M > 0$. Fix $\epsilon > 0$. Choose k large enough such that $\frac{1}{k} < \frac{\epsilon}{2(b-a)}$. Since $D_k \subseteq D$ has measure zero, there exists a countable open cover $\mathcal{O}$ of D_k with total length less than $\frac{\epsilon}{4M}$. For every $x \notin D_k$, we have $\text{osc}_f(x) < 1/k$. By definition of oscillation at a point, there exists an open interval I_x containing x such that $\text{osc}(f, I_x) < 1/k$. Consider the collection $\mathcal{C} = \mathcal{O} \cup \{I_x : x \in [a, b] \setminus D_k\}$. This is an open cover of the compact set $[a, b]$. By lemma 6.2.1, there exists a $\lambda > 0$ such that any interval of length less than λ is contained entirely within some set in $\mathcal{C}$. Let P be a partition with mesh size less than λ. We split the sum $U(f, P) - L(f, P)$ into two groups:

1. **Type 1:** Intervals $[x_{j-1}, x_j]$ contained in some $O \in \mathcal{O}$ (covering D_k).
2. **Type 2:** Intervals $[x_{j-1}, x_j]$ contained in some I_x (where oscillation is small).

For Type 1 intervals, $M_j - m_j \leq 2M$. The total length of these intervals is bounded by the total length of $\mathcal{O}$, which is $\frac{\epsilon}{4M}$.

$$\sum_{j \in \text{Type 1}} (M_j - m_j)\Delta x_j \leq 2M \sum \Delta x_j < 2M \cdot \frac{\epsilon}{4M} = \frac{\epsilon}{2}$$

For Type 2 intervals, the oscillation on the interval is less than $1/k$. The total length is at most $(b - a)$.

$$\sum_{j \in \text{Type 2}} (M_j - m_j)\Delta x_j < \frac{1}{k} \sum \Delta x_j \leq \frac{1}{k}(b - a) < \frac{\epsilon}{2(b - a)}(b - a) = \frac{\epsilon}{2}$$

Combining the sums:

$$U(f,P) - L(f,P) < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon$$

Thus f is Riemann integrable.

■

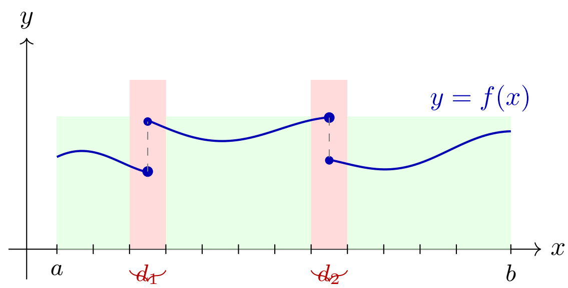


Figure 11.7: Visualisation of the Riemann-Lebesgue Theorem proof. The total width of the red regions is made small to control the large jumps. In the green regions, the oscillation is inherently small.

The partition types are shown in Figure 11.7.

Example 11.7.1. The Dirichlet Function. Consider the function $f : [0,1] \to \mathbb{R}$ defined by:

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

This function is discontinuous at every point in $[0,1]$. The set of discontinuities is $[0,1]$, which has length 1 (not measure zero). Therefore, f is not Riemann integrable. Indeed, for any partition P, every subinterval contains both rationals and irrationals. Thus $M_k = 1$ and $m_k = 0$ for all k.

$$U(f,P) = \sum 1 \cdot \Delta x_k = 1, \quad L(f,P) = \sum 0 \cdot \Delta x_k = 0$$

Since $U(f) \neq L(f)$, it is not integrable.

Example 11.7.2. Thomae's Function. Consider $f : [0,1] \to \mathbb{R}$:

$$f(x) = \begin{cases} 1/q & x = p/q \in \mathbb{Q} \text{ (in lowest terms)} \\ 0 & x \notin \mathbb{Q} \end{cases}$$

This function is continuous at every irrational number and discontinuous at every rational number. The set of discontinuities is $\mathbb{Q} \cap [0,1]$, which is countable and thus has measure zero. By the Riemann-Lebesgue Theorem, Thomae's function is Riemann integrable (and the integral is 0).

11.8 Consequences of the Riemann-Lebesgue Theorem

The investment of effort into establishing the Riemann-Lebesgue Theorem pays dividends. Many properties of the integral, which otherwise require ϵ-δ management with partitions, now follow as set-theoretic consequences.

***Corollary* 11.8.1.** *Integrability of Almost Continuous Functions.* Let $f : [a,b] \to \mathbb{R}$ be a bounded function.

1. If f has finitely many discontinuities, it is Riemann integrable.
2. If f has countably many discontinuities, it is Riemann integrable.

Proof. Finite and countable sets have measure zero. The result follows directly from Theorem 11.7.1. ■

***Corollary* 11.8.2.** *Modification at Finitely Many Points.* Let $f, g : [a, b] \to \mathbb{R}$ be bounded and suppose $f(x) = g(x)$ for all $x \in [a, b]$ except for a finite set. If f is Riemann integrable, then so is g, and

$$\int_a^b f = \int_a^b g.$$

Proof. Set $h = g - f$. Then h is zero except at finitely many points, so h and $|h|$ are Riemann integrable by the previous corollary. Let $M = \sup |h|$. Given $\epsilon > 0$, cover the finitely many points by intervals of total length less than ϵ/M. For a partition refining these endpoints, any Riemann sum of $|h|$ is bounded by M times that total length, hence $\int_a^b |h| < \epsilon$ by monotonicity. Thus $\int_a^b h = 0$, and linearity gives the claim. ■

Algebra of Integrable Functions

We have already established the linearity of the integral. The Riemann-Lebesgue theorem provides a swift proof that the product of integrable functions is integrable, a result that is surprisingly tedious to prove using Darboux sums directly due to sign management.

***Corollary* 11.8.3.** *Product Rule.* Let $f, g \in \mathcal{R}[a, b]$. Then the product $fg \in \mathcal{R}[a, b]$.

Proof. Since f and g are Riemann integrable, they are bounded. Thus fg is bounded. Let D_f, D_g, and D_{fg} denote the sets of discontinuities of f, g, and fg respectively. If f and g are continuous at c, then their product is continuous at c. By contrapositive, if fg is discontinuous at c, then at least one of f or g must be discontinuous there. Thus:

$$D_{fg} \subseteq D_f \cup D_g$$

Since $f, g \in \mathcal{R}[a, b]$, D_f and D_g are sets of measure zero. Their union is a set of measure zero, and the subset D_{fg} is therefore a set of measure zero. By Theorem 11.7.1, fg is integrable. ■

Note. The inclusion $D_{fg} \subseteq D_f \cup D_g$ is strict in general. Consider $f(x) = \operatorname{sgn}(x)$ and $g(x) = \operatorname{sgn}(x)$ on $[-1, 1]$. Both are discontinuous at 0, but the product $f(x)g(x) = 1$ (for $x \neq 0$) extends to a continuous function.

***Corollary* 11.8.4.** *Composition.* Let $f : [a, b] \to [c, d]$ be Riemann integrable and let $\phi : [c, d] \to \mathbb{R}$ be continuous. Then the composition $\phi \circ f \in \mathcal{R}[a, b]$.

Proof. Let $h = \phi \circ f$. Since ϕ is continuous, it maps continuous points to continuous points. Specifically, if f is continuous at x_0, then $\lim_{x \to x_0} f(x) = f(x_0)$. By continuity of ϕ, $\lim_{y \to f(x_0)} \phi(y) = \phi(f(x_0))$. Thus the composition is continuous at x_0. Consequently, $D_{\phi \circ f} \subseteq D_f$. Since D_f has measure zero, so does $D_{\phi \circ f}$. ■

Remark. The reverse composition $f \circ \phi$ need not be integrable. This is a subtle point often missed.

***Corollary* 11.8.5.** *Absolute Integrability.* If $f \in \mathcal{R}[a, b]$, then $|f| \in \mathcal{R}[a, b]$ and:

$$\left| \int_a^b f \right| \leq \int_a^b |f|$$

Proof. Integrability of $|f|$ follows from the previous corollary using the continuous function $\phi(t) = |t|$. The inequality follows from $-|f| \leq f \leq |f|$ and the order-preserving property of the integral. ■

Domain Additivity

We return to the axiomatic foundation laid out at the start of this chapter. We required the integral to satisfy $I_a^b = I_a^c + I_c^b$. We can now prove that the Riemann integral satisfies this property.

Definition 11.8.1. *Indicator Function*. For a subset $S \subseteq [a,b]$, the indicator function $\chi_S : [a,b] \to \mathbb{R}$ is defined by:

$$\chi_S(x) = \begin{cases} 1 & x \in S \\ 0 & x \notin S \end{cases}$$

For an interval $[c,d] \subset [a,b]$, the function $\chi_{[c,d]}$ is discontinuous only at the boundary points $\{c,d\}$ (if they are interior to $[a,b]$). Since finite sets have measure zero, $\chi_{[c,d]}$ is Riemann integrable.

***Theorem* 11.8.1. Additivity over Intervals.** Let $f \in \mathcal{R}[a,b]$ and let $c \in (a,b)$. Then the restrictions $f|_{[a,c]}$ and $f|_{[c,b]}$ are integrable, and:

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

Proof. **Integrability:** The set of discontinuities of the restriction $f|_{[a,c]}$ is $D_f \cap [a,c]$, which is a subset of a measure zero set. Thus the restrictions are integrable.

Equality: We decompose f using indicator functions. Note that for all x:

$$f(x) = f(x)\chi_{[a,c]}(x) + f(x)\chi_{[c,b]}(x) - f(c)\chi_{\{c\}}(x)$$

The term $f(c)\chi_{\{c\}}(x)$ is non-zero only at a single point, so its integral is 0. We may ignore it without affecting the integral values. By corollary 11.8.3, $f\chi_{[a,c]}$ and $f\chi_{[c,b]}$ are integrable on $[a,b]$. By linearity:

$$\int_a^b f = \int_a^b (f\chi_{[a,c]}) + \int_a^b (f\chi_{[c,b]})$$

Consider the first term. The function $g(x) = f(x)\chi_{[a,c]}(x)$ agrees with f on $[a,c]$ and is 0 on $(c,b]$. It is straightforward to verify from the definition of lower/upper sums that adding a zero tail does not change the integral accumulated on the initial segment (modulo the boundary point which contributes nothing). Thus:

$$\int_a^b (f\chi_{[a,c]}) = \int_a^c f$$

Similarly, $\int_a^b (f\chi_{[c,b]}) = \int_c^b f$. The result follows. ∎

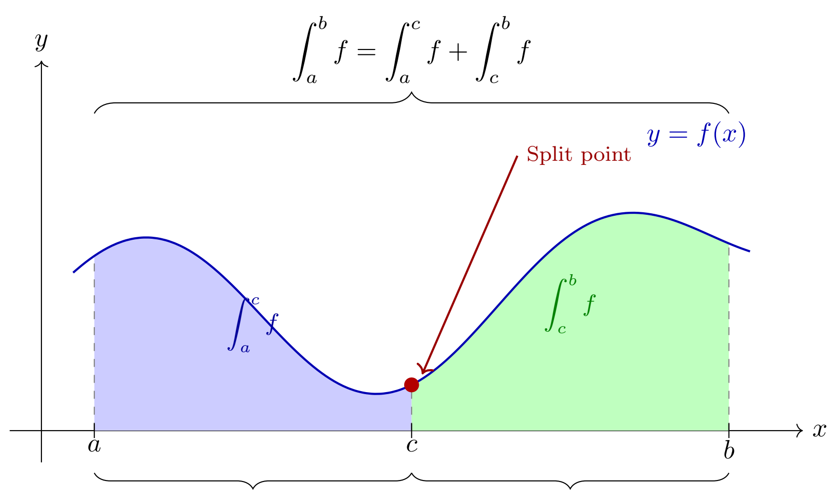


Figure 11.8: Visualisation of the Additivity property.

Remark. (Reversal of Limits). For any $f \in \mathcal{R}[a,b]$, define

$$\int_b^a f = -\int_a^b f, \qquad \int_a^a f = 0.$$

With this convention, the additivity identity

$$\int_c^a f = \int_b^a f + \int_c^b f$$

holds for any $a, b, c \in \mathbb{R}$.

11.9 Closing the Circle: The Fundamental Theorem

We have come full circle.

1. We proposed that an "area function" (integral) should satisfy **Boundedness** and **Additivity**.
2. We showed in Theorem 11.2.1 that *any* such functional must satisfy the Fundamental Theorem of Calculus: $\frac{d}{dx}\int_a^x f = f(x)$.
3. We constructed the Riemann integral $\int_a^b f$ via Darboux sums and proved via the Riemann-Lebesgue theorem that it exists for all continuous functions.
4. We verified above that this constructed integral satisfies Boundedness (from $m(b-a) \leq L(f) \leq U(f) \leq M(b-a)$) and Additivity.

Therefore, the Riemann integral provides the unique solution to the area problem for continuous functions, and it is computed via antidifferentiation.

***Theorem* 11.9.1. Fundamental Theorem of Calculus (First Form).** Let $f \in \mathcal{R}[a,b]$. Define $F(x) = \int_a^x f(t)dt$. Then F is continuous on $[a,b]$. Furthermore, if f is continuous at $c \in [a,b]$, then F is differentiable at c and $F'(c) = f(c)$.

***Theorem* 11.9.2. Fundamental Theorem of Calculus (Second Form).** Let $f \in \mathcal{R}[a,b]$. Suppose there exists a differentiable function G on $[a,b]$ such that $G' = f$. Then:

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

The proof of the First Form is identical to the proof of the Axiomatic Characterisation of the Integral. The Second Form follows from the First if f is continuous; if f is merely integrable, it requires a slightly more delicate argument using the Mean Value Theorem on the difference $G(x_k) - G(x_{k-1})$ in the Riemann sums, but the result holds.

This machinery allows us to evaluate integrals not by infinite summation, but by reversing the differentiation rules derived in the previous chapter.

11.10 Exercises

1. **Explicit Calculation of Riemann Sums.** Let $f(x) = x^2$ on the interval $[0,a]$ where $a > 0$.
   (a) Consider the uniform partition P_n dividing $[0,a]$ into n subintervals of equal width. Calculate the upper sum $U(f, P_n)$ and lower sum $L(f, P_n)$ explicitly using the formula $\sum_{k=1}^n k^2 = \frac{n(n+1)(2n+1)}{6}$.
   (b) Show that $\lim_{n\to\infty}(U(f, P_n) - L(f, P_n)) = 0$.
   (c) Deduce the value of $\int_0^a x^2\,dx$.

2. **Fermat's Method of Integration.** Before Newton and Leibniz, Fermat evaluated $\int_a^b x^p\,dx$ (for $p \neq -1$) using a partition in geometric progression. Let $f(x) = x^p$ on $[0,1]$ for $p > 0$.
   (a) Let $0 < r < 1$. Consider the partition $P = \{\ldots, r^n, \ldots, r^2, r, 1\}$.
   (b) Form the upper sum using the right endpoints. The sum will be a geometric series.
   (c) Evaluate the limit as $r \to 1^-$ to find the area.

3. **Step Functions and Partitions.** Consider the function $h : [0,1] \to \mathbb{R}$ defined by:

$$h(x) = \begin{cases} 1 & 0 \le x < 1 \\ 2 & x = 1 \end{cases}$$

(a) Show that for any partition P of $[0,1]$, the lower sum $L(h,P) = 1$.
(b) Construct a specific partition P such that $U(h,P) < 1 + 10^{-6}$.
(c) Prove using Riemann's Criterion that h is integrable and find its value.

4. **Sequential Criterion for Integrability.** Prove that a bounded function f is Riemann integrable on $[a,b]$ if and only if there exists a sequence of partitions $(P_n)_{n=1}^{\infty}$ such that

$$\lim_{n\to\infty} [U(f,P_n) - L(f,P_n)] = 0.$$

In this case, show that $\int_a^b f = \lim_{n\to\infty} U(f,P_n)$.

5. **Integrability of Monotone Functions.** Let $f : [a,b] \to \mathbb{R}$ be a monotone increasing function.

(a) For a uniform partition P_n of size n, show that:

$$U(f,P_n) - L(f,P_n) = \frac{b-a}{n}(f(b) - f(a)).$$

(b) Deduce that all bounded monotone functions are Riemann integrable, regardless of continuity.

6. **Riemann's Original Definition.** A tagged partition (P,ξ) samples f at points $\xi_k \in [x_{k-1}, x_k]$. The Riemann sum is $S(f,P,\xi) = \sum f(\xi_k)\Delta x_k$. Show that if f is continuous on $[a,b]$, then for every $\epsilon > 0$, there exists $\delta > 0$ such that for any tagged partition P with mesh $\|P\| < \delta$:

$$\left| S(f,P,\xi) - \int_a^b f \right| < \epsilon.$$

7. **The Dirichlet Function.** Let $D(x) = 1$ if $x \in \mathbb{Q} \cap [0,1]$ and 0 otherwise.

(a) Explain why for any partition P of $[0,1]$, $U(D,P) = 1$ and $L(D,P) = 0$.
(b) Conclude that D is not Riemann integrable.
(c) What is the set of discontinuities of D? Does it have measure zero?

8. **Thomae's Function (The Popcorn Function).** Let $t : [0,1] \to \mathbb{R}$ be defined by $t(0) = 1$, and $t(x) = 1/n$ if $x = m/n \in \mathbb{Q}$ is in lowest terms ($n > 0$), and $t(x) = 0$ if $x \notin \mathbb{Q}$.

(a) Use the Riemann-Lebesgue Theorem to conclude t is integrable and evaluate $\int_0^1 t$.

(b) Without the Riemann-Lebesgue theorem, let $\epsilon > 0$. Consider the set $S_\epsilon = \{x : t(x) \ge \epsilon/2\}$. Show S_ϵ is finite. Construct a partition around these points to prove integrability directly.

9. **The Product of Integrable Functions.** Let f, g be integrable on $[a,b]$.

(a) Show that if f is integrable, so is f^2.

(b) Use the identity $fg = \frac{1}{4}((f+g)^2 - (f-g)^2)$ to prove fg is integrable.

10. **Uniform Convergence and Integration.** Let (f_n) be a sequence of integrable functions on $[a,b]$ converging uniformly to f.

(a) Prove that f is integrable.

(b) Prove that $\lim_{n\to\infty} \int_a^b f_n = \int_a^b f$.

(c) Give a counterexample to show that pointwise convergence is insufficient for this conclusion.

11. **Strict Positivity.**

(a) Suppose f is Riemann integrable on $[a, b]$, $f(x) \geq 0$ for all x, and $\int_a^b f = 0$. Must $f(x) = 0$ for all x?

(b) Suppose f is continuous on $[a, b]$, $f(x) \geq 0$, and $\int_a^b f = 0$. Prove that $f(x) = 0$ for all $x \in [a, b]$.

12. Integral Mean Value Theorems.

(a) (First MVT) If f is continuous on $[a, b]$, prove there exists $c \in [a, b]$ such that $\int_a^b f(x)\,dx = f(c)(b - a)$.

(b) (Weighted MVT) If f is continuous and g is integrable with $g(x) \geq 0$ on $[a, b]$, prove there exists $c \in [a, b]$ such that:

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

13. Composition of Functions (Edge Case). Let $f : [a, b] \to [c, d]$ and $g : [c, d] \to \mathbb{R}$ be Riemann integrable.

(a) Construct a counterexample where $g \circ f$ is not integrable.
(b) Prove that if g is continuous, then $g \circ f$ is integrable.
(c) Determine the validity of: If f is continuous and g is integrable, then $g \circ f$ is integrable.

14. Calculus on Fractals. Let $f : [0, 1] \to \mathbb{R}$ be the indicator function of the fat Cantor set (measure > 0, no intervals).

(a) Is f Riemann integrable? Justify your answer using the Riemann-Lebesgue Theorem.
(b) If we construct a function $g(x) = \int_0^x f(t)dt$, is g differentiable? Where?

15. Hölder Continuity and Error Bounds. A function $f : [a, b] \to \mathbb{R}$ is α-Hölder continuous ($0 < \alpha \leq 1$) if there exists $C > 0$ such that $|f(x) - f(y)| \leq C|x - y|^\alpha$ for all x, y.

(a) Show that such a function is Riemann integrable.
(b) Let P_n be a uniform partition with n intervals. Prove the error bound:

$$\left| \int_a^b f(x)\,dx - S(f, P_n, \xi) \right| \leq \frac{C(b - a)^{\alpha+1}}{n^\alpha}.$$

16. Monotonicity of Area. Using only the axioms of Area:

(a) If $S, T \in \mathcal{M}$ and $S \subseteq T$, then $\alpha(S) \leq \alpha(T)$.
(b) If $S_1, S_2, \ldots, S_n$ are disjoint sets in $\mathcal{M}$, then $\alpha(\bigcup_{i=1}^n S_i) = \sum_{i=1}^n \alpha(S_i)$.
(c) Using the Exhaustion Property, prove that a single point and a line segment have area zero.

17. Scaling Invariance. Let $\lambda > 0$. For a set $S \in \mathcal{M}$, define $\lambda S = \{(\lambda x, \lambda y) : (x, y) \in S\}$. Prove from the Normalisation Axiom and Exhaustion that $\alpha(\lambda S) = \lambda^2 \alpha(S)$.

18. Symmetry and the Integral. Using the Boundedness and Additivity axioms:

(a) Prove that if f is an odd continuous function on $[-a, a]$, then $I_{-a}^a(f) = 0$.
(b) Prove that if f is an even continuous function on $[-a, a]$, then $I_{-a}^a(f) = 2I_0^a(f)$.
(c) Suppose f is continuous and $I_a^b(f) = 0$ for all intervals $[a, b]$. Prove that $f(x) = 0$ for all x.

19. The Zero-Gap Partition. Let g be a bounded function on $[a, b]$. Suppose there exists a single partition P such that $L(g, P) = U(g, P)$.

(a) Describe the behaviour of the function g on the subintervals of P.
(b) Prove that such a function is Riemann integrable.
(c) Can a non-constant continuous function satisfy this condition for any partition?

20. Accumulating Discontinuities. Let $f : [0, 1] \to \mathbb{R}$ be defined by:

$$f(x) = \begin{cases} 1 & \text{if } x = 1/n \text{ for some } n \in \mathbb{N} \\ 0 & \text{otherwise} \end{cases}$$

(a) Show that f is discontinuous at infinitely many points. What is the set of discontinuities?

(b) Prove that f is integrable and find $\int_0^1 f$ by isolating the points near zero.

21. Absolute Integrability and Inequalities. Let f be integrable on $[a, b]$.

(a) Prove that $|f|$ is integrable.
(b) Prove that $|\int_a^b f| \leq \int_a^b |f|$.

22. Content Zero and the Cantor Set. A set A has content zero if for every $\epsilon > 0$ it can be covered by a finite union of intervals summing to length less than ϵ.

(a) Show that if the set of discontinuities of f has content zero, then f is integrable.
(b) Show that the Cantor set has content zero.
(c) Let h be the indicator function of the Cantor set. Prove h is integrable and compute its integral.

23. The Cauchy-Schwarz Inequality for Integrals. Let $f, g \in \mathcal{R}[a, b]$.

(a) Show that f^2 and fg are integrable.
(b) Prove that $\left(\int_a^b fg\right)^2 \leq \left(\int_a^b f^2\right)\left(\int_a^b g^2\right)$.

24. First Principles Integration. Calculate $\int_0^b x^3\,dx$ directly from the definition using a uniform partition and the formula $\sum_{k=1}^n k^3 = \left(\frac{n(n+1)}{2}\right)^2$.

25. Geometric Partitioning. Calculate $\int_a^b \frac{1}{x^2}\,dx$ (where $0 < a < b$) using a partition in geometric progression $x_k = aq^k$ where $q = (b/a)^{1/n}$.

26. Manual Sum Computation. Let $f(x) = 1/x$ on $[1, 4]$.

(a) Calculate $L(f, P)$ and $U(f, P)$ for the partition $P = \{1, 1.5, 2, 4\}$.
(b) Verify that refining P by adding $x = 3$ decreases the difference $U - L$.

27. The Derivative of the Integral. Let f be Riemann integrable on $[a, b]$ and define $F(x) = \int_a^x f(t)dt$.

(a) Prove or disprove: F is Lipschitz continuous.
(b) Find an example where f is not continuous at c, but F is still differentiable at c.
(c) Find an example where F is differentiable at c, but $F'(c) \neq f(c)$.

Chapter 12

Sequences and Series of Functions

The preceding chapters established differentiation and integration. To construct transcendental functions from first principles, we use power series, viewing a function as a limit of polynomials and hence as the limit of a sequence of functions. Analytic properties need not pass to the limit under pointwise convergence, so uniform convergence becomes essential.

12.1 Power Series

Definition 12.1.1. ***Power Series***. Let $(a_n)_{n=0}^{\infty}$ be a sequence of real numbers. A power series is a formal expression of the form:

$$\sum_{n=0}^{\infty} a_n x^n = a_0 + a_1 x + a_2 x^2 + \dots$$

We adopt the convention that $x^0 = 1$ for all $x \in \mathbb{R}$ (even $x = 0$), so that the series converges to a_0 when $x = 0$.

While one can define power series centred at an arbitrary point x_0 (i.e., $\sum a_n (x - x_0)^n$), the theory is identical to the case $x_0 = 0$ via a simple translation. We shall therefore restrict our attention to series centred at the origin.

The fundamental question is one of convergence: for which values of x does this infinite sum yield a finite real number? Clearly, it always converges at $x = 0$. The behaviour for $x \neq 0$ is governed by a remarkable rigidity: the set of convergence is necessarily an interval centred at the origin.

***Proposition* 12.1.1.** *Abel's Lemma.* If the power series $\sum a_n x^n$ converges at a point $c \neq 0$, then it converges absolutely for all x satisfying $|x| < |c|$.

Proof. Since $\sum a_n c^n$ converges, the sequence $(a_n c^n)$ is bounded: there exists $M > 0$ with $|a_n c^n| \le M$ for all n. Let $|x| < |c|$ and set $r = |x|/|c| < 1$. Then

$$|a_n x^n| = |a_n c^n| r^n \le M r^n.$$

The geometric series $\sum M r^n$ converges, so $\sum a_n x^n$ converges absolutely by comparison. ∎

Definition 12.1.2. ***Radius of Convergence***. Given a power series $\sum a_n x^n$, we define the radius of convergence $R \in [0, \infty]$ by:

$$R = \sup \left\{ |c| : \sum_{n=0}^{\infty} a_n c^n \text{ converges} \right\}$$

It follows immediately from proposition 12.1.1 that the series converges absolutely for all $x \in (-R, R)$. Furthermore, the series diverges for all $|x| > R$; if it converged at some point x_0 with $|x_0| > R$, Abel's Lemma would imply convergence for all $|y| < |x_0|$, contradicting that R is the supremum. The behaviour at the boundary points $x = \pm R$ is delicate and must be checked on a case-by-case basis.

This construction yields a function $f : (-R, R) \to \mathbb{R}$ defined by $f(x) = \sum_{n=0}^{\infty} a_n x^n$. The central questions of this chapter are:

1. Is f continuous on $(-R, R)$?
2. Is f differentiable? If so, is $f'(x) = \sum n a_n x^{n-1}$?
3. Can we integrate f term-by-term?

A power series is a limit of partial sums, that is, a limit of a sequence of functions. To answer these questions, we must study the general theory of convergence of functions.

Pointwise and Uniform Convergence

Let $S \subseteq \mathbb{R}$ and let (f_n) be a sequence of functions $f_n : S \to \mathbb{R}$. We wish to define what it means for f_n to converge to a limit function f. The most obvious definition is to fix x and take the limit of the sequence of numbers $(f_n(x))$.

Definition 12.1.3. ***Pointwise Convergence***. We say that the sequence (f_n) converges pointwise to f on S if for every $x \in S$:

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

Formally: $\forall x \in S, \forall \epsilon > 0, \exists N \in \mathbb{N}$ such that $n \geq N \implies |f_n(x) - f(x)| < \epsilon$.

Crucially, in pointwise convergence, the integer N depends on both the tolerance ϵ and the location x. We denote this dependency as $N(x, \epsilon)$.

Example 12.1.1. Failure of Continuity preservation. Consider the sequence of functions $f_n : [0, 1] \to \mathbb{R}$ defined by $f_n(x) = x^n$. For any fixed $x \in [0, 1)$, we have $x^n \to 0$. At $x = 1$, we have $f_n(1) = 1^n \to 1$. Thus, the pointwise limit is:

$$f(x) = \begin{cases} 0 & 0 \leq x < 1 \\ 1 & x = 1 \end{cases}$$

Each f_n is a polynomial and hence continuous (in fact, smooth). Yet, the limit function f is discontinuous at $x = 1$. This demonstrates that pointwise convergence is too weak to preserve basic analytic properties.

Why did the continuity break? In the ϵ-δ proof of continuity, we typically estimate $|f(x) - f(y)|$. A natural approach using the sequence (f_n) involves the triangle inequality:

$$|f(x) - f(y)| \leq \underbrace{|f(x) - f_n(x)|}_{\text{small for large } n} + \underbrace{|f_n(x) - f_n(y)|}_{\text{small for small } |x-y|} + \underbrace{|f_n(y) - f(y)|}_{\text{small for large } n} \tag{12.1}$$

To make the first and third terms small simultaneously, we need an n that works for both x and y. If y is close to a "bad" point (like 1 in the example above), the rate of convergence might be arbitrarily slow, forcing n to infinity as $y \to 1$. We need a mode of convergence where the rate is independent of the spatial variable x.

Definition 12.1.4. ***Uniform Convergence***. A sequence of functions (f_n) converges uniformly to f on S if for every $\epsilon > 0$, there exists an integer $N \in \mathbb{N}$ such that for all $n \geq N$ and for all $x \in S$:

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

- **Pointwise:** $\forall \epsilon > 0, \forall x \in S, \exists N \dots$
- **Uniform:** $\forall \epsilon > 0, \exists N, \forall x \in S \dots$

In uniform convergence, N depends only on ϵ, not on x. Geometrically, this means that for $n \geq N$, the graph of f_n lies entirely within a "tube" of radius ϵ around the graph of f (Figure 12.1).

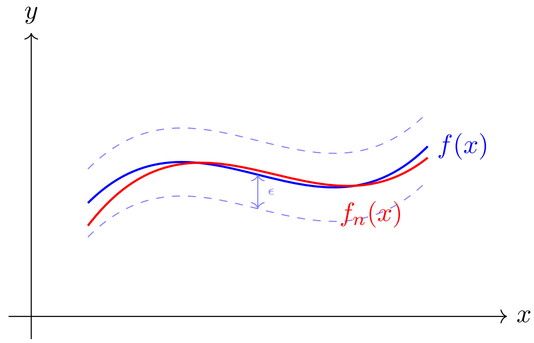


Figure 12.1: Visualisation of Uniform Convergence.

We can now resolve the continuity question. The condition of uniform convergence provides exactly the control needed to execute the $\epsilon/3$ argument suggested in Equation 12.1.

***Theorem* 12.1.1. Uniform Limit Theorem.** Let (f_n) be a sequence of continuous functions defined on a set $S \subseteq \mathbb{R}$. If (f_n) converges uniformly to f on S, then f is continuous on S.

Proof. Let $c \in S$ and $\epsilon > 0$. By definition 12.1.4, there exists N such that $|f_N(y) - f(y)| < \epsilon/3$ for all $y \in S$. By the continuity of f_N, there exists $\delta > 0$ such that $|x - c| < \delta \implies |f_N(x) - f_N(c)| < \epsilon/3$. For $|x - c| < \delta$:

$$|f(x) - f(c)| \leq |f(x) - f_N(x)| + |f_N(x) - f_N(c)| + |f_N(c) - f(c)| < \frac{\epsilon}{3} + \frac{\epsilon}{3} + \frac{\epsilon}{3} = \epsilon$$

■

Remark. This theorem allows us to prove that a function is *not* the uniform limit of a sequence. Returning to $f_n(x) = x^n$ on $[0, 1]$, since the limit function is discontinuous, the convergence cannot be uniform. Indeed, the "error bump" x^n near $x = 1$ does not flatten out; it merely gets squashed against the boundary.

12.2 Cauchy Criterion for Uniform Convergence

In practice, we often do not know the limit function f in advance. For sequences of real numbers, the Cauchy criterion allows us to prove convergence without knowing the limit. An analogous criterion exists for uniform convergence.

***Theorem* 12.2.1. Cauchy Criterion for Uniform Convergence.** A sequence of functions (f_n) defined on S converges uniformly on S if and only if for every $\epsilon > 0$, there exists $N \in \mathbb{N}$ such that for all $m, n \geq N$ and all $x \in S$:

$$|f_n(x) - f_m(x)| < \epsilon$$

Proof. We prove this both ways:

($\implies$): If $f_n \to f$ uniformly, then for n, m large enough, both f_n and f_m are within $\epsilon/2$ of f. By the triangle inequality, they are within ϵ of each other.

($\impliedby$): For each fixed x, the sequence $(f_n(x))$ is a Cauchy sequence of real numbers. By the completeness of $\mathbb{R}$, it converges to some value; call it $f(x)$. This defines the pointwise limit f. To show uniformity, fix $\epsilon > 0$ and choose N such that for all $m, n \geq N$ and all $x \in S$, $|f_n(x) - f_m(x)| < \epsilon/2$. Fixing $n \geq N$ and x, and letting $m \to \infty$ in the inequality, we obtain:

$$|f_n(x) - f(x)| \leq \epsilon/2 < \epsilon$$

Since N was independent of x, the convergence is uniform.

■

For series of functions $\sum g_n(x)$, the Cauchy criterion yields a particularly user-friendly test for uniform convergence, known as the Weierstrass M-Test. This is the primary tool used to establish the regularity of power series.

***Theorem* 12.2.2. Weierstrass M-Test**. Let (g_n) be a sequence of functions on S. Suppose there exists a sequence of non-negative real numbers (M_n) such that:

1. $|g_n(x)| \leq M_n$ for all $x \in S$ and all $n \in \mathbb{N}$.
2. The series of constants $\sum_{n=0}^{\infty} M_n$ converges.

Then the series $\sum_{n=0}^{\infty} g_n(x)$ converges uniformly on S.

Proof. Let $S_k(x) = \sum_{n=0}^{k} g_n(x)$ be the partial sums. For $m > k$:

$$|S_m(x) - S_k(x)| = \left| \sum_{n=k+1}^{m} g_n(x) \right| \leq \sum_{n=k+1}^{m} |g_n(x)| \leq \sum_{n=k+1}^{m} M_n$$

Since $\sum M_n$ converges, the tails of the series can be made arbitrarily small (Cauchy criterion for real series). Thus, for large enough k, the difference $|S_m(x) - S_k(x)|$ is uniformly small for all x. By the Cauchy Criterion for Uniform Convergence, the series converges uniformly. ■

Example 12.2.1. Uniform Convergence of Power Series. Consider a power series $\sum a_n x^n$ with radius of convergence R. Let $0 < \rho < R$. For any $x \in [-\rho, \rho]$, we have $|a_n x^n| \leq |a_n|\rho^n$. Since $\rho < R$, the series $\sum |a_n|\rho^n$ converges (absolute convergence inside the radius). Taking $M_n = |a_n|\rho^n$, the Weierstrass M-Test implies that the power series converges uniformly on the closed interval $[-\rho, \rho]$. Consequently, the limit function $f(x)$ is continuous on $[-\rho, \rho]$. Since ρ can be any value less than R, f is continuous on the open interval $(-R, R)$.

12.3 The Exponential Function

We now connect power series with the exponential function defined earlier. Let $E(x) = e^x$, where e is Euler's number and a^x is defined in the earlier definition of the exponential function. By the theorem Convergence of the Exponential Series, we have

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

for all $x \in \mathbb{R}$.

***Proposition* 12.3.1.** *Differential Characterisation of the Exponential.* The exponential function satisfies $E(0) = 1$ and $E'(x) = E(x)$ for all $x \in \mathbb{R}$. If g is differentiable on $\mathbb{R}$ with $g(0) = 1$ and $g' = g$, then $g = E$.

Proof. The derivative formula $E' = E$ is established in the chapter on differentiation. For uniqueness, let g satisfy the hypotheses. Define $h(x) = g(x)g(-x)$. Then

$$h'(x) = g'(x)g(-x) - g(x)g'(-x) = g(x)g(-x) - g(x)g(-x) = 0,$$

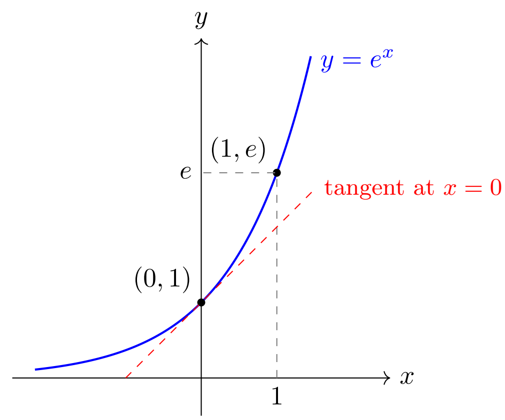
so h is constant and $h(0) = 1$. Hence $g(x)g(-x) = 1$ and g is never zero. Now set $Q(x) = E(x)/g(x)$; then

$$Q'(x) = \frac{E'(x)g(x) - E(x)g'(x)}{g(x)^2} = 0.$$

Thus Q is constant and $Q(0) = 1$, so $g = E$. ■

***Proposition* 12.3.2.** *Algebraic and Order Properties.* The identities in the theorem Properties of Exponentials yield $E(x + y) = E(x)E(y)$ and $E(-x) = 1/E(x)$. In particular $E(x) > 0$ for all x. Since $E' = E > 0$, E is strictly increasing, and since $E'' = E > 0$, E is strictly convex.

We write $E(x)$ as e^x. The graph is shown in Figure 12.2.



The Exponential Function: $f' = f$, $f(0) = 1$.

Figure 12.2: Graph of the exponential function, illustrating convexity and growth.

The Natural Logarithm

Since E is strictly increasing and maps $\mathbb{R}$ onto $(0, \infty)$, it admits an inverse, denoted ln, as defined earlier. By Inverse Function Theorem, ln is differentiable on $(0, \infty)$.

***Proposition* 12.3.3.** *Differentiation of the Logarithm.* The function ln is differentiable on $(0, \infty)$ and its derivative is given by:

$$\frac{d}{dx} \ln x = \frac{1}{x}$$

Proof. Let $f(x) = E(x)$ and $g(y) = \ln y$. By Inverse Function Theorem:

$$g'(y) = \frac{1}{f'(g(y))} = \frac{1}{E(\ln y)} = \frac{1}{y}$$

■

***Corollary* 12.3.1.** *Algebraic Property.* For any $a, x \in (0, \infty)$:

$$\ln(ax) = \ln a + \ln x$$

Proof. This is the product rule in the theorem Properties of Logarithms, specialised to base e. ■

By the logarithm power law, $\ln(x^n) = n \ln x$ for $n \in \mathbb{N}$. Since ln is the inverse of a strictly increasing function, it is strictly increasing. The limits follow from monotonicity and $\ln(1) = 0$:

$$\lim_{x\to\infty} \ln x = \infty, \qquad \lim_{x\to 0^+} \ln x = -\infty.$$

We may also analyse the convexity of the logarithm. Computing the second derivative:

$$\frac{d^2}{dx^2} \ln x = \frac{d}{dx}\left(\frac{1}{x}\right) = -\frac{1}{x^2}$$

Since $-\frac{1}{x^2} < 0$ for all x, the logarithm is strictly concave.

With the logarithm at hand, we use the representation already established for real powers (theorem 8.2.3):

$$a^x = e^{x \ln a} \qquad (a > 0,\ x \in \mathbb{R}).$$

This agrees with integer powers.

The logarithm grows notoriously slowly. We quantify this by comparing it to polynomial growth.

***Theorem* 12.3.1. Growth of the Logarithm**. For any integer $k \geq 1$:

$$\lim_{x \to \infty} \frac{(\ln x)^k}{x} = 0$$

Proof. Let $x = e^z$. As $x \to \infty$, $z \to \infty$. The limit becomes:

$$\lim_{z \to \infty} \frac{z^k}{e^z}$$

We know from the power series expansion of e^z that for $z > 0$, $e^z > \frac{z^{k+1}}{(k+1)!}$. Thus:

$$0 < \frac{z^k}{e^z} < \frac{z^k}{\frac{z^{k+1}}{(k+1)!}} = \frac{(k+1)!}{z}$$

As $z \to \infty$, the upper bound vanishes. By the Squeeze Theorem, the limit is 0. ∎

The limit $\lim_{h \to 0}(1+h)^{1/h} = e$ is established in the chapter on fundamental transcendental limits.

12.4 Trigonometric Functions

We turn now to the circular functions, sine and cosine. Rather than relying on geometric intuition involving triangles (which requires a prior rigorous definition of arc length and angle), we define them analytically via their Maclaurin series.

Definition 12.4.1. *Sine and Cosine*. We define the functions $\sin : \mathbb{R} \to \mathbb{R}$ and $\cos : \mathbb{R} \to \mathbb{R}$ by the power series:

$$\sin x = \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n+1)!} x^{2n+1} = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \dots$$

$$\cos x = \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n)!} x^{2n} = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \dots$$

The convergence of these series for all $x \in \mathbb{R}$ follows easily from the Ratio Test (the limit is 0, similar to the exponential function). Alternatively, one observes that the coefficients are bounded by those of $e^{|x|}$.

Immediate properties follow from the series structure:

1. $\sin(-x) = -\sin x$ (odd function).
2. $\cos(-x) = \cos x$ (even function).
3. $\sin(0) = 0$ and $\cos(0) = 1$.

Differentiating term-by-term (valid within the infinite radius of convergence):

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

$$\frac{d}{dx} \cos x = \sum_{n=1}^{\infty} \frac{(-1)^n (2n) x^{2n-1}}{(2n)!} = \sum_{n=1}^{\infty} \frac{(-1)^n x^{2n-1}}{(2n-1)!} = -\sum_{k=0}^{\infty} \frac{(-1)^k x^{2k+1}}{(2k+1)!} = -\sin x$$

where we re-indexed $k = n - 1$ in the second calculation.

***Proposition* 12.4.1.** *Pythagorean Identity.* For all $x \in \mathbb{R}$, $\sin^2 x + \cos^2 x = 1$.

Proof. Let $g(x) = \sin^2 x + \cos^2 x$. Differentiating:

$$g'(x) = 2\sin x(\sin x)' + 2\cos x(\cos x)' = 2\sin x \cos x - 2\cos x \sin x = 0$$

Thus g is constant. $g(0) = \sin^2 0 + \cos^2 0 = 0 + 1 = 1$. Hence $\sin^2 x + \cos^2 x = 1$ for all x. ∎

To establish the addition formulae and periodicity (and thus link these functions to geometry), we rely on a uniqueness theorem for the defining differential equations.

***Theorem* 12.4.1. Uniqueness of Trigonometric Functions.** Suppose $f, g : \mathbb{R} \to \mathbb{R}$ are differentiable functions satisfying:

$$f'(x) = g(x), \quad g'(x) = -f(x), \quad f(0) = 0, \quad g(0) = 1$$

Then $f(x) = \sin x$ and $g(x) = \cos x$.

Proof. Let $S(x) = \sin x$ and $C(x) = \cos x$. These functions satisfy the conditions. We must show they are the only ones. Consider a general pair f_1, g_1 satisfying the differential relations (but not necessarily the initial conditions yet). Let f, g be the standard sine and cosine. Define two auxiliary functions:

$$A(x) = f(x)g_1(x) - f_1(x)g(x)$$

$$B(x) = f(x)f_1(x) + g(x)g_1(x)$$

Differentiating $A(x)$:

$$\begin{aligned} A'(x) &= f'g_1 + fg_1' - f_1'g - f_1g' \\ &= gg_1 + f(-f_1) - g_1g - f_1(-f) \\ &= gg_1 - ff_1 - gg_1 + ff_1 = 0 \end{aligned}$$

Thus $A(x)$ is constant. Similarly, differentiating $B(x)$:

$$\begin{aligned} B'(x) &= f'f_1 + ff_1' + g'g_1 + gg_1' \\ &= gf_1 + fg_1 + (-f)g_1 + g(-f_1) = 0 \end{aligned}$$

Thus $B(x)$ is constant.

Now, suppose f_1, g_1 satisfy the initial conditions $f_1(0) = 0, g_1(0) = 1$. We evaluate the constants:

$$A(0) = \sin(0) \cdot 1 - 0 \cdot \cos(0) = 0$$

$$B(0) = \sin(0) \cdot 0 + \cos(0) \cdot 1 = 1$$

We obtain the linear system:

$$(\sin x)g_1(x) - f_1(x)(\cos x) = 0$$

$$(\sin x)f_1(x) + (\cos x)g_1(x) = 1$$

Multiply the first by $\cos x$ and the second by $\sin x$ and subtract/add to solve for f_1. It is faster to observe that $B(x) = 1$ implies f_1 and g_1 lie on the unit circle relative to the rotating frame of f, g. Specifically, multiply the second eq by $f = \sin x$: $f^2 f_1 + fgg_1 = f$. Multiply the first eq by $g = \cos x$: $fgg_1 - f_1g^2 = 0 \implies fgg_1 = f_1g^2$. Substitute into the modified second eq: $f^2 f_1 + f_1 g^2 = f \implies f_1(f^2 + g^2) = f$. Since $f^2 + g^2 = 1$, we get $f_1 = f = \sin x$. Similarly $g_1 = g = \cos x$. ∎

This proof technique yields the addition formulae effortlessly by choosing specific functions for f_1 and g_1.

***Theorem* 12.4.2. Addition Formulae.** For all $x, y \in \mathbb{R}$:

1. $\sin(x + y) = \sin x \cos y + \cos x \sin y$
2. $\cos(x + y) = \cos x \cos y - \sin x \sin y$

Proof. Fix $c \in \mathbb{R}$. Let $f_1(x) = \sin(x+c)$ and $g_1(x) = \cos(x+c)$. By the Chain Rule, $f_1'(x) = \cos(x+c) = g_1(x)$ and $g_1'(x) = -\sin(x+c) = -f_1(x)$. Thus f_1, g_1 satisfy the differential relations used in the previous proof. The constants A and B are:

$$A = A(0) = \sin(0)\cos(c) - \sin(c)\cos(0) = -\sin c$$

$$B = B(0) = \sin(0)\sin(c) + \cos(0)\cos(c) = \cos c$$

Returning to the system derived in the proof of uniqueness:

$$(\sin x)g_1(x) - f_1(x)(\cos x) = -\sin c$$

$$(\sin x)f_1(x) + (\cos x)g_1(x) = \cos c$$

Solving this system for $f_1(x)$ (multiply second by $\sin x$, first by $-\cos x$ and add... or essentially invert the rotation matrix):

$$f_1(x) = B\sin x - A\cos x = \cos c \sin x - (-\sin c)\cos x = \sin x \cos c + \cos x \sin c$$

$$g_1(x) = A\sin x + B\cos x = -\sin c \sin x + \cos c \cos x = \cos x \cos c - \sin x \sin c$$

Replacing c with y gives the result. ∎

To define π as the period of these functions, we must locate the zeros of $\cos x$ using theorem 6.4.2.

The Number π and the Geometry of Trigonometric Functions

We have defined the sine and cosine functions via power series and established their algebraic properties, most notably $\sin^2 x + \cos^2 x = 1$ and the addition formulae. However, their geometric significance and periodicity remain to be proved. To do this, we must locate the zeros of the cosine function. This leads to the rigorous definition of the constant π.

From the identity $\sin^2 x + \cos^2 x = 1$, we immediately have the bounds $|\sin x| \leq 1$ and $|\cos x| \leq 1$ for all $x \in \mathbb{R}$. We observe that $\sin(0) = 0$ and $\sin'(0) = \cos(0) = 1$. Since the derivative is continuous and positive at 0, the sine function is strictly increasing in some neighbourhood of 0. Thus, for small $x > 0$, we have $\sin x > 0$.

***Theorem* 12.4.3. Existence of a Root**. There exists a real number $x_0 > 0$ such that $\cos x_0 = 0$.

Proof. We proceed by contradiction. **Assume** that $\cos x \neq 0$ for all $x \in \mathbb{R}$. Since $\cos 0 = 1$ and cosine is continuous, theorem 6.4.2 implies that $\cos x > 0$ for all $x \in \mathbb{R}$. (If it assumed a negative value, it would have to cross zero). Consequently, $\sin'(x) = \cos x > 0$ for all x. This implies that $\sin x$ is a strictly increasing function on $\mathbb{R}$. Since $\sin 0 = 0$, we have $\sin x > 0$ for all $x > 0$.

We now employ the double angle formula derived in the previous section:

$$\cos(2x) = \cos^2 x - \sin^2 x = 1 - 2\sin^2 x$$

Since $\sin x$ is strictly increasing and bounded above by 1, $\cos x$ cannot be bounded below by a positive constant. If $\cos x \geq \frac{1}{2}$ for all x, then by theorem 9.9.2 on $[0, x]$,

$$\sin x - \sin 0 = \sin'(\xi)x \geq \frac{x}{2}$$

for some $\xi \in (0, x)$, contradicting boundedness. Thus, there exists $b > 0$ such that $\cos b < \frac{1}{2}$. It follows that $\sin^2 b = 1 - \cos^2 b > 1 - \frac{1}{4} = \frac{3}{4}$. Now consider the value of cosine at $2b$:

$$\cos(2b) = \cos^2 b - \sin^2 b < \frac{1}{4} - \frac{3}{4} = -\frac{1}{2}$$

This contradicts the assumption that $\cos x > 0$ for all x. Therefore, the assumption is false, and there must exist some x_0 where $\cos x_0 = 0$. ∎

We have established that the set $\mathcal{Z} = \{x > 0 : \cos x = 0\}$ is non-empty. Since cosine is continuous and $\cos 0 = 1$, there is a neighbourhood of 0 where cosine is non-zero. Thus $\inf \mathcal{Z} > 0$.

Definition 12.4.2. ***The Number Pi.*** We define the real number π by:

$$\frac{\pi}{2} = \inf\{x > 0 : \cos x = 0\}$$

Equivalently, π is twice the smallest positive root of the cosine function.

By the continuity of cosine, it follows that $\cos(\frac{\pi}{2}) = 0$. Furthermore, by the definition of the infimum, $\cos x > 0$ for all $x \in [0, \frac{\pi}{2})$. Since $\sin' x = \cos x > 0$ on this interval, the sine function is strictly increasing on $[0, \frac{\pi}{2}]$. Using the Pythagorean identity at the endpoint:

$$\sin^2\left(\frac{\pi}{2}\right) + \cos^2\left(\frac{\pi}{2}\right) = 1 \implies \sin^2\left(\frac{\pi}{2}\right) = 1$$

Since $\sin x$ starts at 0 and increases, we must have $\sin(\frac{\pi}{2}) = 1$.

Having anchored the values at $\frac{\pi}{2}$, the addition formulae allow us to propagate these values to other multiples of π, establishing the periodicity of the functions.

***Proposition* 12.4.2.** *Shift Identities.* For all $x \in \mathbb{R}$:

1. $\cos(x + \frac{\pi}{2}) = -\sin x$
2. $\sin(x + \frac{\pi}{2}) = \cos x$
3. $\cos(x + \pi) = -\cos x$
4. $\sin(x + \pi) = -\sin x$
5. $\cos(x + 2\pi) = \cos x$
6. $\sin(x + 2\pi) = \sin x$

Proof. We prove (1) and (2); the others follow by iteration. Using the addition formula and the values $\cos(\frac{\pi}{2}) = 0, \sin(\frac{\pi}{2}) = 1$:

$$\cos\left(x + \frac{\pi}{2}\right) = \cos x \cos\frac{\pi}{2} - \sin x \sin\frac{\pi}{2} = (\cos x)(0) - (\sin x)(1) = -\sin x$$

$$\sin\left(x + \frac{\pi}{2}\right) = \sin x \cos\frac{\pi}{2} + \cos x \sin\frac{\pi}{2} = (\sin x)(0) + (\cos x)(1) = \cos x$$

For (3), we apply (1) twice: $\cos(x + \pi) = \cos((x + \frac{\pi}{2}) + \frac{\pi}{2}) = -\sin(x + \frac{\pi}{2}) = -(\cos x)$. Similarly for (4). For (5), $\cos(x + 2\pi) = \cos((x + \pi) + \pi) = -\cos(x + \pi) = -(-\cos x) = \cos x$. The proof for (6) is identical. ■

This result confirms that both sine and cosine are periodic with period 2π. Moreover, 2π is the *fundamental* period. If there were a smaller period $0 < p < 2\pi$, it would contradict the sign properties of sine and cosine in the interval $[0, 2\pi]$ derived below.

We can now rigorously analyse the shape of these functions using the derivatives (monotonicity) and second derivatives (convexity).

1. **Interval** $[0, \frac{\pi}{2}]$**:**
   - $\cos x > 0 \implies \sin x$ is strictly increasing (from 0 to 1).
   - $\sin x > 0 \implies \cos x$ is strictly decreasing (from 1 to 0).
   - **Convexity:** $\sin''(x) = -\sin x < 0$, so sin is concave. $\cos''(x) = -\cos x < 0$, so cos is concave.
2. **Interval** $[\frac{\pi}{2}, \pi]$**:** Using $\sin x = \cos(x - \frac{\pi}{2})$ and $\cos x = -\sin(x - \frac{\pi}{2})$:
   - $\sin x$ decreases from 1 to 0 and remains concave.
   - $\cos x$ decreases from 0 to -1 and is convex since $\cos''(x) = -\cos x > 0$.

Extending this analysis to $[0, 2\pi]$ yields the familiar wave patterns in Figure 12.3.

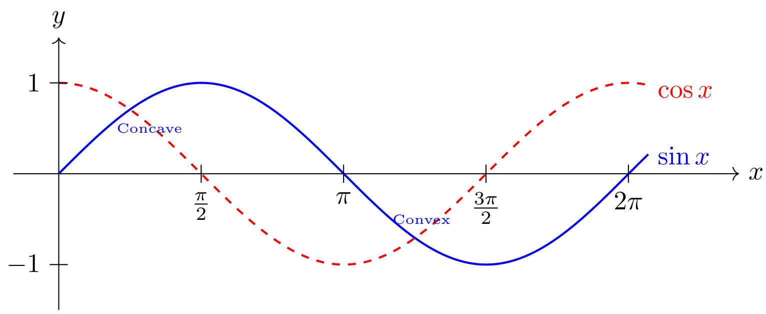


Figure 12.3: Graphs of the trigonometric functions on $[0, 2\pi]$. The sign of the second derivative $f''(x) = -f(x)$ determines the convexity.

The next inequality will be used later to dominate terms in the Basel Problem proof.

***Proposition* 12.4.3.** *Jordan's Inequality.* For $x \in [0, \pi/2]$:

$$\frac{2}{\pi}x \le \sin x \le x$$

Proof. On $[0, \pi/2]$, we have $\sin''(x) = -\sin x \le 0$, so sin is concave. A concave function lies below its tangents and above its chords. The tangent at 0 is $y = x$, giving $\sin x \le x$. The chord from $(0, 0)$ to $(\pi/2, 1)$ is the line $y = \frac{2}{\pi}x$, giving $\sin x \ge \frac{2}{\pi}x$. ■

12.5 The Basel Problem

We compute the sum of $\sum 1/n^2$, whose convergence was established earlier but whose value was left open. This is the Basel problem, named after the hometown of Leonhard Euler. We give a modern proof using the trigonometric identities above and **Tannery's Theorem**, which controls the interchange of limits and infinite sums.

Remark. (Tannery's Theorem). Let $S_n = \sum_{k=0}^{\infty} a_k(n)$. If $\lim\limits_{n\to\infty} a_k(n) = a_k$ for each k, and there exists a sequence M_k such that $|a_k(n)| \le M_k$ for all n, k with $\sum M_k < \infty$, then:

$$\lim_{n\to\infty} \sum_{k=0}^{\infty} a_k(n) = \sum_{k=0}^{\infty} a_k$$

Euler's original factorisation argument was later justified by the Weierstrass Factorisation Theorem. We do not retrace that approach here.

***Theorem* 12.5.1. The Solution to the Basel Problem.**

$$\sum_{n=1}^{\infty} \frac{1}{n^2} = \frac{\pi^2}{6}$$

Proof. The proof proceeds in three stages: establishing a trigonometric recurrence, iterating this to form a finite sum, and taking the limit using Tannery's Theorem.

Step 1: A Trigonometric Identity. We begin with the fundamental identity $\sin^2 x + \cos^2 x = 1$ and the half-angle formula $\sin x = 2\sin(x/2)\cos(x/2)$. Squaring the reciprocal of the half-angle formula yields:

$$\frac{1}{\sin^2 x} = \frac{1}{4\sin^2(x/2)\cos^2(x/2)}$$

Using the identity $\cos^2\theta = \sin^2(\theta + \pi/2)$, we can rewrite the numerator $1 = \sin^2(x/2) + \cos^2(x/2)$ to separate the fraction:

$$\frac{1}{\sin^2 x} = \frac{\sin^2(x/2) + \cos^2(x/2)}{4\sin^2(x/2)\cos^2(x/2)} = \frac{1}{4}\left[\frac{1}{\cos^2(x/2)} + \frac{1}{\sin^2(x/2)}\right]$$

Substituting $\cos^2(x/2) = \sin^2(x/2 + \pi/2)$, we obtain the recurrence:

$$\frac{1}{\sin^2 x} = \frac{1}{4}\left[\frac{1}{\sin^2\frac{x}{2}} + \frac{1}{\sin^2\frac{x+\pi}{2}}\right] \tag{12.2}$$

Step 2: Iteration and Finite Sums. We evaluate this identity at $x = \pi/2$. Since $\sin(\pi/2) = 1$, the left-hand side is 1.

$$1 = \frac{1}{4}\left[\frac{1}{\sin^2\frac{\pi}{4}} + \frac{1}{\sin^2\frac{3\pi}{4}}\right]$$

Observe that $\sin(3\pi/4) = \sin(\pi - \pi/4) = \sin(\pi/4)$. Thus, the terms in the bracket are identical. We apply the identity (12.2) again to each term. For instance, $\frac{1}{\sin^2(\pi/4)}$ splits into terms involving $\pi/8$ and $3\pi/8$. Let us generalise. After n iterations, we obtain a sum of the form:

$$1 = \frac{1}{4^n}\sum_{k=0}^{2^n-1}\frac{1}{\sin^2\left(\frac{(2k+1)\pi}{2^{n+1}}\right)}$$

Due to the symmetry $\sin(\theta) = \sin(\pi - \theta)$, the terms in the second half of the sum (from $k = 2^{n-1}$ to $2^n - 1$) are mirrored repetitions of the terms in the first half. Specifically, the angle for index k and the angle for index $2^n - 1 - k$ sum to π. Thus, we can deduce the sum to k ranging only up to $2^{n-1} - 1$ and multiply by 2:

$$1 = \frac{2}{4^n}\sum_{k=0}^{2^{n-1}-1}\frac{1}{\sin^2\left(\frac{(2k+1)\pi}{2^{n+1}}\right)} \tag{12.3}$$

Step 3: The Limit Process. We wish to take the limit as $n \to \infty$. Let $N = 2^{n+1}$. The equation (12.3) becomes:

$$1 = \frac{2}{4^n}\sum_{k=0}^{2^{n-1}-1}\frac{1}{\sin^2\left(\frac{(2k+1)\pi}{N}\right)}$$

We rearrange the pre-factors to isolate the argument of the sine function. Note that $4^n = N^2/4$.

$$1 = \frac{8}{N^2}\sum_{k=0}^{2^{n-1}-1}\frac{1}{\sin^2\left(\frac{(2k+1)\pi}{N}\right)} = \frac{8}{\pi^2}\sum_{k=0}^{2^{n-1}-1}\frac{1}{(2k+1)^2}\left[\frac{\frac{(2k+1)\pi}{N}}{\sin\left(\frac{(2k+1)\pi}{N}\right)}\right]^2$$

Let $a_k(n)$ denote the k-th term of the sum (where $a_k(n) = 0$ if $k \geq 2^{n-1}$). We seek to apply **Tannery's Theorem** (a discrete analogue of the Dominated Convergence Theorem) to interchange the limit $n \to \infty$ and the summation. The term in the square brackets is of the form $\theta/\sin\theta$ where $\theta_k = \frac{(2k+1)\pi}{2^{n+1}}$. As $n \to \infty$, for any fixed k, $\theta_k \to 0$. We know that $\lim_{\theta\to 0}\frac{\theta}{\sin\theta} = 1$. Thus, the pointwise limit of the summand is:

$$\lim_{n\to\infty} a_k(n) = \frac{1}{(2k+1)^2}\cdot 1^2 = \frac{1}{(2k+1)^2}$$

To justify the interchange, we require a dominating sequence M_k such that $|a_k(n)| \leq M_k$ and $\sum M_k < \infty$. Recall Jordan's Inequality: $\sin x \geq \frac{2}{\pi}x$ for $x \in [0, \pi/2]$. Here, the arguments lie in $(0, \pi/2)$, so $\frac{x}{\sin x} \leq \frac{\pi}{2}$. Consequently, the term in the brackets is bounded by $\pi^2/4$.

$$\left|\frac{1}{(2k+1)^2}\left[\frac{\theta}{\sin\theta}\right]^2\right| \leq \frac{1}{(2k+1)^2}\cdot\frac{\pi^2}{4}$$

The series $\sum \frac{1}{(2k+1)^2}$ converges (by comparison to $\sum 1/k^2$). Thus, the conditions of Tannery's Theorem are satisfied. Taking the limit $n \to \infty$ in (12.3):

$$1 = \frac{8}{\pi^2} \sum_{k=0}^{\infty} \frac{1}{(2k+1)^2}$$

Rearranging gives the sum of the reciprocals of the odd squares:

$$\sum_{k=0}^{\infty} \frac{1}{(2k+1)^2} = \frac{\pi^2}{8}$$

Step 4: Algebraic Conclusion. Let $S = \sum_{n=1}^{\infty} \frac{1}{n^2}$. We separate S into even and odd terms:

$$S = \sum_{k=0}^{\infty} \frac{1}{(2k+1)^2} + \sum_{m=1}^{\infty} \frac{1}{(2m)^2}$$

Substituting the odd sum we just found:

$$S = \frac{\pi^2}{8} + \frac{1}{4} \sum_{m=1}^{\infty} \frac{1}{m^2} = \frac{\pi^2}{8} + \frac{1}{4} S$$

Solving for S:

$$\frac{3}{4} S = \frac{\pi^2}{8} \implies S = \frac{\pi^2}{6}$$

■

Euler later generalised this method to define the Riemann zeta function $\zeta(s) = \sum n^{-s}$, where $\zeta(2) = \pi^2/6$. Its study is central to analytic number theory and the distribution of prime numbers.

12.6 Exercises

1. **Radius of Convergence Calculations.** Determine the radius of convergence R for the following power series $\sum a_n x^n$. In each case, investigate the convergence at the boundary points $x = \pm R$.
   (a) $\sum_{n=1}^{\infty} \frac{x^n}{n^2 3^n}$
   (b) $\sum_{n=1}^{\infty} \frac{(n!)^2}{(2n)!} x^n$
   (c) $\sum_{n=1}^{\infty} \left(1 + \frac{1}{n}\right)^{n^2} x^n$

   Remark. Recall the limit of $(1 + 1/n)^n$.

2. **Algebra of Radii.** Let $\sum a_n x^n$ and $\sum b_n x^n$ have radii of convergence R_a and R_b respectively.
   (a) Prove that the radius of convergence of $\sum (a_n + b_n) x^n$ is at least $\min(R_a, R_b)$. Give an example where it is strictly larger.
   (b) Prove that the radius of convergence of the Hadamard product $\sum (a_n b_n) x^n$ is at least $R_a R_b$.

3. **Term-wise Operations.** Let $f(x) = \sum_{n=0}^{\infty} a_n x^n$ have a radius of convergence $R > 0$.
   (a) Prove that the derived series $\sum_{n=1}^{\infty} n a_n x^{n-1}$ also has radius of convergence R.
   (b) Deduce that f is infinitely differentiable on $(-R, R)$.
   (c) Show that $a_n = \frac{f^{(n)}(0)}{n!}$, confirming that every power series is the Taylor series of its limit function.

4. **The Moving Bump.** Consider the sequence of functions $f_n : [0,1] \to \mathbb{R}$ defined by

$$f_n(x) = \frac{nx}{1 + n^2x^2}.$$

   (a) Find the pointwise limit $f(x) = \lim_{n\to\infty} f_n(x)$ for $x \in [0,1]$.
   (b) Determine the maximum value of $f_n(x)$ on $[0,1]$ using calculus.
   (c) Does (f_n) converge uniformly to f? Justify your answer.

5. **Dini's Theorem.** This theorem provides a converse to the statement "uniform limit of continuous functions is continuous" under specific conditions. Let K be a closed, bounded interval (compact). Let (f_n) be a sequence of continuous functions $f_n : K \to \mathbb{R}$ that converges pointwise to a continuous function f. Suppose further that the sequence is monotonic: for every $x \in K$, $f_{n+1}(x) \le f_n(x)$ for all n. Prove that (f_n) converges uniformly to f.

   Remark. Consider the functions $g_n = f_n - f$. These are continuous, non-negative, and decrease to 0. Use the compactness of K to derive a contradiction if the convergence is not uniform.

6. **Preservation of Boundedness.** Let (f_n) be a sequence of bounded functions on a set S. Suppose $f_n \to f$ uniformly on S.
   (a) Prove that the limit function f is bounded on S.
   (b) Prove that the sequence is uniformly bounded; that is, there exists a constant $M > 0$ such that $|f_n(x)| \le M$ for all $n \in \mathbb{N}$ and all $x \in S$.
   (c) Give a counterexample to (a) if the convergence is only pointwise.

7. **Sequence Definition of the Exponential.** We defined e^x via a series. A common alternative definition is $E(x) = \lim_{n\to\infty}(1 + x/n)^n$.
   (a) Using the derivative of $\ln t$ at $t = 1$, show that $\lim_{h\to 0} \frac{\ln(1+h)}{h} = 1$.
   (b) Let $y_n = (1 + x/n)^n$. Show that $\ln y_n = x \cdot \frac{\ln(1+x/n)}{x/n}$.
   (c) Conclude that $\lim_{n\to\infty} y_n = e^x$.

8. **Functional Characterisation.** Let $f : \mathbb{R} \to \mathbb{R}$ be a differentiable function such that $f(x + y) = f(x)f(y)$ for all x, y and f is not identically zero.
   (a) Prove that $f(0) = 1$ and $f(x) > 0$ for all x.
   (b) By considering the derivative quotient, show that $f'(x) = f'(0)f(x)$.
   (c) Deduce that $f(x) = e^{cx}$ where $c = f'(0)$.

9. **Logarithmic Limits.** Evaluate the following limits using the series expansions of the exponential and logarithm:
   (a) $\lim_{x\to 0} \dfrac{\ln(1+x) - x + x^2/2}{x^3}$
   (b) $\lim_{n\to\infty} n\left(x^{1/n} - 1\right)$ for $x > 0$.

10. **Jordan's Inequality.** In the solution to the Basel problem, we used the inequality $\sin x \ge \frac{2}{\pi}x$ for $x \in [0, \pi/2]$.
    (a) Consider the function $g(x) = \frac{\sin x}{x}$. Show that $g'(x) < 0$ on $(0, \pi/2]$.

    Remark. Use the auxiliary function $h(x) = x\cos x - \sin x$.

    (b) Deduce that $g(x)$ is decreasing and evaluate $g(\pi/2)$ to prove the inequality.
    (c) Interpret this geometrically in terms of the chord and the arc of the unit circle.

11. **The Tangent Function.** Define $\tan x = \frac{\sin x}{\cos x}$ for $x \in (-\pi/2, \pi/2)$.
    (a) Prove that $\tan x$ is a bijection from $(-\pi/2, \pi/2)$ to $\mathbb{R}$.
    (b) Show that its inverse, $\arctan x$, has the power series expansion:

$$\arctan x = \sum_{n=0}^{\infty} \frac{(-1)^n}{2n+1} x^{2n+1} \quad \text{for } |x| < 1.$$

Remark. Expand the derivative $1/(1+x^2)$ as a geometric series and integrate term-by-term.

(c) By taking the limit as $x \to 1^-$ (justified by Abel's Theorem, which you may assume), derive the Leibniz series for $\pi/4$.

12. Minimality of the Period. We established that $\sin(x+2\pi) = \sin x$. Prove that no number $p \in (0, 2\pi)$ satisfies $\sin(x+p) = \sin x$ for all x.

Remark. If such a p exists, consider $x = 0$ to show $\sin p = 0$, and thus p must be a multiple of π. Then check the sign of cosine.

13. Calculus of the Zeta Function. Using the result $\sum_{n=1}^{\infty} \frac{1}{n^2} = \frac{\pi^2}{6}$:

(a) Prove that $\sum_{n=1}^{\infty} \frac{1}{(2n)^2} = \frac{\pi^2}{24}$.

(b) Prove that $\sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n^2} = \frac{\pi^2}{12}$.

Remark. Split the alternating sum into total minus twice the evens.

14. Wallis' Product. Follow these steps to derive the infinite product for π. Let $I_n = \int_0^{\pi/2} \sin^n x \, dx$.

(a) Integrate by parts to prove the recurrence $I_n = \frac{n-1}{n} I_{n-2}$.

(b) Establish that $I_{2n} = \frac{(2n-1)!!}{(2n)!!} \frac{\pi}{2}$ and $I_{2n+1} = \frac{(2n)!!}{(2n+1)!!}$.

(c) Explain why $I_{2n+1} \le I_{2n} \le I_{2n-1}$.

(d) Compute the ratio I_{2n+1}/I_{2n-1} and show it tends to 1 as $n \to \infty$.

(e) Deduce that $\lim_{n\to\infty} \frac{2}{1} \cdot \frac{2}{3} \cdot \frac{4}{3} \cdot \frac{4}{5} \cdots \frac{2n}{2n-1} \cdot \frac{2n}{2n+1} = \frac{\pi}{2}$.

15. Evaluation of $\zeta(4)$. We can extend the Basel method to higher powers.

(a) Start with the identity (12.2): $\csc^2 x = \frac{1}{4}[\csc^2(x/2) + \csc^2((x+\pi)/2)]$.

(b) Square both sides. You will obtain cross terms. Use the symmetry of the sum in (12.3) to handle these.

(c) Apply the limit process again using Tannery's theorem to show that $\sum_{n=0}^{\infty} \frac{1}{(2n+1)^4} = \frac{\pi^4}{96}$.

(d) Deduce that $\sum_{n=1}^{\infty} \frac{1}{n^4} = \frac{\pi^4}{90}$.

16. Uniform Convergence Checks. For each sequence (f_n), determine the pointwise limit f on the given interval and determine if the convergence is uniform.

(a) $f_n(x) = x^n(1-x^n)$ on $[0,1]$.

(b) $f_n(x) = \frac{1}{1+n^2x^2}$ on $[-1,1]$.

(c) $f_n(x) = \sqrt{x^2 + \frac{1}{n}}$ on $\mathbb{R}$.

17. Integration and Limits. Evaluate $\lim_{n\to\infty} \int_0^1 \frac{nx}{1+n^2x^4} \, dx$. Compare this with $\int_0^1 \left(\lim_{n\to\infty} f_n(x)\right) dx$. Does the Integral Limit Theorem apply? Why or why not?

18. Power Series Operations. Given that $\frac{1}{1-x} = \sum_{n=0}^{\infty} x^n$ for $|x| < 1$:

(a) Find the power series for $\frac{1}{(1-x)^2}$ by differentiation.

(b) Find the power series for $\ln(1+x)$ by integration.

(c) Use the result of (a) to sum the series $\sum_{n=1}^{\infty} \frac{n}{2^n}$.

19. Failure of the Converse M-Test. The Weierstrass M-Test is sufficient but not necessary. Consider the series $\sum_{n=1}^{\infty} f_n(x)$ on $[0,1]$ where

$$f_n(x) = \begin{cases} \frac{1}{n} & \text{if } \frac{1}{n+1} < x \le \frac{1}{n} \\ 0 & \text{otherwise} \end{cases}$$

(a) Prove that $\sum f_n(x)$ converges uniformly to a limit function $f(x)$ on $[0,1]$.

(b) Let $M_n = \sup_{[0,1]} |f_n(x)|$. Show that $\sum M_n$ diverges.

20. Smooth but Non-Analytic. In the differentiation chapter, we met the function $f(x) = e^{-1/x^2}$ (with $f(0) = 0$).

(a) Show that f has a Taylor series expansion about $x = 0$, and that this series has an infinite radius of convergence.
(b) Show that this series converges uniformly on $\mathbb{R}$.
(c) Show, however, that the sum of the series is identically zero, and thus does not equal $f(x)$ anywhere except at the origin.
(d) *Conceptual:* Why does this not contradict the theorems regarding power series convergence? (Hint: Read the definition of "represented by a power series" carefully).

Chapter 13

Techniques of Integration

By Axiomatic Characterisation of the Integral, evaluating the definite integral $\int_a^b f$ reduces to finding a function F such that $F' = f$. We systematise this search and give methods for constructing primitives.

13.1 Primitives and the Indefinite Integral

Definition 13.1.1. ***Primitive***. Let $I \subseteq \mathbb{R}$ be an interval and let $f : I \to \mathbb{R}$. A function $F : I \to \mathbb{R}$ is called a primitive (or antiderivative) of f on I if F is differentiable and $F'(x) = f(x)$ for all $x \in I$.

Example 13.1.1. Elementary Primitives. The function $F(x) = x^2$ is a primitive of $f(x) = 2x$ on $\mathbb{R}$. Similarly, $F(x) = \sin x$ is a primitive of $f(x) = \cos x$.

The primitive is not unique. If F is a primitive of f, then $G(x) = F(x) + C$ is also a primitive for any constant $C \in \mathbb{R}$, since the derivative of a constant is zero. By theorem 9.9.2, these are the only ambiguities.

***Proposition* 13.1.1.** *Uniqueness up to Constant.* If F_1 and F_2 are primitives of f on an interval I, then there exists a constant $C \in \mathbb{R}$ such that $F_1(x) = F_2(x) + C$ for all $x \in I$.

Proof. Define $H = F_1 - F_2$. Then $H' = F_1' - F_2' = f - f = 0$ on I. By theorem 9.9.2, H is constant, so $H(x) = C$. ∎

This justifies the following notation for the family of all primitives.

Definition 13.1.2. ***Indefinite Integral***. Let $f : I \to \mathbb{R}$. The indefinite integral of f, denoted $\int f(x)\,dx$, represents the set of all primitives of f on I. If F is any specific primitive, we write:

$$\int f(x)\,dx = F(x) + C$$

where C is an arbitrary integration constant.

Since integration is the inverse of differentiation, we may immediately populate our catalogue of integrals by reversing the standard derivative table.

Function $f(x)$	**Primitive** $\int f(x)dx$
$x^n \quad (n \in \mathbb{Z}, n \neq -1)$	$\dfrac{x^{n+1}}{n+1} + C$
$x^\alpha \quad (\alpha \in \mathbb{R}, \alpha \neq -1, x > 0)$	$\dfrac{x^{\alpha+1}}{\alpha+1} + C$
$\dfrac{1}{x} \quad (x \neq 0)$	$\ln\lvert x\rvert + C$
e^x	$e^x + C$
$\sin x$	$-\cos x + C$
$\cos x$	$\sin x + C$
$\sec^2 x = \dfrac{1}{\cos^2 x}$	$\tan x + C$
$\dfrac{1}{\sqrt{1-x^2}} \quad (\lvert x\rvert < 1)$	$\arcsin x + C$
$\dfrac{1}{1+x^2}$	$\arctan x + C$
$\dfrac{1}{\sqrt{x^2 \pm 1}} \quad (x^2 \pm 1 > 0)$	$\ln\left\lvert x + \sqrt{x^2 \pm 1}\right\rvert + C$

Table 13.1: Table of Standard Primitives.

The linearity of the derivative implies the linearity of the integral.

***Proposition* 13.1.2.** *Linearity.* Let $f, g : I \to \mathbb{R}$ be functions admitting primitives, and let $\alpha, \beta \in \mathbb{R}$. Then:

$$\int (\alpha f(x) + \beta g(x))\, dx = \alpha \int f(x)\, dx + \beta \int g(x)\, dx$$

Proof. Let F and G be primitives of f and g respectively. By linearity of differentiation, $(\alpha F + \beta G)' = \alpha F' + \beta G' = \alpha f + \beta g$. Thus $\alpha F + \beta G$ is the required primitive. ■

Example 13.1.2. Polynomials. Using linearity and the power rule:

$$\int (3 + 5x + 7x^2)\, dx = 3x + \frac{5}{2}x^2 + \frac{7}{3}x^3 + C$$

Integration by Parts

The Product Rule for differentiation states that $(fg)' = f'g + fg'$. Integrating this relation allows us to exchange the differentiation from one function to another.

***Proposition* 13.1.3.** *Integration by Parts.* Let $f, g : I \to \mathbb{R}$ be differentiable functions with continuous derivatives. Then:

$$\int f(x)g'(x)\, dx = f(x)g(x) - \int g(x)f'(x)\, dx \tag{13.1}$$

Or, in the differential notation where $du = f'(x)dx$ and $dv = g'(x)dx$:

$$\int u\, dv = uv - \int v\, du$$

Proof. Integrating the product rule identity $f(x)g'(x) = (f(x)g(x))' - f'(x)g(x)$ immediately yields the result by linearity. ■

This technique is particularly effective when the integrand is a product of functions of different types (e.g., polynomials against exponentials or logarithms), or when dealing with inverse functions.

Example 13.1.3. The Logarithm. We view $\ln x$ as a product $1 \cdot \ln x$. Set $u = \ln x$ and $dv = 1\,dx$. Then $du = \frac{1}{x}dx$ and $v = x$.

$$\int \ln x\,dx = x\ln x - \int x \cdot \frac{1}{x}\,dx = x\ln x - \int 1\,dx = x\ln x - x + C$$

Example 13.1.4. Cyclic Integration. Consider the integrals $I = \int e^{ax}\cos x\,dx$ and $J = \int e^{ax}\sin x\,dx$. Applying parts to I with $u = e^{ax}, dv = \cos x dx$ yields:

$$I = e^{ax}\sin x - a\int e^{ax}\sin x\,dx = e^{ax}\sin x - aJ$$

Applying parts to J similarly yields:

$$J = -e^{ax}\cos x + aI$$

We now have a linear system for I and J. Solving it gives:

$$I = \frac{e^{ax}(\sin x + a\cos x)}{a^2+1} + C$$

Example 13.1.5. Reduction Formulae. Integration by parts often yields recurrence relations. Let $I_n = \int x^n e^x\,dx$ for integers $n \geq 0$. Set $u = x^n$ and $dv = e^x dx$. Then $du = nx^{n-1}dx$ and $v = e^x$.

$$I_n = x^n e^x - n\int x^{n-1}e^x\,dx = x^n e^x - nI_{n-1}$$

Starting from $I_0 = e^x$, we can generate the primitive for any n.

Integration by Substitution

Just as Integration by Parts is the reverse of the Product Rule, Substitution is the reverse of the Chain Rule.

***Proposition* 13.1.4.** *Change of Variables.* Let $\phi : I \to J$ be a differentiable function with continuous derivative, and let $f : J \to \mathbb{R}$ be continuous. Then:

$$\int f(\phi(x))\phi'(x)\,dx = F(\phi(x)) + C$$

where $F = \int f(u)du$. In differential notation, substituting $u = \phi(x)$ yields $du = \phi'(x)dx$, and:

$$\int f(u)\,du = F(u)$$

Proof. Differentiating the right-hand side with respect to x using the Chain Rule gives $F'(\phi(x))\phi'(x) = f(\phi(x))\phi'(x)$, which matches the integrand. ∎

Example 13.1.6. Standard Substitutions.

1. $\int e^{x^2} x\,dx$. Let $u = x^2$, so $du = 2x\,dx$ or $x\,dx = \frac{1}{2}du$.

$$\int e^{x^2} x\,dx = \frac{1}{2}\int e^u du = \frac{1}{2}e^{x^2} + C$$

2. $\int \tan x\,dx = \int \frac{\sin x}{\cos x}dx$. Let $u = \cos x$, so $du = -\sin x dx$.

$$\int \frac{\sin x}{\cos x}dx = -\int \frac{1}{u}du = -\ln|u| + C = \ln|\sec x| + C$$

3. **Trigonometric Powers**: To compute $\int(\sin x)^{2m+1}(\cos x)^n\,dx$, split off one sine term to form the differential of cosine. Let $u = \cos x$, $du = -\sin x dx$. We convert the remaining even power of sine using $\sin^2 x = 1 - \cos^2 x = 1 - u^2$.

Rational Functions and Partial Fractions

A rational function is a ratio of polynomials $R(x) = P(x)/Q(x)$. Since polynomials are easy to integrate, the difficulty lies in the denominator. The method of Partial Fraction Decomposition allows us to express any rational function as a sum of simpler terms ("simple fractions") whose primitives are known.

***Theorem* 13.1.1.** Partial Fraction Decomposition Any rational function $P(x)/Q(x)$ where $\deg(P) < \deg(Q)$ can be written as a sum of terms of the following forms:

1. $\dfrac{A}{(x-r)^k}$ where r is a real root of $Q(x)$.
2. $\dfrac{Bx+C}{((x-\alpha)^2+\beta^2)^k}$ corresponding to complex conjugate roots of $Q(x)$.

The integration of these terms proceeds as follows:

- Terms of type $\frac{1}{x-r}$ integrate to $\ln|x-r|$.
- Terms of type $\frac{1}{(x-r)^k}$ for $k > 1$ integrate to $\frac{-1}{(k-1)(x-r)^{k-1}}$.
- Quadratic terms are handled by completing the square and substituting $u = x - \alpha$, reducing them to forms involving arctan or recurrence relations (see below).

Example 13.1.7. Example of Decomposition. Consider $I = \int \frac{dx}{(x-1)(x^2+1)}$. We posit the decomposition:

$$\frac{1}{(x-1)(x^2+1)} = \frac{A}{x-1} + \frac{Bx+C}{x^2+1}$$

Clearing denominators: $1 = A(x^2+1) + (Bx+C)(x-1)$. Setting $x = 1$ gives $1 = 2A \implies A = 1/2$. Expanding coefficients for x^2 gives $0 = A + B \implies B = -1/2$. The constant term gives $1 = A - C \implies C = -1/2$. Thus:

$$I = \frac{1}{2}\int \frac{dx}{x-1} - \frac{1}{2}\int \frac{x}{x^2+1}dx - \frac{1}{2}\int \frac{1}{x^2+1}dx$$

$$I = \frac{1}{2}\ln|x-1| - \frac{1}{4}\ln(x^2+1) - \frac{1}{2}\arctan x + C$$

Example 13.1.8. Recurrence for Quadratic Denominators. Let $A_n = \int \frac{dx}{(x^2+1)^n}$. Using integration by parts with $u = (x^2+1)^{-n}$ and $dv = dx$:

$$A_n = \frac{x}{(x^2+1)^n} + 2n\int \frac{x^2}{(x^2+1)^{n+1}}dx$$

Writing $x^2 = (x^2+1) - 1$ in the numerator allows us to express the integral in terms of A_n and A_{n+1}:

$$A_n = \frac{x}{(x^2+1)^n} + 2nA_n - 2nA_{n+1}$$

Rearranging yields the reduction formula:

$$A_{n+1} = \frac{1}{2n}\frac{x}{(x^2+1)^n} + \frac{2n-1}{2n}A_n$$

Integration by Parts for Definite Integrals

Integration by parts extends naturally to definite integrals. A direct formulation via theorem 11.9.2 is convenient for reduction formulae.

***Proposition* 13.1.5.** *Definite Integration by Parts.* Let $u, v : [a, b] \to \mathbb{R}$ be continuously differentiable functions. Then:

$$\int_a^b u(x)v'(x)\,dx = \Big[u(x)v(x)\Big]_a^b - \int_a^b v(x)u'(x)\,dx \tag{13.2}$$

Proof. The product uv is continuously differentiable with derivative $(uv)' = u'v + uv'$. Since u' and v' are continuous, these derivatives are integrable. By theorem 11.9.2:

$$\int_a^b (u(x)v'(x) + u'(x)v(x))\,dx = u(b)v(b) - u(a)v(a)$$

Linearity of the integral allows us to rearrange terms to obtain eq. (13.2). ∎

Example 13.1.9. Polynomial Weights. For integers $m, n \geq 0$ set

$$I_{m,n} = \int_{-1}^{1} (x-1)^m (x+1)^n\,dx.$$

If $m = 0$, then

$$I_{0,n} = \int_{-1}^{1} (x+1)^n\,dx = \frac{2^{n+1}}{n+1}.$$

For $m > 0$, write $(x+1)^n = \frac{1}{n+1}\frac{d}{dx}(x+1)^{n+1}$ and integrate by parts to obtain

$$I_{m,n} = -\frac{m}{n+1} I_{m-1,n+1}.$$

Iterating gives

$$I_{m,n} = (-1)^m \frac{m!}{(n+1)(n+2)\cdots(n+m)} I_{0,n+m} = (-1)^m \frac{2^{m+n+1}}{n+m+1} \frac{1}{\binom{n+m}{m}}.$$

Equivalently,

$$I_{m,n} = (-1)^m \frac{2^{m+n+1} m! n!}{(m+n+1)!}.$$

In particular,

$$\int_{-1}^{1} (x^2-1)^n\,dx = (-1)^n \frac{2^{2n+1}}{2n+1} \frac{1}{\binom{2n}{n}}.$$

Example 13.1.10. Wallis Integrals and Product. Let $I_n = \int_0^{\pi/2} \sin^n x\,dx$. Using parts with $u = \sin^{n-1} x$ and $dv = \sin x\,dx$ (for $n \geq 2$):

$$I_n = \left[-\sin^{n-1} x \cos x\right]_0^{\pi/2} + (n-1)\int_0^{\pi/2} \sin^{n-2} x \cos^2 x\,dx$$

The boundary terms vanish. Using $\cos^2 x = 1 - \sin^2 x$, we find:

$$I_n = (n-1)I_{n-2} - (n-1)I_n \implies I_n = \frac{n-1}{n} I_{n-2}$$

Starting from $I_0 = \pi/2$ and $I_1 = 1$, we obtain explicit formulae:

$$I_{2n} = \frac{(2n-1)!!}{(2n)!!}\frac{\pi}{2}, \qquad I_{2n+1} = \frac{(2n)!!}{(2n+1)!!}$$

Since $\sin x \in [0,1]$, the sequence I_n is monotonic decreasing. Thus $I_{2n+1} < I_{2n} < I_{2n-1}$. Dividing by I_{2n+1} and noting that $\lim_{n\to\infty} I_{2n-1}/I_{2n+1} = 1$, we derive the Wallis Product:

$$\frac{\pi}{2} = \lim_{n\to\infty} \left(\frac{2 \cdot 4 \cdot \cdots \cdot (2n)}{1 \cdot 3 \cdot \cdots \cdot (2n-1)}\right)^2 \frac{1}{2n+1} = \prod_{n=1}^{\infty} \frac{4n^2}{4n^2-1}$$

The Wallis product provides the machinery to determine the asymptotic constant in Stirling's approximation for $n!$.

***Theorem* 13.1.2. Stirling's Formula**. As $n \to \infty$,

$$n! \sim \sqrt{2\pi n}\left(\frac{n}{e}\right)^n$$

That is, the ratio of the two sides tends to 1.

Proof. Consider the logarithm of the factorial, $\ln n! = \sum_{k=1}^{n} \ln k$. We compare this sum to the integral $I_n = \int_1^n \ln x\, dx = n \ln n - n + 1$. Using the Trapezoidal Rule on each interval $[k-1, k]$ and the concavity of $\ln x$, the difference between the sum and the integral converges to a constant. Specifically, let $d_n = \ln n! - (n + \frac{1}{2}) \ln n + n$. Then d_n decreases to a limit C. Exponentiating, we find $n! \sim e^C \sqrt{n}(n/e)^n$. To determine e^C, we substitute this approximation into the Wallis product limit derived above. The algebra yields $e^C = \sqrt{2\pi}$. ■

13.2 Taylor's Theorem with Integral Remainder

We previously encountered Taylor's Theorem with the Lagrange form of the remainder (theorem 10.2.1). Integration by parts allows us to derive an exact integral expression for the error term, assuming higher regularity of the function.

***Proposition* 13.2.1.** *Integral Remainder.* Let $f : [a, b] \to \mathbb{R}$ be $n+1$ times continuously differentiable. Let $x_0, x \in [a, b]$. Then:

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

where the remainder is given by:

$$R_n(x) = \frac{1}{n!}\int_{x_0}^{x} f^{(n+1)}(t)(x-t)^n\, dt \tag{13.3}$$

Proof. We proceed by induction using integration by parts. For $n = 0$, the statement is theorem 11.9.1: $f(x) - f(x_0) = \int_{x_0}^{x} f'(t)dt$. Assume the formula holds for $n - 1$:

$$R_{n-1}(x) = \frac{1}{(n-1)!}\int_{x_0}^{x} f^{(n)}(t)(x-t)^{n-1}\, dt$$

Integrate by parts with $u = f^{(n)}(t)$ and $dv = (x-t)^{n-1}dt$. Note that $v = -\frac{(x-t)^n}{n}$.

$$\int_{x_0}^{x} f^{(n)}(t)(x-t)^{n-1}\, dt = \left[-f^{(n)}(t)\frac{(x-t)^n}{n}\right]_{x_0}^{x} + \frac{1}{n}\int_{x_0}^{x} f^{(n+1)}(t)(x-t)^n\, dt$$

The boundary term at x vanishes; the term at x_0 provides the next Taylor coefficient. The integral term becomes $n!R_n(x)$. ■

Example 13.2.1. Logarithmic Series. Let $f(x) = \ln(1-x)$ for $x < 1$. The n-th derivative is $f^{(n)}(t) = -(n-1)!(1-t)^{-n}$. For $x_0 = 0$, the integral remainder is:

$$R_n(x) = -\int_0^x \frac{(x-t)^n}{(1-t)^{n+1}}\, dt$$

For $|x| < 1$, one can bound this integral to show $R_n(x) \to 0$, establishing the convergence of the Taylor series $\ln(1-x) = -\sum x^k/k$. The integral form is particularly useful for analysing convergence at the endpoints (e.g., $x = -1$) where Lagrange's form is less tractable.

Change of Variables in the Riemann Integral

While substitution is a standard tool for finding primitives, its validity for definite Riemann integrals requires careful handling of the domain and the regularity of the transformation.

***Proposition* 13.2.2.** *Change of Variables (Continuous).* Let $f : [a, b] \to \mathbb{R}$ be continuous and let $\phi : [\alpha, \beta] \to [a, b]$ be continuously differentiable. Then:

$$\int_{\alpha}^{\beta} f(\phi(t))\phi'(t)\, dt = \int_{\phi(\alpha)}^{\phi(\beta)} f(x)\, dx \tag{13.4}$$

Proof. Since f is continuous, it possesses a primitive F on $[a, b]$ by theorem 11.9.1. Let $G(t) = F(\phi(t))$. By the Chain Rule, $G'(t) = F'(\phi(t))\phi'(t) = f(\phi(t))\phi'(t)$. Applying theorem 11.9.2 to both sides establishes the equality. ■

However, we often require this theorem for functions that are merely Riemann integrable, not necessarily continuous.

***Proposition* 13.2.3.** *Change of Variables (Monotone).* Let $f : [a, b] \to \mathbb{R}$ be Riemann integrable. Let $\phi : [\alpha, \beta] \to [a, b]$ be a continuously differentiable bijection with $\phi'(t) > 0$ on (α, β) (i.e., ϕ is strictly increasing). Then $(f \circ \phi) \cdot \phi'$ is Riemann integrable on $[\alpha, \beta]$ and eq. (13.4) holds.

Proof. Assume ϕ is strictly increasing (so $\phi' \geq 0$). Let $P = \{t_0, \dots, t_n\}$ be a partition of $[\alpha, \beta]$. Then $Q = \{\phi(t_0), \dots, \phi(t_n)\}$ forms a partition of $[a, b]$. By theorem 9.9.2 applied to ϕ, $\Delta x_k = \phi(t_k) - \phi(t_{k-1}) = \phi'(c_k)\Delta t_k$ for some $c_k \in (t_{k-1}, t_k)$. We can interpret a Riemann sum for the transformed integral using the tags c_k:

$$S((f \circ \phi)\phi', P, \{c_k\}) = \sum f(\phi(c_k))\phi'(c_k)\Delta t_k = \sum f(\xi_k)\Delta x_k$$

where $\xi_k = \phi(c_k)$. This sum resembles a Riemann sum for $\int f$. Careful management of the difference between arbitrary tags and these specific MVT tags allows one to prove the equality of limits. ■

13.3 Improper Integrals

The Riemann integral is defined for bounded functions on compact intervals. We extend it to infinite domains or singularities via limits.

Definition 13.3.1. ***Improper Integral***. Let $f : [a, \omega) \to \mathbb{R}$. Suppose f is Riemann integrable on $[a, x]$ for every $a < x < \omega$. We define the improper integral as:

$$\int_a^{\omega} f(t)\, dt = \lim_{x \to \omega^-} \int_a^x f(t)\, dt$$

provided the limit exists (finite). In this case, we say the integral *converges*. If the limit is $\pm\infty$ or does not exist, the integral *diverges*. This definition applies whether $\omega = \infty$ (infinite interval) or $\omega < \infty$ (finite endpoint where f may be unbounded).

Example 13.3.1. p-Integrals.

1. **Infinite Domain:** $\int_1^{\infty} \frac{1}{x^p} dx$.

$$\int_1^R x^{-p} dx = \begin{cases} [\ln x]_1^R & p = 1 \\ \left[\frac{x^{1-p}}{1-p}\right]_1^R & p \neq 1 \end{cases}$$

   The limit as $R \to \infty$ exists if and only if $p > 1$.
2. **Singularity at Origin:** $\int_0^1 \frac{1}{x^p} dx$. Here the limit as $\epsilon \to 0^+$ exists if and only if $p < 1$.

The linearity of the limit implies that improper integrals are linear. However, dealing with convergence directly via primitives is not always possible. We rely on comparison tests, analogous to those for series.

***Theorem* 13.3.1. Cauchy Criterion for Integrals**. Let $f : [a, \omega) \to \mathbb{R}$. The integral $\int_a^\omega f$ converges if and only if for every $\epsilon > 0$, there exists $c \in (a, \omega)$ such that for all $y > z > c$:

$$\left| \int_z^y f(t)\, dt \right| < \epsilon$$

***Corollary* 13.3.1.** *Comparison Test.* Suppose $0 \le f(x) \le g(x)$ for all $x \in [a, \omega)$.

1. If $\int_a^\omega g$ converges, then $\int_a^\omega f$ converges.
2. If $\int_a^\omega f$ diverges, then $\int_a^\omega g$ diverges.

***Corollary* 13.3.2.** *Limit Comparison Test.* Suppose $f, g \ge 0$ and $\lim_{x\to\omega} \frac{f(x)}{g(x)} = L \in (0, \infty)$. Then $\int_a^\omega f$ and $\int_a^\omega g$ converge or diverge together.

Definition 13.3.2. *Absolute Convergence*. An improper integral $\int_a^\omega f$ is absolutely convergent if $\int_a^\omega |f|$ converges. Absolute convergence implies convergence (by the Cauchy Criterion and the inequality $|\int f| \le \int |f|$).

Example 13.3.2. Gaussian Integral. Consider $\int_0^\infty e^{-x^2} dx$. There is no elementary primitive. We compare with e^{-x}. For $x \ge 1$, $x^2 \ge x$, so $e^{-x^2} \le e^{-x}$. Since $\int_1^\infty e^{-x} dx = 1$ converges, $\int_1^\infty e^{-x^2} dx$ converges. The part on $[0, 1]$ is a standard Riemann integral. Thus the Gaussian integral converges.

13.4 Euler's Gamma Function

The theory of improper integrals allows us to define one of the most important non-elementary functions in analysis: the Gamma function. It extends the factorial function from natural numbers to real (and complex) numbers.

Definition 13.4.1. *Gamma Function*. For $x > 0$, we define the Gamma function $\Gamma(x)$ by the improper integral:

$$\Gamma(x) = \int_0^\infty t^{x-1} e^{-t}\, dt$$

We must first ensure this definition is valid. The integrand $f(t) = t^{x-1}e^{-t}$ has a potential singularity at $t = 0$ if $x < 1$, and an infinite domain.

***Proposition* 13.4.1.** *Properties of the Gamma Function.*

1. The integral defining $\Gamma(x)$ converges for all $x > 0$.
2. **Functional Equation:** $\Gamma(x + 1) = x\Gamma(x)$ for all $x > 0$.
3. **Factorial Property:** For $n \in \mathbb{N}$, $\Gamma(n) = (n - 1)!$.
4. **Scaling:** For $\lambda > 0$, $\displaystyle\int_0^\infty s^{x-1} e^{-\lambda s}\, ds = \frac{\Gamma(x)}{\lambda^x}$.

Proof. **(1) Convergence:** We split the integral at $t = 1$. On $[1, \infty)$, $t^{x-1}e^{-t} \le Ce^{-t/2}$ for sufficiently large t (since polynomial growth is dominated by exponential decay). Since $\int_1^\infty e^{-t/2} dt$ converges, the tail converges by comparison. On $(0, 1]$, $0 < t^{x-1}e^{-t} < t^{x-1}$. The integral $\int_0^1 t^{x-1} dt$ converges if and only if $x - 1 > -1$, i.e., $x > 0$. Thus $\Gamma(x)$ is well-defined for $x > 0$.

(2) Recurrence: Fix $x > 0$ and consider $\Gamma(x + 1)$. We integrate by parts on a finite interval $[\epsilon, R]$ with $u = t^x$ and $dv = e^{-t} dt$:

$$\int_\epsilon^R t^x e^{-t} dt = \left[-t^x e^{-t}\right]_\epsilon^R + x \int_\epsilon^R t^{x-1} e^{-t} dt$$

As $\epsilon \to 0$, $\epsilon^x \to 0$ (since $x > 0$). As $R \to \infty$, $R^x e^{-R} \to 0$. Thus:

$$\Gamma(x+1) = 0 + x\int_0^\infty t^{x-1}e^{-t}dt = x\Gamma(x)$$

(3) Factorials: We proceed by induction. Base case: $\Gamma(1) = \int_0^\infty e^{-t}dt = 1 = 0!$. Inductive step: If $\Gamma(n) = (n-1)!$, then $\Gamma(n+1) = n\Gamma(n) = n(n-1)! = n!$.

(4) Scaling: Use the substitution $t = \lambda s$. Then $dt = \lambda ds$ and limits remain 0 to ∞.

$$\Gamma(x) = \int_0^\infty (\lambda s)^{x-1}e^{-\lambda s}(\lambda\, ds) = \lambda^x \int_0^\infty s^{x-1}e^{-\lambda s}\, ds$$

Dividing by λ^x yields the result. ∎

13.5 Exercises

1. **Mechanical Integration.** Evaluate the following indefinite integrals:
   (a) $\displaystyle\int \frac{dx}{1+e^x}$
   (b) $\displaystyle\int x \arctan x\, dx$
   (c) $\displaystyle\int \frac{dx}{\sqrt{x}+\sqrt[3]{x}}$
   Remark. Use the substitution $x = u^6$.
   (d) $\displaystyle\int \frac{1}{x^4+1}\, dx$
   Remark. Factor the denominator as $(x^2+\sqrt{2}x+1)(x^2-\sqrt{2}x+1)$.

2. **Piecewise Primitives.** Let $f(x) = |x|$ and define $F(x) = \int_{-1}^x f(t)\, dt$.
   (a) Find a piecewise algebraic formula for $F(x)$ valid for all $x \in \mathbb{R}$.
   (b) Determine the points where F is differentiable and compute $F'(x)$. Does $F'(x) = f(x)$ everywhere?
   (c) Repeat the analysis for the function $g(x) = 1$ if $x < 0$ and $g(x) = 2$ if $x \geq 0$. Where does the second derivative of the primitive fail to exist?

3. **Limits of Sums.** Recognise the following limits as definite integrals and evaluate them:
   (a) $\displaystyle\lim_{n\to\infty} \frac{1}{n}\sum_{k=1}^n \left(\frac{k}{n}\right)^3$
   (b) $\displaystyle\lim_{n\to\infty} \sum_{k=1}^n \frac{n}{n^2+k^2}$
   (c) $\displaystyle\lim_{n\to\infty} \frac{1}{n}\sum_{k=1}^n \sin\left(\frac{k\pi}{n}\right)$

4. **Reduction Formulae.** Let $I_n = \int_0^{\pi/2} \sin^n x\, dx$.
   (a) Use integration by parts to show $I_n = \frac{n-1}{n} I_{n-2}$ for $n \geq 2$.
   (b) Evaluate I_6 and I_7.
   (c) **Wallis' Product:** Use the inequalities $\sin^{2n+1} x \leq \sin^{2n} x \leq \sin^{2n-1} x$ on $[0, \pi/2]$ to prove that

   $$\frac{\pi}{2} = \lim_{n\to\infty} \left[\frac{2\cdot 4\cdots(2n)}{1\cdot 3\cdots(2n-1)}\right]^2 \frac{1}{2n+1}.$$

5. **Properties of the Primitive.** Decide whether each statement is true or false, providing a short proof or counterexample.

(a) If $g = h'$ for some function h on $[a,b]$, then g must be continuous on $[a,b]$.
(b) If g is continuous on $[a,b]$, then $g = h'$ for some function h on $[a,b]$.
(c) If $H(x) = \int_a^x h(t)\,dt$ is differentiable at $c \in [a,b]$, then h must be continuous at c.
(d) If f is continuous on $[a,b]$ and $\int_a^x f(t)\,dt = 0$ for all $x \in [a,b]$, then $f(x) = 0$ everywhere on $[a,b]$.

6. **Integration by Parts: Theory.**
   (a) Assume h and k have continuous derivatives on $[a,b]$. Derive the integration-by-parts formula:

$$\int_a^b h(t)k'(t)\,dt = h(b)k(b) - h(a)k(a) - \int_a^b h'(t)k(t)\,dt.$$

   (b) Suppose h is merely differentiable (so h' exists but may not be continuous) but h' is Riemann integrable. Does the formula still hold? Justify your answer using the Product Rule and the Fundamental Theorem of Calculus.

7. **Change of Variable Proof.** Let $g : [a,b] \to \mathbb{R}$ be differentiable with g' continuous. Let $f : [c,d] \to \mathbb{R}$ be continuous, with $g([a,b]) \subseteq [c,d]$.
   (a) Let $F(x) = \int_c^x f(t)\,dt$. Show that the composite function $H(x) = F(g(x))$ is an antiderivative of $f(g(x))g'(x)$.
   (b) Apply the Fundamental Theorem of Calculus to H to prove the change-of-variable formula:

$$\int_a^b f(g(x))g'(x)\,dx = \int_{g(a)}^{g(b)} f(t)\,dt.$$

8. **Total Variation.** Given a function f on $[a,b]$, the total variation is defined as $Vf = \sup_P \sum |f(x_k) - f(x_{k-1})|$ over all partitions P.
   (a) If f is continuously differentiable, use the Fundamental Theorem to show $f(x_k) - f(x_{k-1}) = \int_{x_{k-1}}^{x_k} f'(t)dt$.
   (b) Deduce that $Vf \le \int_a^b |f'(t)|dt$.
   (c) Use the Mean Value Theorem on the partition intervals to establish the reverse inequality, concluding that for C^1 functions, the total variation is the integral of the absolute derivative.

9. **Jump Discontinuities and Differentiability.** Assume f is integrable on $[a,b]$ but has a simple jump discontinuity at $c \in (a,b)$ (i.e., the left and right limits exist but differ).
   (a) Show that the accumulation function $F(x) = \int_a^x f(t)\,dt$ is continuous at c but not differentiable at c.
   (b) Calculate the left and right derivatives of F at c.

10. **Euler's Constant.** Let $L(x) = \int_1^x \frac{1}{t}\,dt$ for $x > 0$.
    (a) Prove the identity $L(xy) = L(x) + L(y)$ using substitution $u = xt$ in the integral definition.
    (b) Let $\gamma_n = \sum_{k=1}^n \frac{1}{k} - L(n)$. Interpret $L(n)$ as the area under the hyperbola $y = 1/x$ and the sum as the area of rectangles circumscribing the curve.
    (c) Prove that the sequence (γ_n) is decreasing and bounded below by 0.
    (d) Conclude that the sequence converges to a limit γ (Euler's Constant).

11. **Convergence of Improper Integrals.** Determine the values of $p \in \mathbb{R}$ for which the following integrals converge:
    (a) $\displaystyle\int_0^1 x^p \ln x\,dx$

    Remark. Integration by parts is useful.

    (b) $\displaystyle\int_0^\infty \frac{x^{p-1}}{1+x}\,dx$

    Remark. Split the integral at $x = 1$ and check behaviour at 0 and ∞.

12. **Gamma Function Identities.** Using the definition $\Gamma(x) = \int_0^\infty t^{x-1}e^{-t}dt$:

(a) Prove that $\Gamma(1/2) = \sqrt{\pi}$.

Remark. Use the substitution $t = u^2$ and the known Gaussian integral $\int_{-\infty}^{\infty} e^{-u^2} du = \sqrt{\pi}$.

(b) Show that $\int_0^1 (\ln(1/x))^{n-1} dx = \Gamma(n)$ for $n > 0$.

13. Irrationality of π. Let $f(x) = \frac{x^n(1-x)^n}{n!}$.

(a) Show that $f(x)$ and all its derivatives take integer values at $x = 0$ and $x = 1$.
(b) Suppose $\pi^2 = a/b$ for integers a, b. Consider the polynomial $P(x) = b^n \sum_{k=0}^{n} (-1)^k \pi^{2n-2k} f^{(2k)}(x)$.
(c) Show that $\frac{d}{dx}[P'(x)\sin(\pi x) - P(x)\pi\cos(\pi x)] = \pi^{2n+2} f(x)\sin(\pi x)$.
(d) Integrate from 0 to 1 to derive a contradiction for large n, proving π^2 (and hence π) is irrational.

14. Continuous Nowhere Differentiable Function. Define $f(x) = \sum_{n=0}^{\infty} \left(\frac{3}{4}\right)^n \phi(4^n x)$, where $\phi(x)$ is the distance from x to the nearest integer.

(a) Prove that the series converges uniformly on $\mathbb{R}$, hence f is continuous.
(b) Fix x_0. For each m, define $h_m = \pm\frac{1}{2}4^{-m}$ (choosing the sign such that no integer lies strictly between $4^m x_0$ and $4^m(x_0 + h_m)$).
(c) Show that the difference quotient $\frac{f(x_0+h_m)-f(x_0)}{h_m}$ diverges as $m \to \infty$.
(d) Conclude f is nowhere differentiable.

Chapter 14

Geometric Applications of the Integral

The definition of the Riemann integral was motivated by the problem of quadrature: assigning a value to the area of a region under a curve. Having established the integral and its fundamental theorems, we return to geometry to compute areas of planar regions and volumes of solids.

14.1 Area of Plane Regions

We restrict our attention to subsets of $\mathbb{R}^2$ with continuous boundary curves. The area of such regions can be computed by slicing them into thin strips, either vertically (integration with respect to x) or horizontally (integration with respect to y).

Regions of Simple Type

Definition 14.1.1. ***Simple Region (Type I)***. A region $D \subseteq \mathbb{R}^2$ is said to be of *simple type with respect to the x-axis* (or a Type I region) if there exists a closed interval $[a, b]$ and continuous functions $f, g : [a, b] \to \mathbb{R}$ such that $g(x) \leq f(x)$ for all $x \in [a, b]$, and:

$$D = \left\{(x, y) \in \mathbb{R}^2 : a \leq x \leq b, \quad g(x) \leq y \leq f(x)\right\}$$

The area of D is defined as:

$$\text{Area}(D) = \int_a^b \big(f(x) - g(x)\big)\, dx$$

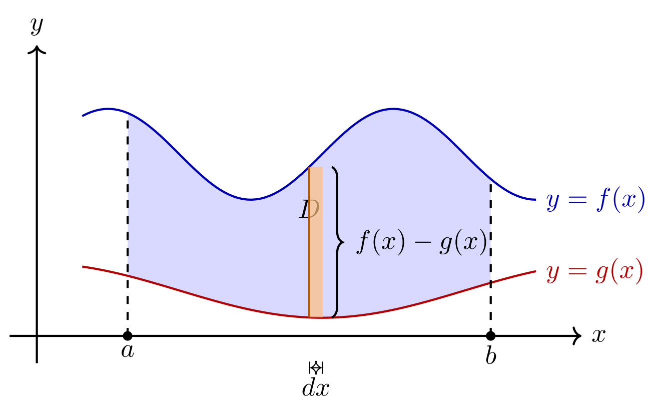


Figure 14.1: A region of simple type with respect to the x-axis (Type I).

Definition 14.1.2. ***Simple Region (Type II)***. A region $D \subseteq \mathbb{R}^2$ is said to be of *simple type with respect to the y-axis* (or a Type II region) if there exists a closed interval $[c, d]$ and continuous functions $\phi, \psi : [c, d] \to \mathbb{R}$ such that $\psi(y) \le \phi(y)$ for all $y \in [c, d]$, and:

$$D = \{(x, y) \in \mathbb{R}^2 : c \le y \le d, \quad \psi(y) \le x \le \phi(y)\}$$

The area of D is defined as:

$$\text{Area}(D) = \int_c^d \big(\phi(y) - \psi(y)\big)\, dy$$

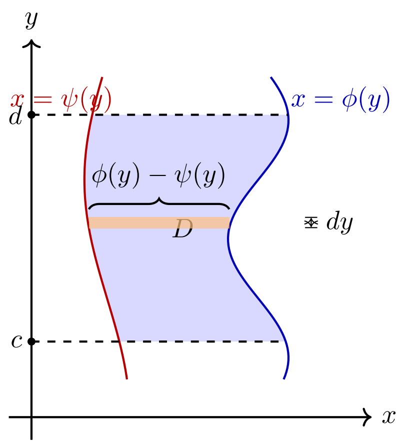


Figure 14.2: A region of simple type with respect to the y-axis (Type II).

Remark. Many regions are simple with respect to both axes in the sense of definitions 14.1.1 and 14.1.2, for example the unit square $[0, 1] \times [0, 1]$. In such cases the choice of variable is a matter of convenience, but algebraic simplicity often dictates a preferred direction.

Calculating Areas Between Curves

When a region is bounded by two curves $y_1 = f(x)$ and $y_2 = g(x)$, the first step is to identify the interval $[a, b]$ by determining the points of intersection, i.e., the roots of $f(x) - g(x) = 0$. By Simple Region (Type I) (definition 14.1.1), if $f \ge g$ on $[a, b]$, then the area is $\int_a^b (f - g)\, dx$.

Example 14.1.1. Parabolic Segment. Determine the area of the region bounded by the curve $y = 6 - x - x^2$ and the x-axis.

Here, $f(x) = 6 - x - x^2$ and $g(x) = 0$. We find the intersection points:

$$6 - x - x^2 = 0 \iff -(x^2 + x - 6) = 0 \iff -(x + 3)(x - 2) = 0$$

The roots are $x = -3$ and $x = 2$. Between these roots, test a point (e.g., $x = 0$): $f(0) = 6 > 0$, so the parabola lies above the axis.

$$\text{Area} = \int_{-3}^{2} (6 - x - x^2)\, dx = \left[6x - \frac{x^2}{2} - \frac{x^3}{3}\right]_{-3}^{2}$$

Evaluating at the endpoints:

$$F(2) = 12 - 2 - \frac{8}{3} = 10 - \frac{8}{3} = \frac{22}{3}$$

$$F(-3) = -18 - \frac{9}{2} - \frac{-27}{3} = -18 - 4.5 + 9 = -13.5 = -\frac{27}{2}$$

$$\text{Area} = \frac{22}{3} - \left(-\frac{27}{2}\right) = \frac{44 + 81}{6} = \frac{125}{6}$$

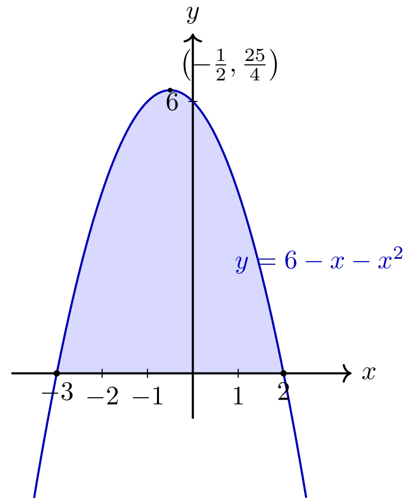


Figure 14.3: Parabolic segment bounded by $y = 6 - x - x^2$ and the x-axis.

Example 14.1.2. Total vs Signed Area. Find the area of the region bounded by $y = \sin x$ and the x-axis for $x \in [0, 2\pi]$.

The definite integral $\int_0^{2\pi} \sin x\, dx = [-\cos x]_0^{2\pi} = -1 - (-1) = 0$. This is the *signed* area. The geometric area is the integral of the absolute value:

$$A = \int_0^{2\pi} |\sin x|\, dx = \int_0^{\pi} \sin x\, dx + \int_{\pi}^{2\pi} (-\sin x)\, dx$$

$$A = [-\cos x]_0^{\pi} - [-\cos x]_{\pi}^{2\pi} = (1 - (-1)) - (-1 - 1) = 2 - (-2) = 4$$

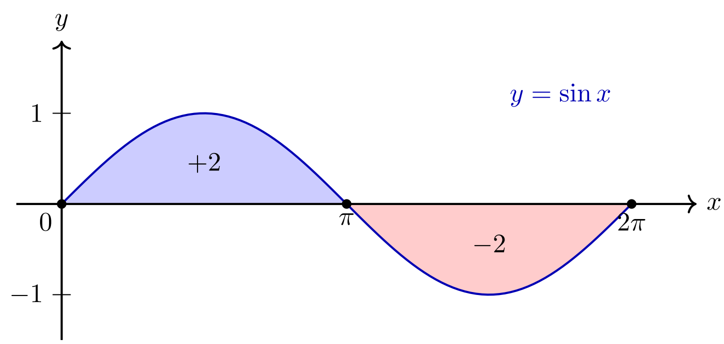


Figure 14.4: Signed and geometric area for $y = \sin x$ on $[0, 2\pi]$. The positive region (blue) has area $+2$, the negative region (red) has area -2. Signed area $= 0$; geometric area $= 4$.

Example 14.1.3. Intersecting Curves. Find the area of the region bounded by the line $y = -x$ and the parabola $y = 2 - x^2$.

Intersection points:

$$-x = 2 - x^2 \implies x^2 - x - 2 = 0 \implies (x - 2)(x + 1) = 0$$

The curves intersect at $x = -1$ and $x = 2$. To determine the relative position, consider $x = 0$: $y_{line} = 0$, $y_{parabola} = 2$. Thus $2 - x^2 \geq -x$ on $[-1, 2]$.

$$A = \int_{-1}^{2} \left[(2 - x^2) - (-x)\right]\, dx = \int_{-1}^{2} (2 + x - x^2)\, dx$$

$$A = \left[2x + \frac{x^2}{2} - \frac{x^3}{3}\right]_{-1}^{2} = \left(4 + 2 - \frac{8}{3}\right) - \left(-2 + \frac{1}{2} + \frac{1}{3}\right) = \frac{9}{2}$$

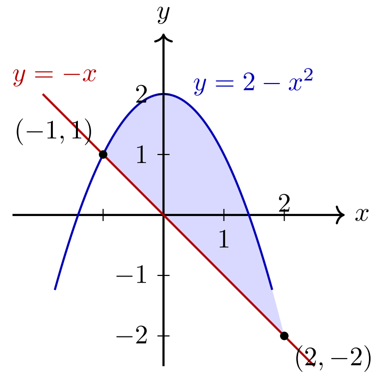


Figure 14.5: Region bounded by $y = -x$ and $y = 2 - x^2$.

Integration with Respect to y Sometimes a region is not Type I but is Type II, or a Type I description forces splitting. In such cases, Type II descriptions give a cleaner integral.

Example 14.1.4. Region Between a Line and a Parabola. Calculate the area of the region bounded by $x = y^2$ and $x = y + 2$.

The curves intersect when $y^2 = y + 2$, so $y = -1$ and $y = 2$. For $y \in [-1, 2]$, the line lies to the right of the parabola, hence by definition 14.1.2:

$$A = \int_{-1}^{2} \left((y+2) - y^2\right) dy = \left[\frac{y^2}{2} + 2y - \frac{y^3}{3}\right]_{-1}^{2} = \frac{9}{2}$$

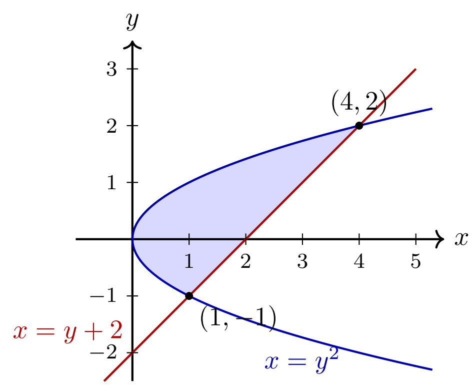


Figure 14.6: Region bounded by $x = y^2$ and $x = y + 2$.

Compound Regions Some regions require splitting into disjoint sub-regions, each of simple type.

Example 14.1.5. Piecewise Boundaries. Find the total area of the regions bounded by the line $5y = x + 6$ and the curve $y = \sqrt{|x|}$.

We treat $x \geq 0$ and $x < 0$ separately. For $x \geq 0$, the curve is $y = \sqrt{x}$ and the line is $y = (x + 6)/5$. Solving $\sqrt{x} = \frac{x+6}{5}$ gives $x = 4, 9$. For $x < 0$, the curve is $y = \sqrt{-x}$. Solving $\sqrt{-x} = \frac{x+6}{5}$ gives $x = -1$. Thus, the intersections occur at $x = -1, 4, 9$. On $[-1, 0]$ and $[0, 4]$ the line lies above the curve, while on $[4, 9]$ the curve lies above the line.

$$A = \int_{-1}^{0} \left(\frac{x+6}{5} - \sqrt{-x}\right) dx + \int_{0}^{4} \left(\frac{x+6}{5} - \sqrt{x}\right) dx + \int_{4}^{9} \left(\sqrt{x} - \frac{x+6}{5}\right) dx$$

Evaluating these integrals yields the total area.

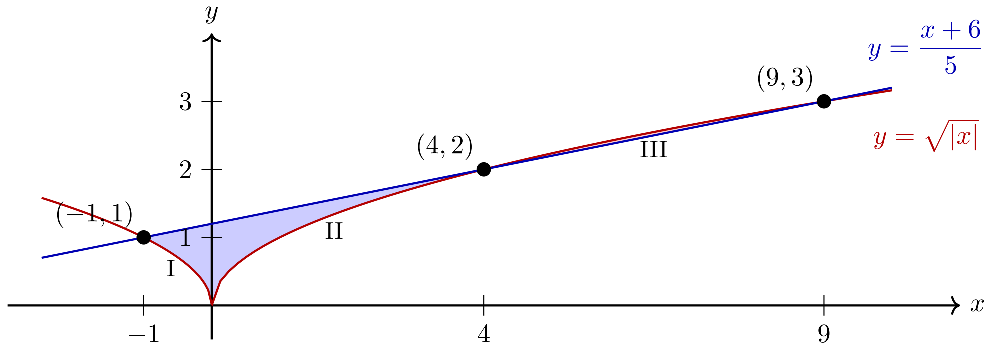


Figure 14.7: Regions bounded by $y = \sqrt{|x|}$ and $y = (x+6)/5$.

***Proposition* 14.1.1.** *Integrability Condition.* If the boundary function is unbounded or undefined at a point, the area must be treated as an improper integral. For instance, the area under $y = x^{-2}$ on $[-1, 1]$ is not given by the formal computation $\int_{-1}^{1} x^{-2} dx = [-x^{-1}]_{-1}^{1}$. The singularity at $x = 0$ requires splitting the integral, and the resulting improper integrals diverge.

14.2 Volumes by Slicing

We now determine volumes of solids. The method extends area computation by approximating the solid with thin slices and passing to a limit.

Consider a solid object S that lies between two parallel planes, which we may take to be perpendicular to the x-axis at $x = a$ and $x = b$. If we slice the solid with a plane perpendicular to the x-axis at a specific point $x \in [a, b]$, we obtain a planar region called the *cross-section* at x. Let $A(x)$ denote the area of this cross-section. If we partition the interval $[a, b]$ into small sub-intervals of width Δx, the volume of the slice of the solid corresponding to the sub-interval $[x_i, x_{i+1}]$ can be approximated by a cylinder (or prism) of base area $A(x_i)$ and height Δx. The total volume V is the limit of the sum of these cylindrical volumes.

Definition 14.2.1. ***Volume by Slicing***. Let S be a solid that lies between $x = a$ and $x = b$. If the cross-sectional area of S in the plane perpendicular to the x-axis at x is a continuous function $A(x)$, then the volume of S is:

$$V = \int_a^b A(x)\, dx$$

Similarly, if the solid is defined between $y = c$ and $y = d$ with cross-sectional area $A(y)$ perpendicular to the y-axis, then $V = \int_c^d A(y)\, dy$.

Remark. This formula is sometimes called Cavalieri's Principle in integral form. It reduces the problem of volume to the problem of determining the area function $A(x)$ and integrating it.

Solids with Known Cross-Sections

This method is particularly powerful for solids where the cross-sections are familiar geometric shapes (squares, circles, triangles).

Example 14.2.1. Volume of a Pyramid. Consider a pyramid with a square base of side length L and height H.

Orient the pyramid so that its apex is at the origin $(0, 0, 0)$ and its central axis lies along the x-axis. The pyramid extends from $x = 0$ to $x = H$. At a distance x from the apex, the cross-section is a square. By similar triangles, the side length $s(x)$ of the square at x is proportional to x. Since $s(H) = L$, we must have

$s(x) = \frac{L}{H}x$. The cross-sectional area is:

$$A(x) = (s(x))^2 = \left(\frac{Lx}{H}\right)^2 = \frac{L^2}{H^2}x^2$$

The volume is:

$$V = \int_0^H \frac{L^2}{H^2}x^2\,dx = \frac{L^2}{H^2}\left[\frac{x^3}{3}\right]_0^H = \frac{L^2}{H^2}\frac{H^3}{3} = \frac{1}{3}L^2H$$

This recovers the elementary formula $V = \frac{1}{3}\text{Base} \times \text{Height}$.

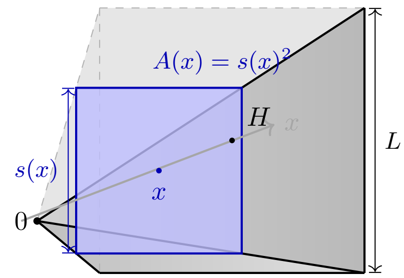


Figure 14.8: Pyramid with square cross-sections.

Example 14.2.2. The Curved Wedge. A wedge is cut from a circular cylinder of radius 3 by two planes. One plane is perpendicular to the axis of the cylinder. The second plane crosses the first at an angle of 45° along a diameter of the cylinder. Find the volume of the wedge.

Let the cylinder be defined by $x^2 + y^2 \leq 9$. The first plane is $z = 0$. The second plane passes through the x-axis (the diameter) at 45°, so its equation is $z = y$ (since $\tan 45^\circ = 1$). The base is the semicircle $y \geq 0$. Slice perpendicular to the x-axis. For a fixed $x \in [-3, 3]$, the cross-section is a triangle. The base of the triangle in the xy-plane extends from the x-axis ($y = 0$) to the circle boundary $y = \sqrt{9 - x^2}$. Let this length be $h(x) = \sqrt{9 - x^2}$. The height of the triangle is determined by the plane $z = y$. At the boundary $y = h(x)$, the height is $z = h(x)$. Thus the cross-section is a right-angled triangle with base $h(x)$ and height $h(x)$.

$$A(x) = \frac{1}{2} \cdot \text{base} \cdot \text{height} = \frac{1}{2}(\sqrt{9 - x^2})(\sqrt{9 - x^2}) = \frac{1}{2}(9 - x^2)$$

The volume is:

$$V = \int_{-3}^{3} \frac{1}{2}(9 - x^2)\,dx$$

By symmetry, this is $2\int_0^3 \frac{1}{2}(9 - x^2)\,dx = \int_0^3 (9 - x^2)\,dx$.

$$V = \left[9x - \frac{x^3}{3}\right]_0^3 = 27 - 9 = 18$$

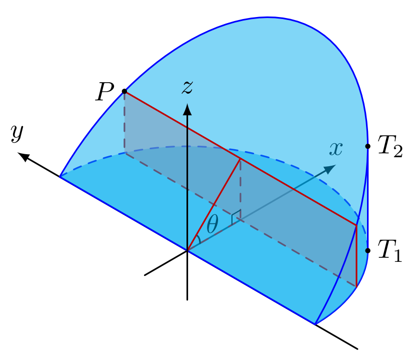


Figure 14.9: Curved wedge cut from a cylinder.

Example 14.2.3. Solid with Circular Cross-Sections. Find the volume of a solid lying between $x = -1$ and $x = 1$, where the cross-sections perpendicular to the x-axis are circular disks whose diameters run from the parabola $y = x^2$ to the parabola $y = 2 - x^2$.

First, determining the geometry of the base. The curves intersect when $x^2 = 2 - x^2 \implies 2x^2 = 2 \implies x = \pm 1$. The region is bounded between these two curves. For any $x \in [-1, 1]$, the diameter of the disk is the vertical distance between the curves:

$$D(x) = (2 - x^2) - (x^2) = 2 - 2x^2$$

The radius is $R(x) = \frac{D(x)}{2} = 1 - x^2$. The cross-sectional area is the area of the circle:

$$A(x) = \pi(R(x))^2 = \pi(1 - x^2)^2 = \pi(1 - 2x^2 + x^4)$$

The volume is:

$$V = \int_{-1}^{1} \pi(1 - 2x^2 + x^4)\, dx$$

Using symmetry (even function):

$$V = 2\pi \int_0^1 (1 - 2x^2 + x^4)\, dx = 2\pi \left[x - \frac{2}{3}x^3 + \frac{1}{5}x^5 \right]_0^1 = \frac{16\pi}{15}$$

Example 14.2.4. Solid with Square Cross-Sections. Find the volume of a solid between $y = -1$ and $y = 1$ (perpendicular to y-axis), where cross-sections are squares whose diagonals run from the semicircle $x = -\sqrt{1 - y^2}$ to $x = \sqrt{1 - y^2}$.

Let us fix a $y \in [-1, 1]$. The cross-section lies in the plane at y. The diagonal of the square spans the distance between the two semicircles.

$$\text{Diagonal } d(y) = \sqrt{1 - y^2} - (-\sqrt{1 - y^2}) = 2\sqrt{1 - y^2}$$

The area of a square with diagonal d is $A = \frac{1}{2}d^2$.

$$A(y) = \frac{1}{2}(2\sqrt{1 - y^2})^2 = \frac{1}{2}(4(1 - y^2)) = 2(1 - y^2)$$

The volume is:

$$V = \int_{-1}^{1} 2(1 - y^2)\, dy = 2\left[y - \frac{y^3}{3} \right]_{-1}^{1} = 2(4/3) = \frac{8}{3}$$

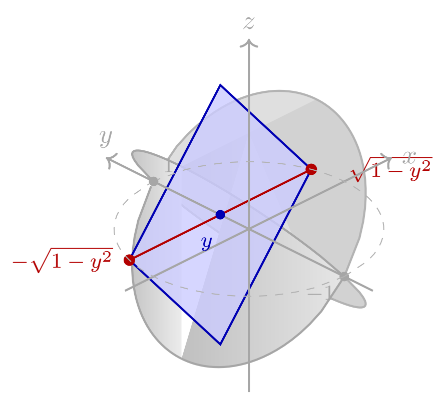


Figure 14.10: Square cross-sections with diagonal on a semicircle.

14.3 Solids of Revolution: The Disk Method

A particularly important class of solids is formed by revolving a planar region about a line. These are called *solids of revolution*. Their radial symmetry simplifies the general slicing method, leading to a specialised technique known as the Disk Method.

Concept and Formula

Consider a planar region bounded by the curve $y = f(x)$, the x-axis, and the vertical lines $x = a$ and $x = b$. If we revolve this region about the x-axis, we generate a solid. Slicing this solid perpendicular to the axis of revolution (the x-axis) at any point $x \in [a, b]$ produces a circular cross-section (a disk). The radius of this disk is the height of the curve at that point, $R(x) = f(x)$. The area of the cross-section is therefore $A(x) = \pi[f(x)]^2$.

Definition 14.3.1. ***The Disk Method***. Let f be continuous and non-negative on $[a, b]$. The volume V of the solid generated by revolving the region under $y = f(x)$ about the x-axis is:

$$V = \int_a^b \pi[f(x)]^2\,dx$$

If the revolution is about the y-axis, and the region is bounded by $x = g(y)$, $y = c$, and $y = d$, the volume is:

$$V = \int_c^d \pi[g(y)]^2\,dy$$

This formula is a direct application of the general slicing method $V = \int A(x)dx$ where $A(x)$ is the area of a circle.

Example 14.3.1. Paraboloid of Revolution. Find the volume of the solid generated by revolving the region bounded by $y = \sqrt{x}$, $x = 0$, and $x = 4$ about the x-axis.

Here $R(x) = \sqrt{x}$. The limits are 0 and 4.

$$V = \int_0^4 \pi(\sqrt{x})^2\,dx = \pi \int_0^4 x\,dx = \pi \left[\frac{x^2}{2}\right]_0^4 = \pi(8 - 0) = 8\pi$$

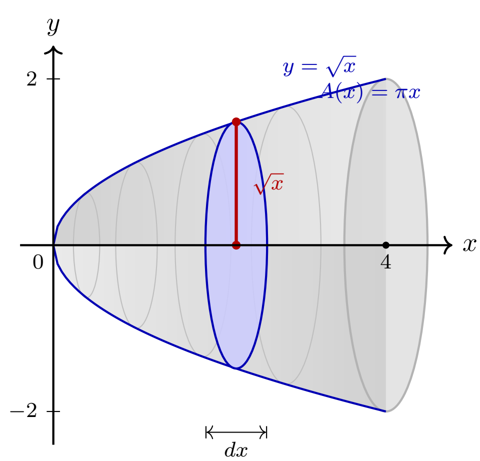


Figure 14.11: Paraboloid formed by revolving $y = \sqrt{x}$ about the x-axis.

Revolving About Other Lines

If the axis of revolution is not a coordinate axis, the radius of the disk is the perpendicular distance from the boundary curve to the axis of rotation.

Example 14.3.2. Shifted Axis. Find the volume of the solid generated by revolving the region bounded by $y = \sqrt{x}$, $y = 1$, and $x = 4$ about the line $y = 1$.

First, identify the region. The curve $y = \sqrt{x}$ intersects the line $y = 1$ at $x = 1$. The region extends to $x = 4$. The axis of revolution is the horizontal line $y = 1$. For any $x \in [1, 4]$, the radius of the disk is the vertical distance between the curve and the axis:

$$R(x) = \sqrt{x} - 1$$

$$V = \int_1^4 \pi(\sqrt{x} - 1)^2 \, dx = \pi \int_1^4 (x - 2\sqrt{x} + 1) \, dx = \pi \left[\frac{x^2}{2} - \frac{4}{3}x^{3/2} + x \right]_1^4 = \pi \left(\frac{4}{3} - \frac{1}{6} \right) = \frac{7\pi}{6}$$

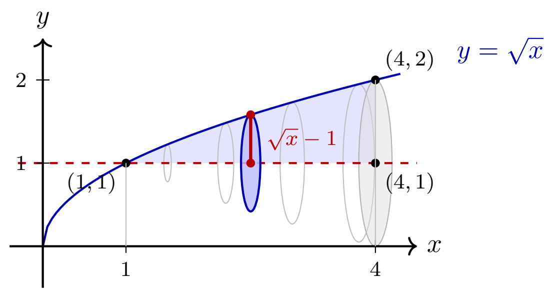


Figure 14.12: Region for revolution about $y = 1$.

Example 14.3.3. Volume of a Sphere. A sphere of radius a can be generated by revolving the semicircle $y = \sqrt{a^2 - x^2}$ (for $x \in [-a, a]$) about the x-axis.

The radius of the slice at x is $R(x) = \sqrt{a^2 - x^2}$.

$$V = \int_{-a}^{a} \pi(\sqrt{a^2 - x^2})^2 \, dx = \pi \int_{-a}^{a} (a^2 - x^2) \, dx$$

Using symmetry:

$$V = 2\pi \int_0^a (a^2 - x^2) \, dx = 2\pi \left[a^2 x - \frac{x^3}{3} \right]_0^a = 2\pi \left(a^3 - \frac{a^3}{3} \right) = 2\pi \left(\frac{2a^3}{3} \right) = \frac{4}{3}\pi a^3$$

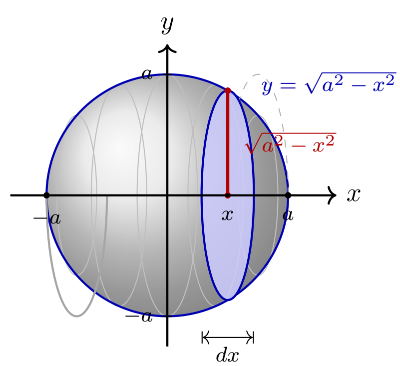


Figure 14.13: Sphere as a solid of revolution.

Disks vs. Washers

In the examples above, the axis of revolution formed a boundary of the region, resulting in solid disks. If the axis of revolution is separated from the region by a gap, the cross-sections become annuli (or "washers"). The area of such a cross-section is the area of the outer circle minus the area of the inner circle:

$$A(x) = \pi[R_{outer}(x)]^2 - \pi[R_{inner}(x)]^2$$

While this is often treated as a separate "Washer Method," it is simply the Disk Method applied to the region between two curves.

Example 14.3.4. Vertical Axis of Revolution. Find the volume generated by revolving the region bounded by $x = 2/y$, $y = 1$, and $y = 4$ about the y-axis.

The revolution is about the y-axis ($x = 0$). The radius is a function of y: $R(y) = 2/y$.

$$V = \int_1^4 \pi \left(\frac{2}{y}\right)^2 dy = 4\pi \int_1^4 y^{-2}\, dy = 4\pi \left[-y^{-1}\right]_1^4 = 4\pi \left(-\frac{1}{4} - (-1)\right) = 4\pi \left(\frac{3}{4}\right) = 3\pi$$

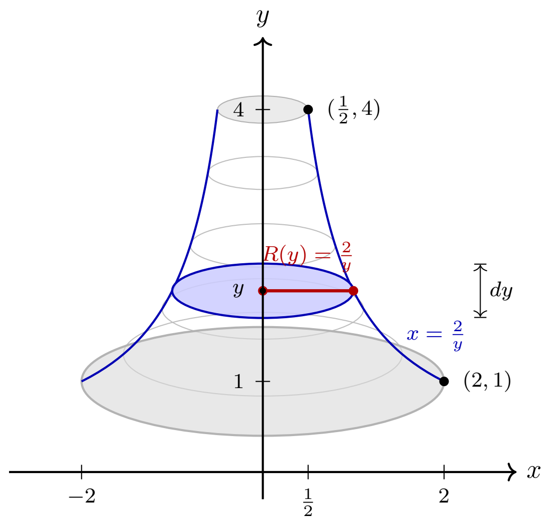


Figure 14.14: Region bounded by $x = 2/y$ and $y \in [1, 4]$.

Example 14.3.5. Implicit Boundaries. Find the volume generated by revolving the region bounded by $y = \tan(\pi x/4)$, $x = 0$, and $x = 1$ about the x-axis.

Here, $R(x) = \tan(\pi x/4)$. Limits are 0 and 1.

$$V = \int_0^1 \pi \tan^2 \left(\frac{\pi x}{4}\right) dx$$

Let $u = \frac{\pi x}{4}$, so $dx = \frac{4}{\pi} du$. Limits become 0 to $\pi/4$.

$$V = \pi \cdot \frac{4}{\pi} \int_0^{\pi/4} \tan^2 u\, du = 4 \int_0^{\pi/4} (\sec^2 u - 1)\, du = 4[\tan u - u]_0^{\pi/4} = 4\left((1 - \pi/4) - 0\right) = 4 - \pi$$

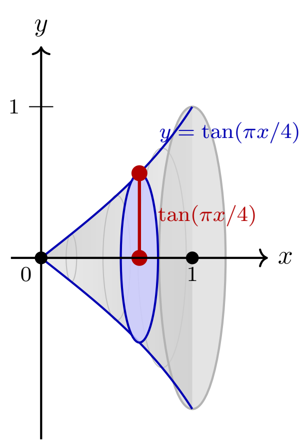


Figure 14.15: Region under $y = \tan(\pi x/4)$ revolved about the x-axis.

Example 14.3.6. Region about $y = \sqrt{2}$. Find the volume generated by revolving the region in the first quadrant bounded above by $y = \sqrt{2}$, below by $y = \sec x \tan x$, and on the left by the y-axis about the line $y = \sqrt{2}$.

The curve meets $y = \sqrt{2}$ at $x = \pi/4$, so $x \in [0, \pi/4]$. The radius is $R(x) = \sqrt{2} - \sec x \tan x$.

$$V = \int_0^{\pi/4} \pi(\sqrt{2} - \sec x \tan x)^2 \, dx$$

Expanding and integrating term by term gives

$$V = \pi \left[2x - 2\sqrt{2} \sec x + \frac{1}{3} \tan^3 x \right]_0^{\pi/4} = \pi \left(\frac{\pi}{2} + 2\sqrt{2} - \frac{11}{3} \right)$$

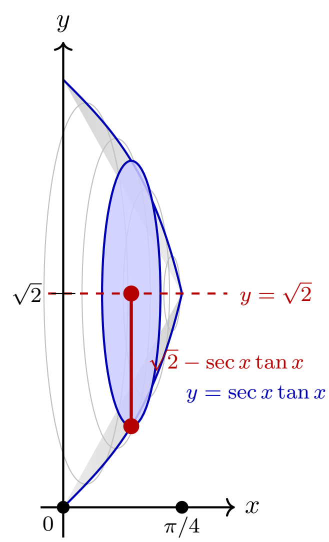


Figure 14.16: Region between $y = \sec x \tan x$ and $y = \sqrt{2}$, revolved about $y = \sqrt{2}$.

Example 14.3.7. Parabola about a Line. Find the volume generated by revolving the region between $y = x^2 + 1$ and $y = 3$ about the line $y = 3$.

The region is bounded above by $y = 3$ and below by the parabola. The axis of revolution is the upper boundary $y = 3$. Radius $R(x)$ is distance from $y = 3$ to $y = x^2 + 1$:

$$R(x) = 3 - (x^2 + 1) = 2 - x^2$$

Intersection points: $x^2 + 1 = 3 \implies x = \pm\sqrt{2}$.

$$V = \int_{-\sqrt{2}}^{\sqrt{2}} \pi(2 - x^2)^2 \, dx$$

By symmetry:

$$V = 2\pi \int_0^{\sqrt{2}} (4 - 4x^2 + x^4) \, dx = 2\pi \left[4x - \frac{4}{3}x^3 + \frac{1}{5}x^5 \right]_0^{\sqrt{2}}$$

$$V = 2\pi \left(4\sqrt{2} - \frac{4}{3}(2\sqrt{2}) + \frac{1}{5}(4\sqrt{2}) \right) = 2\pi\sqrt{2} \left(4 - \frac{8}{3} + \frac{4}{5} \right) = \frac{64\pi\sqrt{2}}{15}$$

14.4 Volumes of Revolution: The Washer Method

In the previous section, we applied the disk method to solids of revolution where the region being revolved was adjacent to the axis of revolution. When there is a gap between the region and the axis, the resulting solid has a hole or cavity. The cross-sections perpendicular to the axis of revolution are no longer disks but annuli (ring-shaped regions), often called "washers".

Concept and Formula

Consider a region bounded by an outer curve $R_{out}(x)$ and an inner curve $R_{in}(x)$ relative to the axis of revolution. When this region is revolved, the cross-sectional area is given by 14.3. The volume is obtained by integrating this area over the appropriate interval.

Definition 14.4.1. ***The Washer Method***. Let a region be bounded by an outer radius $R(x)$ and an inner radius $r(x)$ with respect to the axis of revolution (taken here as horizontal). The volume of the solid generated by revolving this region about the axis is:

$$V = \int_a^b \pi \left([R(x)]^2 - [r(x)]^2\right) \, dx$$

If the revolution is about a vertical axis, we integrate with respect to y:

$$V = \int_c^d \pi \left([R(y)]^2 - [r(y)]^2\right) \, dy$$

Remark. It is crucial to identify the radii correctly. The "outer" radius R is always the distance from the axis of revolution to the *farther* boundary of the region. The "inner" radius r is the distance from the axis to the *nearer* boundary.

Example 14.4.1. Parabola and Line about x-axis. Find the volume generated by revolving the region bounded by $y = x^2 + 1$ and $y = -x + 3$ about the x-axis.

First, determine the intersection points:

$$x^2 + 1 = -x + 3 \implies x^2 + x - 2 = 0 \implies (x + 2)(x - 1) = 0$$

The region spans $x \in [-2, 1]$. In this interval, test $x = 0$: $y_{parabola} = 1$, $y_{line} = 3$. Thus, the line is the outer boundary relative to the x-axis ($y = 0$). Outer radius $R(x) = -x + 3$. Inner radius $r(x) = x^2 + 1$.

$$V = \int_{-2}^{1} \pi \left((-x+3)^2 - (x^2+1)^2\right) dx = \pi \left[-\frac{x^5}{5} - \frac{x^3}{3} - 3x^2 + 8x\right]_{-2}^{1} = \pi \left(\frac{67}{15} - \left(-\frac{284}{15}\right)\right) = \frac{117\pi}{5}$$

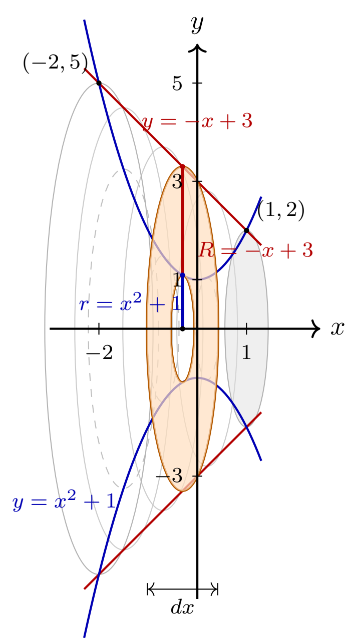


Figure 14.17: Region between $y = x^2 + 1$ and $y = -x + 3$ revolved about the x-axis.

Example 14.4.2. Region Bounded by $y = x^2$ and $y = 2$ about the x-axis. Consider the region in the first quadrant bounded by $y = x^2$ and $y = 2$, revolved about the x-axis.

The curves intersect at $x = \sqrt{2}$. Outer radius $R(x) = 2$ and inner radius $r(x) = x^2$, so

$$V = \int_0^{\sqrt{2}} \pi \left(4 - x^4\right) \, dx = \pi \left[4x - \frac{x^5}{5}\right]_0^{\sqrt{2}} = \frac{16\pi\sqrt{2}}{5}$$

Example 14.4.3. Parabola about $x = -2$. Find the volume of the region in the second quadrant bounded by $y = -x^3$, $y = 0$, and $x = -1$, revolved about the line $x = -2$.

Region: $x \in [-1, 0]$. Since it is in the second quadrant, x is negative. $y = -x^3$ is positive. Revolving about vertical line $x = -2$. We integrate with respect to y. y goes from 0 to $(-1)^3(-1) = 1$. So $y \in [0, 1]$. Express boundaries as functions of y: $y = -x^3 \implies x^3 = -y \implies x = -y^{1/3}$. Left boundary is $x = -1$. Right boundary is $x = -y^{1/3}$. Axis is $x = -2$. Radii are distances from $x = -2$. Outer radius (to the far curve $x = -y^{1/3}$): $R(y) = -y^{1/3} - (-2) = 2 - y^{1/3}$. Inner radius (to the near curve $x = -1$): $r(y) = -1 - (-2) = 1$.

$$V = \pi \int_0^1 \left((2 - y^{1/3})^2 - (1)^2\right) \, dy = \pi \int_0^1 (3 - 4y^{1/3} + y^{2/3}) \, dy = \pi \left[3y - 3y^{4/3} + \frac{3}{5}y^{5/3}\right]_0^1 = \frac{3\pi}{5}$$

Example 14.4.4. Parabola about Horizontal Lines. Consider the region bounded by $y = x^2$ and $y = 1$.

Intersection: $x = \pm 1$. Part (a): Revolve about $y = 2$. Axis is above the region ($y = 2 > 1$). Outer radius (to $y = x^2$): $R(x) = 2 - x^2$. Inner radius (to $y = 1$): $r(x) = 2 - 1 = 1$.

$$V = \pi \int_{-1}^1 \left((2 - x^2)^2 - 1^2\right) \, dx = 2\pi \int_0^1 (4 - 4x^2 + x^4 - 1) \, dx = 2\pi \left[3x - \frac{4}{3}x^3 + \frac{x^5}{5}\right]_0^1 = \frac{56\pi}{15}$$

Part (b): Revolve about $y = -1$. Axis is below the region ($y = -1 < x^2$). Outer radius (to $y = 1$): $R(x) = 1 - (-1) = 2$. Inner radius (to $y = x^2$): $r(x) = x^2 - (-1) = x^2 + 1$.

$$V = \pi \int_{-1}^1 \left(2^2 - (x^2 + 1)^2\right) \, dx = 2\pi \int_0^1 (3 - 2x^2 - x^4) \, dx = 2\pi \left(3 - \frac{2}{3} - \frac{1}{5}\right) = \frac{64\pi}{15}$$

14.5 Cylindrical Shells

The disk and washer methods slice perpendicular to the axis of revolution. In some geometries this forces awkward inversions, so we use the *Method of Cylindrical Shells*, which slices parallel to the axis.

Concept and Formula

Consider a region bounded by $y = f(x)$ and the x-axis on the interval $[a, b]$, revolved about the y-axis. Instead of slicing perpendicular to the y-axis (which would require inverting $f(x)$ to get $x = g(y)$), we slice the region into vertical strips at position x with width dx. When such a strip is revolved about the y-axis, it generates a thin cylindrical shell of radius x, height $f(x)$, and thickness dx. The volume of this shell is approximately its circumference times its height times its thickness: $dV \approx 2\pi x \cdot f(x) \cdot dx$.

Definition 14.5.1. ***The Shell Method***. Let f be continuous and non-negative on $[a, b]$ where $a \geq 0$. The volume V of the solid generated by revolving the region under $y = f(x)$ about the y-axis is:

$$V = \int_a^b 2\pi x f(x) \, dx$$

If revolving a region bounded by $x = g(y)$ about the x-axis, the volume is:

$$V = \int_c^d 2\pi y g(y) \, dy$$

Remark. The term x represents the *shell radius* (distance from the axis) and $f(x)$ represents the *shell height*. If the axis of revolution is shifted to $x = L$, the radius becomes $|x - L|$.

Example 14.5.1. Parabola about y-axis. Find the volume generated by revolving the region bounded by $y = \sqrt{x}$, $x = 0$, and $x = 4$ about the y-axis.

Shell radius: x. Shell height: $\sqrt{x}$. Limits: $x \in [0, 4]$.

$$V = \int_0^4 2\pi x(\sqrt{x})\,dx = 2\pi \int_0^4 x^{3/2}\,dx = 2\pi \left[\frac{2}{5}x^{5/2}\right]_0^4 = \frac{128\pi}{5}$$

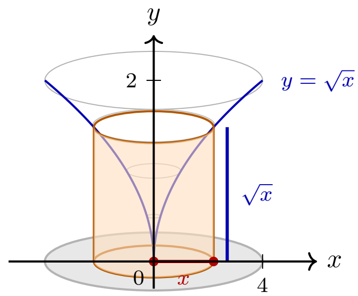


Figure 14.18: Shells for the region under $y = \sqrt{x}$ about the y-axis.

Example 14.5.2. Parabola about Shifted Vertical Axis. Find the volume generated by revolving the region bounded by $y = 3x - x^2$ and the x-axis about the line $x = -1$.

The curve meets the x-axis at $x = 0$ and $x = 3$. Shell radius: $x + 1$. Shell height: $3x - x^2$.

$$V = \int_0^3 2\pi(x+1)(3x - x^2)\,dx = 2\pi \int_0^3 (-x^3 + 2x^2 + 3x)\,dx = 2\pi \left[-\frac{x^4}{4} + \frac{2}{3}x^3 + \frac{3}{2}x^2\right]_0^3 = \frac{45\pi}{2}$$

When revolving about the x-axis (or any horizontal line), we can use horizontal shells (integrating with respect to y).

Example 14.5.3. Curve $x = y^2$ about x-axis. Find the volume generated by revolving the region bounded by $x = y^2$, $x = 0$, and $y = 1$ about the x-axis.

Shell radius: y. Shell height: y^2.

$$V = \int_0^1 2\pi y(y^2)\,dy = 2\pi \int_0^1 y^3\,dy = 2\pi \left[\frac{y^4}{4}\right]_0^1 = \frac{\pi}{2}$$

The disk method gives the same value.

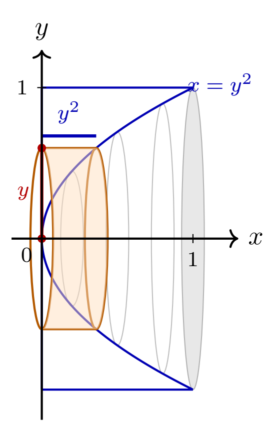


Figure 14.19: Region bounded by $x = y^2$, $x = 0$, and $y = 1$.

Example 14.5.4. Region Bounded by Two Curves. Find the volume generated by revolving the region between $y^2 = x^2$ ($y = x$) and $y^2 = 3x - 2$ ($x = (y^2+2)/3$) in the upper half plane about the line $y = 5$.

Intersection: $x^2 = 3x - 2 \implies x^2 - 3x + 2 = 0 \implies x = 1, 2$. Corresponding y: $y = 1, y = \sqrt{4} = 2$. (Upper half plane). We use horizontal shells. The radius is $5 - y$. The height is the horizontal distance between $x = y$ and $x = (y^2+2)/3$, so $h(y) = y - \frac{y^2+2}{3}$.

$$V = \int_1^2 2\pi(5-y)\left(y - \frac{y^2+2}{3}\right) dy$$

This integral can be evaluated by expansion.

Comparing Methods

For a given problem, one method is often algebraically simpler than the other.

- **Disk/Washer**: Good when the axis of revolution is a boundary of the region or when $y = f(x)$ is easy to square. Slices are perpendicular to the axis.
- **Shells**: Good when the radius is simply x (or y) and the height $f(x)$ is simple, or when inverting $y = f(x)$ is difficult. Slices are parallel to the axis.

Example 14.5.5. Method Selection. Region bounded by $y = x^2$ and $y = 3x^2 - 2$ about the y-axis.

Intersection: $x^2 = 3x^2 - 2 \implies 2x^2 = 2 \implies x = \pm 1$. The region is between the parabolas for $x \in [-1, 1]$. Axis: y-axis (vertical).

Shell Method (dx): Radius x (use symmetry, integrate 0 to 1). Height: $x^2 - (3x^2 - 2) = 2 - 2x^2$.

$$V = \int_0^1 2\pi x(2 - 2x^2)\, dx = 2\pi \int_0^1 (2x - 2x^3)\, dx = 2\pi \left[x^2 - \frac{x^4}{2}\right]_0^1 = 2\pi(1 - 0.5) = \pi$$

Washer Method (dy): Top boundary $y = x^2 \implies x = \sqrt{y}$. Bottom boundary $y = 3x^2 - 2 \implies x = \sqrt{(y+2)/3}$. y ranges from -2 to 1. But the boundaries change at $y = 0$. From $y = -2$ to 0: bounded by $x = \pm\sqrt{(y+2)/3}$. Solid disk. From $y = 0$ to 1: bounded by $x = \sqrt{(y+2)/3}$ (outer) and $x = \sqrt{y}$ (inner).

$$V_1 = \pi \int_{-2}^0 \left(\sqrt{\frac{y+2}{3}}\right)^2 dy = \pi \int_{-2}^0 \frac{y+2}{3} dy = \frac{\pi}{3}\left[\frac{y^2}{2} + 2y\right]_{-2}^0 = \frac{2\pi}{3}$$

$$V_2 = \pi \int_0^1 \left(\frac{y+2}{3} - y\right) dy = \pi \int_0^1 \left(\frac{2}{3} - \frac{2y}{3}\right) dy = \pi\left[\frac{2y}{3} - \frac{y^2}{3}\right]_0^1 = \frac{\pi}{3}$$

Thus $V = V_1 + V_2 = \pi$, it should be easy to see that this is more complex and annoying, so shells are preferable.

14.6 Arc Length

So far, we have used integration to compute areas of regions and volumes of solids. We now turn to a one-dimensional geometric application: finding the length of a curve.

A plane curve can be described as the path traced by a moving point $(x(t), y(t))$ as the parameter t varies over an interval $[a, b]$. We assume the functions $x(t)$ and $y(t)$ are continuously differentiable, defining a *smooth* curve.

$$C = \{(x(t), y(t)) : t \in [a, b]\}$$

We wish to compute the length L of this curve.

Derivation via Linear Approximation

Partition the interval $[a, b]$ into n sub-intervals with endpoints $a = t_0 < t_1 < \cdots < t_n = b$. This partitions the curve into n segments with endpoints $P_i = (x(t_i), y(t_i))$. We approximate the length of the curve by the sum of the lengths of the chords connecting these points:

$$L \approx \sum_{i=1}^{n} \|P_i - P_{i-1}\| = \sum_{i=1}^{n} \sqrt{(x(t_i) - x(t_{i-1}))^2 + (y(t_i) - y(t_{i-1}))^2}$$

By the Mean Value Theorem, $\Delta x_i = x(t_i) - x(t_{i-1}) = x'(c_i)\Delta t_i$ and $\Delta y_i = y'(d_i)\Delta t_i$ for some $c_i, d_i \in [t_{i-1}, t_i]$. As $\Delta t_i \to 0$, the sum approaches the integral of the speed.

Definition 14.6.1. ***Arc Length***. If a smooth curve is defined parametrically by $x = x(t)$ and $y = y(t)$ for $t \in [a, b]$, its length L is:

$$L = \int_a^b \sqrt{\left(\frac{dx}{dt}\right)^2 + \left(\frac{dy}{dt}\right)^2}\, dt$$

If the curve is the graph of a function $y = f(x)$ for $x \in [a, b]$, we use x as the parameter ($t = x, y = f(x)$):

$$L = \int_a^b \sqrt{1 + \left(\frac{dy}{dx}\right)^2}\, dx$$

Remark. The integrand $ds = \sqrt{dx^2 + dy^2}$ is the differential of arc length. The formula $L = \int ds$ unifies the parametric and explicit cases.

Example 14.6.1. Circumference of a Circle. Consider a circle of radius R parametrised by $x(t) = R\cos t, y(t) = R\sin t$ for $t \in [0, 2\pi]$.

Derivatives: $x'(t) = -R\sin t$, $y'(t) = R\cos t$. Speed: $(x')^2 + (y')^2 = R^2\sin^2 t + R^2\cos^2 t = R^2$.

$$L = \int_0^{2\pi} \sqrt{R^2}\, dt = \int_0^{2\pi} R\, dt = 2\pi R$$

Example 14.6.2. Length of a Parabolic Segment. Find the length of the curve $y = x^2$ from $x = 0$ to $x = 1$.

Here $y' = 2x$.

$$L = \int_0^1 \sqrt{1 + (2x)^2}\, dx = \int_0^1 \sqrt{1 + 4x^2}\, dx$$

Let $2x = \tan\theta$, $2dx = \sec^2\theta d\theta$. Limits: $0 \to \arctan 2$.

$$\int \sqrt{1 + \tan^2\theta}\frac{1}{2}\sec^2\theta d\theta = \frac{1}{2}\int \sec^3\theta d\theta$$

Using the reduction formula for $\sec^3\theta$:

$$\frac{1}{2}\left[\frac{1}{2}\sec\theta\tan\theta + \frac{1}{2}\ln|\sec\theta + \tan\theta|\right]$$

Converting back ($\tan\theta = 2x, \sec\theta = \sqrt{1 + 4x^2}$):

$$L = \left[\frac{1}{2}x\sqrt{1 + 4x^2} + \frac{1}{4}\ln|2x + \sqrt{1 + 4x^2}|\right]_0^1 = \frac{1}{2}\sqrt{5} + \frac{1}{4}\ln(2 + \sqrt{5})$$

Example 14.6.3. The Astroid. The astroid is defined by $x = \cos^3 t, y = \sin^3 t$ for $t \in [0, 2\pi]$.

By symmetry, the total length is 4 times the length in the first quadrant ($t \in [0, \pi/2]$). Derivatives: $x'(t) = 3\cos^2 t(-\sin t)$ $y'(t) = 3\sin^2 t(\cos t)$ Sum of squares:

$$(x')^2 + (y')^2 = 9\cos^4 t\sin^2 t + 9\sin^4 t\cos^2 t = 9\sin^2 t\cos^2 t(\cos^2 t + \sin^2 t) = 9\sin^2 t\cos^2 t$$

Square root:

$$\sqrt{(x')^2+(y')^2} = 3|\sin t\cos t| = 3\sin t\cos t \quad (\text{on } [0,\pi/2])$$

Length of one quadrant:

$$L_1 = \int_0^{\pi/2} 3\sin t\cos t\,dt = \frac{3}{2}\int_0^{\pi/2}\sin(2t)\,dt = \frac{3}{2}\left[-\frac{1}{2}\cos(2t)\right]_0^{\pi/2} = -\frac{3}{4}(-1-1) = \frac{3}{2}$$

Total length $L = 4\times\frac{3}{2} = 6$.

Example 14.6.4. The Cycloid. A cycloid is the path traced by a point on the rim of a rolling wheel. Parametrisation: $x = R(t-\sin t), y = R(1-\cos t)$ for $t\in[0,2\pi]$.

Derivatives: $x' = R(1-\cos t), y' = R\sin t$.

$$(x')^2+(y')^2 = R^2(1-2\cos t+\cos^2 t) + R^2\sin^2 t = R^2(2-2\cos t) = 2R^2(1-\cos t)$$

Using the half-angle identity $1-\cos t = 2\sin^2(t/2)$:

$$ds = \sqrt{4R^2\sin^2(t/2)}\,dt = 2R\sin(t/2)\,dt \quad (\text{since } t/2\in[0,\pi])$$

$$L = \int_0^{2\pi} 2R\sin(t/2)\,dt = 2R\left[-2\cos(t/2)\right]_0^{2\pi} = -4R(-1-1) = 8R$$

Arc Length Function

Often we wish to measure distance along a curve from a starting point a. We define the arc length function $s(x)$:

$$s(x) = \int_a^x \sqrt{1+[f'(t)]^2}\,dt$$

By the Fundamental Theorem of Calculus, the rate of change of arc length with respect to x is:

$$\frac{ds}{dx} = \sqrt{1+\left(\frac{dy}{dx}\right)^2}$$

or in differential form $ds^2 = dx^2 + dy^2$.

Example 14.6.5. Finding the Curve from Length. Suppose a smooth curve $y = f(x)$ passing through $(0,0)$ has the property that the arc length from 0 to x is $\sqrt{2}x$ for all $x > 0$. Find $f(x)$.

Given:

$$\int_0^x \sqrt{1+[f'(t)]^2}\,dt = \sqrt{2}x$$

Differentiating both sides with respect to x:

$$\sqrt{1+[f'(x)]^2} = \sqrt{2}$$

Squaring gives $1+[f'(x)]^2 = 2 \implies [f'(x)]^2 = 1 \implies f'(x) = \pm 1$. Since $f(0) = 0$, the curve is either $y = x$ or $y = -x$.

14.7 Surface Area of Revolution

We now consider the surface area generated by rotating a curve about an axis. Unlike volume, where the solid is decomposed into simple cylindrical or washer elements, surface area requires approximating the surface with conical frustums (bands) rather than cylinders.

Derivation via Frustums

Consider a curve $y = f(x)$ for $a \le x \le b$, with $f(x) \ge 0$, rotated about the x-axis. We partition the interval $[a, b]$ into sub-intervals $[x_{i-1}, x_i]$. The segment of the curve between P_{i-1} and P_i is approximated by the chord connecting them. When rotated, this chord sweeps out the lateral surface of a *frustum* of a cone. The surface area of a conical frustum with base radii r_1, r_2 and slant height L is given by $S = \pi(r_1 + r_2)L = 2\pi r_{avg} L$. For the segment from x_{i-1} to x_i, the radii are $f(x_{i-1})$ and $f(x_i)$, and the slant length is $\Delta s_i = \sqrt{(\Delta x_i)^2 + (\Delta y_i)^2}$. Approximating $f(x)$ as constant over the small interval, the area element is $dS = 2\pi f(x)\, ds$.

Definition 14.7.1. ***Area of Surface of Revolution***. Let $y = f(x)$ be a continuously differentiable function on $[a, b]$. The surface area S generated by revolving the curve about the x-axis is:

$$S = \int_a^b 2\pi f(x) \sqrt{1 + [f'(x)]^2}\, dx$$

If the curve is parametrised by $x(t), y(t)$, the formula becomes:

$$S = \int_{t_a}^{t_b} 2\pi y(t) \sqrt{[x'(t)]^2 + [y'(t)]^2}\, dt$$

For revolution about the y-axis, the radius of rotation is x instead of y, so the integrand becomes $2\pi x\, ds$.

Remark. A useful mnemonic is $S = \int 2\pi(\text{radius})\, ds$, where "radius" is the distance from the point on the curve to the axis of revolution, and ds is the arc length differential.

Example 14.7.1. Surface Area of a Sphere. Revolve the semicircle $y = \sqrt{R^2 - x^2}$ about the x-axis.

$f'(x) = \frac{-x}{\sqrt{R^2 - x^2}}$.

$$\sqrt{1 + [f'(x)]^2} = \sqrt{1 + \frac{x^2}{R^2 - x^2}} = \sqrt{\frac{R^2 - x^2 + x^2}{R^2 - x^2}} = \frac{R}{\sqrt{R^2 - x^2}}$$

$$S = \int_{-R}^{R} 2\pi \sqrt{R^2 - x^2} \cdot \frac{R}{\sqrt{R^2 - x^2}}\, dx = \int_{-R}^{R} 2\pi R\, dx = 2\pi R[x]_{-R}^{R} = 4\pi R^2$$

Example 14.7.2. Parabolic Dish. Find the surface area generated by revolving $y = 2\sqrt{x}$ for $1 \le x \le 2$ about the x-axis.

$y' = \frac{1}{\sqrt{x}}$. $ds = \sqrt{1 + \frac{1}{x}}\, dx = \sqrt{\frac{x+1}{x}}\, dx$. Radius of rotation is $y = 2\sqrt{x}$.

$$S = \int_1^2 2\pi(2\sqrt{x}) \sqrt{\frac{x+1}{x}}\, dx = 4\pi \int_1^2 \sqrt{x+1}\, dx = 4\pi \left[\frac{2}{3}(x+1)^{3/2}\right]_1^2 = \frac{8\pi}{3}(3\sqrt{3} - 2\sqrt{2})$$

Example 14.7.3. Revolution about the x-axis. Revolve the arc of $y = \sqrt{2x - 1}$ for $x \in [5/8, 1]$ about the x-axis.

Here $y' = \frac{1}{\sqrt{2x-1}}$, so

$$S = \int_{5/8}^{1} 2\pi y \sqrt{1 + (y')^2}\, dx = \int_{5/8}^{1} 2\pi \sqrt{2x}\, dx = \frac{\pi}{12}\left(16\sqrt{2} - 5\sqrt{5}\right)$$

Example 14.7.4. Gabriel's Horn (The Bugle). Consider the curve $y = 1/x$ for $x \ge 1$, revolved about the x-axis.

Volume: $V = \int_1^\infty \pi(1/x)^2 dx = \pi[-1/x]_1^\infty = \pi$. Finite. Surface area: $y' = -1/x^2$. $ds = \sqrt{1 + 1/x^4} dx$. Since $\sqrt{1 + 1/x^4} > 1$, $S = \int_1^\infty 2\pi(1/x)\sqrt{1 + 1/x^4} dx > \int_1^\infty \frac{2\pi}{x} dx$. The integral $\int_1^\infty \frac{dx}{x} = [\ln x]_1^\infty = \infty$. Thus, the surface area is infinite.

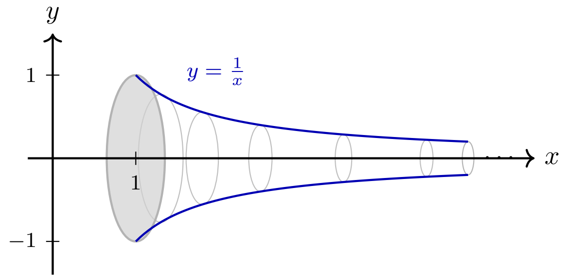


Figure 14.20: Gabriel's Horn generated by $y = 1/x$.

14.8 Exercises

1. **Area between Transcendentals.** Find the area of the region bounded by the curves $y = e^x$, $y = e^{-x}$, and the line $x = 1$.
   (a) Sketch the region and identify the intersection points.
   (b) Set up the integral with respect to x.
   (c) Evaluate the area.

2. **Horizontal vs Vertical Slicing.** Consider the region bounded by $y = \arcsin x$, $y = \arccos x$, and the x-axis.
   (a) Find the point of intersection of the two inverse trigonometric curves.
   (b) Calculate the area by integrating with respect to x. Note that you will need to split the integral.
   (c) Calculate the area by integrating with respect to y. Is this method simpler?

3. **Loop Area.** Find the area enclosed by the loop of the curve $y^2 = x(x-1)^2$.
   (a) Determine the range of x for which the curve is defined.
   (b) Sketch the curve.
   (c) Use symmetry to calculate the total enclosed area.

4. **Rotation about an Oblique Line.** While the text covered horizontal and vertical axes, the principle of slicing perpendicular to the axis holds generally. Find the volume generated by revolving the triangle with vertices $(0,0)$, $(1,0)$, and $(1,1)$ about the line $y = x$.
   (a) Find the perpendicular distance from a point (x, y) on the boundary to the line $y = x$. This is the radius R.
   (b) Parametrise the boundary or integrate along the axis $u = (x+y)/\sqrt{2}$.
   (c) Alternatively, use the theorems of Pappus (if known) or coordinate rotation to simplify the integral.

5. **The Torus.** A torus is generated by revolving the circle $(x-R)^2 + y^2 = r^2$ (where $0 < r < R$) about the y-axis.
   (a) Set up the integral for the volume using the Washer Method (integrating with respect to y).
   (b) Show that the volume is $V = 2\pi^2 R r^2$.

6. **Infinite Solid.** Consider the region bounded by $y = 1/x$, $y = 0$, and $x \geq 1$.
   (a) Show that the volume generated by revolving this region about the x-axis is finite (Gabriel's Horn).
   (b) Show that the volume generated by revolving the same region about the y-axis is infinite.

7. **Shell Method Practice.** Use the shell method to find the volumes of the solids generated by revolving the region bounded by $y = x^2$ and $y = 2 - x^2$ about:
   (a) The y-axis.
   (b) The line $x = 2$.

8. Comparison of Methods. Let R be the region bounded by $y = x^3$, $y = 8$, and $x = 0$.

(a) Set up the integral for the volume of the solid generated by revolving R about the x-axis using the Disk Method.
(b) Set up the integral for the same volume using the Shell Method.
(c) Evaluate both integrals to verify they yield the same result.

9. A "Bead" from a Sphere. A cylindrical hole of radius r is drilled through the centre of a sphere of radius R ($r < R$).

(a) Use the shell method to determine the volume of the remaining solid (the "bead").
(b) Express the result in terms of the height h of the bead ($h = 2\sqrt{R^2 - r^2}$). Note the surprising independence from R.

10. Length of the Catenary. The curve formed by a hanging chain is the catenary $y = a\cosh(x/a)$. Find the length of the catenary from $x = 0$ to $x = b$.

11. Surface Area of the Torus. Using the same circle as in Exercise 5 ($(x - R)^2 + y^2 = r^2$), find the surface area of the torus generated by revolving it about the y-axis.

(a) Parametrise the circle as $x = R + r\cos t, y = r\sin t$.
(b) Use the parametric formula for surface area.

12. Geometric Limits. Consider the arc of the circle $x^2 + y^2 = 1$ in the first quadrant. Approximate its length by inscribing n chords corresponding to equal increments in x. Let L_n be this approximation. Show that $\lim_{n\to\infty} L_n = \pi/2$ by recognising the limit as a Riemann sum for the arc length integral.

13. Complex Boundaries. Find the area of the region lying inside the circle $x^2 + y^2 = 4$ but outside the parabola $y^2 = 3x$.

Remark. Find intersection points, exploit symmetry, and perhaps mix integration variables or methods.

14. Volume of the Ellipsoid. Find the volume of the solid generated by revolving the ellipse $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$ about the x-axis. Show that the result is $\frac{4}{3}\pi ab^2$. By symmetry/analogy, what is the volume if revolved about the y-axis?

15. Surface Area Challenge. Find the surface area generated by revolving the loop of the curve $9ay^2 = x(3a - x)^2$ about the x-axis.

Remark. The loop exists for $x \in [0, 3a]$. Watch out for the derivative calculation.

16. The Orthogonal Trajectories. Consider the family of curves $y = cx^2$.

(a) Find the family of orthogonal trajectories (curves that intersect every member of the first family at right angles).
(b) Sketch the region bounded by $y = x^2, y = 4x^2$ and the orthogonal trajectories through $(1, 1)$ and $(1, 2)$.
(c) (Challenge) Can you compute the area of this curvilinear "rectangle"?

17. Centre of Mass (Centroid). While not explicitly covered, the logic of slicing extends to moments. The moment of a thin slice about the y-axis is $x \cdot dA$. The x-coordinate of the centroid is $\bar{x} = \frac{\int x dA}{\int dA}$. Use this to find the centroid of the semicircular region $y = \sqrt{r^2 - x^2}$.

Remark. This relates to the Theorems of Pappus-Guldinus.

18. Improper Surface Area. We showed Gabriel's Horn (revolving $1/x$ for $x \geq 1$) has finite volume but infinite surface area. Consider the curve $y = e^{-x}$ for $x \geq 0$, revolved about the x-axis.

(a) Compute the volume of the solid. Is it finite?
(b) Set up the integral for the surface area.
(c) Determine if the surface area is finite or infinite.

Remark. Comparison test with e^{-x}.

Chapter 15

Approximation and Convergence

We conclude this course by examining two powerful results that bridge the discrete and the continuous. The first, the Integral Test, connects the summation of a series (a discrete process) to the improper integral of a function (a continuous process). The second, the Weierstrass Approximation Theorem, asserts that the seemingly rigid world of polynomials is flexible enough to approximate any continuous function with arbitrary precision.

15.1 The Integral Test

We have previously established the convergence of the p-series $\sum n^{-p}$ for $p > 1$ and its divergence for $p \leq 1$ using specific comparison arguments. It is natural to ask whether the convergence of a series $\sum f(n)$ can generally be inferred from the behaviour of the function $f(x)$ on the continuous domain $[1, \infty)$.

The intuition is geometric: the sum $\sum f(n)$ represents the total area of rectangles of width 1 and height $f(n)$. If f is decreasing, these rectangles can be used to bound the area under the curve $y = f(x)$, and vice versa.

***Theorem* 15.1.1. The Integral Test**. Let $f : [1, \infty) \to \mathbb{R}$ be a continuous, positive, and decreasing function. Then the series $\sum_{n=1}^{\infty} f(n)$ converges if and only if the improper integral $\int_1^{\infty} f(x)\,dx$ converges.

Proof. Let $a_n = f(n)$. Since f is decreasing, for any $x \in [n, n+1]$, we have:

$$f(n+1) \leq f(x) \leq f(n)$$

Integrating this inequality from n to $n+1$ yields:

$$\int_n^{n+1} f(n+1)\,dx \leq \int_n^{n+1} f(x)\,dx \leq \int_n^{n+1} f(n)\,dx$$

Since $f(n)$ and $f(n+1)$ are constants on this interval, this simplifies to:

$$a_{n+1} \leq \int_n^{n+1} f(x)\,dx \leq a_n$$

($\Longleftarrow$): Suppose the integral converges to a limit I. Summing the left inequality from $n = 1$ to $N - 1$:

$$\sum_{n=1}^{N-1} a_{n+1} \leq \sum_{n=1}^{N-1} \int_n^{n+1} f(x)\,dx = \int_1^N f(x)\,dx \leq \int_1^{\infty} f(x)\,dx = I$$

Thus, the partial sums of the series $\sum_{n=2}^{\infty} a_n$ are bounded above by I. Since terms are positive, the series converges. Consequently, $\sum_{n=1}^{\infty} a_n$ converges.

($\Longrightarrow$): Suppose the series converges to a sum S. Summing the right inequality from $n = 1$ to N:

$$\int_1^{N+1} f(x)\,dx = \sum_{n=1}^{N} \int_n^{n+1} f(x)\,dx \leq \sum_{n=1}^{N} a_n \leq S$$

Since f is positive, the integral function $F(t) = \int_1^t f(x)\,dx$ is increasing. Since it is bounded above by S, $\lim_{t\to\infty} F(t)$ exists.

■

Example 15.1.1. p-series revisited. Consider $f(x) = x^{-p}$ for $p \neq 1$.

$$\int_1^{\infty} x^{-p}\,dx = \lim_{b\to\infty} \left[\frac{x^{1-p}}{1-p}\right]_1^b = \lim_{b\to\infty} \frac{b^{1-p} - 1}{1-p}$$

If $p > 1$, then $1 - p < 0$, so $b^{1-p} \to 0$ and the integral converges. If $p < 1$, then $1 - p > 0$, so $b^{1-p} \to \infty$ and the integral diverges. For $p = 1$, $\int_1^{\infty} x^{-1}\,dx = [\ln x]_1^{\infty} = \infty$. Thus, the Integral Test recovers the convergence criteria for the p-series immediately.

15.2 The Weierstrass Approximation Theorem

We now turn to the final major result of this course. In analysis, polynomials are the simplest functions to manipulate: they are infinitely differentiable, easily integrated, and computable via finite arithmetic operations. The Weierstrass Approximation Theorem assures us that these simple functions are sufficient to capture the behaviour of *any* continuous function on a closed interval.

***Theorem* 15.2.1. Weierstrass Approximation Theorem**. Let $f : [a, b] \to \mathbb{R}$ be a continuous function. For every $\epsilon > 0$, there exists a polynomial $P(x)$ such that:

$$\sup_{x\in[a,b]} |f(x) - P(x)| < \epsilon$$

In other words, the set of polynomials is dense in $C[a, b]$ under the uniform norm.

While there are many proofs of this theorem (some relying on convolutions or the Stone-Weierstrass generalisation), we present a constructive proof due to Bernstein. This approach is particularly illuminating because it provides an explicit formula for the approximating polynomials.

Bernstein Polynomials

Without loss of generality, we may restrict our attention to the interval $[0, 1]$. If f is defined on $[a, b]$, the transformation $g(t) = f(a + t(b - a))$ defines a function on $[0, 1]$.

Definition 15.2.1. *Bernstein Basis Polynomials*. For a fixed $n \in \mathbb{N} \cup \{0\}$, the Bernstein basis polynomials of degree n are defined for $k \in \{0, 1, \ldots, n\}$ as:

$$B_{n,k}(x) = \binom{n}{k} x^k (1 - x)^{n-k}$$

These polynomials arise naturally in probability theory as the probability mass function of a binomial distribution $B(n, x)$. Specifically, if a biased coin with head probability x is tossed n times, $B_{n,k}(x)$ is the probability of observing exactly k heads.

***Lemma* 15.2.1.** *Properties of Bernstein Polynomials.* For all $x \in \mathbb{R}$ and $n \in \mathbb{N}$:

1. $\sum_{k=0}^{n} B_{n,k}(x) = 1$

2. $\sum_{k=0}^{n} kB_{n,k}(x) = nx$
3. $\sum_{k=0}^{n} k(k-1)B_{n,k}(x) = n(n-1)x^2$

Proof. (1) follows immediately from the Binomial Theorem expansion of $(x + (1-x))^n = 1^n = 1$. For (2), we observe the identity $k\binom{n}{k} = n\binom{n-1}{k-1}$.

$$\begin{aligned}\sum_{k=0}^{n} k\binom{n}{k}x^k(1-x)^{n-k} &= \sum_{k=1}^{n} n\binom{n-1}{k-1}x^k(1-x)^{n-k} \\ &= nx\sum_{j=0}^{n-1}\binom{n-1}{j}x^j(1-x)^{(n-1)-j} \quad \text{(setting } j = k-1\text{)} \\ &= nx \cdot 1 = nx\end{aligned}$$

(3) is proved similarly using $k(k-1)\binom{n}{k} = n(n-1)\binom{n-2}{k-2}$. ■

We now construct the approximation. Given $f : [0,1] \to \mathbb{R}$, we define the n-th Bernstein polynomial of f as:

$$B_n(f)(x) = \sum_{k=0}^{n} f\left(\frac{k}{n}\right) B_{n,k}(x)$$

Intuitively, $B_n(f)(x)$ is a weighted average of the values of f. The weight $B_{n,k}(x)$ is concentrated around $k \approx nx$, i.e., $k/n \approx x$. Thus, the sum samples f primarily near x.

To formalise this concentration of measure, we estimate the "variance".

***Lemma* 15.2.2.** *Variance Estimate.* For any $x \in [0,1]$:

$$\sum_{k=0}^{n}\left(\frac{k}{n} - x\right)^2 B_{n,k}(x) \le \frac{1}{4n}$$

Proof. Expanding the square term:

$$\sum_{k=0}^{n}(k - nx)^2 B_{n,k}(x) = \sum_{k=0}^{n}(k^2 - 2nxk + n^2x^2)B_{n,k}(x)$$

Using the identity $k^2 = k(k-1) + k$ and the properties from the lemma 15.2.1:

$$\begin{aligned}\sum k^2 B_{n,k}(x) &= \sum k(k-1)B_{n,k}(x) + \sum kB_{n,k}(x) \\ &= n(n-1)x^2 + nx\end{aligned}$$

Substituting this back:

$$\begin{aligned}\sum(k - nx)^2 B_{n,k}(x) &= [n(n-1)x^2 + nx] - 2nx(nx) + n^2x^2 \cdot 1 \\ &= n^2x^2 - nx^2 + nx - 2n^2x^2 + n^2x^2 \\ &= nx - nx^2 = nx(1-x)\end{aligned}$$

Dividing by n^2:

$$\sum_{k=0}^{n}\left(\frac{k}{n} - x\right)^2 B_{n,k}(x) = \frac{x(1-x)}{n}$$

Since $x(1-x) \le 1/4$ for $x \in [0,1]$, the result follows. ■

We are now equipped to prove the main theorem. We show that $B_n(f)$ converges uniformly to f.

Proof for theorem 15.2.1. Let $\epsilon > 0$. Since f is continuous on the compact interval $[0,1]$, it is bounded (say by M) and uniformly continuous. By uniform continuity, there exists $\delta > 0$ such that for any $x, y \in [0,1]$:

$$|x-y| < \delta \implies |f(x) - f(y)| < \frac{\epsilon}{2}$$

We wish to bound $|f(x) - B_n(f)(x)|$. Using the partition of unity $\sum B_{n,k}(x) = 1$:

$$\left| f(x) - \sum_{k=0}^{n} f\left(\frac{k}{n}\right) B_{n,k}(x) \right| = \left| \sum_{k=0}^{n} \left(f(x) - f\left(\frac{k}{n}\right) \right) B_{n,k}(x) \right| \leq \sum_{k=0}^{n} \left| f(x) - f\left(\frac{k}{n}\right) \right| B_{n,k}(x)$$

We split the sum into two sets of indices: those where k/n is close to x ($S_{close} = \{k : |k/n - x| < \delta\}$) and those where it is far ($S_{far} = \{k : |k/n - x| \geq \delta\}$).

Case 1: Close indices. For $k \in S_{close}$, we have $|x - k/n| < \delta$, so $|f(x) - f(k/n)| < \epsilon/2$.

$$\sum_{k \in S_{close}} \left| f(x) - f\left(\frac{k}{n}\right) \right| B_{n,k}(x) < \frac{\epsilon}{2} \sum_{k \in S_{close}} B_{n,k}(x) \leq \frac{\epsilon}{2} \cdot 1 = \frac{\epsilon}{2}$$

Case 2: Far indices. For $k \in S_{far}$, we use the coarse bound $|f(x) - f(k/n)| \leq |f(x)| + |f(k/n)| \leq 2M$.

$$\sum_{k \in S_{far}} \left| f(x) - f\left(\frac{k}{n}\right) \right| B_{n,k}(x) \leq 2M \sum_{k \in S_{far}} B_{n,k}(x)$$

Notice that for $k \in S_{far}$, we have $\frac{|k/n - x|^2}{\delta^2} \geq 1$. Thus:

$$\sum_{k \in S_{far}} B_{n,k}(x) \leq \sum_{k \in S_{far}} \frac{(k/n - x)^2}{\delta^2} B_{n,k}(x) \leq \frac{1}{\delta^2} \sum_{k=0}^{n} \left(\frac{k}{n} - x\right)^2 B_{n,k}(x)$$

Using the lemma 15.2.2:

$$\frac{1}{\delta^2} \sum_{k=0}^{n} \left(\frac{k}{n} - x\right)^2 B_{n,k}(x) \leq \frac{1}{\delta^2} \cdot \frac{1}{4n} = \frac{1}{4n\delta^2}$$

Thus, the contribution from the far indices is bounded by $2M \cdot \frac{1}{4n\delta^2} = \frac{M}{2n\delta^2}$.

Combining the two parts:

$$|f(x) - B_n(f)(x)| < \frac{\epsilon}{2} + \frac{M}{2n\delta^2}$$

This bound holds for all $x \in [0,1]$. The first term is constant. The second term tends to 0 as $n \to \infty$. We choose N large enough such that for all $n \geq N$, $\frac{M}{2n\delta^2} < \frac{\epsilon}{2}$. Then for all $n \geq N$ and all $x \in [0,1]$:

$$|f(x) - B_n(f)(x)| < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon$$

Thus, $B_n(f)$ converges uniformly to f. ■

This theorem provides a fitting conclusion to our study. We began with the axioms of the real number system, built the limit processes of calculus, and have now shown that these intricate tools allow us to approximate the most general continuous functions by the most elementary ones. This bridge lays the foundation for advanced topics in Functional Analysis and Numerical Analysis.

15.3 Exercises

1. **Series Convergence.** Determine whether the following series converge or diverge using the Integral Test where appropriate.
   (a) $\sum_{n=2}^{\infty} \frac{1}{n \ln n}$
   (b) $\sum_{n=2}^{\infty} \frac{1}{n(\ln n)^2}$
   (c) $\sum_{n=1}^{\infty} ne^{-n^2}$

2. **Error Estimation.** Suppose $\sum a_n$ converges to S. The Integral Test provides bounds on the remainder $R_N = S - S_N = \sum_{n=N+1}^{\infty} a_n$. Let $f(x)$ be the corresponding continuous, decreasing function.
   (a) Prove that $\int_{N+1}^{\infty} f(x)\,dx \le R_N \le \int_{N}^{\infty} f(x)\,dx$.
   (b) How many terms of the series $\sum_{n=1}^{\infty} n^{-3}$ are needed to approximate the sum with an error less than 10^{-3}?

3. **Divergence Speed.** The harmonic series diverges very slowly. Let $H_n = \sum_{k=1}^{n} \frac{1}{k}$.
   (a) Use the Integral Test logic to show $\ln(n+1) < H_n < 1 + \ln n$.
   (b) Estimate how many terms are needed for the sum to exceed 100. (The universe might end before you finish adding them).

4. **Explicit Bernstein Construction.** Let $f(x) = |x - 1/2|$ on $[0, 1]$.
   (a) Write down the Bernstein polynomial $B_2(f)(x)$ explicitly.
   (b) Calculate $B_4(f)(1/2)$.
   (c) Does the sequence $B_n(f)(1/2)$ converge to $f(1/2)$? Check for small n.

5. **Approximation of Derivatives.** Let f be continuously differentiable on $[0, 1]$.
   (a) Calculate the derivative of the Bernstein polynomial $B_n(f)'(x)$.
   (b) Show that it can be written in terms of Bernstein polynomials of degree $n - 1$:

$$B_n(f)'(x) = n \sum_{k=0}^{n-1} \left[f\left(\frac{k+1}{n}\right) - f\left(\frac{k}{n}\right) \right] \binom{n-1}{k} x^k (1-x)^{n-1-k}.$$

   (c) Prove that $B_n(f)'$ converges uniformly to f' on $[0, 1]$.

6. **Simultaneous Approximation.** Let $f \in C^1[0, 1]$. Prove that for every $\epsilon > 0$, there exists a polynomial $P(x)$ such that:

$$|f(x) - P(x)| < \epsilon \quad \text{and} \quad |f'(x) - P'(x)| < \epsilon$$

   for all $x \in [0, 1]$.

   Remark. Approximate f' by a polynomial Q, then integrate.

7. **Limit Evaluation via Series.** Evaluate $\sum_{n=1}^{\infty} \frac{n}{(n+1)!}$.

   Remark. Consider the Taylor series for xe^x.

8. **Approximating Square Roots.** Construct a sequence of polynomials that converges uniformly to $f(x) = \sqrt{x}$ on $[0, 1]$ without using the Bernstein formula directly. Consider the recursive sequence $P_{n+1}(x) = P_n(x) + \frac{1}{2}(x - P_n(x)^2)$ with $P_0(x) = 0$.
   (a) Show that $0 \le P_n(x) \le \sqrt{x}$ for all $x \in [0, 1]$.
   (b) Show that for fixed x, the sequence $P_n(x)$ is increasing.
   (c) Use Dini's Theorem to conclude uniform convergence.

9. Moment Problem. Suppose f is continuous on $[0,1]$ and $\int_0^1 f(x)x^n\,dx = 0$ for all $n = 0, 1, 2, \ldots$.

(a) Use the Weierstrass Approximation Theorem to show that $\int_0^1 f(x)P(x)\,dx = 0$ for any polynomial P.
(b) Show that $\int_0^1 (f(x))^2\,dx = 0$.
(c) Conclude that $f(x) = 0$ for all x.

10. Nowhere Differentiable Approximation. Let f be the continuous nowhere differentiable function defined in the previous chapter. Can f be uniformly approximated by polynomials? Why or why not? What does this say about the "smoothness" of polynomials versus their limits?

11. Separability of $C[0,1]$. A metric space is separable if it contains a countable dense subset. Prove that the space of continuous functions $C[0,1]$ with the uniform metric $d(f,g) = \sup|f-g|$ is separable.

Remark. Consider polynomials with rational coefficients.

12. $\star\star$ **The Covering Lemma.** Let $f : [0,1] \to \mathbb{R}$ be a continuous function. Suppose that $f(0) = f(1) = 0$ and that for every $x \in (0,1)$, $f(x) > 0$. Let K be the graph of f in $\mathbb{R}^2$. Prove that for every $\epsilon > 0$, there exists a polynomial $P(x)$ such that:

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