Advanced Calculus and Linear Algebra I

MATH1071 · Semester 1, 2026

April Kidd

Contents

0 Notation and Logic

Quick reference for the notation used in the rest of these notes. This is not meant to replace a proper logic course; MATH1061 and MATH1081 are better for that. For MATH1071, treat it as a translation table for the symbols that show up in definitions and proofs.

0.1 Mathematical Statements

A **statement** is a sentence that is either true or false. For example, $2 < 5$ is true and $7 = 3$ is false. The phrase "$x < 5$" is not true or false until x has been specified.

Definition 1 (Common logical symbols)

- $P \wedge Q$ means both P and Q are true.
- $P \vee Q$ means at least one of P and Q is true. In mathematics, "or" usually includes the possibility that both are true.
- $\neg P$ means P is false.
- $P \implies Q$ means "if P, then Q".
- $P \iff Q$ means "P if and only if Q"; this says both $P \implies Q$ and $Q \implies P$.

Example 1

The statement

$$x > 2 \wedge x < 5$$

means x lies between 2 and 5. The statement

$$x \leq 2 \vee x \geq 5$$

means x is outside the open interval $(2, 5)$.

0.2 If, Only If, and If and Only If

Students often lose marks by reversing "if" and "only if". Keep the direction explicit.

Definition 2 (Implications)

The statement "P if Q" means

$$Q \implies P.$$

The statement "P only if Q" means

$$P \implies Q.$$

The statement "P if and only if Q" means

$$P \implies Q \quad \text{and} \quad Q \implies P.$$

Example 2

"A number is divisible by 4 only if it is even" means

$$\text{divisible by } \ 4 \implies \text{even.}$$

This is true. The reverse statement is false: 6 is even but not divisible by 4.

Example 3

The equation

$$x^2 = 1$$

is equivalent to

$$x = -1 \lor x = 1.$$

Equivalently,

$$x^2 = 1 \iff x \in \{-1, 1\}.$$

0.3 Quantifiers

Quantifiers say how many objects a statement is talking about.

Definition 3 (For all and there exists)

- $\forall x \in A, P(x)$ means every x in A satisfies $P(x)$.
- $\exists x \in A \text{ such that } P(x)$ means at least one x in A satisfies $P(x)$.
- $\exists! x \in A \text{ such that } P(x)$ means exactly one x in A satisfies $P(x)$.

Example 4

The statement

$$\forall x \in \mathbb{R}, x^2 \geq 0$$

is true. The statement

$$\exists x \in \mathbb{R} \text{ such that } x^2 = 2$$

is true. The statement

$$\exists x \in \mathbb{Q} \text{ such that } x^2 = 2$$

is false.

The order of quantifiers matters.

Example 5

The statement

$$\forall x \in \mathbb{R}, \exists y \in \mathbb{R} \text{ such that } y > x$$

is true: after seeing x, choose $y = x + 1$.

The statement

$$\exists y \in \mathbb{R} \text{ such that } \forall x \in \mathbb{R}, y > x$$

is false: there is no real number larger than every real number.

0.4 Negating Statements

To disprove a statement, write down what its negation says and prove that instead.

Definition 4 (Negating quantifiers)

The negation of

$$\forall x \in A, P(x)$$

is

$$\exists x \in A \ \text{ such that } \ \neg P(x).$$

The negation of

$$\exists x \in A \ \text{ such that } \ P(x)$$

is

$$\forall x \in A, \neg P(x).$$

Example 6
To disprove "every bounded sequence converges", it is enough to give one bounded sequence that does not converge. The sequence $a_n = (-1)^n$ is such a counterexample.

Example 7
The negation of

$$\forall \varepsilon > 0, \exists N \in \mathbb{N} \ \text{ such that } \ n \geq N \implies |a_n - L| < \varepsilon$$

is

$$\exists \varepsilon > 0 \ \text{ such that } \ \forall N \in \mathbb{N}, \exists n \geq N \ \text{ with } \ |a_n - L| \geq \varepsilon.$$

This is the form used when proving a sequence does not converge to L.

0.5 Set Notation

Definition 5 (Set-builder notation)
The notation

$$\{x \in A : P(x)\}$$

means "the set of all x in A such that $P(x)$ is true".

Example 8
The interval $(0, 3)$ can be written as

$$\{x \in \mathbb{R} : 0 < x < 3\}.$$

The set of even integers can be written as

$$\{2k : k \in \mathbb{Z}\}.$$

The symbols $\in$ and $\subseteq$ mean different things. Write $x \in A$ when x is an element of A. Write $A \subseteq B$ when every element of A is also an element of B.

0.6 Common Proof Words

Definition 6 (Proof vocabulary)
- **Assume** introduces a hypothesis being used.
- **Suppose for contradiction** means we temporarily assume the opposite of what we want and derive an impossibility.

- **It follows that** means the next line is a consequence of previous lines.
- **Without loss of generality** means the other cases are genuinely the same after relabelling or symmetry.
- **Counterexample** means one example that shows a universal statement is false.

Important
Do not use "without loss of generality" unless the skipped case is actually the same. For example, if a proof treats $a \le b$, the case $b \le a$ is the same only when the statement is symmetric in a and b.

0.7 Standard Number Sets

The main number systems are:

$$\mathbb{N} = \{1, 2, 3, \ldots\}, \quad \mathbb{Z} = \{\ldots, -2, -1, 0, 1, 2, \ldots\},$$

$$\mathbb{Q} = \left\{ \frac{p}{q} : p \in \mathbb{Z}, q \in \mathbb{N} \right\}, \quad \mathbb{R} = \text{the real numbers}, \quad \mathbb{C} = \text{the complex numbers}.$$

In these notes, $\mathbb{N}$ starts at 1. If a problem uses 0 as a natural number, it should say so.

Note
In limits and proofs, the letters ε and δ usually represent small positive real numbers. The letter N usually represents a large natural number. These are conventions, not new kinds of numbers.

[1]

---

[1] The complex numbers are listed for orientation, but this course works almost entirely with real numbers and will not otherwise use $\mathbb{C}$.

1 Sets, Fields, and Order

The course starts by putting familiar arithmetic on firmer footing. Early MATH1071 is proof-heavy, so it helps to know exactly which properties of the real numbers are being used. Most early arguments use three layers of structure:

- set notation, so that we can say what objects we are discussing;
- field arithmetic, so that addition, multiplication, subtraction, and division behave correctly;
- order and absolute value, so that "close to" can be expressed by inequalities.

The real numbers $\mathbb{R}$ have all of this structure. Later chapters use additional properties of $\mathbb{R}$, especially completeness.

1.1 Sets and Functions

Definition 1 (Set)
A **set** is a collection of objects, called its **elements**. We write $x \in A$ to mean that x is an element of A.

If every element of A is also an element of B, we write $A \subseteq B$ and say that A is a **subset** of B. If $A \subseteq B$ and $A \neq B$, then A is a **proper subset** of B.

Example 1
The set $A = \{1, 2, 3\}$ satisfies $1 \in A$ and $4 \notin A$. Also $\{1, 3\} \subseteq A$.

Sets do not record repeated elements, so $\{1, 1, 2\} = \{1, 2\}$.

Common set operations are:

- $A \cup B = \{x : x \in A \quad \text{or} \quad x \in B\}$, the **union**;
- $A \cap B = \{x : x \in A \quad \text{and} \quad x \in B\}$, the **intersection**;
- $A \setminus B = \{x : x \in A \quad \text{and} \quad x \notin B\}$, the **set difference**;
- $A \times B = \{(a, b) : a \in A, b \in B\}$, the **Cartesian product**.

Definition 2 (Function)
A **function** $f : A \to B$ assigns to each $a \in A$ exactly one element $f(a) \in B$.

The set A is the **domain** and B is the **codomain**. The **range** or **image** of f is

$$f(A) = \{f(a) : a \in A\}.$$

Remark
A formula by itself does not completely specify a function; the domain and codomain matter. For example, $f : \mathbb{R} \to \mathbb{R}$ given by $f(x) = x^2$ and $g : [0, \infty) \to \mathbb{R}$ given by $g(x) = x^2$ have the same formula but different domains.

The standard number systems used in the course are:

Definition 3 (Standard number systems)

- $\mathbb{N} = \{1, 2, 3, ...\}$, the natural numbers;
- $\mathbb{Z} = \{..., -2, -1, 0, 1, 2, ...\}$, the integers;
- $\mathbb{Q} = \left\{\frac{p}{q} : p \in \mathbb{Z}, q \in \mathbb{N}\right\}$, the rational numbers;
- $\mathbb{R}$, the real numbers;

- $\mathbb{C} = \{a + bi : a, b \in \mathbb{R} \text{ and } i^2 = -1\}$, the complex numbers.[2]

We also use the following functions repeatedly.

Definition 4 (Absolute value, floor, and square root)
For $x \in \mathbb{R}$,

$$|x| = \begin{cases} x & \text{if } x \geq 0 \\ -x & \text{if } x < 0 \end{cases}$$

The **floor** $\lfloor x \rfloor$ is the greatest integer less than or equal to x.

For $x \geq 0$, $\sqrt{x}$ is the unique non-negative real number whose square is x.

1.2 Fields

Definition 5 (Field)
A **field** is a set F with two operations, addition and multiplication, such that for all $a, b, c \in F$:

- $a + b \in F$ and $ab \in F$;
- $a + b = b + a$ and $ab = ba$;
- $(a + b) + c = a + (b + c)$ and $(ab)c = a(bc)$;
- there are elements $0, 1 \in F$, with $0 \neq 1$, such that $a + 0 = a$ and $a \cdot 1 = a$;
- for each $a \in F$, there is an element $-a \in F$ such that $a + (-a) = 0$;
- for each $a \in F$ with $a \neq 0$, there is an element $a^{-1} \in F$ such that $aa^{-1} = 1$;
- $a(b + c) = ab + ac$.

The familiar fields are $\mathbb{Q}$, $\mathbb{R}$, and $\mathbb{C}$ with their usual operations. The integers $\mathbb{Z}$ are not a field: $2 \in \mathbb{Z}$, but there is no integer b such that $2b = 1$.

Important
When proving a statement about an arbitrary field, do not use order, decimal expansions, or geometric intuition unless those structures have been assumed. A field proof should only use the field axioms and consequences already proved from them.

1.3 Basic Consequences of the Field Axioms

The following facts are normally used without comment in high-school algebra. In this course, it is worth seeing that they really do follow from the field axioms.

Proposition 1 (Identities and inverses are unique)
Let F be a field.

1. The additive identity is unique.
2. The multiplicative identity is unique.
3. For each $a \in F$, the additive inverse $-a$ is unique.
4. If $a \neq 0$, then the multiplicative inverse a^{-1} is unique.

Proof. Suppose 0 and $0'$ are both additive identities. Then

---

[2] The complex numbers are included here for orientation. This course works almost entirely with real numbers and will not otherwise use $\mathbb{C}$.

$$0 = 0 + 0' = 0',$$

because $0'$ is an additive identity in the first equality and 0 is an additive identity in the second.

The proof for the multiplicative identity is analogous: if 1 and $1'$ are both multiplicative identities, then $1 = 1 \cdot 1' = 1'$.

If u and v are both additive inverses of a, then

$$u = u + 0 = u + (a + v) = (u + a) + v = 0 + v = v.$$

If $a \neq 0$ and u, v are both multiplicative inverses of a, then

$$u = u \cdot 1 = u(av) = (ua)v = 1 \cdot v = v.$$

□

Proposition 2 (Zero and negatives)
Let F be a field and let $a, b \in F$. Then:

1. $a \cdot 0 = 0$;
2. $(-1)(-1) = 1$;
3. $a(-b) = (-a)b = -(ab)$;
4. $(-a)(-b) = ab$.

Proof. Since $0 + 0 = 0$, distributivity gives

$$a \cdot 0 = a(0 + 0) = a \cdot 0 + a \cdot 0.$$

Adding $-(a \cdot 0)$ to both sides gives $0 = a \cdot 0$.

Now $(-1) + 1 = 0$. Multiplying by -1 gives

$$(-1)(-1) + (-1) = 0.$$

Hence $(-1)(-1)$ is the additive inverse of -1, which is 1.

To prove $a(-b) = -(ab)$, observe that

$$ab + a(-b) = a(b + (-b)) = a \cdot 0 = 0.$$

Thus $a(-b)$ is the additive inverse of ab. The proof of $(-a)b = -(ab)$ is the same. Finally,

$$(-a)(-b) = -(a(-b)) = -(-(ab)) = ab.$$

□

Proposition 3 (No zero divisors)
If F is a field and $ab = 0$, then $a = 0$ or $b = 0$.

Proof. If $a = 0$, there is nothing to prove. Suppose $a \neq 0$. Then a^{-1} exists, and

$$b = 1 \cdot b = (a^{-1}a)b = a^{-1}(ab) = a^{-1} \cdot 0 = 0.$$

□

Corollary 4
If $a \in F$ and $a^2 = 1$, then $a \in \{-1, 1\}$.

Proof. Since $a^2 = 1$,

$$0 = a^2 - 1 = (a-1)(a+1).$$

By the no-zero-divisors property, $a - 1 = 0$ or $a + 1 = 0$. Therefore $a = 1$ or $a = -1$. ☐

1.4 Examples of Fields and Non-Fields

Example 2 ($\mathbb{Q}(\sqrt{2})$)
Define

$$\mathbb{Q}(\sqrt{2}) = \left\{a + b\sqrt{2} : a, b \in \mathbb{Q}\right\}.$$

This is a field under the usual addition and multiplication inherited from $\mathbb{R}$.

Addition and multiplication stay inside the set:

$$\left(a + b\sqrt{2}\right) + \left(c + d\sqrt{2}\right) = (a + c) + (b + d)\sqrt{2},$$

and

$$\left(a + b\sqrt{2}\right)\left(c + d\sqrt{2}\right) = (ac + 2bd) + (ad + bc)\sqrt{2}.$$

The identities are $0 = 0 + 0\sqrt{2}$ and $1 = 1 + 0\sqrt{2}$. The additive inverse of $a + b\sqrt{2}$ is $-a - b\sqrt{2}$.

For multiplicative inverses, suppose $a + b\sqrt{2} \neq 0$. Then

$$\frac{1}{a + b\sqrt{2}} = \frac{a - b\sqrt{2}}{a^2 - 2b^2} = \frac{a}{a^2 - 2b^2} - \frac{b}{a^2 - 2b^2}\sqrt{2}.$$

The coefficients are rational. Also $a^2 - 2b^2 \neq 0$: if it were zero and $b \neq 0$, then $\left(\frac{a}{b}\right)^2 = 2$, contradicting the irrationality of $\sqrt{2}$; if $b = 0$, then $a = 0$, contrary to $a + b\sqrt{2} \neq 0$.

Example 3
To write $\left(3 + 4\sqrt{2}\right)^{-1}$ in the form $a + b\sqrt{2}$,

$$\frac{1}{3 + 4\sqrt{2}} = \frac{3 - 4\sqrt{2}}{9 - 32} = -\frac{3}{23} + \frac{4}{23}\sqrt{2}.$$

Example 4 (Integers modulo 3)
The set $\mathrm{Z}_3 = \{0, 1, 2\}$ is a field when addition and multiplication are performed modulo 3.

+	0	1	2
0	0	1	2
1	1	2	0
2	2	0	1

$\cdot$	0	1	2
0	0	0	0
1	0	1	2
2	0	2	1

The additive identity is 0 and the multiplicative identity is 1. Every element has an additive inverse: 0 is its own inverse, while $1 + 2 = 0$ modulo 3. Every nonzero element has a multiplicative inverse: $1^{-1} = 1$ and $2^{-1} = 2$ because $2 \cdot 2 = 1$ modulo 3.

Example 5 (Integers modulo 6)
The set $\mathbb{Z}_6 = \{0, 1, 2, 3, 4, 5\}$ is not a field. In modulo 6 arithmetic,

$$2 \cdot 3 = 6 = 0.$$

But $2 \neq 0$ and $3 \neq 0$. This violates the no-zero-divisors property, so $\mathbb{Z}_6$ cannot be a field.

Remark
More generally, $\mathbb{Z}_p$ is a field when p is prime. If n is composite, then $\mathbb{Z}_n$ has nonzero zero divisors and is not a field.

1.5 Ordered Fields

Field axioms allow addition, subtraction, multiplication, and division. They do not allow us to say that one element is larger than another. For limits and inequalities we need compatible order.

Definition 6 (Ordered field)
An **ordered field** is a field F with a subset $P \subset F$, whose elements are called **positive**, such that:

- for every $a \in F$, exactly one of $a \in P$, $a = 0$, or $-a \in P$ holds;
- if $a, b \in P$, then $a + b \in P$;
- if $a, b \in P$, then $ab \in P$.

We write $a > 0$ when $a \in P$, and $a > b$ when $a - b > 0$.

We write $a < b$ to mean $b > a$, and $a \leq b$ to mean $a < b$ or $a = b$.

Proposition 5 (Order rules)
In an ordered field:

1. If $a > b$ and $b > c$, then $a > c$.
2. If $a > b$, then $a + c > b + c$.
3. If $a > b$ and $c > 0$, then $ac > bc$.
4. If $a > b$ and $c < 0$, then $ac < bc$.
5. If $a \neq 0$, then $a^2 > 0$.
6. $1 > 0$.

Proof. If $a > b$ and $b > c$, then $a - b > 0$ and $b - c > 0$. Since positives are closed under addition,

$$a - c = (a - b) + (b - c) > 0,$$

so $a > c$.

Addition preserves inequalities because $(a + c) - (b + c) = a - b$.

If $c > 0$, then $(a - b)c > 0$, so $ac > bc$. If $c < 0$, then $-c > 0$, so $a(-c) > b(-c)$, and multiplying both sides by -1 reverses the inequality.

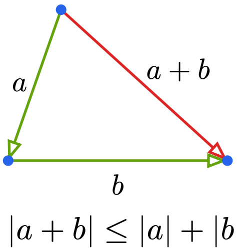
If $a \neq 0$, exactly one of $a > 0$ or $-a > 0$ holds. In the first case $a^2 > 0$; in the second case $(-a)(-a) > 0$, and this is equal to a^2. Since $1 = 1^2$ and $1 \neq 0$, we get $1 > 0$. ☐

1.6 Absolute Value and Distance

Absolute value converts order into distance. Most epsilon arguments in the course are just careful uses of absolute value inequalities.

Proposition 6 (Absolute-value rules)
For all $x, y \in \mathbb{R}$:

1. $|x| \geq 0$, and $|x| = 0$ if and only if $x = 0$.
2. $|xy| = |x||y|$.
3. $|x + y| \leq |x| + |y|$.
4. $||x| - |y|| \leq |x - y|$.



Visual representation of the triangle inequality. The direct path $|a + b|$ cannot be longer than travelling along $|a|$ and then $|b|$.

Proof. The first two statements follow by checking the cases $x \geq 0$ and $x < 0$.

For the triangle inequality, note that $-|x| \leq x \leq |x|$ and $-|y| \leq y \leq |y|$. Adding these inequalities gives

$$-(|x| + |y|) \leq x + y \leq |x| + |y|,$$

so $|x + y| \leq |x| + |y|$.

Since $x = (x - y) + y$, the triangle inequality gives

$$|x| \leq |x - y| + |y|,$$

hence $|x| - |y| \leq |x - y|$. Swapping x and y gives $|y| - |x| \leq |x - y|$. Together, these are exactly

$$||x| - |y|| \leq |x - y|.$$

☐

Corollary 7
For all $x, y \in \mathbb{R}$,

$$|x| - |y| \leq |x - y|$$

and

$$|x| - |y| \leq |x + y|.$$

Proof. The first inequality is one side of the previous proposition. For the second, apply the first inequality with $-y$ in place of y:

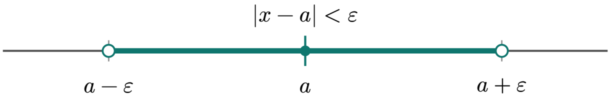
$$|x| - |-y| \leq |x - (-y)| = |x + y|.$$

Since $|-y| = |y|$, the result follows. ☐

Important
The inequality $|x - a| < \varepsilon$ means

$$a - \varepsilon < x < a + \varepsilon.$$

In words: x is within distance ε of a. This is the language of sequence limits and function limits.



The inequality $|x - a| < \varepsilon$ describes the open interval centred at a with radius ε.

1.7 The Archimedean Property

The ordered-field axioms alone do not capture everything special about $\mathbb{R}$. One extra fact used constantly is that the natural numbers are not bounded above inside $\mathbb{R}$.

Definition 7 (Archimedean property)
An ordered field F has the **Archimedean property** if for every $x \in F$, there exists $n \in \mathbb{N}$ such that $n > x$.

Theorem 8 (Archimedean consequences in $\mathbb{R}$)
In $\mathbb{R}$:

1. If $x, y \in \mathbb{R}$ and $x > 0$, then there exists $n \in \mathbb{N}$ such that $nx > y$.
2. If $w > 0$, then there exists $n \in \mathbb{N}$ such that $0 < \frac{1}{n} < w$.

Proof. For the first statement, apply the Archimedean property to $\frac{y}{x}$. There exists $n \in \mathbb{N}$ such that $n > \frac{y}{x}$. Since $x > 0$, multiplying by x gives $nx > y$.

For the second statement, apply the Archimedean property to $\frac{1}{w}$. There exists $n \in \mathbb{N}$ such that $n > \frac{1}{w}$. Since $n > 0$ and $w > 0$, this implies $\frac{1}{n} < w$. Also $\frac{1}{n} > 0$. ☐

1.8 Irrationality Proofs

Irrationality arguments are early examples of proof by contradiction. The key number-theoretic fact used below is Euclid's lemma in a simple form: if a prime p divides a^2, then p divides a.

Proposition 9
The numbers $\sqrt{2}$ and $\sqrt{5}$ are irrational.

Proof. We prove the statement for $\sqrt{2}$. Suppose, for contradiction, that $\sqrt{2} = \frac{p}{q}$ for integers p, q with $q \neq 0$, chosen so that the fraction is in lowest terms. Then

$$p^2 = 2q^2.$$

Thus p^2 is even, so p is even. Write $p = 2k$. Then

$$4k^2 = 2q^2,$$

so $q^2 = 2k^2$. Hence q is even. This contradicts the assumption that $\frac{p}{q}$ was in lowest terms.

The proof for $\sqrt{5}$ is analogous. If $\sqrt{5} = \frac{p}{q}$ in lowest terms, then $p^2 = 5q^2$, so 5 divides p, say $p = 5k$. Then $25k^2 = 5q^2$, so $q^2 = 5k^2$, and 5 divides q. This again contradicts lowest terms. □

2 Bounds, Sequences, and Limits

Limits are where the course first gets properly analytical. Saying that a sequence "gets close" to a number is only intuition; the definition below is the version that can actually be used in a proof.

For most sequence-limit questions, separate two steps:

- What number should the limit be?
- Given an arbitrary tolerance $\varepsilon > 0$, how far along the sequence do we need to go before every later term is inside that tolerance?

2.1 Upper and Lower Bounds

Definition 1 (Bounds)
Let $\Omega \subset \mathbb{R}$.

A number $M \in \mathbb{R}$ is an **upper bound** for Ω if $x \leq M$ for every $x \in \Omega$.

A number $m \in \mathbb{R}$ is a **lower bound** for Ω if $m \leq x$ for every $x \in \Omega$.

The set Ω is **bounded above** if it has an upper bound, **bounded below** if it has a lower bound, and **bounded** if it is both bounded above and bounded below.

Example 1
The interval $(0, 3)$ is bounded above and below. For example, 4 is an upper bound and -1 is a lower bound.

The set $\mathbb{N}$ is bounded below, for instance by 1 or 0, but is not bounded above in $\mathbb{R}$.

Bounds need not be elements of the set. The interval $(0, 3)$ has upper bound 3, but $3 \notin (0, 3)$.

Definition 2 (Supremum and infimum)
Let $\Omega \subset \mathbb{R}$.

A number S is the **supremum** or **least upper bound** of Ω, written $S = \sup \Omega$, if:

1. S is an upper bound for Ω;
2. if M is any upper bound for Ω, then $S \leq M$.

A number I is the **infimum** or **greatest lower bound** of Ω, written $I = \inf \Omega$, if:

1. I is a lower bound for Ω;
2. if m is any lower bound for Ω, then $m \leq I$.

Example 2

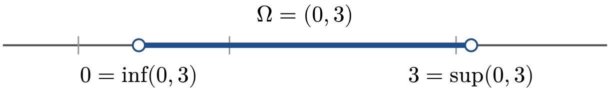
$$\sup(0, 3) = 3, \quad \inf(0, 3) = 0.$$

The set $(0, 3)$ has no maximum or minimum, because neither endpoint belongs to the set.

Example 3
Let

$$\Omega = \{-1\} \cup (0, 3) \cup \{7\} \cup \{13\}.$$

Then $\sup\Omega = 13$ and $\inf\Omega = -1$. In this case the supremum and infimum are also the maximum and minimum.



For $(0,3)$ the supremum and infimum are endpoints approached by the set, even though neither endpoint belongs to it.

Proposition 1
If a supremum or infimum exists, it is unique.

Proof. Suppose S and T are both suprema of Ω. Since S is a supremum and T is an upper bound, $S \le T$. Since T is a supremum and S is an upper bound, $T \le S$. Hence $S = T$.

The proof for infima is the same with the inequalities reversed. □

Theorem 2 (Least upper bound property of $\mathbb{R}$)
Every non-empty subset of $\mathbb{R}$ that is bounded above has a supremum in $\mathbb{R}$.

Remark
This theorem is a completeness property of the real numbers. It is not true in $\mathbb{Q}$. For example,

$$\{q \in \mathbb{Q} : q^2 < 2\}$$

is non-empty and bounded above in $\mathbb{Q}$, but it has no supremum in $\mathbb{Q}$.

Theorem 3 (Epsilon characterisation of the supremum)
Let $\Omega \subset \mathbb{R}$ be non-empty and bounded above. Then $S = \sup\Omega$ if and only if:

1. S is an upper bound for Ω;
2. for every $\varepsilon > 0$, there exists $x \in \Omega$ such that $S - \varepsilon < x \le S$.

Proof. Suppose first that $S = \sup\Omega$. Then S is an upper bound. Let $\varepsilon > 0$. If there were no $x \in \Omega$ with $S - \varepsilon < x$, then every $x \in \Omega$ would satisfy $x \le S - \varepsilon$. Thus $S - \varepsilon$ would be an upper bound smaller than S, contradicting that S is the least upper bound.

Conversely, suppose the two stated conditions hold. Since S is already an upper bound, it remains to show that no smaller number is an upper bound. Let $M < S$. Put $\varepsilon = S - M > 0$. By the second condition, there exists $x \in \Omega$ such that

$$M = S - \varepsilon < x.$$

Hence M is not an upper bound. Therefore every upper bound is at least S, so $S = \sup\Omega$. □

2.2 Sequences

Definition 3 (Sequence)
A **real sequence** is a function $a : \mathbb{N} \to \mathbb{R}$. We usually write a_n instead of $a(n)$, and denote the whole sequence by $(a_n)_{n=1}^{\infty}$ or simply (a_n).

Important
The notation a_n refers to the nth term. The notation (a_n) refers to the whole sequence.

Example 4
The formula $a_n = \frac{1}{n}$ defines the sequence

$$1, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \ldots$$

The terms approach 0, although no term is equal to 0.

Definition 4 (Bounded sequence)
A sequence (a_n) is:

- **bounded above** if there exists $M \in \mathbb{R}$ such that $a_n \leq M$ for all $n \in \mathbb{N}$;
- **bounded below** if there exists $m \in \mathbb{R}$ such that $m \leq a_n$ for all $n \in \mathbb{N}$;
- **bounded** if it is bounded above and bounded below.

Proposition 4
A sequence (a_n) is bounded if and only if there exists $C > 0$ such that $|a_n| \leq C$ for all $n \in \mathbb{N}$.

Proof. If $|a_n| \leq C$ for all n, then $-C \leq a_n \leq C$, so the sequence is bounded.

Conversely, suppose $m \leq a_n \leq M$ for all n. Let $C = \max(|m|, |M|, 1)$. Then $-C \leq a_n \leq C$ for all n, hence $|a_n| \leq C$. $\square$

2.3 Limits of Sequences

Definition 5 (Limit of a sequence)
Let (a_n) be a real sequence and let $L \in \mathbb{R}$. We say that (a_n) **converges to** L, and write

$$\lim_{n \to \infty} a_n = L,$$

if for every $\varepsilon > 0$ there exists $N \in \mathbb{N}$ such that whenever $n \geq N$,

$$|a_n - L| < \varepsilon.$$

If no such L exists, the sequence **diverges**.

In this definition, ε is chosen first. It represents the permitted error. The integer N may depend on ε, but once N is chosen, every later term must work.

Definition 6 (Epsilon neighbourhood)
For $a \in \mathbb{R}$ and $\varepsilon > 0$, the set

$$\{x \in \mathbb{R} : |x - a| < \varepsilon\}$$

is the **epsilon neighbourhood** of a.

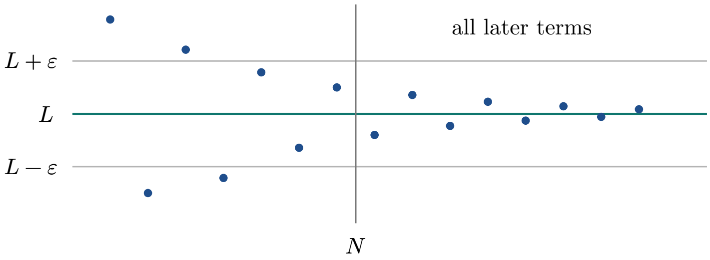
Since

$$|x-a|<\varepsilon$$

is equivalent to

$$a-\varepsilon<x<a+\varepsilon,$$

convergence means that every epsilon neighbourhood of L eventually contains the whole tail of the sequence.



A sequence converges to L if, for every band of width ε around L, all sufficiently late terms lie inside the band.

Example 5 (A basic epsilon proof)
We prove that

$$\lim_{n\to\infty}\frac{1}{n}=0.$$

Let $\varepsilon>0$. Choose $N\in\mathbb{N}$ such that $N>\frac{1}{\varepsilon}$, possible by the Archimedean property. If $n\geq N$, then

$$\left|\frac{1}{n}-0\right|=\frac{1}{n}\leq\frac{1}{N}<\varepsilon.$$

Therefore $\frac{1}{n}\to 0$.

Example 6
We prove that

$$\lim_{n\to\infty}\frac{n}{n+1}=1.$$

Let $\varepsilon>0$. Choose $N\in\mathbb{N}$ such that $N>\frac{1}{\varepsilon}$. If $n\geq N$, then

$$\left|\frac{n}{n+1}-1\right|=\left|\frac{n-n-1}{n+1}\right|=\frac{1}{n+1}\leq\frac{1}{n}\leq\frac{1}{N}<\varepsilon.$$

Example 7
We prove that

$$\lim_{n\to\infty} \frac{1}{\sqrt{n}} = 0.$$

Let $\varepsilon > 0$. Choose $N \in \mathbb{N}$ such that $N > \frac{1}{\varepsilon^2}$. If $n \geq N$, then

$$\left|\frac{1}{\sqrt{n}} - 0\right| = \frac{1}{\sqrt{n}} \leq \frac{1}{\sqrt{N}} < \varepsilon.$$

Example 8
We prove that

$$\lim_{n\to\infty} \frac{1}{n^4+1} = 0.$$

Let $\varepsilon > 0$. Choose $N \in \mathbb{N}$ such that $N > \varepsilon^{-\frac{1}{4}}$. If $n \geq N$, then

$$\left|\frac{1}{n^4+1} - 0\right| = \frac{1}{n^4+1} \leq \frac{1}{n^4} \leq \frac{1}{N^4} < \varepsilon.$$

Note
In written proofs it is enough to say “choose $N \in \mathbb{N}$ such that $N > \frac{1}{\varepsilon}$” rather than writing a floor or ceiling formula. The Archimedean property guarantees such an N exists.

2.4 Uniqueness and Boundedness

Theorem 5 (Uniqueness of limits)
If a sequence converges, then its limit is unique.

Proof. Suppose $a_n \to L$ and $a_n \to M$. Let $\varepsilon > 0$. Since $a_n \to L$, there exists N_1 such that $|a_n - L| < \frac{\varepsilon}{2}$ whenever $n \geq N_1$. Since $a_n \to M$, there exists N_2 such that $|a_n - M| < \frac{\varepsilon}{2}$ whenever $n \geq N_2$.

If $n \geq \max(N_1, N_2)$, then

$$|L - M| = |L - a_n + a_n - M| \leq |L - a_n| + |a_n - M| < \varepsilon.$$

Since this holds for every $\varepsilon > 0$, we must have $|L - M| = 0$, and hence $L = M$. ☐

Theorem 6 (Convergent sequences are bounded)
Every convergent real sequence is bounded.

Proof. Suppose $a_n \to L$. Apply the definition of convergence with $\varepsilon = 1$. There exists $N \in \mathbb{N}$ such that $|a_n - L| < 1$ whenever $n \geq N$.

For $n \geq N$,

$$|a_n| = |a_n - L + L| \leq |a_n - L| + |L| < 1 + |L|.$$

The first $N - 1$ terms form a finite list, so they have a maximum absolute value. Let

$$C = \max(|a_1|, |a_2|, ..., |a_{N-1}|, 1 + |L|).$$

Then $|a_n| \leq C$ for every n, so (a_n) is bounded. □

Important
Boundedness is necessary for convergence, but not sufficient. The sequence $a_n = (-1)^n$ is bounded, but it does not converge.

2.5 Limit Laws

Theorem 7 (Algebra of limits)
Suppose $a_n \to a$, $b_n \to b$, and $\lambda \in \mathbb{R}$. Then:

1. $a_n + b_n \to a + b$;
2. $\lambda a_n \to \lambda a$;
3. $a_n b_n \to ab$;
4. if $b \neq 0$ and $b_n \neq 0$ eventually, then $\frac{a_n}{b_n} \to \frac{a}{b}$.

Proof. For sums, let $\varepsilon > 0$. Choose N_1 such that $|a_n - a| < \frac{\varepsilon}{2}$ for $n \geq N_1$, and choose N_2 such that $|b_n - b| < \frac{\varepsilon}{2}$ for $n \geq N_2$. If $n \geq \max(N_1, N_2)$, then

$$|(a_n + b_n) - (a + b)| \leq |a_n - a| + |b_n - b| < \varepsilon.$$

Scalar multiplication is immediate if $\lambda = 0$. If $\lambda \neq 0$, choose N such that $|a_n - a| < \frac{\varepsilon}{|\lambda|}$ for $n \geq N$.

For products, write[3]

$$a_n b_n - ab = a_{n(b_n - b)} + b(a_n - a).$$

Since (a_n) converges, it is bounded. Choose $C > 0$ such that $|a_n| \leq C$ for all n. Then

$$|a_n b_n - ab| \leq C|b_n - b| + |b||a_n - a|.$$

Choose N_1 so that $C|b_n - b| < \frac{\varepsilon}{2}$ for $n \geq N_1$, and choose N_2 so that $|b||a_n - a| < \frac{\varepsilon}{2}$ for $n \geq N_2$. This proves the product law.

For reciprocals, suppose $b_n \to b \neq 0$. Choose N_1 such that $|b_n - b| < \frac{|b|}{2}$ for $n \geq N_1$. By the reverse triangle inequality,

$$|b_n| \geq |b| - |b_n - b| > \frac{|b|}{2}.$$

Hence, for $n \geq N_1$,

$$\left|\frac{1}{b_n} - \frac{1}{b}\right| = \frac{|b - b_n|}{|b_n||b|} < 2\frac{|b_n - b|}{|b|^2}.$$

Since $b_n \to b$, the final expression can be made smaller than any $\varepsilon > 0$. Thus $\frac{1}{b_n} \to \frac{1}{b}$. The quotient law follows from the product law applied to $a_n \cdot \left(\frac{1}{b_n}\right)$. □

---

[3] This is the standard "add zero" trick: insert a term and subtract it again so the expression splits into pieces we already know how to control. Here

$$a_n b_n - ab = a_n b_n - a_n b + a_n b - ab.$$

This kind of rearrangement appears constantly in analysis.

Example 9
Find

$$\lim_{n\to\infty} \frac{3n^2+2n}{5n^2+3}.$$

Divide numerator and denominator by n^2:

$$\frac{3n^2+2n}{5n^2+3} = \frac{3+\frac{2}{n}}{5+\frac{3}{n^2}}.$$

Since $\frac{1}{n} \to 0$ and $\frac{1}{n^2} \to 0$, the limit laws give

$$\lim_{n\to\infty} \frac{3n^2+2n}{5n^2+3} = \frac{3}{5}.$$

Proposition 8 (Useful permanence facts)
Let $a_n \to a$.

- $|a_n| \to |a|$.
- If $a \geq 0$ and $a_n \geq 0$ eventually, then $\sqrt{a_n} \to \sqrt{a}$.
- If b_n is bounded and $a_n \to 0$, then $a_n b_n \to 0$.
- If $c_n \geq 0$, $c_n \to 0$, and $|b_n - L| \leq c_n$ eventually, then $b_n \to L$.

Proof. The first statement follows from the reverse triangle inequality,

$$||a_n| - |a|| \leq |a_n - a|.$$

For square roots, if $a > 0$ then

$$\left|\sqrt{a_n} - \sqrt{a}\right| = \frac{|a_n - a|}{\sqrt{a_n} + \sqrt{a}},$$

and the denominator is eventually bounded below by $\frac{\sqrt{a}}{2}$. If $a = 0$, then $\left|\sqrt{a_n} - 0\right| = \sqrt{a_n}$ and $a_n \to 0$. Given $\varepsilon > 0$, eventually $0 \leq a_n < \varepsilon^2$, so $\sqrt{a_n} < \varepsilon$.

If $|b_n| \leq M$, then $|a_n b_n| \leq M|a_n| \to 0$. The final statement is exactly the squeeze theorem applied to $-c_n \leq b_n - L \leq c_n$. ☐

Proposition 9 (Polynomial quotients)
If $b_k \neq 0$, then

$$\lim_{n\to\infty} \frac{a_k n^k + a_{k-1} n^{k-1} + \ldots + a_1 n + a_0}{b_k n^k + b_{k-1} n^{k-1} + \ldots + b_1 n + b_0} = \frac{a_k}{b_k}.$$

Proof. Divide numerator and denominator by n^k. Every lower-order term contains a positive power of $\frac{1}{n}$, and therefore converges to 0. ☐

2.6 Squeeze Arguments

Theorem 10 (Squeeze theorem)
Suppose $a_n \leq b_n \leq c_n$ for all sufficiently large n. If $a_n \to L$ and $c_n \to L$, then $b_n \to L$.

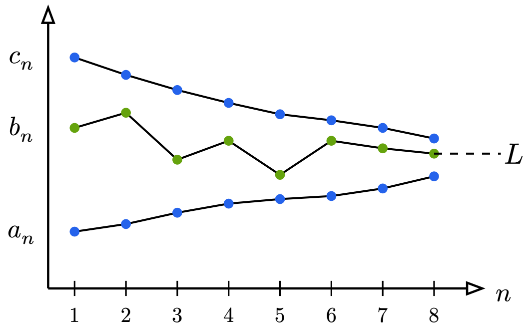


Illustration of the squeeze theorem. The sequence (b_n) is trapped between (a_n) and (c_n), and both bounding sequences converge to L.

Proof. Let $\varepsilon > 0$. Since $a_n \to L$ and $c_n \to L$, there exists N such that for all $n \geq N$,

$$L - \varepsilon < a_n \leq b_n \leq c_n < L + \varepsilon.$$

Therefore $|b_n - L| < \varepsilon$ for all $n \geq N$, so $b_n \to L$. □

Example 10
Since $-1 \leq \sin(n) \leq 1$,

$$-\frac{1}{n^2} \leq \frac{\sin(n)}{n^2} \leq \frac{1}{n^2}.$$

The outer sequences both converge to 0, so

$$\lim_{n\to\infty} \frac{\sin(n)}{n^2} = 0.$$

Example 11
Find

$$\lim_{n\to\infty} \frac{n-1}{n^4+2}.$$

For $n \geq 1$,

$$0 \leq \frac{n-1}{n^4+2} \leq \frac{n}{n^4} = \frac{1}{n^3}.$$

Since $\frac{1}{n^3} \to 0$, the squeeze theorem gives

$$\lim_{n\to\infty} \frac{n-1}{n^4+2} = 0.$$

Example 12
Rationalise to find

$$\lim_{n\to\infty} \left(\sqrt{n+1} - \sqrt{n}\right).$$

We have

$$\sqrt{n+1}-\sqrt{n}=\frac{\left(\sqrt{n+1}-\sqrt{n}\right)\left(\sqrt{n+1}+\sqrt{n}\right)}{\sqrt{n+1}+\sqrt{n}}=\frac{1}{\sqrt{n+1}+\sqrt{n}}.$$

Hence

$$0\leq\sqrt{n+1}-\sqrt{n}\leq\frac{1}{\sqrt{n}},$$

and the right-hand side converges to 0. Therefore the required limit is 0.

Proposition 11
If $a_n \to 0$ and (b_n) is bounded, then $a_n b_n \to 0$.

Proof. Since (b_n) is bounded, choose $C > 0$ such that $|b_n| \leq C$ for all n. Then

$$|a_n b_n| \leq C|a_n|.$$

Since $C|a_n| \to 0$, the squeeze theorem gives $a_n b_n \to 0$. □

2.7 Divergence

To prove a sequence diverges, it is often enough to show that convergence would contradict either uniqueness of limits or boundedness.

Example 13
The sequence $a_n = 3 - (-1)^n$ diverges.

If n is odd, then $a_n = 4$. If n is even, then $a_n = 2$. Suppose $a_n \to L$. Taking $\varepsilon = \frac{1}{2}$, all sufficiently late terms would need to lie in $\left(L-\frac{1}{2}, L+\frac{1}{2}\right)$. But arbitrarily late odd terms equal 4 and arbitrarily late even terms equal 2, which are distance 2 apart. No interval of radius $\frac{1}{2}$ can contain both. Therefore the sequence diverges.

Example 14
The sequence $a_n = n^3$ diverges. If it converged, it would be bounded. But for every $C > 0$, the Archimedean property gives $n \in \mathbb{N}$ with $n > \sqrt[3]{C}$, and then $n^3 > C$. Thus $\left(n^3\right)$ is not bounded.

Example 15
The sequence

$$a_n = \frac{n^6+1}{n^3}$$

diverges, because

$$a_n = n^3 + \frac{1}{n^3}.$$

The term n^3 is unbounded and $\frac{1}{n^3} \geq 0$, so a_n is unbounded. Hence it cannot converge.

Note

Two divergent sequences can have a convergent sum. For example, $a_n = n$ and $b_n = -n$ both diverge, but $a_n + b_n = 0$ for all n.

A convergent sequence multiplied by a divergent sequence may converge or diverge. For example, $(0) \cdot n = 0$ converges, while $1 \cdot n = n$ diverges.

2.8 Standard Limits

The following limits are used often. Some are proved later using monotonicity, exponentials, logarithms, or series.

Theorem 12 (Useful standard limits)

1. If $|\lambda| < 1$, then $\lambda^n \to 0$.
2. If $c > 0$, then $\sqrt[n]{c} \to 1$.
3. $\sqrt[n]{n} \to 1$.

Example 16

Since $0 < \frac{1}{2} < 1$,

$$2^{-n} = \left(\frac{1}{2}\right)^n \to 0.$$

Therefore $\cos^2 \frac{n}{2^n} \to 0$ by the bounded-times-null result, because $0 \leq \cos^2(n) \leq 1$.

3 More Sequences

The previous chapter dealt mostly with sequences that have a single limit. Here we look at what remains useful when a sequence does not settle down: monotonicity, subsequences, upper and lower limits, cluster points, and Cauchy sequences.

3.1 Monotone Sequences

Definition 1 (Monotone sequence)
A sequence (a_n) is:

- **increasing** if $a_{n+1} \geq a_n$ for every $n \in \mathbb{N}$;
- **decreasing** if $a_{n+1} \leq a_n$ for every $n \in \mathbb{N}$;
- **strictly increasing** if $a_{n+1} > a_n$ for every $n \in \mathbb{N}$;
- **strictly decreasing** if $a_{n+1} < a_n$ for every $n \in \mathbb{N}$.

A sequence is **monotone** if it is increasing or decreasing.

The monotone sequences are the easiest sequences to understand: if they are prevented from escaping to infinity, they must converge.

Theorem 1 (Monotone convergence theorem)
A monotone real sequence converges if and only if it is bounded.

Proof. Every convergent sequence is bounded by the convergent sequences are bounded theorem.

Conversely, suppose (a_n) is increasing and bounded. Let

$$\alpha = \sup\{a_n : n \in \mathbb{N}\}.$$

We prove $a_n \to \alpha$. Let $\varepsilon > 0$. Since α is the supremum, there exists $N \in \mathbb{N}$ such that

$$\alpha - \varepsilon < a_N \leq \alpha.$$

Since (a_n) is increasing, every $n \geq N$ satisfies

$$\alpha - \varepsilon < a_N \leq a_n \leq \alpha < \alpha + \varepsilon.$$

Hence $|a_n - \alpha| < \varepsilon$ for every $n \geq N$, so $a_n \to \alpha$.

If (a_n) is decreasing and bounded, apply the increasing case to $(-a_n)$. □

Example 1
Define $a_1 = 2$ and

$$a_{n+1} = \frac{1}{2}(a_n + 6).$$

We show that (a_n) converges.

First, $a_n < 6$ for all n. This is true for $n = 1$. If $a_n < 6$, then

$$a_{n+1} = \frac{1}{2}(a_n + 6) < \frac{1}{2}(6 + 6) = 6.$$

Also,

$$a_{n+1} - a_n = \frac{1}{2}(6 - a_n) > 0,$$

so (a_n) is strictly increasing. It is increasing and bounded above, so it converges by the monotone convergence theorem.

Example 2
Let $a_1 = 1$ and

$$a_{n+1} = \frac{1 + a_n}{2 + a_n}.$$

The fixed points satisfy $L = \frac{1+L}{2+L}$, hence

$$L^2 + L - 1 = 0.$$

Put $\alpha = \frac{\sqrt{5}-1}{2}$, the positive fixed point. If $a_n > \alpha$, then

$$a_{n+1} - \alpha = \frac{(1 - \alpha)(a_n - \alpha)}{2 + a_n} > 0,$$

so the sequence stays above α. Also

$$a_{n+1} < a_n \quad \text{is equivalent to} \quad a_n^2 + a_n - 1 > 0,$$

which holds whenever $a_n > \alpha$. Thus (a_n) is decreasing and bounded below by α, so it converges. Taking limits in the recurrence gives

$$\lim_{n\to\infty} a_n = \alpha.$$

Definition 2 (Limits to infinity)
We write $a_n \to \infty$ if for every $M > 0$ there exists $N \in \mathbb{N}$ such that $a_n > M$ whenever $n \geq N$.

Similarly, $a_n \to -\infty$ if for every $M < 0$ there exists $N \in \mathbb{N}$ such that $a_n < M$ whenever $n \geq N$.

Example 3
The sequence $a_n = n^3$ satisfies $a_n \to \infty$. The sequence $b_n = -n$ satisfies $b_n \to -\infty$.

3.2 Upper and Lower Limits

Even when a bounded sequence does not converge, its tails may still have a highest long-term level and a lowest long-term level.

Definition 3 (Limit superior and limit inferior)
Let (a_n) be a bounded real sequence. Define

$$x_k = \sup\{a_n : n \geq k\}, \quad y_k = \inf\{a_n : n \geq k\}.$$

The **limit superior** and **limit inferior** of (a_n) are

$$\limsup_{n\to\infty} a_n = \lim_{k\to\infty} x_k, \quad \liminf_{n\to\infty} a_n = \lim_{k\to\infty} y_k.$$

The sequence (x_k) is decreasing: as k increases, the tail loses terms, so its supremum cannot go up. Similarly, (y_k) is increasing. Since (a_n) is bounded, both (x_k) and (y_k) are bounded, so the monotone convergence theorem guarantees the two limits exist.

Note
For unbounded sequences, the same tail-supremum and tail-infimum idea is often interpreted in the extended real numbers. For example, if the tail suprema tend to ∞, we write $\limsup a_n = \infty$.

Example 4
Let

$$a_n = \begin{cases} -1 + 3^{-n} & \text{if } \ n \ \text{ is odd} \\ 2^{-n} & \text{if } \ n \ \text{ is even} \end{cases}.$$

The even terms are positive and tend to 0; the odd terms are close to -1 and tend to -1.

For every k, the largest value in the tail occurs at the first even index at least k, so $x_k \to 0$. Hence

$$\limsup_{n\to\infty} a_n = 0.$$

For instance,

$$x_2 = \sup\{2^{-2}, -1 + 3^{-3}, 2^{-4}, \ldots\} = \frac{1}{4}$$

and

$$x_8 = \sup\{2^{-8}, -1 + 3^{-9}, 2^{-10}, \ldots\} = \frac{1}{256}.$$

The smallest values in the tails come from the odd terms and approach -1, so

$$\liminf_{n\to\infty} a_n = -1.$$

Example 5
Some useful quick calculations are:

- If $a_n = \frac{(-1)^n}{n}$, then both even and odd tails tend to 0, so $\limsup a_n = \liminf a_n = 0$.
- If $a_n = (-1)^n \frac{n^4+5}{n^4}$, then the even terms tend to 1 and the odd terms tend to -1, so $\limsup a_n = 1$ and $\liminf a_n = -1$.
- If $a_n = \frac{8-n^3}{n-2}$ for $n \geq 3$, then $a_n \to -\infty$, so the extended upper and lower limits are both $-\infty$.

Example 6
For $a_n = (-1)^n$, every tail contains infinitely many 1s and infinitely many -1s. Therefore

$$\limsup_{n\to\infty} a_n = 1, \quad \liminf_{n\to\infty} a_n = -1.$$

Since these are different, the sequence does not converge.

Proposition 2 (Subadditivity of limsup)
If (a_n) and (b_n) are bounded real sequences, then

$$\limsup_{n\to\infty}(a_n+b_n) \le \limsup_{n\to\infty} a_n + \limsup_{n\to\infty} b_n.$$

Proof. For each k, set

$$A_k = \sup\{a_n : n \ge k\}, \quad B_k = \sup\{b_n : n \ge k\}, \quad C_k = \sup\{a_n + b_n : n \ge k\}.$$

If $n \ge k$, then $a_n \le A_k$ and $b_n \le B_k$, so

$$a_n + b_n \le A_k + B_k.$$

Taking the supremum over $n \ge k$ gives $C_k \le A_k + B_k$. Now take limits as $k \to \infty$. ☐

Theorem 3 (Convergence via limsup and liminf)
A bounded real sequence (a_n) converges to L if and only if

$$\limsup_{n\to\infty} a_n = \liminf_{n\to\infty} a_n = L.$$

Proof. Suppose $a_n \to L$. Let $\varepsilon > 0$. For all sufficiently large n,

$$L - \varepsilon < a_n < L + \varepsilon.$$

Hence, for all sufficiently large k,

$$L - \varepsilon \le y_k \le x_k \le L + \varepsilon.$$

Taking limits gives $\liminf a_n = \limsup a_n = L$.

Conversely, suppose $\limsup a_n = \liminf a_n = L$. Let $\varepsilon > 0$. For all sufficiently large k,

$$L - \varepsilon < y_k \le a_k \le x_k < L + \varepsilon.$$

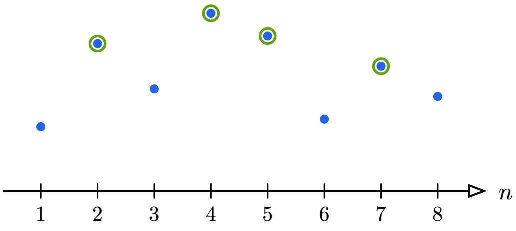
Thus $|a_k - L| < \varepsilon$ for all sufficiently large k, so $a_n \to L$. ☐

3.3 Subsequences

Definition 4 (Subsequence)
Let (a_n) be a sequence. A **subsequence** of (a_n) is a sequence of the form $\left(a_{n_k}\right)$, where

$$n_1 < n_2 < n_3 < \ldots$$

are natural numbers.



A subsequence is formed by choosing terms whose indices continue to increase. Here the selected terms are $a_2, a_4, a_5, a_7, \ldots$.

Theorem 4 (Subsequences of convergent sequences)
If $a_n \to L$, then every subsequence of (a_n) also converges to L.

Proof. Let $\left(a_{n_k}\right)$ be a subsequence and let $\varepsilon > 0$. Since $a_n \to L$, there exists $N \in \mathbb{N}$ such that $|a_n - L| < \varepsilon$ whenever $n \geq N$.

Since $n_k \geq k$, every $k \geq N$ satisfies $n_k \geq N$. Therefore

$$\left|a_{n_k} - L\right| < \varepsilon$$

whenever $k \geq N$, so $a_{n_k} \to L$. □

Example 7
The sequence $a_n = (-1)^n$ has two constant subsequences:

$$a_2, a_4, a_6, \ldots = 1, 1, 1, \ldots$$

and

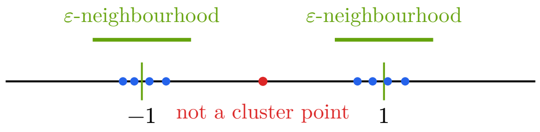
$$a_1, a_3, a_5, \ldots = -1, -1, -1, \ldots.$$

These subsequences converge to different limits, so the original sequence cannot converge by the subsequences of convergent sequences theorem.

3.4 Cluster Points

Definition 5 (Cluster point)
A number $a \in \mathbb{R}$ is a **cluster point** of a sequence (a_n) if for every $\varepsilon > 0$, there are infinitely many indices $n \in \mathbb{N}$ such that

$$|a_n - a| < \varepsilon.$$



A cluster point has infinitely many sequence terms in every epsilon-neighbourhood. A sequence may have more than one cluster point.

Theorem 5 (Cluster points and subsequences)
Let (a_n) be a real sequence and let $a \in \mathbb{R}$.

1. a is a cluster point if and only if for every $\varepsilon > 0$ and every $N \in \mathbb{N}$, there exists $n \geq N$ such that $|a_n - a| < \varepsilon$.
2. a is a cluster point if and only if (a_n) has a subsequence converging to a.

Proof. For (1), if a is a cluster point, every epsilon-neighbourhood contains infinitely many terms, so at least one must occur after any prescribed index N. Conversely, if the stated condition holds, apply it repeatedly with larger and larger N to find infinitely many terms in the epsilon-neighbourhood.

For (2), suppose a is a cluster point. Use (1) with $\varepsilon = 1$ and $N = 1$ to choose n_1. Then use $\varepsilon = \frac{1}{2}$ and $N = n_1 + 1$ to choose $n_2 > n_1$. Continuing in this way gives

$$n_1 < n_2 < n_3 < \ldots$$

and $\left|a_{n_k} - a\right| < \frac{1}{k}$, so $a_{n_k} \to a$.

Conversely, if a subsequence $a_{n_k} \to a$, then every epsilon-neighbourhood of a contains all sufficiently late terms of that subsequence, and hence infinitely many terms of the original sequence. ☐

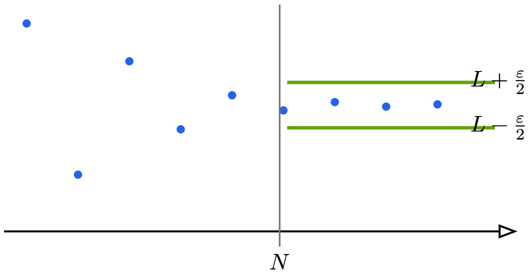
Remark
For a bounded sequence, $\limsup a_n$ and $\liminf a_n$ are respectively the largest and smallest cluster points.

3.5 Cauchy Sequences

Definition 6 (Cauchy sequence)
A sequence (a_n) is **Cauchy** if for every $\varepsilon > 0$ there exists $N \in \mathbb{N}$ such that whenever $m, n \geq N$,

$$|a_n - a_m| < \varepsilon.$$

A convergent sequence is close to its limit. A Cauchy sequence is close to itself: far enough along the sequence, all remaining terms are close to one another.



In a Cauchy sequence, the tail of the sequence eventually fits inside an interval of length ε.

Theorem 6 (Convergent sequences are Cauchy)
Every convergent real sequence is Cauchy.

Proof. Suppose $a_n \to L$ and let $\varepsilon > 0$. Choose $N \in \mathbb{N}$ such that

$$|a_n - L| < \frac{\varepsilon}{2}$$

whenever $n \geq N$. If $m, n \geq N$, then

$$|a_n - a_m| = |a_n - L + L - a_m| \leq |a_n - L| + |a_m - L| < \varepsilon.$$

Therefore (a_n) is Cauchy. □

Lemma 7 (Cauchy sequences are bounded)
Every Cauchy real sequence is bounded.

Proof. Apply the Cauchy definition with $\varepsilon = 1$. There exists $N \in \mathbb{N}$ such that $|a_n - a_N| < 1$ whenever $n \geq N$. Hence, for $n \geq N$,

$$|a_n| \leq |a_n - a_N| + |a_N| < 1 + |a_N|.$$

The finitely many earlier terms are bounded, so the whole sequence is bounded. □

Proposition 8 (Cauchy permanence properties)
Let (a_n) and (b_n) be real sequences.

1. Every subsequence of a Cauchy sequence is Cauchy.
2. If (a_n) and (b_n) are Cauchy, then $(|a_n - b_n|)$ is Cauchy.

Proof. For (1), let $\left(a_{n_k}\right)$ be a subsequence of a Cauchy sequence (a_n). Given $\varepsilon > 0$, choose N such that $|a_n - a_m| < \varepsilon$ whenever $m, n \geq N$. If $j, k \geq N$, then $n_j \geq j \geq N$ and $n_k \geq k \geq N$, so

$$\left|a_{n_j} - a_{n_k}\right| < \varepsilon.$$

For (2), use the reverse triangle inequality. Given $\varepsilon > 0$, choose N_1 such that $|a_n - a_m| < \frac{\varepsilon}{2}$ whenever $m, n \geq N_1$, and choose N_2 such that $|b_n - b_m| < \frac{\varepsilon}{2}$ whenever $m, n \geq N_2$. If $m, n \geq \max(N_1, N_2)$, then

$$||a_n - b_n| - |a_m - b_m|| \leq |(a_n - b_n) - (a_m - b_m)| \leq |a_n - a_m| + |b_n - b_m| < \varepsilon.$$

□

Example 8
The following classifications are typical checks.

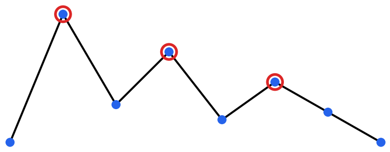
Sequence	Cauchy?	Monotone?	Bounded?
$\frac{1}{2n+1}$	yes, $\to 0$	strictly decreasing	yes
$\frac{n^2+1}{n} = n + \frac{1}{n}$	no	strictly increasing	below only
$\frac{n-2}{3n+4}$	yes, $\to \frac{1}{3}$	strictly increasing	yes
$\frac{2n+1}{n} = 2 + \frac{1}{n}$	yes, $\to 2$	strictly decreasing	yes
$(-1)^n n$	no	no	no

3.6 Monotone Subsequences and Bolzano-Weierstrass

Definition 7 (Peak point)
An index $k \in \mathbb{N}$ is a **peak point** of a sequence (a_n) if

$$a_k > a_n$$

for every $n > k$.



Peak points are terms larger than every later term.

Theorem 9 (Every sequence has a monotone subsequence)
Every real sequence has a monotone subsequence.

Proof. Let (a_n) be a real sequence.

If there are infinitely many peak points, list them as

$$n_1 < n_2 < n_3 < \dots.$$

Since each n_k is a peak point and $n_{k+1} > n_k$, we have

$$a_{n_1} > a_{n_2} > a_{n_3} > \dots,$$

so these terms form a decreasing subsequence.

If there are only finitely many peak points, choose n_1 after the last peak point. Since n_1 is not a peak point, there exists $n_2 > n_1$ such that $a_{n_2} \geq a_{n_1}$. Since n_2 is also not a peak point, there exists $n_3 > n_2$ such that $a_{n_3} \geq a_{n_2}$. Continuing gives an increasing subsequence. ☐

Theorem 10 (Bolzano-Weierstrass theorem)
Every bounded real sequence has a convergent subsequence.

Proof. By the monotone subsequence theorem, every sequence has a monotone subsequence. If the original sequence is bounded, then every subsequence is bounded. Hence the monotone subsequence is bounded, so it converges by the monotone convergence theorem. ☐

Theorem 11 (Cauchy criterion for real sequences)
A real sequence converges if and only if it is Cauchy.

Proof. The forward direction is the convergent sequences are Cauchy theorem.

Conversely, suppose (a_n) is Cauchy. By the Cauchy sequences are bounded lemma, (a_n) is bounded. By the Bolzano-Weierstrass theorem, it has a convergent subsequence $\left(a_{n_k}\right)$, say $a_{n_k} \to L$.

We prove $a_n \to L$. Let $\varepsilon > 0$. Since (a_n) is Cauchy, there exists N_1 such that

$$|a_n - a_m| < \frac{\varepsilon}{2}$$

whenever $m, n \geq N_1$. Since $a_{n_k} \to L$, there exists K such that

$$\left|a_{n_k} - L\right| < \frac{\varepsilon}{2}$$

whenever $k \geq K$.

Choose $k \geq K$ such that $n_k \geq N_1$. Then for every $n \geq N_1$,

$$|a_n - L| \leq \left|a_n - a_{n_k}\right| + \left|a_{n_k} - L\right| < \varepsilon.$$

Hence $a_n \to L$. □

Example 9
The sequence $a_n = (-1)^n$ is bounded but not Cauchy. Taking $\varepsilon = 1$, no matter how far along the sequence we go, there are later even and odd indices m, n with

$$|a_n - a_m| = 2.$$

Thus boundedness alone is not enough for the Cauchy property.

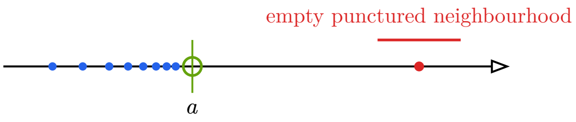
4 Function Limits and Continuity

The sequence chapters studied limits indexed by $n \in \mathbb{N}$. We now study limits indexed by real variables. The main change is that x can approach a point from many nearby values, and the value of the function at the point itself may be irrelevant to the limit.

4.1 Limit Points

Definition 1 (Limit point)
Let $X \subset \mathbb{R}$. A point $a \in \mathbb{R}$ is a **limit point** of X if every punctured neighbourhood of a contains a point of X. Equivalently, for every $\varepsilon > 0$, there exists $x \in X$ such that

$$0 < |x - a| < \varepsilon.$$



A limit point is approached by points of the set. The isolated point on the right is not a limit point.

Example 1
The set $(0, 1]$ has limit points $[0, 1]$. The set $\mathbb{N}$ has no limit points in $\mathbb{R}$. The set $\left\{1, \frac{1}{2}, \frac{1}{3}, \ldots\right\}$ has exactly one limit point, namely 0.

4.2 Limits of Functions

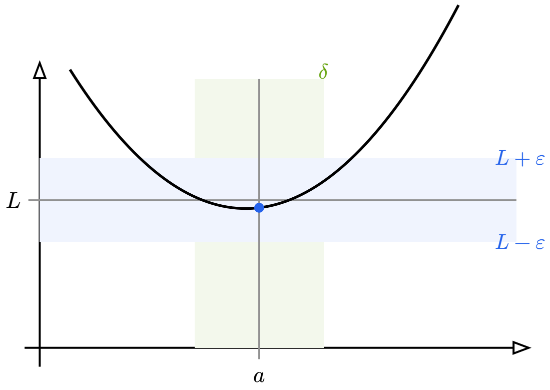
Definition 2 (Limit of a function)
Let $f : X \to \mathbb{R}$ and let a be a limit point of X. We write

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

if for every $\varepsilon > 0$, there exists $\delta > 0$ such that for every $x \in X$,

$$0 < |x - a| < \delta \implies |f(x) - L| < \varepsilon.$$

The condition $0 < |x - a|$ means that the value $f(a)$ does not matter for the limit. A function may have a limit at a even if it is undefined at a, or defined there with the wrong value.



If x lies inside the punctured δ-neighbourhood of a, then $f(x)$ must lie inside the ε-neighbourhood of L.

Example 2
We prove $\lim_{x\to a} x = a$. Let $\varepsilon > 0$ and choose $\delta = \varepsilon$. If $0 < |x - a| < \delta$, then

$$|x - a| < \delta = \varepsilon.$$

Example 3
We prove

$$\lim_{x\to 1}(x^2 + 7x + 12) = 20.$$

Let $\varepsilon > 0$. We need to control

$$\left|x^2 + 7x + 12 - 20\right| = |x - 1||x + 8|.$$

If $|x - 1| < 1$, then $0 < x < 2$, so $|x + 8| < 10$. Choose

$$\delta = \min\left(1, \frac{\varepsilon}{10}\right).$$

Then $0 < |x - 1| < \delta$ implies

$$\left|x^2 + 7x + 12 - 20\right| < 10|x - 1| < \varepsilon.$$

Example 4
We prove

$$\lim_{x\to 2} \frac{x^3 - 8}{x - 2} = 12.$$

For $x \neq 2$,

$$\frac{x^3 - 8}{x - 2} = x^2 + 2x + 4.$$

Hence

$$\left|\frac{x^3 - 8}{x - 2} - 12\right| = \left|x^2 + 2x - 8\right| = |x - 2||x + 4|.$$

If $|x - 2| < 1$, then $1 < x < 3$, so $|x + 4| < 7$. Taking

$$\delta = \min\left(1, \frac{\varepsilon}{7}\right)$$

proves the limit.

Example 5
We prove

$$\lim_{x\to 3}\frac{2}{x+3}=\frac{1}{3}.$$

For x near 3,

$$\left|\frac{2}{x+3}-\frac{1}{3}\right|=\left|\frac{6-x-3}{3(x+3)}\right|=\frac{|x-3|}{3|x+3|}.$$

If $|x-3|<1$, then $x>2$, so $|x+3|>5$. Hence

$$\left|\frac{2}{x+3}-\frac{1}{3}\right|<\frac{|x-3|}{15}.$$

Taking $\delta=\min(1,15\varepsilon)$ proves the limit.

Theorem 1 (Algebra of function limits)
Suppose

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

Then

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

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

and, if $M\neq 0$,

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

Proof. The proof follows the same estimates as the sequence limit laws. For products, add zero:

$$f(x)g(x)-LM=f(x)(g(x)-M)+M(f(x)-L).$$

Since $f(x)\to L$, the values of $f(x)$ are bounded for x close enough to a. The remaining terms are controlled by the two limits. The quotient law follows by first showing

$$\frac{1}{g(x)}\to\frac{1}{M}.$$

□

Theorem 2 (Limit of a composite function)
Suppose

$$\lim_{x\to a}g(x)=b \quad \text{and} \quad \lim_{y\to b}f(y)=L.$$

If $g(x)\neq b$ for all x sufficiently close to a with $x\neq a$, then

$$\lim_{x\to a}f(g(x))=L.$$

Proof. Let $\varepsilon > 0$. Since $f(y) \to L$ as $y \to b$, choose $\eta > 0$ such that

$$0 < |y - b| < \eta \quad \Longrightarrow \quad |f(y) - L| < \varepsilon.$$

Since $g(x) \to b$ as $x \to a$, choose $\delta > 0$ so that

$$0 < |x - a| < \delta \quad \Longrightarrow \quad |g(x) - b| < \eta.$$

Shrink δ if needed so that $g(x) \neq b$ in the same punctured neighbourhood. Then

$$0 < |g(x) - b| < \eta,$$

so $|f(g(x)) - L| < \varepsilon$. □

Corollary 3 (Continuous outer functions preserve limits)
If $g(x) \to b$ as $x \to a$ and f is continuous at b, then

$$f(g(x)) \to f(b).$$

Example 6
Algebraic simplification gives

$$\lim_{x\to 2} \frac{x^2 - x - 2}{x^2 - 4} = \lim_{x\to 2} \frac{(x-2)(x+1)}{(x-2)(x+2)} = \frac{3}{4}.$$

The cancellation is valid because limits ignore the point $x = 2$ itself.

Example 7
Some quick limits from the same rules are

$$\lim_{x\to 3} \frac{2}{x+3} = \frac{1}{3} \quad \text{and} \quad \lim_{x\to 0} \left(-\frac{x^4}{4}\right) \cos\left(\frac{1}{x^3+2}\right) = 0.$$

The first uses continuity of $\frac{2}{x+3}$ at 3. For the second, use $|\cos(t)| \le 1$:

$$\left| \left(-\frac{x^4}{4}\right) \cos\left(\frac{1}{x^3+2}\right) \right| \le \frac{x^4}{4} \to 0.$$

Example 8
Another removable singularity is

$$\lim_{x\to 1} \frac{x^2 + x - 2}{x^2 - 4x + 3} = \lim_{x\to 1} \frac{(x+2)(x-1)}{(x-1)(x-3)} = -\frac{3}{2}.$$

Example 9
The limit

$$\lim_{x\to 2} \frac{x^2 - 4}{|x - 2|}$$

does not exist. If $x > 2$, then the expression is $x + 2 \to 4$. If $x < 2$, then it is $-(x + 2) \to -4$. The one-sided limits are different.

4.3 One-Sided and Infinite Limits

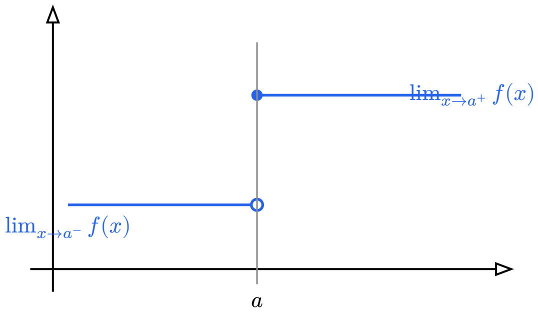
Definition 3 (One-sided limits)
We write $\lim_{x\to a^+} f(x) = L$ if for every $\varepsilon > 0$, there exists $\delta > 0$ such that

$$0 < x - a < \delta \implies |f(x) - L| < \varepsilon.$$

Similarly, $\lim_{x\to a^-} f(x) = L$ means the same statement with $0 < a - x < \delta$.

Proposition 4 (Two-sided limits and one-sided limits)
The two-sided limit $\lim_{x\to a} f(x) = L$ exists if and only if both one-sided limits exist and

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



A jump discontinuity has different left and right limits.

Definition 4 (Limits involving infinity)
We write $\lim_{x\to\infty} f(x) = L$ if for every $\varepsilon > 0$, there exists $M > 0$ such that

$$x > M \implies |f(x) - L| < \varepsilon.$$

We write $\lim_{x\to a} f(x) = \infty$ if for every $M > 0$, there exists $\delta > 0$ such that

$$0 < |x - a| < \delta \implies f(x) > M.$$

The definitions of $x \to -\infty$ and $f(x) \to -\infty$ are analogous.

Example 10
Divide by the highest power of x:

$$\lim_{x\to\infty} \frac{x^2 + x}{x^2 + 4x + 7} = \lim_{x\to\infty} \frac{1 + \frac{1}{x}}{1 + \frac{4}{x} + \frac{7}{x^2}} = 1.$$

More generally, if the numerator and denominator have the same degree, the limit at infinity is the ratio of leading coefficients.

4.4 Squeeze and Sequential Criteria

Theorem 5 (Squeeze theorem for functions)
Suppose $f(x) \leq h(x) \leq g(x)$ for all x sufficiently close to a, except possibly at a. If

$$\lim_{x\to a} f(x) = \lim_{x\to a} g(x) = L,$$

then

$$\lim_{x\to a} h(x) = L.$$

The same result holds for one-sided limits and limits at infinity.

Example 11
Since $-1 \le \sin\left(\frac{1}{x}\right) \le 1$ for $x \ne 0$,

$$-x^2 \le x^2 \sin\left(\frac{1}{x}\right) \le x^2.$$

Both outside functions tend to 0 as $x \to 0$, so $x^2 \sin\left(\frac{1}{x}\right) \to 0$ by the squeeze theorem for functions.

Example 12
At $x = 2$,

$$|2x-1| = 2x-1 \quad \text{and} \quad |2x+1| = 2x+1.$$

Hence

$$\lim_{x\to 2} \frac{|2x-1| - |2x+1|}{x} = \lim_{x\to 2} \frac{-2}{x} = -1.$$

Theorem 6 (Sequential criterion for function limits)
Let a be a limit point of X. Then $\lim_{x\to a} f(x) = L$ if and only if for every sequence (x_n) in X with $x_n \ne a$ and $x_n \to a$, we have

$$f(x_n) \to L.$$

Proof. If $f(x) \to L$ and $x_n \to a$, then the epsilon-delta condition for f eventually applies to x_n, so $f(x_n) \to L$.

Conversely, suppose the sequential condition holds but $f(x)$ does not tend to L. Then there exists $\varepsilon_0 > 0$ such that for every $\delta > 0$, some $x \in X$ satisfies

$$0 < |x-a| < \delta \quad \text{and} \quad |f(x) - L| \ge \varepsilon_0.$$

Taking $\delta = \frac{1}{n}$ gives a sequence (x_n) with $x_n \to a$ but $f(x_n)$ not tending to L, a contradiction.
☐

Note
The standard limits

$$\lim_{x\to 0} \frac{\sin(x)}{x} = 1, \quad \lim_{x\to 0} \frac{1-\cos(x)}{x} = 0, \quad \lim_{x\to 0} \frac{1-\cos(x)}{x^2} = \frac{1}{2}$$

may be used in this course.

Example 13
Using $\frac{\sin(x)}{x} \to 1$ and $\cos(x) \to 1$,

$$\lim_{x\to 0} \frac{\sin(x)\cos(x)}{x} = \lim_{x\to 0} \left(\frac{\sin(x)}{x}\right) \cos(x) = 1.$$

4.5 Continuity

Definition 5 (Continuity at a point)
Let $f : X \to \mathbb{R}$ and let $a \in X$. The function f is **continuous at** a if for every $\varepsilon > 0$, there exists $\delta > 0$ such that for every $x \in X$,

$$|x - a| < \delta \quad\Longrightarrow\quad |f(x) - f(a)| < \varepsilon.$$

If a is also a limit point of X, this is equivalent to

$$\lim_{x\to a} f(x) = f(a).$$

Definition 6 (Continuity on a set)
A function $f : X \to \mathbb{R}$ is **continuous on** X if it is continuous at every point of X.

Theorem 7 (Algebra of continuous functions)
If f and g are continuous at a, then $f + g$ and fg are continuous at a. If $g(a) \neq 0$, then the quotient function is continuous at a.

Proof. This is the algebra of function limits with $L = f(a)$ and $M = g(a)$. ☐

Corollary 8 (Polynomials and rational functions)
Every polynomial is continuous on $\mathbb{R}$. A rational function $\frac{p(x)}{q(x)}$ is continuous at every point where $q(x) \neq 0$.

Example 14
To make

$$f(x) = \begin{cases} x^2 + 3x + 5 & \text{if } \ x > 1 \\ a & \text{if } \ x = 1 \\ 12x - 3 & \text{if } \ x < 1 \end{cases}$$

continuous at 1, the three values must agree. The left and right limits are both 9, so we need $a = 9$.

Example 15
The function

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

is not continuous at 0. Since $f(0) = 0$, take $\varepsilon = \frac{1}{2}$. For every $\delta > 0$, choose $x = \frac{\delta}{2} > 0$. Then $|x - 0| < \delta$, but

$$|f(x) - f(0)| = 1 \geq \frac{1}{2}.$$

Example 16
For

$$f(x) = \begin{cases} x^2 - 3 & \text{if } x > 0 \\ 0 & \text{if } x = 0 \\ \sin(x) & \text{if } x < 0 \end{cases}$$

the limit at 0 does not exist. The right-hand limit is -3, while the left-hand limit is 0.

Example 17
Let

$$f(x) = \begin{cases} 2x + |2x - 3| & \text{if } x \neq 2 \\ a & \text{if } x = 2. \end{cases}$$

To make f continuous at 2, choose

$$a = \lim_{x \to 2} (2x + |2x - 3|) = 4 + 1 = 5.$$

Proposition 9 (Continuity of absolute value)
If f is continuous at a, then $|f|$ is continuous at a.

Proof. By the reverse triangle inequality,

$$||f(x)| - |f(a)|| \leq |f(x) - f(a)|.$$

The right-hand side tends to 0 as $x \to a$, so $|f(x)| \to |f(a)|$. ☐

Example 18
The converse is false. The function

$$f(x) = \begin{cases} 1 & \text{if } x \in \mathbb{Q} \\ -1 & \text{if } x \text{ is irrational} \end{cases}$$

is continuous nowhere because every interval contains both rational and irrational points, but $|f(x)| = 1$ is continuous everywhere.

Proposition 10 (Squeeze continuity at zero)
Suppose $|f(x)| \leq |g(x)|$ for all x near 0, and g is continuous at 0 with $g(0) = 0$. If $f(0) = 0$, then f is continuous at 0.

Proof. Since g is continuous at 0, for every $\varepsilon > 0$ there exists $\delta > 0$ such that $|x| < \delta$ implies $|g(x)| < \varepsilon$. Hence

$$|f(x) - f(0)| = |f(x)| \leq |g(x)| < \varepsilon.$$

☐

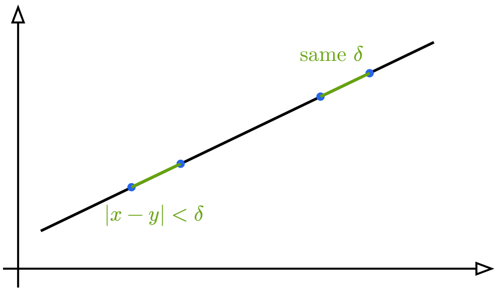
The special case $|f(x)| \leq |x|$ follows by taking $g(x) = x$.

4.6 Uniform Continuity

Definition 7 (Uniform continuity)
Let $I \subset \mathbb{R}$. A function $f : I \to \mathbb{R}$ is **uniformly continuous on** I if for every $\varepsilon > 0$, there exists $\delta > 0$ such that for all $x, y \in I$,

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

Continuity at a point allows δ to depend on the point. Uniform continuity requires one δ to work across the whole set.



Uniform continuity asks for one horizontal tolerance that works everywhere on the interval.

Proposition 11 (Uniform continuity implies continuity)
If f is uniformly continuous on I, then f is continuous on I.

Proof. Fix $a \in I$. In the definition of uniform continuity, take $y = a$. The same δ proves continuity at a. □

Example 19
The function $f(x) = x$ is uniformly continuous on $\mathbb{R}$: choose $\delta = \varepsilon$.

Example 20
The function $f(x) = 3x + 7$ is uniformly continuous on $\mathbb{R}$. Given $\varepsilon > 0$, choose $\delta = \frac{\varepsilon}{3}$. If $|x - y| < \delta$, then

$$|f(x) - f(y)| = 3|x - y| < \varepsilon.$$

Example 21
The function $f(x) = \frac{1}{x}$ is uniformly continuous on $\left[\frac{1}{3}, \infty\right)$. If $x, y \geq \frac{1}{3}$, then $|xy| \geq \frac{1}{9}$, so

$$\left|\frac{1}{x} - \frac{1}{y}\right| = \frac{|x - y|}{|xy|} \leq 9|x - y|.$$

Taking $\delta = \frac{\varepsilon}{9}$ proves uniform continuity.

Example 22
The function $f(x) = x^2$ is uniformly continuous on $[0, 3]$. If $x, y \in [0, 3]$, then

$$\left|x^2 - y^2\right| = |x - y||x + y| \leq 6|x - y|.$$

Taking $\delta = \frac{\varepsilon}{6}$ proves uniform continuity.

Example 23
The function $f(x) = x^2$ is not uniformly continuous on $(0, \infty)$. Suppose it were. Take $\varepsilon = 1$ and let $\delta > 0$ be supplied by uniform continuity. Choose $x > \frac{1}{\delta}$ and set $y = x + \frac{\delta}{2}$. Then $|x - y| < \delta$, but

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

a contradiction.

Theorem 12 (Continuous functions on closed intervals are uniformly continuous)
If $f : [a, b] \to \mathbb{R}$ is continuous, then f is uniformly continuous on $[a, b]$.

Proof. Suppose not. Then there exists $\varepsilon_0 > 0$ such that for every $\delta > 0$, there are $x, y \in [a, b]$ with

$$|x - y| < \delta \quad \text{but} \quad |f(x) - f(y)| \geq \varepsilon_0.$$

Taking $\delta = \frac{1}{n}$ gives sequences (x_n) and (y_n) in $[a, b]$ such that

$$|x_n - y_n| < \frac{1}{n} \quad \text{and} \quad |f(x_n) - f(y_n)| \geq \varepsilon_0.$$

By the Bolzano-Weierstrass theorem, (x_n) has a convergent subsequence $x_{n_k} \to c \in [a, b]$. Since

$$\left|y_{n_k} - c\right| \leq \left|y_{n_k} - x_{n_k}\right| + \left|x_{n_k} - c\right|,$$

we also have $y_{n_k} \to c$.

Continuity gives $f\left(x_{n_k}\right) \to f(c)$ and $f\left(y_{n_k}\right) \to f(c)$, so

$$\left|f\left(x_{n_k}\right) - f\left(y_{n_k}\right)\right| \to 0,$$

contradicting $\left|f\left(x_{n_k}\right) - f\left(y_{n_k}\right)\right| \geq \varepsilon_0$ for every k. □

Proposition 13 (Continuous functions preserve Cauchy sequences on closed intervals)
Let $f : [a, b] \to \mathbb{R}$ be continuous. If (a_n) is a Cauchy sequence in $[a, b]$, then $(f(a_n))$ is Cauchy.

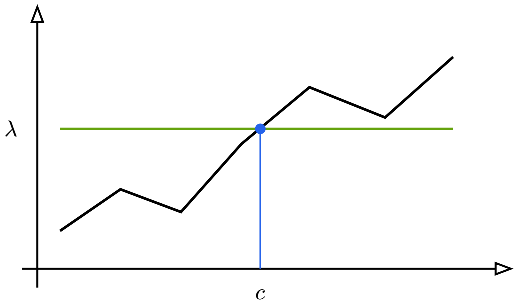
Proof. By the Cauchy criterion, (a_n) converges to some $L \in \mathbb{R}$. Since every $a_n \in [a, b]$ and $[a, b]$ is closed, $L \in [a, b]$. Continuity gives $f(a_n) \to f(L)$, so $(f(a_n))$ is Cauchy by the convergent sequences are Cauchy theorem. □

4.7 Intermediate and Extreme Values

Theorem 14 (Intermediate value theorem)
Let $f : [a, b] \to \mathbb{R}$ be continuous. If λ lies between $f(a)$ and $f(b)$, then there exists $c \in [a, b]$ such that

$$f(c) = \lambda.$$

If λ lies strictly between $f(a)$ and $f(b)$, then c may be chosen in (a, b).



A continuous function must hit every height between its endpoint values.

Example 24
Every real cubic

$$p(x) = x^3 + ax^2 + bx + c$$

has a real root. Indeed, the leading term dominates: $p(x) \to \infty$ as $x \to \infty$ and $p(x) \to -\infty$ as $x \to -\infty$. Choose R large enough that $p(-R) < 0 < p(R)$. Since p is continuous on $[-R, R]$, the intermediate value theorem gives some $r \in (-R, R)$ with $p(r) = 0$.

Example 25
There is a real number x such that

$$x^5 + 34x + 13 = 100.$$

Let $f(x) = x^5 + 34x - 87$. Then f is continuous,

$$f(1) = -52 < 0 \quad \text{and} \quad f(2) = 13 > 0.$$

Hence the intermediate value theorem gives a root in $(1, 2)$.

Example 26
There is a real number x such that $\sin x = x - 1$. Let

$$f(x) = \sin x - x + 1.$$

Then $f(0) = 1 > 0$ and $f(2) = \sin 2 - 1 < 0$. Since f is continuous, the intermediate value theorem gives a root in $(0, 2)$.

Definition 8 (Extrema)
Let $f : X \to \mathbb{R}$. The function f attains a **maximum** at $c \in X$ if

$$f(c) \geq f(x)$$

for all $x \in X$. It attains a **minimum** at $c \in X$ if

$$f(c) \leq f(x)$$

for all $x \in X$. Maxima and minima are called **extrema**.

Theorem 15 (Extreme value theorem)
If $f : [a, b] \to \mathbb{R}$ is continuous, then f attains a maximum and a minimum on $[a, b]$.

Note
Closed and bounded matters. The function $f(x) = x$ on $(0, 1)$ has no maximum or minimum. The function $f(x) = x^2$ on $[1, \infty)$ has a minimum but no maximum.

5 Differentiation

Differentiation measures the best linear approximation to a function at a point. The main topics here are derivative rules, the link between differentiability and continuity, and the mean value theorem.

5.1 Derivatives

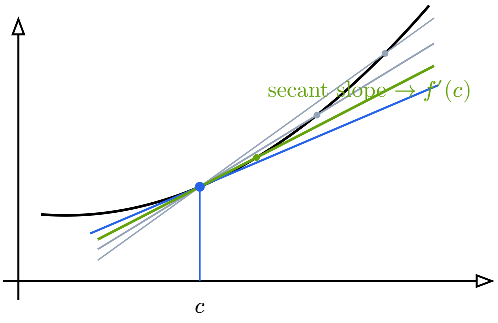
Definition 1 (Derivative at a point)
Let $f : I \to \mathbb{R}$, where I is an interval, and let c be an interior point of I. We say that f is **differentiable at** c if the limit

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

exists. Equivalently,

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

The number $f'(c)$ is the **derivative** of f at c.



The derivative is the limiting slope of secant lines through nearby points.

Example 1
For $f(x) = x^2$, we have

$$f'(c) = \lim_{h \to 0} \frac{(c+h)^2 - c^2}{h} = \lim_{h \to 0} \frac{2ch + h^2}{h} = 2c.$$

Example 2
The function $f(x) = |x|$ is not differentiable at 0. The right difference quotient is

$$\lim_{h \to 0^+} \frac{|h|}{h} = 1,$$

while the left difference quotient is

$$\lim_{h \to 0^-} \frac{|h|}{h} = -1.$$

Example 3
Let

$$f(x) = \begin{cases} (x+a)^2 + 1 & \text{if } \ x \geq 0 \\ 2x + 5 & \text{if } \ x < 0. \end{cases}$$

Continuity at 0 requires

$$a^2 + 1 = 5,$$

so $a = 2$ or $a = -2$. For differentiability at 0, compare one-sided derivatives. The left derivative is 2, while the right derivative is $2a$. Thus differentiability would require $a = 1$. Therefore neither continuous value of a makes the function differentiable at 0.

Proposition 1 (Differentiability implies continuity)
If f is differentiable at c, then f is continuous at c.

Proof. For $x \neq c$,

$$f(x) - f(c) = \left(\frac{f(x) - f(c)}{x - c} \right) (x - c).$$

As $x \to c$, the first factor tends to $f'(c)$ and the second factor tends to 0. Hence $f(x) - f(c) \to 0$, so $f(x) \to f(c)$. □

Note
The converse is false. For example, $f(x) = |x|$ is continuous at 0 but not differentiable there.

5.2 Derivative Laws

Theorem 2 (Derivative rules)
Suppose f and g are differentiable at c, and let $\lambda \in \mathbb{R}$.

$$(f+g)'(c) = f'(c) + g'(c),$$

$$(\lambda f)'(c) = \lambda f'(c),$$

$$(fg)'(c) = f'(c)g(c) + f(c)g'(c).$$

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. The sum and scalar rules follow directly from the limit definition. For products, add zero:

$$f(x)g(x) - f(c)g(c) = f(x)(g(x) - g(c)) + g(c)(f(x) - f(c)).$$

Divide by $x - c$ and take $x \to c$. Differentiability gives continuity, so $f(x) \to f(c)$. The quotient rule follows by applying the product rule to $f\left(\frac{1}{g}\right)$. □

Theorem 3 (Chain rule)
Let g be differentiable at c, and let f be differentiable at $g(c)$. Then $f \circ g$ is differentiable at c, and

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

Proof. Put $b = g(c)$. Define

$$q(y) = \begin{cases} \frac{f(y)-f(b)}{y-b} & \text{if} \quad y \neq b \\ f'(b) & \text{if} \quad y = b. \end{cases}$$

Since f is differentiable at b, we have $q(y) \to f'(b)$ as $y \to b$, so q is continuous at b. For x near c,

$$f(g(x)) - f(g(c)) = q(g(x))(g(x) - g(c)).$$

Therefore

$$\frac{f(g(x)) - f(g(c))}{x - c} = q(g(x))\frac{g(x) - g(c)}{x - c}.$$

Taking $x \to c$ gives $q(g(x)) \to q(b) = f'(b)$ and $\frac{g(x)-g(c)}{x-c} \to g'(c)$. □

Example 4
The standard derivative rules give:

$$\left(\frac{\mathrm{d}}{\mathrm{d}}x\right)(\sin x \cos x) = \cos^2 x - \sin^2 x,$$

$$\left(\frac{\mathrm{d}}{\mathrm{d}}x\right) \log(x^2 + 1) = \frac{2x}{x^2 + 1},$$

$$\left(\frac{\mathrm{d}}{\mathrm{d}}x\right)(e^x(x^5 + x^6 + 7)) = e^x(x^5 + x^6 + 7) + e^x(5x^4 + 6x^5).$$

Example 5
More chain-rule examples are

$$\left(\frac{\mathrm{d}}{\mathrm{d}}x\right)\sqrt{x^3 + 3x} = \frac{3x^2 + 3}{2\sqrt{x^3 + 3x}},$$

and

$$\left(\frac{\mathrm{d}}{\mathrm{d}}x\right) \arccos(x^2) = -2\frac{x}{\sqrt{1 - x^4}}.$$

5.3 Inverse Functions

Theorem 4 (Inverse derivative formula)
Let f be strictly monotone and continuous on an interval I, and suppose f is differentiable at $c \in I$ with $f'(c) \neq 0$. If f^{-1} is the inverse function on $f(I)$, then f^{-1} is differentiable at $f(c)$ and

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

Example 6
Since $\tan'(x) = \sec^2 x = 1 + \tan^2 x$ on $(-\frac{\pi}{2}, \frac{\pi}{2})$, the inverse derivative formula gives

$$\left(\frac{\mathrm{d}}{\mathrm{d}}x\right) \arctan(x) = \frac{1}{1+x^2}.$$

Therefore

$$\left(\frac{\mathrm{d}}{\mathrm{d}}x\right)(2x \arctan(x^3)) = 2\arctan(x^3) + 2x\left(3\frac{x^2}{1+x^6}\right).$$

5.4 Higher Derivatives and L'Hopital's Rule

Definition 2 (Higher derivatives)
If f' is differentiable, its derivative is the **second derivative** f''. Continuing recursively gives f''', $f^{(4)}$, and so on.

Theorem 5 (L'Hopital's rule)
Let f and g be differentiable near a, except possibly at a, with $g'(x) \neq 0$ near a. Suppose either

$$f(x) \to 0 \quad \text{and} \quad g(x) \to 0$$

or

$$|f(x)| \to \infty \quad \text{and} \quad |g(x)| \to \infty$$

as $x \to a$. If

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

then

$$\lim_{x \to a} \frac{f(x)}{g(x)} = L.$$

One-sided and infinite endpoint versions are used in the same way.

Example 7

$$\lim_{x \to 0} \frac{e^{2x}-1}{e^x - 1} = \lim_{x \to 0} \frac{2e^{2x}}{e^x} = 2.$$

Example 8

$$\lim_{x \to 0^+} x \log x = \lim_{x \to 0^+} \frac{\log x}{\frac{1}{x}}.$$

This is an $-\frac{\infty}{\infty}$ form, so by L'Hopital's rule,

$$\lim_{x\to 0^+}\frac{\log x}{\frac{1}{x}} = \lim_{x\to 0^+}\frac{\frac{1}{x}}{-\frac{1}{x^2}} = \lim_{x\to 0^+} -x = 0.$$

Example 9

$$\lim_{x\to 1}\frac{5\log x}{x-1} = \lim_{x\to 1}\frac{\frac{5}{x}}{1} = 5.$$

5.5 Local and Global Extrema

Definition 3 (Critical point)
Let f be defined on an interval I. An interior point $c \in I$ is a **critical point** of f if $f'(c) = 0$ or if $f'(c)$ does not exist.

Definition 4 (Local extremum)
A function f has a **local maximum** at c if there exists an open interval U containing c such that

$$f(c) \geq f(x)$$

for every $x \in U$ in the domain of f. Local minimum is defined similarly with $\leq$. Either is called a **local extremum**.

Theorem 6 (Fermat's theorem)
If f has a local extremum at an interior point c and f is differentiable at c, then

$$f'(c) = 0.$$

Proof. Suppose f has a local minimum at c. For x near c with $x > c$,

$$\frac{f(x) - f(c)}{x - c} \geq 0.$$

For x near c with $x < c$, the denominator is negative while $f(x) - f(c) \geq 0$, so

$$\frac{f(x) - f(c)}{x - c} \leq 0.$$

Since the two one-sided limits of the difference quotient are equal to $f'(c)$, we must have $f'(c) = 0$. The maximum case is analogous. □

Example 10
For

$$f(x) = x^4 - 4x^3 + 4x^2 + 3 = x^2(x-2)^2 + 3$$

on $[-1, 1]$,

$$f'(x) = 4x^3 - 12x^2 + 8x = 4x(x-1)(x-2).$$

The critical points in $[-1, 1]$ are 0 and 1. Checking endpoints and critical points:

$$f(-1) = 12, \quad f(0) = 3, \quad f(1) = 4.$$

Hence the global maximum is 12 at $x = -1$, and the global minimum is 3 at $x = 0$.

Example 11
For $f(x) = 1 - x^2$ on $(-2, 2]$, the only critical point is 0. We have $f(0) = 1$, so the global maximum is 1 at 0. There is no global minimum: as $x \to -2^+$, $f(x) \to -3$, but -2 is not in the interval.

Example 12
For $f(x) = e^x(x^3 - 2x)$ on $[0, 5]$,

$$f'(x) = e^x(x^3 + 3x^2 - 2x - 2).$$

The critical points are the roots of

$$x^3 + 3x^2 - 2x - 2 = 0$$

in $[0, 5]$. These generally need a numerical or graphing check. Once found, compare their function values with $f(0)$ and $f(5)$ to determine the global extrema.

5.6 Rolle's Theorem and the Mean Value Theorem

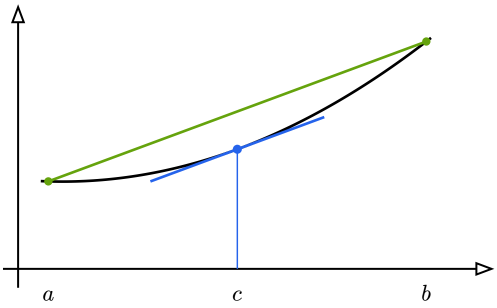
Theorem 7 (Rolle's theorem)
Suppose $f : [a, b] \to \mathbb{R}$ is continuous on $[a, b]$, differentiable on (a, b), and $f(a) = f(b)$. Then there exists $c \in (a, b)$ such that

$$f'(c) = 0.$$

Proof. By the extreme value theorem, f attains a maximum and a minimum on $[a, b]$. If both occur only at the endpoints, then f is constant and any $c \in (a, b)$ works. Otherwise, f has an interior local extremum, so Fermat's theorem gives $f'(c) = 0$. □

Theorem 8 (Mean value theorem)
Suppose $f : [a, b] \to \mathbb{R}$ is continuous on $[a, b]$ and differentiable on (a, b). Then there exists $c \in (a, b)$ such that

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



The mean value theorem says some tangent slope equals the secant slope over the interval.

Proof. Define the line through the endpoints,

$$\ell(x) = f(a) + \left(\frac{f(b) - f(a)}{b - a}\right)(x - a),$$

and set $g(x) = f(x) - \ell(x)$. Then $g(a) = g(b) = 0$. By Rolle's theorem, there exists $c \in (a, b)$ with $g'(c) = 0$. Therefore

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

☐

Corollary 9 (Monotonicity from the derivative)
Let f be continuous on $[a, b]$ and differentiable on (a, b).

- If $f'(x) \geq 0$ for all $x \in (a, b)$, then f is increasing on $[a, b]$.
- If $f'(x) \leq 0$ for all $x \in (a, b)$, then f is decreasing on $[a, b]$.
- If $f'(x) = 0$ for all $x \in (a, b)$, then f is constant on $[a, b]$.

Proof. If $x < y$, apply the mean value theorem to f on $[x, y]$. There exists $c \in (x, y)$ such that

$$f(y) - f(x) = f'(c)(y - x).$$

The sign of $f'(c)$ determines the sign of $f(y) - f(x)$. ☐

Theorem 10 (Second derivative test)
Let f be differentiable in a neighbourhood of c, and suppose $f'(c) = 0$ and $f''(c)$ exists.

- If $f''(c) > 0$, then f has a local minimum at c.
- If $f''(c) < 0$, then f has a local maximum at c.
- If $f''(c) = 0$, the test is inconclusive.

Proof. Suppose $f''(c) > 0$. Since

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

there is a neighbourhood of c in which

$$\frac{f'(x) - f'(c)}{x - c} > 0$$

for $x \neq c$. As $f'(c) = 0$, this means $f'(x) < 0$ just to the left of c and $f'(x) > 0$ just to the right of c. By the monotonicity corollary, f decreases into c and increases after c, so c is a local minimum. The case $f''(c) < 0$ is the same with the inequalities reversed.

When $f''(c) = 0$, the examples x^4, $-x^4$, and x^3 at 0 show that a local minimum, a local maximum, or neither may occur. ☐

Corollary 11 (Bounded derivative gives uniform continuity)
If $f : I \to \mathbb{R}$ is differentiable on an interval I and there is $M \geq 0$ such that

$$|f'(x)| \leq M$$

for all $x \in I$, then f is uniformly continuous on I.

Proof. For $x, y \in I$, the mean value theorem gives some c between x and y such that

$$|f(x) - f(y)| = |f'(c)||x - y| \leq M|x - y|.$$

Choose $\delta = \frac{\varepsilon}{M}$ if $M > 0$, and any $\delta > 0$ if $M = 0$. □

Example 13
Since $|\cos x| \leq 1$ for all x,

$$|\sin x - \sin y| \leq |x - y|$$

for all $x, y \in \mathbb{R}$ by the previous corollary.

Example 14
The function $\log x$ is uniformly continuous on $[1, \infty)$ because

$$|(\log x)'| = \frac{1}{x} \leq 1$$

on that interval.

6 Riemann Integration

The Riemann integral makes signed area precise using lower and upper rectangles. A bounded function is integrable when these lower and upper approximations can be made arbitrarily close.

6.1 Partitions and Sums

Definition 1 (Partition)
Let $a < b$. A **partition** of $[a, b]$ is a finite set

$$P = \{x_0, x_1, \dots, x_n\}$$

such that

$$a = x_0 < x_1 < \dots < x_n = b.$$

The subintervals of P are $[x_{i-1}, x_i]$, and their lengths are

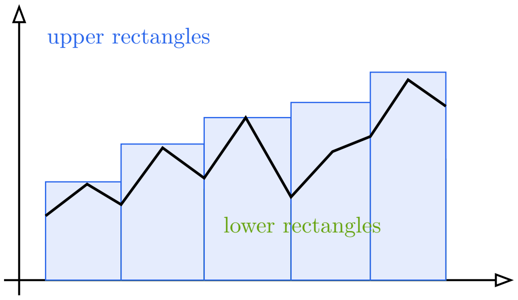
$$\Delta x_i = x_i - x_{i-1}.$$

Definition 2 (Lower and upper sums)
Let $f : [a, b] \to \mathbb{R}$ be bounded, and let $P = \{x_0, \dots, x_n\}$ be a partition of $[a, b]$. For each i, define

$$m_i = \inf\{f(x) : x \in [x_{i-1}, x_i]\}, \quad M_i = \sup\{f(x) : x \in [x_{i-1}, x_i]\}.$$

The **lower sum** and **upper sum** of f with respect to P are

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



Lower sums use the smallest value on each subinterval; upper sums use the largest value.

Proposition 1 (Basic bounds for sums)
If $|f(x)| \le M$ for every $x \in [a, b]$, then every partition P satisfies

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

Proof. Since $-M \le m_i \le M_i \le M$ on every subinterval,

$$-M\Delta x_i \le m_i \Delta x_i \le M_i \Delta x_i \le M\Delta x_i.$$

Summing over i gives the result. ☐

6.2 Refinements

Definition 3 (Refinement)
If P and P' are partitions of $[a, b]$, then P' is a **refinement** of P if

$$P \subset P'.$$

In other words, P' is obtained by adding extra partition points to P.

Lemma 2 (Refinements improve lower and upper sums)
If P' refines P, then

$$L(f, P) \leq L(f, P') \quad \text{and} \quad U(f, P') \leq U(f, P).$$

Proof. It is enough to consider adding one point c inside a subinterval $[x_{i-1}, x_i]$. Let

$$m = \inf_{[x_{i-1}, x_i]} f,$$

and let m_1, m_2 be the corresponding infima on $[x_{i-1}, c]$ and $[c, x_i]$. Since smaller intervals have fewer points,

$$m \leq m_1 \quad \text{and} \quad m \leq m_2.$$

Hence

$$m(x_i - x_{i-1}) \leq m_1(c - x_{i-1}) + m_2(x_i - c).$$

This shows the lower sum cannot decrease. The upper-sum argument is the same, using suprema and reversing the inequality. □

Corollary 3 (Lower sums are below upper sums)
If P and Q are any two partitions of $[a, b]$, then

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

Proof. Let $R = P \cup Q$. Then R refines both P and Q. By the refinement lemma,

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

□

6.3 Integrability

Definition 4 (Lower and upper integrals)
Let $f : [a, b] \to \mathbb{R}$ be bounded. The **lower integral** and **upper integral** are

$$\underline{\int_a^b} f = \sup\{L(f, P) : P \text{ is a partition of } [a, b]\},$$

and

$$\overline{\int_a^b} f = \inf\{U(f, P) : P \text{ is a partition of } [a, b]\}.$$

Definition 5 (Riemann integrable)
A bounded function $f : [a, b] \to \mathbb{R}$ is **Riemann integrable** on $[a, b]$ if

$$\underline{\int_a^b} f = \overline{\int_a^b} f.$$

In this case, their common value is denoted

$$\int_a^b f(x)\,\mathrm{d}x.$$

Theorem 4 (Integrability criterion)
A bounded function $f : [a, b] \to \mathbb{R}$ is Riemann integrable if and only if for every $\varepsilon > 0$, there exists a partition P of $[a, b]$ such that

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

Proof. Suppose f is integrable. By the definitions of supremum and infimum, choose partitions P_1 and P_2 such that

$$\int_a^b f - L(f, P_1) < \frac{\varepsilon}{2}, \quad U(f, P_2) - \int_a^b f < \frac{\varepsilon}{2}.$$

Let $P = P_1 \cup P_2$. By refinement,

$$U(f, P) - L(f, P) \le U(f, P_2) - L(f, P_1) < \varepsilon.$$

Conversely, if such partitions exist, then

$$0 \le \overline{\int_a^b} f - \underline{\int_a^b} f \le U(f, P) - L(f, P)$$

can be made smaller than every $\varepsilon > 0$. Hence the upper and lower integrals are equal.[4] □

Example 1
The constant function $f(x) = c$ is integrable on $[a, b]$, and

$$\int_a^b c\,\mathrm{d}x = c(b - a).$$

Indeed, every lower sum and every upper sum equals $c(b - a)$.

Example 2
The function

$$f(x) = \begin{cases} 1 \text{ if } & x \in \mathbb{Q} \\ 0 \text{ if } & x \text{ is irrational} \end{cases}$$

---

[4] This criterion is often called Darboux's criterion in analysis texts, because it is phrased using upper and lower sums.

is not integrable on $[0,1]$. Every subinterval contains both rational and irrational points, so for every partition P,

$$L(f,P)=0 \quad \text{and} \quad U(f,P)=1.$$

The gap can never be made small.

6.4 Computing Sums

Example 3
Let

$$f(x)=\begin{cases}4 \text{ if } & 0\le x<3\\ 0 \text{ if } & x=3\\ 1 \text{ if } & 3<x\le 5.\end{cases}$$

For

$$P_n=\left\{0,3-\frac{1}{n},3+\frac{1}{n},5\right\},$$

the lower sum uses heights $4,0,1$, and the upper sum uses heights $4,4,1$. Hence

$$L(f,P_n)=4\left(3-\frac{1}{n}\right)+0\left(\frac{2}{n}\right)+1\left(2-\frac{1}{n}\right)=14-\frac{5}{n},$$

while

$$U(f,P_n)=4\left(3-\frac{1}{n}\right)+4\left(\frac{2}{n}\right)+1\left(2-\frac{1}{n}\right)=14+\frac{3}{n}.$$

Therefore

$$U(f,P_n)-L(f,P_n)=\frac{8}{n}\to 0,$$

so f is integrable and

$$\int_0^5 f(x)\,\mathrm{d}x=14.$$

Example 4
Let $b>0$, let $f(x)=x^2$ on $[0,b]$, and use the equal partition $x_i=i\frac{b}{n}$. Since x^2 is increasing on $[0,b]$,

$$L(f,P_n)=\sum_{i=1}^{n}\left((i-1)\frac{b}{n}\right)^2\left(\frac{b}{n}\right)=\frac{b^3}{n^3}\sum_{i=1}^{n}(i-1)^2$$

and

$$U(f,P_n)=\sum_{i=1}^{n}\left(i\frac{b}{n}\right)^2\left(\frac{b}{n}\right)=\frac{b^3}{n^3}\sum_{i=1}^{n}i^2.$$

Using

$$\sum_{i=1}^{n} i^2 = n(n+1)\frac{2n+1}{6},$$

both sums tend to $\frac{b^3}{3}$. Hence

$$\int_0^b x^2 \, \mathrm{d}x = \frac{b^3}{3}.$$

6.5 Existence of Integrals

Theorem 5 (Continuous functions are integrable)
Every continuous function $f : [a,b] \to \mathbb{R}$ is Riemann integrable.

Proof. Since f is continuous on $[a,b]$, it is uniformly continuous by the closed interval uniform continuity theorem. Let $\varepsilon > 0$. Choose $\delta > 0$ such that

$$|x - y| < \delta \implies |f(x) - f(y)| < \frac{\varepsilon}{b-a}.$$

Choose a partition P whose subinterval lengths are all less than δ. On each subinterval, the difference between the supremum and infimum is at most $\frac{\varepsilon}{b-a}$. Hence

$$U(f,P) - L(f,P) \le \sum_{i=1}^{n} \frac{\varepsilon}{b-a} \Delta x_i = \varepsilon.$$

The integrability criterion proves integrability. ☐

Theorem 6 (Finite discontinuities)
If $f : [a,b] \to \mathbb{R}$ is bounded and continuous except at finitely many points, then f is Riemann integrable.

Proof. We prove the idea for one discontinuity $c \in (a,b)$. The endpoint cases and finitely many discontinuities are handled by repeating the same construction.

Let $|f(x)| \le M$ and let $\varepsilon > 0$. If $M = 0$, then $f = 0$ and there is nothing to prove. Otherwise choose $\eta > 0$ so small that $c - \eta, c + \eta \in [a,b]$ and

$$4M\eta < \frac{\varepsilon}{2}.$$

On the middle interval $[c - \eta, c + \eta]$, the upper-lower contribution is at most

$$(M - (-M)) \cdot 2\eta = 4M\eta < \frac{\varepsilon}{2}.$$

On the two closed intervals $[a, c - \eta]$ and $[c + \eta, b]$, the function is continuous, hence integrable by the continuous integrability theorem. Choose partitions on those intervals so their combined upper-lower gaps are less than $\frac{\varepsilon}{2}$. Adding the points $c - \eta$ and $c + \eta$ gives a partition of $[a,b]$ with total gap less than ε. ☐

6.6 Integral Properties

Theorem 7 (Basic properties of the integral)
Suppose f and g are integrable on $[a,b]$.

1. If $\lambda \in \mathbb{R}$, then λf is integrable and
$$\int_a^b \lambda f(x)\,\mathrm{d}x = \lambda \int_a^b f(x)\,\mathrm{d}x.$$
2. The function $f+g$ is integrable and
$$\int_a^b (f(x)+g(x))\,\mathrm{d}x = \int_a^b f(x)\,\mathrm{d}x + \int_a^b g(x)\,\mathrm{d}x.$$
3. If $f(x) \le g(x)$ for all $x \in [a,b]$, then
$$\int_a^b f(x)\,\mathrm{d}x \le \int_a^b g(x)\,\mathrm{d}x.$$
4. If $c \in [a,b]$, then
$$\int_a^b f(x)\,\mathrm{d}x = \int_a^c f(x)\,\mathrm{d}x + \int_c^b f(x)\,\mathrm{d}x.$$
5. The function $|f|$ is integrable and
$$\left|\int_a^b f(x)\,\mathrm{d}x\right| \le \int_a^b |f(x)|\,\mathrm{d}x.$$

Example 5
If f is integrable and $f(x) \ge 0$ on $[a,b]$, then
$$\int_a^b f(x)\,\mathrm{d}x \ge 0$$
by the order property, comparing f with the zero function.

Example 6
Let $a > 0$. On $[a, a+1]$, the function $f(x) = \frac{1}{x}$ satisfies
$$\frac{1}{a+1} \le \frac{1}{x}.$$
Integrating and using the order property gives
$$\frac{1}{a+1} \le \int_a^{a+1} \frac{1}{x}\,\mathrm{d}x = \log\left(1+\frac{1}{a}\right).$$

Proposition 8 (Zero integral of a non-negative continuous function)
Let $f : [a,b] \to \mathbb{R}$ be continuous and suppose $f(x) \ge 0$ for all x. If

$$\int_a^b f(x)\,\mathrm{d}x = 0,$$

then $f(x) = 0$ for all $x \in [a, b]$.

Proof. Suppose $f(c) > 0$ for some c. By continuity, there is a small interval around c on which $f(x) \geq \frac{f(c)}{2}$. The integral over that small interval is positive, and $f \geq 0$ elsewhere, so the integral over $[a, b]$ is positive. This contradicts the assumption. □

6.7 Fundamental Theorem of Calculus

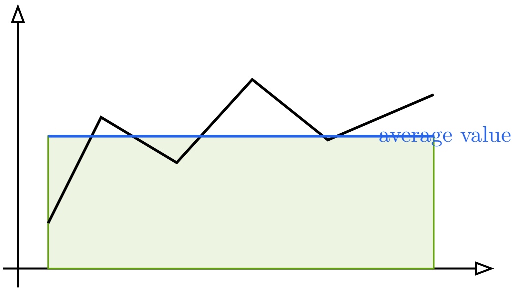
Theorem 9 (Mean value theorem for integrals)
If $f : [a, b] \to \mathbb{R}$ is continuous, then there exists $c \in [a, b]$ such that

$$\int_a^b f(x)\,\mathrm{d}x = f(c)(b - a).$$

Proof. By the extreme value theorem, f attains a minimum m and maximum M on $[a, b]$. The order property gives

$$m(b - a) \leq \int_a^b f(x)\,\mathrm{d}x \leq M(b - a).$$

Dividing by $b - a$ gives a number between m and M. By the intermediate value theorem, $f(c)$ equals that number for some $c \in [a, b]$. □



The average value of a continuous function is attained somewhere on the interval.

Theorem 10 (Fundamental theorem of calculus I)
Let $f : [a, b] \to \mathbb{R}$ be continuous and define

$$F(x) = \int_a^x f(t)\,\mathrm{d}t.$$

Then F is differentiable on (a, b) and

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

Proof. Fix $x \in (a, b)$. For small $h \neq 0$ with $x + h \in [a, b]$,

$$\frac{F(x+h)-F(x)}{h}=\frac{1}{h}\int_x^{x+h} f(t)\,\mathrm{d}t.$$

By the mean value theorem for integrals, this average equals $f(c_h)$ for some c_h between x and $x+h$. As $h \to 0$, we have $c_h \to x$, so by continuity $f(c_h) \to f(x)$. Hence $F'(x) = f(x)$. □

Theorem 11 (Fundamental theorem of calculus II)
If f is continuous on $[a,b]$ and A is continuous on $[a,b]$ with $A'(x) = f(x)$ for every $x \in (a,b)$, then

$$\int_a^b f(x)\,\mathrm{d}x = A(b) - A(a).$$

Proof. Let

$$F(x) = \int_a^x f(t)\,\mathrm{d}t.$$

By the first fundamental theorem, $F'(x) = f(x) = A'(x)$. Hence $(F - A)' = 0$ on (a,b), so $F - A$ is constant by the mean value theorem. Therefore

$$F(b) - F(a) = A(b) - A(a).$$

Since $F(a) = 0$ and $F(b) = \int_a^b f(x)\,\mathrm{d}x$, the result follows. □

7 Integration Techniques and Applications

The fundamental theorem turns many definite integrals into antiderivative calculations. The techniques below are the ones that appear most often in tutorials and exams.

7.1 Variable Limits

Proposition 1 (Leibniz rule for variable endpoints)
Let f be continuous and let u, v be differentiable. If

$$F(x) = \int_{u(x)}^{v(x)} f(t)\,\mathrm{d}t,$$

then

$$F'(x) = f(v(x))v'(x) - f(u(x))u'(x).$$

Proof. Fix an antiderivative A of f. By the fundamental theorem of calculus,

$$F(x) = A(v(x)) - A(u(x)).$$

Differentiate using the chain rule. □

Example 1
Let

$$F(x) = \int_{x^2}^{\log x} \tan(t)\,\mathrm{d}t.$$

Then

$$F'(x) = \frac{\tan(\log x)}{x} - \tan(x^2) \cdot 2x.$$

Example 2

Note
Found in the 2026 tutorials.

The Eulerian logarithmic integral is

$$\mathrm{Li}(x) = \int_2^x \frac{1}{\log t}\,\mathrm{d}t.$$

For $x > 2$,

$$\mathrm{Li}'(x) = \frac{1}{\log x},$$

and

$$\mathrm{Li}''(x) = -\frac{1}{x(\log x)^2}.$$

Example 3

Note
Found in the 2026 tutorials.

Let

$$g(x) = \int_0^x \frac{\sin t}{t+1}\,\mathrm{d}t.$$

We show $g(x) > 0$ for every $x > 0$. On each full period, the positive half dominates the following negative half because $\frac{1}{t+1}$ is decreasing:

$$\int_{2k\pi}^{(2k+1)\pi} \frac{\sin t}{t+1}\,\mathrm{d}t + \int_{(2k+1)\pi}^{(2k+2)\pi} \frac{\sin t}{t+1}\,\mathrm{d}t > 0.$$

Inside a negative half-period, the partial negative contribution is still smaller in magnitude than the preceding positive half-period. Hence every initial integral from 0 to x is positive.

7.2 The Logarithm from an Integral

[5]

Define, for $x > 0$,

$$L(x) = \int_1^x \frac{1}{t}\,\mathrm{d}t.$$

By the fundamental theorem, $L'(x) = \frac{1}{x}$ and $L(1) = 0$.

Proposition 2 (Integral logarithm laws)
For $x, y > 0$ and $\alpha \in \mathbb{R}$,

$$L(x \cdot y) = L(x) + L(y), \quad L(x^\alpha) = \alpha L(x).$$

Consequently, if e is defined by $L(e) = 1$, then $L(e^x) = x$.

Proof. Fix $y > 0$ and define

$$H(x) = L(xy) - L(x).$$

Then

$$H'(x) = \frac{y}{xy} - \frac{1}{x} = 0.$$

Hence H is constant. Since $H(1) = L(y)$, we get $L(x \cdot y) = L(x) + L(y)$.

For powers, first define $K(x) = L(x^\alpha) - \alpha L(x)$ for $x > 0$. Differentiating gives $K'(x) = 0$, and $K(1) = 0$, so $K = 0$. □

---

[5] In this course, log means the natural logarithm. We avoid writing ln; Mr Gauss and Ole Warnaar will haunt you if you use ln instead of log.

7.3 Integration by Parts

Theorem 3 (Integration by parts)
If u and v are differentiable, then

$$\int u(x)v'(x)\,\mathrm{d}x = u(x)v(x) - \int u'(x)v(x)\,\mathrm{d}x.$$

For a definite integral,

$$\int_a^b u(x)v'(x)\,\mathrm{d}x = [u(x)v(x)]_a^b - \int_a^b u'(x)v(x)\,\mathrm{d}x.$$

Proof. Integrate the product rule

$$(uv)' = u'v + uv'.$$

□

Note
A useful choice heuristic for u is logarithmic, inverse trigonometric, algebraic, trigonometric, exponential, in that order. It is only a heuristic; the algebra decides.

Example 4
For $\int x\log x\,\mathrm{d}x$, take $u = \log x$ and $v' = x$. Then $u' = \frac{1}{x}$ and $v = \frac{x^2}{2}$, so

$$\int x\log x\,\mathrm{d}x = x^2\log\frac{x}{2} - \int \frac{x}{2}\,\mathrm{d}x = x^2\log\frac{x}{2} - \frac{x^2}{4} + C.$$

Example 5
Let

$$I = \int e^{2x}\sin x\,\mathrm{d}x.$$

Integrating by parts twice gives

$$I = \frac{e^{2x}\sin x}{2} - \frac{1}{2}\int e^{2x}\cos x\,\mathrm{d}x,$$

and

$$\int e^{2x}\cos x\,\mathrm{d}x = \frac{e^{2x}\cos x}{2} + \frac{1}{2}\int e^{2x}\sin x\,\mathrm{d}x.$$

Hence

$$I = \frac{e^{2x}\sin x}{2} - \frac{e^{2x}\cos x}{4} - \frac{I}{4},$$

so

$$\int e^{2x}\sin x\,\mathrm{d}x = e^{2x}\frac{2\sin x - \cos x}{5} + C.$$

Proposition 4 (Repeated integral identity)
If f is continuous, then

$$\int_0^x \left(\int_0^t f(u)\,\mathrm{d}u \right) \mathrm{d}t = \int_0^x f(u)(x-u)\,\mathrm{d}u.$$

Proof. Let $G(t) = \int_0^t f(u)\,\mathrm{d}u$. Integrate by parts with $u = G(t)$ and $v' = 1$:

$$\int_0^x G(t)\,\mathrm{d}t = [tG(t)]_0^x - \int_0^x tf(t)\,\mathrm{d}t = x\int_0^x f(t)\,\mathrm{d}t - \int_0^x tf(t)\,\mathrm{d}t.$$

Combine the two integrals. ☐

7.4 Substitution

Theorem 5 (Integration by substitution)
Let g be continuously differentiable and let f be continuous. Then

$$\int_a^b f(g(x))g'(x)\,\mathrm{d}x = \int_{g(a)}^{g(b)} f(u)\,\mathrm{d}u.$$

Proof. Let F be an antiderivative of f. Then $(F \circ g)' = (f \circ g)g'$. Apply the fundamental theorem to both sides. ☐

Example 6
For $\int \log(\cos x)\tan x\,\mathrm{d}x$, take $u = \log(\cos x)$. Then

$$\mathrm{d}u = -\tan x\,\mathrm{d}x.$$

Hence

$$\int \log(\cos x)\tan x\,\mathrm{d}x = -\int u\,\mathrm{d}u = -\frac{1}{2}(\log(\cos x))^2 + C.$$

Example 7

$$\int \sin\left(\frac{x}{2}\right)\cos\left(\frac{x}{2}\right)\mathrm{d}x = \int 2\sin(u)\cos(u)\,\mathrm{d}u = \int \sin(2u)\,\mathrm{d}u = -\frac{1}{2}\cos x + C,$$

where $u = \frac{x}{2}$.

Example 8
For

$$\int \frac{\cos\left(x^{\frac{1}{3}}\right)}{x^{\frac{1}{3}}}\,\mathrm{d}x,$$

take $u = x^{\frac{1}{3}}$, so $x = u^3$ and $\mathrm{d}x = 3u^2\,\mathrm{d}u$. Then

$$\int \frac{\cos\left(x^{\frac{1}{3}}\right)}{x^{\frac{1}{3}}}\,\mathrm{d}x = 3\int u\cos u\,\mathrm{d}u = 3u\sin u + 3\cos u + C.$$

Therefore the antiderivative is

$$3x^{\frac{1}{3}}\sin\left(x^{\frac{1}{3}}\right)+3\cos\left(x^{\frac{1}{3}}\right)+C.$$

Example 9

$$\int_1^e \frac{\sin(\log x)}{x}\,\mathrm{d}x=\int_0^1 \sin u\,\mathrm{d}u=1-\cos 1.$$

7.5 Partial Fractions

Definition 1 (Rational function)
A **rational function** is a quotient $\frac{p(x)}{q(x)}$ of polynomials. When the degree of p is at least the degree of q, first use polynomial division. Then factor the denominator and decompose into simpler fractions.

Example 10
Compute

$$\int \frac{2x+1}{x^2+7x+12}\,\mathrm{d}x.$$

Since

$$x^2+7x+12=(x+3)(x+4),$$

write

$$\frac{2x+1}{(x+3)(x+4)}=\frac{A}{x+3}+\frac{B}{x+4}.$$

Then

$$2x+1=A(x+4)+B(x+3),$$

giving $A=-5$ and $B=7$. Therefore

$$\int \frac{2x+1}{x^2+7x+12}\,\mathrm{d}x=-5\log|x+3|+7\log|x+4|+C.$$

Example 11
For

$$\int \frac{e^x}{e^{2x}+e^x+1}\,\mathrm{d}x,$$

take $u=e^x$, so $\mathrm{d}u=e^x\,\mathrm{d}x$. Then

$$\int \frac{1}{u^2+u+1}\,\mathrm{d}u.$$

Complete the square:

$$u^2+u+1=\left(u+\frac{1}{2}\right)^2+\frac{3}{4}.$$

Hence

$$\int \frac{e^x}{e^{2x}+e^x+1}\,\mathrm{d}x = \frac{2}{\sqrt{3}}\arctan\left(\frac{2e^x+1}{\sqrt{3}}\right) + C.$$

7.6 Trigonometric Substitution

Note
Found in the 2026 tutorials.

The common substitutions are:

- use $x = a\sin\theta$ for $\sqrt{a^2 - x^2}$;
- use $x = a\tan\theta$ for $\sqrt{a^2 + x^2}$;
- use $x = a\sec\theta$ for $\sqrt{x^2 - a^2}$.

Example 12

$$\int \frac{1}{\sqrt{9-x^2}}\,\mathrm{d}x = \arcsin\left(\frac{x}{3}\right) + C.$$

This follows from $x = 3\sin\theta$.

Example 13
To compute

$$\int \frac{x}{\sqrt{4-4x^2}}\,\mathrm{d}x,$$

it is faster to use $u = 4 - 4x^2$. Since $\mathrm{d}u = -8x\,\mathrm{d}x$,

$$\int \frac{x}{\sqrt{4-4x^2}}\,\mathrm{d}x = -\frac{1}{8}\int u^{-\frac{1}{2}}\,\mathrm{d}u = -\frac{1}{4}\sqrt{4-4x^2} + C.$$

7.7 Definite Integrals and Areas

Note
The examples in this section are tutorial-style practice.

Example 14
To compute

$$\int_0^3 |3x-4|\,\mathrm{d}x,$$

split at $x = \frac{4}{3}$:

$$\int_0^{\frac{4}{3}} (4-3x)\,\mathrm{d}x + \int_{\frac{4}{3}}^3 (3x-4)\,\mathrm{d}x = \frac{8}{3} + \frac{25}{6} = \frac{41}{6}.$$

Example 15
For

$$\int_1^5 \frac{2x^5 - x + 3}{x^2}\,\mathrm{d}x,$$

simplify first:

$$\frac{2x^5 - x + 3}{x^2} = 2x^3 - \frac{1}{x} + \frac{3}{x^2}.$$

Therefore

$$\int_1^5 \left(2x^3 - \frac{1}{x} + 3x^{-2}\right)\mathrm{d}x = \left[\frac{x^4}{2} - \log x - \frac{3}{x}\right]_1^5.$$

Definition 2 (Area between curves)
If $f(x) \geq g(x)$ on $[a, b]$, then the area between the curves $y = f(x)$ and $y = g(x)$ is

$$\int_a^b (f(x) - g(x))\,\mathrm{d}x.$$

If the curves cross, split the interval at their intersection points.

7.8 Improper Integrals

Definition 3 (Improper integrals)
If f is integrable on $[a + \varepsilon, b]$ for every $\varepsilon > 0$, define

$$\int_a^b f(x)\,\mathrm{d}x = \lim_{\varepsilon \to 0^+} \int_{a+\varepsilon}^b f(x)\,\mathrm{d}x,$$

provided the limit exists. This handles an unbounded integrand at a.

If f is integrable on $[a, R]$ for every $R > a$, define

$$\int_a^\infty f(x)\,\mathrm{d}x = \lim_{R \to \infty} \int_a^R f(x)\,\mathrm{d}x,$$

provided the limit exists. The definitions at b^- and $-\infty$ are analogous.

Example 16
For $p \neq 1$,

$$\int_1^R \frac{1}{x^p}\,\mathrm{d}x = \frac{R^{1-p} - 1}{1 - p}.$$

Hence

$$\int_1^\infty \frac{1}{x^p}\,\mathrm{d}x$$

converges exactly when $p > 1$. For $p = 1$, the integral is $\log R$, which diverges.

Example 17

$$\int_0^1 \frac{1}{\sqrt{x}}\,\mathrm{d}x = \lim_{\varepsilon\to 0^+}\int_\varepsilon^1 x^{-\frac{1}{2}}\,\mathrm{d}x = \lim_{\varepsilon\to 0^+}\left(2-2\sqrt{\varepsilon}\right) = 2.$$

Example 18

The improper integral

$$\int_0^\infty \cos x\,\mathrm{d}x$$

diverges, because

$$\int_0^R \cos x\,\mathrm{d}x = \sin R$$

has no limit as $R \to \infty$.

8 Series and Taylor Expansions

A series is an infinite sum, but convergence is still a statement about sequences: the sequence of partial sums must converge. Most of this chapter is about deciding whether that happens without explicitly finding the sum.

8.1 Series and Partial Sums

Definition 1 (Series)
Let (a_n) be a sequence. The expression

$$\sum_{n=1}^{\infty} a_n = a_1 + a_2 + a_3 + \ldots$$

is called a **series**. The Nth **partial sum** is

$$S_N = \sum_{n=1}^{N} a_n.$$

Definition 2 (Convergent series)
The series $\sum_{n=1}^{\infty} a_n$ **converges** to S if the partial sums satisfy

$$S_N \to S.$$

In that case we write

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

If the partial sums do not converge, the series **diverges**.

Theorem 1 (Term test)
If $\sum_{n=1}^{\infty} a_n$ converges, then

$$a_n \to 0.$$

Proof. Since $a_n = S_n - S_{n-1}$ and both partial-sum sequences tend to the same limit S,

$$a_n \to S - S = 0.$$

□

Note
The converse is false. The harmonic series has terms $\frac{1}{n} \to 0$ but still diverges.

Theorem 2 (Cauchy criterion for series)
The series $\sum_{n=1}^{\infty} a_n$ converges if and only if for every $\varepsilon > 0$, there exists $N \in \mathbb{N}$ such that whenever $q > p \geq N$,

$$\left| \sum_{n=p+1}^{q} a_n \right| < \varepsilon.$$

Proof. This is exactly the Cauchy criterion for the partial sums, since

$$S_q - S_p = \sum_{n=p+1}^{q} a_n.$$

☐

Example 1
The harmonic series $\sum_{n=1}^{\infty} \frac{1}{n}$ diverges. If it converged, the Cauchy criterion with $\varepsilon = \frac{1}{3}$ would give N such that every tail past N has size less than $\frac{1}{3}$. But

$$\sum_{n=N+1}^{2N} \frac{1}{n} > \sum_{n=N+1}^{2N} \frac{1}{2N} = \frac{1}{2},$$

a contradiction.

8.2 First Examples

Proposition 3 (Geometric series)
Let $r \in \mathbb{R}$. The geometric series $\sum_{n=0}^{\infty} r^n$ converges if and only if $|r| < 1$. In that case,

$$\sum_{n=0}^{\infty} r^n = \frac{1}{1-r}.$$

Proof. For $r \neq 1$,

$$S_N = 1 + r + \ldots + r^N = \frac{1 - r^{N+1}}{1-r}.$$

This converges exactly when $r^{N+1} \to 0$, which is equivalent to $|r| < 1$. If $r = 1$, the partial sums diverge. If $r = -1$, the terms do not tend to 0. ☐

Example 2
For $|x| < 1$,

$$\frac{1}{1-x} = \sum_{n=0}^{\infty} x^n.$$

This identity is the source of many Maclaurin series later in the chapter.

Proposition 4 (Telescoping series)
If most terms cancel in the partial sums, compute the partial sums explicitly before using a convergence test.

Example 3
Since

$$\frac{1}{n^2+3n+2} = \frac{1}{(n+1)(n+2)} = \frac{1}{n+1} - \frac{1}{n+2},$$

the partial sums telescope:

$$\sum_{n=1}^{N} \frac{1}{n^2+3n+2} = \frac{1}{2} - \frac{1}{N+2}.$$

Hence

$$\sum_{n=1}^{\infty} \frac{1}{n^2+3n+2} = \frac{1}{2}.$$

8.3 Comparison Tests

Theorem 5 (Comparison test)
Suppose $0 \le a_n \le b_n$ for all sufficiently large n.

- If $\sum b_n$ converges, then $\sum a_n$ converges.
- If $\sum a_n$ diverges, then $\sum b_n$ diverges.

Proof. For non-negative terms, the partial sums are increasing. If $\sum b_n$ converges, then the partial sums of $\sum a_n$ are increasing and bounded above by the partial sums of $\sum b_n$ plus finitely many earlier terms. Hence they converge by the monotone convergence theorem. □

Theorem 6 (Limit comparison test)
Suppose $a_n > 0$, $b_n > 0$, and

$$\lim_{n\to\infty} \frac{a_n}{b_n} = c$$

for some $0 < c < \infty$. Then $\sum a_n$ and $\sum b_n$ either both converge or both diverge.

Proof. Since $\frac{a_n}{b_n} \to c > 0$, eventually

$$\frac{c}{2} \le \frac{a_n}{b_n} \le 2c.$$

Thus $(\frac{c}{2})b_n \le a_n \le 2cb_n$ eventually. Apply the comparison test in both directions. □

Example 4
The series

$$\sum_{n=1}^{\infty} \frac{n^2+2}{n^4+5}$$

converges by limit comparison with $\frac{1}{n^2}$, since

$$\frac{\frac{n^2+2}{n^4+5}}{\frac{1}{n^2}} = \frac{n^4+2n^2}{n^4+5} \to 1.$$

Example 5
The series

$$\sum_{n=1}^{\infty} n^2 \log \frac{n}{n^3+1}$$

diverges, since the terms are comparable to $\frac{\log n}{n}$, and

$$\sum_{n=2}^{\infty} \frac{\log n}{n}$$

diverges by the integral test.

Example 6
Consider

$$\sum_{n=N}^{\infty} \frac{e^{-an+\frac{1}{n}}}{n},$$

where N is fixed and $a \in \mathbb{R}$. If $a > 0$, then the terms are eventually bounded by a constant multiple of $\frac{e^{-an}}{n}$, so the series converges by comparison with a geometric series. If $a = 0$, then $\frac{e^{\frac{1}{n}}}{n} \geq \frac{1}{n}$, so the series diverges. If $a < 0$, the terms do not tend to 0. Hence the series diverges exactly when $a \leq 0$.

8.4 Integral Test and p-Series

Theorem 7 (Integral test)
Let $f : [1, \infty) \to \mathbb{R}$ be continuous, non-negative, and decreasing. Then

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

and

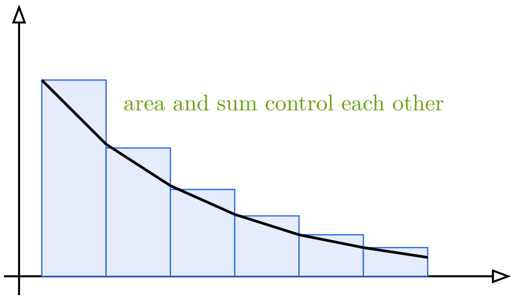
$$\int_{1}^{\infty} f(x)\,\mathrm{d}x$$

converge or diverge together.

Proof. For a decreasing function, the rectangles of height $f(n)$ trap the integral:

$$\sum_{n=2}^{N} f(n) \leq \int_{1}^{N} f(x)\,\mathrm{d}x \leq \sum_{n=1}^{N-1} f(n).$$

These inequalities show that the partial sums are bounded exactly when the improper integrals are bounded. □



For a decreasing positive function, unit-width rectangles compare the series with the area under the graph.

Proposition 8 (p-series)
The series

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

converges if and only if $p > 1$.

Proof. If $p \leq 0$, the terms do not tend to 0. If $0 < p \leq 1$, compare with the harmonic series when $p < 1$, and use the harmonic series when $p = 1$. If $p > 1$, the integral test gives convergence because

$$\int_1^{\infty} \frac{1}{x^p} \, \mathrm{d}x$$

converges exactly when $p > 1$. □

Example 7
The series

$$\sum_{n=2}^{\infty} \frac{1}{n(\log n)^p}$$

converges if and only if $p > 1$. This follows from the integral test using $u = \log x$:

$$\int_2^{\infty} \frac{1}{x(\log x)^p} \, \mathrm{d}x = \int_{\log 2}^{\infty} \frac{1}{u^p} \, \mathrm{d}u.$$

8.5 Root and Ratio Tests

Theorem 9 (Root test)
Let a_n be a sequence and set

$$\rho = \limsup_{n \to \infty} \sqrt[n]{|a_n|}.$$

If $\rho < 1$, then $\sum a_n$ converges absolutely. If $\rho > 1$, then $\sum a_n$ diverges. If $\rho = 1$, the test is inconclusive.

Example 8
For

$$\sum_{n=1}^{\infty} \frac{n^n}{3^{1+2n}},$$

the root test gives

$$\sqrt[n]{\frac{n^n}{3^{1+2n}}} = \frac{n}{9\sqrt[n]{3}} \to \infty.$$

Hence the series diverges.

Example 9
For

$$\sum_{n=1}^{\infty} \left(\frac{5n - 3n^3}{7n^3 + 2}\right)^n,$$

the root test gives

$$\left|\frac{5n - 3n^3}{7n^3 + 2}\right| \to \frac{3}{7} < 1,$$

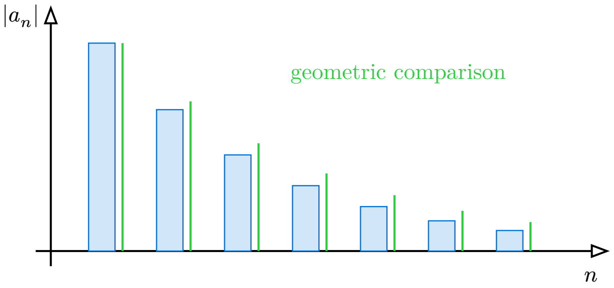
so the series converges absolutely.

Theorem 10 (Ratio test)
Suppose $a_n \neq 0$ eventually and

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

exists.

- If $r < 1$, then $\sum a_n$ converges absolutely.
- If $r > 1$, then $\sum a_n$ diverges.
- If $r = 1$, the test is inconclusive.



Ratio and root tests compare the tail of a series with a geometric tail. If the eventual shrink factor is below 1, the tail is controlled by a convergent geometric series.

Example 10
For

$$\sum_{n=1}^{\infty} \frac{(2n)!}{5^n n!},$$

the ratio is

$$\frac{a_{n+1}}{a_n} = \frac{(2n+2)(2n+1)}{5(n+1)} = 2\frac{2n+1}{5} \to \infty.$$

Hence the series diverges.

Example 11
For

$$\sum_{n=1}^{\infty} \frac{n!}{(2n)!},$$

the ratio is

$$\frac{a_{n+1}}{a_n} = \frac{n+1}{(2n+2)(2n+1)} = \frac{1}{2(2n+1)} \to 0.$$

Hence the series converges.

8.6 Alternating and Absolute Convergence

Theorem 11 (Alternating series test)
Suppose $b_n \geq 0$, $b_{n+1} \leq b_n$ for all sufficiently large n, and $b_n \to 0$. Then

$$\sum_{n=1}^{\infty} (-1)^{n+1} b_n$$

converges.

Proof. The even and odd partial sums are monotone and bound each other. Their difference is one term b_n, which tends to 0, so they converge to the same limit. ☐

Definition 3 (Absolute and conditional convergence)
The series $\sum a_n$ is **absolutely convergent** if $\sum |a_n|$ converges. It is **conditionally convergent** if $\sum a_n$ converges but $\sum |a_n|$ diverges.

Proposition 12 (Absolute convergence implies convergence)
If $\sum |a_n|$ converges, then $\sum a_n$ converges.

Proof. This follows from the Cauchy criterion and the triangle inequality:

$$\left| \sum_{n=p+1}^{q} a_n \right| \leq \sum_{n=p+1}^{q} |a_n|.$$

☐

Corollary 13 (Absolute convergence estimates)
If $\sum |a_n|$ converges, then

$$\left| \sum_{n=1}^{\infty} a_n \right| \leq \sum_{n=1}^{\infty} |a_n|.$$

If (b_n) is bounded, then $\sum a_n b_n$ also converges absolutely.

Proof. The first inequality follows by applying the finite triangle inequality to the partial sums and taking limits. If $|b_n| \leq M$, then

$$\sum |a_n b_n| \leq M \sum |a_n|,$$

so $\sum a_n b_n$ converges by comparison. □

Example 12
The series

$$\sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{\sqrt[4]{n}}$$

converges by the alternating series test, but it does not converge absolutely because $\sum \frac{1}{n^{\frac{1}{4}}}$ diverges. Hence it converges conditionally.

Example 13
The series

$$\sum_{n=1}^{\infty} \sin \frac{n}{\sqrt{n^4+1}}$$

converges absolutely, since

$$\left|\sin \frac{n}{\sqrt{n^4+1}}\right| \le \frac{1}{n^2}$$

and $\sum \frac{1}{n^2}$ converges.

8.7 Cauchy Condensation

Note
Starred result from the 2026 tutorials.

Theorem 14 (Cauchy condensation test)
Let (a_n) be non-negative and decreasing. Then

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

converges if and only if

$$\sum_{k=0}^{\infty} 2^k a_{2^k}$$

converges.

Proof. Group the original series into dyadic blocks:

$$a_1 + (a_2 + a_3) + (a_4 + \ldots + a_7) + \ldots.$$

Since (a_n) is decreasing, the block from 2^k to $2^{k+1} - 1$ is trapped between constant multiples of $2^k a_{2^k}$ and $2^k a_{2^{k+1}}$. These comparisons show the two series converge or diverge together. □

Example 14
Applying condensation to $a_n = \frac{1}{n^p}$ gives

$$2^k a_{2^k} = \frac{2^k}{(2^k)^p} = \left(2^{1-p}\right)^k.$$

This geometric series converges exactly when $p > 1$, recovering the p-series test.

8.8 Taylor and Maclaurin Series

Definition 4 (Taylor series)
Suppose f has derivatives of all orders at a. The **Taylor series** of f centred at a is

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

When $a = 0$, this is called the **Maclaurin series**.

Note
A Taylor series need not equal the original function without further hypotheses. In this course, the standard expansions below may be used on their intervals of convergence.

Proposition 15 (Standard Maclaurin series)

$$e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!}, \quad x \in \mathbb{R}.$$

$$\sin x = \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n+1}}{(2n+1)!}, \quad x \in \mathbb{R}.$$

$$\cos x = \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n}}{(2n)!}, \quad x \in \mathbb{R}.$$

$$\frac{1}{1-x} = \sum_{n=0}^{\infty} x^n, \quad |x| < 1.$$

$$\log(1+x) = \sum_{n=1}^{\infty} (-1)^{n-1} \frac{x^n}{n}, \quad -1 < x \le 1.$$

Example 15
Since

$$\sin \pi = \pi - \frac{\pi^3}{3!} + \frac{\pi^5}{5!} - \frac{\pi^7}{7!} + \dots = 0,$$

the series

$$\pi - \frac{\pi^3}{3!} + \frac{\pi^5}{5!} - \frac{\pi^7}{7!} + \dots$$

sums to 0.

Example 16
Since

$$\cos e = 1 - \frac{e^2}{2!} + \frac{e^4}{4!} - \frac{e^6}{6!} + \dots,$$

the series

$$1 - \frac{e^2}{2!} + \frac{e^4}{4!} - \frac{e^6}{6!} + \dots$$

sums to $\cos e$.

Example 17
Using $\log(1+x)$ at $x = 1$,

$$1-\frac{1}{2}+\frac{1}{3}-\frac{1}{4}+\ldots=\log 2.$$

Example 18
For $\log(4+x^2)$,

$$\log(4+x^2)=\log 4+\log\left(1+\frac{x^2}{4}\right)=2\log 2+\sum_{n=1}^{\infty}(-1)^{n-1}\frac{x^{2n}}{n4^n}.$$

This expansion is valid for $|x|<2$, and also at $x=\pm 2$ by the endpoint convergence of the logarithm series.

Example 19
Since

$$\log\left((1+x)^2\right)=2\log(1+x),$$

for $x>-1$, we have

$$\log\left((1+x)^2\right)=2\sum_{n=1}^{\infty}(-1)^{n-1}\frac{x^n}{n}.$$

The usual interval from the logarithm series is $-1<x\le 1$.

Example 20
Subtracting the two logarithm series gives

$$\log\left(\frac{1+x}{1-x}\right)=\log(1+x)-\log(1-x)=2\sum_{k=0}^{\infty}\frac{x^{2k+1}}{2k+1},\quad |x|<1.$$

Only the odd powers remain.

Example 21
Using $\sin^2 x=\frac{1-\cos(2x)}{2}$,

$$\sin^2 x=\frac{1}{2}\left(1-\sum_{n=0}^{\infty}(-1)^n\frac{(2x)^{2n}}{(2n)!}\right).$$

The constant terms cancel, leaving a series starting with x^2.

Example 22
Multiplication by a polynomial is usually cleaner than differentiating a product. For

$$f(x)=(\sin x)(x^2+3),$$

use $\sin x=\sum_{n=0}^{\infty}(-1)^n\frac{x^{2n+1}}{(2n+1)!}$ and multiply by x^2+3:

$$f(x)=3\sum_{n=0}^{\infty}(-1)^n\frac{x^{2n+1}}{(2n+1)!}+\sum_{n=0}^{\infty}(-1)^n\frac{x^{2n+3}}{(2n+1)!}.$$

Example 23
From the geometric series,

$$\frac{1}{1+u} = \sum_{n=0}^{\infty} (-1)^n u^n, \quad |u| < 1.$$

Taking $u = e^{-x}$ gives

$$\frac{1}{1+e^{-x}} = \sum_{n=0}^{\infty} (-1)^n e^{-nx}, \quad x > 0.$$

8.9 Shifted Standard Series

Proposition 16 (Useful Taylor series at zero)
For fixed $a \neq 0$, the following follow from the geometric series and standard binomial series, with the usual radius restrictions. For the logarithm formula, assume $a < 0$ so that $x - a$ is real and positive near 0.

$$\frac{1}{x-a} = -\frac{1}{a}\sum_{n=0}^{\infty}\left(\frac{x}{a}\right)^n,$$

$$\log(x-a) = \log(-a) - \sum_{n=1}^{\infty}\frac{x^n}{na^n}$$

$$\frac{1}{\sqrt{1-x}} = \sum_{n=0}^{\infty}\binom{2n}{n}\frac{x^n}{4^n}, \quad \frac{1}{\sqrt{1-x^2}} = \sum_{n=0}^{\infty}\binom{2n}{n}\frac{x^{2n}}{4^n},$$

and

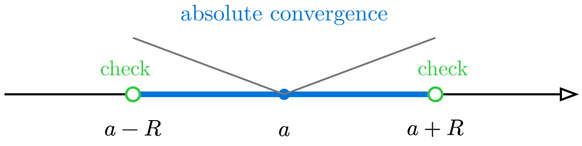
$$\arcsin x = \sum_{n=0}^{\infty}\binom{2n}{n}\frac{x^{2n+1}}{4^n(2n+1)}.$$

8.10 Radius of Convergence

Definition 5 (Power series and radius of convergence)
A **power series** centred at a is a series of the form

$$\sum_{n=0}^{\infty} c_n (x-a)^n.$$

Its **radius of convergence** is the number $R \in [0,\infty]$ such that the series converges absolutely for $|x-a| < R$ and diverges for $|x-a| > R$. Endpoints must be checked separately.



A radius of convergence gives an open interval of guaranteed convergence. The endpoints $a - R$ and $a + R$ must be checked separately.

Example 24
For

$$\sum_{n=0}^{\infty} \frac{x^n}{n!},$$

the ratio test gives

$$\left| \frac{x^{n+1}}{(n+1)!} \cdot \frac{n!}{x^n} \right| = \frac{|x|}{n+1} \to 0.$$

Thus the radius of convergence is ∞.

9 Systems of Linear Equations and Matrices

Linear algebra starts here with systems of equations. Matrix notation is useful because the same row operations solve the system, describe the solution set, and later test invertibility.

9.1 Linear Systems

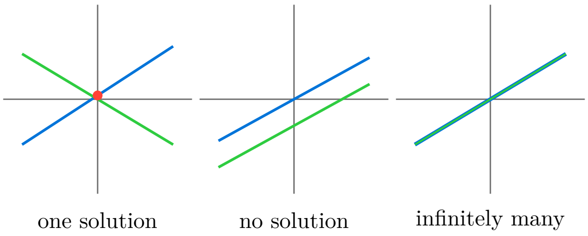
Definition 1 (Linear equation)
A **linear equation** in variables $x_1, \dots, x_n$ has the form

$$a_1 x_1 + a_2 x_2 + \dots + a_n x_n = b,$$

where $a_1, \dots, a_n, b \in \mathbb{R}$ are fixed scalars. A **linear system** is a finite list of linear equations in the same variables.

Definition 2 (Solution set)
A **solution** of a system is an ordered tuple $(x_1, \dots, x_n)$ satisfying every equation. The **solution set** is the set of all solutions.

Note
A linear system has exactly one of the following behaviours: no solution, exactly one solution, or infinitely many solutions. The row-reduction process below is how we tell which case we are in.



In $\mathbb{R}^2$, a system of two linear equations asks where two lines meet.

Example 1
The system

$$2x + y + 8z = 0, \quad x + 2y + z = 3, \quad 3x + 7y + z = 11$$

has infinitely many solutions. Row reduction gives

$$x + 5z = -1, \quad y - 2z = 2.$$

Taking $z = t$,

$$(x, y, z) = (-1 - 5t, 2 + 2t, t), \quad t \in \mathbb{R}.$$

9.2 Matrices and Augmented Matrices

Definition 3 (Matrix)
An $m \times n$ **matrix** is a rectangular array of scalars with m rows and n columns:

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}.$$

The entry in row i and column j is denoted a_{ij}.

Definition 4 (Coefficient and augmented matrices)
The system $Ax = b$ has **coefficient matrix** A and **augmented matrix** $[A \mid b]$. For example,

$$\begin{pmatrix} 2 & 1 & 8 \\ 1 & 2 & 1 \\ 3 & 7 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 3 \\ 11 \end{pmatrix}$$

corresponds to

$$[A \mid b] = \begin{pmatrix} 2 & 1 & 8 & 0 \\ 1 & 2 & 1 & 3 \\ 3 & 7 & 1 & 11 \end{pmatrix}.$$

Definition 5 (Elementary row operations)
The elementary row operations are:

- swap two rows,
- multiply a row by a non-zero scalar,
- add a scalar multiple of one row to another row.

These operations do not change the solution set of the corresponding linear system.

Proposition 1 (Row equivalence preserves solutions)
If one augmented matrix is obtained from another by elementary row operations, then the two augmented matrices represent systems with the same solution set.

Proof. Swapping equations does not change which tuples satisfy all equations. Multiplying an equation by a non-zero scalar gives an equivalent equation. Replacing one equation by itself plus a multiple of another is reversible by subtracting the same multiple, so it also preserves the solution set. ☐

Note
Each elementary row operation is itself an invertible transformation of the rows. This is why row reduction can simplify a system without losing information: every step can be undone by another elementary row operation.

Definition 6 (Elementary matrix)
An **elementary matrix** is obtained by applying one elementary row operation to an identity matrix.

Proposition 2 (Row operations as matrix multiplication)
Applying an elementary row operation to an $m \times n$ matrix is the same as left-multiplying by an invertible $m \times m$ elementary matrix.

Proof. Apply the row operation to I_m and call the result E. For any $m \times n$ matrix A, the product EA performs that same operation on the rows of A. The row operation can be undone, so the inverse operation gives an elementary matrix F with $FE = EF = I_m$. Hence E is invertible. ☐

9.3 Echelon Form

Definition 7 (Row echelon form)
A matrix is in **row echelon form** if:

- every non-zero row is above every zero row,
- each leading non-zero entry is strictly to the right of the leading non-zero entry in the row above,
- entries below each leading entry are zero.

A leading non-zero entry is called a **pivot**.

Definition 8 (Reduced row echelon form)
A matrix is in **reduced row echelon form** if it is in row echelon form, every pivot is 1, and every pivot is the only non-zero entry in its column.

Proposition 3 (Reading a reduced system)
In a reduced augmented matrix:

- a row of the form $0 = c$ with $c \neq 0$ means the system is inconsistent;
- a variable in a pivot column is a pivot variable;
- a variable in a non-pivot column is a free variable.

Example 2
Row reduction gives

$$\begin{pmatrix} 2 & -2 & 4 & 0 \\ 1 & 4 & -3 & 3 \\ 4 & -3 & 7 & 1 \end{pmatrix} \sim \begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

The last row says $0 = 1$, so the system is inconsistent.

Example 3
Row reduction gives

$$\begin{pmatrix} 2 & 3 & -2 & 1 & -2 \\ 1 & 1 & 3 & 5 & 1 \\ 2 & 4 & 5 & 7 & 9 \end{pmatrix} \sim \begin{pmatrix} 1 & 0 & 0 & 3 & -6 \\ 0 & 1 & 0 & -1 & 4 \\ 0 & 0 & 1 & 1 & 1 \end{pmatrix}.$$

Writing $x_4 = t$, the solution set is

$$(x_1, x_2, x_3, x_4) = (-6 - 3t, 4 + t, 1 - t, t), \quad t \in \mathbb{R}.$$

9.4 Homogeneous Systems

Definition 9 (Homogeneous system)
A system of the form

$$Ax = 0$$

is called **homogeneous**.

Proposition 4 (Homogeneous systems are always consistent)
Every homogeneous system has at least the zero solution $x = 0$.

Note
The zero solution is called the **trivial solution**. A homogeneous system has non-trivial solutions exactly when at least one variable is free after row reduction.

Definition 10 (Kernel and rank)
For a matrix A, the **kernel** is the solution set of the homogeneous system:

$$\ker A = \{x : Ax = 0\}.$$

The **rank** of A is the number of non-zero rows in a row echelon form of A.

Example 4
For

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 0 & 0 & 1 \\ 2 & 4 & 6 \end{pmatrix},$$

the homogeneous system $Ax = 0$ gives $z = 0$ and $x + 2y = 0$. Hence

$$(x, y, z) = t(-2, 1, 0), \quad t \in \mathbb{R}.$$

Example 5
For

$$A = \begin{pmatrix} 1 & 0 & 0 & -7 & 8 \\ 0 & 1 & 0 & 3 & 2 \\ 0 & 0 & 1 & 1 & -5 \end{pmatrix},$$

the homogeneous system has free variables $x_4 = s$ and $x_5 = t$. The solution set is

$$(x_1, x_2, x_3, x_4, x_5) = s(7, -3, -1, 1, 0) + t(-8, -2, 5, 0, 1).$$

9.5 Matrix Arithmetic

Definition 11 (Matrix addition and scalar multiplication)
Matrices of the same size are added entrywise:

$$(A + B)_{ij} = a_{ij} + b_{ij}.$$

Scalars multiply matrices entrywise:

$$(cA)_{ij} = ca_{ij}.$$

Definition 12 (Matrix multiplication)
If A is $m \times n$ and B is $n \times p$, then AB is the $m \times p$ matrix with entries

$$(AB)_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj}.$$

Note
Matrix multiplication is defined only when the inner dimensions match. It is associative and distributive, but generally not commutative: usually $AB \neq BA$.

Example 6
Let

$$D = \begin{pmatrix} x & 0 \\ 0 & y \end{pmatrix}, \quad E = \begin{pmatrix} a & b \\ c & d \end{pmatrix}.$$

Then

$$DE = \begin{pmatrix} xa & xb \\ yc & yd \end{pmatrix}, \quad ED = \begin{pmatrix} ax & by \\ cx & dy \end{pmatrix}.$$

These are equal for every E exactly when $x = y$. Otherwise D commutes only with matrices satisfying $(x - y)b = 0$ and $(y - x)c = 0$.

9.6 Matrix Transformations

Matrices and linear transformations are two ways of describing the same finite-dimensional behaviour. In this course we mostly use the standard bases of $\mathbb{R}^n$ and $\mathbb{R}^m$, so a linear map can be recorded by the matrix whose columns are its outputs on the standard basis vectors.

Definition 13 (Standard basis)
The **standard basis** of $\mathbb{R}^n$ is $e_1, \ldots, e_n$, where e_j has a 1 in the j-th position and 0 elsewhere.

Definition 14 (Matrix transformation)
If A is an $m \times n$ matrix, it defines a function

$$T_A : \mathbb{R}^n \to \mathbb{R}^m, \quad T_A(x) = Ax.$$

This function is linear:

$$T_A(cx + dy) = cT_A(x) + dT_A(y).$$

Proposition 5 (Matrices represent linear transformations)
If $T : \mathbb{R}^n \to \mathbb{R}^m$ is linear, then

$$T(x) = Ax$$

where the columns of A are $T(e_1), \ldots, T(e_n)$.

Proof. Write $x = x_1 e_1 + \ldots + x_n e_n$. By linearity,

$$T(x) = x_1 T(e_1) + \ldots + x_n T(e_n),$$

which is exactly the matrix product Ax when the columns of A are $T(e_1), \ldots, T(e_n)$. □

Note
For a general finite-dimensional vector space, one first chooses a basis and then records coordinates relative to that basis. The standard-basis case above is the version needed for $\mathbb{R}^n$.

Proposition 6 (Composition corresponds to multiplication)
If A is $m \times n$ and B is $n \times p$, then

$$T_A \circ T_B = T_{AB}.$$

Proof. For $x \in \mathbb{R}^p$,

$$(T_A \circ T_B)(x) = T_A(Bx) = A(Bx) = (AB)x.$$

□

9.7 Inverse Matrices

Definition 15 (Identity matrix)
The $n \times n$ **identity matrix** is

$$I_n = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 1 \end{pmatrix}.$$

It satisfies $I_n A = A$ and $AI_n = A$ whenever the products are defined.

Definition 16 (Inverse matrix)
A square matrix A is **invertible** if there is a matrix B such that

$$AB = BA = I.$$

In that case B is the **inverse** of A and is written A^{-1}.

Proposition 7 (Uniqueness of inverses)
A square matrix has at most one inverse.

Proof. If B and C are both inverses of A, then

$$B = BI = B(AC) = (BA)C = IC = C.$$

□

Proposition 8 (One-sided inverses)
Let A, B, C be $n \times n$ matrices. If

$$AB = CA = I_n,$$

then $B = C$. Hence $B = C = A^{-1}$.

Proof. We have

$$B = I_n B = (CA)B = C(AB) = CI_n = C.$$

Thus the same matrix is both a left and right inverse for A. □

Proposition 9 (Inverse of a product)
If A and B are invertible $n \times n$ matrices, then AB is invertible and

$$(AB)^{-1} = B^{-1}A^{-1}.$$

Also, $\left(A^{-1}\right)^{-1} = A$.

Proof. Since

$$(AB)(B^{-1}A^{-1}) = A(BB^{-1})A^{-1} = AA^{-1} = I_n$$

and similarly

$$(B^{-1}A^{-1})(AB) = I_n,$$

the inverse of AB is $B^{-1}A^{-1}$. The statement $\left(A^{-1}\right)^{-1} = A$ follows directly from the definition of inverse. ☐

Proposition 10 (Finding inverses by row reduction)
A square matrix A is invertible exactly when $[A \mid I]$ row-reduces to $[I \mid B]$. In that case $B = A^{-1}$.

Proof. Row operations on $[A \mid I]$ are left multiplication by elementary matrices. A sequence of row operations is therefore left multiplication by some invertible matrix E.

If the reduction gives $[I \mid B]$, then $EA = I$ and $EI = B$. Thus $B = E$ and $BA = I$. Since E is invertible, $A = E^{-1}$, so $B = A^{-1}$.

Conversely, if A is invertible, row reduction cannot get stuck before a pivot appears in every column; otherwise there would be a non-zero vector x with $Ax = 0$, contradicting $x = A^{-1}Ax = 0$. Hence A row-reduces to I, and applying the same row operations to the right half gives A^{-1}. ☐

Example 7
Let

$$A = \begin{pmatrix} 0 & 1 & 2 \\ 1 & 0 & 3 \\ 0 & 0 & 1 \end{pmatrix}.$$

Row-reducing $[A \mid I]$ gives

$$\begin{pmatrix} 0 & 1 & 2 & 1 & 0 & 0 \\ 1 & 0 & 3 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 \end{pmatrix} \sim \begin{pmatrix} 1 & 0 & 0 & 0 & 1 & -3 \\ 0 & 1 & 0 & 1 & 0 & -2 \\ 0 & 0 & 1 & 0 & 0 & 1 \end{pmatrix}.$$

Hence

$$A^{-1} = \begin{pmatrix} 0 & 1 & -3 \\ 1 & 0 & -2 \\ 0 & 0 & 1 \end{pmatrix}.$$

Theorem 11 (Invertibility criteria)
For an $n \times n$ matrix A, the following statements are equivalent.

- A is invertible.
- A has a right inverse: there exists a matrix B such that $AB = I_n$.
- $\operatorname{rank}(A) = n$.
- $\ker A = \{0\}$.

Proof. If A is invertible and $x \in \ker A$, then

$$x = A^{-1}Ax = A^{-1}0 = 0,$$

so $\ker A = \{0\}$.

If $\ker A = \{0\}$, then the homogeneous system has no free variable. Thus every column has a pivot, so $\operatorname{rank}(A) = n$.

If $\operatorname{rank}(A) = n$, then $Ax = b$ has a unique solution for every $b \in \mathbb{R}^n$. In particular, for each standard basis vector e_i, choose the unique vector x^i with $Ax^i = e_i$. Putting these vectors as the columns of a matrix B gives $AB = I_n$, so A has a right inverse.

Finally, if $AB = I_n$, then $\ker B = \{0\}$ because $Bx = 0$ implies $x = ABx = 0$. By the previous implications, B has a right inverse C. Since $AB = I_n$ and $BC = I_n$, the one-sided inverse result gives $A = C$, so A is the inverse of B and hence is invertible. □

Proposition 12 (Transpose and inverse)
If A is invertible, then A^T is invertible and

$$\left(A^T\right)^{-1} = \left(A^{-1}\right)^T.$$

Proof. Since $AA^{-1} = I$ and $A^{-1}A = I$, taking transposes gives

$$\left(A^{-1}\right)^T A^T = I, \quad A^T\left(A^{-1}\right)^T = I.$$

□

9.8 Euclidean Vectors

Note
Found in the 2026 tutorials. This section collects the vector-geometry tools used in tutorial questions.

Definition 17 (Dot product and norm)
For vectors

$$u = \begin{pmatrix} u_1 \\ \dots \\ u_n \end{pmatrix} \quad \text{and} \quad v = \begin{pmatrix} v_1 \\ \dots \\ v_n \end{pmatrix}$$

in $\mathbb{R}^n$, the **dot product** is

$$u \cdot v = \sum_{i=1}^{n} u_i v_i.$$

The **norm** or **length** of u is

$$\|u\| = \sqrt{u \cdot u}.$$

Definition 18 (Orthogonality and angle)
Two vectors u and v are **orthogonal** if $u \cdot v = 0$. If u and v are non-zero, the angle θ between them is determined by

$$\cos\theta = \frac{u \cdot v}{\|u\|\|v\|}.$$

Theorem 13 (Cauchy-Schwarz inequality)
For all $u, v \in \mathbb{R}^n$,

$$|u \cdot v| \leq \|u\|\|v\|.$$

Proof. If $v = 0$, then both sides are 0. Otherwise, for every $t \in \mathbb{R}$,

$$0 \leq \|u - tv\|^2 = \|u\|^2 - 2t(u \cdot v) + t^2\|v\|^2.$$

Choose $t = \frac{u \cdot v}{\|v\|^2}$. Substitution gives

$$0 \leq \|u\|^2 - \frac{(u \cdot v)^2}{\|v\|^2}.$$

Multiplying by $\|v\|^2$ gives $(u \cdot v)^2 \leq \|u\|^2\|v\|^2$, and taking square roots gives the result. □

Theorem 14 (Vector triangle inequality)
For all $u, v \in \mathbb{R}^n$,

$$\|u + v\| \leq \|u\| + \|v\|.$$

Proof. Squaring the left-hand side gives

$$\|u + v\|^2 = \|u\|^2 + 2u \cdot v + \|v\|^2.$$

By Cauchy-Schwarz,

$$2u \cdot v \leq 2|u \cdot v| \leq 2\|u\|\|v\|.$$

Hence

$$\|u+v\|^2 \leq (\|u\|+\|v\|)^2.$$

Both sides are non-negative, so taking square roots proves the result. □

Proposition 15 (Parallelogram law)
For all $u, v \in \mathbb{R}^n$,

$$\|u+v\|^2 + \|u-v\|^2 = 2\|u\|^2 + 2\|v\|^2.$$

Proof. Expand both squared norms using the dot product:

$$\|u+v\|^2 = \|u\|^2 + 2u \cdot v + \|v\|^2,$$

$$\|u-v\|^2 = \|u\|^2 - 2u \cdot v + \|v\|^2.$$

Adding cancels the dot-product terms. □

Corollary 16 (Equal lengths from orthogonal diagonals)
If $u+v$ and $u-v$ are orthogonal, then $\|u\| = \|v\|$.

Proof. Orthogonality gives

$$0 = (u+v) \cdot (u-v) = \|u\|^2 - \|v\|^2.$$

Hence $\|u\|^2 = \|v\|^2$, and norms are non-negative. □

Definition 19 (Projection)
If $a \neq 0$, the projection of b onto a is

$$\operatorname{proj}_a b = \frac{b \cdot a}{a \cdot a} a.$$

The component of b orthogonal to a is

$$\operatorname{orth}_a b = b - \operatorname{proj}_a b.$$

Proposition 17 (Projection residual is orthogonal)
If $a \neq 0$, then $\operatorname{orth}_a b$ is orthogonal to a.

Proof. Compute

$$a \cdot (b - \operatorname{proj}_a b) = a \cdot b - a \cdot \left(\frac{b \cdot a}{a \cdot a} a \right) = a \cdot b - (b \cdot a) = 0.$$

□

Definition 20 (Cross product in three dimensions)
For vectors

$$u = \begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix} \quad \text{and} \quad v = \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix},$$

the **cross product** is

$$u \times v = \begin{pmatrix} u_2v_3 - u_3v_2 \\ u_3v_1 - u_1v_3 \\ u_1v_2 - u_2v_1 \end{pmatrix}.$$

It is orthogonal to both u and v.

Example 8
Let

$$A = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}, \quad B = \begin{pmatrix} 7 \\ 2 \\ 1 \end{pmatrix}, \quad C = \begin{pmatrix} 1 \\ 3 \\ 3 \end{pmatrix}.$$

Then

$$B - A = \begin{pmatrix} 6 \\ 0 \\ -2 \end{pmatrix}, \quad C - A = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}.$$

Since $(B-A)\cdot(C-A) = 0$, the angle at A is $\frac{\pi}{2}$.

Example 9
A vector orthogonal to the plane through A, B, and C is

$$(B - A) \times (C - A) = \begin{pmatrix} 6 \\ 0 \\ -2 \end{pmatrix} \times \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 2 \\ 0 \\ 6 \end{pmatrix}.$$

The length of the projection of $B - A$ onto $C - A$ is

$$\frac{|(B-A)\cdot(C-A)|}{\|C - A\|} = 0.$$

Example 10
Since

$$C - B = \begin{pmatrix} -6 \\ 1 \\ 2 \end{pmatrix},$$

a unit vector parallel to $C - B$ is

$$\frac{1}{\sqrt{41}}\begin{pmatrix}-6\\1\\2\end{pmatrix}.$$

A vector orthogonal to both $C - B$ and the z-axis is

$$\begin{pmatrix}-6\\1\\2\end{pmatrix} \times \begin{pmatrix}0\\0\\1\end{pmatrix} = \begin{pmatrix}1\\6\\0\end{pmatrix},$$

so a unit vector with this property is

$$\frac{1}{\sqrt{37}}\begin{pmatrix}1\\6\\0\end{pmatrix}.$$

10 Determinants, Eigenvalues, and Eigenvectors

Determinants measure signed volume-scaling for square matrices, so they give a scalar test for invertibility. Eigenvalues and eigenvectors record directions on which a matrix acts by simple scaling. These ideas are tightly connected: eigenvalues are found by applying a determinant to $A - \lambda I$.

10.1 Determinants

If A is an $n \times n$ matrix, the columns of A are the images of the standard basis vectors under the matrix transformation T_A. These vectors span a fundamental parallelepiped. Geometrically,

$$\mathrm{Vol}(\mathrm{FP}_A) = |\det A|.$$

If this volume is zero, the transformation has collapsed space into a lower-dimensional object, so it cannot be invertible.

Definition 1 (Determinant)
The **determinant** assigns a scalar $\det A$ to each square matrix A. It measures the signed version of the volume-scaling described above; below we prove that a square matrix is invertible exactly when its determinant is non-zero.

The determinant is defined recursively. If $A = (a)$ is 1×1, then $\det A = a$. If A is $n \times n$ with $n \geq 2$, delete row 1 and column j to obtain the $(n-1) \times (n-1)$ matrix A_{1j}. Then

$$\det A = \sum_{j=1}^{n} (-1)^{1+j} a_{1j} \det A_{1j}.$$

For

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix},$$

the determinant is

$$\det A = ad - bc.$$

Definition 2 (Minors and cofactors)
Let A be an $n \times n$ matrix. Write A_{ij} for the matrix obtained from A by deleting row i and column j. The (i,j)-**minor** is

$$M_{ij} = \det A_{ij}.$$

The (i,j)-**cofactor** is

$$C_{ij} = (-1)^{i+j} M_{ij}.$$

Theorem 1 (Cofactor expansion)
The determinant of an $n \times n$ matrix can be expanded along any row or any column:

$$\det A = \sum_{j=1}^{n} a_{ij} C_{ij}$$

for a fixed row i, and

$$\det A = \sum_{i=1}^{n} a_{ij} C_{ij}$$

for a fixed column j.

Note
Cofactor expansion is usually best when a row or column has many zeroes. For larger matrices, row operations are usually faster.

Example 1
Expanding a 3×3 determinant along the first row gives

$$\det \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix} = a \det \begin{pmatrix} e & f \\ h & i \end{pmatrix} - b \det \begin{pmatrix} d & f \\ g & i \end{pmatrix} + c \det \begin{pmatrix} d & e \\ g & h \end{pmatrix}.$$

Hence

$$\det \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix} = a(ei - fh) - b(di - fg) + c(dh - eg).$$

Proposition 2 (Determinants and row operations)
Let A be a square matrix.

- Swapping two rows multiplies the determinant by -1.
- Multiplying one row by c multiplies the determinant by c.
- Adding a multiple of one row to another row does not change the determinant.

Note
These rules are the computational backbone of determinants in this course. The recursive definition is useful for explaining what the determinant is; row operations are usually better for computing large determinants.

Proposition 3 (Triangular determinant)
If A is upper triangular or lower triangular, then $\det A$ is the product of the diagonal entries.

Example 2
For

$$A = \begin{pmatrix} x & 1 & 2 \\ 1 & x & 3 \\ 0 & 0 & 1 \end{pmatrix},$$

cofactor expansion along the last row gives

$$\det A = \det \begin{pmatrix} x & 1 \\ 1 & x \end{pmatrix} = x^2 - 1.$$

Hence A is non-invertible exactly when $x = \pm 1$.

10.2 Determinant Laws

Theorem 4 (Basic determinant laws)
If A and B are $n \times n$ matrices, then

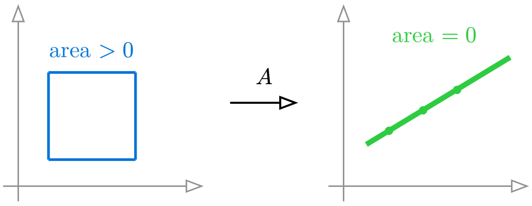
$$\det(AB) = \det(A)\det(B), \quad \det(A^T) = \det(A).$$

If A is invertible, then

$$\det(A^{-1}) = \frac{1}{\det(A)}.$$

Theorem 5 (Invertibility and determinant)
A square matrix A is invertible if and only if

$$\det A \neq 0.$$



When $\det A = 0$*, the associated linear map collapses area: two-dimensional input can land on a line, so the map cannot be invertible.*

Example 3
Suppose $\det(AB^2) = 4$ and $\det(A^2B) = 2$. Let $x = \det A$ and $y = \det B$. Then

$$xy^2 = 4, \quad x^2y = 2.$$

Dividing gives $\frac{y}{x} = 2$, so $y = 2x$. Substituting into $x^2y = 2$ gives $2x^3 = 2$, hence

$$\det A = 1, \quad \det B = 2.$$

Example 4
If $A^2 = A$, then

$$\det(A)^2 = \det(A).$$

Hence $\det(A)(\det(A) - 1) = 0$, so $\det(A) = 0$ or $\det(A) = 1$.

Note
The orthogonal-matrix facts in this subsection are from the 2026 tutorials.

Definition 3 (Orthogonal matrix)
A square matrix Q is **orthogonal** if

$$Q^{-1} = Q^T.$$

Proposition 6 (Determinant of an orthogonal matrix)
If Q is orthogonal, then $\det Q = \pm 1$.

Proof. Since $Q^T Q = I$,

$$\det(Q^T Q) = \det I.$$

Using the determinant laws, this gives $(\det Q)^2 = 1$. ☐

Proposition 7 (Products of orthogonal matrices)
If X and Y are orthogonal matrices of the same size, then XY is orthogonal.

Proof. Since $X^{-1} = X^T$ and $Y^{-1} = Y^T$,

$$(XY)^{-1} = Y^{-1}X^{-1} = Y^T X^T = (XY)^T.$$

☐

Proposition 8 (Orthogonal matrices preserve dot products)
If Q is orthogonal, then

$$(Qx) \cdot (Qy) = x \cdot y$$

for all $x, y \in \mathbb{R}^n$. Hence orthogonal matrices preserve lengths and angles.

Proof. Since $Q^T Q = I$,

$$(Qx) \cdot (Qy) = x^T Q^T Q y = x^T y = x \cdot y.$$

☐

Note
The trace, symmetry, and nilpotent-matrix facts below are from the 2026 tutorials.

Definition 4 (Trace)
The **trace** of a square matrix is the sum of its diagonal entries:

$$\operatorname{tr}(A) = a_{11} + \dots + a_{nn}.$$

Definition 5 (Symmetric and skew-symmetric matrices)
A square matrix A is **symmetric** if $A^T = A$, and **skew-symmetric** if $A^T = -A$.

Proposition 9 (Useful transpose facts)
For any square matrix A:

- $A + A^T$ is symmetric;
- if A is skew-symmetric, then $\operatorname{tr}(A) = 0$;
- if A is skew-symmetric and has odd size, then $\det A = 0$.

Proof. First,

$$\left(A + A^T\right)^T = A^T + A = A + A^T.$$

If A is skew-symmetric, then every diagonal entry satisfies $a_{ii} = -a_{ii}$, so $a_{ii} = 0$ and $\operatorname{tr}(A) = 0$. If A is $n \times n$ and n is odd, then

$$\det A = \det\left(A^T\right) = \det(-A) = (-1)^n \det A = -\det A,$$

so $\det A = 0$. ☐

Definition 6 (Nilpotent matrix)
A square matrix A is **nilpotent** if $A^k = 0$ for some $k \in \mathbb{N}$.

Proposition 10 (Nilpotent matrices are singular)
If A is nilpotent, then $\det A = 0$.

Proof. If $A^k = 0$, then

$$(\det A)^k = \det\left(A^k\right) = \det 0 = 0.$$

Hence $\det A = 0$. ☐

Proposition 11 (Vandermonde determinant)

Note
Starred result from the 2026 tutorials.

For

$$V_n = \begin{pmatrix} 1 & x_1 & x_1^2 & \dots & x_1^{n-1} \\ 1 & x_2 & x_2^2 & \dots & x_2^{n-1} \\ \dots & \dots & \dots & \dots & \dots \\ 1 & x_n & x_n^2 & \dots & x_n^{n-1} \end{pmatrix},$$

we have

$$\det V_n = \prod_{1 \le i < j \le n} \left(x_j - x_i\right).$$

Proof. The determinant is a polynomial in x_n of degree at most $n - 1$. If $x_n = x_i$ for some $i < n$, then two rows are equal, so the determinant is 0. Hence

$$\prod_{i=1}^{n-1} (x_n - x_i)$$

divides $\det V_n$ as a polynomial in x_n. Comparing the coefficient of x_n^{n-1} gives the determinant of V_{n-1}. Induction gives the formula. ☐

10.3 Characteristic Polynomials

Definition 7 (Eigenvalue and eigenvector)
Let A be a square matrix. A scalar λ is an **eigenvalue** of A if there is a non-zero vector v such that

$$Av = \lambda v.$$

Such a vector v is an **eigenvector** for λ.

Definition 8 (Eigenspace)
The **eigenspace** for an eigenvalue λ is

$$E_\lambda = \ker(A - \lambda I) = \{v : (A - \lambda I)v = 0\}.$$

The zero vector belongs to the eigenspace, but is not called an eigenvector.

Theorem 12 (Characteristic equation)
A scalar λ is an eigenvalue of A if and only if

$$\det(A - \lambda I) = 0.$$

The polynomial $\det(A - \lambda I)$ is the **characteristic polynomial** of A.

Proof. The equation $Av = \lambda v$ is equivalent to $(A - \lambda I)v = 0$. This has a non-zero solution exactly when $A - \lambda I$ is non-invertible, which is equivalent to $\det(A - \lambda I) = 0$ by invertibility and determinant. □

Example 5
For

$$A = \begin{pmatrix} 5 & 0 & 0 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{pmatrix},$$

we have

$$\det(A - \lambda I) = \det \begin{pmatrix} 5-\lambda & 0 & 0 \\ 1 & 2-\lambda & 1 \\ 1 & 1 & 2-\lambda \end{pmatrix}$$

$$= (5-\lambda)((2-\lambda)^2 - 1) = (5-\lambda)(3-\lambda)(1-\lambda).$$

Hence the eigenvalues are 5, 3, and 1.

10.4 Finding Eigenvectors

Proposition 13 (How to find eigenvectors)
Once λ is known, solve the homogeneous system

$$(A - \lambda I)v = 0.$$

The non-zero solutions are the eigenvectors for λ.

Example 6
Continuing the previous example, row reduction gives:

$$E_5 = \operatorname{span}\left\{\begin{pmatrix}2\\1\\1\end{pmatrix}\right\},\quad E_3 = \operatorname{span}\left\{\begin{pmatrix}0\\1\\1\end{pmatrix}\right\},\quad E_1 = \operatorname{span}\left\{\begin{pmatrix}0\\-1\\1\end{pmatrix}\right\}.$$

Thus each eigenspace has dimension 1.

Example 7
Let

$$A = \begin{pmatrix}a & 2 & 1\\ -1 & 2 & 1\\ 2 & -2 & -1\end{pmatrix}.$$

Then

$$\det(A-\lambda I) = (a-\lambda)\lambda(\lambda-1).$$

The eigenvalues are therefore among 0, 1, and a. Corresponding eigenvectors are

$$\lambda = 0 : \begin{pmatrix}0\\1\\-2\end{pmatrix},\quad \lambda = 1 : \begin{pmatrix}-1\\a\\-(a+1)\end{pmatrix},\quad \lambda = a : \begin{pmatrix}-a\\1\\-2\end{pmatrix}.$$

If $a \notin \{0,1\}$, these give three distinct one-dimensional eigenspaces. If $a = 0$ or $a = 1$, there are two distinct eigenvalues, each with a one-dimensional eigenspace.

10.5 Useful Eigenvalue Facts

Proposition 14 (Shifted eigenvalues)
If λ is an eigenvalue of A with eigenvector v, then $\lambda + c$ is an eigenvalue of $A + cI$ with the same eigenvector.

Proof. Since $Av = \lambda v$,

$$(A + cI)v = Av + cv = (\lambda + c)v.$$

☐

Proposition 15 (Powers and inverses)
Suppose $Av = \lambda v$ with $v \neq 0$.

- For $n \in \mathbb{N}$, $A^n v = \lambda^n v$.
- If A is invertible, then $\lambda \neq 0$ and $A^{-1}v = \lambda^{-1}v$.

Proof. The power statement follows by induction. If A is invertible, then

$$v = A^{-1}Av = A^{-1}\lambda v,$$

so $A^{-1}v = \lambda^{-1}v$. ☐

Proposition 16 (Transpose has the same eigenvalues)
A square matrix A and its transpose A^T have the same eigenvalues.

Proof. Their characteristic polynomials are equal:

$$\det(A^T - \lambda I) = \det((A - \lambda I)^T) = \det(A - \lambda I).$$

Hence they have the same roots. □

Proposition 17 (Eigenvectors for distinct eigenvalues)
Eigenvectors corresponding to distinct eigenvalues are linearly independent.

Proof. For two eigenvectors, suppose $c_1 v_1 + c_2 v_2 = 0$, where

$$Av_1 = \lambda_1 v_1, \quad Av_2 = \lambda_2 v_2, \quad \lambda_1 \neq \lambda_2.$$

Applying A gives $c_1 \lambda_1 v_1 + c_2 \lambda_2 v_2 = 0$. Subtracting $\lambda_2(c_1 v_1 + c_2 v_2) = 0$ gives

$$c_1(\lambda_1 - \lambda_2)v_1 = 0.$$

Since $v_1 \neq 0$ and $\lambda_1 \neq \lambda_2$, $c_1 = 0$. Then $c_2 = 0$.

The same idea proves the general finite case by induction: apply A to a dependence relation, subtract one eigenvalue times the original relation, and use the induction hypothesis on the remaining eigenvectors. □

Proposition 18 (Positive quadratic forms and eigenvalues)
Let A be a symmetric 2×2 matrix. If

$$x^T A x > 0$$

for every non-zero $x \in \mathbb{R}^2$, then every eigenvalue of A is positive.

Proof. If $Av = \lambda v$ with $v \neq 0$, then

$$v^T A v = v^{T(\lambda v)} = \lambda(v^T v) = \lambda \|v\|^2.$$

The left-hand side is positive and $\|v\|^2 > 0$, so $\lambda > 0$. □

11 Vector Spaces

Note

This chapter is mostly based on 2026 tutorial material and the advertised Week 13 topics. Treat it as exam-prep and tutorial extension material unless a result is also stated in lecture.

The same linear-algebra ideas apply to vectors in $\mathbb{R}^n$, matrices, polynomials, and solution sets of homogeneous systems. The shared language is vector spaces, subspaces, span, independence, basis, and dimension.

11.1 Vector Spaces and Subspaces

Definition 1 (Vector space)

A **vector space over** $\mathbb{R}$ is a set V whose elements can be added and multiplied by real scalars, satisfying the usual rules: commutativity and associativity of addition, a zero vector, additive inverses, distributivity, and scalar associativity.

Example 1

The following are vector spaces over $\mathbb{R}$:

- $\mathbb{R}^n$ with usual vector addition and scalar multiplication,
- $M_{m\times n}(\mathbb{R})$, the set of $m \times n$ real matrices,
- $P_{n(\mathbb{R})}$, the set of real polynomials of degree at most n.

Definition 2 (Subspace)

A subset $U \subseteq V$ is a **subspace** of V if:

- $0 \in U$,
- whenever $u, v \in U$, then $u + v \in U$,
- whenever $u \in U$ and $c \in \mathbb{R}$, then $cu \in U$.

Proposition 1 (Subspace test)

A non-empty subset $U \subseteq V$ is a subspace if and only if

$$cu + dv \in U$$

for all $u, v \in U$ and all $c, d \in \mathbb{R}$.

Example 2

Note

Found in the 2026 tutorials.

The set of $n \times n$ upper triangular matrices is a subspace of $M_{n\times n}(\mathbb{R})$: the zero matrix is upper triangular, sums of upper triangular matrices are upper triangular, and scalar multiples are upper triangular.

Example 3

Note
Found in the 2026 tutorials.

The set of $n \times n$ orthogonal matrices is not a subspace. It contains I, but $2I$ is not orthogonal because

$$(2I)^{T(2I)} = 4I \neq I.$$

Example 4
The set

$$\left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3 : x + 2yz = 6 \right\}$$

is not a subspace. It does not contain 0, and the condition is not linear because of the product yz.

11.2 Nullspaces

Definition 3 (Nullspace)
The **nullspace** of a matrix A is

$$\ker A = \{x : Ax = 0\}.$$

It is also called the **kernel** of A.

Proposition 2 (Nullspaces are subspaces)
For every matrix A, the nullspace $\ker A$ is a subspace.

Proof. Since $A0 = 0$, we have $0 \in \ker A$. If $u, v \in \ker A$ and $c, d \in \mathbb{R}$, then

$$A(cu + dv) = cAu + dAv = 0.$$

Hence $cu + dv \in \ker A$. □

Proposition 3 (Nullspace inclusion)

Note
Found in the 2026 tutorials.

If A and B are square matrices, then

$$\ker B \subseteq \ker(AB).$$

Proof. If $x \in \ker B$, then $Bx = 0$. Therefore

$$ABx = A0 = 0,$$

so $x \in \ker(AB)$. □

11.3 Span

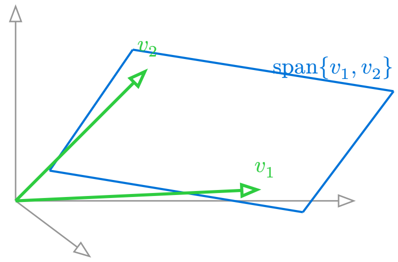
Definition 4 (Linear combination)
A **linear combination** of vectors $v_1, ..., v_k$ is a vector of the form

$$c_1 v_1 + ... + c_k v_k,$$

where $c_1, ..., c_k \in \mathbb{R}$.

Definition 5 (Span)
The **span** of vectors $v_1, ..., v_k$ is the set of all their linear combinations:

$$\operatorname{span}\{v_1, ..., v_k\} = \{c_1 v_1 + ... + c_k v_k : c_1, ..., c_k \in \mathbb{R}\}.$$



Two non-parallel vectors in $\mathbb{R}^3$ span the plane through the origin containing them.

Proposition 4 (Spans are subspaces)
For any vectors $v_1, ..., v_k \in V$, the set $\operatorname{span}\{v_1, ..., v_k\}$ is a subspace of V.

Example 5
The set

$$\left\{ \begin{pmatrix} 5t \\ -3t \\ t \\ t \end{pmatrix} \in \mathbb{R}^4 : t \in \mathbb{R} \right\}$$

is

$$\operatorname{span}\left\{ \begin{pmatrix} 5 \\ -3 \\ 1 \\ 1 \end{pmatrix} \right\}.$$

It is a one-dimensional subspace of $\mathbb{R}^4$.

11.4 Linear Independence

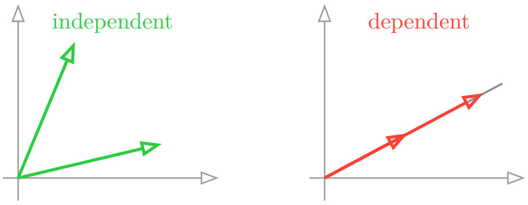
Definition 6 (Linear independence)
Vectors $v_1, ..., v_k$ are **linearly independent** if the equation

$$c_1 v_1 + ... + c_k v_k = 0$$

implies

$$c_1 = \ldots = c_k = 0.$$

If there is a non-trivial solution, the vectors are **linearly dependent**.



In $\mathbb{R}^2$, two vectors are linearly independent exactly when neither lies on the line spanned by the other.

Proposition 5 (Subsets and supersets)
- Every subset of a linearly independent set is linearly independent.
- Every superset of a linearly dependent set is linearly dependent.

Proof. For the first statement, a non-trivial dependence relation among a subset would also be a non-trivial dependence relation among the original set by giving the missing vectors coefficient 0. For the second statement, keep the existing non-trivial dependence relation and again give every added vector coefficient 0. ☐

Example 6
The matrices

$$A = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad B = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad C = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$$

are linearly independent in $M_{2\times 2}(\mathbb{R})$. If

$$c_1 A + c_2 B + c_3 C = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix},$$

then the left-hand side is

$$\begin{pmatrix} c_1 & c_2 \\ c_3 & 0 \end{pmatrix},$$

so comparing entries gives $c_1 = c_2 = c_3 = 0$.

Proposition 6 (Eigenvectors with distinct eigenvalues)

Note
Found in the 2026 tutorials.

Eigenvectors corresponding to distinct eigenvalues are linearly independent.

Note
This is the vector-space version of the eigenvector independence result from the previous chapter.

11.5 Bases and Dimension

Definition 7 (Basis)
A list of vectors $(v_1, \ldots, v_k)$ is a **basis** for a vector space V if:

- the vectors span V,
- the vectors are linearly independent.

Definition 8 (Dimension)
If V has a finite basis, then the number of vectors in any basis of V is called the **dimension** of V, denoted $\dim(V)$.

Example 7
The set

$$U = \left\{ \begin{pmatrix} a \\ b \\ c \\ d \\ e \end{pmatrix} \in \mathbb{R}^5 : a - c - d = 0 \right\}$$

is a subspace. Since $a = c + d$,

$$\begin{pmatrix} a \\ b \\ c \\ d \\ e \end{pmatrix} = b \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} + c \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} + d \begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} + e \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}.$$

These four vectors form a basis, so $\dim(U) = 4$.

Example 8

Note
Found in the 2026 tutorials.

Let

$$V = \{p(x) \in P_7(\mathbb{R}) : p(1) = 0\}.$$

This is a vector space: if $p(1) = q(1) = 0$, then $(cp + dq)(1) = 0$. A clean basis is

$$x - 1, \quad x(x - 1), \quad x^2(x - 1), \quad \ldots, \quad x^6(x - 1).$$

Hence $\dim(V) = 7$.

Example 9

Note
Found in the 2026 tutorials.

The four vectors

$$\begin{pmatrix} x \\ 1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ x \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ x \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 1 \\ x \end{pmatrix}$$

form the columns of the matrix

$$A = \begin{pmatrix} x & 1 & 1 & 1 \\ 1 & x & 1 & 1 \\ 1 & 1 & x & 1 \\ 1 & 1 & 1 & x \end{pmatrix} = (x-1)I + J,$$

where J is the all-ones matrix. The vector

$$\begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}$$

is scaled by $x+3$, while every vector whose entries sum to 0 is scaled by $x-1$. Hence

$$\det A = (x+3)(x-1)^3.$$

The vectors fail to form a basis exactly when $x = 1$ or $x = -3$. If $x = 1$, all four vectors are equal, so the span has dimension 1. If $x = -3$, the span has dimension 3.

11.6 Sums and Intersections of Subspaces

Note
Found in the 2026 tutorials.

Definition 9 (Sum of subspaces)
If U and V are subspaces of a vector space W, their sum is

$$U + V = \{u + v : u \in U, v \in V\}.$$

Proposition 7 (Sums and intersections)
If U and V are subspaces of W, then $U + V$ and $U \cap V$ are subspaces of W.

Proof. For $U + V$, use the subspace test:

$$c(u_1 + v_1) + d(u_2 + v_2) = (cu_1 + du_2) + (cv_1 + dv_2) \in U + V.$$

For $U \cap V$, if x, y lie in both spaces, then $cx + dy$ lies in both spaces. □

Proposition 8 (Union warning)
If neither U nor V is contained in the other, then $U \cup V$ is not a subspace.

Proof. Choose $u \in U$ with $u \notin V$ and $v \in V$ with $v \notin U$. If $u + v \in U$, then $v = (u + v) - u \in U$, a contradiction. If $u + v \in V$, then $u = (u + v) - v \in V$, a contradiction. Thus $U \cup V$ is not closed under addition. □

Theorem 9 (Dimension formula)
If U and V are finite-dimensional subspaces, then

$$\dim(U + V) = \dim(U) + \dim(V) - \dim(U \cap V).$$

Example 10
If U and V are two-dimensional subspaces of $\mathbb{R}^4$, then $\dim(U \cap V)$ can be 0, 1, or 2. By the dimension formula,

$$\dim(U + V) = 4 - \dim(U \cap V),$$

so $\dim(U + V)$ can be 4, 3, or 2 respectively.

Example 11
If U and V are distinct three-dimensional subspaces of $\mathbb{R}^4$, then $\dim(U \cap V) = 2$. Indeed, $U + V \subseteq \mathbb{R}^4$, so

$$\dim(U \cap V) = 6 - \dim(U + V) \geq 2.$$

Since $U \neq V$, the intersection cannot have dimension 3. Hence it has dimension 2.

11.7 Matrix Subspaces

Note
Found in the 2026 tutorials.

Example 12
Let

$$W = \{X \in M_{2\times 2}(\mathbb{R}) : X = X^T, \operatorname{tr}(X) = 0\}.$$

A matrix in W has the form

$$\begin{pmatrix} a & b \\ b & -a \end{pmatrix} = a\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} + b\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

Hence a basis is

$$\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix},$$

and $\dim(W) = 2$.

Example 13
Let

$$W' = \left\{ \begin{pmatrix} 0 & a \\ 0 & b \end{pmatrix} : a, b \in \mathbb{R} \right\}.$$

For the subspace W above,

$$W \cap W' = \left\{ \begin{pmatrix} 0 & a \\ 0 & b \end{pmatrix} : \begin{pmatrix} 0 & a \\ 0 & b \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ a & b \end{pmatrix} \text{ and } b = 0 \right\} = \{0\}.$$

Also every matrix

$$\begin{pmatrix} p & q \\ r & s \end{pmatrix}$$

can be written as a sum of an element of W and an element of W':

$$\begin{pmatrix} p & q \\ r & s \end{pmatrix} = \begin{pmatrix} p & r \\ r & -p \end{pmatrix} + \begin{pmatrix} 0 & q-r \\ 0 & s+p \end{pmatrix}.$$

Hence $M_{2\times 2}(\mathbb{R}) = W + W'$.

Example 14

The set of 2×2 real matrices A satisfying $A^2 = 0$ is not a subspace. For example,

$$A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad B = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$$

both satisfy $A^2 = B^2 = 0$, but

$$(A+B)^2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \neq 0.$$

11.8 Linear Maps

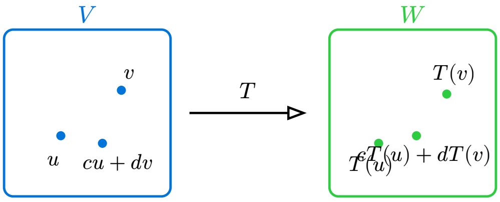
Chapter 9 described linear maps $\mathbb{R}^n \to \mathbb{R}^m$ using matrices in the standard basis. The same definition works for any vector spaces; matrices are what appear after bases have been chosen.

Definition 10 (Linear map)

A function $T : V \to W$ between vector spaces is **linear** if

$$T(cu + dv) = cT(u) + dT(v)$$

for all $u, v \in V$ and $c, d \in \mathbb{R}$.



A linear map sends linear combinations in the domain to the same linear combinations of the images.

Note

To prove a map is linear, it is enough to check additivity $T(u+v) = T(u) + T(v)$ and homogeneity $T(cu) = cT(u)$. To disprove linearity, find one counterexample to either rule.

Example 15

Define $T : M_{2\times 2}(\mathbb{R}) \to M_{2\times 2}(\mathbb{R})$ by

$$T(A) = 3A^T + \operatorname{tr}(A)\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}.$$

This map is linear because transpose and trace are linear:

$$T(cA + dB) = 3(cA + dB)^T + \operatorname{tr}(cA + dB)\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} = cT(A) + dT(B).$$

Definition 11 (Eigenvector of a linear map)

If $T : V \to V$ is linear, a non-zero vector $v \in V$ is an **eigenvector** of T with eigenvalue λ if

$$T(v) = \lambda v.$$

Example 16

For the map above,

$$T\left(\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}\right) = 3\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix},$$

so

$$\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

is an eigenvector with eigenvalue 3. But

$$T(I) = 3I + 2\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix},$$

which is not a scalar multiple of I, so I is not an eigenvector.

12 Revision Guide

This chapter is a compact checklist for pre-exam revision. It is not a substitute for the definitions, proofs, and examples in the earlier chapters.

12.1 Key Results and Theorems

Note

The following results are central to the course and are natural candidates for statement, proof, or application questions:

- monotone convergence theorem,
- Bolzano-Weierstrass theorem,
- Cauchy convergence criterion for sequences,
- intermediate value theorem,
- extreme value theorem,
- Rolle's theorem and the mean value theorem,
- fundamental theorem of calculus,
- comparison criterion for series.

For each theorem, practise writing the assumptions before the conclusion. A correct conclusion with missing assumptions is usually not a correct theorem statement.

Example 1 (Monotone convergence theorem)

A complete statement is:

If (a_n) is monotone increasing and bounded above, then (a_n) converges and

$$\lim_{n\to\infty} a_n = \sup\{a_n : n \in \mathbb{N}\}.$$

Similarly, if (a_n) is monotone decreasing and bounded below, then it converges to its infimum.

Example 2 (Bolzano-Weierstrass theorem)

A complete statement is:

Every bounded sequence in $\mathbb{R}$ has a convergent subsequence.

A standard proof route is: find a monotone subsequence, observe that it is still bounded, then apply the monotone convergence theorem to that subsequence.

Example 3 (Fundamental theorem of calculus)

A complete FTC statement must specify the regularity assumptions. One common version is:

If $f : [a, b] \to \mathbb{R}$ is continuous and

$$F(x) = \int_a^x f(t)\,\mathrm{d}t,$$

then F is differentiable on (a, b) and $F'(x) = f(x)$.

12.2 Proof Routines

Proposition 1 (Epsilon proof for sequences)
To prove $a_n \to L$ from the definition:

1. Start with $\varepsilon > 0$.
2. Estimate $|a_n - L|$ by something simple.
3. Choose N so that the simple expression is $< \varepsilon$ whenever $n \geq N$.
4. Finish by writing the implication $n \geq N \Rightarrow |a_n - L| < \varepsilon$.

Example 4
To prove $\frac{1}{\sqrt{n}} \to 0$, let $\varepsilon > 0$ and choose $N > \frac{1}{\varepsilon^2}$. If $n \geq N$, then

$$\left|\frac{1}{\sqrt{n}} - 0\right| = \frac{1}{\sqrt{n}} \leq \frac{1}{\sqrt{N}} < \varepsilon.$$

Proposition 2 (Epsilon-delta proof for functions)
To prove $\lim_{x\to a} f(x) = L$:

1. Start with $\varepsilon > 0$.
2. Rewrite or bound $|f(x) - L|$ in terms of $|x - a|$.
3. Choose δ so the bound is $< \varepsilon$.
4. Finish with $0 < |x - a| < \delta \Rightarrow |f(x) - L| < \varepsilon$.

Example 5
To show $\lim_{x\to 2} x^2 = 4$, restrict $|x - 2| < 1$, so $1 < x < 3$ and $|x + 2| < 5$. Then

$$|x^2 - 4| = |x - 2||x + 2| < 5|x - 2|.$$

Choose $\delta = \min\left(1, \frac{\varepsilon}{5}\right)$.

Proposition 3 (Continuity proof pattern)
To prove a function is continuous at a, either use the definition directly, or combine known continuous functions using limit laws. To prove discontinuity, find one-sided limits or a sequence $x_n \to a$ such that $f(x_n)$ does not tend to $f(a)$.

Proposition 4 (Subspace proof pattern)
To prove U is a subspace, show:

$$0 \in U, \quad u, v \in U, c, d \in \mathbb{R} \Rightarrow cu + dv \in U.$$

To disprove it, it is enough to show one failure: no zero vector, not closed under addition, or not closed under scalar multiplication.

12.3 Series Questions

Note
For a series $\sum a_n$, first check $a_n \to 0$. If not, the series diverges. If yes, choose a test based on the form of the term.

Term shape	Likely tool

r^n, constants to the n	geometric series, root test
factorials	ratio test
powers of n	comparison or limit comparison
$\frac{1}{n(\log n)^p}$	integral test
alternating signs	alternating series test, then absolute convergence
monotone positive terms with powers of 2	Cauchy condensation; starred in the 2026 tutorials

Example 6
For

$$\sum_{n=1}^{\infty} \frac{(2n)!}{5^n n!},$$

factorials suggest the ratio test:

$$\frac{a_{n+1}}{a_n} = \frac{(2n+2)(2n+1)}{5(n+1)} \to \infty.$$

Hence the series diverges.

Example 7
For

$$\sum_{n=2}^{\infty} \frac{1}{n(\log n)^p},$$

the integral test gives

$$\int_2^{\infty} \frac{1}{x(\log x)^p}\, \mathrm{d}x = \int_{\log 2}^{\infty} \frac{1}{u^p}\, \mathrm{d}u.$$

Therefore the series converges exactly when $p > 1$.

12.4 Integration Questions

Note
For a definite or improper integral:

- check whether the integrand is continuous on the interval;
- identify singular endpoints or infinite endpoints;
- for variable limits, use the fundamental theorem before simplifying;
- for antiderivatives, try substitution before integration by parts if a derivative is visibly present.

Example 8
If

$$F(x) = \int_{\sin x}^{1} e^{t^2}\,\mathrm{d}t,$$

then the upper limit is constant and the lower limit depends on x. By the variable-limit rule,

$$F'(x) = -e^{\sin^2 x}\cos x.$$

Example 9
To decide

$$\int_2^\infty \frac{1}{x^2-1}\,\mathrm{d}x,$$

use partial fractions:

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

Hence

$$\int_2^R \frac{1}{x^2-1}\,\mathrm{d}x = \frac{1}{2}\log\left(\frac{R-1}{R+1}\right) + \frac{1}{2}\log 3,$$

which tends to $\frac{1}{2}\log 3$.

12.5 Linear Algebra Questions

Note
The determinant, eigenvalue, and vector-space material below is represented strongly in the 2026 tutorials and recent exams. The current lecture transcript set does not include the full Week 12–13 determinant/eigenvalue/vector-space lectures.

Question type	Method
Solve $Ax = b$	row-reduce the augmented matrix $[A \mid b]$
Solve $Ax = 0$	row-reduce and express the free variables as parameters
Find A^{-1}	row-reduce $[A \mid I]$ to $[I \mid A^{-1}]$
Build a matrix for T	use the columns $T(e_1), \ldots, T(e_n)$
Compute a large determinant	use row operations or expand along a sparse row or column
Find eigenvalues	solve $\det(A - \lambda I) = 0$
Find eigenvectors	solve $(A - \lambda I)v = 0$ for each eigenvalue
Test a subspace	check $0 \in U$ and closure under linear combinations

Note

The dot product, cross product, projections, orthogonal matrices, symmetric/skew-symmetric matrices, trace, nilpotent matrices, and Vandermonde determinant are tutorial-derived tools in these notes. Vandermonde is starred in the 2026 tutorials.

Example 10

For

$$A = \begin{pmatrix} 5 & 0 & 0 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{pmatrix},$$

the characteristic polynomial is

$$\det(A - \lambda I) = (5 - \lambda)\big((2 - \lambda)^2 - 1\big).$$

Thus $\lambda = 5, 3, 1$. Then solve $(A - \lambda I)v = 0$ separately for each eigenvalue.

Example 11

Let

$$V = \{p(x) \in P_7(\mathbb{R}) : p(1) = 0\}.$$

The condition $p(1) = 0$ is linear, so V is a subspace. Every polynomial in V has a factor $x - 1$, and degree at most 7, so

$$p(x) = (x - 1)q(x), \quad \deg q \leq 6.$$

A basis is

$$x - 1, x(x - 1), x^2(x - 1), \ldots, x^6(x - 1),$$

and $\dim V = 7$.

12.6 Common Traps

Note

- A theorem cannot be used unless its hypotheses are checked.
- A sequence can have terms tending to 0 while the series still diverges.
- Absolute convergence implies convergence, but convergence need not imply absolute convergence.
- A Taylor series formula must be used inside its interval of convergence.
- The zero vector is in every eigenspace, but is not an eigenvector.
- A set described by a non-linear condition is usually not a subspace.
- In this course, write log for the natural logarithm.

12.7 Complete Answer Standards

For computation questions, write enough working that the method is clear. For proof questions, state the relevant assumptions and make each implication explicit.

Example 12

A complete answer to “show the series converges absolutely” should include:

1. the absolute-value series,
2. the comparison or test being used,
3. the conclusion that the absolute-value series converges,
4. the conclusion that the original series converges absolutely.

Example 13

A complete answer to “find an eigenspace” should include:

1. the eigenvalue λ,
2. the matrix $A - \lambda I$,
3. the reduced homogeneous system,
4. the span of the non-zero eigenvectors.