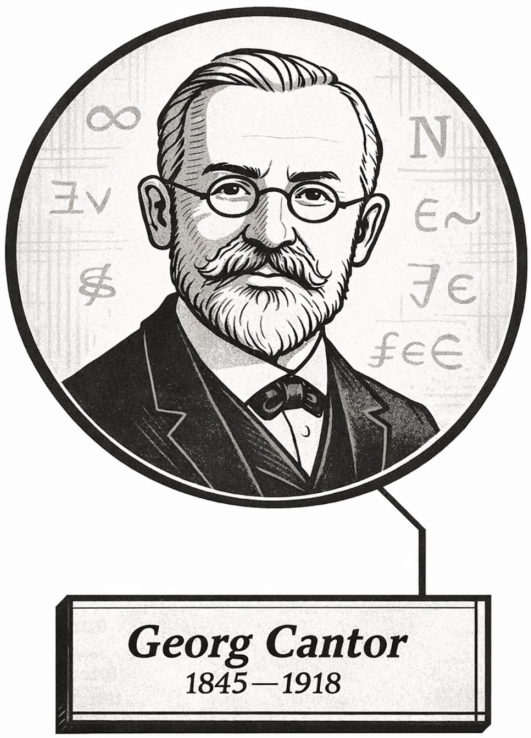
Learning Abstract Mathematics

A Structured Journey Through Logic, Sets, Algebra, and Analysis

William E. Sollers III

March 2, 2026

"Proof is the price of admission."



To infinity and beyond!

Georg Cantor (probably)

Contents

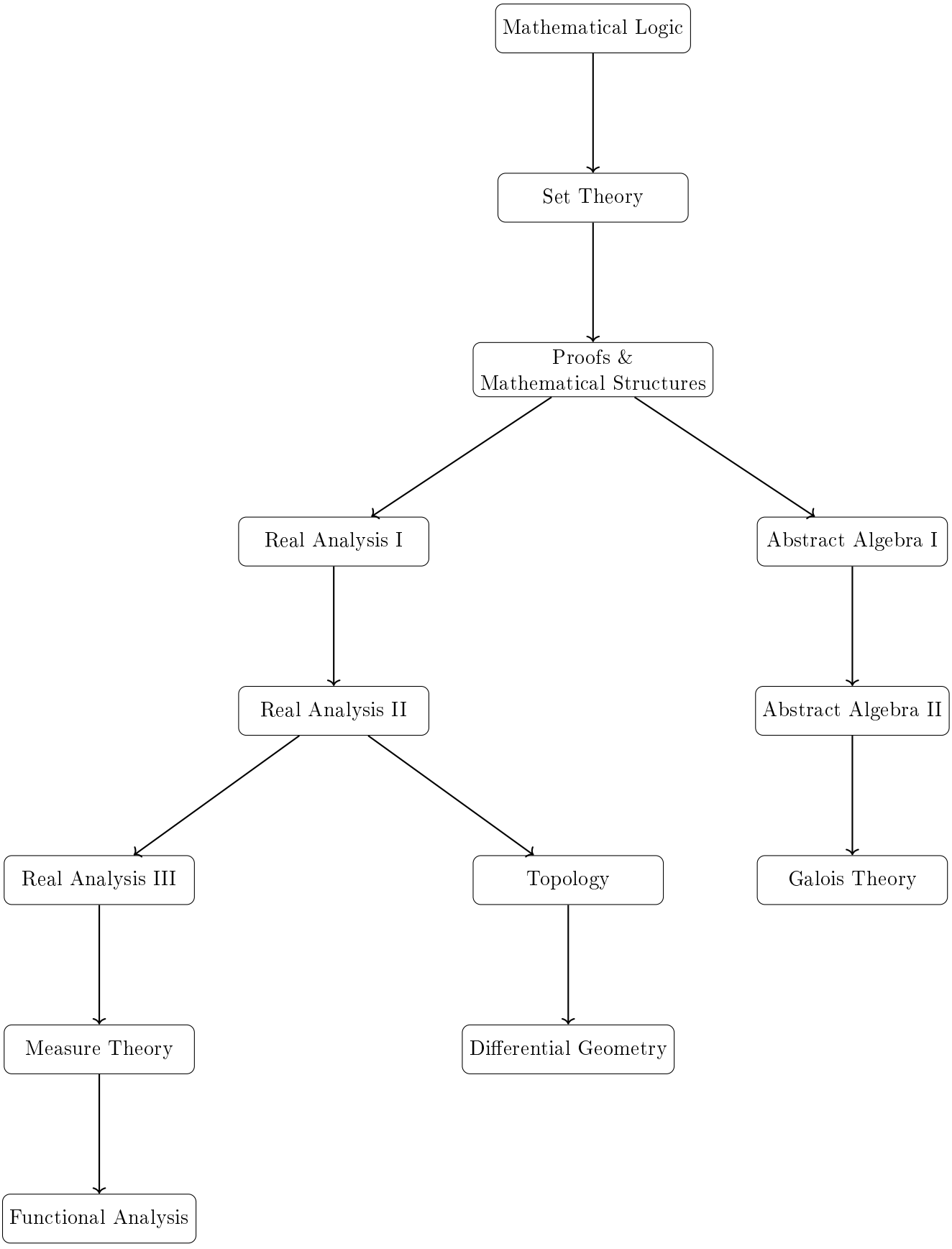


Figure 1: Standard dependency graph for undergraduate abstract mathematics

Part I

Mathematical Logic

Chapter 1

Propositional Logic

1.1 Notes

Where You Are in the Journey

Propositional Logic $\rightarrow$ Predicate Calculus $\rightarrow$ Sets & Functions $\rightarrow$ Proof Techniques $\rightarrow$ Axiom Systems $\rightarrow$ $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$ $\rightarrow$ Real Analysis $\rightarrow$ $\cdots$

How we got here. This is the starting point. Mathematics needs a precise language for reasoning — a system for forming statements, combining them with connectives, and determining their truth mechanically. Propositional logic provides that foundation.

What this chapter builds. We develop the syntax of propositional formulas, their semantics via truth assignments, and the proof theory of natural deduction. The central results — soundness, completeness, and compactness — characterise what is provable and what follows logically.

Where this leads. Predicate logic extends this language with quantifiers, allowing us to talk about objects and their properties rather than just truth values. The metatheoretic pattern (syntax, semantics, soundness, completeness) established here repeats at every level of the logical hierarchy.

Structural Roadmap

Each major topic is organised as:

$$\text{Syntax} \longrightarrow \text{Semantics} \longrightarrow \text{Proof Theory} \longrightarrow \text{Metatheory}$$

The global progression is:

1. Syntax: propositional variables, connectives, well-formed formulas

2. Semantics: truth assignments, validity, logical equivalence, normal forms
3. Proof theory: inference rules, natural deduction, derivability
4. Metatheory: functional completeness, compactness, Craig's interpolation
5. Reference: fallacies and summary tables

Remark 1.1 (Primary source). The primary driver is Bjørndahl's Logic and Proof, supplemented by Suppes' Introduction to Logic.

1.1.1 Syntax of Propositional Logic

Syntax Toolkit Quick Reference

Concept	Meaning	Detail
Propositional variable	Atomic symbol with no internal structure	Def
Logical connective	Symbol forming compound formulas	Def
Well-formed formula	Recursively constructed expression	Def
Unique readability	Every wff has exactly one parse tree	Thm
Parse tree	Tree representation of formula structure	Def
Subformula	Constituents of a formula	Def
Formula depth	Nesting depth of connectives	Def
Operator precedence	Disambiguation convention	Def

Definition (Propositional Variable)

A propositional variable (or atomic proposition) is a primitive symbol representing a statement that is either true or false.
The set of all propositional variables is denoted

$$\mathsf{Prop} = \{P_1, P_2, P_3, \dots\}$$

or informally by letters $P, Q, R, S, \dots$

Remark 1.2 (English reading). A propositional variable is the smallest meaningful unit in propositional logic. It names a proposition "It is raining," "The number is prime" but has no further decomposable structure. Atomicity is the defining feature: unlike predicate logic, we cannot look inside a variable and ask who or what it talks about.

Remark 1.3 (Fully formal statement). Prop is a countably infinite set of symbols, pairwise distinct. A propositional variable is any element $P_i \in \mathsf{Prop}$. No propositional variable is a connective, a parenthesis, or a formula built from other symbols.

Remark 1.4 (Consequence for proof strategy). Proofs about all propositional variables are trivial base cases in structural induction: if $\varphi \in \mathsf{Prop}$, the property holds by the atomic case of the induction.

Definition (Logical Connectives)

The logical connectives of propositional logic are:

Symbol	Name	Arity	Reading
$\neg$	Negation	Unary	"not P"
$\wedge$	Conjunction	Binary	"P and Q"
$\vee$	Disjunction	Binary	"P or Q"
$\rightarrow$	Conditional	Binary	"if P then Q"
$\leftrightarrow$	Biconditional	Binary	"P if and only if Q"

Remark 1.5 (English reading). Connectives are the glue. They take propositions as input and produce propositions as output. $\neg$ flips a truth value; binary connectives combine two truth values into one according to a fixed truth table.

Remark 1.6 (Notation variants across sources). $\sim$ or ! for $\neg$; &, $\cdot$, or && for $\wedge$; |, + for $\vee$; $\supset$, $\Rightarrow$ for $\rightarrow$; $\equiv$, $\Leftrightarrow$ for $\leftrightarrow$. Bjørndahl uses $\neg, \wedge, \vee, \rightarrow$. Suppes & Hill use similar. Always check the source's symbol table.

Remark 1.7 (Logical implication). Five connectives suffice for propositional logic, but not all are independent. $\leftrightarrow$ is definable as $(P \rightarrow Q) \wedge (Q \rightarrow P)$; $\rightarrow$ is definable as $\neg P \vee Q$. The standard five are chosen for readability, not minimality.

Definition (Well-Formed Formula)

Let $\mathcal{L}$ be a propositional language. The set $\mathsf{WFF}_{\mathcal{L}}$ of well-formed formulas is the smallest set satisfying:

1. Atomic: Every $P \in \mathsf{Prop}$ is a wff.
2. Negation: If φ is a wff, so is $\neg\varphi$.
3. Binary: If φ, ψ are wffs, so are $(\varphi \wedge \psi)$, $(\varphi \vee \psi)$, $(\varphi \rightarrow \psi)$, $(\varphi \leftrightarrow \psi)$.
4. Closure: Nothing else is a wff.

Remark 1.8 (English reading). A wff is any expression that can be produced by starting with atomic variables and repeatedly applying connectives according to the rules above. The closure clause (rule 4) is crucial: it rules out nonsense strings like $\rightarrow P\wedge$ that no finite application of rules 1–3 can produce.

Remark 1.9 (Fully quantified form). WFF is the intersection of all sets X satisfying: (i) $\mathsf{Prop} \subseteq X$; (ii) $\varphi \in X \Rightarrow \neg\varphi \in X$; (iii) $\varphi, \psi \in X \Rightarrow (\varphi \circ \psi) \in X$ for each binary connective $\circ$. This is the standard inductive definition / least fixed-point construction.

Remark 1.10 (Consequence structural induction). Because WFF is inductively defined, structural induction applies: to prove property $\mathcal{P}(\varphi)$ for all wffs, prove it for atomic φ, and show that each formation rule preserves $\mathcal{P}$. This is the standard proof technique for syntactic results in logic.

Definition 1.11 (Atomic and Molecular Formulas). A formula is atomic if it is a propositional variable. A formula is molecular (or compound) if it is not atomic.

Theorem (Unique Readability)

Every well-formed formula φ can be constructed in exactly one way. Formally, exactly one of the following holds:

1. $\varphi \in \mathsf{Prop}$ (atomic case).
2. $\varphi = \neg\psi$ for a unique wff ψ.
3. $\varphi = (\psi \circ \chi)$ for a unique binary connective $\circ$ and unique wffs ψ, χ.

Remark 1.12 (Why this theorem is a theorem). The recursive definition of wffs gives many ways to generate formulas. Unique readability asserts that no two generation paths produce the same string. This is non-trivial: it fails for most context-free grammars. Propositional logic avoids ambiguity by requiring explicit parentheses around every binary connective application.

Remark 1.13 (Consequence well-defined semantics). Truth evaluation $v(\varphi)$ is defined recursively by cases on the structure of φ. If φ had two structures, $v(\varphi)$ might be ill-defined. Unique readability guarantees that the recursive definition of truth is coherent.

Remark 1.14 (Consequence structural induction is safe). Structural induction proceeds by case analysis on the unique form of φ. Without unique readability, cases could overlap and the induction would break.

Definition (Parse Tree)

The parse tree (or formation tree) of a wff φ is a labeled binary tree where:

- Each leaf is labeled with a propositional variable.
- Each internal node is labeled with a connective.
- A node labeled $\neg$ has one child; a node labeled with a binary connective has two children.

Remark 1.15 (Intuition). The parse tree externalizes the unique recursive structure of a formula. Reading it top-down reconstructs the formula; reading it bottom-up shows how sub-results combine into the whole truth value.

Remark 1.16 (Consequence). Unique readability guarantees every wff has a unique parse tree. This makes structural induction synonymous with tree induction: induction on the height of the parse tree.

Definition (Subformula)

The set of subformulas of φ, written $\mathrm{Sub}(\varphi)$, is defined recursively:

1. φ atomic: $\mathrm{Sub}(\varphi) = \{\varphi\}$.
2. $\varphi = \neg\psi$: $\mathrm{Sub}(\varphi) = \{\varphi\} \cup \mathrm{Sub}(\psi)$.
3. $\varphi = (\psi \circ \chi)$: $\mathrm{Sub}(\varphi) = \{\varphi\} \cup \mathrm{Sub}(\psi) \cup \mathrm{Sub}(\chi)$.

A proper subformula is any element of $\mathrm{Sub}(\varphi)$ other than φ itself.

Remark 1.17 (Intuition). Every node in the parse tree of φ corresponds to exactly one subformula. The subformulas are the "parts" of φ visible at each level of nesting.

Remark 1.18 (Consequence). Many logical properties tautology, satisfiability are defined on whole formulas but proved by induction on subformulas. A property that holds for all subformulas of φ and is preserved by connectives holds for φ.

Definition (Formula Depth)

The depth (or complexity) of a formula φ:

1. φ atomic: $\text{depth}(\varphi) = 0$.
2. $\varphi = \neg\psi$: $\text{depth}(\varphi) = \text{depth}(\psi) + 1$.
3. $\varphi = (\psi \circ \chi)$: $\text{depth}(\varphi) = \max\{\text{depth}(\psi), \text{depth}(\chi)\} + 1$.

Remark 1.19 (Intuition). Depth measures how deeply nested the connectives are equivalently, the height of the parse tree. A depth-0 formula has no connectives; a depth-n formula has a subformula at nesting level n.

Remark 1.20 (Consequence strong induction). Induction on depth is the canonical way to prove properties of all wffs when you need the inductive hypothesis to apply to all formulas of strictly smaller depth, not just immediate subformulas.

Definition (Operator Precedence)

Standard precedence (highest binds tightest):

1. Parentheses (override all)
2. $\neg$ (negation)
3. $\wedge$ (conjunction)
4. $\vee$ (disjunction)
5. $\to$ (conditional, right-associative)
6. $\leftrightarrow$ (biconditional)

Remark 1.21 (English reading). Precedence is a convention for omitting parentheses without creating ambiguity. $\neg P \wedge Q$ means $(\neg P) \wedge Q$ because $\neg$ binds tighter than $\wedge$. The conditional is right-associative: $P \to Q \to R$ means $P \to (Q \to R)$.

Remark 1.22 (Common error). $\neg P \wedge Q$ is not $\neg(P \wedge Q)$. Negation applies to the smallest formula immediately to its right. When in doubt, use explicit parentheses.

Example 1.23. Using standard precedence:

- $\neg P \wedge Q$ means $(\neg P) \wedge Q$.
- $P \vee Q \wedge R$ means $P \vee (Q \wedge R)$ if $\wedge > \vee$.
- $P \to Q \vee R$ means $P \to (Q \vee R)$.

1.1.2 Additional Connectives

Additional Connectives Quick Reference

Symbol	Name	Equivalent	Detail
$P \oplus Q$	XOR (Exclusive Or)	$\neg(P \leftrightarrow Q)$	Def
$P \uparrow Q$	NAND (Sheffer stroke)	$\neg(P \wedge Q)$	Def
$P \downarrow Q$	NOR (Peirce arrow)	$\neg(P \vee Q)$	Def
$\{\uparrow\}$	NAND alone is adequate		Prop
$\{\downarrow\}$	NOR alone is adequate		Prop

P	Q	$P \oplus Q$	$P \uparrow Q$	$P \downarrow Q$
T	T	F	F	F
T	F	T	T	F
F	T	T	T	F
F	F	F	T	T

Definition (Exclusive Or XOR)

The exclusive or, denoted $P \oplus Q$ (also $P \veebar Q$), is true when exactly one of P, Q is true:

$$P \oplus Q \;\equiv\; \neg(P \leftrightarrow Q)$$

Equivalently: $(P \vee Q) \wedge \neg(P \wedge Q)$, and also $(P \wedge \neg Q) \vee (\neg P \wedge Q)$.

Remark 1.24 (English reading). XOR captures the natural-language "either or but not both." It differs from inclusive or ($\vee$) only in the (T, T) row: $P \vee Q$ is true there, $P \oplus Q$ is false.

Remark 1.25 (Fully quantified properties). For all P, Q, R:

$$\begin{aligned}
P \oplus Q &\equiv Q \oplus P \quad \text{(commutativity)} \\
(P \oplus Q) \oplus R &\equiv P \oplus (Q \oplus R) \quad \text{(associativity)} \\
P \oplus \bot &\equiv P \quad \text{(identity)} \\
P \oplus \top &\equiv \neg P \quad \text{(negation)} \\
P \oplus P &\equiv \bot \quad \text{(self-inverse)} \\
P \wedge (Q \oplus R) &\equiv (P \wedge Q) \oplus (P \wedge R) \quad \text{(distributivity of } \wedge \text{ over } \oplus)
\end{aligned}$$

The self-inverse law makes $(\mathbb{B}, \oplus)$ a group a fact that resurfaces in abstract algebra.

Remark 1.26 (Consequence). XOR is definable from standard connectives, so it adds no expressive power. It is included for notational convenience and because it has a clean algebraic structure (Boolean ring addition).

Definition (NAND Sheffer Stroke)

The NAND connective (Sheffer stroke), denoted $P \uparrow Q$ or $P|Q$, is false only when both operands are true:

$$P \uparrow Q \equiv \neg(P \wedge Q)$$

Remark 1.27 (English reading). "Not both P and Q." NAND is true in every row of the truth table except when both inputs are true. It is the most common single connective in digital circuit design (NAND gates are universal).

Remark 1.28 (Fully quantified expressing all connectives via NAND).

$$\begin{aligned} \neg P &\equiv P \uparrow P \\ P \wedge Q &\equiv (P \uparrow Q) \uparrow (P \uparrow Q) \\ P \vee Q &\equiv (P \uparrow P) \uparrow (Q \uparrow Q) \\ P \to Q &\equiv P \uparrow (Q \uparrow Q) \end{aligned}$$

Proposition 1.29 (Non-Associativity of NAND). NAND is not associative: $(P \uparrow Q) \uparrow R \not\equiv P \uparrow (Q \uparrow R)$.

Remark 1.30 (Consequence). Non-associativity means NAND expressions require careful parenthesization. Despite this, $\{\uparrow\}$ alone is functionally complete see the Functional Completeness section.

Definition (NOR Peirce Arrow)

The NOR connective (Peirce arrow), denoted $P \downarrow Q$, is true only when both operands are false:

$$P \downarrow Q \equiv \neg(P \vee Q)$$

Remark 1.31 (English reading). "Neither P nor Q." NOR is the dual of NAND (swap $\wedge \leftrightarrow \vee$ and $\top \leftrightarrow \bot$). It is also a universal gate in digital circuits.

Remark 1.32 (Fully quantified expressing all connectives via NOR).

$$\begin{aligned} \neg P &\equiv P \downarrow P \\ P \vee Q &\equiv (P \downarrow Q) \downarrow (P \downarrow Q) \\ P \wedge Q &\equiv (P \downarrow P) \downarrow (Q \downarrow Q) \\ P \to Q &\equiv ((P \downarrow P) \downarrow Q) \downarrow ((P \downarrow P) \downarrow Q) \end{aligned}$$

Proposition 1.33 (Non-Associativity of NOR). NOR is not associative: $(P \downarrow Q) \downarrow R \not\equiv P \downarrow (Q \downarrow R)$.

Remark 1.34 (Duality with NAND). NAND and NOR are De Morgan duals: replacing each by the other and swapping $\wedge \leftrightarrow \vee$ transforms any NAND circuit into a NOR circuit computing the same function. They share identical functional completeness properties.

1.1.3 Semantics of Propositional Logic

Semantics Toolkit Quick Reference

Concept	Meaning	Detail
Truth assignment	Function $v : \mathsf{Prop} \to \{\mathsf{T}, \mathsf{F}\}$	Def
Truth under v	$v \models \varphi$ iff $v(\varphi) = \mathsf{T}$	Def
Tautology	True under all assignments ($\models \varphi$)	Def
Contradiction	False under all assignments	Def
Satisfiable	True under at least one assignment	Def
Contingency	Neither tautology nor contradiction	Def
Logical consequence	$\Gamma \models \varphi$	Def
Logical equivalence	$\varphi \equiv \psi$ ($\varphi \leftrightarrow \psi$ is a tautology)	Def

Definition (Truth Assignment)

A truth assignment (or valuation) is a function

$$v : \mathsf{Prop} \to \{\mathsf{T}, \mathsf{F}\}.$$

Every truth assignment extends uniquely to $\hat{v} : \mathsf{WFF} \to \{\mathsf{T}, \mathsf{F}\}$ by:

1. $\hat{v}(P) = v(P)$ for $P \in \mathsf{Prop}$.
2. $\hat{v}(\neg\varphi) = \mathsf{T}$ iff $\hat{v}(\varphi) = \mathsf{F}$.
3. $\hat{v}(\varphi \wedge \psi) = \mathsf{T}$ iff $\hat{v}(\varphi) = \hat{v}(\psi) = \mathsf{T}$.
4. $\hat{v}(\varphi \vee \psi) = \mathsf{T}$ iff $\hat{v}(\varphi) = \mathsf{T}$ or $\hat{v}(\psi) = \mathsf{T}$.
5. $\hat{v}(\varphi \to \psi) = \mathsf{T}$ iff $\hat{v}(\varphi) = \mathsf{F}$ or $\hat{v}(\psi) = \mathsf{T}$.
6. $\hat{v}(\varphi \leftrightarrow \psi) = \mathsf{T}$ iff $\hat{v}(\varphi) = \hat{v}(\psi)$.

We write $v(\varphi)$ for $\hat{v}(\varphi)$.

Remark 1.35 (English reading). A truth assignment is an interpretation: it gives a definite truth value to every atomic proposition, and the extension rules then compute the truth value of every compound formula mechanically. There is no ambiguity once v is fixed.

Remark 1.36 (Fully quantified). $\hat{v}$ is the unique function $\mathsf{WFF} \to \{\mathsf{T}, \mathsf{F}\}$ extending v and compatible with the recursive structure of wffs. Uniqueness follows from unique readability: each wff has exactly one form, so the recursive cases never conflict.

Remark 1.37 (Truth table connection). A formula with n distinct variables has exactly 2^n distinct truth assignments. The truth table lists all of them, one per row.

Truth Tables for Standard Connectives

P	Q	$\neg P$	$P \wedge Q$	$P \vee Q$	$P \to Q$	$P \leftrightarrow Q$
T	T	F	T	T	T	T
T	F	F	F	T	F	F
F	T	T	F	T	T	F
F	F	T	F	F	T	T

Definition (Satisfaction)

A truth assignment v satisfies formula φ, written $v \models \varphi$, if $v(\varphi) = \mathsf{T}$.
v satisfies a set Γ if $v \models \varphi$ for every $\varphi \in \Gamma$.

Remark 1.38 (Intuition). $v \models \varphi$ is the semantic turnstile: v makes φ true. The symbol $\models$ without a subscript on the left will be reused for logical consequence ($\Gamma \models \varphi$) — context determines which reading.

Definitions (Formula Classification)

Tautology — φ is a tautology (written $\models \varphi$) if $v \models \varphi$ for every truth assignment v.

Contradiction — φ is a contradiction (or unsatisfiable) if $v \not\models \varphi$ for every v.

Satisfiable — φ is satisfiable if there exists some v with $v \models \varphi$.

Contingency — φ is a contingency if it is neither a tautology nor a contradiction.

Remark 1.39 (Mutual exclusivity and exhaustiveness). Every formula belongs to exactly one category:

Category	True under all v?	True under some v?
Tautology	Yes	Yes (trivially)
Contradiction	No	No
Contingency	No	Yes

Remark 1.40 (Key logical implication). A formula is satisfiable if and only if it is not a contradiction. This equivalence is the bridge between satisfiability and unsatisfiability, and is the basis for refutation proofs (prove φ by deriving a contradiction from $\neg\varphi$).

Example 1.41. • $P \vee \neg P$ — tautology (law of excluded middle).

- $P \wedge \neg P$ — contradiction.
- $P \to Q$ — contingency (false when $P = \mathsf{T}$, $Q = \mathsf{F}$).

Definition (Logical Consequence)

φ is a logical consequence of Γ, written $\Gamma \models \varphi$, if every truth assignment satisfying Γ also satisfies φ.
Formally: $\forall v,\ (v \models \Gamma) \Rightarrow (v \models \varphi)$.

Remark 1.42 (English reading). $\Gamma \models \varphi$ means: there is no way to make all of Γ true while making φ false. The premise set Γ semantically forces φ.

Remark 1.43 (Special cases). • $\emptyset \models \varphi$ iff φ is a tautology.

- $\Gamma \models \varphi$ for any Γ containing a contradiction.
- $\{P\} \models Q$ iff $P \to Q$ is a tautology (tautological implication).

Remark 1.44 (Connection to proof systems). The fundamental theorem of propositional logic: $\Gamma \models \varphi \iff \Gamma \vdash \varphi$ (soundness and completeness). The left side is semantic; the right side is syntactic. They coincide for propositional logic.

Definition (Logical Equivalence)

φ and ψ are logically equivalent, written $\varphi \equiv \psi$, if they have the same truth value under every truth assignment. Equivalently:

- $\varphi \models \psi$ and $\psi \models \varphi$, or
- $\varphi \leftrightarrow \psi$ is a tautology, or
- φ and ψ have identical truth tables.

Remark 1.45 (Intuition). Logically equivalent formulas say the same thing they agree on every possible state of affairs. They can be freely substituted for one another in any context.

Remark 1.46 (Equivalence is an equivalence relation). $\equiv$ is reflexive ($\varphi \equiv \varphi$), symmetric, and transitive. Moreover, it is a congruence: if $\varphi \equiv \psi$, then $(\varphi \circ \chi) \equiv (\psi \circ \chi)$ for any connective $\circ$ and formula χ. This substitution property is what makes equivalences useful for formula simplification.

1.1.4 Logical Equivalences

Logical Equivalences Quick Reference

Name	Equivalence(s)	Detail
Double Negation	$\neg\neg P \equiv P$	
De Morgan (pair)	$\neg(P \wedge Q) \equiv \neg P \vee \neg Q$; $\neg(P \vee Q) \equiv \neg P \wedge \neg Q$	
Commutativity	$P \wedge Q \equiv Q \wedge P$; $P \vee Q \equiv Q \vee P$	
Associativity	$(P \wedge Q) \wedge R \equiv P \wedge (Q \wedge R)$; and $\vee$	
Distributivity	$P \wedge (Q \vee R) \equiv (P \wedge Q) \vee (P \wedge R)$; and dual	
Idempotence	$P \wedge P \equiv P$; $P \vee P \equiv P$	
Absorption	$P \wedge (P \vee Q) \equiv P$; $P \vee (P \wedge Q) \equiv P$	
Identity	$P \wedge \top \equiv P$; $P \vee \bot \equiv P$	
Domination	$P \vee \top \equiv \top$; $P \wedge \bot \equiv \bot$	
Negation Laws	$P \vee \neg P \equiv \top$; $P \wedge \neg P \equiv \bot$	
Material Implication	$P \to Q \equiv \neg P \vee Q$	
Contraposition	$P \to Q \equiv \neg Q \to \neg P$	
Exportation	$(P \wedge Q) \to R \equiv P \to (Q \to R)$	
Neg. Conditional	$\neg(P \to Q) \equiv P \wedge \neg Q$	
Biconditional Exp.	$P \leftrightarrow Q \equiv (P \to Q) \wedge (Q \to P)$	
Duality Principle	$\varphi \equiv \psi \Rightarrow \varphi^d \equiv \psi^d$	Thm

Core Equivalences

Proposition 1.47 (Double Negation). $\neg\neg P \equiv P$

Remark 1.48 (Intuition). Negating twice returns to the original truth value. In classical logic, $\neg\neg P$ and P are interchangeable everywhere. In intuitionistic logic this fails $\neg\neg P \Rightarrow P$ is not provable without the law of excluded middle.

Remark 1.49 (Proof strategy). Truth table: 2 rows, 2 columns; inspect directly.

Proposition 1.50 (De Morgan's Laws).

$$\neg(P \wedge Q) \equiv \neg P \vee \neg Q$$
$$\neg(P \vee Q) \equiv \neg P \wedge \neg Q$$

Remark 1.51 (Intuition). Negation distributes through conjunction/disjunction by flipping the connective. "Not both" = "not the first or not the second." "Neither" = "not the first and not the second."

Remark 1.52 (Consequence). De Morgan's laws are the primary tool for pushing negations inward (step 2 of NNF conversion) and for expanding $\neg$ through complex formulas.

Remark 1.53 (Logical implication duality). The two De Morgan laws are duals of each other under the duality principle (swap $\wedge \leftrightarrow \vee$, $\top \leftrightarrow \bot$).

Proposition 1.54 (Commutativity). $P \wedge Q \equiv Q \wedge P$; $\quad P \vee Q \equiv Q \vee P$

Remark 1.55 (Intuition). Order of operands does not matter for $\wedge$ and $\vee$. Note: $\rightarrow$ is not commutative ($P \rightarrow Q \not\equiv Q \rightarrow P$ in general).

Proposition 1.56 (Associativity). $(P \wedge Q) \wedge R \equiv P \wedge (Q \wedge R)$; $\quad (P \vee Q) \vee R \equiv P \vee (Q \vee R)$

Remark 1.57 (Intuition). Parenthesization does not affect truth value for chains of $\wedge$ or $\vee$. This justifies writing $P \wedge Q \wedge R$ without parentheses.

Remark 1.58 (Common error). $\rightarrow$ is right-associative by convention but not associative as an equivalence: $(P \rightarrow Q) \rightarrow R \not\equiv P \rightarrow (Q \rightarrow R)$ in general.

Proposition 1.59 (Distributivity). $P \wedge (Q \vee R) \equiv (P \wedge Q) \vee (P \wedge R)$; $\quad P \vee (Q \wedge R) \equiv (P \vee Q) \wedge (P \vee R)$

Remark 1.60 (Intuition). $\wedge$ distributes over $\vee$ and $\vee$ distributes over $\wedge$ unlike in arithmetic where only multiplication distributes over addition, here both distribute over the other.

Remark 1.61 (Consequence). Distributivity is the key step in CNF and DNF conversion. DNF uses $\wedge$ over $\vee$; CNF uses $\vee$ over $\wedge$.

Proposition 1.62 (Idempotence). $P \wedge P \equiv P$; $\quad P \vee P \equiv P$

Remark 1.63 (Intuition). Repeating a formula under $\wedge$ or $\vee$ adds no new information. Useful for simplification.

Proposition 1.64 (Absorption). $P \wedge (P \vee Q) \equiv P$; $\quad P \vee (P \wedge Q) \equiv P$

Remark 1.65 (Intuition). The stronger condition absorbs the weaker. If P is true, then $P \vee Q$ is certainly true, so $P \wedge (P \vee Q)$ is just P.

Proposition 1.66 (Identity Laws). $P \wedge \top \equiv P$; $\quad P \vee \bot \equiv P$

Remark 1.67 (Intuition). $\top$ is the identity for $\wedge$; $\bot$ is the identity for $\vee$. They play the role of 1 and 0 in Boolean algebra.

Proposition 1.68 (Domination Laws). $P \vee \top \equiv \top$; $\quad P \wedge \bot \equiv \bot$

Remark 1.69 (Intuition). $\top$ dominates $\vee$; $\bot$ dominates $\wedge$. A disjunction containing $\top$ is always true; a conjunction containing $\bot$ is always false.

Proposition 1.70 (Negation Laws). $P \vee \neg P \equiv \top$ (Law of Excluded Middle); $\quad P \wedge \neg P \equiv \bot$ (Law of Non-Contradiction)

Remark 1.71 (Logical significance). These are the two fundamental laws of classical logic. Excluded middle ($P \vee \neg P$) is rejected by intuitionists. Non-contradiction is accepted universally. They are duals of each other.

Conditional Equivalences

Proposition 1.72 (Material Implication). $P \to Q \;\equiv\; \neg P \vee Q$

Remark 1.73 (Intuition). The conditional is only false when the antecedent is true and the consequent false. In all other cases including when P is false it is vacuously true. This often surprises newcomers.

Remark 1.74 (Consequence). Material implication eliminates $\to$ in favor of $\neg$ and $\vee$, the first step in NNF conversion.

Proposition 1.75 (Contraposition). $P \to Q \;\equiv\; \neg Q \to \neg P$

Remark 1.76 (Proof strategy). Proof by contraposition: to prove $P \to Q$, it is equivalent to prove $\neg Q \to \neg P$. Choose whichever direction is easier to argue.

Proposition 1.77 (Exportation / Importation). $(P \wedge Q) \to R \;\equiv\; P \to (Q \to R)$

Remark 1.78 (Intuition). A proof from two premises can be restructured as a chain of single-premise implications. This is the basis for currying in functional programming and natural deduction.

Proposition 1.79 (Negation of Conditional). $\neg(P \to Q) \;\equiv\; P \wedge \neg Q$

Remark 1.80 (Intuition). The only way a conditional is false is if the antecedent holds but the consequent fails.

Remark 1.81 (Consequence). Used in refutation proofs: to show $P \to Q$, assume $\neg(P \to Q) \equiv P \wedge \neg Q$ and derive a contradiction.

Proposition 1.82 (Biconditional Expansion). $P \leftrightarrow Q \;\equiv\; (P \to Q) \wedge (Q \to P)$

Remark 1.83 (Proof strategy). To prove $P \leftrightarrow Q$, prove both $P \to Q$ and $Q \to P$ separately. This is the standard "iff proof" structure.

The Duality Principle

Definition (Dual Formula)

The dual of φ, written φ^d, is obtained by simultaneously replacing $\wedge \leftrightarrow \vee$ and $\top \leftrightarrow \bot$, leaving all variables and negations unchanged.

Theorem (Duality Principle)

If $\varphi \equiv \psi$ is a logical equivalence involving only $\neg, \wedge, \vee, \top, \bot$, then $\varphi^d \equiv \psi^d$.

Remark 1.84 (Intuition). Every valid equivalence has a "mirror image" obtained by swapping $\wedge \leftrightarrow \vee$ and $\top \leftrightarrow \bot$. De Morgan's laws are each other's duals. Identity and Domination are each other's duals.

Remark 1.85 (Consequence). Duality halves the number of equivalences to memorize. Learn one form; the dual comes for free.

Remark 1.86 (Formally). For any formula φ with variables $P_1, \ldots, P_n$: $\varphi^d \equiv \neg\varphi[\neg P_1/P_1, \ldots, \neg P_n/P_n]$, where $\neg P_i/P_i$ denotes simultaneous substitution of $\neg P_i$ for P_i.

1.1.5 Normal Forms

Normal Forms Quick Reference

Form	Structure	Detail
Literal	P or $\neg P$	Def
Clause (disjunctive)	Disjunction of literals	Def
NNF	Negations only on variables; uses $\neg, \wedge, \vee$ only	Def
CNF	Conjunction of disjunctions of literals $\bigwedge_i \bigvee_j L_{ij}$	Def
DNF	Disjunction of conjunctions of literals $\bigvee_i \bigwedge_j L_{ij}$	Def
NNF existence	Every formula $\equiv$ some NNF formula	Prop
CNF existence	Every formula $\equiv$ some CNF formula	Prop
DNF existence	Every formula $\equiv$ some DNF formula	Prop

Conversion pipeline: Original formula $\xrightarrow{\text{elim } \to,\leftrightarrow}$ NNF $\xrightarrow{\text{distribute } \vee/\wedge}$ CNF or DNF

Definitions (Literal and Clause)

A literal is a propositional variable (P, a positive literal) or the negation of one ($\neg P$, a negative literal).
A (disjunctive) clause is a disjunction of literals: $L_1 \vee L_2 \vee \cdots \vee L_k$.
A conjunctive clause (or term) is a conjunction of literals: $L_1 \wedge L_2 \wedge \cdots \wedge L_k$.

Remark 1.87 (Why literals matter). CNF and DNF are defined in terms of literals, not arbitrary formulas. Literals are the atoms of normal form theory. Every formula in NNF is a tree whose leaves are literals.

Definition (Negation Normal Form NNF)

A formula is in negation normal form (NNF) if:

1. Negation ($\neg$) is applied only to propositional variables (no negations of compound formulas).
2. The only connectives used are $\neg$, $\wedge$, $\vee$.

Proposition 1.88 (Every Formula Has an NNF). Every propositional formula is logically equivalent to a formula in NNF.

Remark 1.89 (Conversion procedure).

1. Eliminate $\rightarrow$: replace $P \rightarrow Q$ with $\neg P \vee Q$.
2. Eliminate $\leftrightarrow$: replace $P \leftrightarrow Q$ with $(P \wedge Q) \vee (\neg P \wedge \neg Q)$.
3. Push $\neg$ inward via De Morgan and double negation: $\neg(P \wedge Q) \rightarrow \neg P \vee \neg Q$; $\neg(P \vee Q) \rightarrow \neg P \wedge \neg Q$; $\neg\neg P \rightarrow P$.

Remark 1.90 (Why NNF first). NNF is the prerequisite for CNF/DNF conversion. Once in NNF, the only remaining structural operations are distributing $\wedge$ over $\vee$ (for CNF) or $\vee$ over $\wedge$ (for DNF).

Definition (Conjunctive Normal Form CNF)

A formula is in conjunctive normal form (CNF) if it is a conjunction of disjunctive clauses:

$$\bigwedge_{i=1}^{m} \bigvee_{j=1}^{n_i} L_{ij}$$

where each L_{ij} is a literal.

Proposition 1.91 (Every Formula Has a CNF). Every propositional formula is logically equivalent to a formula in CNF.

Remark 1.92 (CNF conversion from NNF). Distribute $\vee$ over $\wedge$ repeatedly: $P \vee (Q \wedge R) \equiv (P \vee Q) \wedge (P \vee R)$. Repeat until no $\vee$ sits above a $\wedge$ in the parse tree.

Remark 1.93 (Truth-table method for CNF). For each row where the formula is false, form a disjunctive clause using positive literals for variables assigned F and negative literals for variables assigned T. Take the conjunction of all such clauses.

Remark 1.94 (Consequence resolution). CNF is the required input format for the resolution proof procedure. Every satisfiability solver (SAT solver) works on CNF.

Example 1.95. $(P \vee Q) \wedge (\neg P \vee R) \wedge (Q \vee \neg R)$ is in CNF.

Definition (Disjunctive Normal Form DNF)

A formula is in disjunctive normal form (DNF) if it is a disjunction of conjunctive clauses:

$$\bigvee_{i=1}^{m} \bigwedge_{j=1}^{n_i} L_{ij}$$

where each L_{ij} is a literal.

Proposition 1.96 (Every Formula Has a DNF). Every propositional formula is logically equivalent to a formula in DNF.

Remark 1.97 (DNF conversion from NNF). Distribute $\wedge$ over $\vee$ repeatedly: $P \wedge (Q \vee R) \equiv (P \wedge Q) \vee (P \wedge R)$.

Remark 1.98 (Truth-table method for DNF). For each row where the formula is true, form a conjunctive clause using positive literals for variables assigned T and negative literals for variables assigned F. Take the disjunction of all such clauses.

Remark 1.99 (CNF vs. DNF trade-off). CNF and DNF can both be exponentially larger than the original formula. Neither is always more compact. The choice depends on application: resolution needs CNF; model enumeration benefits from DNF.

1.1.6 Inference Rules

Inference Rules Quick Reference

Rule	Premises	Conclusion	Detail
Modus Ponens (MP)	$P \to Q$, P	Q	
Modus Tollens (MT)	$P \to Q$, $\neg Q$	$\neg P$	
Hypothetical Syllogism	$P \to Q$, $Q \to R$	$P \to R$	
Disjunctive Syllogism	$P \vee Q$, $\neg P$	Q	
Conj. Introduction	P, Q	$P \wedge Q$	
Conj. Elimination	$P \wedge Q$	P (or Q)	
Disj. Introduction	P	$P \vee Q$	
Disj. Elimination	$P \vee Q$, $[P \vdash R]$, $[Q \vdash R]$	R	
Cond. Introduction (CP)	$[P \vdash Q]$	$P \to Q$	
Bicond. Introduction	$P \to Q$, $Q \to P$	$P \leftrightarrow Q$	
Bicond. Elimination	$P \leftrightarrow Q$	$P \to Q$ (or $Q \to P$)	
Neg. Introduction (RAA)	$[P \vdash \bot]$	$\neg P$	
Neg. Elimination	$[\neg P \vdash \bot]$	P	
Double Neg. Elim.	$\neg\neg P$	P	
Constructive Dilemma	$P \vee Q$, $P \to R$, $Q \to S$	$R \vee S$	
Destructive Dilemma	$P \to R$, $Q \to S$, $\neg R \vee \neg S$	$\neg P \vee \neg Q$	

Conditional Rules

Definition 1.100 (Modus Ponens). From a conditional and its antecedent, infer the consequent:

$$\frac{P \to Q \qquad P}{Q}$$

Remark 1.101 (Intuition). "If it rains the ground is wet. It is raining. Therefore, the ground is wet." The most fundamental inference rule. All proof systems for classical logic include MP or an equivalent.

Remark 1.102 (Fully quantified). $\forall P, Q : [(P \to Q) \wedge P] \models Q$.

Remark 1.103 (Common error Affirming the Consequent). From $P \to Q$ and Q, one cannot infer P. This is the fallacy of affirming the consequent.

Definition 1.104 (Modus Tollens). From a conditional and the negation of its consequent, infer the negation of the antecedent:

$$\frac{P \to Q \qquad \neg Q}{\neg P}$$

Remark 1.105 (Intuition). "If it rains the ground is wet. The ground is not wet. Therefore, it did not rain." MT is logically equivalent to applying MP to the contrapositive $\neg Q \to \neg P$.

Remark 1.106 (Common error Denying the Antecedent). From $P \to Q$ and $\neg P$, one cannot infer $\neg Q$.

Definition 1.107 (Hypothetical Syllogism). From two conditionals chained, infer the composed conditional:

$$\frac{P \to Q \qquad Q \to R}{P \to R}$$

Remark 1.108 (Intuition). Transitivity of implication. Chains of reasoning can be composed into a single step.

Remark 1.109 (Consequence). Hypothetical syllogism makes $\to$ transitive, allowing arbitrarily long inference chains to be compressed.

Definition 1.110 (Conditional Introduction (Conditional Proof)). If assuming P in a subproof leads to Q, then discharge the assumption and infer $P \to Q$:

$$\frac{[P \vdash Q]}{P \to Q}$$

Remark 1.111 (Intuition). The standard method for proving a conditional: assume the antecedent, derive the consequent, discharge. This corresponds to exportation: a proof from hypothesis P of Q is exactly a proof of $P \to Q$.

Conjunction Rules

Definition 1.112 (Conjunction Introduction). From two formulas, infer their conjunction:

$$\frac{P \qquad Q}{P \wedge Q}$$

Definition 1.113 (Conjunction Elimination). From a conjunction, infer either conjunct:

$$\frac{P \wedge Q}{P} \qquad \frac{P \wedge Q}{Q}$$

Remark 1.114 (Intuition). Introduction: combining two independent truths. Elimination: extracting a component from a joint assertion.

Disjunction Rules

Definition 1.115 (Disjunction Introduction). From a proposition, infer any disjunction containing it:

$$\frac{P}{P \vee Q}$$

Remark 1.116 (Intuition). "It's raining, so it's raining or snowing." Adding a disjunct weakens the claim.

Definition 1.117 (Disjunction Elimination (Proof by Cases)). From a disjunction, if the same conclusion R follows from each disjunct, infer R:

$$\frac{P \vee Q \qquad [P \vdash R] \qquad [Q \vdash R]}{R}$$

Remark 1.118 (Proof strategy). Case analysis: prove R holds in case P; prove R holds in case Q; since one of P, Q must hold, conclude R.

Biconditional Rules

Definition 1.119 (Biconditional Introduction).

$$\frac{P \to Q \qquad Q \to P}{P \leftrightarrow Q}$$

Definition 1.120 (Biconditional Elimination).

$$\frac{P \leftrightarrow Q}{P \to Q} \qquad \frac{P \leftrightarrow Q}{Q \to P}$$

Remark 1.121 (Proof strategy). Every iff proof has two halves: prove ($\Rightarrow$) and prove ($\Leftarrow$).

Negation Rules

Definition 1.122 (Negation Introduction (Reductio ad Absurdum)). If assuming P leads to contradiction $\bot$, infer $\neg P$:

$$\frac{[P \vdash \bot]}{\neg P}$$

Definition 1.123 (Negation Elimination (Indirect Proof)). If assuming $\neg P$ leads to $\bot$, infer P:

$$\frac{[\neg P \vdash \bot]}{P}$$

Definition 1.124 (Double Negation Elimination).

$$\frac{\neg\neg P}{P}$$

Remark 1.125 (Classical vs. intuitionistic logic). Negation elimination and double negation elimination are classical rules; they fail in intuitionistic logic, where $\neg\neg P \Rightarrow P$ is not provable without the law of excluded middle. All three rules are equivalent to LEM in a suitable sense.

Dilemma Rules

Definition 1.126 (Constructive Dilemma).

$$\frac{P \vee Q \qquad P \to R \qquad Q \to S}{R \vee S}$$

Definition 1.127 (Destructive Dilemma).

$$\frac{P \to R \qquad Q \to S \qquad \neg R \vee \neg S}{\neg P \vee \neg Q}$$

Remark 1.128 (Intuition). Constructive dilemma applies two implications to a disjunction, deriving a disjunction of the two conclusions. Destructive dilemma is its contrapositive form.

1.1.7 Resolution

Resolution Quick Reference

Concept	Key fact	Detail
Resolution rule	From $C_1 \vee L$ and $C_2 \vee \neg L$, derive $C_1 \vee C_2$	Def
Empty clause $\square$	Always false; represents a contradiction	Def
Resolution refutation	Derive $\square$ from Γ proves Γ unsatisfiable	Def
Soundness	$\Gamma \vdash_{\text{res}} \square \Rightarrow \Gamma$ unsat.	Thm
Completeness	Γ unsat. $\Rightarrow \Gamma \vdash_{\text{res}} \square$	Thm
Validity testing	$\Gamma \models \varphi$ iff $\Gamma \cup \{\neg\varphi\}$ derives $\square$	
Strategies	Unit res., set of support, linear, input	

Definition (Resolution Rule)

Given two clauses containing complementary literals L and $\neg L$:

$$\frac{C_1 \vee L \qquad C_2 \vee \neg L}{C_1 \vee C_2}$$

The derived clause $C_1 \vee C_2$ is the resolvent. The literal L is resolved upon (eliminated). C_1, C_2 are (possibly empty) disjunctions of literals.

Remark 1.129 (English reading). Resolution combines two clauses that disagree on exactly one variable, producing a clause that contains everything both clauses said, minus the disagreement. It is the only inference rule in the resolution proof system.

Remark 1.130 (Fully quantified). For all clauses C_1, C_2 and literal L: $(C_1 \vee L) \wedge (C_2 \vee \neg L) \models (C_1 \vee C_2)$. Resolution is sound: the resolvent is a semantic consequence of the parent clauses.

Remark 1.131 (Requires CNF). Resolution applies only to clauses (disjunctions of literals). The input set Γ must be in CNF. Converting to CNF is always possible (by the CNF existence theorem).

Example 1.132. Resolving $(P \vee Q)$ and $(\neg P \vee R)$ on P:

$$\frac{P \vee Q \qquad \neg P \vee R}{Q \vee R}$$

Resolving (P) and $(\neg P)$: derive the empty clause $\square$.

Definition (Empty Clause)

The empty clause, denoted $\square$ (or $\bot$), is the disjunction of no literals. It is always false (unsatisfiable by any truth assignment).

Remark 1.133 (Intuition). Deriving $\square$ is the resolution analogue of deriving a contradiction. Since $\square$ is false under every assignment, its derivation from Γ proves Γ cannot be satisfied.

Definition (Resolution Refutation)

A resolution refutation of a set of clauses Γ is a finite sequence of resolution steps deriving $\square$ from Γ.
If a resolution refutation exists, then Γ is unsatisfiable.

Theorem (Soundness and Completeness of Resolution)

A set of clauses Γ is unsatisfiable if and only if the empty clause $\square$ can be derived from Γ by resolution:

$$\Gamma \text{ unsatisfiable} \iff \Gamma \vdash_{\text{res}} \square$$

Remark 1.134 (Soundness direction ($\Rightarrow$)). If $\square$ is derived, then Γ is unsatisfiable. This follows because each resolution step is semantically valid (preserves consequence), and $\square$ is always false.

Remark 1.135 (Completeness direction ($\Leftarrow$)). If Γ is unsatisfiable, resolution will eventually produce $\square$. Resolution is refutation-complete: it can detect every unsatisfiable set, but it is not designed to enumerate all consequences.

Remark 1.136 (Logical implication). To test $\Gamma \models \varphi$:

1. Negate: $\neg\varphi$.
2. Convert $\Gamma \cup \{\neg\varphi\}$ to CNF.
3. Run resolution. If $\square$ is derived, then $\Gamma \models \varphi$.

Validity reduces to unsatisfiability, which resolution decides.

Remark 1.137 (Resolution strategies). Naive resolution may generate many irrelevant clauses. Strategies restrict which pairs are resolved:

- Unit resolution: Always resolve with a unit clause (single literal).
- Set of support: At least one parent must come from the negated goal (set of support). Keeps the search goal-directed.
- Linear resolution: Each step uses the most recently derived clause as one parent.
- Input resolution: At least one parent is always an original input clause.

All strategies preserve completeness for refutation.

1.1.8 Proof Systems

Proof Systems Quick Reference

Concept	Meaning	Detail
Derivability $\Gamma \vdash \varphi$	φ provable from Γ syntactically	Def
Soundness	$\Gamma \vdash \varphi \Rightarrow \Gamma \models \varphi$	Def
Completeness	$\Gamma \models \varphi \Rightarrow \Gamma \vdash \varphi$	Def
$\vdash \leftrightarrow \models$	For PL: derivable iff semantically valid	Thm
Semantic ($\models$)	Truth-assignment based; model-theoretic	
Syntactic ($\vdash$)	Rule-application based; proof-theoretic	

Definition (Derivability)

φ is derivable from Γ in a proof system $\mathcal{P}$, written $\Gamma \vdash \varphi$, if there exists a finite sequence of formulas ending in φ, each of which is either:

- an element of Γ (an assumption), or
- obtained from earlier formulas by an inference rule of $\mathcal{P}$.

Remark 1.138 (English reading). Derivability is a purely syntactic notion: it depends on the proof system's rules and makes no reference to truth or models. A derivation is a finite certificate that can be checked mechanically, step by step.

Remark 1.139 (Proof systems for propositional logic). Common proof systems include:

- Natural deduction introduction and elimination rules for each connective; subproofs; discharge of assumptions. Closest to how mathematicians actually reason.
- Sequent calculus operates on sequents $\Gamma \vdash \Delta$; left and right rules; useful for proof theory.
- Hilbert systems many axiom schemas, few rules (typically MP only). Compact to define, tedious to use.
- Resolution single rule; requires CNF input; optimized for automated theorem proving.

All systems for classical propositional logic prove the same theorems.

Definitions (Soundness and Completeness)

A proof system is sound if every derivable formula is semantically valid:

$$\Gamma \vdash \varphi \;\Rightarrow\; \Gamma \models \varphi.$$

A proof system is complete if every semantically valid formula is derivable:

$$\Gamma \models \varphi \;\Rightarrow\; \Gamma \vdash \varphi.$$

Remark 1.140 (Why soundness matters). Soundness prevents the proof system from proving false things. Without soundness, a formal proof is worthless it might "prove" $P \wedge \neg P$.

Remark 1.141 (Why completeness matters). Completeness ensures the proof system is powerful enough: every logical truth has a proof. Without completeness, the system might miss valid arguments.

Theorem (Soundness and Completeness of Propositional Logic)

For any standard proof system for classical propositional logic (natural deduction, Hilbert system, sequent calculus, resolution):

$$\Gamma \vdash \varphi \iff \Gamma \models \varphi.$$

Remark 1.142 (Logical significance). This equivalence is the central metatheorem of propositional logic. It means the syntactic and semantic notions of entailment coincide perfectly. A formula is provable if and only if it is true under all models.

Remark 1.143 (Contrast with first-order logic). In first-order logic, completeness still holds (Gödel's completeness theorem, 1930). But for arithmetic (Peano arithmetic), Gödel's incompleteness theorems show that some true statements are unprovable. Propositional logic avoids incompleteness because it is decidable truth tables provide a decision procedure.

Notion	Symbol	Nature
Tautological implication	$P \models_{\text{taut}} Q$	Semantic
Logical consequence	$\Gamma \models \varphi$	Semantic, model-theoretic
Derivability	$\Gamma \vdash \varphi$	Syntactic, proof-theoretic

1.1.9 Functional Completeness

Functional Completeness Quick Reference

Concept	Key fact	Detail
Truth function	$f : \{T, F\}^n \to \{T, F\}$; exactly 2^{2^n} of them	Def
Functional completeness	Every truth function expressible from S	Def
Adequate set	Synonym for functionally complete set	Def
Minimal adequate set	No proper subset is adequate	Def
$\{\neg, \wedge\}$	Adequate expresses $\vee$ via De Morgan	Thm
$\{\neg, \vee\}$	Adequate	Thm
$\{\neg, \to\}$	Adequate	Thm
$\{\uparrow\}$ NAND	Adequate (singleton!)	Thm
$\{\downarrow\}$ NOR	Adequate (singleton!)	Thm
$\{\wedge, \vee\}$	Not adequate cannot express $\neg$	Thm
Post's theorem	Iff criterion for adequacy via 5 classes	Thm

n-ary truth functions: $n = 1 \Rightarrow 4$ functions; $n = 2 \Rightarrow 16$; $n = 3 \Rightarrow 256$

Definition (Truth Function)

An n-ary truth function is a function

$$f : \{\mathsf{T}, \mathsf{F}\}^n \to \{\mathsf{T}, \mathsf{F}\}.$$

There are exactly 2^{2^n} distinct n-ary truth functions.

Remark 1.144 (English reading). A truth function is the abstract input-output behavior of a formula: it maps each possible combination of truth values for n variables to a single output. Two formulas with the same truth table compute the same truth function.

Remark 1.145 (Counting). For n inputs, there are 2^n possible input rows. Each row independently maps to T or F, giving 2^{2^n} total functions. For $n = 1$: 4 functions (constant-T, identity, negation, constant-F). For $n = 2$: 16 functions, including $\wedge, \vee, \to, \leftrightarrow, \oplus, \uparrow, \downarrow$, and 9 others.

Remark 1.146 (Consequence). Functional completeness asks: does a set of connectives generate all 2^{2^n} truth functions for every n? If yes, no expressible truth function is beyond the reach of the set.

Example 1.147. The 4 unary truth functions for input P:

P	$\bot$	P	$\neg P$	$\top$
T	F	T	F	T
F	F	F	T	T

Definition (Functional Completeness)

A set S of logical connectives is functionally complete (or adequate) if for every $n \geq 1$ and every n-ary truth function f, there exists a formula φ built using only connectives from S whose truth table computes f.

Remark 1.148 (Intuition). S is adequate if the logic built from S can describe every possible truth-functional relationship. No truth function escapes it.

Remark 1.149 (Fully quantified). S is functionally complete $\iff \forall n \geq 1,\ \forall f : \{T, F\}^n \to \{T, F\},\ \exists \varphi \in \mathsf{WFF}(S),\ \forall v : \{T, F\}^n,\ v(\varphi) = f(v)$.

Definition (Adequate Set / Minimal Adequate Set)

An adequate set is a functionally complete set. A minimal adequate set is an adequate set no proper subset of which is adequate.

Theorem (Standard Functionally Complete Sets)

The following sets of connectives are functionally complete:

1. $\{\neg, \wedge\}$
2. $\{\neg, \vee\}$
3. $\{\neg, \to\}$
4. $\{\bot, \to\}$ (constant false + conditional)
5. $\{\neg, \wedge, \vee, \to, \leftrightarrow\}$ (all standard connectives)

Remark 1.150 (Proof sketch for $\{\neg, \wedge\}$). Every truth function can be expressed in DNF using $\neg$, $\wedge$, $\vee$. Since $P \vee Q \equiv \neg(\neg P \wedge \neg Q)$ (De Morgan), $\vee$ is definable from $\{\neg, \wedge\}$. Therefore $\{\neg, \wedge\}$ expresses all DNF formulas, hence all truth functions.

Remark 1.151 (Why $\{\neg, \wedge\}$ is minimal). Remove either element: $\{\wedge\}$ alone cannot express negation (it preserves falsity); $\{\neg\}$ alone cannot combine propositions (it never increases the number of variables). Neither singleton is adequate. So $\{\neg, \wedge\}$ has no redundant member.

Theorem (NAND is Functionally Complete)

$\{\uparrow\}$ (NAND alone) is functionally complete.

Theorem (NOR is Functionally Complete)

$\{\downarrow\}$ (NOR alone) is functionally complete.

Remark 1.152 (Consequence circuit design). NAND and NOR are each singleton adequate sets. In digital hardware, any logic circuit can be built entirely from NAND gates or entirely from NOR gates. This is why NAND/NOR are called universal gates.

Theorem (Non-Adequate Sets)

The following sets are not functionally complete:

1. $\{\wedge, \vee\}$ cannot express $\neg$ (or any function that is false when all inputs are true).
2. $\{\to\}$ alone cannot express functions false when all inputs are false.
3. $\{\neg\}$ alone unary only; cannot build any binary function.
4. $\{\leftrightarrow, \neg\}$ cannot express $\wedge$ (this set generates only affine/XOR-based functions under close examination).

Remark 1.153 (How to prove non-adequacy). Identify a closure property shared by all functions in S that is not shared by some target truth function. Common properties: truth-preserving ($f(T, \ldots, T) = T$), falsity-preserving ($f(F, \ldots, F) = F$), monotonic, affine (XOR-representable), self-dual. If all members of S share the property, S cannot express functions lacking it. For $\{\wedge, \vee\}$: both are truth-preserving and falsity-preserving, so the set cannot express $\neg$ (which is neither).

Theorem (Post's Functional Completeness Theorem)

A set S of Boolean functions is functionally complete if and only if, for each of the following five Post classes, S contains at least one function not in that class:

1. $\mathbf{T_0}$: Truth-preserving $f(T, \ldots, T) = T$.
2. $\mathbf{T_1}$: Falsity-preserving $f(F, \ldots, F) = F$.
3. $\mathbf{M}$: Monotonic $x_i \leq y_i \; \forall i \Rightarrow f(\mathbf{x}) \leq f(\mathbf{y})$ (where $F < T$).
4. $\mathbf{A}$: Affine (linear) f expressible as $a_0 \oplus a_1 x_1 \oplus \cdots \oplus a_n x_n$.
5. $\mathbf{S}$: Self-dual $f(x_1, \ldots, x_n) = \neg f(\neg x_1, \ldots, \neg x_n)$.

Remark 1.154 (How to apply Post's theorem). For each of the 5 classes, check whether S contains a function outside that class. If yes for all 5, S is adequate. If S misses any class entirely (all members of S lie within some class), S is not adequate.

Remark 1.155 (Logical significance). Post's theorem (Emil Post, 1941) gives a complete, finite, mechanical criterion for adequacy. It is one of the deepest results of classical propositional logic, characterizing the entire lattice of clones of Boolean functions.

Example 1.156 (Applying Post's theorem to $\{\neg, \wedge\}$).

Class	$\neg \in$ class?	$\wedge \in$ class?	Some member outside?
T_0 truth-pres.	No ($\neg T = F$)	Yes	Yes ($\neg$)
T_1 falsity-pres.	No ($\neg F = T$)	Yes	Yes ($\neg$)
M monotone	No	Yes	Yes ($\neg$)
A affine	Yes ($\neg P = T \oplus P$)	No	Yes ($\wedge$)
S self-dual	Yes	No	Yes ($\wedge$)

All 5 classes have a member of S outside them. By Post's theorem, $\{\neg, \wedge\}$ is functionally complete. ✓

1.1.10 Compactness Theorem

Compactness Quick Reference

Concept	Key fact	Detail
Compactness theorem	Γ sat. $\iff$ every finite $\Gamma_0 \subseteq \Gamma$ sat.	Thm
Consequence form	$\Gamma \models \varphi \Rightarrow$ some finite $\Gamma_0 \models \varphi$	Cor
Graph coloring app.	Infinite graph k-colorable iff all finite subgraphs are	Ex
Proof methods	Via completeness; direct construction; ultraproducts	

Theorem (Compactness of Propositional Logic)

A set of formulas Γ is satisfiable if and only if every finite subset of Γ is satisfiable.
Equivalently: if Γ is unsatisfiable, some finite subset of Γ is already unsatisfiable.

Remark 1.157 (English reading). Compactness says that satisfiability is a finitary property: an infinite set of premises can only be contradictory if some finite fragment is already contradictory. You can never need infinitely many premises to derive a contradiction; finitely many always suffice.

Remark 1.158 (Fully quantified form). $\forall \Gamma \subseteq \mathsf{WFF} : \big(\exists v,\ v \models \Gamma\big) \iff \big(\forall \Gamma_0 \subseteq_{\text{fin}} \Gamma,\ \exists v,\ v \models \Gamma_0\big)$.

Remark 1.159 (Which direction is trivial). ($\Rightarrow$): If v satisfies Γ, it satisfies every subset. Trivial.

($\Leftarrow$): The substantive direction. If every finite subset is satisfiable, is the whole set? Intuition says "yes" but it is not obvious an infinite set might encode constraints that collectively rule out every assignment even though no finite fragment does.

Remark 1.160 (Why this is a metatheorem). Compactness is a statement about the entire logic, not about any particular formula. It belongs to metatheory it tells us something about the structure of propositional logic itself.

Remark 1.161 (Proof methods). Three standard approaches:

1. Via completeness: If every finite subset of Γ is satisfiable, no finite subset derives $\bot$. By completeness, Γ does not derive $\bot$. By soundness-completeness, Γ is consistent, hence satisfiable. (Circular-looking but rigorous with careful bookkeeping.)
2. Direct construction (König's lemma): Build a satisfying assignment by extending partial assignments level by level through all formulas in Γ, using compactness of $\{T, F\}^\omega$ (Tychonoff).
3. Ultraproducts: Take an ultraproduct of the satisfying assignments for all finite subsets.

Corollary 1.162 (Compactness of Logical Consequence). If $\Gamma \models \varphi$, then there exists a finite subset $\Gamma_0 \subseteq \Gamma$ with $\Gamma_0 \models \varphi$.

Equivalently: logical consequence from an infinite premise set always depends on only finitely many premises.

Remark 1.163 (Practical implication). In any propositional proof, one uses only finitely many formulas from Γ. Compactness tells us this is not a limitation we never need infinitely many premises for a single conclusion.

Example 1.164 (Graph Coloring). Claim: An infinite graph G is k-colorable if and only if every finite subgraph of G is k-colorable.

Encoding: For each vertex v and color $i \in \{1, \ldots, k\}$, introduce propositional variable $C_{v,i}$ ("vertex v has color i"). Add formulas:

- For each v: $C_{v,1} \vee \cdots \vee C_{v,k}$ (every vertex has some color).
- For each v, distinct i, j: $\neg(C_{v,i} \wedge C_{v,j})$ (at most one color).
- For each edge (u, v), each i: $\neg(C_{u,i} \wedge C_{v,i})$ (adjacent vertices differ).

Each finite subset of this formula set involves only finitely many vertices a finite subgraph. If every finite subgraph is k-colorable, every finite subset is satisfiable. By compactness, the full set is satisfiable, giving a k-coloring of G.

Remark 1.165 (Why compactness is surprising here). The graph coloring argument works even for uncountably infinite graphs. Compactness is doing real work: it converts local, finite satisfiability into global satisfiability for a system with infinitely many variables.

1.1.11 Craig's Interpolation Theorem

Interpolation Quick Reference

Concept	Key fact	Detail
Common language	Variables appearing in both φ and ψ	Def
Interpolant	θ s.t. $\varphi \models \theta \models \psi$ with vars in both	Thm
Craig's theorem	If $\varphi \models \psi$ (shared var), interpolant exists	Thm
Trivial cases	$\varphi \equiv \bot$: use $\bot$; $\psi \equiv \top$: use $\top$	
Significance	Modularity of reasoning; Beth definability; program verification	

Definition (Common Language)

The common language of formulas φ and ψ is the set of propositional variables appearing in both φ and ψ:

$$\mathsf{Var}(\varphi) \cap \mathsf{Var}(\psi).$$

Remark 1.166 (Intuition). Two formulas may talk about overlapping but distinct vocabularies. The common language is the shared vocabulary the concepts both formulas discuss. Interpolation says: when φ forces ψ, the reason can always be expressed entirely in the shared vocabulary.

Theorem (Craig's Interpolation Theorem)

Let φ and ψ be propositional formulas with $\varphi \models \psi$ and $\mathsf{Var}(\varphi) \cap \mathsf{Var}(\psi) \neq \emptyset$.
Then there exists a formula θ (an interpolant) such that:

1. $\varphi \models \theta$
2. $\theta \models \psi$
3. $\mathsf{Var}(\theta) \subseteq \mathsf{Var}(\varphi) \cap \mathsf{Var}(\psi)$

Remark 1.167 (English reading). If φ entails ψ, there is an intermediate formula θ that: (1) follows from φ; (2) implies ψ; and (3) uses only variables common to both. The interpolant θ captures exactly the logically relevant content that φ passes to ψ, in their shared vocabulary.

Remark 1.168 (Fully quantified form). $\forall \varphi, \psi \in \mathsf{WFF}$: if $\varphi \models \psi$ and $\mathsf{Var}(\varphi) \cap \mathsf{Var}(\psi) \neq \emptyset$, then $\exists \theta \in \mathsf{WFF}$ such that $\varphi \models \theta$, $\theta \models \psi$, and $\mathsf{Var}(\theta) \subseteq \mathsf{Var}(\varphi) \cap \mathsf{Var}(\psi)$.

Remark 1.169 (Trivial cases).
- If φ is a contradiction, take $\theta = \bot$ (any ψ follows from $\bot$, and $\bot$ uses no variables).
- If ψ is a tautology, take $\theta = \top$ ($\varphi \models \top$ always, and $\top \models \psi$ when ψ is a tautology).

- If $\mathsf{Var}(\varphi) \cap \mathsf{Var}(\psi) = \emptyset$ and $\varphi \models \psi$, then φ must be a contradiction or ψ a tautology (since no shared variable can transmit information). The hypothesis requires at least one shared variable to avoid this.

Example 1.170. Let $\varphi = P \wedge Q$ and $\psi = P \vee R$. Then $\varphi \models \psi$ (since $P \wedge Q$ implies P, which implies $P \vee R$). Shared variable: P.

Interpolant: $\theta = P$, since $P \wedge Q \models P$ and $P \models P \vee R$, and $\mathsf{Var}(\theta) = \{P\} \subseteq \{P, Q\} \cap \{P, R\}$. ✓

Example 1.171. Let $\varphi = P \to Q$ and $\psi = \neg Q \to \neg P$ (contrapositive). They are logically equivalent, so $\varphi \models \psi$. Shared variables: $\{P, Q\}$.

Interpolants include $\theta = P \to Q$, $\theta = \neg P \vee Q$, or $\theta = \top$ (since ψ is a tautology given $\varphi \equiv \psi$).

Remark 1.172 (Logical and mathematical significance). Craig's theorem (William Craig, 1957) has several important consequences:

1. Modularity of reasoning: Whenever $\varphi \models \psi$, the "reason" can be expressed using only the shared vocabulary. This justifies modular reasoning in large formal systems you never need to bring in private vocabulary of one module to pass a conclusion to another.
2. Beth definability: A predicate is implicitly definable in a theory if and only if it is explicitly definable. In propositional terms: if a variable is determined by the rest of a formula, an explicit definition exists. Interpolation is the key proof tool for this.
3. Constructive proof via sequent calculus: Interpolation can be proved constructively by tracing a cut-elimination proof in the sequent calculus. The interpolant is built simultaneously with the proof.
4. Program verification and model checking: Interpolants are used to compute abstractions of reachability proofs in software model checking. Given a failing execution trace, an interpolant separates initial states from bad states.

Remark 1.173 (Interpolation in first-order logic). Craig's theorem extends to first-order logic: if $\varphi \models \psi$ in FOL, an interpolant exists in their common language. The proof is harder, requiring cut elimination in first-order sequent calculus. In higher-order logic, interpolation can fail.

1.1.12 Common Errors and Fallacies

Errors and Fallacies Quick Reference

Error	Form	Valid?	Detail
Modus Ponens	$P \to Q,\ P\ \therefore Q$	Yes	
Modus Tollens	$P \to Q,\ \neg Q\ \therefore \neg P$	Yes	
Affirming Consequent	$P \to Q,\ Q\ \therefore P$	No	
Denying Antecedent	$P \to Q,\ \neg P\ \therefore \neg Q$	No	
Conditional misread	$P \to Q$ is false when P true, Q false only		
Inclusive vs. exclusive $\vee$	Logic uses inclusive or		
Missing rows	n vars $\Rightarrow 2^n$ rows needed		
Precedence errors	$\neg P \wedge Q \neq \neg(P \wedge Q)$		

Formal Fallacies

Definition (Affirming the Consequent)

The following inference is invalid:

$$\frac{P \to Q \qquad Q}{P \quad \text{(INVALID)}}$$

Remark 1.174 (Why it fails). $P \to Q$ says: whenever P is true, Q must be true. It says nothing about what happens when Q is true. Q may hold for reasons entirely unrelated to P. Countermodel: $P = F$, $Q = T$. Then $P \to Q$ is true and Q is true, but P is false.

Remark 1.175 (English example). "If it rains, the ground is wet. The ground is wet. Therefore, it rained." Invalid: a sprinkler could explain the wet ground.

Remark 1.176 (Fully quantified diagnostic). $\not\models [(P \to Q) \wedge Q] \to P$. Check: set $P = F$, $Q = T$. The antecedent is $T \wedge T = T$; the consequent is F. The whole formula is false under this assignment. Not a tautology.

Definition (Denying the Antecedent)

The following inference is invalid:

$$\frac{P \to Q \qquad \neg P}{\neg Q \quad \text{(INVALID)}}$$

Remark 1.177 (Why it fails). $P \to Q$ only guarantees Q when P holds. When P is false, Q may be true or false the conditional is silent. Countermodel: $P = F$, $Q = T$.

Remark 1.178 (English example). "If it rains, the ground is wet. It did not rain. Therefore, the ground is not wet." Invalid: a sprinkler could still make the ground wet.

Valid Inference Patterns for Comparison

Remark 1.179. The two valid patterns involving a conditional:

Pattern	Form	Valid?	Why
Modus Ponens	$P \to Q,\ P\ \therefore Q$	Yes	$\models [(P \to Q) \wedge P] \to Q$
Modus Tollens	$P \to Q,\ \neg Q\ \therefore \neg P$	Yes	Contrapositive of MP
Affirming Consequent	$P \to Q,\ Q\ \therefore P$	No	Countermodel: $P = F, Q = T$
Denying Antecedent	$P \to Q,\ \neg P\ \therefore \neg Q$	No	Countermodel: $P = F, Q = T$

The two fallacies are easy to confuse because they look like the valid patterns with one element swapped.

Common Truth Table Errors

Remark 1.180 (Misreading the conditional). The conditional $P \to Q$ is false only in the case $(P, Q) = (T, F)$. In all other cases it is true, including when P is false. This "vacuous truth" is counterintuitive but essential. The mistake is treating the conditional as false whenever the antecedent is false.

Remark 1.181 (Confusing $\to$ with $\leftrightarrow$). $P \to Q$ is not symmetric. $P \to Q$ and $Q \to P$ are different (they are converses). Treating a conditional as a biconditional is a form of affirming the consequent.

Remark 1.182 (Inclusive vs. exclusive or). In formal logic, $\vee$ is inclusive: $P \vee Q$ is true when both P and Q are true. Natural language "or" is often exclusive ("tea or coffee" usually means one but not both). When formalizing English, always check which reading is intended.

Remark 1.183 (Forgetting rows). A formula with n distinct variables has 2^n rows. For 3 variables: 8 rows. Omitting rows leads to incorrect classification (e.g., calling a contingency a tautology because you only checked favorable assignments).

Remark 1.184 (Operator precedence errors). • $\neg P \wedge Q$ means $(\neg P) \wedge Q$, not $\neg(P \wedge Q)$.

- $P \vee Q \wedge R$ means $P \vee (Q \wedge R)$ if $\wedge$ binds tighter.
- When in doubt: add explicit parentheses.

Fallacy Checklist

Potential error	Diagnostic question
Affirming the consequent	Was Q used to infer P from $P \to Q$?
Denying the antecedent	Was $\neg P$ used to infer $\neg Q$ from $P \to Q$?
Conditional misread	Was the conditional treated as false when the antecedent is false?
Confusing direction	Was $P \to Q$ treated as equivalent to $Q \to P$?
Missing rows	Were all 2^n truth-table rows considered?
Precedence error	Were all connectives properly parenthesized before evaluating?
Inclusive/exclusive confusion	Was $\vee$ read as exclusive when inclusive was intended?

1.1.13 Summary Tables

All Connectives Truth Tables

P	Q	$\neg P$	$P \wedge Q$	$P \vee Q$	$P \to Q$	$P \leftrightarrow Q$	$P \oplus Q$	$P \uparrow Q$
T	T	F	T	T	T	T	F	F
T	F	F	F	T	F	F	T	T
F	T	T	F	T	T	F	T	T
F	F	T	F	F	T	T	F	T

$P \downarrow Q = \neg(P \vee Q)$: T only in the F, F row.

Formula Classification

Class	True under	Satisfiable?	Example
Tautology	All assignments	Yes (trivially)	$P \vee \neg P$
Contradiction	No assignment	No	$P \wedge \neg P$
Contingency	Some, not all	Yes	$P \to Q$
Satisfiable	≥ 1 assignment	(broader category)	P, $P \vee Q$

Remark 1.185 (Key relationship). Satisfiable = not a contradiction. Every tautology is satisfiable; every contingency is satisfiable. The three primary classes (tautology, contradiction, contingency) are mutually exclusive and exhaustive.

Logical Equivalences Complete Reference

Name	Equivalence
Double Negation	$\neg\neg P \equiv P$
De Morgan (1)	$\neg(P \wedge Q) \equiv \neg P \vee \neg Q$
De Morgan (2)	$\neg(P \vee Q) \equiv \neg P \wedge \neg Q$
Commutativity	$P \wedge Q \equiv Q \wedge P; \quad P \vee Q \equiv Q \vee P$
Associativity	$(P \wedge Q) \wedge R \equiv P \wedge (Q \wedge R)$ (and for $\vee$)
Distributivity (1)	$P \wedge (Q \vee R) \equiv (P \wedge Q) \vee (P \wedge R)$
Distributivity (2)	$P \vee (Q \wedge R) \equiv (P \vee Q) \wedge (P \vee R)$
Idempotence	$P \wedge P \equiv P; \quad P \vee P \equiv P$
Absorption	$P \wedge (P \vee Q) \equiv P; \quad P \vee (P \wedge Q) \equiv P$
Identity	$P \wedge \top \equiv P; \quad P \vee \bot \equiv P$
Domination	$P \vee \top \equiv \top; \quad P \wedge \bot \equiv \bot$
Excluded Middle	$P \vee \neg P \equiv \top$
Non-Contradiction	$P \wedge \neg P \equiv \bot$
Material Implication	$P \to Q \equiv \neg P \vee Q$
Contraposition	$P \to Q \equiv \neg Q \to \neg P$
Exportation	$(P \wedge Q) \to R \equiv P \to (Q \to R)$
Neg. Conditional	$\neg(P \to Q) \equiv P \wedge \neg Q$
Bicond. Expansion	$P \leftrightarrow Q \equiv (P \to Q) \wedge (Q \to P)$
Bicond. Disjunction	$P \leftrightarrow Q \equiv (P \wedge Q) \vee (\neg P \wedge \neg Q)$
Neg. Biconditional	$\neg(P \leftrightarrow Q) \equiv P \oplus Q \equiv P \leftrightarrow \neg Q$

Inference Rules Valid vs. Invalid Patterns

Rule	Form	Valid?
Modus Ponens	$P \to Q,\ P\ \therefore Q$	Yes
Modus Tollens	$P \to Q,\ \neg Q\ \therefore \neg P$	Yes
Hypothetical Syllog.	$P \to Q,\ Q \to R\ \therefore P \to R$	Yes
Disjunctive Syllog.	$P \vee Q,\ \neg P\ \therefore Q$	Yes
Conj. Intro	$P,\ Q\ \therefore P \wedge Q$	Yes
Disj. Intro	$P\ \therefore P \vee Q$	Yes
Affirm. Consequent	$P \to Q,\ Q\ \therefore P$	No
Denying Antecedent	$P \to Q,\ \neg P\ \therefore \neg Q$	No

Functionally Complete Sets

Set	Adequate?	Remark
$\{\neg, \wedge\}$	Yes	Minimal
$\{\neg, \vee\}$	Yes	Minimal
$\{\neg, \to\}$	Yes	Minimal
$\{\bot, \to\}$	Yes	Minimal
$\{\uparrow\}$ (NAND)	Yes	Minimal singleton
$\{\downarrow\}$ (NOR)	Yes	Minimal singleton
$\{\neg, \wedge, \vee, \to, \leftrightarrow\}$	Yes	Not minimal
$\{\wedge, \vee\}$	No	Cannot express $\neg$
$\{\to\}$	No	Cannot express falsity-preserving functions
$\{\neg\}$	No	Unary only
$\{\neg, \leftrightarrow\}$	No	Cannot express $\wedge$

Key Metatheorems

Theorem	Statement
Soundness	If $\Gamma \vdash \varphi$ then $\Gamma \models \varphi$
Completeness	If $\Gamma \models \varphi$ then $\Gamma \vdash \varphi$
Sound+Complete	$\Gamma \vdash \varphi \iff \Gamma \models \varphi$
Compactness	Γ satisfiable $\iff$ every finite $\Gamma_0 \subseteq \Gamma$ is satisfiable
Craig's Interp.	If $\varphi \models \psi$ (shared vars), $\exists \theta$ with $\varphi \models \theta \models \psi$ and $\mathsf{Var}(\theta) \subseteq \mathsf{Var}(\varphi) \cap \mathsf{Var}(\psi)$
Post's Theorem	S adequate $\iff$ for each of 5 Post classes, some member of S lies outside it
Unique Readability	Every wff has a unique parse tree

Propositional Logic vs. Predicate Logic

Aspect	Propositional Logic	Predicate Logic
Atomic formulas	Propositional variables (P, Q, R)	Predicate symbols applied to terms
Internal structure	None	Terms, variables, functions
Quantifiers	None	$\forall$, $\exists$
Semantics	Truth assignments $v : \mathsf{Prop} \to \{T, F\}$	Structures + variable assignments
Decision problem	Decidable (truth tables, $O(2^n)$)	Undecidable in general
Compactness	Yes	Yes (Gödel, 1930)
Interpolation	Yes	Yes
Soundness/Completeness	Yes	Yes (Gödel, 1930)
Incompleteness	No (decidable)	Yes for arithmetic (Gödel, 1931)

1.2 Proofs

1.3 Capstone

1.4 Capstone Assessment: Propositional Logic

Purpose. This capstone assesses mastery of propositional logic as a formal system. All problems must be solved using only propositional reasoning: logical connectives, equivalence, implication, and formal proof techniques. No quantifiers or set-theoretic arguments are permitted.

Instructions. For each problem, write a complete and rigorous proof. You may use truth tables, equivalence transformations, or natural deduction, but your method must be explicit and logically justified.

Problem 1 Nontrivial Implication Structure

Prove that the following implication is logically valid:

$$\big[(p \to q) \land (r \to s) \land (p \lor r)\big] \;\to\; (q \lor s).$$

Your proof must make explicit where each assumption is used.

Problem 2 Hidden Equivalence

Prove that the following two formulas are logically equivalent:

$$(p \to q) \lor (q \to p) \quad \text{and} \quad (p \leftrightarrow q) \lor (\neg p \land \neg q).$$

You may not appeal to intuition; your argument must rely on formal equivalence rules or semantic reasoning.

Problem 3 Disguised Contradiction

Show that the formula

$$(p \land (p \to q)) \land \neg q$$

is a contradiction.

Your proof must explicitly identify the source of inconsistency.

Problem 4 Implication Elimination Challenge

Prove that the formula

$$(p \to (q \to r))$$

is logically equivalent to

$$(p \wedge q) \to r.$$

Your proof must use only propositional equivalences and must not rely on informal reasoning.

Problem 5 Semantic Consequence

Prove that the set of formulas

$$\{\, p \to q,\ q \to r,\ \neg r \,\}$$

logically implies $\neg p$.

That is, show that every valuation that makes all three premises true must also make $\neg p$ true.

Completion Criterion. You pass the Propositional Logic capstone if all five proofs are:

- logically correct,
- explicitly justified at each step, and
- written without appeal to informal semantic intuition.

Successful completion certifies readiness to proceed to Predicate Calculus.

Chapter 2

Predicate Logic

2.1 Notes

Where You Are in the Journey

Propositional Logic $\to$ Predicate Calculus $\to$ Sets & Functions $\to$ Proof Techniques $\to$ Axiom Systems $\to$ $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$ $\to$ Real Analysis $\to$ $\cdots$

How we got here. Propositional logic gave us a language for combining truth values with connectives. But it cannot express "every integer has a successor" or "there exists a prime greater than 100" — statements that quantify over objects. Predicate logic provides the missing machinery.

What this chapter builds. We extend propositional syntax with variables, constants, function symbols, predicate symbols, and quantifiers. The semantics are now structures: a domain of objects plus interpretations of every symbol. The proof theory adds quantifier introduction and elimination rules. Soundness and completeness (Gödel's theorem) are the major metatheorems.

Where this leads. Every subsequent chapter uses predicate logic as its underlying language. The axiomatic construction of $\mathbb{N}$ is a predicate-logic theory. Set theory, ring theory, and topology are all first-order theories.

Structural Roadmap

The same four-layer architecture as propositional logic applies:

$$\text{Syntax} \longrightarrow \text{Semantics} \longrightarrow \text{Proof Theory} \longrightarrow \text{Metatheory}$$

The global progression is:

1. Syntax: terms, atomic formulas, well-formed formulas, free and bound variables, substitution

2. Semantics: structures, variable assignments, satisfaction, validity, logical equivalence, substitution lemmas
3. Quantifiers: universal, existential, unique existential, bounded quantifiers, negation, commutation, prenex normal form
4. Proof theory: quantifier rules (UI, UG, EI, EG), equality, soundness and completeness
5. Translation: English to predicate logic, scope, square of opposition
6. Reference: quantifier fallacies, comparison tables

Remark 2.1 (Primary source). The primary driver is Bjørndahl's Logic and Proof, supplemented by Gerstein's Introduction to Mathematical Structures and Proofs.

2.1.1 Syntax of First-Order Logic: Terms and Formulas

Terms and Formulas Quick Reference

Concept	Meaning	Detail
Variable	Symbol ranging over domain elements	Def
Term	Syntactic object denoting a domain element	Def
Atomic formula	Predicate applied to terms	Def
Well-formed formula	Recursively constructed formula	Def
Molecular formula	Non-atomic wff	Def

Definition (Variable)

A variable is a syntactic symbol that ranges over elements of a fixed domain of discourse. Variables serve as placeholders in formulas and do not refer to specific objects until they are assigned values or bound by quantifiers.

Remark 2.2 (English reading). Variables are the unknowns of predicate logic. They pick out no particular object on their own; their value is supplied either by a variable assignment (semantics) or by a quantifier (syntax).

Remark 2.3 (Consequence). Because variables have no fixed denotation, the truth of a formula containing free variables depends on what values are assigned to those variables. This dependence is tracked formally by the variable assignment function s.

Definition (Term)

A term is a syntactic expression intended to denote an object in the domain of discourse.
The set of terms of a formal language is defined recursively:

1. Variables. Every variable is a term.
2. Constants. Every constant symbol is a term.
3. Function application. If f is an n-ary function symbol and $t_1, \ldots, t_n$ are terms, then $f(t_1, \ldots, t_n)$ is a term.
4. Closure. No expression is a term unless it can be obtained by finitely many applications of rules (1)–(3).

Remark 2.4 (English reading). Terms are the noun phrases of predicate logic: they name (or describe) objects in the domain. Constants name fixed objects; variables name arbitrary ones; function symbols build complex names from simpler ones.

Remark 2.5 (Fully quantified form). Terms are purely syntactic objects. Under an interpretation, each term denotes an element of the domain of discourse, but the term itself is not an object of the domain.

Definition (Atomic Formula)

An atomic formula is a well-formed formula obtained by applying an n-ary predicate symbol to n terms.
If P is an n-ary predicate symbol and $t_1, \ldots, t_n$ are terms, then

$$P(t_1, \ldots, t_n)$$

is an atomic formula.
Atomic formulas contain no logical connectives or quantifiers and serve as the base case for the recursive definition of well-formed formulas.

Remark 2.6 (English reading). Atomic formulas are the simplest complete statements: they assert that a predicate (property or relation) holds of specific objects. They are the predicate-logic counterpart of propositional variables.

Definition (Well-Formed Formula)

The set of well-formed formulas (wffs) of a first-order language is defined recursively:

1. Atomic formulas. Every atomic formula is a well-formed formula.
2. Negation. If φ is a formula, then $\neg\varphi$ is a formula.
3. Binary connectives. If φ and ψ are formulas and $\circ \in \{\wedge, \vee, \rightarrow, \leftrightarrow\}$, then $(\varphi \circ \psi)$ is a formula.
4. Quantification. If φ is a formula and x is a variable, then $\forall x\, \varphi$ and $\exists x\, \varphi$ are formulas.
5. Closure. No expression is a formula unless it can be obtained by finitely many applications of rules (1)–(4).

Remark 2.7 (English reading). The wff definition says exactly which strings of symbols count as grammatical sentences of first-order logic. Everything in the language is built bottom-up from atomic formulas by the four formation rules.

Remark 2.8 (Distinction from propositional logic). Rule (4) is new: quantifiers bind variables and create formulas from formulas. This extra layer is what makes predicate logic more expressive than propositional logic.

Definition (Molecular Formula)

A molecular formula is a well-formed formula that is not atomic.
Equivalently, a formula is molecular if it is formed from one or more atomic formulas by the application of logical connectives or quantifiers.

Remark 2.9 (English reading). Atomic formulas express basic properties or relations. Molecular formulas express compound statements built from atomic formulas using the logical apparatus of the language.

2.1.2 Syntax: Free Variables, Bound Variables, and Substitution

Variables and Substitution Quick Reference

Concept	Meaning	Detail
Scope of a quantifier	Formula governed by the quantifier	Def
Bound/free occurrence	Whether x is captured by a quantifier	Def
Free variable $\mathrm{FV}(\varphi)$	Variables whose values affect truth	Def
Sentence	Formula with $\mathrm{FV}(\varphi) = \varnothing$	Def
Substitution $\varphi[t/x]$	Replace free x by term t	Def
Free for substitution	No variable in t becomes bound	Def
Alpha-equivalence	Renaming bound variables	Def

Definition (Scope of a Quantifier)

Let φ be a formula of a first-order language.
If φ is of the form $\forall x\,\psi$ or $\exists x\,\psi$, then the formula ψ is called the scope of the quantifier.
The quantifier is said to bind all occurrences of the variable x that appear within its scope.

Remark 2.10 (English reading). The scope is the reach of a quantifier the subformula it governs. Scope is syntactically determined by the parenthesisation of the formula, not by proximity to the quantifier symbol.

Example 2.11. In $(\forall x\,P(x)) \wedge Q(x)$, the scope of $\forall x$ is $P(x)$ only. The occurrence of x in $Q(x)$ falls outside the scope and is free.

Definition (Bound and Free Occurrences)

An occurrence of a variable x in a formula φ is bound if it lies within the scope of a quantifier $\forall x$ or $\exists x$.
An occurrence of x is free if it is not bound by any quantifier in φ.

Remark 2.12 (English reading). The same variable can have both bound and free occurrences in a single formula. Each occurrence is classified independently by checking whether it falls within the scope of a binding quantifier.

Definition (Free Variables of a Formula)

The set $\mathrm{FV}(\varphi)$ of free variables of a formula φ is defined recursively:

1. If $\varphi = P(t_1, \ldots, t_n)$ is atomic, then $\mathrm{FV}(\varphi) = \bigcup_{i=1}^{n} \mathrm{Var}(t_i)$, where $\mathrm{Var}(t_i)$ is the set of variables in term t_i.
2. $\mathrm{FV}(\neg\varphi) = \mathrm{FV}(\varphi)$.
3. $\mathrm{FV}(\varphi \circ \psi) = \mathrm{FV}(\varphi) \cup \mathrm{FV}(\psi)$ for any binary connective $\circ$.
4. $\mathrm{FV}(\forall x\, \varphi) = \mathrm{FV}(\varphi) \setminus \{x\}$.
5. $\mathrm{FV}(\exists x\, \varphi) = \mathrm{FV}(\varphi) \setminus \{x\}$.

Remark 2.13 (English reading). $\mathrm{FV}(\varphi)$ collects exactly those variables whose values can influence the truth of φ under a structure. Quantifying over x removes x from the free set because the quantifier takes responsibility for ranging over all (or some) values of x.

Remark 2.14 (Common error). Terms themselves contain only free variables. Variables become bound only through quantification in formulas. The interpretation of a formula depends exactly on the values assigned to its free variables.

Definition (Sentence)

A sentence (or closed formula) is a formula with no free variables.
Formally, φ is a sentence if and only if $\mathrm{FV}(\varphi) = \varnothing$.

Remark 2.15 (Consequence). For sentences, truth depends only on the structure, not on the variable assignment. This makes sentences the natural objects to be called true or false in a model, without qualification.

Definition (Substitution Notation)

Let φ be a formula, x a variable, and t a term.
The expression $\varphi[t/x]$ denotes the formula obtained from φ by replacing every free occurrence of x with the term t, leaving bound occurrences of x unchanged.

Definition (Free for Substitution)

A term t is free for x in φ if no free occurrence of x in φ lies within the scope of a quantifier $\forall y$ or $\exists y$ where y is a variable occurring in t.
Equivalently, t is free for x in φ if the substitution $\varphi[t/x]$ does not result in any variable in t becoming bound.

Remark 2.16 (Capture-avoiding substitution). A substitution $\varphi[t/x]$ is admissible only if t is free for x in φ. If this condition is violated, a variable occurring in t may become bound after substitution, silently changing the meaning of the formula. This is called variable capture.

Remark 2.17 (Common error). Variable capture is one of the most common syntactic mistakes in predicate logic. Always check that the term being substituted introduces no variables that fall inside a binding quantifier in the target formula.

Definition (Alpha-Equivalence)

Formulas that differ only by the names of bound variables are logically equivalent. This is called alpha-equivalence (or alphabetic variance):

$$\forall x\, \varphi \;\equiv\; \forall y\, \varphi[y/x] \quad \text{(provided } y \text{ is not free in } \varphi\text{)}.$$

Remark 2.18 (Consequence). Alpha-equivalence is the formal basis for the bound variable renaming used in prenex normal form conversion and in any proof where variable capture must be avoided. Two alpha-equivalent formulas are interchangeable in all contexts.

2.1.3 Syntax: Formula Depth

Formula Depth Quick Reference

Concept	Meaning	Detail
Formula depth	Max formation steps from atomic formulas	Def

Definition (Formula Depth)

The depth (or complexity) of a formula φ, denoted $\text{depth}(\varphi)$, is a natural number defined recursively:

1. Atomic case. If φ is atomic, then $\text{depth}(\varphi) = 0$.
2. Negation. If $\varphi = \neg\psi$, then $\text{depth}(\varphi) = \text{depth}(\psi) + 1$.
3. Binary connectives. If $\varphi = (\psi \circ \chi)$ for $\circ \in \{\wedge, \vee, \to, \leftrightarrow\}$, then
$$\text{depth}(\varphi) = \max\{\text{depth}(\psi), \text{depth}(\chi)\} + 1.$$
4. Quantifiers. If $\varphi = \forall x\, \psi$ or $\varphi = \exists x\, \psi$, then $\text{depth}(\varphi) = \text{depth}(\psi) + 1$.

Remark 2.19 (English reading). The depth of a formula measures the maximum number of logical formation steps required to build the formula from atomic formulas. It corresponds to the height of the formula's syntactic parse tree.

Remark 2.20 (Proof strategy). Formula depth is the standard measure used in structural induction proofs over first-order formulas. The base case is depth 0 (atomic formulas); the inductive step handles each formation rule and assumes the result holds for sub-formulas of strictly smaller depth.

2.1.4 Semantics: Structures and Variable Assignments

Structures Quick Reference

Concept	Meaning	Detail
First-order language $\mathcal{L}$	Signature specifying symbols	Def
Structure $\mathcal{M} = \langle D, I \rangle$	Interpretation of $\mathcal{L}$	Def
Variable assignment s	Function assigning domain elements to variables	Def
Modified assignment $s[x \mapsto d]$	Change s at x only	Def
Term interpretation $[\![t]\!]_{\mathcal{M},s}$	Semantic value of a term	Def

Definition (First-Order Language)

A first-order language $\mathcal{L}$ consists of:

- a set of constant symbols,
- a set of function symbols, each with a specified arity, and
- a set of predicate symbols, each with a specified arity.

Every first-order language also includes a countable set of variables and the logical symbols $\neg, \wedge, \vee, \to, \leftrightarrow, \forall, \exists$, and (optionally) $=$.

Remark 2.21 (English reading). A first-order language is the alphabet of a formal system. It specifies which non-logical names are available (constants, functions, predicates) but says nothing yet about what those names mean. Meaning is supplied by a structure.

Definition (Structure)

Let $\mathcal{L}$ be a first-order language. A structure (or interpretation) for $\mathcal{L}$ is a pair

$$\mathcal{M} = \langle D, I \rangle,$$

where D is a nonempty set called the domain of discourse (or universe), and I is an interpretation function assigning:

- to each constant symbol c: an element $I(c) \in D$,
- to each n-ary function symbol f: a function $I(f) : D^n \to D$,
- to each n-ary predicate symbol P: a relation $I(P) \subseteq D^n$.

Remark 2.22 (English reading). A structure gives the language its meaning. The domain specifies the universe of objects being reasoned about, and the interpretation function assigns each non-logical symbol a concrete mathematical object.

Remark 2.23 (Nonempty domain convention). Throughout classical first-order logic, the domain D is required to be nonempty. Without this, the inference $\forall x\, \varphi \Rightarrow \exists x\, \varphi$ (subalternation) would fail: a universal statement is vacuously true in an empty domain while the existential statement is false.

Remark 2.24 (Source note). Some texts use the term interpretation and others use structure; the two are synonymous. The notation $\mathcal{M}$, $\mathfrak{A}$, or $\mathbf{A}$ all appear in the literature.

Definition (Variable Assignment)

Let $\mathcal{M} = \langle D, I \rangle$ be a structure for $\mathcal{L}$. A variable assignment is a function

$$s : \mathsf{Var} \to D$$

that assigns to each variable an element of the domain.

Remark 2.25 (English reading). A structure assigns meaning to non-logical symbols; a variable assignment assigns meaning to variables. Together they determine the truth value of every formula in the language.

Definition (Modified Assignment)

Let s be a variable assignment, x a variable, and $d \in D$. The modified assignment $s[x \mapsto d]$ is defined by

$$s[x \mapsto d](y) = \begin{cases} d & \text{if } y = x, \\ s(y) & \text{if } y \neq x. \end{cases}$$

Remark 2.26 (Consequence). Modified assignments are the formal device for evaluating quantified formulas. The clause $\mathcal{M}, s \models \forall x\, \varphi$ checks $\mathcal{M}, s[x \mapsto d] \models \varphi$ for every $d \in D$, testing the formula as x ranges over all domain elements.

Definition (Interpretation of a Term)

The interpretation of a term t in $\mathcal{M}$ under s, denoted $[\![t]\!]_{\mathcal{M},s}$, is defined recursively:

1. If t is a variable x: $[\![x]\!]_{\mathcal{M},s} = s(x)$.
2. If t is a constant c: $[\![c]\!]_{\mathcal{M},s} = I(c)$.
3. If $t = f(t_1, \ldots, t_n)$: $[\![f(t_1, \ldots, t_n)]\!]_{\mathcal{M},s} = I(f)\big([\![t_1]\!]_{\mathcal{M},s}, \ldots, [\![t_n]\!]_{\mathcal{M},s}\big)$.

Remark 2.27 (English reading). This is the semantic counterpart of the recursive term definition. Each formation rule for terms has a corresponding clause that computes a domain element from the interpretations of its components.

2.1.5 Semantics: Satisfaction and Truth

Satisfaction and Truth Quick Reference

Concept	Meaning	Detail
Satisfaction $\mathcal{M}, s \models \varphi$	Formula true in $\mathcal{M}$ under s	Def
Truth in a structure $\mathcal{M} \models \varphi$	True for every assignment	Def
Validity / logical truth	True in every structure	Def
Satisfiability	True in some structure	Def

Definition (Satisfaction)

Let $\mathcal{M} = \langle D, I \rangle$ be a structure, s a variable assignment, and φ a formula. The relation $\mathcal{M}, s \models \varphi$ (read "$\mathcal{M}$ satisfies φ under s") is defined recursively:

1. Atomic. $\mathcal{M}, s \models P(t_1, \ldots, t_n) \iff (\llbracket t_1 \rrbracket_{\mathcal{M},s}, \ldots, \llbracket t_n \rrbracket_{\mathcal{M},s}) \in I(P)$.
2. Equality. $\mathcal{M}, s \models (t_1 = t_2) \iff \llbracket t_1 \rrbracket_{\mathcal{M},s} = \llbracket t_2 \rrbracket_{\mathcal{M},s}$.
3. Negation. $\mathcal{M}, s \models \neg\varphi \iff \mathcal{M}, s \not\models \varphi$.
4. Conjunction. $\mathcal{M}, s \models (\varphi \wedge \psi) \iff \mathcal{M}, s \models \varphi$ and $\mathcal{M}, s \models \psi$.
5. Disjunction. $\mathcal{M}, s \models (\varphi \vee \psi) \iff \mathcal{M}, s \models \varphi$ or $\mathcal{M}, s \models \psi$.
6. Implication. $\mathcal{M}, s \models (\varphi \to \psi) \iff \mathcal{M}, s \not\models \varphi$ or $\mathcal{M}, s \models \psi$.
7. Biconditional. $\mathcal{M}, s \models (\varphi \leftrightarrow \psi) \iff (\mathcal{M}, s \models \varphi \Leftrightarrow \mathcal{M}, s \models \psi)$.
8. Universal quantification. $\mathcal{M}, s \models \forall x\, \varphi \iff$ for all $d \in D$, $\mathcal{M}, s[x \mapsto d] \models \varphi$.
9. Existential quantification. $\mathcal{M}, s \models \exists x\, \varphi \iff$ there exists $d \in D$ such that $\mathcal{M}, s[x \mapsto d] \models \varphi$.

Remark 2.28 (English reading). Satisfaction is the bridge between syntax and semantics. It tells us, for each formula formation rule, what it means for the formula to hold. Clauses (1)–(7) mirror the propositional connectives; clauses (8)–(9) are the new quantifier clauses unique to first-order logic.

Remark 2.29 (Fully quantified form for universal quantification). The clause $\mathcal{M}, s \models \forall x\, \varphi$ holds if and only if for every element d of the domain D, the formula φ is satisfied under the assignment that agrees with s everywhere except that it sends x to d. This is what it means for a property to hold universally.

Definition (Truth in a Structure)

A sentence φ is true in a structure $\mathcal{M}$, written $\mathcal{M} \models \varphi$, if $\mathcal{M}, s \models \varphi$ for every variable assignment s.
A sentence φ is false in $\mathcal{M}$ if $\mathcal{M} \not\models \varphi$.

Remark 2.30 (English reading). For sentences (formulas with no free variables), the choice of assignment is irrelevant all assignments agree. So $\mathcal{M} \models \varphi$ is well-defined without reference to any particular s.

Definition (Validity and Satisfiability)

A formula φ is:

- valid (or a logical truth) if $\mathcal{M}, s \models \varphi$ for every structure $\mathcal{M}$ and every assignment s;
- satisfiable if $\mathcal{M}, s \models \varphi$ for some $\mathcal{M}$ and some s;
- unsatisfiable (or a contradiction) if it is not satisfiable.

Remark 2.31 (Consequence). A formula is valid if no structure can make it false. These are the logical

truths of predicate logic the formulas provable by logic alone, without any special assumptions about the domain.

2.1.6 Semantics: Models, Theories, and Logical Consequence

Models and Theories Quick Reference

Concept	Meaning	Detail
Model $\mathcal{M} \models \Gamma$	Structure satisfying all sentences of Γ	Def
Theory	Set of sentences; consistent if it has a model	Def
Logical consequence $\Gamma \models \varphi$	True in every model of Γ	Def
Logical equivalence $\varphi \equiv \psi$	Mutually entailing formulas	Def

Definition (Model)

Let φ be a sentence and $\mathcal{M}$ a structure. We say that $\mathcal{M}$ is a model of φ if $\mathcal{M} \models \varphi$.
More generally, if Γ is a set of sentences, then $\mathcal{M}$ is a model of Γ if $\mathcal{M} \models \gamma$ for every $\gamma \in \Gamma$. In this case, we write $\mathcal{M} \models \Gamma$.

Definition (Theory)

A theory is a set of sentences.
A theory T is satisfiable (or consistent) if it has at least one model.
A theory T is complete if for every sentence φ in the language, either $T \models \varphi$ or $T \models \neg\varphi$.

Remark 2.32 (Source note). In some texts, a theory is required to be closed under logical consequence: if $T \models \varphi$ then $\varphi \in T$. In other texts, a theory is simply any set of axioms from which consequences are derived. Both usages appear in the literature.

Definition (Logical Consequence)

Let Γ be a set of formulas and φ a formula. We say that φ is a logical consequence of Γ, written $\Gamma \models \varphi$, if for every structure $\mathcal{M}$ and every variable assignment s:

$$\text{if } \mathcal{M}, s \models \gamma \text{ for all } \gamma \in \Gamma, \text{ then } \mathcal{M}, s \models \varphi.$$

Equivalently, every model of Γ is a model of φ.

Remark 2.33 (English reading). Logical consequence says: if all the hypotheses in Γ are true in some structure, then φ must be true there too. This is the semantic notion of entailment.

Remark 2.34 (Notation). The notation $\models \varphi$ (empty left-hand side) means φ is valid: it is a logical consequence of the empty set of premises.

Definition (Logical Equivalence)

Two formulas φ and ψ are logically equivalent, written $\varphi \equiv \psi$, if each is a logical consequence of the other:

$$\varphi \equiv \psi \quad \Longleftrightarrow \quad (\varphi \models \psi \text{ and } \psi \models \varphi).$$

Equivalently, φ and ψ have the same truth value in every structure under every variable assignment.

Remark 2.35 (Consequence). Logical equivalence is the semantic counterpart of provable bi-implication. Two logically equivalent formulas are interchangeable in any context without changing truth values.

2.1.7 Semantics: Predicates, Relations, and Substitution Lemmas

Predicates and Substitution Lemmas Quick Reference

Concept	Meaning	Detail
Predicate	Open formula or truth-valued function on domain	Def
Substitution Lemma (terms)	Assign vs. substitute, same result	Lem
Substitution Lemma (formulas)	Syntactic substitution = semantic update	Lem

Definition (Predicate)

A predicate is an expression containing one or more variables that represents a property or relation and becomes a proposition when all its variables are instantiated.
Formally, a predicate has two equivalent views:

1. Syntactic view. A predicate is an open formula $\varphi(x_1, \ldots, x_n)$ that has no truth value until specific objects are substituted for its variables.
2. Semantic view. Given a domain D, an n-ary predicate determines a function $P : D^n \to \{\mathsf{T}, \mathsf{F}\}$.

Remark 2.36 (Predicates vs. relations). A predicate is a syntactic or semantic device: syntactically, an open formula; semantically, a truth-valued function. A relation is a purely set-theoretic object: an n-ary relation on D is a subset $R \subseteq D^n$. Under an interpretation, predicates and relations correspond via $P(a_1, \ldots, a_n)$ is true $\Longleftrightarrow$ $(a_1, \ldots, a_n) \in R$. Thus predicates belong to the language of logic, while relations belong to the structures interpreting that language.

Lemma 2.37 (Substitution Lemma for Terms). Let $\mathcal{M}$ be a structure, s a variable assignment, x a variable, and $d \in D$. For any term t,

$$[\![t]\!]_{\mathcal{M},\, s[x \mapsto d]} = [\![t[d/x]]\!]_{\mathcal{M},\, s}.$$

Remark 2.38 (English reading). Evaluating a term after modifying an assignment is the same as substituting the value directly into the term and then evaluating. This shows that syntactic substitution and semantic update commute for terms.

Lemma 2.39 (Substitution Lemma for Formulas). Let φ be a formula, t a term free for x in φ. Then for any structure $\mathcal{M}$ and assignment s,

$$\mathcal{M}, s \models \varphi[t/x] \quad \Longleftrightarrow \quad \mathcal{M},\, s\big[x \mapsto [\![t]\!]_{\mathcal{M},s}\big] \models \varphi.$$

Remark 2.40 (English reading). Satisfying φ with x replaced by t is the same as satisfying φ under the assignment that sends x to the value of t. This lemma formally connects syntactic substitution with semantic evaluation and is essential for proving the soundness of Universal Instantiation.

Remark 2.41 (Proof strategy). Both lemmas are proved by structural induction on the term (Lemma for terms) or formula (Lemma for formulas). The key cases are variables and quantifiers; the remaining cases follow directly from the inductive hypothesis.

2.1.8 Quantifiers: Universal, Existential, and Bounded

Basic Quantifiers Quick Reference

Symbol	Reading	Meaning	Detail
$\forall x\,\varphi$	For all x, φ	True for every domain element	Def
$\exists x\,\varphi$	There exists x such that φ	True for some domain element	Def
$\exists! x\,\varphi$	There exists exactly one x such that φ	Existence + uniqueness	Def
$\forall x \in A\,\varphi$	For all x in A, φ	Restricted universal	Def
$\exists x \in A\,\varphi$	There exists x in A such that φ	Restricted existential	Def

Definition (Universal Quantifier)

The universal quantifier, denoted $\forall$, is a logical operator that binds a variable and asserts that a formula holds for all elements of the domain of discourse.
If φ is a formula and x is a variable, then $\forall x\,\varphi$ is a formula, read "for all x, φ."

Remark 2.42 (English reading). The universal quantifier expresses a global condition over the entire domain. It does not claim the domain is nonempty on its own, but by the nonempty domain convention, there is always at least one element to instantiate.

Definition (Existential Quantifier)

The existential quantifier, denoted $\exists$, is a logical operator that binds a variable and asserts that a formula holds for at least one element of the domain of discourse.
If φ is a formula and x is a variable, then $\exists x\,\varphi$ is a formula, read "there exists an x such that φ."

Remark 2.43 (English reading). The existential quantifier expresses a local condition. It does not specify which element witnesses the claim; it merely asserts that such a witness exists somewhere in the domain.

Definition (Unique Existential Quantifier)

The unique existential quantifier $\exists!$ asserts that there exists exactly one element satisfying a given property. It is an abbreviation:

$$\exists! x\,\varphi \;:=\; \exists x\,\big(\varphi \wedge \forall y\,(\varphi[y/x] \to y = x)\big),$$

where y is a variable not occurring in φ.
Equivalently,

$$\exists! x\,\varphi \;\equiv\; \exists x\,\varphi \;\wedge\; \forall x\,\forall y\,\big((\varphi \wedge \varphi[y/x]) \to x = y\big).$$

Remark 2.44 (English reading). $\exists!x\,\varphi$ combines existence (at least one x satisfies φ) and uniqueness (at most one such x exists). It is not a primitive quantifier but a convenient abbreviation.

Example 2.45. "There is exactly one even prime number" formalizes as $\exists!x\,(P(x) \wedge E(x))$, where $P(x)$ means "x is prime" and $E(x)$ means "x is even."

Definition (Bounded Quantifiers)

Let A be a set (or a predicate defining a set). The bounded quantifiers are abbreviations:

$$\begin{aligned} \forall x \in A\,\varphi &:= \forall x\,(x \in A \to \varphi), \\ \exists x \in A\,\varphi &:= \exists x\,(x \in A \wedge \varphi). \end{aligned}$$

Remark 2.46 (Asymmetry). Bounded universal uses implication ($\to$), while bounded existential uses conjunction ($\wedge$). This is not arbitrary: if the set A is empty, $\forall x \in A\,\varphi$ is vacuously true (there is no x to violate the implication), while $\exists x \in A\,\varphi$ is false (no witness exists).

Remark 2.47 (Common error). Writing $\forall x \in A\,\varphi$ as $\forall x\,(x \in A \wedge \varphi)$ is wrong: this would assert that every element is in A, not just that every element of A satisfies φ.

Theorem 2.48 (Negation of Bounded Quantifiers). Let A be a set and φ a formula. Then:

$$\begin{aligned} \neg(\forall x \in A\,\varphi) &\equiv \exists x \in A\,\neg\varphi, \\ \neg(\exists x \in A\,\varphi) &\equiv \forall x \in A\,\neg\varphi. \end{aligned}$$

Remark 2.49. The domain restriction $x \in A$ is preserved under negation because it functions as a constraint on which elements are quantified over, not as part of the claim being negated.

2.1.9 Quantifiers: Quantifier Laws

Quantifier Laws Quick Reference

Law	Equivalence	Condition	Detail
Negation (1)	$\neg\forall x\,\varphi \equiv \exists x\,\neg\varphi$		Thm
Negation (2)	$\neg\exists x\,\varphi \equiv \forall x\,\neg\varphi$		Thm
Same-type commutation	$\forall x\,\forall y\,\varphi \equiv \forall y\,\forall x\,\varphi$	same type	Thm
Distribution $\forall/\wedge$	$\forall x\,(\varphi \wedge \psi) \equiv (\forall x\,\varphi) \wedge (\forall x\,\psi)$		Thm
Distribution $\exists/\vee$	$\exists x\,(\varphi \vee \psi) \equiv (\exists x\,\varphi) \vee (\exists x\,\psi)$		Thm
Vacuous $\forall$	$\forall x\,\varphi \equiv \varphi$	$x \notin \mathrm{FV}(\varphi)$	Thm
Renaming	$\forall x\,\varphi \equiv \forall y\,\varphi[y/x]$	y not free in φ	Thm

Proposition 2.50 (Quantifier Negation Laws). For any formula φ:

$$\begin{aligned} \neg\forall x\,\varphi &\equiv \exists x\,\neg\varphi, \\ \neg\exists x\,\varphi &\equiv \forall x\,\neg\varphi. \end{aligned}$$

Remark 2.51 (English reading). To negate a universally quantified statement, swap $\forall$ for $\exists$ and negate the inner formula. To negate an existential, swap $\exists$ for $\forall$ and negate. The negation pushes through the quantifier, flipping it.

Remark 2.52 (Procedure for negating nested quantifiers). When negating a statement with multiple quantifiers, push the negation inward past every quantifier, flipping each one, until it reaches the atomic predicate. For example:

$$\neg(\forall x\, \exists y\, \varphi) \equiv \exists x\, \neg(\exists y\, \varphi) \equiv \exists x\, \forall y\, \neg\varphi.$$

Remark 2.53 (General schema for negating quantified implications). For $Q_1 x_1 \cdots Q_n x_n(\varphi \to \psi)$ where each $Q_i \in \{\forall, \exists\}$:

$$\neg\big(Q_1 x_1 \cdots Q_n x_n(\varphi \to \psi)\big) \;\equiv\; Q_1' x_1 \cdots Q_n' x_n(\varphi \wedge \neg\psi),$$

where $Q_i' = \exists$ when $Q_i = \forall$, and $Q_i' = \forall$ when $Q_i = \exists$.

Remark 2.54 (Common error). Negating only the predicate without flipping the quantifier is incorrect: $\neg(\forall x\, P(x)) \not\equiv \forall x\, \neg P(x)$. The right side asserts that every element fails P, which is much stronger than merely asserting that not every element satisfies P.

Proposition 2.55 (Quantifier Commutation). Quantifiers of the same type commute:

$$\begin{aligned} \forall x\, \forall y\, \varphi &\;\equiv\; \forall y\, \forall x\, \varphi, \\ \exists x\, \exists y\, \varphi &\;\equiv\; \exists y\, \exists x\, \varphi. \end{aligned}$$

However, quantifiers of different types do not commute in general:

$$\forall x\, \exists y\, \varphi \;\not\equiv\; \exists y\, \forall x\, \varphi.$$

Remark 2.56 (Common error). Swapping $\forall$ and $\exists$ changes the meaning. The statement $\exists y\, \forall x\, \varphi$ requires a single uniform witness y that works for all x, whereas $\forall x\, \exists y\, \varphi$ allows a different y for each x.

Proposition 2.57 (Quantifier Distribution). The following distribution equivalences hold:

$$\begin{aligned} \forall x\, (\varphi \wedge \psi) &\;\equiv\; (\forall x\, \varphi) \wedge (\forall x\, \psi), \\ \exists x\, (\varphi \vee \psi) &\;\equiv\; (\exists x\, \varphi) \vee (\exists x\, \psi). \end{aligned}$$

The following do not hold in general:

$$\begin{aligned} \forall x\, (\varphi \vee \psi) &\;\not\equiv\; (\forall x\, \varphi) \vee (\forall x\, \psi), \\ \exists x\, (\varphi \wedge \psi) &\;\not\equiv\; (\exists x\, \varphi) \wedge (\exists x\, \psi). \end{aligned}$$

Additionally, when $x \notin \mathrm{FV}(\psi)$:

$$\begin{aligned} \forall x\, (\varphi \wedge \psi) &\;\equiv\; (\forall x\, \varphi) \wedge \psi, \\ \exists x\, (\varphi \vee \psi) &\;\equiv\; (\exists x\, \varphi) \vee \psi. \end{aligned}$$

Remark 2.58 (Mnemonic). Universal distributes over conjunction ($\forall/\wedge$ match because both are "all-or-nothing"); existential distributes over disjunction ($\exists/\vee$ match because both assert "at least one").

Proposition 2.59 (Vacuous Quantification). If x does not occur free in φ, then:

$$\begin{aligned} \forall x\, \varphi &\;\equiv\; \varphi, \\ \exists x\, \varphi &\;\equiv\; \varphi. \end{aligned}$$

Remark 2.60 (English reading). A quantifier is vacuous if its bound variable does not appear free in the scope. Such quantifiers may be freely added or removed without changing meaning.

Proposition 2.61 (Renaming Bound Variables). If y does not occur in φ, then:

$$\begin{aligned} \forall x\, \varphi &\equiv \forall y\, \varphi[y/x], \\ \exists x\, \varphi &\equiv \exists y\, \varphi[y/x]. \end{aligned}$$

Remark 2.62 (Consequence). This is a direct consequence of alpha-equivalence (Definition 2.1.2). Renaming bound variables is the standard technique for avoiding variable capture when applying substitution rules.

2.1.10 Quantifiers: Prenex Normal Form

Prenex Normal Form Quick Reference

Concept	Meaning	Detail
Prenex normal form (PNF)	All quantifiers at front, quantifier-free matrix	Def
PNF Theorem	Every formula is equivalent to a PNF formula	Thm
Quantifier-pulling rules	Equivalences for extracting quantifiers	Def

Definition (Prenex Normal Form)

A first-order formula is in prenex normal form (PNF) if it has the shape

$$Q_1 x_1\, Q_2 x_2 \cdots Q_n x_n \;\; \psi,$$

where each $Q_i \in \{\forall, \exists\}$ and ψ is quantifier-free. The string $Q_1 x_1 \cdots Q_n x_n$ is the quantifier prefix; ψ is the matrix.

Theorem (Prenex Normal Form Theorem)

Every first-order formula is logically equivalent to a formula in prenex normal form.

Remark 2.63 (Proof strategy conversion procedure). Given a formula φ, the following four steps produce an equivalent PNF:

1. Eliminate $\leftrightarrow$ and $\rightarrow$. Rewrite using only $\neg, \wedge, \vee$: $(A \rightarrow B) \equiv (\neg A \vee B)$.
2. Push negations inward (Negation Normal Form). Use De Morgan's laws and quantifier-negation laws until $\neg$ applies only to atomic formulas.
3. Standardize bound variables apart. Rename bound variables so that no variable is quantified twice and no bound variable coincides with any free variable.
4. Pull quantifiers outward. Apply the quantifier-pulling equivalences (below) to move all quantifiers to the front.

Remark 2.64 (Quantifier-pulling equivalences). The following hold when $x \notin \mathrm{FV}(\psi)$:

Equivalence	Side condition
$(\forall x\, \phi) \wedge \psi \equiv \forall x\, (\phi \wedge \psi)$	$x \notin \mathrm{FV}(\psi)$
$(\exists x\, \phi) \vee \psi \equiv \exists x\, (\phi \vee \psi)$	$x \notin \mathrm{FV}(\psi)$
$(\forall x\, \phi) \vee \psi \equiv \forall x\, (\phi \vee \psi)$	$x \notin \mathrm{FV}(\psi)$
$(\exists x\, \phi) \wedge \psi \equiv \exists x\, (\phi \wedge \psi)$	$x \notin \mathrm{FV}(\psi)$

Example 2.65 (PNF conversion). Normalize $\varphi := \neg\big(\forall x\, P(x) \to \exists y\, Q(y)\big)$.

Step 1 (eliminate $\to$): $\varphi \equiv \neg\big(\neg\forall x\, P(x) \vee \exists y\, Q(y)\big)$.

Step 2 (push $\neg$ inward): $\varphi \equiv (\forall x\, P(x)) \wedge (\forall y\, \neg Q(y))$.

Step 3 (standardize apart): Variables are already distinct.

Step 4 (pull quantifiers outward): $\varphi \equiv \forall x\, \forall y\, (P(x) \wedge \neg Q(y))$.

2.1.11 Quantifiers: Logical Strength and Quantifier Order

Logical Strength Quick Reference

Concept	Meaning	Detail
Entailment $A \models B$	A is true implies B is true	Def
Logical strength	A stronger than B if $A \models B$ but $B \not\models A$	Def
Model-set viewpoint	$A \models B \iff \mathrm{Mod}(A) \subseteq \mathrm{Mod}(B)$	Def
Quantifier alternation	Alternating $\forall/\exists$ in prefix	Def

Definition (Logical Entailment)

Let A and B be sentences. We write $A \models B$ and say that A entails B if in every structure in which A is true, B is also true. Equivalently, $A \models B$ means $A \to B$ is valid.

Definition (Logical Strength)

We say that A is (logically) stronger than B if $A \models B$ but $B \not\models A$.
A and B are logically equivalent if $A \models B$ and $B \models A$.

Remark 2.66 (English reading). "Stronger" means "harder to satisfy" (fewer models make it true). If A is stronger than B, then A rules out more possibilities: every world where A holds is a world where B holds, but not conversely.

Definition (Model-Set Viewpoint)

For a statement A, let $\mathrm{Mod}(A)$ denote the class of all structures in which A is true. Then

$$A \models B \quad \iff \quad \mathrm{Mod}(A) \subseteq \mathrm{Mod}(B).$$

Remark 2.67 (English reading). Strength is set inclusion of model classes. "A is stronger than B" is literally $\mathrm{Mod}(A) \subsetneq \mathrm{Mod}(B)$: A satisfies fewer structures than B.

Definition (Quantifier Alternation)

A quantifier alternation occurs when $Q_i \neq Q_{i+1}$ for some i in the prefix $Q_1 x_1 \cdots Q_k x_k$ of a formula. For example, $\forall x\, \exists y\, \Phi(x, y)$ has one alternation, while $\forall x\, \forall y\, \Phi(x, y)$ has none.

Remark 2.68 (Why alternation matters for strength). $\forall x\, \exists y\, \Phi(x, y)$ requires a rule assigning a (possibly x-dependent) witness y to each x. In contrast, $\exists y\, \forall x\, \Phi(x, y)$ requires a single uniform witness y that works for all x. Uniformity is strictly stronger than dependence.

Remark 2.69 (Strength hierarchy). The following entailments hold:

1. $\forall x\, \Phi(x) \models \exists x\, \Phi(x)$ (universal is stronger).
2. $\exists y\, \forall x\, \Phi(x, y) \models \forall x\, \exists y\, \Phi(x, y)$ (uniform witness is stronger than dependent witness).

Neither converse holds in general.

Example 2.70 (Swap failure on $\mathbb{R}$). Let the domain be $\mathbb{R}$ and $\Phi(x, y) = (y > x)$.

$\forall x\, \exists y\, (y > x)$ is true: given any x, take $y = x + 1$.

$\exists y\, \forall x\, (y > x)$ is false: no single real exceeds all reals.

Hence $\exists y\, \forall x$ is strictly stronger than $\forall x\, \exists y$.

Example 2.71 (Universal vs. existential). Let the domain be $\mathbb{R}$ and $\Phi(x) = (x^2 \geq 0)$.

$\forall x\, (x^2 \geq 0)$ entails $\exists x\, (x^2 \geq 0)$, but not conversely. Thus the universal statement is strictly stronger.

Remark 2.72 (Logical strength increases when you:). • add conjuncts ($A \wedge B$ is stronger than A),

- replace $\exists$ with $\forall$,
- move existential quantifiers outward (requiring uniform witnesses),
- introduce quantifier alternation.

2.1.12 Proof Theory: Quantifier Inference Rules

Quantifier Inference Rules Quick Reference

Rule	Schema	Key restriction	Detail
UI	$\forall x\, \varphi \Rightarrow \varphi[t/x]$	t free for x	Def
UG	$\varphi \Rightarrow \forall x\, \varphi$	x not free in assumptions	Def
EI	$\exists x\, \varphi \Rightarrow \varphi[c/x]$	c fresh witness	Def
EG	$\varphi[t/x] \Rightarrow \exists x\, \varphi$	none	Def

Definition (Universal Instantiation — UI)

From a universally quantified statement, infer any instance obtained by substituting a term for the bound variable:

$$\forall x\, \varphi \;\Rightarrow\; \varphi[t/x].$$

Remark 2.73 (Side condition). The term t must be free for x in φ. If t contains a variable that would become bound after substitution, the inference is invalid.

Definition (Universal Generalization — UG)

From a formula, infer a universally quantified statement:

$$\varphi \;\Rightarrow\; \forall x\, \varphi.$$

Remark 2.74 (Restriction on UG). The variable x must not occur free in any undischarged assumption on which φ depends. This restriction ensures that x is truly arbitrary — not tied to a specific hypothesis about x.

Remark 2.75 (Arbitrary element argument). To prove $\forall x\, \varphi(x)$, fix an arbitrary element x (introducing no special assumptions about x), prove $\varphi(x)$, and then apply UG. The variable x must remain unconstrained throughout.

Definition (Existential Instantiation — EI)

From an existentially quantified statement, infer an instance with a fresh constant (witness):

$$\exists x\, \varphi \;\Rightarrow\; \varphi[c/x].$$

Remark 2.76 (Witness discipline for EI). The constant c must be fresh: it must not appear in φ, in any undischarged assumption, or in the final conclusion of the proof. The constant represents an arbitrary but fixed witness to the existential claim, not a specific known object.

Remark 2.77 (Common error). Using a witness constant introduced by EI outside its allowed scope, or choosing c before establishing the existential premise, invalidates the proof.

Definition (Existential Generalization — EG)

From a formula containing a term, infer an existential statement:

$$\varphi[t/x] \;\Rightarrow\; \exists x\, \varphi.$$

Remark 2.78 (No side conditions). EG has no side conditions: any term t may be replaced by an existentially quantified variable. To prove $\exists x\, \varphi(x)$, explicitly exhibit a term t such that $\varphi(t)$ holds, then apply EG.

Remark 2.79 (Counterexample argument). To refute $\forall x\, \varphi(x)$: produce a single term t such that $\neg\varphi(t)$ holds. To refute $\exists x\, \varphi(x)$: show that $\varphi(t)$ fails for every term t in the domain.

2.1.13 Proof Theory: Equality Rules

Equality Rules — Quick Reference

Rule	Schema	Status	Detail
Reflexivity (EqI)	$t = t$	Primitive rule	Def
Equality Elim (EqE)	$t_1 = t_2,\ \varphi[t_1/x] \Rightarrow \varphi[t_2/x]$	Primitive rule	Def
Symmetry	$t_1 = t_2 \Rightarrow t_2 = t_1$	Derived	Thm
Transitivity	$t_1 = t_2,\ t_2 = t_3 \Rightarrow t_1 = t_3$	Derived	Thm
Term substitution	$f(\dots, t_1, \dots) = f(\dots, t_2, \dots)$	Derived	Thm
Predicate substitution	$P(\dots, t_1, \dots) \Leftrightarrow P(\dots, t_2, \dots)$	Derived	Thm

Definition (Equality Introduction — Reflexivity)

For any term t, one may infer $t = t$.

Remark 2.80 (English reading). Reflexivity requires no premises and holds for all terms under all interpretations. It is the foundational rule that allows equality reasoning to get off the ground.

Definition (Equality Elimination — Substitution of Equals)

From $t_1 = t_2$ and $\varphi[t_1/x]$, one may infer $\varphi[t_2/x]$.

Remark 2.81 (English reading). Equality elimination permits replacing a term by an equal term in any formula position, provided the substitution is capture-avoiding. This formalizes the intuition that equal things can be substituted for each other.

Theorem 2.82 (Symmetry of Equality). From $t_1 = t_2$, one may infer $t_2 = t_1$.

Theorem 2.83 (Transitivity of Equality). From $t_1 = t_2$ and $t_2 = t_3$, one may infer $t_1 = t_3$.

Theorem 2.84 (Term Substitution under Equality). If $t_1 = t_2$, then for any function symbol f, $f(\dots, t_1, \dots) = f(\dots, t_2, \dots)$.

Theorem 2.85 (Predicate Substitution under Equality). If $t_1 = t_2$ and P is an n-ary predicate symbol, then $P(\dots, t_1, \dots) \ \Leftrightarrow\ P(\dots, t_2, \dots)$.

Remark 2.86 (Summary). These derived rules collectively express that functions and predicates respect equality: equal inputs produce equal outputs (for functions) or equivalent propositions (for predicates). Together they amount to the principle of Leibniz substitutivity.

Remark 2.87 (Common error). When substituting equals for equals, the substitution must be made consistently. Replacing a term by an equal in one occurrence but not another in the same formula can produce incorrect conclusions.

2.1.14 Proof Theory: Soundness and Completeness

Soundness and Completeness Quick Reference

Property	Direction	Detail
Soundness	$\Gamma \vdash \varphi \Rightarrow \Gamma \models \varphi$	Def
Completeness	$\Gamma \models \varphi \Rightarrow \Gamma \vdash \varphi$	Def
Soundness Theorem (FOL)	All standard FOL rules are sound	Thm
Gödel's Completeness Theorem	FOL is complete	Thm

Definition (Soundness)

A proof system is sound if every provable formula is valid:

$$\Gamma \vdash \varphi \implies \Gamma \models \varphi.$$

Definition (Completeness)

A proof system is complete if every valid formula is provable:

$$\Gamma \models \varphi \implies \Gamma \vdash \varphi.$$

Theorem 2.88 (Soundness of First-Order Logic). The standard inference rules for first-order logic (UI, UG, EI, EG, and the propositional rules) are sound: they preserve truth in all structures.

If $\Gamma \vdash \varphi$, then $\Gamma \models \varphi$.

Theorem (Gödel's Completeness Theorem)

First-order logic is complete: every logically valid formula is provable.

$$\Gamma \models \varphi \implies \Gamma \vdash \varphi.$$

Equivalently: if a set of sentences Γ is consistent (has no proof of contradiction), then Γ has a model.

Remark 2.89 (English reading). Soundness ensures proofs do not lead us astray: we cannot prove false statements from true premises. Completeness ensures proofs are powerful enough: every statement that is true in all models can be established by a formal proof.

Together, soundness and completeness show that syntactic provability ($\vdash$) and semantic entailment ($\models$) coincide for first-order logic.

Remark 2.90 (Completeness vs. incompleteness). Gödel's completeness theorem (1930) should not be confused with his incompleteness theorems (1931). The completeness theorem concerns first-order logic as a proof system. The incompleteness theorems concern the limitations of formal theories capable of expressing arithmetic, and are a separate result.

2.1.15 Translation: English to Logic and Scope Ambiguity

Translation Quick Reference

English pattern	Logical form	Negation	Detail
Everyone has P	$\forall x\, P(x)$	$\exists x\, \neg P(x)$	Table
Someone has P	$\exists x\, P(x)$	$\forall x\, \neg P(x)$	Table
Everyone likes something	$\forall x\, \exists y\, L(x, y)$	$\exists x\, \forall y\, \neg L(x, y)$	Table
Someone likes everything	$\exists x\, \forall y\, L(x, y)$	$\forall x\, \exists y\, \neg L(x, y)$	Table
Scope ambiguity	Two readings differ by quantifier order		Table

Remark 2.91 (Translation discipline). When translating a symbolic formula into natural language, respect the original quantifier nesting of the formula. Do not reorder quantifiers unless explicitly asked to normalize or rewrite the expression.

English	Logical Form	Negation
Everyone has property P	$\forall x\, P(x)$	$\exists x\, \neg P(x)$
Someone has property P	$\exists x\, P(x)$	$\forall x\, \neg P(x)$
Everyone likes something	$\forall x\, \exists y\, L(x, y)$	$\exists x\, \forall y\, \neg L(x, y)$
Someone likes everything	$\exists x\, \forall y\, L(x, y)$	$\forall x\, \exists y\, \neg L(x, y)$
Every student passed every exam	$\forall x\, \forall y\, P(x, y)$	$\exists x\, \exists y\, \neg P(x, y)$
Some student passed every exam	$\exists x\, \forall y\, P(x, y)$	$\forall x\, \exists y\, \neg P(x, y)$

Remark 2.92 (Scope determines meaning). Natural language often leaves quantifier scope ambiguous. Formal logic resolves this by fixing a precise quantifier order. Changing the order of quantifiers generally changes the meaning of the statement.

English	Logical Form	Meaning
Everyone loves someone	$\forall x\, \exists y\, L(x, y)$	Each person may love a different person.
	$\exists y\, \forall x\, L(x, y)$	There is one person whom everyone loves.
Every student passed an exam	$\forall x\, \exists y\, P(x, y)$	Each student passed at least one (possibly different) exam.
	$\exists y\, \forall x\, P(x, y)$	There is a single exam that all students passed.
A teacher knows every student	$\exists x\, \forall y\, K(x, y)$	There is one teacher who knows all students.
	$\forall y\, \exists x\, K(x, y)$	Every student is known by at least one teacher.
Every function has a zero	$\forall f\, \exists x\, Z(f, x)$	Each function has at least one (possibly different) zero.
	$\exists x\, \forall f\, Z(f, x)$	There is a single point that is a zero of every function.

2.1.16 Translation: Square of Opposition

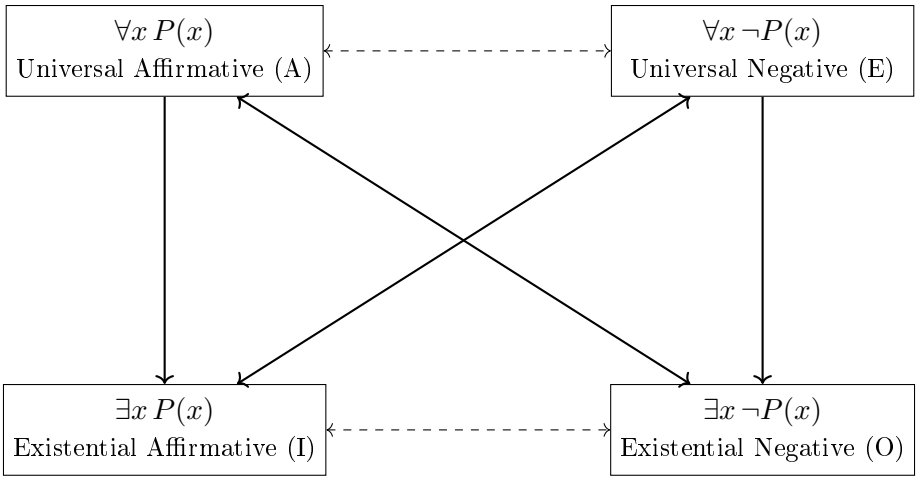
Square of Opposition Quick Reference

Form	Name	Formula	Detail
A	Universal Affirmative	$\forall x\, P(x)$	Def
E	Universal Negative	$\forall x\, \neg P(x)$	Def
I	Existential Affirmative	$\exists x\, P(x)$	Def
O	Existential Negative	$\exists x\, \neg P(x)$	Def

Definition (Quantified Opposition Forms)

Let $P(x)$ be a formula with one free variable. The four standard quantified forms are:

$$\begin{aligned} \text{Universal Affirmative (A):} &\quad \forall x\, P(x) \\ \text{Universal Negative (E):} &\quad \forall x\, \neg P(x) \\ \text{Existential Affirmative (I):} &\quad \exists x\, P(x) \\ \text{Existential Negative (O):} &\quad \exists x\, \neg P(x) \end{aligned}$$



Remark 2.93 (Logical relations in first-order logic with nonempty domains). • Contradictories: A and O ($\forall x\, P(x)$ vs $\exists x\, \neg P(x)$); E and I ($\forall x\, \neg P(x)$ vs $\exists x\, P(x)$). These pairs cannot both be true and cannot both be false.

- Subalternation: A entails I ($\forall x\, P(x) \Rightarrow \exists x\, P(x)$); E entails O ($\forall x\, \neg P(x) \Rightarrow \exists x\, \neg P(x)$). These hold in classical first-order logic with nonempty domains.
- Contraries and subcontraries do not generally hold in first-order logic without existential presuppositions.

Remark 2.94 (Modern status of the square). In classical first-order logic, only contradiction and subalternation are logically valid relations. The traditional notions of contrariety and subcontrariety rely on existential assumptions and are not preserved in general model-theoretic semantics.

2.1.17 Reference: Common Errors and Fallacies

Errors and Fallacies Quick Reference

Error category	Where it arises	Detail
Quantifier fallacies	Negation, scope, and swap errors	Table
Common incorrect negations	Forgetting to flip quantifiers	Table
Inference rule errors	Misuse of UI, UG, EI, EG	Rem

Fallacy	Diagnostic Question
Failure to flip quantifier under negation	Was $\forall$ changed to $\exists$ (or vice versa) when negating?
Negating predicate only	Was only the predicate negated while the quantifier was left unchanged?
Partial quantifier negation	When negating nested quantifiers, were all quantifiers flipped?
Quantifier order confusion	Was the order of quantifiers changed without justification?
Illicit quantifier swap	Were $\forall$ and $\exists$ commuted when they are not of the same type?
Variable capture	Did substitution introduce a variable that became bound unintentionally?
Illegal existential instantiation	Was a witness chosen before an existential statement was established?
Illicit universal generalization	Was a universally quantified conclusion drawn from a statement depending on a specific object?
Scope ambiguity	Was the scope of a quantifier unclear or implicitly extended beyond its syntactic bounds?
Vacuous quantification misuse	Was a quantifier added or removed even though the variable does not occur free in the formula?

Remark 2.95. Quantifier errors almost always stem from ignoring scope or treating quantifiers as informal linguistic modifiers. Every quantifier introduces a binding context that must be respected syntactically before semantic reasoning is applied.

Original	Incorrect Negation	Why It Is Wrong
$\forall x\, P(x)$	$\forall x\, \neg P(x)$	Fails to flip the quantifier; asserts everyone fails P.
$\exists x\, P(x)$	$\exists x\, \neg P(x)$	Negates the predicate but not the existential claim.
$\forall x\, \exists y\, P(x, y)$	$\exists x\, \exists y\, \neg P(x, y)$	Only the outer quantifier was flipped.
$\exists x\, \forall y\, P(x, y)$	$\forall x\, \forall y\, \neg P(x, y)$	Overstrengthened the negation.
$\forall x\, (P(x) \to Q(x))$	$\forall x\, (P(x) \wedge \neg Q(x))$	Correct negation is $\exists x\, (P(x) \wedge \neg Q(x))$.

Remark 2.96 (Negation discipline). When negating a quantified statement, push the negation inward across every quantifier, flipping each one, until it reaches the atomic predicate. Negating only the predicate or only one quantifier changes the logical claim in an incorrect direction.

Remark 2.97 (Common inference rule errors). • Choosing a witness before establishing an existential premise.

- Using a witness constant introduced by EI outside its allowed scope.
- Generalizing universally over a variable that depends on a special assumption.
- Substituting inside the scope of a quantifier without checking for variable capture.
- Replacing a term by an equal term in one occurrence but not another.

Remark 2.98. A useful informal check: if a proof step would still make sense after replacing $\forall$ with "for most" or $\exists$ with "maybe", the step is almost certainly invalid.

2.1.18 Reference: Summary Tables

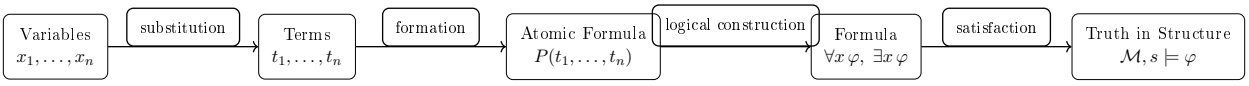
Comparison: Propositional vs. Predicate Logic

Aspect	Propositional Logic	Predicate Logic
Atomic formulas	Propositional variables (e.g. P, Q)	Predicate symbols applied to terms (e.g. $P(x)$, $R(x, y)$)
Internal structure	No internal structure	Terms, variables, constants, functions
Dependence on variables	None	May contain free variables
Semantic interpretation	Assigned a truth value directly	True or false relative to a structure and assignment
Use of quantifiers	Not available	Essential component
Semantic evaluation	Truth tables	Satisfaction in a structure

Quantifier Rules Summary

Rule Name	Schema	Conditions / Notes
Universal Instantiation	$\forall x\, \varphi \Rightarrow \varphi[t/x]$	t must be free for x in φ
Existential Generalization	$\varphi[t/x] \Rightarrow \exists x\, \varphi$	t may be any term
Existential Instantiation	$\exists x\, \varphi \Rightarrow \varphi[c/x]$	c is a new constant symbol
Universal Generalization	$\varphi \Rightarrow \forall x\, \varphi$	x not free in any undischarged assumption
Quantifier Negation	$\neg\forall x\, \varphi \equiv \exists x\, \neg\varphi$	
	$\neg\exists x\, \varphi \equiv \forall x\, \neg\varphi$	
Quantifier Commutation	$\forall x\, \forall y\, \varphi \equiv \forall y\, \forall x\, \varphi$	Same quantifier type only
	$\exists x\, \exists y\, \varphi \equiv \exists y\, \exists x\, \varphi$	
Vacuous Quantification	$\forall x\, \varphi \equiv \varphi$	x not free in φ
	$\exists x\, \varphi \equiv \varphi$	
Renaming Bound Variables	$\forall x\, \varphi \equiv \forall y\, \varphi[y/x]$	y not free in φ

Syntax to Semantics Overview



Remark 2.99. Variables and terms belong purely to syntax. Predicates form atomic formulas. Logical connectives and quantifiers build complex formulas. A structure and variable assignment then determine whether a formula is satisfied, with quantifiers ranging over the domain of discourse.

2.2 Proofs

2.3 Capstone

2.4 Capstone Assessment: Predicate Calculus

Purpose. This capstone assesses mastery of first-order predicate logic, including quantifiers, identity, scope, semantic consequence, and formal proof structure. All arguments must be expressed using first-order logical reasoning. No set-theoretic arguments may be used unless explicitly stated.

Instructions. Each problem requires a complete and rigorous proof. You must explicitly justify quantifier introduction and elimination steps. Appeals to intuition or informal paraphrase are not sufficient.

Problem 1 Quantifier Order Sensitivity

Prove that the following implication is logically valid:

$$\forall x(P(x) \to Q(x)) \ \wedge \ \exists x\, P(x) \ \to \ \exists x\, Q(x).$$

Your proof must explicitly identify where existential instantiation and universal instantiation are applied.

Problem 2 Failure of Converse

Show that the converse of the statement in Problem 1 is not logically valid. That is, show that

$$\exists x\, Q(x) \ \to \ \exists x\, P(x)$$

does not follow from

$$\forall x(P(x) \to Q(x)).$$

Your argument must be semantic (valuation/model-based), not syntactic.

Problem 3 Scope and Negation

Prove that the following two formulas are logically equivalent:

$$\neg\forall x\, P(x) \quad \text{and} \quad \exists x\, \neg P(x).$$

Your proof must make explicit use of the semantics of quantifiers.

Problem 4 Identity Reasoning

Assume identity is part of the language. Prove that the following implication is logically valid:

$$x = y \;\rightarrow\; (P(x) \leftrightarrow P(y)).$$

Your proof must not assume substitutivity without justification.

Problem 5 Semantic Consequence

Show that the set of formulas

$$\{\, \forall x(P(x) \rightarrow Q(x)),\; \forall x(Q(x) \rightarrow R(x)),\; \exists x\, P(x) \,\}$$

logically implies

$$\exists x\, R(x).$$

Your proof must argue that every structure satisfying the premises also satisfies the conclusion.

Completion Criterion. You have mastered predicate calculus if all five proofs:

- use quantifier rules correctly,
- respect variable scope and dependency,
- distinguish syntax from semantics, and
- are written without hidden assumptions.

Successful completion certifies readiness to proceed to set-theoretic foundations.

Chapter 3

Sets, Relations, and Functions

3.1 Notes

Where You Are in the Journey

Propositional Logic $\rightarrow$ Predicate Calculus $\rightarrow$ Sets & Functions $\rightarrow$ Proof Techniques $\rightarrow$ Axiom Systems $\rightarrow$ $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$ $\rightarrow$ Real Analysis $\rightarrow$ $\cdots$

How we got here. Logic gave us a precise language. We now need mathematical objects to reason about. Sets are the universe of mathematical objects; functions are the structure-preserving maps between them. Together they provide the language of all subsequent mathematics.

What this chapter builds. We develop naïve and axiomatic set theory: membership, subsets, operations, power sets, and families. Relations — including equivalence relations and partial orders — give us structure on sets. Functions capture how sets relate to each other: injections, surjections, bijections, and cardinality.

Where this leads. Every algebraic structure is a set with additional operations. Every proof about $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{R}$ uses the function and relation vocabulary established here. Equivalence relations appear again in the integer and rational constructions; order relations appear throughout analysis.

Structural Roadmap

Each major topic is organised as:

$$\text{Definitions} \longrightarrow \text{Main Theorems} \longrightarrow \text{Consequences and Structural Insight}$$

The global progression is:

1. Sets: membership, subsets, operations (union, intersection, difference), power sets, the empty

set

2. Relations: binary relations, properties (reflexive, symmetric, transitive), composition
3. Families: indexed families, arbitrary unions and intersections
4. Equivalence relations: equivalence classes, partitions, quotient sets
5. Functions: domain, codomain, image, injection, surjection, bijection, composition, inverse
6. Order: partial orders, total orders, well-orders, Zorn's lemma

Remark 3.1 (Primary source). The primary driver is Suppes, Introduction to Logic, supplemented by Gerstein, Introduction to Mathematical Structures and Proofs.

3.1.1 Sets, Membership, and ZFC Axioms

Foundations Quick Reference

Concept	Meaning	Detail
Set, membership	Primitive notions; meaning fixed by axioms	Def
Extensionality	Sets equal iff same elements	Ax
Empty Set	Exists a set with no elements	Ax
Pairing	Any two sets can be collected	Ax
Union	Elements of sets in a family	Ax
Power Set	All subsets form a set	Ax
Infinity	An infinite set exists	Ax
Separation	Subsets by definable property	Ax
Replacement	Functional image of a set is a set	Ax
Foundation	No infinite descending $\in$-chains	Ax
Choice	Selection function exists	Ax

Definition (Set and Membership)

In axiomatic set theory, the notions of set and membership are primitive.

- A set is an object.
- Membership is a binary relation, denoted by $\in$, between objects.

If x is an object and A is a set, the statement $x \in A$ is read as "x is an element of A." No definition of "set" or "$\in$" is given in more basic terms. Their meaning is determined entirely by the axioms governing them.

Remark 3.2 (English reading). Primitive means we do not define these in terms of simpler concepts. Rather, we fix their meaning implicitly by stating the axioms they must obey. This is the standard approach in formal mathematics: rules of behaviour replace informal definitions.

Remark 3.3 (Consequence). All subsequent notions—subsets, ordered pairs, relations, functions, number systems—are definitions introduced within this axiomatic framework. Every proved result is a theorem derived logically from the axioms via the rules of inference.

Axiom System (ZFC)

Axiom of Extensionality.
Two sets are equal iff they have the same elements.

$$\forall A \forall B \Big(A = B \iff \forall x \, (x \in A \leftrightarrow x \in B) \Big).$$

Axiom of Empty Set.
There exists a set with no elements.

$$\exists A \forall x \, (x \notin A).$$

Axiom of Pairing.
For any two sets, there exists a set containing exactly those two sets.

$$\forall A \forall B \exists C \forall x \, (x \in C \leftrightarrow (x = A \vee x = B)).$$

Axiom of Union.
For any family of sets, there exists a set containing exactly the elements of those sets.

$$\forall A \exists U \forall x \, \big(x \in U \leftrightarrow \exists B \, (B \in A \wedge x \in B) \big).$$

Axiom of Power Set.
For any set, there exists a set of all its subsets.

$$\forall A \exists P \forall x \, (x \in P \leftrightarrow x \subseteq A).$$

Axiom of Infinity.
There exists an infinite set.

$$\exists A \Big(\varnothing \in A \wedge \forall x \, (x \in A \to x \cup \{x\} \in A) \Big).$$

Axiom Schema of Separation.
Given a set and a property, there exists a subset containing exactly the elements satisfying that property. For any formula $\varphi(x)$,

$$\forall A \exists B \forall x \, (x \in B \leftrightarrow (x \in A \wedge \varphi(x))).$$

Axiom Schema of Replacement.
If $\varphi(x, y)$ defines a functional relation on a set A, then the image of A under φ is a set. For any formula $\varphi(x, y)$,

$$\forall A \Big(\big(\forall x \in A \, \exists! y \, \varphi(x, y) \big) \to \exists B \forall y \, \big(y \in B \leftrightarrow \exists x \in A \, \varphi(x, y) \big) \Big).$$

Axiom of Foundation.
Every nonempty set has an $\in$-minimal element.

$$\forall A \Big(A \neq \varnothing \to \exists x \in A \, (x \cap A = \varnothing) \Big).$$

Axiom of Choice.
For any family of nonempty sets, there exists a selection function.

Remark 3.4 (Role of the axioms). The axioms are not statements to be proved, but rules specifying how $\in$ behaves and which sets are permitted to exist. From this point onward, set theory functions as the underlying language of mathematics: all reasoning about mathematical objects is ultimately grounded in these axioms, even when they are not cited explicitly.

Remark 3.5 (Separation vs. Replacement). Separation is a schema: one axiom instance per definable property φ. It carves out subsets of an already-existing set. Replacement is strictly stronger: it can produce sets whose elements are not already contained in any pre-existing set, allowing the construction of large stages of the cumulative hierarchy.

Remark 3.6 (Axiom of Choice and right inverses). The Axiom of Choice is equivalent to many statements used later in analysis, including Zorn's Lemma, the well-ordering theorem, and the statement that every surjective function has a right inverse. It is independent of the other ZFC axioms.

3.1.2 Set Constructions and Operations

Set Operations Quick Reference

Concept	Notation	Membership condition	Detail
Empty set	$\varnothing$	$x \in \varnothing$ never	Def
Subset	$A \subseteq B$	$x \in A \Rightarrow x \in B$	Def
Proper subset	$A \subsetneq B$	$A \subseteq B$ and $A \neq B$	Def
Set equality	$A = B$	same elements both ways	Def
Union	$A \cup B$	in A or in B	Def
Intersection	$A \cap B$	in A and in B	Def
Set difference	$A \setminus B$	in A but not B	Def
Symmetric diff.	$A \triangle B$	in exactly one of A, B	Def
Complement	$A^c = U \setminus A$	in U but not A	Def
Cartesian product	$A \times B$	all ordered pairs (a, b)	Def
Power set	$\mathcal{P}(A)$	all subsets of A	Def
De Morgan	$(A \cup B)^c = A^c \cap B^c$	complement of union	Thm

Definition (Empty Set)

The unique set with no elements is called the empty set and is denoted $\varnothing$.

Remark 3.7 (Vacuous truth). Statements of the form $\forall x \in \varnothing,\ P(x)$ are vacuously true: there is no element $x \in \varnothing$ for which $P(x)$ could fail. This is not a special convention but a consequence of how universal quantification is defined.

Definition (Subset)

Let A and B be sets. We say A is a subset of B, written $A \subseteq B$, if every element of A is also an element of B:

$$A \subseteq B \iff \forall x\,(x \in A \to x \in B).$$

Definition (Proper Subset)

A is a proper subset of B, written $A \subsetneq B$, if $A \subseteq B$ and $A \neq B$.

Remark 3.8 (Notation convention). Some authors write $A \subset B$ for a proper subset; others use it to mean $A \subseteq B$. In these notes, $\subseteq$ always denotes subset (possibly equal) and $\subsetneq$ always denotes proper subset.

Definition (Set Equality)

Two sets A and B are equal, written $A = B$, if they have the same elements:

$$A = B \iff \forall x\,(x \in A \leftrightarrow x \in B).$$

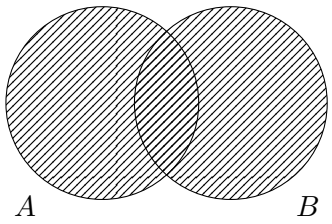
Equivalently, $A = B \iff (A \subseteq B \wedge B \subseteq A)$.

Remark 3.9 (Proof strategy). In practice, to prove $A = B$ one shows mutual inclusion: first $A \subseteq B$ (take arbitrary $x \in A$ and deduce $x \in B$), then $B \subseteq A$. This two-step structure appears throughout set-theoretic proofs.

Definition (Union)

The union of A and B, denoted $A \cup B$, is the set of all elements belonging to at least one of the two sets:

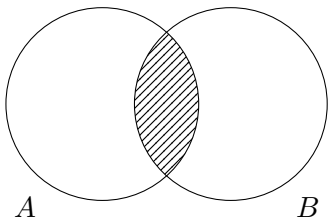
$$A \cup B \;=\; \{\, x \mid x \in A \vee x \in B \,\}.$$



Definition (Intersection)

The intersection of A and B, denoted $A \cap B$, is the set of all elements common to both:

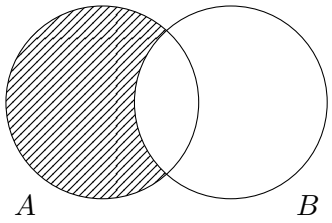
$$A \cap B \;=\; \{\, x \mid x \in A \wedge x \in B \,\}.$$



Definition (Set Difference)

The set difference $A \setminus B$ is the set of elements in A but not in B:

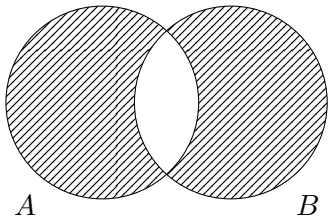
$$A \setminus B \;=\; \{\, x \mid x \in A \wedge x \notin B \,\}.$$



Definition (Symmetric Difference)

The symmetric difference $A \triangle B$ is the set of elements belonging to exactly one of A and B:

$$A \triangle B \;=\; (A \setminus B) \cup (B \setminus A) \;=\; \{\, x \mid (x \in A \vee x \in B) \wedge x \notin A \cap B \,\}.$$



Remark 3.10 (Algebraic structure of $\triangle$). The symmetric difference is commutative and associative. Under $\triangle$, the power set $\mathcal{P}(U)$ forms an abelian group with identity $\varnothing$ and every element its own inverse ($A \triangle A = \varnothing$).

Definition (Complement)

Let U be a fixed universe and $A \subseteq U$. The complement of A relative to U, denoted A^c, is:

$$A^c \;=\; U \setminus A.$$

Definition (Cartesian Product)

The Cartesian product of sets A and B is:

$$A \times B = \{\, (a, b) \mid a \in A \text{ and } b \in B \,\}.$$

Remark 3.11 (Ordered pairs). The ordered pair (a, b) is formally defined as $\{\{a\}, \{a, b\}\}$. This Kuratowski encoding ensures the key property: $(a, b) = (c, d)$ iff $a = c$ and $b = d$.

Remark 3.12 (Non-commutativity). In general, $A \times B \neq B \times A$. The product is however associative up to canonical isomorphism: $(A \times B) \times C \cong A \times (B \times C)$.

Definition (Power Set)

The power set of A, denoted $\mathcal{P}(A)$, is the set of all subsets of A:

$$\mathcal{P}(A) := \{ S \mid S \subseteq A \}.$$

Remark 3.13 (Size). If A has n elements, then $|\mathcal{P}(A)| = 2^n$. For infinite A, Cantor's theorem shows $|\mathcal{P}(A)| > |A|$ strictly.

Theorem (De Morgan's Laws)

Let U be a universe and $A, B \subseteq U$. Then

$$(A \cup B)^c = A^c \cap B^c \quad \text{and} \quad (A \cap B)^c = A^c \cup B^c.$$

Remark 3.14 (Logical analogy). De Morgan's laws mirror the propositional logic identities $\neg(P \vee Q) \equiv \neg P \wedge \neg Q$ and $\neg(P \wedge Q) \equiv \neg P \vee \neg Q$ under the correspondence $\cap \leftrightarrow \wedge$, $\cup \leftrightarrow \vee$, $A^c \leftrightarrow \neg A$.

Definition 3.15 (Set Duality). Two set-theoretic expressions over a fixed universe U are dual if one is obtained from the other by simultaneously replacing $\cup \leftrightarrow \cap$ and $\varnothing \leftrightarrow U$, with complements unchanged.

Corollary 3.16 (Principle of Set Duality). Any identity involving $\cup$, $\cap$, $\varnothing$, and U that holds for all subsets of a universe remains valid when each operation and constant is replaced by its dual.

Remark 3.17 (Using duality in proofs). To prove a statement involving unions and intersections, it often suffices to prove one version; the dual statement follows immediately.

Example 3.18 (Distributive law via duality). The two distributive laws

$$A \cap (B \cup C) = (A \cap B) \cup (A \cap C) \quad \text{and} \quad A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$$

are duals of each other. Proving either one and applying the Principle of Set Duality yields the other immediately.

Remark 3.19 (Transition to relations). The Cartesian product provides a way to encode ordered information. Relations and functions will be defined as special subsets of Cartesian products, so they require no new foundational objects.

3.1.3 Algebraic Laws of Set Operations

Set Algebra Laws Quick Reference

Law	Union form	Intersection form	Detail
Commutativity	$A \cup B = B \cup A$	$A \cap B = B \cap A$	Thm
Associativity	$(A \cup B) \cup C = A \cup (B \cup C)$	analogous	Thm
Distributivity	$A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$	dual	Thm
Identity	$A \cup \varnothing = A$	$A \cap U = A$	Thm
Absorption	$A \cup (A \cap B) = A$	$A \cap (A \cup B) = A$	Thm
Involution	$(A^c)^c = A$	—	Thm

Remark 3.20 (Algebraic structure of set operations). The operations $\cup$, $\cap$, and complement satisfy algebraic laws analogous to those of logical connectives. Together they endow $\mathcal{P}(U)$ with the structure of a Boolean algebra. These laws justify manipulation of set expressions in later proofs, particularly those involving equivalence classes, partitions, and functions.

Theorem 3.21 (Commutativity of Union and Intersection). Let A, B be sets. Then

$$A \cup B = B \cup A \quad \text{and} \quad A \cap B = B \cap A.$$

Theorem 3.22 (Associativity of Union and Intersection). Let A, B, C be sets. Then

$$(A \cup B) \cup C = A \cup (B \cup C) \quad \text{and} \quad (A \cap B) \cap C = A \cap (B \cap C).$$

Theorem 3.23 (Distributive Laws). Let A, B, C be sets. Then

$$A \cap (B \cup C) = (A \cap B) \cup (A \cap C),$$

$$A \cup (B \cap C) = (A \cup B) \cap (A \cup C).$$

Theorem 3.24 (Identity and Absorption Laws). Let A, B be sets and U a universe with $A \subseteq U$. Then

$$A \cup \varnothing = A, \qquad A \cap U = A,$$

$$A \cup (A \cap B) = A, \qquad A \cap (A \cup B) = A.$$

Theorem 3.25 (Involution of Complement). For any $A \subseteq U$,

$$(A^c)^c = A.$$

Remark 3.26 (Non-commutative and non-associative operations). Set difference $\setminus$ is neither commutative nor associative:

$$A \setminus B \neq B \setminus A, \qquad (A \setminus B) \setminus C \neq A \setminus (B \setminus C) \quad \text{in general.}$$

It interacts with union and intersection via:

$$A \setminus (B \cup C) = (A \setminus B) \cap (A \setminus C), \qquad A \setminus (B \cap C) = (A \setminus B) \cup (A \setminus C).$$

These follow from $A \setminus B = A \cap B^c$ together with De Morgan's laws.

Remark 3.27 (Cartesian product). The Cartesian product $\times$ is not commutative, is associative only up to canonical isomorphism, and distributes over union in each coordinate:

$$A \times (B \cup C) = (A \times B) \cup (A \times C).$$

3.1.4 Ordered Pairs and Relations

Relations Quick Reference

Concept	Meaning	Detail
Ordered pair	$(a, b) := \{\{a\}, \{a, b\}\}$; identity iff both coords equal	Def
Cartesian product	$A \times B$: all ordered pairs from A and B	Def
Relation	Any subset $R \subseteq A \times B$	Def

Definition (Ordered Pair)

Let a and b be sets. The ordered pair (a, b) is defined (Kuratowski) as

$$(a, b) := \{\{a\}, \{a, b\}\}.$$

Theorem 3.28 (Uniqueness of Ordered Pairs). For any sets a, b, c, d,

$$(a, b) = (c, d) \iff (a = c \wedge b = d).$$

Remark 3.29 (Why this encoding works). The purpose of the Kuratowski definition is purely to guarantee the uniqueness theorem above. Once $(a, b) = (c, d) \iff a = c \wedge b = d$ is established, we may treat ordered pairs as a primitive with this property and forget the encoding.

Definition (Cartesian Product as foundation for relations)

The Cartesian product of sets A and B is:

$$A \times B := \{\, (a, b) \mid a \in A \text{ and } b \in B \,\}.$$

Definition (Relation)

A relation from A to B is any subset $R \subseteq A \times B$.
If $(a, b) \in R$ we write $a\, R\, b$ and say "a is related to b." When $A = B$ we say R is a relation on A.

Remark 3.30 (Relations as structured sets). Relations introduce no new foundational objects: they are simply sets of ordered pairs, constructed using operations already available. Properties of relations can therefore be studied with ordinary set-theoretic reasoning.

3.1.5 Properties of Relations

Relational Properties Quick Reference

Property	Formal definition	Canonical example	Detail
Reflexive	$\forall a,\ (a, a) \in R$	$=$ on any set	Def
Irreflexive	$\forall a,\ (a, a) \notin R$	$<$ on $\mathbb{N}$	Def
Symmetric	$(a, b) \in R \Rightarrow (b, a) \in R$	"same age as"	Def
Antisymmetric	$(a, b), (b, a) \in R \Rightarrow a = b$	$\leq$ on $\mathbb{N}$	Def
Asymmetric	$(a, b) \in R \Rightarrow (b, a) \notin R$	$<$ on $\mathbb{N}$	Def
Transitive	$(a, b), (b, c) \in R \Rightarrow (a, c) \in R$	$\leq$ on $\mathbb{N}$	Def
Total (Connex)	$(a, b) \in R \vee (b, a) \in R$	$\leq$ on $\mathbb{R}$	Def
Structural classes formed by combining properties:			
Equivalence	Reflexive + Symmetric + Transitive	$=$ on any set	Def
Preorder	Reflexive + Transitive	$\leq$ on $\mathbb{N}$	Def
Partial order	Reflexive + Antisymmetric + Transitive	$\subseteq$ on $\mathcal{P}(A)$	Def
Total order	Partial order + Total	$\leq$ on $\mathbb{R}$	Def

Definition 3.31 (Reflexive Relation). A relation R on A is reflexive if every element is related to itself:

$$\forall a \in A,\ (a, a) \in R.$$

Definition 3.32 (Irreflexive Relation). R is irreflexive if no element is related to itself:

$$\forall a \in A,\ (a, a) \notin R.$$

Remark 3.33 (Independence of reflexive and irreflexive). A relation may be neither reflexive nor irreflexive (if some but not all elements satisfy $(a, a) \in R$). However, a relation cannot be both reflexive and irreflexive unless $A = \varnothing$.

Definition 3.34 (Symmetric Relation). R is symmetric if related pairs commute:

$$\forall a, b \in A,\ (a, b) \in R \to (b, a) \in R.$$

Definition 3.35 (Antisymmetric Relation). R is antisymmetric if mutual relation implies equality:

$$\forall a, b \in A,\ \big((a, b) \in R \land (b, a) \in R\big) \to a = b.$$

Definition 3.36 (Asymmetric Relation). R is asymmetric if related pairs never commute:

$$\forall a, b \in A,\ (a, b) \in R \to (b, a) \notin R.$$

Remark 3.37 (Asymmetric vs. antisymmetric). Every asymmetric relation is antisymmetric, but not conversely. $\leq$ on $\mathbb{R}$ is antisymmetric but not asymmetric (since $1 \leq 1$). Strict order $<$ is both asymmetric and irreflexive.

Definition 3.38 (Transitive Relation). R is transitive if it extends along chains:

$$\forall a, b, c \in A,\ \big((a, b) \in R \land (b, c) \in R\big) \to (a, c) \in R.$$

Definition 3.39 (Total (Connex) Relation). R is total (or connex) if every pair is comparable:

$$\forall a, b \in A,\ (a, b) \in R \lor (b, a) \in R.$$

Definition (Equivalence Relation)

R is an equivalence relation on A if it is reflexive, symmetric, and transitive.

Remark 3.40 (Role in mathematics). Equivalence relations axiomatize the idea of "sameness up to a chosen criterion." They partition sets into equivalence classes and are the basis for quotient constructions throughout algebra, topology, and analysis.

Definition (Preorder)

R is a preorder on A if it is reflexive and transitive.

Definition (Partial Order)

R is a partial order on A if it is reflexive, antisymmetric, and transitive.

Definition (Total Order)

R is a total order on A if it is a partial order and is total.

Remark 3.41 (Using structural properties in proofs). When R is asserted to be an equivalence relation or partial order, this is shorthand for the conjunction of its defining properties. In proofs, cite only the specific component needed: "since R is a partial order, antisymmetry implies ..." rather than restating all properties.

Example 3.42 (Using a structural property in a proof). Let R be an equivalence relation on A. For any $a, b \in A$, $(a, b) \in R \Rightarrow [a] = [b]$, where $[a] = \{x \in A \mid (a, x) \in R\}$.

Remark 3.43 (Properties are logically independent). No basic property implies another in general. A relation may be transitive without being reflexive, or symmetric without being transitive. The structural classes are distinguished precisely by requiring specific combinations.

3.1.6 Indexed Families of Sets

Indexed Families Quick Reference

Concept	Meaning	Detail
Indexed family	Function $F : I \to \mathcal{P}(U)$; written $\{A_i\}_{i \in I}$	Def
Indexed union	$\bigcup_{i \in I} A_i$: in at least one A_i	Def
Indexed intersection	$\bigcap_{i \in I} A_i$: in every A_i	Def
Pairwise disjoint	$i \neq j \Rightarrow A_i \cap A_j = \varnothing$	Def
Cover	$\bigcup_{C \in \mathcal{C}} C = A$	Def
Arbitrary product	$\prod_{i \in I} A_i$: choice functions $f : I \to \bigcup A_i$	Def

Remark 3.44 (Motivation). Many set-theoretic constructions require not just pairs of sets but infinite families. Indexed families provide the formal language for partitions, equivalence classes, unions over countably or uncountably many sets, and products indexed by arbitrary index sets.

Definition (Indexed Family of Sets)

Let I be a set (the index set) and U a universe. An indexed family of sets is a function

$$F : I \to \mathcal{P}(U),$$

with $F(i)$ typically written A_i. The family is denoted $\{A_i\}_{i \in I}$.

Remark 3.45 (Family vs. set of sets). An indexed family is formally a function, not a set of sets. Different indices may correspond to the same set: $i \neq j$ does not imply $A_i \neq A_j$. This distinction matters for equivalence class constructions.

Definition (Indexed Union)

The union of the indexed family $\{A_i\}_{i \in I}$ is:

$$\bigcup_{i \in I} A_i \;:=\; \{\, x \mid \exists i \in I \text{ such that } x \in A_i \,\}.$$

Definition (Indexed Intersection)

For $I \neq \varnothing$, the intersection of $\{A_i\}_{i \in I}$ is:

$$\bigcap_{i \in I} A_i \;:=\; \{\, x \mid \forall i \in I,\; x \in A_i \,\}.$$

Remark 3.46 (Why $I \neq \varnothing$ is required). If $I = \varnothing$, then $\forall i \in I, x \in A_i$ holds vacuously for every x, making the intersection the entire universe U. This is normally left undefined or requires specifying a background universe explicitly.

Definition (Pairwise Disjoint Family)

The family $\{A_i\}_{i \in I}$ is pairwise disjoint if

$$\forall i, j \in I,\; i \neq j \to A_i \cap A_j = \varnothing.$$

Definition 3.47 (Cover). A collection $\mathcal{C} \subseteq \mathcal{P}(A)$ is a cover of A if

$$\bigcup_{C \in \mathcal{C}} C = A.$$

3.1.7 Arbitrary Cartesian Products

Definition (Arbitrary Cartesian Product)

Let $\{A_i\}_{i \in I}$ be an indexed family of sets. The Cartesian product is:

$$\prod_{i \in I} A_i \;:=\; \left\{ f : I \to \bigcup_{i \in I} A_i \;\middle|\; \forall i \in I,\; f(i) \in A_i \right\}.$$

Remark 3.48 (Elements as choice functions). An element of $\prod_{i \in I} A_i$ is a choice function: a function that assigns to each index i an element of the corresponding set A_i. For finite $I = \{1, \ldots, n\}$, this reduces to the familiar n-tuple $(a_1, \ldots, a_n)$.

Remark 3.49 (Axiom of Choice connection). For infinite I, the product $\prod_{i \in I} A_i$ is nonempty iff a choice function exists. This existence is not guaranteed by the other ZFC axioms: it is equivalent to the Axiom of Choice. Thus the Axiom of Choice is precisely the assertion that arbitrary products of nonempty sets are nonempty.

3.1.8 Equivalence Classes and Partitions

Equivalence and Partitions Quick Reference

Concept	Meaning	Detail
Equivalence class	$[a]_R = \{x \in A \mid (a,x) \in R\}$	Def
Quotient set	A/R: all equivalence classes	Def
Index of R	Cardinality of A/R	Def
Canonical surjection	$\pi : A \to A/R$, $\pi(a) = [a]$	Def
Partition	Nonempty, disjoint, covering collection of subsets	Def
Rep. independence	$[a] = [b] \iff (a,b) \in R$	Lem
Equiv.–partition correspondence	Bijection between equiv. relations and partitions	Thm

Definition (Equivalence Class)

Let R be an equivalence relation on A. For $a \in A$, the equivalence class of a is:

$$[a]_R := \{\, x \in A \mid (a,x) \in R \,\}.$$

When R is clear from context, we write $[a]$.

Remark 3.50 (English reading). $[a]_R$ is the set of all elements that R declares "the same as a." Two elements lie in the same class iff they are related: this is precisely the Representative Independence Lemma below.

Definition (Quotient Set)

The quotient set of A by R is the set of all equivalence classes:

$$A/R := \{\, [a]_R \mid a \in A \,\}.$$

Definition 3.51 (Index of an Equivalence Relation). The index of R on A is the cardinality $|A/R|$, i.e. the number of equivalence classes.

Definition (Canonical Surjection)

The canonical surjection (quotient map) is the function

$$\pi : A \to A/R, \qquad \pi(a) := [a].$$

Remark 3.52 (Properties of π). π is surjective by construction. Elements $a, b \in A$ satisfy $\pi(a) = \pi(b)$ iff $(a,b) \in R$. The canonical surjection is the prototype for all quotient constructions: it reappears as the quotient homomorphism in algebra and the quotient map in topology.

Definition (Partition)

A partition of A is a collection $\mathcal{P}$ of subsets of A such that:

(i) every block is nonempty: $\forall P \in \mathcal{P},\ P \neq \varnothing$;

(ii) distinct blocks are disjoint: $\forall P, Q \in \mathcal{P},\ P \neq Q \to P \cap Q = \varnothing$;

(iii) the blocks cover A: $\bigcup_{P \in \mathcal{P}} P = A$.

The sets $P \in \mathcal{P}$ are called the blocks of the partition.

Remark 3.53 (Partition vs. cover). Every partition of A is a cover of A whose members are nonempty and pairwise disjoint. Partitions are precisely the covers satisfying the disjointness condition.

Lemma 3.54 (Representative Independence Lemma). Let R be an equivalence relation on A. For any $a, b \in A$,

$$[a] = [b] \iff (a, b) \in R.$$

Remark 3.55 (Consequence for quotient maps). This lemma is the key fact underlying well-definedness of functions on quotient sets: a function $f : A \to B$ defined by $f([a]) = \cdots$ is well-defined iff the formula gives the same output for all representatives of $[a]$.

Theorem 3.56 (Equivalence Relations and Partitions). Let A be a set.

(i) If R is an equivalence relation on A, then A/R is a partition of A.

(ii) If $\mathcal{P}$ is a partition of A, then the relation $R_{\mathcal{P}}$ defined by

$$(a, b) \in R_{\mathcal{P}} \iff \exists P \in \mathcal{P} \text{ with } a \in P \text{ and } b \in P$$

is an equivalence relation on A.

(iii) These constructions are inverse: $R_{A/R} = R$ and $A/R_{\mathcal{P}} = \mathcal{P}$.

Remark 3.57 (Significance). This theorem establishes a bijection between equivalence relations on A and partitions of A. The two perspectives—"same block" (partition) and "related" (equivalence relation)—are interchangeable and each is more natural in different contexts.

Example 3.58 (Extremal equivalence relations). On any set A:

1. The equality relation $(a, b) \in R \iff a = b$ gives singleton classes $[a] = \{a\}$ and the finest partition of A.
2. The universal relation $R = A \times A$ gives $[a] = A$ for all a, and the coarsest partition (one block).

All other equivalence relations on A lie strictly between these extremes.

3.1.9 Functions

Functions Quick Reference

Concept	Meaning	Detail
Function	Left-total, right-unique relation	Def
Domain / Codomain	$\mathrm{dom}(f) = A$, $\mathrm{cod}(f) = B$ for $f : A \to B$	Def
Image of function	$\mathrm{im}(f) = \{f(a) \mid a \in A\}$	Def
Image of set	$f(S) = \{f(a) \mid a \in S\}$ for $S \subseteq A$	Def
Preimage	$f^{-1}(T) = \{a \in A \mid f(a) \in T\}$ for $T \subseteq B$	Def
Fiber	$f^{-1}(\{b\})$: preimage of a singleton	Def
Graph	$\{(a, b) \in A \times B \mid b = f(a)\}$	Def
Injective	Distinct inputs $\Rightarrow$ distinct outputs	Def
Surjective	Every codomain element is achieved	Def
Bijective	Injective and surjective	Def
Identity	$\mathrm{id}_A(a) = a$	Def
Inclusion map	$\iota : A \hookrightarrow B$, $\iota(a) = a$ for $A \subseteq B$	Def
Composition	$(g \circ f)(a) = g(f(a))$	Def
Inverse	Defined for bijections; $f^{-1} \circ f = \mathrm{id}_A$	Def
Left / Right inverse	Section and retraction	Def
Restriction	$f\|_C : C \to B$	Def
Extension	$g : A' \to B$ with $g\|_A = f$	Def

Definition (Function)

Let A and B be sets. A function from A to B is a relation $f \subseteq A \times B$ such that:

(i) (Existence / left-total) for every $a \in A$, there exists $b \in B$ with $(a, b) \in f$;

(ii) (Uniqueness / right-unique) if $(a, b_1) \in f$ and $(a, b_2) \in f$ then $b_1 = b_2$.

If $(a, b) \in f$ we write $f(a) = b$.

Remark 3.59 (English reading). A function is a rule assigning to each input exactly one output. The two conditions formalize "defined everywhere" (existence) and "single-valued" (uniqueness).

Remark 3.60 (Function as relation). Every function is a relation, but not every relation is a function. The two conditions that distinguish functions are left-totality and right-uniqueness.

Definition (Domain and Codomain)

If f is a function from A to B, we write $f : A \to B$, where A is the domain $\mathrm{dom}(f)$ and B the codomain $\mathrm{cod}(f)$.

Definition (Image of a Function)

The image (or range) of $f : A \to B$ is:

$$\mathrm{im}(f) \;:=\; \{\, b \in B \mid \exists a \in A,\ f(a) = b \,\}.$$

The image may be a proper subset of the codomain.

Definition (Image of a Set)

For $S \subseteq A$, the image of S under f is:

$$f(S) \;:=\; \{\, f(a) \mid a \in S \,\}.$$

Note $f(A) = \mathrm{im}(f)$.

Definition (Preimage)

For $T \subseteq B$, the preimage (inverse image) of T under f is:

$$f^{-1}(T) \;:=\; \{\, a \in A \mid f(a) \in T \,\}.$$

Remark 3.61 (Preimage notation warning). $f^{-1}(T)$ does not denote an inverse function. It is defined for any function f, regardless of whether f is injective or bijective.

Definition (Fiber)

The fiber of f over $b \in B$ is the preimage of the singleton:

$$f^{-1}(\{b\}) \;=\; \{\, a \in A \mid f(a) = b \,\}.$$

Remark 3.62 (Fibers partition the domain). The collection $\{f^{-1}(\{b\}) \mid b \in \mathrm{im}(f)\}$ is a partition of A. Every function thus induces an equivalence relation $a_1 \sim_f a_2 \iff f(a_1) = f(a_2)$, whose equivalence classes are precisely the fibers.

Definition (Graph of a Function)

The graph of $f : A \to B$ is:

$$\mathrm{Graph}(f) \;:=\; \{\, (a, b) \in A \times B \mid b = f(a) \,\}.$$

Remark 3.63 (Function as graph). A relation $G \subseteq A \times B$ is the graph of a function $f : A \to B$ iff G is left-total and right-unique. Functions may therefore be identified with their graphs: a function is a special kind of relation.

Definition (Injective Function)

$f : A \to B$ is injective (one-to-one) if distinct elements of A have distinct images:

$$\forall a_1, a_2 \in A,\; f(a_1) = f(a_2) \implies a_1 = a_2.$$

Definition (Surjective Function)

$f : A \to B$ is surjective (onto) if every element of B is achieved:

$$\forall b \in B,\; \exists a \in A \text{ such that } f(a) = b.$$

Equivalently, $\mathrm{im}(f) = B$.

Definition (Bijective Function)

$f : A \to B$ is bijective if it is both injective and surjective.

Remark 3.64 (Fiber interpretation). Injectivity: each fiber has at most one element. Surjectivity: every fiber is nonempty. Bijectivity: every fiber has exactly one element.

Definition (Identity and Inclusion)

The identity function on A is $\mathrm{id}_A : A \to A$, $\mathrm{id}_A(a) = a$.
If $A \subseteq B$, the inclusion map is $\iota : A \hookrightarrow B$, $\iota(a) = a$.

Remark 3.65 (Identity vs. inclusion). Both send each element to itself, but the inclusion map has a strictly larger codomain when $A \subsetneq B$. The inclusion map is always injective.

Definition 3.66 (Constant Function). $f : A \to B$ is constant if there exists $b_0 \in B$ with $f(a) = b_0$ for all $a \in A$.

3.1.10 Composition, Inverses, and Set-Image Laws

Composition & Inverses Quick Reference

Result	Statement	Proof method	Detail
Associativity of $\circ$	$h \circ (g \circ f) = (h \circ g) \circ f$	element-wise	Thm
Identity and $\circ$	$f \circ \mathrm{id}_A = \mathrm{id}_B \circ f = f$	element-wise	Thm
Inj/Surj under $\circ$	Closed under composition; partial converses	element-wise	Thm
Inverse characterization	$f^{-1} \circ f = \mathrm{id}_A$, $f \circ f^{-1} = \mathrm{id}_B$	bijectivity	Thm
Inverse of composition	$(g \circ f)^{-1} = f^{-1} \circ g^{-1}$	verify both sides	Thm
One-sided inverses	Left inverse $\iff$ injective; right $\iff$ surjective		Thm
Preimage preserves ops	f^{-1} commutes with $\cup, \cap, \setminus, c$	element-wise	Thm
Image and set ops	Images preserve $\cup$; contain $\subseteq$ for $\cap$	counterexample	Thm

Definition (Composition)

Let $f : A \to B$ and $g : B \to C$. The composition $g \circ f : A \to C$ is:

$$(g \circ f)(a) := g\big(f(a)\big) \quad \text{for all } a \in A.$$

Remark 3.67 (Non-commutativity). Composition is generally not commutative: $g \circ f \neq f \circ g$ even when both are defined.

Theorem 3.68 (Associativity of Composition). For $f : A \to B$, $g : B \to C$, $h : C \to D$:

$$h \circ (g \circ f) = (h \circ g) \circ f.$$

Theorem 3.69 (Identity and Composition). For any $f : A \to B$:

$$f \circ \mathrm{id}_A = f \quad \text{and} \quad \mathrm{id}_B \circ f = f.$$

Theorem 3.70 (Injectivity and Surjectivity Under Composition). Let $f : A \to B$ and $g : B \to C$.

(i) If f and g are injective, then $g \circ f$ is injective.

(ii) If f and g are surjective, then $g \circ f$ is surjective.

(iii) If $g \circ f$ is injective, then f is injective.

(iv) If $g \circ f$ is surjective, then g is surjective.

Remark 3.71 (Partial converses). In (iii), g need not be injective. In (iv), f need not be surjective. Bijectivity of $g \circ f$ does not imply bijectivity of either f or g individually.

Definition (Inverse Function)

Let $f : A \to B$ be bijective. The inverse function is $f^{-1} : B \to A$ defined by: $f^{-1}(b) = a$ iff $f(a) = b$.

Theorem 3.72 (Characterization of Inverse Functions). Let $f : A \to B$ be bijective. Then

$$f^{-1} \circ f = \mathrm{id}_A \quad \text{and} \quad f \circ f^{-1} = \mathrm{id}_B .$$

Conversely, a function admits an inverse iff it is bijective.

Theorem 3.73 (Inverse of a Composition). Let $f : A \to B$ and $g : B \to C$ be bijective. Then $g \circ f$ is bijective and

$$(g \circ f)^{-1} = f^{-1} \circ g^{-1}.$$

Definition (Left and Right Inverses)

Let $f : A \to B$.
A left inverse of f is a function $g : B \to A$ with $g \circ f = \mathrm{id}_A$.
A right inverse (or section) of f is a function $h : B \to A$ with $f \circ h = \mathrm{id}_B$.

Theorem 3.74 (One-Sided Inverses and Function Properties). Let $f : A \to B$.

(i) f has a left inverse $\iff$ f is injective (and $A \neq \varnothing$ or $A = B = \varnothing$).

(ii) f has a right inverse $\iff$ f is surjective.

(iii) If f has both a left inverse g and a right inverse h, then $g = h$ and f is bijective.

Remark 3.75 (Axiom of Choice). The existence of a right inverse for every surjective function is equivalent to the Axiom of Choice.

Definition (Restriction and Extension)

Let $f : A \to B$ and $C \subseteq A$. The restriction of f to C is $f|_C : C \to B$, $f|_C(a) = f(a)$ for all $a \in C$.
Let $A \subseteq A'$, and let $f : A \to B$. A function $g : A' \to B$ is an extension of f if $g|_A = f$.

Remark 3.76 (Extensions not unique). An extension agrees with f on all of A but may be defined on a larger domain A'. Extensions are generally not unique; uniqueness requires additional constraints (such as continuity in analysis or linearity in algebra), leading to important results like the Tietze Extension Theorem and the Hahn–Banach Theorem.

Theorem 3.77 (Preimages Preserve Set Operations). Let $f : A \to B$ and $S, T \subseteq B$. Then

(i) $f^{-1}(S \cup T) = f^{-1}(S) \cup f^{-1}(T)$,

(ii) $f^{-1}(S \cap T) = f^{-1}(S) \cap f^{-1}(T)$,

(iii) $f^{-1}(S \setminus T) = f^{-1}(S) \setminus f^{-1}(T)$,

(iv) $f^{-1}(S^c) = (f^{-1}(S))^c$.

Remark 3.78 (Why preimages are well-behaved). Preimages preserve all set-theoretic operations. This property makes preimages central in topology and measure theory, where properties such as continuity and measurability are defined via preimages.

Theorem 3.79 (Images and Set Operations). Let $f : A \to B$ and $S, T \subseteq A$. Then

(i) $f(S \cup T) = f(S) \cup f(T)$,

(ii) $f(S \cap T) \subseteq f(S) \cap f(T)$,

(iii) $f(S \setminus T) \supseteq f(S) \setminus f(T)$.

Equality holds in (ii) and (iii) for all S, T iff f is injective.

Remark 3.80 (Asymmetry between images and preimages). Forward images preserve unions but generally not intersections or complements. This asymmetry between images and preimages is fundamental and explains why preimages appear more naturally in topology and measure theory.

3.1.11 Ordered Sets

Ordered Sets Quick Reference

Concept	Meaning	Detail
Ordered set	$(A, \leq)$: set with a partial order	Def
Strict order	$a < b \iff a \leq b \wedge a \neq b$	Def
Comparable / incomparable	$a \leq b$ or $b \leq a$ / neither	Def
Total (linear) order	Partial order with all pairs comparable	Def
Upper / lower bound	$u \geq s$ / $\ell \leq s$ for all $s \in S$	Def
Minimal / maximal	No smaller/larger element in S	Def
Least / greatest	Smaller/larger than all elements of S	Def
Order-preserving	$a \leq b \Rightarrow f(a) \leq' f(b)$	Def
Order isomorphism	Bijective order-preserving map with order-preserving inverse	Def
Well-ordered set	Every nonempty subset has a least element	Def
Chain / antichain	All comparable / none comparable	Def
Initial segment	Downward-closed subset	Def

Definition (Ordered Set)

An ordered set (or partially ordered set, poset) is a pair $(A, \leq)$ where $\leq$ is a partial order on A: a relation that is reflexive, antisymmetric, and transitive.

Remark 3.81 (English reading). An ordered set is a set equipped with a specified way to compare elements, but without requiring all pairs to be comparable.

Definition (Strict Order)

Let $(A, \leq)$ be an ordered set. The strict order $<$ is defined by:

$$a < b \iff (a \leq b \wedge a \neq b).$$

Remark 3.82 (Strict and non-strict are equivalent). The relation $<$ is irreflexive and transitive. Conversely, given any strict partial order $<$, one recovers a non-strict order by setting $a \leq b \iff (a < b \vee a = b)$. The two formulations carry exactly the same information.

Definition 3.83 (Comparable and Incomparable Elements). In an ordered set $(A, \leq)$, elements $a, b \in A$ are comparable if $a \leq b$ or $b \leq a$; they are incomparable if neither holds.

Remark 3.84 (Incomparable elements). Incomparable elements can only occur in partial orders. In a total order, every pair is comparable by definition.

Definition (Total / Linear Order)

An ordered set $(A, \leq)$ is a total order (or linear order) if every pair of elements is comparable:

$$\forall a, b \in A,\ a \leq b \vee b \leq a.$$

Remark 3.85 (Examples). $(\mathbb{R}, \leq)$ is a total order. $(\mathcal{P}(A), \subseteq)$ is a partial order that is generally not total.

Definition (Upper and Lower Bounds)

Let $(A, \leq)$ be an ordered set and $S \subseteq A$.
An element $u \in A$ is an upper bound of S if $\forall s \in S,\ s \leq u$.
An element $\ell \in A$ is a lower bound of S if $\forall s \in S,\ \ell \leq s$.

Remark 3.86 (Consequence). Bounds need not exist and need not be unique. They need not lie in S itself. Bounds are central to the definition of completeness for ordered fields.

Definition (Minimal, Maximal, Least, Greatest Elements)

Let $S \subseteq A$ in an ordered set $(A, \leq)$.
$m \in S$ is a minimal element of S if $\nexists s \in S$ with $s < m$.
$M \in S$ is a maximal element of S if $\nexists s \in S$ with $M < s$.
$\ell \in S$ is the least element if $\forall s \in S,\ \ell \leq s$.
$g \in S$ is the greatest element if $\forall s \in S,\ s \leq g$.

Remark 3.87 (Minimal vs. least). A least element is below every element of S; it must be unique if it exists. A minimal element merely has nothing below it within S; there may be many. Every least element is minimal, but not conversely.

Definition (Order-Preserving Map and Isomorphism)

Let $(M, \leq)$ and $(M', \leq')$ be partially ordered sets.
A function $f : M \to M'$ is order-preserving (or monotone) if

$$\forall a, b \in M, \quad a \leq b \implies f(a) \leq' f(b).$$

f is an order isomorphism if it is bijective and

$$\forall a, b \in M, \quad a \leq b \iff f(a) \leq' f(b).$$

Remark 3.88 (Order-isomorphic posets). An order isomorphism preserves and reflects the order structure exactly, including comparability, minimal and maximal elements, and bounds. Two posets are order-isomorphic if such a map exists; they are structurally identical from the order-theoretic viewpoint.

Definition (Well-Ordered Set)

An ordered set $(A, <)$ is well-ordered if every nonempty subset $S \subseteq A$ has a least element:

$$\forall S \subseteq A, \ \big(S \neq \varnothing \ \Rightarrow \ \exists m \in S \text{ s.t. } m \leq s \ \forall s \in S\big).$$

Remark 3.89 (Well-order implies total order). Every well-ordered set is totally ordered. The well-ordering condition is strictly stronger: it requires a least element in every nonempty subset, not just in A as a whole.

Example 3.90 (Well-order examples). $(\mathbb{N}, \leq)$ is well-ordered. $(\mathbb{Z}, \leq)$ is not: the set $\{\ldots, -3, -2, -1\}$ has no least element. $(\mathbb{R}, \leq)$ is not: $(0, 1)$ has no least element.

Remark 3.91 (Connection to ordinal numbers). An ordinal number is defined as an equivalence class of well-ordered sets under order isomorphism: it measures the order type of a well-ordered set, not merely its cardinality.

Definition 3.92 (Chain and Antichain). A subset $C \subseteq A$ of a poset $(A, \leq)$ is a chain if every pair of elements in C is comparable. It is an antichain if no two distinct elements of C are comparable.

Definition 3.93 (Initial Segment). A subset $I \subseteq A$ is an initial segment of $(A, \leq)$ if

$$a \in I \text{ and } b \leq a \ \Rightarrow \ b \in I.$$

Remark 3.94 (Transition). Ordered sets provide the abstract framework for order structures on the real numbers, function spaces, and metric spaces. In later sections, order interacts with topology and analysis through intervals, monotone functions, and the completeness axiom for $\mathbb{R}$.

3.1.12 Order Extensions

Order Extensions Quick Reference

Concept	Meaning	Detail
Preorder	Reflexive, transitive (not necessarily antisymmetric)	Def
Loset	Linear order = complete partial order	Def
Symmetric part	$xIy \iff xRy \wedge yRx$	Def
Asymmetric part	$xPy \iff xRy \wedge \neg(yRx)$	Def
Transitive closure	Smallest transitive relation containing R	Def
Extension	$\succsim'$ agrees with $\succsim$ on strict and weak ranks	Def
OWC	$xT(\succsim)y \Rightarrow \neg(y \succ x)$	Def
Szpilrajn's Thm	Every partial order extends to a linear order	Thm

Definition (Symmetric and Asymmetric Parts)

For any binary relation R on X, define:

- the symmetric part: $xIy \iff xRy$ and yRx
- the asymmetric part: $xPy \iff xRy$ but not yRx

Note that $R = P \cup I$.

Remark 3.95 (Preference-theoretic reading). In decision theory, R encodes a preference relation: xRy means "x is at least as good as y." Then I is the indifference relation (equally good) and P is the strict preference relation (strictly better). This decomposition is fundamental to utility representation theory.

Definition (Preorder and Loset)

A binary relation R on X is a preorder if it is reflexive and transitive.

A loset (linearly ordered set) is a pair (X, R) where R is a linear order: a complete partial order (R is reflexive, transitive, antisymmetric, and complete).

Remark 3.96 (Preorder vs. partial order). A preorder need not be antisymmetric: two distinct elements may satisfy xRy and yRx simultaneously (they are "indifferent" but not equal). Adding antisymmetry promotes a preorder to a partial order. Every partial order is a preorder, but not conversely.

Definition (Maximal Elements and Upper Bounds)

Let $\succsim$ be a binary relation on X.

- The set of maximal elements of X is

$$\mathrm{Max}(X, \succsim) = \{x \in X \mid y \succsim x \text{ for no } y \in X \text{ with } y \neq x\}.$$

- The set of upper bounds of X is

$$M(X, \succsim) = \{x \in X \mid x \succsim y \text{ for all } y \in X\}.$$

Remark 3.97 (Maximal vs. greatest). A maximal element has nothing strictly above it; an upper bound (greatest element) is above everything. In a partial order, the greatest element is unique if it exists and is always maximal, but maximal elements need not be greatest. In a linear order, maximal and greatest coincide.

Definition (Transitive Closure)

The transitive closure $T(R)$ of a binary relation R on X is the smallest transitive relation containing R: that is, $T(R)$ is transitive, $xRy \Rightarrow xT(R)y$, and no strictly smaller relation is both transitive and contains R.

Remark 3.98 (Constructive description). Define $R^0 = R$ and $xR^m y$ if there exist $z_1, \ldots, z_m \in X$ such that $xRz_1R \cdots Rz_mRy$. Then

$$T(R) = R \cup \bigcup_{m \in \mathbb{N}} R^m.$$

Existence follows from the fact that the intersection of any collection of transitive relations is transitive, so the intersection of all transitive supersets of R is the smallest such relation.

Definition (Extension of a Preorder)

Let $\succsim$ be a preorder on X. A preorder $\succsim'$ is an extension of $\succsim$ if

$$x \succsim y \;\Rightarrow\; x \succsim' y, \qquad x \succ y \;\Rightarrow\; x \succ' y.$$

A complete extension is an extension that is also complete.

Remark 3.99 (What extensions do). An extension of $\succsim$ "fills in" the comparisons left undecided by $\succsim$ without reversing any existing strict comparison. The goal is to promote a partial ranking to a total ranking while respecting the original judgements.

Theorem (Szpilrajn, 1930)

For any nonempty set X and partial order $\succsim$ on X, there exists a linear order that is an extension of $\succsim$.

Remark 3.100 (Proof sketch via Hausdorff Maximum Principle). Let T_X be the set of all partial orders on X extending $\succsim$, ordered by inclusion. By the Hausdorff Maximum Principle (equivalent to AC), T_X has a maximal chain; its union $\succsim^*$ is a partial order extending $\succsim$. If $\succsim^*$ were not complete,

one could enlarge it via transitive closure, contradicting maximality. Hence $\succsim^*$ is a linear order extending $\succsim$.

Dependence on AC. The theorem is equivalent to the Axiom of Choice for infinite sets; no proof avoiding AC is known.

Corollary 3.101 (Complete preorder extension). For any nonempty set X and preorder $\succsim$ on X, there exists a complete preorder that is an extension of $\succsim$.

Remark 3.102 (Proof). Pass to the quotient $X/\sim$ (where $\sim$ is the symmetric part), apply Szpilrajn's Theorem to obtain a linear order on $X/\sim$, then pull back to a complete preorder on X.

Definition and Proposition (Only Weak Cycles)

Let $\succsim$ be a binary relation on X with asymmetric part $\succ$ and transitive closure $T(\succsim)$. We say $\succsim$ satisfies only weak cycles (OWC) if

$$xT(\succsim)y \;\Rightarrow\; \neg(y \succ x).$$

Proposition. A binary relation $\succsim$ on a nonempty set X can be extended to a complete preorder if and only if it satisfies OWC.

Remark 3.103 (Intuition for OWC). OWC prohibits the following: there is a "chain" of weak preferences $x_1 \succsim x_2 \succsim \cdots \succsim x_n$ but a strict reversal $x_n \succ x_1$. Such a configuration cannot be extended to a complete preorder because the chain forces $x_1 \succsim' x_n$ in any extension, but strict preference $x_n \succ x_1$ in the extension would mean $x_n \succsim' x_1$ but not $x_1 \succsim' x_n$ contradicting $x_1 \succsim' x_n$.

Remark 3.104 (Transition to AC). Szpilrajn's Theorem relies on Zorn's Lemma / Hausdorff Maximum Principle, which are equivalent to the Axiom of Choice. The next section develops these equivalences directly.

3.1.13 Hasse Diagrams, Supremum and Infimum, and the Duality Principle

Hasse Diagrams, Sup/Inf, Duality Quick Reference

Concept	Meaning	Detail
Hasse diagram	Graph encoding a poset; edges only for cover relations	Def
Cover relation	$a \lessdot b$: $a < b$ with nothing strictly between	Def
Supremum (join)	Least upper bound: $\sup S = \min\{\, u \mid u \geq s \ \forall s \in S \,\}$	Def
Infimum (meet)	Greatest lower bound: $\inf S = \max\{\, \ell \mid \ell \leq s \ \forall s \in S \,\}$	Def
Dual poset	$(A, \geq)$ obtained by reversing the order on $(A, \leq)$	Def
Duality principle	Every theorem about posets yields a theorem in the dual	Prop

Definition (Cover Relation)

Let $(A, \leq)$ be a poset and $a, b \in A$. We say b covers a, written $a \lessdot b$, if $a < b$ and there is no $c \in A$ with $a < c < b$.

Remark 3.105 (Intuition). The cover relation strips the partial order down to its "essential edges." If $a \lessdot b$, then b is the immediate successor of a: you cannot insert anything between them. In a finite poset every strict relation $a < b$ factors as a finite chain of cover steps.

Definition (Hasse Diagram)

The Hasse diagram of a finite poset $(A, \leq)$ is the directed graph whose vertices are the elements of A, with an upward edge from a to b whenever $a \lessdot b$. Three graphical conventions apply:

(i) Reflexive loops ($a \leq a$) are suppressed.

(ii) Edges implied by transitivity are suppressed: if $a \lessdot b$ and $b \lessdot c$, no direct edge from a to c is drawn.

(iii) Direction is encoded by height: $a \lessdot b$ is drawn with b strictly above a, so arrowheads are unnecessary.

Remark 3.106 (Why Hasse diagrams work). Because we keep only cover edges, the diagram is uncluttered while still encoding the full order: $a \leq b$ holds in $(A, \leq)$ if and only if there is an upward path in the Hasse diagram from a to b (or $a = b$).

Example 3.107 (Divisibility on $\{1, 2, 3, 4, 6, 12\}$). Order the set $D = \{1, 2, 3, 4, 6, 12\}$ by divisibility ($a \leq b$ iff $a \mid b$). The covers are:

$$1 \lessdot 2, \quad 1 \lessdot 3, \quad 2 \lessdot 4, \quad 2 \lessdot 6, \quad 3 \lessdot 6, \quad 4 \lessdot 12, \quad 6 \lessdot 12.$$

The Hasse diagram places 1 at the bottom, 12 at the top, 2 and 3 on the next level up from 1, 4 and 6 above those, and 12 at the apex. The edge $1 \to 4$ is not drawn because $1 \lessdot 2 \lessdot 4$ already connects them through 2.

Example 3.108 (Power set $\mathcal{P}(\{a, b, c\})$ under inclusion). The Hasse diagram has four levels: $\varnothing$ at the bottom; the three singletons $\{a\}, \{b\}, \{c\}$ one level up; the three two-element sets above those; and $\{a, b, c\}$ at the top. Each singleton is covered by the two two-element sets containing it. This diagram is the Boolean lattice B_3.

Remark 3.109 (Limitation to finite posets). For infinite posets such as $(\mathbb{N}, \leq)$, the cover relation is well-defined ($n \lessdot n + 1$) but the full diagram cannot be drawn. In analysis, Hasse diagrams serve as a conceptual aid for finite or schematic examples rather than as a proof tool.

Definition (Supremum and Infimum)

Let $(A, \leq)$ be a poset and $S \subseteq A$.
The supremum (or least upper bound) of S, written $\sup S$, is an element $u^* \in A$ satisfying:

(i) u^* is an upper bound of S: $\forall s \in S,\ s \leq u^*$;

(ii) u^* is least among upper bounds: if u is any upper bound of S then $u^* \leq u$.

The infimum (or greatest lower bound) of S, written $\inf S$, is an element $\ell^* \in A$ satisfying:

(i) ℓ^* is a lower bound of S: $\forall s \in S,\ \ell^* \leq s$;

(ii) ℓ^* is greatest among lower bounds: if ℓ is any lower bound of S then $\ell \leq \ell^*$.

When they exist, $\sup S$ and $\inf S$ are unique.

Remark 3.110 (Uniqueness). If u^* and v^* are both suprema of S, then u^* is an upper bound so $v^* \leq u^*$, and v^* is an upper bound so $u^* \leq v^*$. Antisymmetry gives $u^* = v^*$. The same argument applies to infima.

Remark 3.111 (Sup and inf need not lie in S). The set $S = (0,1) \subseteq \mathbb{R}$ has $\sup S = 1$ and $\inf S = 0$, neither of which belongs to S. Contrast with $\max S$ and $\min S$, which are the supremum and infimum when they happen to lie in S. Every maximum is a supremum, but not conversely.

Remark 3.112 (Sup and inf need not exist). In the poset $(\mathbb{Q}, \leq)$, the set $\{r \in \mathbb{Q} : r^2 < 2\}$ is bounded above (e.g. by 2) but has no supremum in $\mathbb{Q}$: the candidate $\sqrt{2}$ is not rational. This failure is precisely what the completeness axiom for $\mathbb{R}$ addresses: it asserts that every nonempty subset of $\mathbb{R}$ that is bounded above does have a supremum in $\mathbb{R}$.

Remark 3.113 (Connection to greatest and least elements). $\sup S$ is the least element of the set of upper bounds of S in A; $\inf S$ is the greatest element of the set of lower bounds of S in A. This is why they are alternatively called the join and meet of S in lattice theory.

Example 3.114 (Sup and inf in the divisibility poset). In the divisibility poset $(D, |)$ with $D = \{1, 2, 3, 4, 6, 12\}$, take $S = \{4, 6\}$. The upper bounds of S (elements divisible by both 4 and 6) are $\{12\}$, so $\sup S = 12$. The lower bounds (elements dividing both 4 and 6) are $\{1, 2\}$; the greatest of these is 2, so $\inf S = 2$. Note: $\sup S = \operatorname{lcm}(4, 6)$ and $\inf S = \gcd(4, 6)$, revealing that gcd and lcm are order-theoretic concepts.

Definition (Dual Poset)

Let $(A, \leq)$ be a poset. The dual poset is $(A, \geq)$, where $a \geq b$ is defined to mean $b \leq a$. The dual is also written $(A, \leq^{\mathrm{op}})$.

Remark 3.115 (Intuition). The dual poset is obtained by flipping the Hasse diagram upside down. Every structural feature is preserved but reflected: what was at the top is now at the bottom. Minimal elements become maximal, upper bounds become lower bounds, and the supremum becomes the infimum.

Proposition 3.116 (Duality Principle). Let Φ be any first-order statement about a poset $(A, \leq)$ expressed using only the relation $\leq$. Let Φ^* be the statement obtained from Φ by replacing every occurrence of $\leq$ with $\geq$ (equivalently, working in the dual poset). Then:

$$(A, \leq) \models \Phi \iff (A, \geq) \models \Phi.$$

In particular, if Φ is a theorem about all posets, then so is Φ^*.

Remark 3.117 (Practical use). The duality principle means theorems come in pairs at no extra cost. Some instances you will use repeatedly in analysis:

- Supremum $\leftrightarrow$ Infimum. Every theorem about sup yields a theorem about inf by duality. For instance, if sup of a bounded-above set exists, duality (applied to the negation of the order) gives that inf of a bounded-below set exists and this is exactly how the infimum is deduced from the completeness axiom for $\mathbb{R}$.
- Minimal $\leftrightarrow$ Maximal. A statement about minimal elements in a poset dualises to a statement about maximal elements.

- Well-ordering. $(\mathbb{N}, \leq)$ is well-ordered (every nonempty subset has a least element); the dual $(\mathbb{N}, \geq)$ is not well-ordered (the whole set has no greatest element). Duality preserves the form of the definition but does not preserve well-ordering; this asymmetry is a key feature, not a failure of duality.

Remark 3.118 (Transition). With Hasse diagrams, sup/inf, and duality in hand, the order vocabulary of this chapter is complete. The next payoff is in Volume II: when $\mathbb{R}$ is constructed, the completeness axiom is the assertion that every nonempty bounded-above subset of $\mathbb{R}$ has a supremum — an order-theoretic condition that cannot be satisfied by $\mathbb{Q}$, as the example above shows.

3.1.14 Induced Orders and Order Embeddings

Induced Orders and Order Embeddings — Quick Reference

Concept	Meaning	Detail
Induced preorder	$x \leq_f y \Leftrightarrow f(x) \leq' f(y)$; always a preorder	Def
Induced partial order	Induced order is a partial order iff f is injective	Prop
f-indistinguishable	$x \sim_f y$: both $x \leq_f y$ and $y \leq_f x$	Def
Order embedding	Injective, order-preserving, and order-reflecting	Def
Embedding vs. isomorphism	Embedding is iso onto image; isomorphism is surjective too	Prop
Suborder	Restriction of $\leq'$ to a subset $S \subseteq B$	Def

Key results:

$\leq_f$ preorder: always. Antisymmetry: iff f injective.
f order embedding $\Rightarrow (A, \leq_f) \cong (f(A), \leq')$.

The simplest way to put an order on a set A is to compare elements of A indirectly through a function into an already-ordered set. This construction, called the induced order or pullback order, is ubiquitous: it explains how the usual order on $\mathbb{Z}$ restricts to $\mathbb{N}$, how functions on a common domain acquire a pointwise order, and how subsets inherit the order of the ambient space.

Definition (Induced Order)

Let $(B, \leq')$ be a partially ordered set and let $f : A \to B$ be a function. The order induced by f (or pullback order along f) is the relation $\leq_f$ on A defined by:

$$x \leq_f y \iff f(x) \leq' f(y).$$

Remark 3.119 (Intuition). Two elements of A are compared by sending them through f and comparing their images in $(B, \leq')$. The order on A is entirely inherited from B via f: we never compare elements of A directly, only their f-values.

Remark 3.120 (Fully quantified form). $\forall x, y \in A, \quad x \leq_f y \iff f(x) \leq' f(y)$.

Proposition 3.121 ($\leq_f$ is always a preorder). For any function $f : A \to B$ and partial order $(B, \leq')$, the relation $\leq_f$ is a preorder on A: it is reflexive and transitive.

Remark 3.122 (Proof strategy). Both properties are inherited directly from $(B, \leq')$ by pulling back through f: reflexivity at x uses reflexivity at $f(x)$, and transitivity along x, y, z uses transitivity along $f(x), f(y), f(z)$.

Definition (f-Indistinguishable Elements)

Let $f : A \to B$ and $\leq_f$ be the induced order. Two elements $x, y \in A$ are f-indistinguishable, written $x \sim_f y$, if both $x \leq_f y$ and $y \leq_f x$, i.e. if

$$f(x) \leq' f(y) \quad \text{and} \quad f(y) \leq' f(x).$$

Remark 3.123 (Intuition). Two elements are f-indistinguishable when they map to the same position in the $\leq'$-order not necessarily to the same point, but to mutually comparable points. When $\leq'$ is a partial order (antisymmetric), $x \sim_f y$ forces $f(x) = f(y)$, making f non-injective the only obstruction to antisymmetry.

Proposition 3.124 ($\leq_f$ is a partial order iff f is injective). Let $(B, \leq')$ be a partially ordered set and $f : A \to B$. The induced order $\leq_f$ is a partial order on A if and only if f is injective.

Remark 3.125 (Common error). It is tempting to think the induced order is automatically a partial order "because $\leq'$ is one." It is not. The issue is antisymmetry: if f collapses two distinct points $x \neq y$ to the same or comparable images, we get $x \leq_f y$ and $y \leq_f x$ with $x \neq y$, violating antisymmetry. Reflexivity and transitivity lift freely; antisymmetry does not.

Example 3.126 (Non-injective f destroys antisymmetry). Let $B = (\mathbb{N}, \leq)$ and let $f : \{a, b\} \to \mathbb{N}$ be the constant function $f(a) = f(b) = 0$. Then $a \leq_f b$ (since $0 \leq 0$) and $b \leq_f a$ (since $0 \leq 0$), but $a \neq b$. So $\leq_f$ is not antisymmetric, confirming that non-injectivity forces a failure.

Example 3.127 (Injective f gives a genuine partial order). Let $B = (\mathcal{P}(\{1, 2, 3\}), \subseteq)$ and let $A = \{x, y, z\}$ with $f(x) = \{1\}$, $f(y) = \{2\}$, $f(z) = \{1, 2\}$. Then f is injective. The induced order has $x \leq_f z$ (since $\{1\} \subseteq \{1, 2\}$) and $y \leq_f z$ (since $\{2\} \subseteq \{1, 2\}$), but x and y are incomparable. This is a genuine partial order on A.

Definition (Suborder / Restriction)

Let $(B, \leq')$ be a partially ordered set and $S \subseteq B$. The suborder on S (or induced order on S) is the relation $\leq'_S$ defined by:

$$x \leq'_S y \iff x \leq' y, \quad \text{for } x, y \in S.$$

Remark 3.128 (Connection to induced orders). The suborder on $S \subseteq B$ is exactly the order induced by the inclusion map $\iota : S \hookrightarrow B$, $\iota(x) = x$. Since ι is injective, the suborder is always a partial order whenever $(B, \leq')$ is. This justifies speaking of "S with the inherited order" without further verification.

Definition (Order Embedding)

Let $(A, \leq)$ and $(B, \leq')$ be partially ordered sets. A function $f : A \to B$ is an order embedding if for all $x, y \in A$:

$$x \leq y \iff f(x) \leq' f(y).$$

Remark 3.129 (Two directions, one condition). The ($\Rightarrow$) direction says f is order-preserving (see Definition (Order-Preserving Map)). The ($\Leftarrow$) direction — $f(x) \leq' f(y) \Rightarrow x \leq y$ — is called order-reflection. An order embedding both preserves and reflects the order, so f is a perfect local copy of $(A, \leq)$ inside $(B, \leq')$.

Remark 3.130 (Relation to the induced order). If $f : A \to B$ is an order embedding from $(A, \leq)$ to $(B, \leq')$, then $\leq$ coincides with the induced order $\leq_f$:

$$x \leq y \iff f(x) \leq' f(y) \iff x \leq_f y.$$

In other words, an order embedding is exactly an injective function whose induced order agrees with the existing order on A.

Proposition 3.131 (Order embeddings are injective). Every order embedding $f : (A, \leq) \to (B, \leq')$ is injective.

Remark 3.132 (Proof strategy). Injectivity falls out of the reflection direction: if two elements have the same image, reflection forces them to be mutually $\leq$-related, and antisymmetry collapses them to equality.

Proposition 3.133 (Order embedding is isomorphism onto image). If $f : (A, \leq) \to (B, \leq')$ is an order embedding, then f is an order isomorphism from $(A, \leq)$ to the suborder $(f(A), \leq'_{f(A)})$.

Remark 3.134 (Consequence). This proposition is the order-theoretic analogue of the first isomorphism theorem: an order embedding always realises $(A, \leq)$ as an isomorphic copy of a sub-poset of $(B, \leq')$. When f is also surjective, the embedding becomes a full order isomorphism (see Definition (Order Isomorphism)).

Example 3.135 (Standard embeddings). Each of the following is an order embedding:

(i) The inclusion $\mathbb{N} \hookrightarrow \mathbb{Z}$ with standard $\leq$ on both: $m \leq n \iff m \leq n$ trivially. The natural numbers sit inside the integers as an isomorphic copy of $(\mathbb{N}, \leq)$.

(ii) The function $f : \mathbb{N} \to \mathbb{N}$, $f(n) = 2n$, embeds $(\mathbb{N}, \leq)$ into itself: $m \leq n \iff 2m \leq 2n$. The image is the even numbers with the inherited order.

(iii) The map $A \mapsto A$ from $(\mathcal{P}(X), \subseteq)$ to itself is an order isomorphism (the identity), but any injective $g : \mathcal{P}(X) \to \mathcal{P}(X)$ that preserves and reflects $\subseteq$ is an order embedding.

Remark 3.136 (Transition). Induced orders and embeddings are the order-theoretic counterpart of substructures in algebra. Just as a subgroup is a subset closed under the group operations, a sub-poset is a subset with the inherited order — and an order embedding is the morphism that witnesses this. In analysis, these ideas appear when restricting an order from $\mathbb{R}$ to subsets (intervals, sequences, function spaces), and when the completeness property is transferred between ordered structures.

3.1.15 Zorn's Lemma and the Axiom of Choice

AC Equivalences Quick Reference

Statement	Role
Axiom of Choice (AC)	Choice function from any collection of nonempty sets
Product form of AC	Cartesian product of nonempty sets is nonempty
Zorn's Lemma	Poset with every chain bounded $\Rightarrow$ maximal element
Hausdorff Maximum Principle	Every poset has a $\subseteq$-maximal chain
Well-Ordering Theorem	Every set admits a well-ordering

All five statements are equivalent (over ZF set theory).

Axiom of Choice

Let $\mathcal{A}$ be any collection of nonempty sets. Then there exists a function

$$f : \mathcal{A} \to \bigcup \mathcal{A}$$

such that $f(A) \in A$ for all $A \in \mathcal{A}$.

Equivalent form. The Cartesian product of an arbitrary collection of nonempty sets is nonempty.

Remark 3.137 (Why AC is non-constructive). For finite collections, the axiom is provable without extra assumptions (see SRF-STO-C08-S8.1). For infinite collections, no construction can produce the choice function: AC asserts its existence without exhibiting it. This non-constructive character is both AC's power and the source of the foundational controversy surrounding it.

Zorn's Lemma

Let $(P, \leq)$ be a poset in which every chain (loset) has an upper bound in P. Then P has at least one maximal element.

Remark 3.138 (Reading Zorn's Lemma). The hypothesis "every chain has an upper bound" does not require the bound to lie in the chain itself. The conclusion gives a maximal element: an element above which nothing in P sits. In a partial order, there may be many maximal elements; Zorn guarantees at least one.

Remark 3.139 (Typical proof pattern using Zorn's Lemma). To show an object of type X exists with a maximal property:

1. Partially order candidate objects by inclusion (or extension).
2. Verify every chain of candidates has an upper bound (usually its union).
3. Conclude by Zorn that a maximal candidate exists.
4. Show the maximal candidate has the desired property (often: if it lacked it, it could be enlarged, contradicting maximality).

This pattern recurs throughout algebra and analysis: maximal ideals, algebraic bases (Hamel bases), maximal consistent sets, and Szpilrajn's Theorem.

Hausdorff Maximum Principle

In every poset $(P, \leq)$ there exists a $\subseteq$-maximal chain; that is, a chain $C \subseteq P$ such that no chain $C' \subseteq P$ strictly contains C.

Remark 3.140 (Relation to Zorn). The Hausdorff Maximum Principle is an immediate corollary of Zorn's Lemma: take the poset of chains ordered by inclusion; every chain of chains is bounded by its union; Zorn gives a maximal chain. Conversely, Hausdorff implies Zorn. Both are equivalent to AC over ZF.

Remark 3.141 (Application: Szpilrajn via Hausdorff). In the proof of Szpilrajn's Theorem (Thm), let T_X be the set of all partial orders on X extending $\succsim$, ordered by $\subseteq$. By the Hausdorff Maximum Principle, T_X has a maximal chain A; its union $\succsim^* = \bigcup A$ is a partial order extending $\succsim$. If $\succsim^*$ were

not total, there would exist incomparable x, y and one could extend $\succsim^*$ by adding (x, y) (closing under transitivity), contradicting maximality of A. Hence $\succsim^*$ is a linear order. $\square$

Remark 3.142 (Transition). Zorn's Lemma is the workhorse for existence proofs throughout abstract algebra and analysis. It first appears in the project context here and at the proof stubs in §11 of the Chapter VIII exercises.

3.2 Proofs

3.3 Capstone

3.4 Capstone Assessment: Sets, Relations, and Functions

Purpose. This capstone assesses mastery of elementary set theory, relations, and functions as used in analysis and abstract mathematics. All proofs must be written using precise definitions and logical reasoning.

Instructions. Each problem requires a complete proof. You must explicitly invoke definitions (e.g. function, equivalence relation) when they are used. No appeal to diagrams or intuition is permitted.

Problem 1 Equivalence Relations

Let $\sim$ be a relation on a set A. Prove that $\sim$ is an equivalence relation if and only if its equivalence classes form a partition of A.

Your proof must establish both directions.

Problem 2 Images and Preimages

Let $f : A \to B$ be a function and let $S \subseteq A$. Prove that

$$S \subseteq f^{-1}(f(S)).$$

Give an example where equality does not hold.

Problem 3 Injectivity and Left Inverses

Prove that a function $f : A \to B$ is injective if and only if there exists a function $g : B \to A$ such that

$$g \circ f = \mathrm{id}_A.$$

Your proof must clearly indicate where injectivity is used.

Problem 4 Composition of Relations

Let $R \subseteq A \times B$ and $S \subseteq B \times C$ be relations. Prove that if both R and S are transitive relations (on their respective domains), then their composition need not be transitive.

Your proof must include a concrete counterexample.

Problem 5 Order Relations

Let $(A, \leq)$ be a partially ordered set. Prove that $\leq$ is antisymmetric if and only if

$$(x \leq y \ \wedge \ y \leq x) \ \rightarrow \ x = y$$

holds for all $x, y \in A$.

Your proof must explicitly use the definition of antisymmetry.

Completion Criterion. You have mastered sets, relations, and functions if all five proofs:

- correctly invoke definitions,
- handle element-wise reasoning rigorously,
- distinguish relations from functions,
- and avoid implicit assumptions.

Successful completion certifies readiness to proceed to foundations of the real line and completeness.

Chapter 4

Axiom Systems

4.1 Notes

Where You Are in the Journey

Propositional Logic $\rightarrow$ Predicate Calculus $\rightarrow$ Sets & Functions $\rightarrow$ Proof Techniques $\rightarrow$ Axiom Systems $\rightarrow$ Natural Numbers ($\mathbb{N}$) $\rightarrow$ $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$ $\rightarrow$ $\cdots$

How we got here. Proof techniques gave us the tools to reason rigorously — direct proof, contradiction, induction. We now turn those tools on mathematics itself: can we build the natural numbers from nothing but logical axioms? This chapter answers yes.

What this chapter builds. We lay down the five Peano axioms, which pin down the structure of zero and the successor function. We verify that small numerals $(1, 2, 3, \ldots)$ exist, confirm that the axioms rule out pathological models, and prove the Recursion Theorem — the logical guarantee that recursive definitions actually produce well-defined functions. These are the logical primitives: the minimum needed for $\mathbb{N}$ to exist.

Where this leads. Volume II, Natural Numbers takes over immediately: it uses the Peano axioms and the Recursion Theorem as its foundation and builds the full arithmetic of $\mathbb{N}$ — addition, order, multiplication, and exponentiation — from scratch. Nothing from that development is assumed here.

Structural Roadmap

This chapter covers the logical skeleton of $\mathbb{N}$ only. Arithmetic (addition, order, multiplication, exponentiation) lives in Volume II, Natural Numbers.

$$\text{Axioms} \longrightarrow \text{Numerals} \longrightarrow \text{Induction principle} \longrightarrow \text{Recursion theorem}$$

The global progression is:

1. Peano axioms (P1–P5): the five axioms that characterise $\mathbb{N}$ up to isomorphism. Includes the induction axiom and a discussion of what the axioms exclude.
2. Numerals: defining $1 := 0{+}{+}$, $2 := 1{+}{+}$, ... and verifying they are distinct natural numbers.
3. Recursion theorem (Tao Prop. 2.1.16): for any set X, base value $a \in X$, and function $f : X \to X$, there exists a unique $g : \mathbb{N} \to X$ satisfying $g(0) = a$ and $g(n{+}{+}) = f(g(n))$. This theorem is what makes recursive definitions legitimate.

Remark 4.1 (Scope boundary). The Peano axioms tell us $\mathbb{N}$ exists and has the right logical shape. The Recursion Theorem tells us recursive definitions are valid. Everything else — what $\mathbb{N}$ can do arithmetically — is in Volume II.

Remark 4.2 (Primary source). Tao, Analysis I, Chapter 2, ğ2.1.

4.1.0.1 The Peano Axioms

The Peano Axioms

P1. $0 \in \mathbb{N}$ (zero is a natural number)
Logical form: $0 \in \mathbb{N}$

P2. $n \in \mathbb{N} \Rightarrow n{+}{+} \in \mathbb{N}$ (closure under successor)
Logical form: $\forall n \in \mathbb{N},\ n{+}{+} \in \mathbb{N}$

P3. $n{+}{+} \neq 0$ for all $n \in \mathbb{N}$ (zero is not a successor)
Logical form: $\forall n \in \mathbb{N},\ n{+}{+} \neq 0$

P4. $n{+}{+} = m{+}{+} \Rightarrow n = m$ (successor is injective)
Logical form: $\forall n, m \in \mathbb{N},\ n{+}{+} = m{+}{+} \Rightarrow n = m$

P5. Induction: if $P(0)$ is true and $P(n) \Rightarrow P(n{+}{+})$ for all $n \in \mathbb{N}$, then $P(n)$ is true for all $n \in \mathbb{N}$.
Logical form: $\big[P(0) \wedge \forall n \in \mathbb{N},\ P(n) \Rightarrow P(n{+}{+})\big] \Rightarrow \forall n \in \mathbb{N},\ P(n)$

Remark 4.3 (Why five axioms?). P1–P2 build $\mathbb{N}$ upward from 0. P3 prevents wrap-around (e.g. $3{+}{+} = 0$). P4 prevents collapse (e.g. $4{+}{+} = 2{+}{+}$ with $4 \neq 2$). P5 excludes rogue elements (e.g. $0.5 \in \mathbb{N}$) by asserting $\mathbb{N}$ is the smallest set satisfying P1–P2.

Remark 4.4 (P5 is second-order). P5 quantifies over properties P, not just elements. This makes it an axiom schema (one axiom per property P), not a single first-order axiom.

Remark 4.5 (Axioms define structure, not objects). The Peano axioms do not say what 0 or $n{+}{+}$ are — they say what properties they have. 0 could be the empty set $\varnothing$, or the singleton $\{\varnothing\}$, or an abstract symbol with no set-theoretic content. None of this matters. What matters is that the five axioms are satisfied. Any two systems satisfying P1–P5 are isomorphic: there is a unique structure-preserving bijection between them. This is why we speak of the natural numbers rather than a natural number system — all valid models are interchangeable for mathematical purposes. This is what it means for an axiom system to be categorical (in second-order logic). In first-order

logic the situation is different: first-order Peano arithmetic admits nonstandard models containing "infinite" natural numbers not reachable from 0 by successors. Second-order induction (P5 as stated) rules these out.

Remark 4.6 (Tao's notation). Tao writes $n{+}{+}$ for the successor of n (increment), following computer-language convention. We adopt this notation throughout. $n{+}{+}$ and $S(n)$ refer to the same object.

Definition 4.7 (Standard numerals). We define:

$$1 := 0{+}{+}, \quad 2 := 1{+}{+} = (0{+}{+}){+}{+}, \quad 3 := 2{+}{+}, \quad \text{etc.}$$

In general, each standard numeral is the successor of the previous one.

Proposition 4.8 (Tao 2.1.4). 3 is a natural number.

Proposition 4.9 (Tao 2.1.6 — $4 \neq 0$). 4 is not equal to 0.

Remark 4.10 (Why prove $4 \neq 0$?). Without P3, a wrap-around system where $3{+}{+} = 0$ (like a 4-clock) satisfies P1 and P2. P3 rules this out, and the proof above shows P3 doing its work explicitly.

Proposition 4.11 (Tao 2.1.8 — $6 \neq 2$). 6 is not equal to 2.

Remark 4.12 (Pattern). Distinctness proofs always proceed by descending through P4 until reaching $4 \neq 0$ (established via P3). P3 provides the "ground" that stops the descent.

Theorem (Recursion on $\mathbb{N}$)

Let X be a set, let $a \in X$, and let $f : X \to X$ be a function. Then there exists a unique function $g : \mathbb{N} \to X$ such that:

$$g(0) = a \quad \text{and} \quad g(n{+}{+}) = f(g(n)) \quad \text{for all } n \in \mathbb{N}.$$

Remark 4.13 (Reading the theorem). The theorem has two parts. Existence: the recursive rule actually defines a function — there are no gaps (every n gets a value) and no contradictions (no n gets two different values). Uniqueness: there is only one such function. Two functions satisfying the same recursive rule must agree everywhere, proved by induction on n.

Remark 4.14 (Tao's version). Tao states this as Proposition 2.1.16 in a slightly more general form: instead of a single function $f : X \to X$, he allows a family $f_n : X \to X$ indexed by n. The version above is the standard form and suffices for all applications here.

Remark 4.15 (Why all five axioms are needed). • P1 anchors the recursion: $g(0) = a$ is well-defined.

- P2 guarantees $g(n{+}{+})$ is always reachable from $g(n)$.
- P3 ensures $g(0)$ is never overwritten by a step $g(n{+}{+})$, since $n{+}{+} \neq 0$ for all n.
- P4 ensures no two steps conflict: if $m \neq n$ then $m{+}{+} \neq n{+}{+}$, so $g(m{+}{+})$ and $g(n{+}{+})$ are defined independently.
- P5 (induction) proves g is defined at every $n \in \mathbb{N}$, not just the ones we can reach by hand.

Remove any one axiom and the construction breaks.

Remark 4.16 (Recursive definitions as the engine of arithmetic). Every arithmetic operation on $\mathbb{N}$ is an instance of this theorem with $X = \mathbb{N}$:

- Addition $(+m)$: $g(0) = m$, $f(x) = x{+}{+}$. This gives $g(n) = n + m$.
- Multiplication $(\times m)$: $g(0) = 0$, $f(x) = x + m$. This gives $g(n) = n \times m$.
- Exponentiation $(m^{\cdot})$: $g(0) = 1$, $f(x) = x \times m$. This gives $g(n) = m^n$.

Each new operation is defined in terms of the previous one. The recursion theorem guarantees all three are well-defined and unique.

4.1.1 von Neumann Numerals

4.1.1.1 Definitions and Theorems

Remark 4.17 (Purpose). This section isolates the set-theoretic toolkit sufficient to define the von Neumann numerals

$$0 := \varnothing, \qquad n^+ := n \cup \{n\},$$

and to prove the key structural facts:

$$m < n \iff m \in n,$$

and each numeral is transitive and well-ordered by $\in$.

Axiom 4.18 (Extensionality). For all sets A, B,

$$\big(\forall x\,(x \in A \leftrightarrow x \in B)\big) \;\Rightarrow\; A = B.$$

Axiom 4.19 (Empty Set). There exists a set with no elements, denoted $\varnothing$.

Axiom 4.20 (Pairing). For any sets a, b there exists $\{a, b\}$.

Definition 4.21 (Singleton). For any set a, define $\{a\} := \{a, a\}$ (exists by Pairing).

Axiom 4.22 (Union). For any set A there exists a set $\bigcup A$ such that

$$x \in \bigcup A \;\leftrightarrow\; \exists y \in A\ (x \in y).$$

Definition 4.23 (Binary union). For sets a, b, define

$$a \cup b := \bigcup \{a, b\}.$$

Proposition 4.24 (Basic properties of union). For all sets a, b and all x,

$$x \in a \cup b \;\leftrightarrow\; (x \in a \vee x \in b).$$

Definition 4.25 (von Neumann successor). For any set n, define its von Neumann successor by

$$S_V(n) := n \cup \{n\}.$$

Definition 4.26 (von Neumann numerals). Define recursively:

$$0_V := \varnothing, \qquad (n+1)_V := S_V(n_V) = n_V \cup \{n_V\}.$$

The sets n_V are called the von Neumann numerals.

Proposition 4.27 (First few von Neumann numerals).

$$0_V = \varnothing, \quad 1_V = \{0_V\} = \{\varnothing\}, \quad 2_V = \{0_V, 1_V\}, \quad 3_V = \{0_V, 1_V, 2_V\}, \ \ldots$$

More generally, for $n \geq 1$,

$$n_V = \{0_V, 1_V, \ldots, (n-1)_V\}.$$

Definition 4.28 (Transitive set). A set A is transitive if

$$\forall x \, (x \in A \Rightarrow x \subseteq A).$$

Equivalently, $\forall x \, \forall y \, ((y \in x \in A) \Rightarrow y \in A)$.

Proposition 4.29 (Each von Neumann numeral is transitive). For every n, the set n_V is transitive.

Proposition 4.30 (Successor is injective). For all sets a, b,

$$S_V(a) = S_V(b) \ \Rightarrow \ a = b.$$

Remark 4.31 (About Foundation). Proposition 4.33 is usually proved inside ZFC using the Axiom of Foundation (Regularity) to exclude membership cycles. If you are working in a foundation-free setting, injectivity of successor may require an extra assumption.

Proposition 4.32 (Order by membership). For natural numbers m, n,

$$m < n \iff m_V \in n_V.$$

Proposition 4.33 (Well-ordering by $\in$). The relation $\in$ well-orders the set $\mathbb{N}_V := \{n_V : n \in \mathbb{N}\}$, and this well-order is isomorphic to $(\mathbb{N}, <)$ via $n \mapsto n_V$.

4.1.1.2 Consequences

Corollary 4.34 (Each numeral is the set of its predecessors). For every n,

$$n_V = \{m_V : m < n\}.$$

Corollary 4.35 (Monotonicity). If $m < n$, then $m_V \subsetneq n_V$.

4.1.2 Zermelo Numerals

4.1.2.1 Definitions and Theorems

Remark 4.36 (Purpose). This section records a minimal set-theoretic toolkit (axioms, derived constructions, and lemmas) sufficient to define the Zermelo numerals

$$0 := \varnothing, \qquad n^+ := \{n\} \quad (n \in \mathbb{N}).$$

The goal is not to redevelop all of ZFC, but to isolate the ingredients used in this construction.

Axiom 4.37 (Extensionality). For all sets A, B,

$$\big(\forall x\,(x \in A \leftrightarrow x \in B)\big) \;\Rightarrow\; A = B.$$

Axiom 4.38 (Empty Set). There exists a set with no elements. We denote one such set by $\varnothing$:

$$\exists E\ \forall x\ (x \notin E).$$

Axiom 4.39 (Pairing). For any sets a, b there exists a set $\{a, b\}$ whose elements are exactly a and b:

$$\forall a\,\forall b\,\exists P\ \forall x\ (x \in P \leftrightarrow (x = a \vee x = b)).$$

Definition 4.40 (Singleton). For any set a, define the singleton of a by

$$\{a\} := \{a, a\},$$

whose existence follows from the Pairing Axiom (Axiom 4.42).

Proposition 4.41 (Uniqueness of the empty set). If E and E' have no elements, then $E = E'$.

Proposition 4.42 (Singleton is well-defined). For each set a there exists a unique set whose only element is a, denoted $\{a\}$.

Definition 4.43 (Zermelo successor). For any set n, define its Zermelo successor by

$$S_Z(n) := \{n\}.$$

Definition 4.44 (Zermelo numerals). Define recursively:

$$0_Z := \varnothing, \qquad (n+1)_Z := S_Z(n_Z) = \{n_Z\}.$$

The sets n_Z are called the Zermelo numerals.

Proposition 4.45 (First few Zermelo numerals).

$$0_Z = \varnothing, \quad 1_Z = \{\varnothing\}, \quad 2_Z = \{\{\varnothing\}\}, \quad 3_Z = \{\{\{\varnothing\}\}\}, \ \ldots$$

Proposition 4.46 (Injectivity of Zermelo successor). For all sets a, b,

$$S_Z(a) = S_Z(b) \;\Rightarrow\; a = b.$$

Proposition 4.47 (Distinctness of Zermelo numerals). If $m \neq n$ as natural numbers, then $m_Z \neq n_Z$ as sets. Equivalently, the map $n \mapsto n_Z$ is injective.

Proposition 4.48 (Membership pattern). For Zermelo numerals, we have:

$$m_Z \in n_Z \iff n = m + 1.$$

In particular, each n_Z has exactly one element when $n \geq 1$.

Remark 4.49 (Why Zermelo numerals are useful). Zermelo numerals are the iterated singleton tower. They are convenient when you want a very rigid notion of "n steps": each numeral (except 0) has exactly one element. They are less convenient for encoding order by membership (that role is played better by von Neumann numerals).

4.1.2.2 Consequences

Corollary 4.50 (No numeral contains itself). For every n, $n_Z \notin n_Z$.

Corollary 4.51 (Strict chain under membership). For $n \geq 1$,

$$(n-1)_Z \in n_Z, \quad (n-2)_Z \in (n-1)_Z, \ \ldots, \ 0_Z \in 1_Z.$$

4.2 Proofs

Tao 2.1.4 — 3 is a natural number

Remark 4.52 (Return). ← Back to Proposition (Tao 2.1.4) in Notes

Proposition 4.53 (Tao 2.1.4). 3 is a natural number.

Proof. Given. The standard numerals defined by $1 := 0{+}{+}$, $2 := 1{+}{+}$, $3 := 2{+}{+}$ (Definition 4.7). Goal. To show $3 \in \mathbb{N}$.

By P1, $0 \in \mathbb{N}$. By P2, applied to 0: $0{+}{+} \in \mathbb{N}$, i.e. $1 \in \mathbb{N}$. By P2, applied to 1: $1{+}{+} \in \mathbb{N}$, i.e. $2 \in \mathbb{N}$. By P2, applied to 2: $2{+}{+} \in \mathbb{N}$, i.e. $3 \in \mathbb{N}$.

Hence, $3 \in \mathbb{N}$. as required. □

Remark 4.54 (Proof shape). This is a chain of three applications of P2. Each step takes a known natural number and produces its successor. The chain is anchored by P1, which supplies the base element $0 \in \mathbb{N}$.

Remark 4.55 (Generalisation). The same pattern shows every standard numeral n is a natural number: by induction, $0 \in \mathbb{N}$ (P1) and if $n \in \mathbb{N}$ then $n{+}{+} \in \mathbb{N}$ (P2), so all numerals are natural numbers. Proposition 2.1.4 is just this argument made explicit for $n = 3$.

Tao 2.1.6 — $4 \neq 0$

Remark 4.56 (Return). ← Back to Proposition (Tao 2.1.6) in Notes

Proposition 4.57 (Tao 2.1.6 — $4 \neq 0$). 4 is not equal to 0.

Proof. Given. The numeral $4 := 3{+}{+}$ (Definition 4.7) and the Peano axiom P3: $n{+}{+} \neq 0$ for all $n \in \mathbb{N}$. Goal. To show $4 \neq 0$.

By Proposition 4.8, $3 \in \mathbb{N}$. By P3, applied to $n := 3$: $3{+}{+} \neq 0$. By definition, $4 = 3{+}{+}$.

Hence, $4 \neq 0$. as required. □

Remark 4.58 (The role of P3). Axiom P3 is precisely the statement that no natural number's successor is 0. Without it, a cyclic system such as $\{0, 1, 2, 3\}$ with $3{+}{+} = 0$ would satisfy P1 and P2 but fail this proposition. P3 is the axiom that rules out wrap-around.

Remark 4.59 (Dependency). This proof depends on Proposition 4.8 (so that $3 \in \mathbb{N}$ and P3 may be applied) and on Axiom P3. It is used in turn by Proposition 4.13: the descent argument for $6 \neq 2$ eventually bottoms out here.

Tao 2.1.8 — $6 \neq 2$

Remark 4.60 (Return). ← Back to Proposition (Tao 2.1.8) in Notes

Proposition 4.61 (Tao 2.1.8 — $6 \neq 2$). 6 is not equal to 2.

Proof. Given. The numerals $6 := 5{+}{+}$ and $2 := 1{+}{+}$, and the Peano axioms P3 and P4. Goal. To show $6 \neq 2$.

Strategy. We apply P4 repeatedly to reduce the inequality to the known result $4 \neq 0$ (Proposition 4.10).

Step 1. Suppose, toward a contradiction, that $6 = 2$. By definition, $6 = 5{+}{+}$ and $2 = 1{+}{+}$, so

$$5{+}{+} = 1{+}{+}.$$

Step 2. By P4, applied to the equation $5{+}{+} = 1{+}{+}$: $5 = 1$. Since $5 = 4{+}{+}$ and $1 = 0{+}{+}$, this gives $4{+}{+} = 0{+}{+}$.

Step 3. By P4, applied to $4{+}{+} = 0{+}{+}$: $4 = 0$.

Step 4. But by Proposition 4.10, $4 \neq 0$. This is a contradiction.

Hence, the assumption $6 = 2$ is false, so $6 \neq 2$. as required. □

Remark 4.62 (Proof shape). This is a proof by contradiction. The key move is applying P4 (injectivity of successor) twice to reduce $6 = 2$ to $4 = 0$, which has already been ruled out.

Remark 4.63 (The descent pattern). Every inequality $n \neq m$ between standard numerals with $n > m$ is proved by the same descending argument: strip one successor from each side via P4, repeat until one side reaches 0, then apply P3 (via Proposition 4.10 or P3 directly) to obtain the contradiction. The number of P4-applications equals $\min(n, m)$ when both are positive. Here $\min(6, 2) - \min(6, 2) = 0$ after adjusting, and we need exactly two applications.

Remark 4.64 (Dependencies). This proof uses Proposition 4.10 ($4 \neq 0$) as its contradiction target, and Axioms P3 and P4.

4.3 Capstone

Chapter 5

Proof Techniques

5.1 Notes

Where You Are in the Journey

Propositional Logic $\rightarrow$ Predicate Calculus $\rightarrow$ Sets & Functions $\rightarrow$ Proof Techniques $\rightarrow$ Real Analysis $\rightarrow$ Algebraic Structures $\rightarrow$ Linear Algebra $\rightarrow \cdots$

How we got here. Logic gave us the language of valid inference and sets gave us the objects we reason about. This chapter asks the operational question: how do we build a valid mathematical argument from start to finish?

What this chapter builds. A complete architecture for proof writing: how to read a statement and identify its proof strategy immediately; how to execute that strategy step by step; how to use induction to reason about well-ordered and recursive structures; and how to deploy specific algebraic tactics when progress stalls.

Where this leads. Every proof in every subsequent chapter uses this toolkit. Epsilon-delta analysis uses direct proof and contradiction. Algebraic structure proofs use the uniqueness and satisfy-and-cite patterns. Induction appears throughout number theory, algebra, and combinatorics.

Structural Roadmap

This chapter is organized in six sections of increasing specificity.

$$\text{Architecture} \longrightarrow \text{Construction} \longrightarrow \text{Structures} \longrightarrow \text{Induction} \longrightarrow \text{Tactics} \longrightarrow \text{Reference}$$

Sections 1–4 should be read sequentially on first pass. Sections 5–6 are reference material to consult when working a proof and progress stalls.

Remark 5.1 (Four distinct questions). Architecture answers: what kind of proof is this and what skeleton should it have? Construction answers: how do I execute that skeleton line by line? Structures answers: what do the standard proof patterns look like in practice? Tactics answers: once a proof is in progress, what specific moves are available? These are four different questions. Conflating them is the most common source of proof-writing paralysis.

5.1.0.1 Proof Architecture

The Two Questions Before You Write Anything. Before writing a single symbol, ask two questions in order.

> Question 1. What is the logical form of the statement?
> Read the claim and identify its outermost logical structure: universal, conditional, biconditional, existential, uniqueness, equality, or structural assertion.
>
> Question 2. What proof strategy does this form suggest?
> The form determines the skeleton of the proof mechanically, before any mathematics is done.

Remark 5.2 (Why order matters). Question 2 cannot be answered before Question 1. The most common proof error is reaching for a familiar strategy — contradiction is the usual default — without checking whether it matches the statement form. Read the statement first. The form is not preliminary; it is the first mathematical act of the proof.

Example 5.3 (Applying the two questions). Statement: The identity element of a group is unique.

Q1. Logical form: $\forall G$ (group), $\forall e, e' \in G$, [both satisfy the identity axiom] $\Rightarrow e = e'$. This is a universal statement with a uniqueness conclusion.

Q2. Strategy: introduce an arbitrary group G, assume two identity elements e and e' exist, show $e = e'$. This is the assume-two-and-compare pattern (see Section 5.1.0.3).

The strategy is forced entirely by the form. No creative insight is required to select it.

Statement Forms and Their Proof Strategies. The following table maps every common statement form to the proof strategy it demands. This table should be consulted at the start of every proof until the correspondences are automatic.

Statement Form	Proof Strategy	Opening Move
$\forall x,\ P(x)$	Introduce arbitrary element	"Let x be arbitrary."
$P \Rightarrow Q$	Direct proof	"Assume P."
$P \Rightarrow Q$	Contrapositive	"Assume $\neg Q$."
$\neg Q \Rightarrow \neg P$ (same as above)	Contradiction	"Assume P and $\neg Q$."
$P \Leftrightarrow Q$	Two directions	Prove $P \Rightarrow Q$, then $Q \Rightarrow P$.
$\exists x,\ P(x)$	Construct a witness	"Define $x := \ldots$ and verify $P(x)$."
$\exists! x,\ P(x)$	Existence + uniqueness	Construct witness; then assume-two-and-compare.
$A = B$ (sets)	Double inclusion	Prove $A \subseteq B$ and $B \subseteq A$.
$a = b$ (elements)	Satisfy-and-cite or algebra	Show a satisfies the defining property of b.
$P(n)$ for all $n \in \mathbb{N}$	Induction	Base case; inductive step.
"X is unique"	Assume-two-and-compare	"Let x, y both satisfy the definition."

Remark 5.4 (Multiple strategies for one form). Some statement forms admit more than one strategy. $P \Rightarrow Q$ can be proved directly, by contrapositive, or by contradiction. The table lists the most natural default. Direct proof should be attempted first; contrapositive and contradiction are reached for when the hypothesis is hard to use forward or when the negation of the conclusion is more tractable.

Remark 5.5 (The equality forms are different). Notice that $A = B$ (set equality) and $a = b$ (element equality) suggest different strategies. Set equality uses double inclusion because sets are defined by their members. Element equality in an algebraic context uses satisfy-and-cite: show that a satisfies the same defining property as b, then invoke a previously proved uniqueness theorem. See Section 5.1.0.1.

The Five Proof Archetypes. At the highest level of compression, every mathematical proof reduces to one of five archetypes. Everything else is refinement of these five.

The Five Archetypes

1. Construct something. Exhibit a concrete object and verify it has the required property. Used for all existence claims.
2. Assume two and compare. Assume two objects both satisfy a definition; derive that they are equal. Used for all uniqueness claims.
3. Take a minimal element. Apply the well-ordering principle to a nonempty set; exploit the minimality of the least element. Used in number theory, induction, and contradiction arguments.
4. Induct. Verify a base case; show the property propagates to successors. Used whenever the claim concerns all natural numbers or a recursively defined structure.
5. Chase an arbitrary element. Take an arbitrary element of one set or satisfying one property; track it through definitions to land in the target. Used for set inclusions, function properties, and structural claims. Also may be used for direct proofs when the logic flows in a straight line.

Remark 5.6 (Using the archetypes). When starting a proof, identify which archetype applies before consulting the detailed strategy map. The archetype tells you the global shape of the argument. The statement map tells you the specific opening move. The construction algorithm (Section 5.1.0.2) tells you how to fill in each line.

Remark 5.7 (Archetypes can nest). A single proof may use more than one archetype. A common pattern in algebra is to construct a witness (archetype 1) and then invoke a uniqueness theorem (archetype 2) to conclude that the constructed object equals a previously named one. This nesting is the satisfy-and-cite tactic developed in Section 5.1.0.1.

The Satisfy-and-Cite Pattern. The most commonly missing pattern at the undergraduate level is one that does not appear in any of the standard proof archetypes by name, yet appears constantly in abstract algebra and analysis.

Definition 5.8 (Satisfy-and-Cite Pattern). To prove that $a = b$ in a context where b is a uniquely determined object (the unique identity, the unique inverse, the unique limit, the unique fixed point), proceed as follows:

1. Show that a satisfies the defining property of b.
2. Cite the previously proved uniqueness theorem for b.
3. Conclude $a = b$.

Remark 5.9 (Why this is different from assume-two-and-compare). The assume-two-and-compare pattern proves uniqueness: it shows that any two objects satisfying a definition are equal.

The satisfy-and-cite pattern uses a previously proved uniqueness result: it shows that a specific object a equals a specific named object b by showing a satisfies b's defining property, then invoking uniqueness to collapse a to b.

Assume-two-and-compare appears in the proof that the identity is unique. Satisfy-and-cite appears in every proof that uses that uniqueness theorem afterward.

Example 5.10 (Socks-shoes property: $(ab)^{-1} = b^{-1}a^{-1}$). We want to show $(ab)^{-1} = b^{-1}a^{-1}$.

The inverse of ab is the unique element x such that $(ab)x = e$ and $x(ab) = e$.

We compute:

$$(ab)(b^{-1}a^{-1}) = a(bb^{-1})a^{-1} = aea^{-1} = aa^{-1} = e.$$

Similarly $(b^{-1}a^{-1})(ab) = e$.

So $b^{-1}a^{-1}$ satisfies the defining property of $(ab)^{-1}$. By uniqueness of inverses (previously proved), $(ab)^{-1} = b^{-1}a^{-1}$. □

Notice: we never assumed two inverses existed and compared them. We showed one specific element satisfies the inverse definition, then cited uniqueness.

Example 5.11 (Ring theorem: $a \cdot (-b) = -(ab)$). We want to show $a(-b) = -(ab)$.

The additive inverse $-(ab)$ is the unique element x such that $ab + x = 0$.

We compute:

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

So $a(-b)$ satisfies the defining property of $-(ab)$. By uniqueness of additive inverses, $a(-b) = -(ab)$. □

Same pattern: compute, satisfy, cite.

Remark 5.12 (The pattern across mathematics). Satisfy-and-cite appears wherever unique objects are defined:

- Algebra: inverses, identities, quotients, kernels
- Analysis: limits, suprema, infima, fixed points
- Linear algebra: zero vector, additive inverses, linear transformations determined by a basis

Each time a uniqueness theorem is proved, it creates a new tool that can be activated by the satisfy-and-cite pattern for every future proof in that structure.

5.1.0.2 The Proof Construction Algorithm

Remark 5.13 (Purpose of this section). This section records a general proof-writing procedure intended to be applied consciously until it becomes internalized. It answers the question: once the architecture is chosen (Section 5.1.0.1), how is the argument executed line by line?

Step 0: Classify the Statement. Determine the logical form of the claim. Common forms include:

- Universal: $\forall x\ P(x)$

- Conditional: $P \to Q$
- Biconditional: $P \leftrightarrow Q$
- Existential: $\exists x\, P(x)$
- Equality: $X = Y$
- Set equality: $A = B$
- Structural: "R is an equivalence relation", "f is injective"

The logical form determines the overall structure of the proof.

Step 1: Restate the Givens. Explicitly record what is given. Introduce all sets, relations, and functions, and note any structural properties assumed.

> Let A be a set and let R be an equivalence relation on A.

This licenses later use of reflexivity, symmetry, and transitivity without reintroducing them each time.

Step 2: Identify Objects and Their Types. Before reasoning begins, identify the type of each object: is it an element, a set, a function, or a relation? What is its ambient universe?

$$x \in A, \quad f : A \to B, \quad R \subseteq A \times A.$$

Many logical errors arise from confusing equality, membership, and inclusion. Every object used in a proof must have a declared type.

Step 3: Introduce Arbitrary Elements. If the claim is universal, immediately introduce arbitrary elements.

$$\text{Let } a, b \in A \text{ be arbitrary.}$$

This step enables general reasoning and avoids illegal specialization.

Step 4: Expand the Goal. Rewrite the conclusion using definitions rather than named concepts.

- To prove f injective: expand the definition of injectivity.
- To prove $A = B$: prove $A \subseteq B$ and $B \subseteq A$.
- To prove $x \in A \cup B$: rewrite as $x \in A \vee x \in B$.

Never attempt to prove a named concept directly without first expanding its definition.

Step 5: Introduce Helper Objects. Introduce auxiliary objects needed for the argument: witnesses, intermediate elements, bounds, images under functions.

$$\text{Let } x \in A, \quad \text{let } P_a \text{ be the block containing } a, \quad \text{let } b := f(a).$$

No object should appear in a proof without being explicitly introduced.

Step 6: Apply One Definition or Property at a Time. Each step follows from exactly one of:

- a definition,
- a hypothesis,
- a previously proved theorem,
- a basic logical rule.

If progress stalls, identify which definition has not yet been unpacked.

Step 7: Use Forward and Backward Reasoning. Proofs often alternate between two modes:

- Forward: deduce consequences from the hypotheses.
- Backward: rewrite the goal to determine what would suffice to prove it.

Backward reasoning is effective when the conclusion involves nested definitions.

Step 8: Handle Cases Explicitly. If a statement splits into cases, enumerate and exhaust all of them.

$$\text{Either } x \in A \text{ or } x \notin A.$$

Each case is treated separately; together they cover all possibilities.

Step 9: Close the Argument. Once the desired conclusion is reached, state it explicitly.

- "Thus $x \in B$, so $A \subseteq B$."
- "Hence f is injective."
- "Therefore the two sets are equal."

Step 10: Signal Completion. Conclude with □ or an explicit statement that the proof is complete.

Legal Moves.

- An existential witness may not be chosen before existence is proved.
- An arbitrary element may only be introduced under universal scope.
- The conclusion may never be assumed.
- Definitions must be applied in full, not partially.

Line-by-Line Discipline. For each line, silently check:

(i) Has every object in this line been defined?

(ii) Which definition, hypothesis, or theorem justifies this step?

(iii) Does this step move the argument closer to the stated goal?

Stop Conditions. A proof is complete when:

- a universal claim has been shown for an arbitrary element;
- an existential claim has produced a valid witness;
- a set equality has established both inclusions;
- each direction of a biconditional has been proved.

Remark 5.14 (On internalizing the algorithm). This procedure is intentionally explicit and mechanical. With practice, these steps become automatic and are applied subconsciously. Experienced mathematicians follow the same process but omit intermediate steps once correctness is assured. Until fluency is achieved, longer proofs with explicit steps are preferable to shorter proofs with implicit gaps.

5.1.0.3 Proof Structures

Direct Proof. A direct proof of $P \Rightarrow Q$ assumes P and derives Q through a chain of valid inferences.

Template: Direct Proof

Given. P.
Goal. Q.
Assume P.
⋮ [expand definitions; derive consequences]
Therefore Q. □

Remark 5.15 (When to use direct proof). Direct proof is the default. Reach for it first. It is appropriate whenever the hypothesis gives you something concrete to work with and the conclusion follows by expanding definitions and applying theorems. Reserve contrapositive and contradiction for cases where the hypothesis is hard to use forward.

Example 5.16 (Direct proof in a group). Claim. In a group G, if $a^2 = e$ then $a = a^{-1}$.

Proof. Assume $a^2 = e$, i.e., $aa = e$. Multiply both sides on the right by a^{-1}:

$$aa \cdot a^{-1} = e \cdot a^{-1}.$$

By associativity, $a(aa^{-1}) = a^{-1}$, so $ae = a^{-1}$, so $a = a^{-1}$. □

Proof by Contrapositive. The contrapositive of $P \Rightarrow Q$ is $\neg Q \Rightarrow \neg P$. These are logically equivalent. A proof by contrapositive proves $\neg Q \Rightarrow \neg P$ directly.

Template: Contrapositive

Goal. $P \Rightarrow Q$.
Equivalent goal. $\neg Q \Rightarrow \neg P$.

Assume $\neg Q$.

⋮ [derive $\neg P$]
Therefore $\neg P$. Hence $P \Rightarrow Q$. □

Remark 5.17 (When to use contrapositive). Use contrapositive when the negation of Q is more informative or easier to manipulate than P itself. Injectivity proofs are a canonical example: to prove $f(a) = f(b) \Rightarrow a = b$, it is often cleaner to prove the contrapositive $a \neq b \Rightarrow f(a) \neq f(b)$.

Remark 5.18 (Contrapositive vs contradiction). Contrapositive: assume $\neg Q$, derive $\neg P$, done. Contradiction: assume both P and $\neg Q$, derive any absurdity, done. Contrapositive is cleaner when it works because the assumption is $\neg Q$ alone, not $P \wedge \neg Q$. Try contrapositive before reaching for contradiction.

Example 5.19 (Contrapositive for divisibility). Claim. If n^2 is even then n is even.

Proof (contrapositive). Assume n is odd. Then $n = 2k + 1$ for some $k \in \mathbb{Z}$.

$$n^2 = (2k+1)^2 = 4k^2 + 4k + 1 = 2(2k^2 + 2k) + 1,$$

which is odd. Hence n odd $\Rightarrow$ n^2 odd, i.e., n^2 even $\Rightarrow$ n even. □

Proof by Contradiction. To prove P, assume $\neg P$ and derive a statement that contradicts a known truth: a hypothesis, an axiom, or a previously proved theorem.

Template: Contradiction

Goal. P.

Suppose for contradiction that $\neg P$.

⋮ [derive a contradiction]
This contradicts [name the violated fact].
Therefore P. □

Remark 5.20 (When to use contradiction). Contradiction is best reserved for claims whose negation has strong consequences. It is the natural tool for:

- Irrationality proofs ($\neg P$ gives a fraction in lowest terms and leads to a parity contradiction).
- Infinitude of primes ($\neg P$ gives a finite list; constructing a number not on the list contradicts completeness).
- Statements involving nonexistence.

Avoid using contradiction as a default. When a direct proof or contrapositive works, it is cleaner.

Remark 5.21 (Name the contradiction explicitly). A well-written contradiction proof always names what is contradicted. "This contradicts [G2: identity axiom]" is better than "contradiction." Naming the violated fact confirms that the contradiction is genuine and locates the logical break.

Example 5.22 ($\sqrt{2}$ is irrational). Proof. Suppose for contradiction that $\sqrt{2} \in \mathbb{Q}$. Write $\sqrt{2} = p/q$ with $p, q \in \mathbb{Z}$, $q \neq 0$, and $\gcd(p, q) = 1$. Then $2q^2 = p^2$, so p^2 is even, so p is even (write $p = 2k$). Then $2q^2 = 4k^2$, so $q^2 = 2k^2$, so q is even. But then $\gcd(p, q) \geq 2$, contradicting $\gcd(p, q) = 1$. □

Proof by Cases. If the hypotheses or conclusion naturally divide into exhaustive and mutually exclusive alternatives, handle each case separately.

Template: Proof by Cases

Goal: Prove Q, given that $P_1 \vee P_2 \vee \cdots \vee P_k$ is exhaustive.

Verify exhaustiveness: Every possibility falls under some P_i.

Case 1: Assume P_1. [Argue.] Hence Q.

Case 2: Assume P_2. [Argue.] Hence Q.

⋮

Case k: Assume P_k. [Argue.] Hence Q.

Since $P_1 \vee P_2 \vee \cdots \vee P_k$ is exhaustive and Q holds in each case, Q holds. □

Remark 5.23 (The exhaustiveness obligation). A case proof is only valid if the cases together cover every possibility. Always verify this. Common exhaustive splits:

- n even or n odd
- $a > 0$, $a = 0$, or $a < 0$
- $a \mid b$ or $a \nmid b$
- $x \in A$ or $x \notin A$

Remark 5.24 (Cases can reduce to earlier cases). In a two-case proof, if Case 2 is symmetric to Case 1, it is acceptable to write: "Case 2 is symmetric." But only when the symmetry is genuinely complete — when the argument for Case 2 is obtained from Case 1 by renaming variables.

Example 5.25 (Parity of $n(n+1)$). Claim. For every $n \in \mathbb{Z}$, $n(n+1)$ is even.

Proof. Case 1: n is even. Write $n = 2k$. Then $n(n+1) = 2k(n+1)$, which is even.

Case 2: n is odd. Write $n+1 = 2k$. Then $n(n+1) = n \cdot 2k$, which is even.

In both cases $n(n+1)$ is even. Since every integer is even or odd, the proof is complete. □

Existence and Uniqueness. The statement $\exists!\, x\ P(x)$ has two parts: existence ($\exists x\ P(x)$) and uniqueness ($\exists!\, x$). They are proved separately.

Template: Existence and Uniqueness ($\exists!\, x\ P(x)$)

Step 1 Existence.
Define $x :=$ [explicit construction or witness].
Verify that $P(x)$ holds.

Step 2 Uniqueness (assume-two-and-compare).
Let x and y be arbitrary objects both satisfying P.
[Apply the definition of x with input y, and the definition of y with input x. The two applications collapse.]
Conclude $x = y$.

Step 3 Conclude.
By Steps 1 and 2, there exists exactly one object satisfying P. □

Remark 5.26 (Existence must come first). The uniqueness argument assumes that at least one object satisfying P exists; otherwise "x and y both satisfy P" has no content. Always prove existence before uniqueness.

Remark 5.27 (The key move in assume-two-and-compare). The proof that $x = y$ almost always uses x and y against each other: apply the definition of x as an identity (or inverse, or fixed point) with y as the input, and simultaneously apply the definition of y with x as the input. The two applications collapse to give $x = y$.

This move appears in: uniqueness of group identity, uniqueness of inverses, uniqueness of the zero vector, uniqueness of limits, uniqueness of the supremum.

Example 5.28 (Uniqueness of group identity). Proof. Suppose e and e' both satisfy the identity axiom in G. Then:

$$e = e \cdot e' = e'.$$

The first equality uses e' as an identity; the second uses e as an identity. Hence $e = e'$. □

Remark 5.29 (Uniqueness creates a tool). Every uniqueness theorem, once proved, activates the satisfy-and-cite tactic (Section 5.1.0.1). Proving uniqueness is not just an end in itself — it is the creation of a new proof tool that applies to all subsequent arguments in the same structure.

5.1.0.4 Mathematical Induction

Why Induction Works. Induction feels like a technique. It is actually a theorem. It follows from the axiomatic structure of $\mathbb{N}$.

Remark 5.30 (The Peano axiom behind induction). The Peano axioms characterize $\mathbb{N}$ as the smallest inductive set: a set containing 0 and closed under the successor function S. Formally, Axiom P5 states:

If $A \subseteq \mathbb{N}$ satisfies (i) $0 \in A$ and (ii) $n \in A \Rightarrow S(n) \in A$, then $A = \mathbb{N}$.

The induction principle is a direct restatement: if $P(0)$ holds and $P(n) \Rightarrow P(S(n))$, then defining $A = \{n \in \mathbb{N} : P(n)\}$ gives an inductive set, hence $A = \mathbb{N}$, hence $P(n)$ holds for all n. The full treatment is in the Axiom Systems chapter.

Remark 5.31 (Induction and well-ordering are equivalent). The induction principle and the well-ordering principle for $\mathbb{N}$ are logically equivalent. The well-ordering principle states: every nonempty subset of $\mathbb{N}$ has a least element.

This equivalence means every induction argument can be rephrased as a minimal counterexample argument, and vice versa. Both are available as proof tools; the choice between them is a matter of which formulation makes the argument cleaner. The equivalence is proved in Section 5.1.0.4.

Remark 5.32 (The domino picture is misleading). The common intuition "induction is like an infinite row of dominoes" is useful but imprecise. It suggests that induction requires each domino to knock over the next, which captures the inductive step but misses the point that the base case is what starts the chain. More importantly, it does not explain why the argument is valid — it is valid because $\mathbb{N}$ is defined to be the smallest inductive set, not because of any physical metaphor. The domino picture is a mnemonic, not a proof.

Theorem: Weak Induction

Let $P(n)$ be a statement depending on $n \in \mathbb{N}$. Suppose:

1. Base case. $P(0)$ is true.
2. Inductive step. For all $n \in \mathbb{N}$, $P(n) \Rightarrow P(n+1)$.

Then $P(n)$ is true for all $n \in \mathbb{N}$.

Template: Weak Induction Proof

Let $P(n)$ denote the statement: [write it out explicitly].

Base case ($n =$ [start]). [Verify P(start) directly.] Hence P(start) holds.

Inductive step. Let $n \in \mathbb{N}$ and assume $P(n)$ holds. [This is the inductive hypothesis — state it explicitly.] We show $P(n+1)$ holds.
[Work. Use $P(n)$ at least once.]
Hence $P(n+1)$ holds.

By induction, $P(n)$ holds for all $n \geq$ [start]. □

Remark 5.33 (The base case is not trivial). The base case is logically necessary, not a formality. A proof that omits it is invalid, regardless of how clean the inductive step is. However, the base case is usually the easiest part. Do not spend more than a line or two on it.

Remark 5.34 (State the inductive hypothesis explicitly). The single most important discipline in induction proofs is writing the inductive hypothesis as a full sentence before the inductive step

begins. "Assume $P(n)$ holds" is not sufficient. Write: "Assume $\sum_{k=1}^{n} k = \frac{n(n+1)}{2}$." This forces you to know what you are allowed to use, and it is the step most students skip.

Example 5.35 (Sum formula). Proposition. For all $n \geq 1$: $\displaystyle\sum_{k=1}^{n} k = \frac{n(n+1)}{2}$.

Proof. Let $P(n)$ denote the statement $\displaystyle\sum_{k=1}^{n} k = \frac{n(n+1)}{2}$.

Base case ($n = 1$). $\displaystyle\sum_{k=1}^{1} k = 1 = \frac{1 \cdot 2}{2}$. Hence $P(1)$ holds.

Inductive step. Let $n \geq 1$ and assume $\displaystyle\sum_{k=1}^{n} k = \frac{n(n+1)}{2}$. (Inductive hypothesis.)

We show $\displaystyle\sum_{k=1}^{n+1} k = \frac{(n+1)(n+2)}{2}$.

$$\begin{aligned}\sum_{k=1}^{n+1} k &= \left(\sum_{k=1}^{n} k\right) + (n+1) \\ &= \frac{n(n+1)}{2} + (n+1) && \text{(by inductive hypothesis)} \\ &= (n+1)\left(\frac{n}{2} + 1\right) = \frac{(n+1)(n+2)}{2}.\end{aligned}$$

Hence $P(n+1)$ holds.

By induction, $P(n)$ holds for all $n \geq 1$. □

Example 5.36 (Divisibility). Proposition. For all $n \geq 0$, $3 \mid (4^n - 1)$.

Proof. Let $P(n)$ denote "$3 \mid (4^n - 1)$".

Base case ($n = 0$). $4^0 - 1 = 0 = 3 \cdot 0$. Hence $3 \mid 0$, so $P(0)$ holds.

Inductive step. Assume $3 \mid (4^n - 1)$, i.e., $4^n - 1 = 3m$ for some $m \in \mathbb{Z}$.

We compute:

$$4^{n+1} - 1 = 4 \cdot 4^n - 1 = 4(4^n - 1) + 4 - 1 = 4 \cdot 3m + 3 = 3(4m + 1).$$

Hence $3 \mid (4^{n+1} - 1)$, so $P(n+1)$ holds.

By induction, $P(n)$ holds for all $n \geq 0$. □

Remark 5.37 (Anatomy of both examples). In both examples, the inductive step follows a single pattern: rewrite $P(n+1)$ to expose $P(n)$ inside it, then substitute the inductive hypothesis. This rewriting step is the core skill of weak induction. When it is not obvious how to expose $P(n)$, the statement $P(n)$ may need to be strengthened — see Section 5.1.0.4.

Choosing $P(n)$: The Hard Part. The statement you induct on is a choice, not given data. For textbook sum formulas, $P(n)$ is handed to you. In original proofs, you must invent it. Sometimes the obvious choice does not work and must be strengthened.

Remark 5.38 (The strengthening principle). If the inductive step fails with $P(n)$, the fix is usually to add more to $P(n)$, not to try a different proof strategy. This is counterintuitive: the stronger claim $P'(n)$ (which implies $P(n)$) is easier to induct on because the inductive hypothesis $P'(n)$ gives you more to work with.

Example 5.39 (Strengthening is necessary). Goal. Show that the Fibonacci sequence F_n defined by $F_1 = F_2 = 1$, $F_{n+2} = F_{n+1} + F_n$ satisfies $F_n < 2^n$ for all $n \geq 1$.

Naive attempt. Let $P(n)$: "$F_n < 2^n$".

Inductive step: assume $F_n < 2^n$; show $F_{n+1} < 2^{n+1}$.

We have $F_{n+1} = F_n + F_{n-1}$. We know $F_n < 2^n$ by hypothesis. But what is F_{n-1}? We have no bound on F_{n-1} from $P(n)$ alone. The step is stuck.

Fix: use strong induction. Assume $F_k < 2^k$ for all $k \leq n$. Then:

$$F_{n+1} = F_n + F_{n-1} < 2^n + 2^{n-1} < 2^n + 2^n = 2^{n+1}.$$

The step goes through because the stronger hypothesis provides $F_{n-1} < 2^{n-1}$. This is the canonical signal that strong induction is needed.

Example 5.40 (Strengthening the claim itself). Goal. Prove that for all $n \geq 1$, the sum $S_n = \sum_{k=1}^{n} \frac{1}{k(k+1)}$ equals $1 - \frac{1}{n+1}$.

What $P(n)$ to choose? The statement is handed to us: $P(n)$: $S_n = 1 - \frac{1}{n+1}$. Here the step is: $S_{n+1} = S_n + \frac{1}{(n+1)(n+2)}$. Substitute $S_n = 1 - \frac{1}{n+1}$ and simplify. The inductive step works because $P(n)$ gives the exact closed form needed.

Lesson. When the claim is a closed-form identity, $P(n)$ is the identity itself. The step is always: split off the last term, apply the inductive hypothesis, simplify algebraically.

Remark 5.41 (Diagnostic questions for choosing $P(n)$). When stuck choosing $P(n)$:

1. What exactly do I need to prove for $n = 0, 1, 2, 3$? Write it out. The pattern of $P(n)$ will appear.
2. Does the inductive step require knowing $P(n-1)$ as well as $P(n)$? If yes, switch to strong induction.
3. Does the inductive step produce a bound or equation that is slightly weaker than $P(n+1)$? If yes, strengthen $P(n)$.

Choosing $P(n)$ in Axiomatic and Recursive Settings. The examples above assume $P(n)$ is a formula handed to you or readable from a sequence. In axiomatic settings (e.g., Tao Chapter 2), the goal is

stated as a universal claim about a recursively defined operation, and $P(n)$ must be constructed from the logical form of the goal. Two rules govern this construction.

Remark 5.42 (Rule 1: $P(n)$ mirrors the logical form of the goal). Every universal claim has the form $\forall n,\ \Phi(n)$. Strip the universal quantifier. What remains is $P(n)$.

The shape of $P(n)$ is determined by Φ:

Goal form	Shape of $P(n)$
$\forall n,\ f(n) = g(n)$	Equality: $P(n) :\equiv f(n) = g(n)$
$\forall n,\ A(n) \Rightarrow B(n)$	Conditional: $P(n) :\equiv A(n) \Rightarrow B(n)$
$\forall n,\ A(n) \Leftrightarrow B(n)$	Biconditional: $P(n) :\equiv A(n) \Leftrightarrow B(n)$
$\forall n,\ \varphi(n)$	Simple property: $P(n) :\equiv \varphi(n)$

Do not add or remove hypotheses. Do not simplify. Transcribe the logical form exactly.

Remark 5.43 (Rule 2: Induct on the recursive variable). When the goal involves a recursively defined operation, identify which variable the definition recurses on. Induct on that variable and hold all others fixed.

To find the recursive variable: locate the definition of the operation and identify which variable is reduced in the recursive clause.

Definition	Recursive clause	Induct on
Addition (Tao A1–A2)	$(n{+}{+}) + m := (n + m){+}{+}$	n
Multiplication (Tao M1–M2)	$(n{+}{+}) \times m := (n \times m) + m$	n
Exponentiation (Tao)	$m^{n{+}{+}} := m^n \times m$	n

The non-recursive variable is held fixed throughout the induction and treated as an arbitrary element of $\mathbb{N}$.

Example 5.44 (Constructing $P(n)$ for a conditional goal). Goal. Prove Lemma 2.3.3: for all $n, m \in \mathbb{N}$, if $n > 0$ and $m > 0$ then $nm > 0$.

Step 1 Read the logical form. The goal is a universal conditional: $\forall n,\ (n > 0 \wedge m > 0) \Rightarrow nm > 0$. Strip the quantifier:

$$P(n) :\equiv \text{if } n > 0 \text{ and } m > 0 \text{ then } n \times m > 0.$$

Step 2 Identify the recursive variable. The operation is multiplication. M2 recurses on n: $(n{+}{+}) \times m := (n \times m) + m$. Induct on n, hold m fixed.

Step 3 Check the base case form. $P(0)$ states: if $0 > 0$ and $m > 0$ then $0 \times m > 0$. The hypothesis $0 > 0$ is false, so $P(0)$ is vacuously true. A conditional $P(n)$ will often produce a vacuous base case when the hypothesis excludes $n = 0$.

Remark 5.45 (Vacuous base cases are a signal, not a problem). When $P(n)$ is a conditional and its hypothesis fails at the base value, the base case is vacuously true by definition. This is expected and requires no further argument. It signals that the claim only has content for values of n where the hypothesis holds — positivity, nonzero-ness, a bound, or any other precondition that excludes the base case. The induction then establishes the claim for all n where the hypothesis is satisfied.

Remark 5.46 (Diagnostic questions for axiomatic settings). When constructing $P(n)$ in an axiomatic proof:

1. What is the logical form of the goal? Write it out with explicit quantifiers before choosing $P(n)$.
2. Which operation is involved? Find its recursive definition and identify the recursive variable.
3. Does the base case require $n = 0$ to satisfy a hypothesis? If the hypothesis fails at $n = 0$, expect a vacuous base case.
4. After writing $P(n{+}{+})$, can I apply the recursive definition to expose $P(n)$ or an earlier lemma? If not, the wrong variable may have been chosen.

Theorem: Strong Induction

Let $P(n)$ be a statement depending on $n \in \mathbb{N}$. Suppose that for every $n \in \mathbb{N}$:

$$\Big(\forall k < n,\ P(k)\Big) \Rightarrow P(n).$$

Then $P(n)$ holds for all $n \in \mathbb{N}$.

Remark 5.47 (The base case in strong induction). When $n = 0$, the hypothesis $\forall k < 0,\ P(k)$ is vacuously true (no $k < 0$ exists in $\mathbb{N}$). So the step for $n = 0$ reduces to showing $P(0)$ unconditionally — this is the base case, and it is implicit in the universal statement rather than listed separately. In practice, still write the base case explicitly.

Remark 5.48 (When to use strong induction). Use strong induction when the proof of $P(n + 1)$ requires knowing $P(k)$ for some $k < n$ that is not specifically $k = n$.

The diagnostic: in the inductive step, write $P(n + 1) = f(P(?), P(?), \ldots)$ and ask what values of ? you need. If the answer includes anything other than n, use strong induction.

Canonical situations:

- Recursive sequences: $a_{n+1} = a_n + a_{n-1}$ (need $n - 1$).
- Prime factorization: $n + 1 = ab$ requires $P(a)$ and $P(b)$ where $a, b < n + 1$ but are otherwise arbitrary.
- Any argument that breaks $n + 1$ into parts.

Template: Strong Induction Proof

Let $P(n)$ denote: [write it out].

Base case ($n =$ [start]). [Verify directly.]

Inductive step. Let $n >$ [start] and assume $P(k)$ holds for all [start] $\leq k < n$. (Strong inductive hypothesis.) We show $P(n)$ holds.
[Work. Use $P(k)$ for some specific $k < n$.]
Hence $P(n)$ holds.
By strong induction, $P(n)$ holds for all $n \geq$ [start]. □

Example 5.49 (Every integer $n \geq 2$ has a prime factor). Proof. Let $P(n)$: "n has a prime factor".

Base case ($n = 2$). 2 is prime, so 2 is a prime factor of itself.

Inductive step. Let $n > 2$ and assume $P(k)$ holds for all $2 \leq k < n$.

Case 1: n is prime. Then n is its own prime factor.

Case 2: n is composite. Then $n = ab$ for some $2 \leq a, b < n$. By the strong inductive hypothesis applied to a, a has a prime factor p. Then $p \mid a$ and $a \mid n$, so $p \mid n$.

In both cases n has a prime factor. By strong induction, $P(n)$ holds for all $n \geq 2$. □

Remark 5.50 (Why weak induction cannot do this directly). In the prime factor proof, the composite case splits n into $a < n$. We need $P(a)$, not $P(n-1)$. Weak induction only provides $P(n-1)$, which is useless if $a \neq n-1$. Strong induction provides $P(k)$ for all $k < n$, covering whatever value a takes.

Well-Ordering Principle

Every nonempty subset of $\mathbb{N}$ has a least element.

Remark 5.51 (Well-ordering as a proof engine). The well-ordering principle drives the minimal counterexample argument, which is the well-ordering dual of induction:

1. Suppose $P(n)$ fails for some n.
2. Let $S = \{n \in \mathbb{N} : \neg P(n)\}$. By assumption $S \neq \emptyset$.
3. By well-ordering, S has a least element m.
4. Derive a contradiction: show either $P(m)$ holds (contradicting $m \in S$) or some $k < m$ satisfies $k \in S$ (contradicting minimality of m).

This is structurally the same as induction but runs in the opposite direction: instead of climbing up from a base case, you descend to a minimal failure and derive a contradiction.

Proposition 5.52 (Equivalence of induction and well-ordering). The following are equivalent over the axioms of $\mathbb{N}$ without P5:

1. The induction principle (weak induction).

2. The well-ordering principle.

3. The strong induction principle.

Remark 5.53 (Proof sketch of equivalence). Induction ⇒ Well-ordering. Suppose $S \subseteq \mathbb{N}$ has no least element. Define $P(n)$: "$n \notin S$". Then $P(0)$ holds (else $0 \in S$ would be the least element). If $P(k)$ holds for all $k \leq n$ then $n+1 \notin S$ (else $n+1$ would be a least element). By induction, $P(n)$ holds for all n, so $S = \emptyset$.

Well-ordering ⇒ Induction. Suppose the base case and inductive step hold. Let $S = \{n : \neg P(n)\}$. If $S \neq \emptyset$, let m be its least element. Then $m \neq 0$ (base case), so $m \geq 1$, so $m - 1 < m$ satisfies $P(m-1)$ (minimality of m). But by the inductive step, $P(m-1) \Rightarrow P(m)$, contradicting $m \in S$. Hence $S = \emptyset$.

Template: Minimal Counterexample

Suppose for contradiction that $P(n)$ fails for some $n \in \mathbb{N}$.
Let $S = \{n \in \mathbb{N} : \neg P(n)\}$. By assumption $S \neq \emptyset$. By well-ordering, S has a least element m.
[Derive a contradiction using minimality of m.]
This contradicts [minimality of m / the base case]. Therefore $P(n)$ holds for all $n \in \mathbb{N}$. □

Example 5.54 (Minimal counterexample: divisibility). Claim. For all $n \geq 1$, $2^n > n$.

Proof (minimal counterexample). Suppose the claim fails. Let $S = \{n \geq 1 : 2^n \leq n\}$. By assumption $S \neq \emptyset$; let m be its least element.

Then $m \geq 2$ (since $2^1 = 2 > 1$), so $m - 1 \geq 1$ and $m - 1 < m$, hence $m - 1 \notin S$, so $2^{m-1} > m - 1$. Then:

$$2^m = 2 \cdot 2^{m-1} > 2(m-1) = 2m - 2 \geq m$$

(since $m \geq 2$), contradicting $m \in S$. Therefore $S = \emptyset$ and the claim holds. □

Structural Induction (Preview). Induction on $\mathbb{N}$ is a special case of a more general principle that applies to any recursively defined structure.

Remark 5.55 (The general principle). Whenever a set X is defined by:

1. a collection of base cases (atomic elements), and

2. a collection of constructor rules (ways of building larger elements from smaller ones),

structural induction is available: to prove $P(x)$ for all $x \in X$, prove P for all base cases, and prove that if P holds for the parts then P holds for the whole.

Example 5.56 (Induction on algebraic expressions). Define an algebraic term over a ring R by:

- Base: any $r \in R$ is a term.
- Constructor: if s, t are terms, so are $s + t$ and $s \cdot t$.

To prove a property P holds for all terms:

1. Prove $P(r)$ for all $r \in R$ (base cases).
2. Prove: if $P(s)$ and $P(t)$ hold, then $P(s + t)$ holds.
3. Prove: if $P(s)$ and $P(t)$ hold, then $P(s \cdot t)$ holds.

This is structural induction on the recursive definition of terms.

Remark 5.57 (Where structural induction appears in this project). • Axiom Systems: induction on $\mathbb{N}$ is proved from Peano P5, which is exactly the structural induction principle for $(\mathbb{N}, 0, S)$.

- Abstract Algebra: properties of polynomials over a ring are proved by induction on degree (structural induction on the recursive definition of polynomials).
- Computer Science: correctness of recursive algorithms is proved by structural induction on the input data structure (lists, trees, expressions).

The full development of structural induction appears in the Abstract Algebra chapter.

Common Induction Mistakes.

Remark 5.58 (Mistake 1: Missing base case). A proof that establishes the inductive step but not the base case proves nothing. The inductive step alone says: if $P(n)$ holds for some n, then $P(n + 1)$ holds. Without a base case, there is no n for which $P(n)$ is known to hold, so the chain never starts.

Classic fake theorem: "All horses are the same colour." Proof attempt: Base case $n = 1$: one horse is trivially the same colour as itself. Inductive step: [flawed argument]. The flaw is in the inductive step, but the lesson is that even a correct base case does not rescue a bad step. Scrutinize both parts independently.

Remark 5.59 (Mistake 2: Circular inductive step). The inductive step must prove $P(n + 1)$ using only $P(n)$ (or earlier cases in strong induction) and previously established facts. It is illegal to assume $P(n + 1)$ anywhere in the proof of $P(n + 1)$.

This error is most common when the student "derives" both sides of an equation and meets in the middle, inadvertently assuming the conclusion. Always work in one direction: from $P(n)$, derive $P(n + 1)$.

Remark 5.60 (Mistake 3: Wrong or missing inductive hypothesis). The most common error in practice is writing "assume $P(n)$" without stating what $P(n)$ says. This is a sign that $P(n)$ has not been written down as a full sentence. Write it out. If you cannot state $P(n)$ as a complete mathematical sentence, you do not yet know what you are inducting on.

Remark 5.61 (Mistake 4: Off-by-one base case). Induction proves $P(n)$ for all $n \geq n_0$, where n_0 is the base case. If the claim is only true for $n \geq 2$, the base case $n = 0$ may fail, and the proof should start at $n = 2$.

Check: what is the smallest n for which the claim is true? That is your base case. Do not default to $n = 0$ or $n = 1$ without verifying.

Remark 5.62 (Mistake 5: Using strong induction when weak suffices). Strong induction is not more powerful than weak induction in the sense that they prove the same class of statements. But a proof that uses the full strength of the strong hypothesis when only $P(n)$ is needed is unnecessarily complicated. Use weak induction unless the step genuinely requires knowing $P(k)$ for some $k < n$ other than $k = n - 1$.

Remark 5.63 (Mistake 6: Forgetting vacuity in strong induction). In strong induction, the case $n = 0$ has a vacuously true hypothesis $\forall k < 0,\ P(k)$, since no natural number is less than 0. This means $P(0)$ must be proved directly, not derived from the inductive hypothesis. Many students write "by the inductive hypothesis applied to $n - 1$" when $n = 0$ has no predecessor. The base case is always proved from scratch.

5.1.0.5 Algebraic Tactics

The DU/TA/AM Framework. Every line of an algebraic proof is justified by exactly one of three sources. This framework formalizes the tagging system used throughout these notes.

Definition 5.64 (Proof line tags). Each step in a three-column proof carries one of three tags:

- DU (Definition / Unpack). The step applies a definition or unpacks a named concept into its explicit conditions. Example: expanding "e is an identity" into "$ea = ae = a$ for all $a \in G$".
- TA (Theorem / Axiom). The step cites a previously proved proposition or a structural axiom of the system (group axiom, ring axiom, field axiom, or Peano axiom). Example: citing G1 (associativity) to regroup $(ab)c = a(bc)$.
- AM (Algebraic Manipulation). The step performs a calculation whose justification is itself a combination of DU and TA moves, compressed for readability. Example: simplifying $a \cdot e = a$ in one step after the identity axiom has already been cited.

Remark 5.65 (Why tag every line?). The DU/TA/AM system enforces the most important discipline in proof writing: every step must be justified by something. A line with no tag is a line whose justification has not been identified. When a proof stalls, examining the tags reveals which definition has not been unpacked or which axiom has not been cited.

At the learning stage, every line should carry an explicit tag. As fluency increases, AM steps can absorb multiple tagged moves. But the underlying DU and TA structure is always present even when not written.

Remark 5.66 (The hierarchy of tags). DU and TA are the atomic justifications. AM is a shorthand for a sequence of DU and TA steps that have been verified and compressed. A fully formal proof contains only DU and TA steps. The three-column format used in these notes puts the statement in the middle column and the tag in the left column, forcing explicit justification of every line.

Catalog of Algebraic Tactics. The following tactics are the move-level tools available once a proof is in progress and the architecture has been chosen. Each entry states the tactic, its precondition (what must already be in place), and its effect (what it achieves).

Tactic	Precondition	Effect
Multiply by inverse	$a \neq 0$ in a field; or $a \in G$ a group element	Multiply both sides by a^{-1}; use associativity to collapse $a^{-1}a = e$; isolates the target variable.
Cancellation	$ac = bc$ or $ca = cb$ in a group or integral domain	Conclude $a = b$. In a group: multiply by c^{-1}. In an integral domain: cite no-zero-divisors.
Distributivity bridge	Multiplication by 0, -1, or a negative element	Rewrite the product as a sum using distributivity, then use additive group structure to collapse. The bridge crosses from $\times$ to $+$.
Add zero	Need to introduce a term without changing the expression	Write $a = a + 0 = a + (b + (-b))$; regroup to expose needed structure.
Satisfy-and-cite	A uniqueness theorem for some object b is in scope	Show a satisfies b's defining property; cite uniqueness; conclude $a = b$. Avoids assume-two-and-compare entirely.
Double negation	$-(-a)$ appears or is the goal	$-a$ is the unique element x with $a + x = 0$; $-(-a)$ satisfies this with $x = a$; by uniqueness, $-(-a) = a$.
Absorb into axiom	A sub-expression matches the left-hand side of an axiom	Rewrite using the axiom. Tag as TA.

Remark 5.67 (Distributivity bridge in detail). The distributivity bridge is the central tactic for ring and field zero/negative proofs. The pattern is always:

1. Introduce $a \cdot 0$ or $a \cdot (-b)$.
2. Write $0 = x + (-x)$ or $-b = b + (-b) + (-b)$ — use the additive inverse definition to expand.
3. Apply distributivity: $a(x + y) = ax + ay$.
4. Use uniqueness of additive inverse to identify the result.

Every instance of "$a \cdot 0 = 0$" and "$a(-b) = -(ab)$" uses this bridge. Once seen once, it is immediately recognizable.

Remark 5.68 (Cancellation: two different theorems). "Cancellation" names two distinct results that must not be conflated:

- Group cancellation: in any group, $ac = bc$ implies $a = b$. Proved by multiplying by c^{-1}, which exists by the group inverse axiom.
- Domain cancellation: in an integral domain, $ac = bc$ and $c \neq 0$ imply $a = b$. Proved by rewriting as $c(a - b) = 0$ and citing no zero divisors.

The preconditions are different. In a group, c need not be specified as nonzero (inverses always exist). In an integral domain, $c \neq 0$ is essential.

5.1.0.6 Proof Strategies Reference

Lookup Table: Stuck Situations and Strategies.

Remark 5.69 (How to use this table). When progress on a proof stalls, locate the column that describes your situation. The middle column names the strategy. The right column gives the specific first move and a cross-reference to a worked example.

Stuck because...	Strategy	First move / See
I need to prove P but don't know where to start	Two-question check	Write the logical form of P. Consult the statement map. §5.1.0.1
I need to show two things are equal and one is uniquely defined	Satisfy-and-cite	Show the first satisfies the definition of the second. Cite the uniqueness theorem. §5.1.0.1
I need to show something is the only object of its kind	Assume-two-and-compare	"Let x, y both satisfy the definition. Show $x = y$." §5.1.0.3
I have a conditional $P \Rightarrow Q$ and P is hard to use	Contrapositive	"Assume $\neg Q$." Derive $\neg P$. §5.1.0.3
I have a conditional and the conclusion is a negation or impossible	Contradiction	"Assume P and $\neg Q$." Derive any contradiction. §5.1.0.3
The claim splits naturally by cases	Case analysis	List all cases explicitly. Verify they are exhaustive. §5.1.0.3
The claim is about all $n \in \mathbb{N}$	Induction	Write $P(n)$ as a full sentence. Prove base case. §5.1.0.4
The inductive step needs $P(k)$ for some $k < n$ (not just $n-1$)	Strong induction	Replace "assume $P(n)$" with "assume $P(k)$ for all $k < n$". §5.1.0.4
The claim is about all $n \geq n_0$ and feels easier to negate	Minimal counterexample	"Suppose $S = \{n : \neg P(n)\} \neq \emptyset$. Let $m = \min S$." §5.1.0.4
I need to prove $a \cdot 0 = 0$ or $a(-b) = -(ab)$	Distributivity bridge	Write $0 = x + (-x)$; apply distributivity; use uniqueness of additive inverse. §5.1.0.5
I have $ac = bc$ and need $a = b$	Cancellation	Group: multiply by c^{-1}. Domain: write $(a-b)c = 0$; cite no zero divisors. §5.1.0.5
I need to set two sets equal	Double inclusion	Prove $A \subseteq B$: chase an arbitrary $x \in A$. Then prove $B \subseteq A$. §5.1.0.1
A definition has not been unpacked	DU move	Identify the named concept and write out its definition. Tag the line DU. §5.1.0.5
A step uses an axiom or theorem by name	TA move	Cite the axiom or theorem explicitly by label. Tag the line TA. §5.1.0.5

Remark 5.70 (Cross-references to worked examples). The following table maps each strategy to a worked example in these notes where it is the primary tool.

Strategy	Primary worked examples
Assume-two-and-compare	Prop. 17.16, Prop. 17.18
Satisfy-and-cite	Prop. 17.22, Prop. 17.38
Distributivity bridge	Prop. 17.36, Prop. 17.38
Cancellation (group)	Prop. 17.20
Cancellation (domain)	Prop. 17.44
Multiply by inverse + no zero divisors	Prop. 17.52
Weak induction	Sum formula, divisibility examples (§5.1.0.4)
Strong induction	Prime factor theorem (§5.1.0.4)
Minimal counterexample	$2^n > n$ example (§5.1.0.4)
Contrapositive	n^2 even $\Rightarrow$ n even (§5.1.0.3)
Contradiction	$\sqrt{2} \notin \mathbb{Q}$ (§5.1.0.3)

Chapter 6

Model Theory

Where You Are in the Journey

Propositional Logic $\rightarrow$ Predicate Calculus $\rightarrow$ Sets & Functions $\rightarrow$ Proof Techniques $\rightarrow$ Axiom Systems $\rightarrow$ Model Theory $\rightarrow \cdots$

How we got here. Predicate logic gave us the syntax and proof theory of first-order languages. Model theory studies the relationship between formal theories and their interpretations (models): which structures satisfy a given set of axioms?

What this chapter will build. The compactness theorem, the Löwenheim–Skolem theorems, and the study of complete and categorical theories.

Where this leads. Model theory provides the metamathematical foundations for understanding how algebraic structures (groups, fields, rings) relate to their axioms.

Status: Planned

Coming Soon

Notes, proofs, and exercises will appear here in a future revision.

6.1 Notes

To be populated.

6.2 Proofs

To be populated.

6.3 Capstone

To be populated.

Chapter 7

Type Theory

Where You Are in the Journey

Propositional Logic $\rightarrow$ Predicate Calculus $\rightarrow$ Sets & Functions $\rightarrow$ Proof Techniques $\rightarrow$ Model Theory $\rightarrow$ Type Theory $\rightarrow \cdots$

How we got here. Set theory provides one foundation for mathematics. Type theory provides an alternative, assigning every expression a type to avoid paradoxes and to make proof-checking mechanical.

What this chapter will build. Simply typed lambda calculus, the Curry–Howard correspondence (proofs as programs), and dependent type theory as a foundation for mechanised proof assistants (Lean, Coq).

Where this leads. Modern proof assistants (Lean, Coq, Agda) are built on type theory. Understanding type theory illuminates what it means to formalise mathematics completely.

Status: Planned

Coming Soon

Notes, proofs, and exercises will appear here in a future revision.

7.1 Notes

To be populated.

7.2 Proofs

To be populated.

7.3 Capstone

To be populated.

Part II

Foundations of Formal Number Systems

Chapter 8

Natural Numbers ($\mathbb{N}$)

8.1 Notes

Where You Are in the Journey

Propositional Logic $\rightarrow$ Sets & Functions $\rightarrow$ Proof Techniques $\rightarrow$ Axiom Systems $\rightarrow$ Natural Numbers ($\mathbb{N}$) $\rightarrow$ Integers ($\mathbb{Z}$) $\rightarrow$ Rationals ($\mathbb{Q}$) $\rightarrow$ Reals ($\mathbb{R}$) $\rightarrow \cdots$

How we got here. The Peano axioms in Volume I established that $\mathbb{N}$ exists: five axioms pin down the structure of zero and the successor function, and the recursion theorem guarantees that recursive definitions actually produce well-defined functions. This chapter takes those logical primitives and builds arithmetic — the concrete mathematics of counting, addition, multiplication, and order.

What this chapter builds. We follow Tao's Analysis I, Chapter 2 (ğğ2.2–2.3). Addition is defined by recursion on the left argument and every algebraic property (commutativity, associativity, cancellation) is proved from the definition alone. Order on $\mathbb{N}$ is defined additively and the full trichotomy is established. Multiplication is then defined recursively and its properties derived. Strong induction and the well-ordering principle round out the toolkit.

Where this leads. The integers $\mathbb{Z}$ are constructed by adjoining additive inverses to $\mathbb{N}$. The cancellation law for natural-number addition — proved here — carries the entire integer construction forward. Every algebraic structure built later inherits the patterns established in this chapter.

Structural Roadmap

The development follows Tao's construction of $\mathbb{N}$ arithmetic from the Peano axioms. The architecture at every step is:

$$\text{Definition (recursive)} \longrightarrow \text{Well-definedness} \longrightarrow \text{Lemmas} \longrightarrow \text{Propositions} \longrightarrow \text{Corollaries}$$

Nothing beyond the Peano axioms and what has already been proved is assumed. The available toolkit grows with each result.

The global progression is:

1. Addition (Tao ğ2.2, Def. 2.2.1): defined by recursion on the left argument. Key results: zero on the right, successor on the right, commutativity, associativity, cancellation.
2. Order on $\mathbb{N}$ (Tao ğ2.2, Def. 2.2.11): $n \leq m$ iff $\exists a \in \mathbb{N},\ m = n + a$. Key results: transitivity, antisymmetry, trichotomy, monotonicity under addition.
3. Strong induction (Tao P2.2.14): the principle of strong induction and backwards induction, derived from the Peano induction axiom.
4. Multiplication (Tao ğ2.3, Def. 2.3.1): defined by recursion. Key results: commutativity, associativity, distributivity, no zero divisors.
5. Exponentiation (Mendelson ğ2.6): defined recursively. Basic exponent laws proved.

Remark 8.1 (What stays in Axiom Systems). The Peano axioms (P1–P5), numeral definitions ($1 := 0{+}{+}$, etc.), and the recursion theorem live in Volume I, Axiom Systems. That chapter is the logical foundation; this chapter is the arithmetic development built on top of it.

Remark 8.2 (Primary sources). Tao, Analysis I, Chapter 2 (primary); Mendelson, Number Systems and the Foundations of Analysis, Chapter 2 (exponentiation and additional structural perspective).

8.1.0.1 Addition

Definition 2.2.1 (Addition)

Let $m \in \mathbb{N}$ be fixed. Addition is defined recursively on the left argument:

$$\text{A1:}\quad 0 + m \ :=\ m$$

$$\text{A2:}\quad (n{+}{+}) + m \ :=\ (n + m){+}{+}$$

Remark 8.3 (The definition is asymmetric). A1 and A2 recurse on the left argument only. $3 + 5$ means: increment 5 three times. $5 + 3$ means: increment 3 five times. That both give 8 is a theorem (commutativity), not a definition. Never assume $a + b = b + a$ until it is proved.

Remark 8.4 (Consequence for proof strategy). When proving an identity involving addition by induction, the induction variable must be in the left position of the outermost addition on both sides, so that A2 fires cleanly.

Addition Toolkit Quick Reference

Label	Statement	Proof method
A1	$0 + m = m$	Definition
A2	$(n{+}{+}) + m = (n + m){+}{+}$	Definition
L2.2.2	$n + 0 = n$	Induction on n
L2.2.3	$n + (m{+}{+}) = (n + m){+}{+}$	Induction on n
P2.2.4	$n + m = m + n$	Induction on n, uses L2.2.2, L2.2.3
P2.2.5	$(a + b) + c = a + (b + c)$	Induction on a, uses A1, A2
P2.2.6	$a + b = a + c \Rightarrow b = c$	Induction on a, uses P4
Def 2.2.7	n positive $\iff n \neq 0$	Definition
P2.2.8	a positive, $b \in \mathbb{N} \Rightarrow a + b$ positive	Induction on b, uses P3
C2.2.9	$a + b = 0 \Rightarrow a = 0$ and $b = 0$	Contradiction + P2.2.8
L2.2.10	a positive $\Rightarrow \exists!\ b,\ b{+}{+} = a$	Induction (existence); P4 (uniqueness)

Lemma 8.5 (Tao 2.2.2). For all $n \in \mathbb{N}$, $\quad n + 0 = n$.

Remark 8.6. A1 gives $0 + m = m$ (zero on the left). This lemma gives $n + 0 = n$ (zero on the right). The asymmetry of the definition means these are different facts. Commutativity cannot be invoked to pass from one to the other because commutativity has not been proved yet.

Lemma 8.7 (Tao 2.2.3). For all $n, m \in \mathbb{N}$, $\quad n + (m{+}{+}) = (n + m){+}{+}$.

Remark 8.8. A2 gives $(n{+}{+})+m = (n+m){+}{+}$ (successor on the left). This lemma gives $n+(m{+}{+}) = (n + m){+}{+}$ (successor on the right). Again, a different fact requiring its own proof.

Corollary 8.9 (Successor as Addition of One). For all $n \in \mathbb{N}$, $\quad n{+}{+} = n + 1$.

Proposition 8.10 (Tao 2.2.4 Commutativity). For all $n, m \in \mathbb{N}$, $\quad n + m = m + n$.

Proposition 8.11 (Tao 2.2.5 Associativity). For all $a, b, c \in \mathbb{N}$, $\quad (a + b) + c = a + (b + c)$.

Remark 8.12 (Induction variable). Induct on a (the left argument of the outermost addition on both sides). A2 fires on both sides simultaneously only when a is the variable.

Proposition 8.13 (Tao 2.2.6 Cancellation). Let $a, b, c \in \mathbb{N}$ with $a + b = a + c$. Then $b = c$.

Remark 8.14 (Virtual subtraction). Subtraction does not exist in $\mathbb{N}$ yet. This proposition provides a "virtual subtraction": it lets us cancel a from both sides of an equation without ever invoking $-$. It will later be used to define subtraction on $\mathbb{Z}$.

Definition 8.15 (Tao 2.2.7 Positive natural numbers). A natural number n is positive if and only if $n \neq 0$.

Proposition 8.16 (Tao 2.2.8). If a is positive and $b \in \mathbb{N}$, then $a + b$ is positive (and hence $b + a$ is positive by commutativity).

Corollary 8.17 (Tao 2.2.9). If $a, b \in \mathbb{N}$ and $a + b = 0$, then $a = 0$ and $b = 0$.

Lemma 8.18 (Tao 2.2.10 Predecessor). Let a be a positive natural number. Then there exists exactly one $b \in \mathbb{N}$ such that $b{+}{+} = a$.

Remark 8.19 (Two-part proof). Existence: induction on a. Base case is vacuous ($a = 0$ excluded by positivity). Inductive step: witness $b = a$ satisfies $b{+}{+} = a{+}{+}$ directly; inductive hypothesis is not needed.

Uniqueness: if $b{+}{+} = a$ and $c{+}{+} = a$ then $b{+}{+} = c{+}{+}$, so $b = c$ by P4. One axiom citation closes uniqueness entirely.

Definition 8.20 (Tao 2.2.11 Order on $\mathbb{N}$). Let $n, m \in \mathbb{N}$.

$$n \geq m \quad (\text{equivalently } m \leq n) \quad \iff \quad \exists\, a \in \mathbb{N},\ n = m + a.$$

$$n > m \quad (\text{equivalently } m < n) \quad \iff \quad n \geq m \text{ and } n \neq m.$$

Remark 8.21 (Unpacking $\geq$). $n \geq m$ is an existence claim: you must exhibit a witness a. $n > m$ adds a negation: the witness a must be nonzero, i.e., positive. Always unpack the definition before writing any proof involving order.

Order Toolkit Quick Reference

Label	Statement	Key tool
P2.2.12(a)	$a \geq a$ (reflexive)	Witness $a = a + 0$ by L2.2.2
P2.2.12(b)	$a \geq b,\ b \geq c \Rightarrow a \geq c$ (transitive)	Combine witnesses; P2.2.5
P2.2.12(c)	$a \geq b,\ b \geq a \Rightarrow a = b$ (anti-symmetric)	P2.2.6, C2.2.9
P2.2.12(d)	$a \geq b \iff a + c \geq b + c$ (add preserves order)	P2.2.5, P2.2.4, P2.2.6
P2.2.12(e)	$a < b \iff a{+}{+} \leq b$	L2.2.10, L2.2.3, A2
P2.2.12(f)	$a < b \iff b = a + d$ for some positive d	Def 2.2.7, C2.2.9
P2.2.13	Trichotomy: exactly one of $a < b,\ a = b,\ a > b$	Induction + P2.2.12

Proposition 8.22 (Tao 2.2.12 Basic properties of order). Let $a, b, c \in \mathbb{N}$.

(a) Reflexive. $a \geq a$.

(b) Transitive. If $a \geq b$ and $b \geq c$, then $a \geq c$.

(c) Anti-symmetric. If $a \geq b$ and $b \geq a$, then $a = b$.

(d) Addition preserves order. $a \geq b \iff a + c \geq b + c$.

(e) $a < b \iff a{+}{+} \leq b$.

(f) $a < b \iff b = a + d$ for some positive d.

Remark 8.23 (Strategy for each part). • (a): existence proof. Witness is immediate from L2.2.2.

- (b): combine two witnesses using P2.2.5 to get a new witness.
- (c): combine witnesses to get $a = a + (n + m)$, cancel via P2.2.6, then apply C2.2.9 to conclude $n = m = 0$.
- (d): biconditional — two directions, each an existence argument plus P2.2.5/P2.2.4 for regrouping, P2.2.6 for the backward direction.

- (e): forward direction uses L2.2.10 to extract a predecessor, then L2.2.3 and A2 to rewrite. Backward direction: exhibit witness and invoke C2.2.9 and Def 2.2.7.
- (f): unpacks $<$ directly using Def 2.2.7.

Proposition 8.24 (Tao 2.2.13 Trichotomy). For all $a, b \in \mathbb{N}$, exactly one of the following holds:

$$a < b, \qquad a = b, \qquad a > b.$$

Remark 8.25 (Two obligations). Must show: (i) at most one holds (mutual exclusion), and (ii) at least one holds (exhaustion by induction on a).

Proposition 2.2.14 (Strong Principle of Induction)

Let $m_0 \in \mathbb{N}$, and let $P(m)$ be a property of natural numbers. Suppose that for each $m \geq m_0$:

$$\big[\forall m_0 \leq m' < m,\ P(m') \text{ is true}\big] \implies P(m) \text{ is true.}$$

Then $P(m)$ is true for all $m \geq m_0$.

Remark 8.26 (What is different from ordinary induction). In ordinary induction (P5), the inductive step assumes $P(n)$ and proves $P(n{+}{+})$ — you get to use exactly one prior case. In strong induction, the inductive step assumes $P(m')$ for all $m' < m$ and proves $P(m)$ — you get to use every prior case at once. This is essential when the proof of $P(m)$ requires not just $P(m-1)$ but some earlier case $P(k)$ for $k \ll m$, as frequently occurs in number theory, recursion, and later real analysis arguments.

Remark 8.27 (The Q-trick: how the proof works). The standard proof reduces strong induction to ordinary induction by a single definitional move. Define:

$$Q(n) \ :\Longleftrightarrow\ \forall m,\ m_0 \leq m < n \Rightarrow P(m).$$

Then $Q(n)$ says: "P holds for everything below n (from m_0 onward)." Crucially:

- $Q(n)$ is vacuously true when $n < m_0$ or $n = m_0$, because the condition $m_0 \leq m < n$ has no solutions.
- The hypothesis of P2.2.14 says exactly: if $Q(m)$ holds then $P(m)$ holds (for $m \geq m_0$).
- This means $Q(n) \Rightarrow Q(n{+}{+})$ by ordinary induction, since $Q(n{+}{+})$ extends $Q(n)$ by one more case ($m = n$).
- By P5, $Q(n)$ holds for all $n \in \mathbb{N}$.
- Therefore $P(m)$ holds for all $m \geq m_0$.

The Q-trick is worth memorizing: wrap the target property in a universal quantifier over all prior values.

Remark 8.28 (Base case is free). Notice $P(m_0)$ follows immediately: the hypothesis of the implication ($P(m')$ true for all $m_0 \leq m' < m_0$) is vacuously satisfied, so $P(m_0)$ holds. You do not need a separate base case argument — it is absorbed into the inductive step at $m = m_0$.

Remark 8.29 (Common values of m_0). In practice $m_0 = 0$ or $m_0 = 1$. The general m_0 matters for statements like "for all $n \geq 2$, n has a prime factorization" where the base case $n = 0$ or $n = 1$ would be vacuous or trivial.

Exercise 2.2.6 (Backwards Induction)

Let $n \in \mathbb{N}$, and let $P(m)$ be a property of natural numbers such that:

(i) $P(n)$ is true.

(ii) Whenever $P(m{+}{+})$ is true, $P(m)$ is also true.

Then $P(m)$ is true for all $m \leq n$.

Remark 8.30 (Direction of travel). Ordinary induction climbs up: $P(0) \to P(1) \to P(2) \to \cdots$ Backwards induction descends down: $P(n) \to P(n-1) \to \cdots \to P(0)$. The descent is finite (stops at 0), which is why it works. An infinite descent in the other direction would not terminate.

Remark 8.31 (Proof strategy: induct on n, not on m). The clever move in the proof is to induct on the ceiling n, not the property variable m. Fix P satisfying condition (ii). Define:

$$R(n) \iff \big[\, P(n) \Rightarrow \forall m \leq n,\ P(m) \,\big].$$

Show $R(n)$ holds for all n by ordinary induction. Then whenever $P(n)$ is given, $R(n)$ delivers $P(m)$ for all $m \leq n$.

Remark 8.32 (When backwards induction appears). Backwards induction arises in finite combinatorics and optimization: to show a property holds at every stage of a finite process, establish it at the end and propagate backward. It also appears in game theory (backward induction in finite games) and in proofs that finite decreasing sequences must terminate.

8.1.0.2 Multiplication

Definition 2.3.1 (Multiplication)

Let $m \in \mathbb{N}$ be fixed. Multiplication is defined recursively on the left argument:

$$\text{M1:} \quad 0 \times m := 0$$

$$\text{M2:} \quad (n{+}{+}) \times m := (n \times m) + m$$

Remark 8.33 (Multiplication as iterated addition). M2 says: multiplying $n{+}{+}$ by m is the same as multiplying n by m and then adding one more copy of m. Thus $n \times m$ adds m to itself n times:

$$0 \times m = 0, \quad 1 \times m = 0 + m, \quad 2 \times m = 0 + m + m, \quad \ldots$$

Remark 8.34 (Same asymmetry as addition). M1 and M2 recurse on the left argument only. Commutativity $n \times m = m \times n$ is again a theorem, not a definition.

Definition 8.35 (Exponentiation Tao 2.3.11). Let $m \in \mathbb{N}$. Exponentiation is defined recursively:

$$m^0 := 1, \qquad m^{n{+}{+}} := m^n \times m.$$

In particular, $0^0 := 1$.

Multiplication Toolkit Quick Reference

Label	Statement	Key tool
M1	$0 \times m = 0$	Definition
M2	$(n{+}{+}) \times m = (n \times m) + m$	Definition
L2.3.2	$n \times m = m \times n$	Induction
L2.3.3	$nm = 0 \iff n = 0$ or $m = 0$	P2.2.8 (contrapositive)
P2.3.4	$a(b+c) = ab + ac$ (distributive law)	Induction on c
P2.3.5	$(a \times b) \times c = a \times (b \times c)$	Induction
P2.3.6	$a < b,\ c > 0 \Rightarrow ac < bc$	P2.2.12(f), P2.3.4
C2.3.7	$ac = bc,\ c \neq 0 \Rightarrow a = b$	P2.2.13 + P2.3.6
P2.3.9	Euclidean algorithm: $\exists m, r,\ n = mq + r,\ 0 \leq r < q$	Induction

Lemma 8.36 (Tao 2.3.2 Commutativity). For all $n, m \in \mathbb{N}$, $n \times m = m \times n$.

Lemma 8.37 (Tao 2.3.3 No zero divisors). For all $n, m \in \mathbb{N}$, $n \times m = 0$ if and only if $n = 0$ or $m = 0$.

In particular, if n and m are both positive, then nm is positive.

Proposition 8.38 (Tao 2.3.4 Distributive law). For all $a, b, c \in \mathbb{N}$:

$$a(b+c) = ab + ac \qquad \text{and} \qquad (b+c)a = ba + ca.$$

Remark 8.39 (Proof strategy). The second identity follows from the first by commutativity (L2.3.2). For the first, induct on c (the right summand), keeping a and b fixed. M2 fires on the left side; the inductive hypothesis closes the step.

Proposition 8.40 (Tao 2.3.5 Associativity). For all $a, b, c \in \mathbb{N}$, $(a \times b) \times c = a \times (b \times c)$.

Proposition 8.41 (Tao 2.3.6 Multiplication preserves order). If $a < b$ and c is positive, then $ac < bc$.

Corollary 8.42 (Tao 2.3.7 Cancellation for multiplication). Let $a, b, c \in \mathbb{N}$ with $ac = bc$ and $c \neq 0$. Then $a = b$.

Remark 8.43 (Virtual division). This is the multiplicative analogue of the additive cancellation law (P2.2.6). It provides "virtual division" before division is defined, and will be used to define $\mathbb{Q}$ later.

Proposition 8.44 (Tao 2.3.9 Euclidean algorithm). Let $n \in \mathbb{N}$ and let q be positive. Then there exist $m, r \in \mathbb{N}$ such that:

$$0 \leq r < q \qquad \text{and} \qquad n = mq + r.$$

Remark 8.45. This is the seed of number theory: every natural number has a unique quotient and remainder when divided by a positive number.

Theorem 6.3 Basic Properties of Exponentiation (?Ch. 2, §2.6, Theorem 6.3)

For all $x, y, z \in \mathbb{N}$:

(a) $x^{y+z} = x^y \times x^z$ (b) $(x^y)^z = x^{y \times z}$ (c) $(x \times y)^z = x^z \times y^z$

Remark 8.46 (Recall the definition). Exponentiation on $\mathbb{N}$ is defined recursively (Tao Def. 2.3.11, Mendelson Thm. 6.1):

$$x^0 := 1, \qquad x^{n++} := x^n \times x.$$

So x^{n++} appends one more factor of x on the right. All three parts of Thm 6.3 are proved by induction on z.

Remark 8.47 (Why induct on z?). z is the variable that appears in the exponent on the left-hand side in a form that interacts directly with the recursive definition. Specifically:

- In (a), x^{y+z} recurses on z via $y + z{+}{+} = (y + z){+}{+}$.
- In (b), $(x^y)^z$ recurses on z directly.
- In (c), $(x \times y)^z$ recurses on z directly.

Fix all other variables and let z be the induction variable.

Remark 8.48 (Part (a): $x^{y+z} = x^y \times x^z$ the exponent addition law). This is the multiplicative analogue of the distributive law. It says: raising x to a sum of exponents is the same as multiplying the separate powers.

Proof sketch (induct on z).

- Base case $z = 0$: $x^{y+0} = x^y = x^y \times 1 = x^y \times x^0$. Uses L2.2.2 ($y + 0 = y$) and the definition $x^0 = 1$.
- Inductive step: assume $x^{y+z} = x^y \times x^z$. Then:

$$x^{y+z{+}{+}} = x^{(y+z){+}{+}} = x^{y+z} \times x \overset{\text{IH}}{=} (x^y \times x^z) \times x = x^y \times (x^z \times x) = x^y \times x^{z{+}{+}}.$$

  Uses L2.2.3 ($y + z{+}{+} = (y + z){+}{+}$), the definition of $x^{(y+z){+}{+}}$, the inductive hypothesis, and associativity (P2.3.5).

Remark 8.49 (Part (b): $(x^y)^z = x^{y \times z}$ the power of a power law). Exponentiating a power collapses to multiplication of exponents.

Proof sketch (induct on z).

- Base case $z = 0$: $(x^y)^0 = 1 = x^0 = x^{y \times 0}$. Uses M1 ($y \times 0 = 0$) and the definition $x^0 = 1$.
- Inductive step: assume $(x^y)^z = x^{y \times z}$. Then:

$$(x^y)^{z{+}{+}} = (x^y)^z \times x^y \overset{\text{IH}}{=} x^{y \times z} \times x^y = x^{y \times z + y} = x^{y \times z{+}{+}}.$$

  Uses the definition of $(x^y)^{z{+}{+}}$, the inductive hypothesis, part (a) to combine $x^{y \times z} \times x^y = x^{y \times z + y}$, and M2 ($(y \times z{+}{+}) = y \times z + y$).

Remark 8.50 (Part (c): $(x \times y)^z = x^z \times y^z$ the product power law). A product raised to a power distributes over the factors.

Proof sketch (induct on z).

- Base case $z = 0$: $(x \times y)^0 = 1 = 1 \times 1 = x^0 \times y^0$.
- Inductive step: assume $(x \times y)^z = x^z \times y^z$. Then:

$$(x \times y)^{z++} = (x \times y)^z \times (x \times y) \overset{\text{IH}}{=} (x^z \times y^z) \times (x \times y).$$

  Rearrange using commutativity and associativity to get $(x^z \times x) \times (y^z \times y) = x^{z++} \times y^{z++}$. The rearrangement requires four applications of P2.3.5 (associativity) and one of L2.3.2 (commutativity).

Remark 8.51 (Also useful: Mendelson Lemma 6.2 ($1^x = 1$)). Proved by induction on x: base $1^0 = 1$; step $1^{x++} = 1^x \times 1 = 1 \times 1 = 1$. This is needed as a lemma in the uniqueness argument for the definition of exponentiation and is useful in its own right.

8.2 Proofs

Chapter 9

Integers ($\mathbb{Z}$)

Where You Are in the Journey

Propositional Logic → Predicate Calculus → Sets & Functions → Proof Techniques → Axiom Systems & Natural Numbers → The Integers → Rationals & Ordered Fields → Real Line Foundations → ⋯

How we got here. The natural numbers gave us addition, multiplication, and a well-founded order, but subtraction is not available: for $a < b$ in $\mathbb{N}$, the expression $a - b$ is meaningless. Every subsequent number system depends on fixing this, so we pause here to do it rigorously.

What this chapter builds. We follow two complementary sources. Tao's Analysis I, Chapter 4 constructs $\mathbb{Z}$ from $\mathbb{N} \times \mathbb{N}$ via formal pairs and proves all arithmetic properties directly. Mendelson's Number Systems and the Foundations of Analysis, Chapters 3–4 performs the same construction but frames the result in the language of abstract algebra: rings, integral domains, and ordered integral domains. Together they give both the computational fluency and the structural vocabulary that real analysis and abstract algebra require.

Where this leads. The ring and integral-domain structure established here is not special to $\mathbb{Z}$: the same axioms, the same order properties, and the same absolute-value toolkit reappear verbatim for $\mathbb{Q}$ and $\mathbb{R}$. Everything algebraic built later inherits the patterns established here.

Structural Roadmap

Both sources follow the same logical spine:

Equivalence Relation ⟶ Definitions ⟶ Well-Definedness ⟶ Ring Laws ⟶ Order ⟶ Absolute Value

Nothing is assumed beyond what has been proved for $\mathbb{N}$. The available toolkit at each step is exactly what precedes it — most critically, the cancellation law for natural-number addition, which carries the entire construction forward.

The global progression is:

1. The equivalence relation. Integers are equivalence classes of pairs $(a, b) \in \mathbb{N} \times \mathbb{N}$ under $(a, b) \sim (c, d) \iff a + d = c + b$. Transitivity requires cancellation in $\mathbb{N}$.

2. Addition and multiplication. Defined on representatives; well-definedness verified separately for each operation (Tao L4.1.3; Mendelson Lemmas 2.1, 2.3).

3. Ring laws. The nine identities of Tao's P4.1.6 (= Mendelson Thms 2.2, 2.4) are exactly the axioms of a commutative ring with unit. All proofs expand pairs and reduce to arithmetic in $\mathbb{N}$.

4. Integral domain. No zero divisors (P4.1.8 / Thm 2.4(v)) and the cancellation law (C4.1.9 / Thm 3.3) elevate $\mathbb{Z}$ from a ring to an integral domain. Mendelson's §3.3 names and defines these structures abstractly.

5. Order. Tao defines order directly via $\mathbb{N}$-differences (Def 4.1.10). Mendelson axiomatises an ordered integral domain (Def 4.1, axioms O1–O5) and instantiates it for $\mathbb{Z}$ via a positivity set (Thm 4.3, Cor 4.6). The same five axioms recur unchanged for $\mathbb{R}$.

6. Absolute value. Mendelson's Theorem 4.8 (11 parts, for any ordered integral domain) provides a permanent toolkit — triangle inequality, reverse triangle inequality, multiplicativity — that requires no re-proof when we reach $\mathbb{Q}$ and $\mathbb{R}$. Tao defers absolute value to the rational chapter.

7. Induction boundary. Induction fails for $\mathbb{Z}$ (Tao Ex 4.1.8): there is no smallest element to anchor a descent below 0. Strong induction on $\mathbb{N}$ remains available for the positive integers inside $\mathbb{Z}$; Mendelson Thm 4.7 recovers a Peano system from $\mathcal{P}_{\mathbb{Z}}$ to make this precise.

9.0.0.1 Tao Construction

Remark 9.1 (Why the integers are needed). The natural number system $\mathbb{N}$ with addition and multiplication has reached its limits: subtraction is not available. Given $a, b \in \mathbb{N}$ with $a < b$, the expression $a - b$ has no value in $\mathbb{N}$. To fix this, we pass to a larger system $\mathbb{Z}$ in which every natural number has an additive inverse.

Remark 9.2 (The circularity problem). The naive approach define an integer as a difference $a - b$ of two natural numbers is circular, because the symbol $-$ is exactly what we are trying to construct. Tao resolves this by introducing a placeholder symbol $--$ with no arithmetic meaning, using $a—b$ purely as notation for a formal pair (a, b). Once subtraction is defined at the end of the construction, we verify that $a—b$ coincides with $a - b$ and discard the scaffolding.

Remark 9.3 (Why not define integers as signed naturals?). One might define an integer as either a positive natural number, zero, or the negation of a positive natural number. Lemma 4.1.5 (trichotomy) shows this classification is correct, but using it as the definition forces case-splits for every arithmetic operation: negative $\times$ negative, positive + negative of different sizes, etc. The verification of the ring laws becomes enormously messy. The equivalence-class construction pays one upfront cost and avoids all subsequent case-explosion.

Definition 4.1.1 (Integers)

An integer is an expression of the form $a—b$, where a and b are natural numbers.
Two integers are equal:

$$a—b \;=\; c—d \quad \iff \quad a+d \;=\; c+b.$$

The set of all integers is denoted $\mathbb{Z}$.

Remark 9.4 (Reading the equality condition). The condition $a+d=c+b$ is the cross-addition criterion that would hold for genuine differences: $a-b=c-d \iff a+d=c+b$. It uses only addition in $\mathbb{N}$, which is already available, and avoids any reference to subtraction.

Remark 9.5 (Equality must be verified as a legitimate equivalence relation). Four axioms must be checked: reflexivity, symmetry, transitivity, and substitution. Reflexivity and symmetry are immediate from the definition. Transitivity requires the cancellation law for natural numbers (Proposition 2.2.6): the key load-bearing result from Chapter 2. Substitution cannot be verified until operations on $\mathbb{Z}$ are defined, and must be re-checked for each operation.

Definition 4.1.2 (Addition and Multiplication)

$$(a—b)+(c—d) \;:=\; (a+c)—(b+d)$$
$$(a—b)\times(c—d) \;:=\; (ac+bd)—(ad+bc)$$

Remark 9.6 (Motivation for the multiplication formula). Expanding the product of genuine differences: $(a-b)(c-d)=ac-ad-bc+bd=(ac+bd)-(ad+bc)$. The formula is forced by this foreknowledge.

Lemma 9.7 (Tao 4.1.3 Addition and multiplication are well-defined). Let $a,b,a',b',c,d\in\mathbb{N}$. If $(a—b)=(a'—b')$, then:

$$(a—b)+(c—d)=(a'—b')+(c—d),$$
$$(a—b)\times(c—d)=(a'—b')\times(c—d),$$

and similarly when replacing $(c—d)$ by an equal integer.

Remark 9.8 (Why well-definedness must be checked). The operations are defined in terms of representatives (a,b) of an equivalence class. If equal inputs (i.e., equal integers with different representatives) gave different outputs, the operation would not be a function on $\mathbb{Z}$ it would be a function on representations, which is meaningless. This check is the substitution axiom for each operation.

Remark 9.9 (Embedding $\mathbb{N}$ in $\mathbb{Z}$). The integers of the form $n—0$ behave identically to the natural numbers:

$$(n—0)+(m—0)=(n+m)—0, \qquad (n—0)\times(m—0)=nm—0.$$

Furthermore $(n—0)=(m—0)$ if and only if $n=m$. We therefore identify $n\equiv n—0$, embedding $\mathbb{N}$ into $\mathbb{Z}$. In particular $0=0—0$ and $1=1—0$.

Definition 4.1.4 (Negation)

$-(a—b) \;:=\; b—a.$
In particular, for a positive natural number $n=n—0$:

$$-n \;:=\; 0—n.$$

Remark 9.10 (Well-definedness of negation). Negation must also be checked as well-defined: if $(a—b) = (a'—b')$ then $-(a—b) = -(a'—b')$. This is Exercise 4.1.2.

Lemma 9.11 (Tao 4.1.5 Trichotomy of integers). Let x be an integer. Then exactly one of the following holds:

(a) $x = 0$.

(b) $x = n$ for some positive natural number n.

(c) $x = -n$ for some positive natural number n.

If (b) holds we call x a positive integer; if (c) holds we call x a negative integer.

Remark 9.12 (Proof sketch). Existence of one case: $x = a—b$. By trichotomy of $\mathbb{N}$ (P2.2.13), either $a > b$, $a = b$, or $a < b$. Each case yields (b), (a), or (c) respectively. Mutual exclusion uses P2.2.8 and Proposition 2.2.6.

Proposition 4.1.6 (Laws of Algebra for $\mathbb{Z}$)

Let $x, y, z \in \mathbb{Z}$. Then:

$x + y = y + x$	(commutativity of addition)
$(x + y) + z = x + (y + z)$	(associativity of addition)
$x + 0 = 0 + x = x$	(additive identity)
$x + (-x) = (-x) + x = 0$	(additive inverse)
$xy = yx$	(commutativity of multiplication)
$(xy)z = x(yz)$	(associativity of multiplication)
$x \cdot 1 = 1 \cdot x = x$	(multiplicative identity)
$x(y + z) = xy + xz$	(left distributive law)
$(y + z)x = yx + zx$	(right distributive law)

Remark 9.13 (This is the commutative ring axioms). These nine identities are exactly the definition of a commutative ring. Without the identity $xy = yx$ the remaining eight would give a ring. Note: these properties were proved for $\mathbb{N}$, but $\mathbb{Z} \supsetneq \mathbb{N}$, so the proofs must be redone.

Remark 9.14 (Proof strategy). Write $x = (a—b)$, $y = (c—d)$, $z = (e—f)$ and expand both sides in terms of natural number arithmetic. This is far cleaner than case-splitting on sign via Lemma 4.1.5. Tao demonstrates the method on associativity of multiplication.

Definition 9.15 (Subtraction). For integers x, y:

$$x - y \;:=\; x + (-y).$$

Remark 9.16 (Recovering the $--$ notation). Once subtraction is defined, one checks that for natural numbers a, b:

$$a - b = (a—0) + (0—b) = a—b.$$

The placeholder $--$ is now equal to genuine subtraction; the scaffolding is removed and $--$ is discarded.

Proposition 9.17 (Tao 4.1.8 No zero divisors). Let $a, b \in \mathbb{Z}$ with $ab = 0$. Then $a = 0$ or $b = 0$ (or both).

Corollary 9.18 (Tao 4.1.9 Cancellation law for $\mathbb{Z}$). Let $a, b, c \in \mathbb{Z}$ with $ac = bc$ and $c \neq 0$. Then $a = b$.

Definition 4.1.10 (Ordering of $\mathbb{Z}$)

Let $n, m \in \mathbb{Z}$.

$$n \geq m \iff n = m + a \text{ for some } a \in \mathbb{N}.$$

$$n > m \iff n \geq m \text{ and } n \neq m.$$

Remark 9.19 (Consistency with $\mathbb{N}$). This definition is verbatim the same as Definition 2.2.11 for $\mathbb{N}$. Since the embedding $n \equiv n—0$ is consistent with addition, the two orderings agree on natural numbers.

Lemma 9.20 (Tao 4.1.11 Properties of order on $\mathbb{Z}$). Let $a, b, c \in \mathbb{Z}$. Then:

(a) $a > b$ if and only if $a - b$ is a positive natural number.

(b) (Addition preserves order.) If $a > b$, then $a + c > b + c$.

(c) (Positive multiplication preserves order.) If $a > b$ and c is a positive integer, then $ac > bc$.

(d) (Negation reverses order.) If $a > b$, then $-a < -b$.

(e) (Transitivity.) If $a > b$ and $b > c$, then $a > c$.

(f) (Trichotomy.) Exactly one of $a > b$, $a < b$, $a = b$ holds.

Remark 9.21 (Strategy: derive from part (a)). Part (a) reformulates $>$ in terms of positivity. Parts (b)–(f) all follow from (a) by translating into statements about positive natural numbers, where the analogous results are already known.

Remark 9.22 (Exercise 4.1.8 Induction does not apply to $\mathbb{Z}$). Axiom P5 (induction) does not carry over to the integers. Specifically: there exists a property $P(n)$ of integers such that $P(0)$ is true and $P(n) \Rightarrow P(n{+}{+})$ for all $n \in \mathbb{Z}$, yet $P(n)$ fails for some $n \in \mathbb{Z}$. This is because $\mathbb{Z}$ has no smallest element there is nothing to anchor a descent below 0. The situation becomes worse for $\mathbb{Q}$ and $\mathbb{R}$.

Integer Toolkit Tao §4.1 Quick Reference

Label	Statement	Proof method
Def 4.1.1	$a—b = c—d \iff a + d = c + b$	Definition
Def 4.1.2	Addition and multiplication formulas	Definition
L4.1.3	Operations are well-defined	Expand; use $a + b' = a' + b$
Def 4.1.4	$-(a—b) := b—a$	Definition
L4.1.5	Trichotomy of $\mathbb{Z}$	Cases $a > b$, $a = b$, $a < b$
P4.1.6	Nine ring laws	Expand via representatives
Def 4.1.7	$x - y := x + (-y)$	Definition
P4.1.8	No zero divisors	Uses L2.3.3
C4.1.9	Cancellation: $ac = bc, c \neq 0 \Rightarrow a = b$	Uses P4.1.8 or C2.3.7+L4.1.5
Def 4.1.10	Order: $n \geq m \iff n = m + a$, $a \in \mathbb{N}$	Definition
L4.1.11	Six order properties	Use part (a) as bridge

9.0.0.2 Mendelson Construction

ğ3.1 The Equivalence Relation on $P \times P$

Remark 9.23 (Starting point). Mendelson begins from $P \times P$, the set of all ordered pairs of positive integers (his notation for $\mathbb{N} \setminus \{0\}$). Tao uses $\mathbb{N} \times \mathbb{N}$ (including zero). The idea is identical: a pair (n, j) represents the "formal difference" $n - j$.

Definition 3.1 (Equivalence relation on $P \times P$)

For natural numbers n, k, j, i, define a relation $\sim$ on ordered pairs by

$$(n, j) \sim (k, i) \quad \Longleftrightarrow \quad n + i = k + j.$$

Theorem 9.24 (Mendelson 1.1 — $\sim$ is an equivalence relation). For all natural numbers h, i, j, k, m, n:

(R) $(h, i) \sim (h, i)$ (Reflexivity)

(S) $(h, i) \sim (j, k) \Rightarrow (j, k) \sim (h, i)$ (Symmetry)

(T) $\big[(h, i) \sim (j, k) \wedge (j, k) \sim (m, n)\big] \Rightarrow (h, i) \sim (m, n)$ (Transitivity)

Remark 9.25 (Transitivity uses cancellation). As in Tao, transitivity of $\sim$ requires the cancellation law for addition in $\mathbb{N}$. This is the same load-bearing step in both constructions.

Definition 3.2 (The integers $\mathbb{Z}$)

$\mathbb{Z}$ is the set of all equivalence classes of $P \times P$ under $\sim$. Elements of $\mathbb{Z}$ are called integers. Distinguished elements: $0_{\mathbb{Z}} = [(1, 1)], \quad 1_{\mathbb{Z}} = [(2, 1)]$.

Remark 9.26 (Notation comparison). Mendelson writes $[(n, j)]$ for an equivalence class; Tao writes $a—b$ for the same object. Mendelson's notation makes the set-theoretic content explicit. Tao's notation is cleaner for computation.

ğ3.2 Addition and Multiplication

Lemma 9.27 (Mendelson 2.1 — Addition well-defined). If $(n, j) \sim (n_1, j_1)$ and $(k, i) \sim (k_1, i_1)$, then $(n + k,\ j + i) \sim (n_1 + k_1,\ j_1 + i_1)$.

Definition 3.3 (Addition)

For $\alpha, \beta \in \mathbb{Z}$ with representatives $(n, j) \in \alpha$ and $(k, i) \in \beta$:

$$\alpha +_{\mathbb{Z}} \beta = [(n + k,\ j + i)].$$

Lemma 2.1 guarantees independence of representatives.

Theorem 9.28 (Mendelson 2.2 — Properties of addition). For all $\alpha, \beta, \gamma \in \mathbb{Z}$:

(i) Commutativity: $\alpha +_{\mathbb{Z}} \beta = \beta +_{\mathbb{Z}} \alpha$.

(ii) Associativity: $\alpha +_{\mathbb{Z}} (\beta +_{\mathbb{Z}} \gamma) = (\alpha +_{\mathbb{Z}} \beta) +_{\mathbb{Z}} \gamma$.

(iii) Additive identity: $\alpha +_{\mathbb{Z}} 0_{\mathbb{Z}} = \alpha$.

(iv) Unique additive inverse: $\exists!\, \delta \in \mathbb{Z}$ such that $\alpha +_{\mathbb{Z}} \delta = 0_{\mathbb{Z}}$.

Remark 9.29 (Uniqueness of the inverse). Part (iv) asserts both existence and uniqueness of $-\alpha$. Tao's P4.1.6 asserts existence only; uniqueness follows later from C4.1.9. Mendelson builds uniqueness into the statement.

Lemma 9.30 (Mendelson 2.3 — Multiplication well-defined). If $(n, j) \sim (n_1, j_1)$ and $(k, i) \sim (k_1, i_1)$, then $(nk + ji,\ jk + ni) \sim (n_1k_1 + j_1i_1,\ j_1k_1 + n_1i_1)$.

Definition 3.4 (Multiplication)

For $\alpha, \beta \in \mathbb{Z}$ with representatives $(n, j) \in \alpha$ and $(k, i) \in \beta$:

$$\alpha \times_{\mathbb{Z}} \beta = [(nk + ji,\ jk + ni)].$$

Lemma 2.3 guarantees well-definedness.

Theorem 9.31 (Mendelson 2.4 — Properties of multiplication). For all $\alpha, \beta, \gamma \in \mathbb{Z}$:

(i) Commutativity: $\alpha \times_{\mathbb{Z}} \beta = \beta \times_{\mathbb{Z}} \alpha$.

(ii) Associativity: $\alpha \times_{\mathbb{Z}} (\beta \times_{\mathbb{Z}} \gamma) = (\alpha \times_{\mathbb{Z}} \beta) \times_{\mathbb{Z}} \gamma$.

(iii) Distributivity: $\alpha \times_{\mathbb{Z}} (\beta +_{\mathbb{Z}} \gamma) = (\alpha \times_{\mathbb{Z}} \beta) +_{\mathbb{Z}} (\alpha \times_{\mathbb{Z}} \gamma)$.

(iv) Multiplicative identity: $\alpha \times_{\mathbb{Z}} 1_{\mathbb{Z}} = \alpha$.

(v) No zero divisors: $\alpha \neq 0_{\mathbb{Z}} \wedge \beta \neq 0_{\mathbb{Z}} \Rightarrow \alpha \times_{\mathbb{Z}} \beta \neq 0_{\mathbb{Z}}$.

Remark 9.32 (Structural consequence). Theorems 2.2 and 2.4 establish that $(\mathbb{Z}, +_{\mathbb{Z}}, \times_{\mathbb{Z}})$ is an integral domain: a commutative ring with unit and no zero divisors. Mendelson makes this explicit in §3.3 via abstract definitions.

ğ3.3 Rings and Integral Domains (Abstract Theory)

Remark 9.33 (Abstract vs concrete). Mendelson develops ring theory abstractly before applying it to $\mathbb{Z}$. Tao works concretely throughout, naming the ring structure only in Remark 4.1.7.

Definition 3.5 (Ring; integral domain)

A ring $(R, +, \times)$: $(R, +)$ is an abelian group; $\times$ is associative; $\times$ distributes over $+$ on both sides. A ring is commutative if $xy = yx$; has a unit if $\exists\, 1$ with $x \cdot 1 = 1 \cdot x = x$. A nonzero x is a zero divisor if $\exists$ nonzero y with $xy = 0$. An integral domain is a commutative ring with unit, $0 \neq 1$, and no zero divisors.

Theorem 9.34 (Mendelson 3.3 — Cancellation ⇔ no zero divisors). In a commutative ring with unit:

$$(\forall x, y, z:\ xy = xz \wedge x \neq 0 \Rightarrow y = z) \iff \text{no zero divisors.}$$

Theorem 9.35 (Mendelson 3.4 — Trivial ring). In a ring with unit: $0 = 1 \iff R$ is a singleton.

ğ4 Ordered Integral Domains and Order on $\mathbb{Z}$

Remark 9.36 (Strategy). Mendelson proves order theory for any integral domain first, then instantiates it for $\mathbb{Z}$. This gives order on $\mathbb{Q}$ and $\mathbb{R}$ for free later.

Definition 4.1 (Ordered integral domain)

$(R, +, \times, <)$ is an ordered integral domain if $(R, +, \times)$ is an integral domain ($0 \neq 1$) and $<$ satisfies:

(O1) $x \not< x$	Irreflexivity
(O2) $x < y \wedge y < z \Rightarrow x < z$	Transitivity
(O3) $x < y \vee x = y \vee y < x$	Trichotomy
(O4) $x < y \Rightarrow x + z < y + z$	Addition-monotone
(O5) $x < y \wedge 0 < z \Rightarrow xz < yz$	Positive-mult-monotone

x is positive if $0 < x$; negative if $x < 0$.

Theorem 9.37 (Mendelson 4.1 — Consequences of order axioms). In any ordered integral domain: (i) exactly one of $x < y$, $x = y$, $y < x$ holds; (ii) $x < y \wedge u < v \Rightarrow x + u < y + v$; (iii) $0 < z \wedge xz < yz \Rightarrow x < y$.

Theorem 9.38 (Mendelson 4.2 — Order and positivity). In any ordered integral domain: (i) $x < y \iff y - x$ positive; (ii) $x < y \iff x - y$ negative; (iii) $x < y \iff -y < -x$; (iv) sum and product of positives are positive; (v) product of two negatives is positive.

Theorem 9.39 (Mendelson 4.3 — Positivity set construction). Let $(R, +, \times)$ be an integral domain, $0 \neq 1$. If $\mathcal{P} \subseteq R$ satisfies: $0 \notin \mathcal{P}$; $\forall x:\ x \in \mathcal{P} \vee x = 0 \vee -x \in \mathcal{P}$; $\mathcal{P}$ closed under $+$ and $\times$; then $(R, +, \times, <)$ with $x < y \iff y - x \in \mathcal{P}$ is an ordered integral domain and $\mathcal{P}$ is its positive set.

Remark 9.40 (Why this matters). Theorem 4.3 reduces verifying five order axioms to four positivity conditions. Mendelson uses it to order $\mathbb{Z}$ via $\mathcal{P}_{\mathbb{Z}}$ without checking (O1)–(O5) directly.

Lemma 9.41 (Mendelson 4.4 — Positivity is class-invariant). For $\alpha \in \mathbb{Z}$ and any two representatives $(n, j), (k, i) \in \alpha$: $j < n \iff i < k$.

Definition 4.5 (Positivity set $\mathcal{P}_{\mathbb{Z}}$)

$$\mathcal{P}_{\mathbb{Z}} = \{\alpha \in \mathbb{Z} : \forall (n, j) \in \alpha,\ j < n\}.$$

Lemma 4.4 ensures this is well-defined. Intuitively: α is positive iff its representative has second coordinate strictly less than first, i.e. $n - j > 0$.

Lemma 9.42 (Mendelson 4.5). $\mathcal{P}_{\mathbb{Z}}$ satisfies the four conditions of Theorem 4.3.

Corollary 9.43 (Mendelson 4.6). $(\mathbb{Z}, +_{\mathbb{Z}}, \times_{\mathbb{Z}}, <_{\mathbb{Z}})$ is an ordered integral domain with positive set $\mathcal{P}_{\mathbb{Z}}$.

Theorem 9.44 (Mendelson 4.7 — Recovery of Peano system). Let $T(x) = x +_{\mathbb{Z}} 1_{\mathbb{Z}}$ for $x \in \mathcal{P}_{\mathbb{Z}}$. Then $(\mathcal{P}_{\mathbb{Z}}, T, 1_{\mathbb{Z}})$ is a Peano system.

Remark 9.45 (Significance). Starting from a Peano system for $\mathbb{N}$, we construct $\mathbb{Z}$, and then prove the positive integers inside $\mathbb{Z}$ form a new Peano system consistent with the original. Tao does not prove this explicitly.

ğ4 Absolute Value

Definition 4.6 (Absolute value)

In any ordered integral domain: $|x| = x$ if $0 \le x$; $|x| = -x$ if $x < 0$.

Theorem 9.46 (Mendelson 4.8 — Properties of absolute value). In any ordered integral domain, $|\cdot|$ satisfies: (1) $|x| \ge 0$; (2) $|x| = 0 \iff x = 0$; (3) $|-x| = |x|$; (4) $|x - y| = |y - x|$; (5) $|xy| = |x||y|$; (6) $-|x| \le x \le |x|$; (7) $|z| < u \iff -u < z < u$; (8) $|z| \le u \iff -u \le z \le u$; (9) $u \ge v \wedge u \ge -v \Rightarrow u \ge |v|$; (10) $|x + y| \le |x| + |y|$ (triangle inequality); (11) $|x - y| \ge \big||x| - |y|\big|$ (reverse triangle inequality).

ğ5 Division, Divisibility, Primes

Remark 9.47 (Scope). This material is not in Tao Ch 4 and is not a prerequisite for real analysis. It is valuable background for number theory and abstract algebra. See the comparison table for coverage decisions.

Theorem 9.48 (Mendelson 5.1 — Euclidean division). For any integer $\alpha > 1$ and any integer β, there exist unique integers q, r with $\beta = q\alpha + r$ and $0 \le r < \alpha$.

Definition 9.49 (Divisibility). $\alpha \mid \beta \iff \exists \gamma \in \mathbb{Z} : \beta = \alpha\gamma$.

Theorem 9.50 (Mendelson 5.2–5.9). Standard divisibility properties, existence and Bézout form of gcd, characterisation of relative primeness, infinitude of primes (Euclid), Euclid's lemma ($\rho \mid \alpha\beta \Rightarrow \rho \mid \alpha$ or $\rho \mid \beta$), and the fundamental theorem of arithmetic (unique prime factorisation up to order and sign).

ğ6 Integers Modulo n (Optional)

Remark 9.51 (Scope). Congruence mod n as an equivalence relation, $\mathbb{Z}_n$ as a commutative ring with unit; integral domain iff n is prime. Not required for the real analysis track.

ğ7–9 Integer Action, Embedding Theory, Uniqueness of $\mathbb{Z}$

Remark 9.52 (Scope). Chapters 7–9 are graduate-level algebra. Key results: Thm 8.8: every characteristic-0 integral domain contains a copy of $\mathbb{Z}$. Thm 9.5: $\mathcal{D} \cong \mathbb{Z}$ iff characteristic 0 and no proper subdomains. Thm 9.8: any two well-ordered integral domains are isomorphic — $\mathbb{Z}$ is unique up to isomorphism. These results are beyond what is needed for real analysis but are worth knowing as statements.

Definition 9.53 (Well-ordered integral domain). An ordered integral domain is well-ordered if every nonempty subset of its positive elements contains a least element.

Theorem 9.54 (Mendelson 9.8 — Uniqueness of $\mathbb{Z}$). Any two well-ordered integral domains are isomorphic. In particular, $\mathbb{Z}$ is, up to isomorphism, the unique well-ordered integral domain.

9.0.0.3 Tao vs. Mendelson: Comparison Table

Tao vs. Mendelson: Integer Theory Statement-by-Statement Comparison

Remark (How to read this table). Each row is one mathematical concept. Coverage recommendations: Core = do both; Tao only = Tao suffices, Mendelson adds abstraction only; Men. only = Mendelson covers this, Tao does not; Skip = beyond real analysis scope for now.

Block 1: The Construction of $\mathbb{Z}$

Concept	Tao	Mendelson	Same idea?	Coverage
Formal pairs as integers	Def 4.1.1: $a—b$, $a—b = c—d \iff a+d = c+b$	Def 3.1–3.2: $(n,j) \sim (k,i) \iff n+i = k+j$; $\mathbb{Z}$ = equiv. classes	Yes identical relation, different notation	Core
Equality is an equivalence relation	Ex 4.1.1 (reflexivity, symmetry); transitivity proved in text	Thm 1.1 (all three parts)	Yes	Core
Distinguished elements $0, 1$	$0 = 0—0$, $1 = 1—0$ (implicit in Def 4.1.1)	$0_{\mathbb{Z}} = [(1,1)]$, $1_{\mathbb{Z}} = [(2,1)]$ (explicit definition)	Yes same objects	Core
Embedding $\mathbb{N} \hookrightarrow \mathbb{Z}$	Remark after Def 4.1.2: $n \equiv n—0$	Thm 4.7 + Cor 9.7 (recovered as Peano system)	Tao informal; Mendelson proves formally	Core

Block 2: Addition

Concept	Tao	Mendelson	Same?	Coverage
Addition formula	Def 4.1.2: $(a—b) + (c—d) := (a + c)—(b + d)$	Def 3.3 (after Lem 2.1): $\alpha +_{\mathbb{Z}} \beta = [(n + k, j + i)]$	Yes	Core
Well-definedness of addition	L4.1.3 (combines add. and mult.)	Lem 2.1 (addition separately)	Yes Mendelson separates the two	Core
Commutativity	P4.1.6 (i)	Thm 2.2 (i)	Yes	Core
Associativity	P4.1.6 (ii)	Thm 2.2 (ii)	Yes	Core
Additive identity	P4.1.6 (iii): $x + 0 = 0 + x = x$	Thm 2.2 (iii): $\alpha + 0_{\mathbb{Z}} = \alpha$	Yes	Core
Additive inverse (negation)	Def 4.1.4: $-(a—b) := b—a$; P4.1.6 (iv) existence	Thm 2.2 (iv): unique δ with $\alpha + \delta = 0$	Tao: existence; Men.: existence + uniqueness in one statement	Core
Negation well-defined	Ex 4.1.2	(Implicit in Lem 2.1 / Thm 2.2)	Yes Tao makes it an explicit exercise	Core

Block 3: Multiplication

Concept	Tao	Mendelson	Same?	Coverage
Multiplication formula	Def 4.1.2: $(a—b)(c—d) := (ac+bd)—(ad+bc)$	Def 3.4 (after Lem 2.3): $\alpha \times_{\mathbb{Z}} \beta = [(nk+ji, jk+ni)]$	Yes	Core
Well-definedness of multiplication	L4.1.3	Lem 2.3	Yes	Core
Commutativity	P4.1.6 (v)	Thm 2.4 (i)	Yes	Core
Associativity	P4.1.6 (vi); proved in text as model calculation	Thm 2.4 (ii)	Yes	Core
Multiplicative identity	P4.1.6 (vii)	Thm 2.4 (iv)	Yes	Core
Distributive law	P4.1.6 (viii, ix)	Thm 2.4 (iii)	Yes	Core
$(-1) \times a = -a$	Ex 4.1.3	Not stated separately (follows from ring axioms)	Tao only as exercise	Core (Tao)
No zero divisors	P4.1.8	Thm 2.4 (v)	Yes	Core
Cancellation law	C4.1.9	Thm 3.3 (as equivalence, in any comm. ring with unit)	Mendelson more general	Core
$\mathbb{Z}$ is a commutative ring / integral domain	Rem 4.1.7 (named but not defined abstractly)	Def 3.5 + Thms 2.2, 2.4 (built up formally)	Mendelson makes it explicit	Core

Block 4: Order

Concept	Tao	Mendelson	Same?	Coverage
Order definition	Def 4.1.10: $n \geq m \iff n = m + a$, $a \in \mathbb{N}$	Def 4.1 (abstract OID axioms) + $<_{\mathbb{Z}}$ via $\mathcal{P}_{\mathbb{Z}}$ (Cor 4.6)	Same result; Mendelson builds via positivity set	Core
Order irreflexivity	(Implicit in Def 4.1.10 and L4.1.11(f))	Explicitly O1 in Def 4.1	Men. more explicit	Core
Trichotomy of integers	L4.1.5: every integer is positive, zero, or negative (one only)	Thm 4.1(i) + Def 4.1 O3	Yes	Core
$a > b \iff a - b$ positive	L4.1.11(a)	Thm 4.2(i)	Yes	Core
Addition preserves order	L4.1.11(b)	Def 4.1 O4 (axiom); Thm 4.1(ii) (addition of inequalities)	Yes	Core
Positive mult. preserves order	L4.1.11(c)	Def 4.1 O5 (axiom)	Yes	Core
Negation reverses order	L4.1.11(d)	Thm 4.2(iii)	Yes	Core
Transitivity	L4.1.11(e)	Def 4.1 O2 (axiom)	Yes	Core
Order trichotomy (strict)	L4.1.11(f)	Thm 4.1(i)	Yes	Core
Positivity set construction (Thm 4.3)	—	Thm 4.3: $\mathcal{P}$ satisfying 4 conditions $\Rightarrow$ OID	Men. only	Men. only (worth knowing)
Product of negatives is positive	(Derivable from P4.1.6 + L4.1.11)	Thm 4.2(v) (stated explicitly)	Tao implicit; Men. explicit	Core
Recovery of Peano system in $\mathbb{Z}$	—	Thm 4.7: $(\mathcal{P}_{\mathbb{Z}}, T, 1_{\mathbb{Z}})$ is a Peano system	Men. only	Men. only (important conceptually)

Block 5: Absolute Value

Concept	Tao	Mendelson	Same?	Coverage
Definition of $\|x\|$	(Defined for rationals in Def 4.3.1, not integers separately)	Def 4.6 (for any ordered integral domain)	Men. gives it for $\mathbb{Z}$ directly; Tao defers to $\mathbb{Q}$	Men. only (do it here)
Triangle inequality $\|x+y\| \leq \|x\|+\|y\|$	P4.3.3(b) (for rationals)	Thm 4.8(10)	Yes Mendelson proves it for $\mathbb{Z}$	Core
Multiplicativity $\|xy\| = \|x\|\|y\|$	P4.3.3(d) (for rationals)	Thm 4.8(5)	Yes	Core
Reverse triangle inequality	(Implicit)	Thm 4.8(11) explicitly	Men. explicit	Core
Full 11-part absolute value theorem	Spread across P4.3.3 (for $\mathbb{Q}$)	Thm 4.8 (complete, for any OID)	Men. more systematic	Men. only (do once, use forever)

Block 6: Induction and Structural Observations

Concept	Tao	Mendelson	Same?	Coverage
Induction fails for $\mathbb{Z}$	Ex 4.1.8: explicit counterexample	Implied by lack of least element; Thm 9.6 addresses positives only	Tao makes it an explicit exercise	Core (Tao)
Subtraction definition	$x-y := x+(-y)$ (after Def 4.1.4)	Standard (follows from additive inverse)	Yes	Core
$0 \neq 1$ in $\mathbb{Z}$	Implicit (P2.2.8 + L4.1.5)	Thm 3.4 (explicit: $0=1 \iff$ trivial ring)	Men. explicit	Core

Block 7: Material in Mendelson Only — Coverage Decision Required

Concept	Mendelson ref.	Prereq for RA?	Recommendation
Abstract ring and integral domain definitions	ğ3.3, Def 3.5	No (but clarifying)	Include notes; no proof sheets needed
Cancellation $\Leftrightarrow$ no zero divisors (abstract)	Thm 3.3	No	Notes only; proof optional
Trivial ring theorem ($0 = 1 \Rightarrow$ singleton)	Thm 3.4	No	Skip proof; note the statement
Abstract ordered integral domain axioms	Def 4.1	Indirectly (same axioms used for $\mathbb{R}$)	Include these axioms recur throughout real analysis
Positivity set construction (Thm 4.3)	Thm 4.3	No (used internally)	Notes only elegant but not exercised directly
Peano recovery (Thm 4.7)	Thm 4.7	No	Notes only; important for logical completeness
Euclidean division (ğ5, Thm 5.1)	Thm 5.1	No	Proof sheet classical and useful
Divisibility properties (ğ5, Thm 5.2)	Thm 5.2 (11 parts)	No	Proof sheet good algebraic practice
GCD existence and Bézout (Thm 5.3)	Thm 5.3	No	Proof sheet important
Infinitude of primes (Thm 5.6)	Thm 5.6	No	Proof sheet Euclid's classic argument
Fundamental theorem of arithmetic (Thm 5.9)	Thm 5.9	No	Proof sheet core number theory
Modular arithmetic (ğ6)	Thms 6.1–6.6	No	Skip for now; revisit in algebra track
Integer action on domains (ğ7, Thms 7.1–7.4)	Thms 7.1–7.4	No	Skip graduate algebra
$N_{\mathcal{D}}$, $Z_{\mathcal{D}}$ embedding theory (ğ8)	Thms 8.1–8.9	No	Skip graduate algebra
Subdomain theory (ğ9, Thms 9.1–9.4)	Thms 9.1–9.4	No	Skip graduate algebra
Characterisation of $\mathbb{Z}$ (Thm 9.5)	Thm 9.5	No	Notes only worth knowing as a statement
Well-ordered integral domain (Def, Cor 9.7)	Def + Cor 9.7	No	Notes only
Uniqueness of $\mathbb{Z}$ up to isomorphism (Thm 9.8)	Thm 9.8	No	Note the statement philosophically important

Summary: Recommended Coverage

Core (both sources): Equality, addition, multiplication, well-definedness, ring laws, no zero divisors, cancellation, order, trichotomy, six order properties.

Mendelson adds (do notes, consider proof sheets): Abstract ring/integral domain definitions (§3.3); abstract OID axioms (Def 4.1) — these recur in real analysis; absolute value on $\mathbb{Z}$ (Thm 4.8) — Tao defers to $\mathbb{Q}$; positivity set construction (Thm 4.3) — elegant technique; Peano recovery (Thm 4.7) — logical closure; number theory: Euclidean division, divisibility, GCD, primes (ğ5).

Skip for now (graduate algebra): Modular arithmetic (ğ6), integer action on domains (ğ7), $N_{\mathcal{D}}/Z_{\mathcal{D}}$ embedding theory (ğ8–9). Know the statements of Thms 9.5 and 9.8 (uniqueness of $\mathbb{Z}$).

9.1 Proofs

Chapter 10

Rational Numbers ($\mathbb{Q}$)

Where You Are in the Journey

$$\mathbb{N} \;\to\; \text{Integers } (\mathbb{Z}) \;\to\; \text{Rationals } (\mathbb{Q}) \;\to\; \text{Real Numbers } (\mathbb{R}) \;\to\; \cdots$$

How we got here. The integers gave us additive inverses, making subtraction always possible. But division remains problematic: $1 \div 2$ has no integer solution. The rationals adjoin multiplicative inverses for all non-zero integers, giving us a field.

What this chapter will build. The construction of $\mathbb{Q}$ from $\mathbb{Z}$ via equivalence classes of pairs, the field axioms, the dense order, and the incompleteness of $\mathbb{Q}$ (existence of $\sqrt{2}$ gaps).

Where this leads. The incompleteness of $\mathbb{Q}$ is the motivation for constructing $\mathbb{R}$. The Dedekind and Cauchy constructions both start from $\mathbb{Q}$.

Status: Planned

Coming Soon

Notes, proofs, and exercises will appear here in a future revision.

10.1 Notes

To be populated.

10.2 Proofs

To be populated.

10.3 Capstone

To be populated.

Chapter 11

Real Numbers ($\mathbb{R}$)

11.1 Notes

Where You Are in the Journey

Propositional Logic $\rightarrow$ Predicate Calculus $\rightarrow$ Sets & Functions $\rightarrow$ Proof Techniques $\rightarrow$ Axiom Systems $\rightarrow$ $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{Q}$ $\rightarrow$ Real Numbers ($\mathbb{R}$) $\rightarrow$ Real Analysis $\rightarrow$ $\cdots$

How we got here. The rational numbers gave us a dense ordered field, but they contain gaps: there is no $q \in \mathbb{Q}$ with $q^2 = 2$. The real number system fills these gaps. Two rigorous constructions achieve this — Dedekind cuts and Cauchy sequences — and both yield the same unique complete ordered field.

What this chapter builds. The axiomatic structure of $\mathbb{R}$: field axioms, order axioms, the completeness axiom, and their immediate consequences — the Archimedean property, density of $\mathbb{Q}$, and the interval and bound theory that underpins all of analysis. We also study how $\mathbb{R}$ is constructed from $\mathbb{Q}$ via Dedekind cuts and Cauchy sequences.

Where this leads. Real analysis in Volume III takes $\mathbb{R}$ as given and develops sequences, series, convergence, and their limiting behaviour. The field and order axioms established here propagate unchanged into metric spaces, normed spaces, and abstract algebra.

Structural Roadmap

This chapter covers the algebraic and analytic axioms of $\mathbb{R}$ and the constructions that justify its existence. Sequence and series theory is developed in Volume III, Real Analysis.

Axioms $\longrightarrow$ Order & Bounds $\longrightarrow$ Completeness $\longrightarrow$ Constructions

The global progression is:

1. Field and order axioms
2. Intervals and convexity
3. Bounds: upper/lower bounds, maximum, minimum
4. Extremal values: supremum and infimum
5. Completeness: the least upper bound property and its consequences
6. Constructions of $\mathbb{R}$: Dedekind cuts
7. Constructions of $\mathbb{R}$: Cauchy sequences
8. Interval arithmetic

Remark 11.1 (Primary sources). Axiomatic development follows Abbott, Understanding Analysis and Ross, Elementary Analysis. The Cauchy sequence construction follows Tao, Analysis I, Chapters 5–6. The Dedekind cut construction follows Rudin, Principles of Mathematical Analysis, Appendix to Chapter 1.

11.1.1 Axioms of the Real Numbers

Axioms Of The Reals Quick Reference	
Core items	Key definitions/results introduced in this file.
How to use	Read the boxed items first; proofs and consequences follow.
Dependencies	Refer back to earlier sections as needed.

The real numbers $\mathbb{R}$ form a totally ordered field. This structure consists of:

- Field axioms (algebraic structure),
- Order axioms (order structure),
- Completeness axiom (analytic structure).

11.1.1.1 Basic Definitions

Remark 11.2. A field is a set equipped with two binary operations, addition ($+$) and multiplication ($\cdot$), satisfying closure, associativity, commutativity, identity, inverse, and distributive laws.

An ordered field is a field equipped with a total order compatible with the algebraic operations.

11.1.1.2 Main Theorems (Axioms)

Additive Axioms

Axiom A1 (Additive Closure).

$$\forall x \forall y \, (x, y \in \mathbb{R} \to x + y \in \mathbb{R})$$

Axiom A2 (Additive Commutativity).

$$\forall x \forall y \, (x + y = y + x)$$

Axiom A3 (Additive Associativity).

$$\forall x \forall y \forall z \, ((x + y) + z = x + (y + z))$$

Axiom A4 (Additive Identity).

$$\exists 0 \, \forall x \, (x + 0 = x)$$

Axiom A5 (Additive Inverse).

$$\forall x \, \exists y \, (x + y = 0)$$

Multiplicative Axioms

Axiom M1 (Multiplicative Closure).

$$\forall x \forall y \, (x, y \in \mathbb{R} \to x \cdot y \in \mathbb{R})$$

Axiom M2 (Multiplicative Commutativity).

$$\forall x \forall y \, (x \cdot y = y \cdot x)$$

Axiom M3 (Multiplicative Associativity).

$$\forall x \forall y \forall z \, ((x \cdot y) \cdot z = x \cdot (y \cdot z))$$

Axiom M4 (Multiplicative Identity).

$$\exists 1 \, (1 \neq 0 \wedge \forall x \, (x \cdot 1 = x))$$

Axiom M5 (Multiplicative Inverse).

$$\forall x \, (x \neq 0 \to \exists y \, (x \cdot y = 1))$$

Distributive Axiom

Axiom D (Distributivity).

$$\forall x \forall y \forall z \, (x \cdot (y + z) = x \cdot y + x \cdot z)$$

Linear Order Axioms

O1 (Reflexivity).

$$\forall x \in \mathbb{R}, \quad x \leq x$$

O2 (Antisymmetry).

$$\forall x, y \in \mathbb{R}, \quad (x \leq y \wedge y \leq x) \to x = y$$

O3 (Transitivity).

$$\forall x, y, z \in \mathbb{R}, \quad (x \leq y \wedge y \leq z) \to x \leq z$$

O4 (Totality / Comparability).

$$\forall x, y \in \mathbb{R}, \quad x \leq y \ \vee \ y \leq x$$

The strict order is defined by:

$$x < y \quad \text{iff} \quad x \leq y \text{ and } x \neq y.$$

Compatibility with Field Operations

O5 (Additive Monotonicity).

$$\forall x, y, z \in \mathbb{R}, \quad x \leq y \to x + z \leq y + z$$

O6 (Multiplicative Monotonicity for Nonnegative Factors).

$$\forall x, y, z \in \mathbb{R}, \quad (x \leq y \land 0 \leq z) \to xz \leq yz$$

O7 (Positivity of the Unit).

$$0 < 1$$

Remark 11.3. Axiom O6 implies that multiplication by a negative number reverses inequalities; this will later be proved as a theorem.

11.1.1.3 Consequences

The logical implication of this entire section is:

$$\text{Field Axioms} \Rightarrow \text{Algebraic Structure}$$

$$\text{Field} + \text{Order Axioms} \Rightarrow \text{Ordered Field}$$

To uniquely characterize $\mathbb{R}$ among ordered fields, one must additionally assume:

$$\text{Completeness.}$$

Remark 11.4 (Logical Structure).

$$\text{Field} \Rightarrow \text{Ordered Field} \Rightarrow \text{Complete Ordered Field.}$$

The real numbers $\mathbb{R}$ are the unique (up to isomorphism) complete ordered field.

11.1.2 Intervals in the Real Numbers

Intervals Quick Reference	
Core items	Key definitions/results introduced in this file.
How to use	Read the boxed items first; proofs and consequences follow.
Dependencies	Refer back to earlier sections as needed.

Intervals are fundamental subsets of the real line defined using the order relation on $\mathbb{R}$. Let $a, b \in \mathbb{R}$ with $a < b$.

11.1.2.1 Basic Definitions

Bounded Intervals

Definition (Open Interval)

The open interval from a to b is the set

$$(a, b) := \{x \in \mathbb{R} : a < x < b\}.$$

$$\forall x \, (x \in (a, b) \leftrightarrow a < x \wedge x < b)$$

Definition (Closed Interval)

The closed interval from a to b is the set

$$[a, b] := \{x \in \mathbb{R} : a \leq x \leq b\}.$$

$$\forall x \, (x \in [a, b] \leftrightarrow a \leq x \wedge x \leq b)$$

Definition (Half-Open Intervals)

The left-closed, right-open interval is

$$[a, b) := \{x \in \mathbb{R} : a \leq x < b\}.$$

The left-open, right-closed interval is

$$(a, b] := \{x \in \mathbb{R} : a < x \leq b\}.$$

Unbounded Intervals

Definition (Open Rays)

The open rays determined by $a \in \mathbb{R}$ are

$$(a, \infty) := \{x \in \mathbb{R} : x > a\}, \qquad (-\infty, a) := \{x \in \mathbb{R} : x < a\}.$$

Definition (Closed Rays)

The closed rays determined by $a \in \mathbb{R}$ are

$$[a, \infty) := \{x \in \mathbb{R} : x \geq a\}, \qquad (-\infty, a] := \{x \in \mathbb{R} : x \leq a\}.$$

Degenerate and Trivial Intervals

Definition (Degenerate Interval)

If $a = b$, the closed interval

$$[a, a] = \{a\}$$

is called a degenerate interval.

Definition (Empty Interval)

If $a > b$, the set

$$(a, b) = \varnothing$$

is called an empty interval.

11.1.2.2 Main Theorems

Theorem 11.5 (Characterization of Intervals). A subset $I \subseteq \mathbb{R}$ is an interval if and only if whenever $x, z \in I$ and $x < y < z$, then $y \in I$. Equivalently,

$$\forall x, z \in I,\ x < z \Rightarrow \big(\forall y \in \mathbb{R},\ (x < y < z \Rightarrow y \in I)\big).$$

Remark 11.6. Intervals are precisely the subsets of $\mathbb{R}$ with the property that if $x < y < z$ and x, z belong to the set, then y also belongs to the set.

11.1.2.3 Consequences

The logical implication of this section is:

$$\text{Order Structure of } \mathbb{R} \Rightarrow \text{Intervals} \Rightarrow \text{Convex subsets of } \mathbb{R}.$$

Intervals encode the idea of no gaps between points of a set. They are the basic building blocks for:

- neighborhoods,
- open sets,
- topology of $\mathbb{R}$,
- and completeness arguments (e.g. nested intervals).

Remark 11.7 (Logical Structure). The major structural flow is:

$$\text{Field Axioms} \Rightarrow \text{Order Axioms} \Rightarrow \text{Definition of Intervals} \Rightarrow \text{Convexity Property} \Rightarrow \text{Topological Structure of } \mathbb{R}.$$

11.1.3 Bounds and Extremal Values

Bounds Quick Reference

Core items	Key definitions/results introduced in this file.
How to use	Read the boxed items first; proofs and consequences follow.
Dependencies	Refer back to earlier sections as needed.

11.1.3.1 Basic Definitions

Definition (Upper Bound)

Let $A \subseteq \mathbb{R}$. A number $u \in \mathbb{R}$ is an upper bound for A if

$$\forall a \in A, \quad a \leq u.$$

Definition (Lower Bound)

Let $A \subseteq \mathbb{R}$. A number $\ell \in \mathbb{R}$ is a lower bound for A if

$$\forall a \in A, \quad \ell \leq a.$$

Definition (Bounded Above / Below)

A set $A \subseteq \mathbb{R}$ is said to be:

- bounded above if it has at least one upper bound;
- bounded below if it has at least one lower bound;
- bounded if it is both bounded above and bounded below.

Definition (Maximum)

Let $A \subseteq \mathbb{R}$. A number $m \in A$ is a maximum of A, written $m = \max A$, if

$$\forall a \in A, \quad a \leq m.$$

Definition (Minimum)

Let $A \subseteq \mathbb{R}$. A number $m \in A$ is a minimum of A, written $m = \min A$, if

$$\forall a \in A, \quad m \leq a.$$

Remark 11.8. Every maximum is an upper bound that is also a member of A; every minimum is a lower bound that is also a member of A. In particular, $\max A$ and $\min A$ need not exist, but when they do they are unique. The concepts of supremum and infimum generalise this: they are the least upper bound and greatest lower bound respectively, and need not belong to A. Both are unique when they exist (see Proposition 11.16).

Proposition 11.9 (Uniqueness of Supremum and Infimum). Let $A \subseteq \mathbb{R}$ be nonempty. If $\sup A$ exists, it is unique. If $\inf A$ exists, it is unique.

Definition (Supremum (Least Upper Bound))

Let $A \subseteq \mathbb{R}$ be nonempty and bounded above. A number $s \in \mathbb{R}$ is the supremum of A, written $s = \sup A$, if:

1. s is an upper bound for A, and
2. if u is any upper bound for A, then $s \leq u$.

Definition (Infimum (Greatest Lower Bound))

Let $A \subseteq \mathbb{R}$ be nonempty and bounded below. A number $s \in \mathbb{R}$ is the infimum of A, written $s = \inf A$, if:

1. s is a lower bound for A, and
2. if ℓ is any lower bound for A, then $\ell \leq s$.

11.1.3.2 Equivalent Formulations

The definition of supremum requires checking all upper bounds. The following proposition gives an equivalent condition that is often easier to apply in practice, replacing the universal quantifier over upper bounds with a local ε-witness condition.

Proposition 11.10 (ε-Characterization of Supremum and Infimum). Let $A \subseteq \mathbb{R}$ be nonempty. Both of the following hold.

1. If A is bounded above and $s \in \mathbb{R}$, then $s = \sup A$ if and only if both of the following conditions hold:
   (a) s is an upper bound for A, and
   (b) for every $\varepsilon > 0$, there exists $a \in A$ such that $s - \varepsilon < a$.
2. If A is bounded below and $s \in \mathbb{R}$, then $s = \inf A$ if and only if both of the following conditions hold:
   (a) s is a lower bound for A, and
   (b) for every $\varepsilon > 0$, there exists $a \in A$ such that $a < s + \varepsilon$.

Corollary 11.11 (ε-Approximation). Let $A \subseteq \mathbb{R}$ be nonempty.

1. If $s = \sup A$, then $\forall \varepsilon > 0,\ \exists a \in A$ such that $s - \varepsilon < a \leq s$.
2. If $s = \inf A$, then $\forall \varepsilon > 0,\ \exists a \in A$ such that $s \leq a < s + \varepsilon$.

11.1.3.3 Summary Table

Remark 11.12 (Predicate vs. Existence: Structural Template). Each bound-type concept is defined by a predicate $P_A(x)$ on $\mathbb{R}$.

- Predicate form. The statement "x is a P-object of A" means $P_A(x)$ holds for a specific identified x. To prove this, exhibit x and verify $P_A(x)$ directly.
- Existence form. The statement "A has a P-object" means $\exists x\, P_A(x)$. Once established, fix such an x and reason from $P_A(x)$.

Concept	Predicate $P_A(x)$	Predicate Form	Existence Form
Upper bound	$\forall a \in A,\ a \leq x$	x is an upper bound of A	A is bounded above
Lower bound	$\forall a \in A,\ a \geq x$	x is a lower bound of A	A is bounded below
Maximum	$x \in A \wedge \forall a \in A,\ a \leq x$	$x = \max A$	A has a maximum
Minimum	$x \in A \wedge \forall a \in A,\ a \geq x$	$x = \min A$	A has a minimum
Supremum	$(\forall a \in A,\ a \leq x) \wedge (\forall \varepsilon > 0,\ \exists a \in A,\ a > x - \varepsilon)$	$x = \sup A$	$\sup A$ exists
Infimum	$(\forall a \in A,\ a \geq x) \wedge (\forall \varepsilon > 0,\ \exists a \in A,\ a < x + \varepsilon)$	$x = \inf A$	$\inf A$ exists

Table 11.1: Rows grouped by logical complexity: bounds (universal only), extrema (membership + universal), least/greatest bounds (universal + approximation).

11.1.3.4 Consequences

Remark 11.13 (Logical Structure). The structural progression is:

$$\text{Upper/Lower bounds} \Rightarrow \text{Boundedness} \Rightarrow \text{Supremum/Infimum} \Rightarrow \text{Maximum/Minimum}.$$

All later extremal and limit arguments depend on this hierarchy.

11.1.4 Bounds and Extremal Values

Bounds Extremal Values Quick Reference

Core items	Key definitions/results introduced in this file.
How to use	Read the boxed items first; proofs and consequences follow.
Dependencies	Refer back to earlier sections as needed.

Remark 11.14 (Predicate vs. Existence: Structural Template). Each bound-type concept (maximum, minimum, upper bound, supremum, etc.) is defined by a predicate on $\mathbb{R}$. Given $A \subseteq \mathbb{R}$, let $P_A(x)$ denote such a predicate.

Predicate form. The statement "x is a P-object of A" means $P_A(x)$ holds for a specific, identified element $x \in \mathbb{R}$. To prove this, one exhibits the element and verifies each condition in $P_A(x)$.

Existence form. The statement "A has a P-object" means $\exists x\, P_A(x)$. Once established, one may fix such an x and reason about it (existential instantiation).

Notational glosses.

- "x is a P-object of A" abbreviates $P_A(x)$.
- "A has a P-object" abbreviates $\exists x\, P_A(x)$.

Proof strategy summary.

Goal	Form	Strategy
$P_A(x)$	Predicate	Exhibit x; verify $P_A(x)$ directly.
$\exists x\, P_A(x)$	Existence	Construct or name a candidate; verify $P_A(x)$.
Hypothesis: $\exists x\, P_A(x)$	—	Fix such an x; use $P_A(x)$ in further reasoning.

Table 11.2: *
The hypothesis row reflects existential instantiation, not a proof goal.

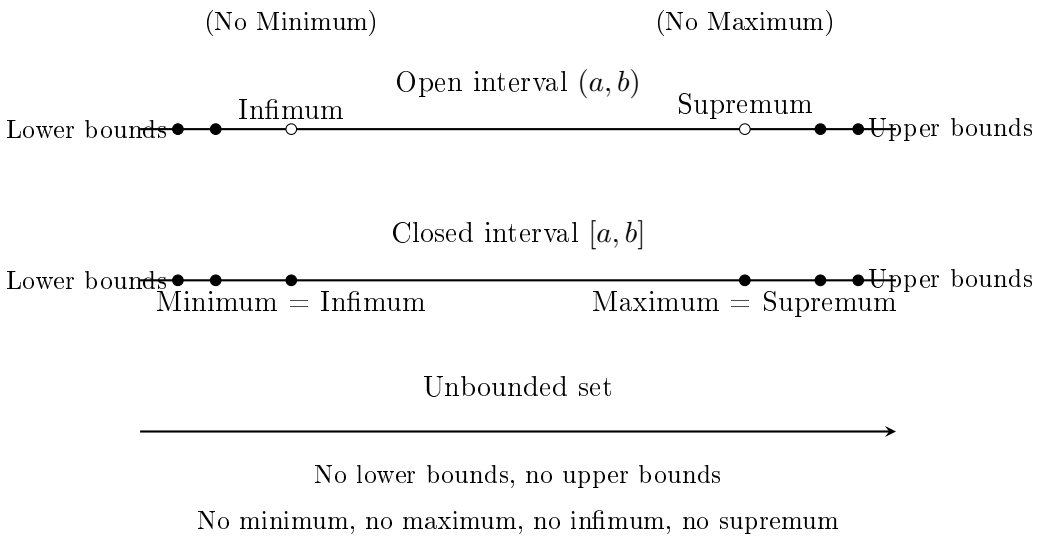


Figure 11.1: Bounds, Extrema, Infimum, and Supremum for Subsets of $\mathbb{R}$.

Concept	Predicate $P_A(x)$	Predicate Form	Existence Form
Upper bound	$\forall a \in A,\ a \leq x$	x is an upper bound of A	A is bounded above
Lower bound	$\forall a \in A,\ a \geq x$	x is a lower bound of A	A is bounded below
Maximum	$x \in A \ \wedge\ \forall a \in A,\ a \leq x$	$x = \max A$	A has a maximum
Minimum	$x \in A \ \wedge\ \forall a \in A,\ a \geq x$	$x = \min A$	A has a minimum
Supremum	$(\forall a \in A,\ a \leq x) \wedge (\forall \varepsilon > 0,\ \exists a \in A,\ a > x - \varepsilon)$	$x = \sup A$	$\sup A$ exists
Infimum	$(\forall a \in A,\ a \geq x) \wedge (\forall \varepsilon > 0,\ \exists a \in A,\ a < x + \varepsilon)$	$x = \inf A$	$\inf A$ exists

Table 11.3: *

Rows are grouped by logical complexity: bounds (universal only), extrema (membership + universal), least/greatest bounds (universal + approximation). The predicate column gives the condition that must hold for a specific $x \in \mathbb{R}$; the existence form asserts $\exists x\ P_A(x)$.

Definition (Upper Bound)

Let $A \subseteq \mathbb{R}$. A number $u \in \mathbb{R}$ is an upper bound for A if

$$\forall a \in A, \quad a \leq u.$$

Definition (Lower Bound)

Let $A \subseteq \mathbb{R}$. A number $\ell \in \mathbb{R}$ is a lower bound for A if

$$\forall a \in A, \quad \ell \leq a.$$

Definition (Bounded Above / Below)

A set $A \subseteq \mathbb{R}$ is said to be:

- bounded above if it has at least one upper bound;
- bounded below if it has at least one lower bound;
- bounded if it is both bounded above and bounded below.

Definition (Maximum)

Let $A \subseteq \mathbb{R}$. A number $m \in A$ is a maximum of A, written $m = \max A$, if

$$\forall a \in A, \quad a \leq m.$$

Definition (Minimum)

Let $A \subseteq \mathbb{R}$. A number $m \in A$ is a minimum of A, written $m = \min A$, if

$$\forall a \in A, \quad m \leq a.$$

Remark 11.15. Every maximum is an upper bound that is also a member of A; every minimum is a lower bound that is also a member of A. In particular, $\max A$ and $\min A$ need not exist, but when they do they are unique. The concepts of supremum and infimum generalise this: they are the least upper bound and greatest lower bound respectively, and need not belong to A. Both are unique when they exist (see Proposition 11.16).

Proposition 11.16 (Uniqueness of Supremum and Infimum). Let $A \subseteq \mathbb{R}$ be nonempty. If $\sup A$ exists, it is unique. If $\inf A$ exists, it is unique.

Definition (Supremum (Least Upper Bound))

Let $A \subseteq \mathbb{R}$ be nonempty and bounded above. A number $s \in \mathbb{R}$ is the supremum of A, written $s = \sup A$, if:

1. s is an upper bound for A, and
2. if u is any upper bound for A, then $s \leq u$.

Definition (Infimum (Greatest Lower Bound))

Let $A \subseteq \mathbb{R}$ be nonempty and bounded below. A number $s \in \mathbb{R}$ is the infimum of A, written $s = \inf A$, if:

1. s is a lower bound for A, and
2. if ℓ is any lower bound for A, then $\ell \leq s$.

11.1.4.1 Equivalent Formulations

The definition of supremum requires checking all upper bounds. The following proposition gives an equivalent condition that is often easier to apply in practice, replacing the universal quantifier over upper bounds with a local ε-witness condition.

Proposition 11.17 (ε-Characterization of Supremum and Infimum). Let $A \subseteq \mathbb{R}$ be nonempty. Both of the following hold.

1. If A is bounded above and $s \in \mathbb{R}$, then $s = \sup A$ if and only if both of the following conditions hold:

(a) s is an upper bound for A, and

(b) for every $\varepsilon > 0$, there exists $a \in A$ such that $s - \varepsilon < a$.

2. If A is bounded below and $s \in \mathbb{R}$, then $s = \inf A$ if and only if both of the following conditions hold:

(a) s is a lower bound for A, and

(b) for every $\varepsilon > 0$, there exists $a \in A$ such that $a < s + \varepsilon$.

Corollary 11.18 (ε-Approximation). Let $A \subseteq \mathbb{R}$ be nonempty.

1. If $s = \sup A$, then $\forall \varepsilon > 0,\ \exists a \in A$ such that $s - \varepsilon < a \leq s$.
2. If $s = \inf A$, then $\forall \varepsilon > 0,\ \exists a \in A$ such that $s \leq a < s + \varepsilon$.

11.1.5 Completeness of the Real Numbers

Completeness Quick Reference	
Core items	Key definitions/results introduced in this file.
How to use	Read the boxed items first; proofs and consequences follow.
Dependencies	Refer back to earlier sections as needed.

11.1.5.1 Completeness Axiom

Remark 11.19 (Why completeness is needed). The ordered field axioms alone do not prevent "holes" in the number line. The rationals $\mathbb{Q}$ satisfy all field and order axioms, yet fail completeness. Consider the set

$$S = \{x \in \mathbb{Q} : x^2 < 2\}.$$

This set is nonempty (e.g. $1 \in S$) and bounded above in $\mathbb{Q}$ (e.g. 2 is an upper bound). Yet $\sup S$ does not exist as a rational number: the candidate $\sqrt{2}$ is irrational. The set S has no least upper bound in $\mathbb{Q}$ — a hole sits exactly where $\sqrt{2}$ should be.

The Completeness Axiom asserts that $\mathbb{R}$ has no such holes: every nonempty bounded-above set has a supremum inside $\mathbb{R}$. This single axiom is what distinguishes $\mathbb{R}$ from $\mathbb{Q}$, and it underlies every major theorem in analysis.

Axiom (Completeness Axiom (Least Upper Bound Property))

Every nonempty subset of $\mathbb{R}$ that is bounded above has a supremum in $\mathbb{R}$.
Equivalently: if $S \subseteq \mathbb{R}$ is nonempty and bounded above, then $\sup S$ exists as a real number.

Remark 11.20 (Logical form).

$$\forall S \left((S \subseteq \mathbb{R} \wedge S \neq \varnothing \wedge \exists M \in \mathbb{R}\ \forall x \in S\ (x \leq M)) \to \exists s \in \mathbb{R}\ (s = \sup S) \right).$$

Remark 11.21 (Equivalent formulations). Over the ordered-field axioms, completeness is equivalent to each of the following:

- Every nonempty set bounded below has an infimum.
- Every Cauchy sequence in $\mathbb{R}$ converges.
- (Nested Interval Property) Every nested sequence of nonempty closed bounded intervals has nonempty intersection.

11.1.5.2 Nested Interval Property

Theorem 11.22 (Nested Interval Property). Let $\{[a_n, b_n]\}_{n\in\mathbb{N}}$ be nonempty closed intervals such that

$$[a_{n+1}, b_{n+1}] \subseteq [a_n, b_n] \quad \text{for all } n.$$

Then

$$\bigcap_{n=1}^{\infty} [a_n, b_n] \neq \varnothing.$$

If additionally $b_n - a_n \to 0$, the intersection consists of exactly one point.

11.1.5.3 Archimedean Property

Definition

$\mathbb{R}$ satisfies the Archimedean property if

$$\forall x \in \mathbb{R}\ \exists n \in \mathbb{N}\ (n > x).$$

Equivalently:

$$\forall x > 0\ \forall y \in \mathbb{R}\ \exists n \in \mathbb{N}\ (nx > y).$$

Remark 11.23 (Logical form).

$$\forall x\ \exists n\ (n > x). \qquad \forall x > 0\ \forall y\ \exists n\ (nx > y).$$

Theorem 11.24 (Archimedean Property). $\mathbb{R}$ satisfies the Archimedean property.

Corollary 11.25 (Archimedean Property Corollary). If $x > 0$ and $y \in \mathbb{R}$, then $\exists n \in \mathbb{N}$ such that $nx > y$.

11.1.5.4 Integer Part

Lemma 11.26 (Floor lemma). For every $x \in \mathbb{R}$ there exists a unique $m \in \mathbb{Z}$ such that

$$m \leq x < m + 1.$$

11.1.5.5 Density

Definition

A subset A of a linearly ordered set X is dense in X if

$$\forall a < b \, \exists c \in A \, (a < c < b).$$

Remark 11.27 (Logical form for $\mathbb{Q}$ dense in $\mathbb{R}$).

$$\forall a < b \, \exists q \in \mathbb{Q} \, (a < q < b).$$

Theorem 11.28 (Density of $\mathbb{Q}$). $\mathbb{Q}$ is dense in $\mathbb{R}$.

Corollary 11.29 (Irrationals Are Dense). Between any two distinct real numbers lies an irrational.

Corollary 11.30 (Density of Irrationals). The irrationals are dense in $\mathbb{R}$.

11.1.5.6 Existence of Square Roots

Theorem 11.31 (Existence of Square Roots). For every $a \geq 0$ there exists a unique $x \geq 0$ such that $x^2 = a$.

Remark 11.32 (Logical form).

$$\forall a \geq 0 \, \exists! \, x \geq 0 \, (x^2 = a).$$

Structural Summary

$$\text{Field Axioms} \Rightarrow \text{Order Axioms} \Rightarrow \text{Completeness Axiom}$$

Completeness yields, in logical order:

$$\text{Nested Interval Property} \Rightarrow \text{Archimedean Property} \Rightarrow \text{Floor Lemma} \Rightarrow \text{Density of } \mathbb{Q} \Rightarrow \text{Existence of } \sqrt{a}$$

Completeness is the property that prevents holes in $\mathbb{R}$. It is equivalent to the Cauchy Criterion and underlies every major limit theorem in real analysis.

11.1.6 Dedekind Cut Construction of $\mathbb{R}$

Dedekind Cuts Quick Reference

Concept	Meaning	Detail
Dedekind cut	Partition $(A \mid B)$ of $\mathbb{Q}$	Def
Real number	An equivalence class of Dedekind cuts	Def
Completeness	Every non-empty bounded cut has a sup	Thm

Definition (Dedekind Cut)

A Dedekind cut is a pair (A, B) with $A \cup B = \mathbb{Q}$, $A \cap B = \emptyset$, $A \neq \emptyset$, $B \neq \emptyset$, satisfying:

1. If $p \in A$ and $q < p$ then $q \in A$ (downward closed),
2. A has no greatest element.

Remark 11.33 (English reading). A Dedekind cut slices the rationals into a lower set A and upper set B so that A contains all rationals "below" the cut point and has no maximum. Each cut corresponds to exactly one real number.

Definition (Real Number via Dedekind Cut)

A real number is a Dedekind cut (A, B) of $\mathbb{Q}$. The set $\mathbb{R}$ is defined as the collection of all Dedekind cuts, equipped with order: $(A_1, B_1) \leq (A_2, B_2)$ iff $A_1 \subseteq A_2$.

Remark 11.34 (Status). This section is a stub. Full content arithmetic on cuts, the proof that $\mathbb{R}$ is a complete ordered field, and the uniqueness theorem will be developed in a future revision. Primary source: Rudin, Principles of Mathematical Analysis, Appendix to Chapter 1.

Theorem (Dedekind Completeness)

The ordered field $(\mathbb{R}, +, \cdot, \leq)$ constructed from Dedekind cuts is complete: every non-empty subset of $\mathbb{R}$ that is bounded above has a least upper bound in $\mathbb{R}$.

Remark 11.35 (Significance). This is the point of the entire construction: $\mathbb{Q}$ has "gaps" (e.g. no rational satisfies $x^2 = 2$), and Dedekind cuts fill exactly those gaps, yielding a complete ordered field.

11.1.7 Cauchy Sequence Construction of $\mathbb{R}$

Cauchy Completion Quick Reference

Concept	Meaning	Detail
Cauchy sequence over $\mathbb{Q}$	Sequence with terms arbitrarily close	Def
Equivalence of Cauchy seqs	Same limit behaviour	Def
Real number as equivalence class	$\mathbb{R} = \mathbb{Q}^{\text{Cauchy}}/\sim$	Def

Definition (Cauchy Sequence over $\mathbb{Q}$)

A sequence (a_n) of rationals is Cauchy if

$$\forall \varepsilon \in \mathbb{Q}_{>0},\ \exists N \in \mathbb{N},\ \forall m, n \geq N, \quad |a_m - a_n| < \varepsilon.$$

Definition (Equivalence of Cauchy Sequences)

Two Cauchy sequences (a_n) and (b_n) over $\mathbb{Q}$ are equivalent, written $(a_n) \sim (b_n)$, if

$$\lim_{n \to \infty} (a_n - b_n) = 0 \quad (\text{in } \mathbb{Q}).$$

Definition (Real Number via Cauchy Sequences)

A real number is an equivalence class $[(a_n)]$ of Cauchy sequences of rationals under $\sim$. The set $\mathbb{R}$ is defined as the quotient $\mathbb{R} = \mathcal{C}(\mathbb{Q})/\sim$, where $\mathcal{C}(\mathbb{Q})$ denotes the set of all Cauchy sequences of rationals.

Remark 11.36 (Status). This section is a stub. Full content arithmetic on equivalence classes, verification of field axioms, proof of completeness, and comparison with the Dedekind construction will be developed in a future revision. Primary source: Tao, Analysis I, Chapters 5–6.

Remark 11.37 (Comparison with Dedekind cuts). Both constructions yield the same object up to isomorphism. The Dedekind approach is more geometric (cutting the line); the Cauchy approach is more analytic (completing via limits). Tao develops the Cauchy approach; Rudin the Dedekind approach.

11.1.8 Interval Arithmetic

Interval Arithmetic Quick Reference

Operation	Rule	Detail
Addition	$[a,b]+[c,d]=[a+c,\ b+d]$	Def
Subtraction	$[a,b]-[c,d]=[a-d,\ b-c]$	Def
Multiplication	$[a,b]\cdot[c,d]$ (case analysis)	Def
Containment	$[a,b]\subseteq[c,d] \iff c\le a, b\le d$	Def

Definition (Interval Addition)

For closed intervals $[a,b]$ and $[c,d]$, their sum is

$$[a,b]+[c,d] \;=\; [a+c,\ b+d].$$

Definition (Interval Subtraction)

For closed intervals $[a,b]$ and $[c,d]$, their difference is

$$[a,b]-[c,d] \;=\; [a-d,\ b-c].$$

Definition (Interval Multiplication)

For closed intervals $[a,b]$ and $[c,d]$, their product is

$$[a,b]\cdot[c,d] \;=\; \big[\min(ac,ad,bc,bd),\ \max(ac,ad,bc,bd)\big].$$

Definition (Interval Containment)

$[a,b]\subseteq[c,d]$ if and only if $c\le a$ and $b\le d$.

Remark 11.38 (Motivation). Interval arithmetic provides a framework for rigorous numerical computation with guaranteed error bounds. A real number $x\in[a,b]$ means the computed value may lie

anywhere in the interval; operations propagate these bounds.

Remark 11.39 (Status). This section is a stub. A full treatment the sub-distributivity law, dependency problem, applications to root-finding, and connections to compactness will be developed in a future revision.

11.5 Flashcards — Real Line Foundations (Avery 5388)

Axiom of Completeness

Archimedean Property

Upper Bound (Set)

Lower Bound (Set)

Bounded (Set)

Maximum

Minimum

Supremum (LUB)

ε-Characterization of Supremum

Infimum (GLB)

Sequence

Bounded Sequence

NOT Bounded (Sequence)

Convergence

NOT Convergent to L

Cauchy Sequence

NOT Cauchy

Subsequence

Index Growth ($n_k \geq k$)

Subsequential Limit

Monotone Subsequence Theorem

Convergence via Even/Odd Subsequences

Subsequence of a Subsequence

Tail Set

$s_n = \sup_{k \geq n} a_k$

$\limsup\ a_n$ (definition)

$i_n = \inf_{k \geq n} a_k$

$\liminf\ a_n$ (definition)

Uniqueness of Limits

Convergent $\Rightarrow$ Bounded

Subsequence Inherits Limit

Bolzano–Weierstrass

Convergent $\Rightarrow$ Cauchy

Cauchy $\Rightarrow$ Bounded

Cauchy $\Leftrightarrow$ Convergent (in $\mathbb{R}$)

$\liminf a_n \leq \limsup a_n$

Convergence $\Leftrightarrow \liminf = \limsup$

Monotone Convergence Theorem

Squeeze Theorem

Divergence Criterion

Monotone Sequence (definitions)

Order Limit Theorem

Algebra of Limits

lim sup Characterization

lim inf Characterization

Convergence is a Tail Property

Finite Modification Theorem

Proof Sketch: Convergent $\Rightarrow$ Bounded

Proof Sketch: Uniqueness of Limits

Proof Sketch: Cauchy $\Rightarrow$ Bounded

Proof Sketch: Bolzano–Weierstrass

Proof Sketch: Convergence $\Leftrightarrow \liminf = \limsup$

Plain English. A real number u is an upper bound of a set $A \subseteq \mathbb{R}$ if every element of A is $\leq u$.

Fully quantified.

$$\forall A \subseteq \mathbb{R}\ \forall u \in \mathbb{R}\ \big(u \text{ is an upper bound of } A \iff \forall a \in A\ (a \leq u)\big).$$

Plain English. $\mathbb{N}$ is unbounded in $\mathbb{R}$: no real number is an upper bound for $\mathbb{N}$. Equivalently, $1/n$ can be made smaller than any positive real.

Fully quantified (two equivalent forms).

$$\text{(i) } \forall x \in \mathbb{R}\ \exists n \in \mathbb{N}\ (n > x).$$

$$\text{(ii) } \forall \varepsilon > 0\ \exists n \in \mathbb{N}\ \left(\tfrac{1}{n} < \varepsilon\right).$$

Proof mechanism. Follows from the Axiom of Completeness: if $\mathbb{N}$ were bounded above, $s = \sup \mathbb{N}$ would exist, but $s - 1$ cannot be an upper bound, giving $n > s - 1$, so $n + 1 > s$ — contradiction.

Key use. Form (ii) justifies the "$\varepsilon = 1/n$" move in proofs.

Plain English. Every nonempty subset of $\mathbb{R}$ that is bounded above has a least upper bound in $\mathbb{R}$. This is what separates $\mathbb{R}$ from $\mathbb{Q}$.

Fully quantified.

$$\forall A \subseteq \mathbb{R},\ \Big(A \neq \emptyset\ \wedge\ \exists u \in \mathbb{R}\ \forall a \in A\ (a \leq u)\Big)$$

$$\to\ \exists s \in \mathbb{R}\ \Big[(\forall a \in A,\ a \leq s)\ \wedge$$

$$\forall u \in \mathbb{R}\big((\forall a \in A,\ a \leq u) \to s \leq u\big)\Big].$$

Why it matters. BW, MCT, Cauchy⇔Convergent all depend on this. It fails in $\mathbb{Q}$: $\{q \in \mathbb{Q} : q^2 < 2\}$ has no supremum in $\mathbb{Q}$.

Plain English. m is the maximum of A if $m \in A$ and every $a \in A$ satisfies $a \leq m$.

Fully quantified.

$$\forall A \subseteq \mathbb{R}\ \forall m \in \mathbb{R}\ \Big(m = \max A \iff (m \in A) \wedge \forall a \in A\ (a \leq m)\Big).$$

Plain English. A set A is bounded if it has both an upper bound and a lower bound.

Fully quantified.

$$\begin{aligned}\forall A \subseteq \mathbb{R}\ \Big(&A \text{ bounded} \iff \\ &\exists u \in \mathbb{R}\ \forall a \in A\ (a \leq u) \\ &\wedge\ \exists \ell \in \mathbb{R}\ \forall a \in A\ (\ell \leq a)\Big).\end{aligned}$$

Plain English. A real number ℓ is a lower bound of a set $A \subseteq \mathbb{R}$ if every element of A is $\geq \ell$.

Fully quantified.

$$\forall A \subseteq \mathbb{R}\ \forall \ell \in \mathbb{R}\ \big(\ell \text{ is a lower bound of } A \iff \forall a \in A\ (\ell \leq a)\big).$$

Plain English. $s = \sup A$ can always be approximated from inside A: no matter how small $\varepsilon > 0$ is, some element of A lies within ε of s.

Fully quantified.

$$s = \sup A \;\Rightarrow\; \forall \varepsilon > 0\ \exists a \in A,\ s - \varepsilon < a \leq s.$$

Equivalently: $s - \varepsilon$ is not an upper bound of A for any $\varepsilon > 0$.

Symmetric form for infimum.

$$t = \inf A \;\Rightarrow\; \forall \varepsilon > 0\ \exists a \in A,\ t \leq a < t + \varepsilon.$$

Key use. This is the engine of nearly every completeness proof: MCT, BW (bisection), and Cauchy⇔Convergent all invoke it.

Plain English. $s = \sup A$ if s is an upper bound of A, and every upper bound u satisfies $s \leq u$.

Fully quantified.

$$\forall A \subseteq \mathbb{R}\ \forall s \in \mathbb{R}\ \Big(s = \sup A \iff (\forall a \in A\ (a \leq s)) \wedge \forall u \in \mathbb{R}\big[(\forall a \in A\ (a \leq u)) \to s \leq u\big]\Big).$$

ε-approx move (used in proofs).

$$s = \sup A \;\Rightarrow\; \forall \varepsilon > 0\ \exists a \in A\ (s - \varepsilon < a).$$

Plain English. m is the minimum of A if $m \in A$ and every $a \in A$ satisfies $m \leq a$.

Fully quantified.

$$\forall A \subseteq \mathbb{R}\ \forall m \in \mathbb{R}\ \Big(m = \min A \iff (m \in A) \wedge \forall a \in A\ (m \leq a)\Big).$$

Plain English. A sequence is bounded if all its terms stay within some fixed distance from 0.

Fully quantified.

$$\forall a : \mathbb{N} \to \mathbb{R} \left(a \text{ bounded } \iff \exists M > 0 \; \forall n \in \mathbb{N}, \; |a_n| \leq M \right).$$

Plain English. A sequence in $\mathbb{R}$ is a function from $\mathbb{N}$ to $\mathbb{R}$.

Fully quantified.

$$\forall a \left(a \text{ is a real sequence } \iff a : \mathbb{N} \to \mathbb{R} \right).$$

Plain English. $t = \inf A$ if t is a lower bound of A, and every lower bound ℓ satisfies $\ell \leq t$.

Fully quantified.

$$\begin{aligned} \forall A \subseteq \mathbb{R} \; \forall t \in \mathbb{R} \Big(& t = \inf A \iff \\ & (\forall a \in A \; (t \leq a)) \wedge \\ & \forall \ell \in \mathbb{R} \big[(\forall a \in A \; (\ell \leq a)) \to \ell \leq t \big] \Big). \end{aligned}$$

ε-approx move (used in proofs).

$$t = \inf A \; \Rightarrow \; \forall \varepsilon > 0 \; \exists a \in A \; (a < t + \varepsilon).$$

Plain English. a_n does not converge to L if some tolerance $\varepsilon > 0$ is violated infinitely often.

Fully quantified (negation).

$$\forall a : \mathbb{N} \to \mathbb{R}\ \forall L \in \mathbb{R}$$
$$\Big(\neg(a_n \to L) \iff \exists \varepsilon > 0\ \forall N \in \mathbb{N}\ \exists n \geq N,\ |a_n - L| \geq \varepsilon\Big).$$

Plain English. $a_n \to L$ means the terms eventually get arbitrarily close to L and stay close thereafter.

Fully quantified.

$$\forall a : \mathbb{N} \to \mathbb{R}\ \forall L \in \mathbb{R}$$
$$\Big(a_n \to L \iff \forall \varepsilon > 0\ \exists N \in \mathbb{N}\ \forall n \geq N,\ |a_n - L| < \varepsilon\Big).$$

Plain English. A sequence is not bounded if every proposed bound is exceeded in absolute value by some term.

Fully quantified (negation).

$$\forall a : \mathbb{N} \to \mathbb{R}\ \Big(\neg(a \text{ bounded}) \iff \forall M > 0\ \exists n \in \mathbb{N},\ |a_n| > M\Big).$$

Plain English. A subsequence is formed by selecting terms using a strictly increasing index map.

Fully quantified.

$$\forall a : \mathbb{N} \to \mathbb{R}\ \forall b : \mathbb{N} \to \mathbb{R}$$
$$\Big(b \text{ is a subseq. of } a \iff \exists \sigma : \mathbb{N} \to \mathbb{N}$$
$$\big(\forall k < \ell,\ \sigma(k) < \sigma(\ell)\big) \wedge \big(\forall k,\ b_k = a_{\sigma(k)}\big)\Big).$$

Plain English. Not Cauchy means there is some $\varepsilon > 0$ such that, no matter how far out you go, two later terms are at least ε apart.

Fully quantified (negation).

$$\forall a : \mathbb{N} \to \mathbb{R}\ \Big(\neg(a \text{ Cauchy}) \iff$$
$$\exists \varepsilon > 0\ \forall N \in \mathbb{N}\ \exists m, n \geq N,\ |a_m - a_n| \geq \varepsilon\Big).$$

Plain English. A sequence is Cauchy if its terms eventually get arbitrarily close to each other.

Fully quantified.

$$\forall a : \mathbb{N} \to \mathbb{R}\ \Big(a \text{ Cauchy} \iff$$
$$\forall \varepsilon > 0\ \exists N \in \mathbb{N}\ \forall m, n \geq N,\ |a_m - a_n| < \varepsilon\Big).$$

Plain English. Every real sequence has a monotone subsequence.

Fully quantified.

$$\forall a : \mathbb{N} \to \mathbb{R}\ \exists b \text{ subseq. of } a$$
$$\Big(\forall k,\ b_k \le b_{k+1}\Big) \vee \Big(\forall k,\ b_{k+1} \le b_k\Big).$$

Proof idea (peak indices). Call n a peak if $a_n \ge a_m$ for all $m > n$. If infinitely many peaks exist, they form a decreasing subsequence. If only finitely many, past the last peak every index can be extended to a strictly increasing subsequence.

Why it matters. Gives an alternative proof of BW: every bounded sequence has a monotone subseq., which by MCT converges.

Plain English. L is a subsequential limit of (a_n) if some subsequence of (a_n) converges to L. The set of all subsequential limits is written $\mathcal{L}(a_n)$.

Fully quantified.

$$L \text{ is a subsequential limit of } (a_n) \iff$$
$$\exists \sigma : \mathbb{N} \to \mathbb{N}\ \big(\forall k < \ell,\ \sigma(k) < \sigma(\ell)\big)$$
$$\wedge\ \forall \varepsilon > 0\ \exists K\ \forall k \ge K,\ |a_{\sigma(k)} - L| < \varepsilon.$$

Key fact. If $a_n \to L$ then $\mathcal{L}(a_n) = \{L\}$ a convergent sequence has exactly one subsequential limit.

Plain English. If (n_k) is strictly increasing in $\mathbb{N}$, then $n_k \ge k$ for all k.

Fully quantified.

$$\forall (n_k) : \mathbb{N} \to \mathbb{N}$$
$$\Big(\forall k < \ell,\ n_k < n_\ell\Big) \to \Big(\forall k \in \mathbb{N},\ n_k \ge k\Big).$$

Proof. Induction: $n_1 \ge 1$; if $n_k \ge k$ then $n_{k+1} \ge n_k + 1 \ge k + 1$.

Plain English. The n-tail of (a_n) is the set of all terms from index n onward.

Fully quantified.

$$\forall a : \mathbb{N} \to \mathbb{R}\ \forall n \in \mathbb{N},$$
$$T_n(a) = \{a_k : k \in \mathbb{N} \wedge k \geq n\}.$$

Plain English. A subsequence of a subsequence is itself a subsequence of the original sequence.

Fully quantified.

$$\forall a, b, c : \mathbb{N} \to \mathbb{R},$$
$$\big(b \text{ subseq. of } a\big) \wedge \big(c \text{ subseq. of } b\big) \to \big(c \text{ subseq. of } a\big).$$

Proof. If $b_k = a_{n_k}$ with (n_k) strictly increasing, and $c_j = b_{k_j}$ with (k_j) strictly increasing, then $c_j = a_{n_{k_j}}$ and (n_{k_j}) is strictly increasing (composition of strictly increasing maps).

Key use. Applying BW twice: a bounded sequence has a convergent subsequence; any subsequence of that is also a subsequence of the original.

Plain English. If the even-indexed and odd-indexed subsequences both converge to the same limit, the full sequence converges to that limit.

Fully quantified.

$$\forall a : \mathbb{N} \to \mathbb{R}\ \forall L \in \mathbb{R},$$
$$\big(a_{2n} \to L \ \wedge\ a_{2n+1} \to L\big) \to a_n \to L.$$

Special case of the Partition Convergence Principle: if $\mathbb{N}$ is split into finitely many infinite parts and the subsequence on each part converges to the same L, then $a_n \to L$.

Plain English. i_n is the infimum of the n-tail: the greatest lower bound of all terms from index n onward.

Fully quantified.

$$\forall a : \mathbb{N} \to \mathbb{R}\ \forall n \in \mathbb{N},$$
$$i_n = \inf\{a_k : k \in \mathbb{N} \wedge k \geq n\}.$$

Note: (i_n) is increasing (larger n = fewer terms to inf over).

Plain English. $\limsup a_n$ is the limit of the tail suprema; it captures the eventual upper oscillation level.

Definition via s_n.

$$\text{Let } s_n = \sup\{a_k : k \geq n\}.$$
$$\limsup_{n\to\infty} a_n := \lim_{n\to\infty} s_n \quad \text{(extended reals)}.$$

(s_n) is decreasing, so the limit always exists in $[-\infty, +\infty]$.

Plain English. s_n is the supremum of the n-tail: the least upper bound of all terms from index n onward.

Fully quantified.

$$\forall a : \mathbb{N} \to \mathbb{R}\ \forall n \in \mathbb{N},$$
$$s_n = \sup\{a_k : k \in \mathbb{N} \wedge k \geq n\}.$$

Note: (s_n) is decreasing (larger n = fewer terms to sup over).

Plain English. Every convergent real sequence is bounded.

Fully quantified.

$$\forall a : \mathbb{N} \to \mathbb{R}\ \forall L \in \mathbb{R},$$
$$(a_n \to L) \to \exists M > 0\ \forall n \in \mathbb{N},\ |a_n| \leq M.$$

Plain English. A sequence cannot converge to two different real numbers.

Fully quantified.

$$\forall a : \mathbb{N} \to \mathbb{R}\ \forall L, M \in \mathbb{R},$$
$$\big(a_n \to L\ \wedge\ a_n \to M\big) \to L = M.$$

Plain English. $\liminf a_n$ is the limit of the tail infima; it captures the eventual lower oscillation level.

Definition via i_n.

$$\text{Let } i_n = \inf\{a_k : k \geq n\}.$$
$$\liminf_{n\to\infty} a_n := \lim_{n\to\infty} i_n \quad \text{(extended reals)}.$$

(i_n) is increasing, so the limit always exists in $[-\infty, +\infty]$.

Plain English. Every convergent sequence is Cauchy.

Fully quantified.

$$\forall a : \mathbb{N} \to \mathbb{R}\ \forall L \in \mathbb{R},$$
$$(a_n \to L) \to (a \text{ is Cauchy}).$$

Plain English. Every bounded real sequence has a convergent subsequence.

Fully quantified.

$$\forall a : \mathbb{N} \to \mathbb{R},$$
$$\Big(\exists M > 0\ \forall n,\ |a_n| \leq M\Big)$$
$$\to \exists b : \mathbb{N} \to \mathbb{R}\ \exists L \in \mathbb{R}$$
$$\big(b \text{ is a subsequence of } a\ \wedge\ b_n \to L\big).$$

Plain English. If a sequence converges, every subsequence converges to the same limit.

Fully quantified.

$$\forall a : \mathbb{N} \to \mathbb{R}\ \forall L \in \mathbb{R}\ \forall b : \mathbb{N} \to \mathbb{R},$$
$$\big(a_n \to L \wedge b \text{ is a subsequence of } a\big) \to (b_n \to L).$$

Plain English. The eventual bottom level is never above the eventual top level.

Fully quantified.

$$\forall a : \mathbb{N} \to \mathbb{R},$$
$$\liminf_{n\to\infty} a_n \leq \limsup_{n\to\infty} a_n.$$

Proof idea. Since $i_n \leq s_n$ for all n, taking limits preserves the inequality.

Plain English. In $\mathbb{R}$, a sequence converges if and only if it is Cauchy. This is completeness.

Fully quantified.

$$\forall a : \mathbb{N} \to \mathbb{R},$$
$$(a \text{ Cauchy}) \leftrightarrow \exists L \in \mathbb{R}\ (a_n \to L).$$

Plain English. Every Cauchy sequence is bounded.

Fully quantified.

$$\forall a : \mathbb{N} \to \mathbb{R},$$
$$(a \text{ Cauchy}) \to \exists M > 0\ \forall n \in \mathbb{N},\ |a_n| \leq M.$$

Plain English. If a sequence is trapped between two sequences that share the same limit, it must converge to that same limit.

Fully quantified.

$$\begin{aligned}&\forall a, b, c : \mathbb{N} \to \mathbb{R}\ \forall L \in \mathbb{R},\\&\Big(\forall n,\ a_n \le b_n \le c_n\Big) \wedge (a_n \to L) \wedge (c_n \to L)\\&\quad \to\ b_n \to L.\end{aligned}$$

Proof idea. For $n \ge \max(N_a, N_c)$: $L - \varepsilon < a_n \le b_n \le c_n < L + \varepsilon$.

Plain English. A monotone bounded sequence converges. The limit is the supremum (increasing case) or infimum (decreasing case) of its range.

Fully quantified.

$$\begin{aligned}&\forall a : \mathbb{N} \to \mathbb{R},\\&\Big(\forall n,\ a_n \le a_{n+1}\Big) \wedge \Big(\exists M\ \forall n,\ a_n \le M\Big)\\&\quad \to\ a_n \to \sup\{a_n : n \in \mathbb{N}\}.\\&\Big(\forall n,\ a_{n+1} \le a_n\Big) \wedge \Big(\exists M\ \forall n,\ M \le a_n\Big)\\&\quad \to\ a_n \to \inf\{a_n : n \in \mathbb{N}\}.\end{aligned}$$

Why it matters. Equivalent to BW, Cauchy⇔Convergent, and the Nested Interval Property all are manifestations of completeness.

Plain English. A sequence converges if and only if liminf and limsup agree; then they equal the limit.

Fully quantified.

$$\begin{aligned}&\forall a : \mathbb{N} \to \mathbb{R},\\&\Big(\exists L \in \mathbb{R}\ (a_n \to L)\Big)\\&\quad \leftrightarrow \Big(\liminf_{n\to\infty} a_n = \limsup_{n\to\infty} a_n \in \mathbb{R}\Big).\end{aligned}$$

Plain English. If $x_n \le y_n$ for all n and both sequences converge, then their limits preserve the inequality.

Fully quantified.

$$\begin{aligned}&\forall (x_n), (y_n) : \mathbb{N} \to \mathbb{R}\ \forall x, y \in \mathbb{R},\\ &\Big(\forall n,\ x_n \le y_n\Big) \wedge (x_n \to x) \wedge (y_n \to y)\\ &\quad \to x \le y.\end{aligned}$$

Warning. Strict inequality $x_n < y_n$ does not guarantee $x < y$ (e.g. $x_n = 0 < 1/n = y_n$, but both $\to 0$).

Why it matters. Foundation for the Squeeze Theorem and most order-based convergence arguments.

Plain English. A sequence is increasing if each term is $\le$ the next; strictly increasing if each term is $<$ the next. Decreasing and strictly decreasing are symmetric. A sequence is monotone if it is either increasing or decreasing.

Fully quantified.

$$\begin{aligned}&(a_n) \text{ increasing} \iff \forall n,\ a_n \le a_{n+1}.\\ &(a_n) \text{ strictly increasing} \iff \forall n,\ a_n < a_{n+1}.\\ &(a_n) \text{ decreasing} \iff \forall n,\ a_{n+1} \le a_n.\\ &(a_n) \text{ strictly decreasing} \iff \forall n,\ a_{n+1} < a_n.\\ &(a_n) \text{ monotone} \iff (a_n) \text{ increasing or decreasing.}\end{aligned}$$

Plain English. A sequence diverges if and only if it is unbounded or it has two subsequences converging to different limits.

Fully quantified.

$$\begin{aligned}&\forall a : \mathbb{N} \to \mathbb{R},\\ &\neg(\exists L \in \mathbb{R},\ a_n \to L) \iff\\ &\quad \Big(\forall M > 0\ \exists n,\ |a_n| > M\Big)\\ &\quad \vee\ \Big(\exists b, c \text{ subseqs of } a,\ \exists L \ne L',\\ &\qquad b_n \to L\ \wedge\ c_n \to L'\Big).\end{aligned}$$

Practical use. To show (a_n) diverges: exhibit two subsequences with different limits (e.g. even/odd indices for $(-1)^n$).

Plain English. $\ell = \liminf a_n$ means: ℓ is undershot infinitely often, and is an eventual lower bound.

Fully quantified. $\ell = \liminf_{n\to\infty} a_n \iff$

(i) $\forall \varepsilon > 0\ \exists^{\infty} n,\ a_n < \ell + \varepsilon.$

(ii) $\forall \varepsilon > 0\ \exists N\ \forall n \geq N,\ a_n > \ell - \varepsilon.$

Consequence. $\liminf a_n$ is the smallest subsequential limit of (a_n). There always exists a subsequence $a_{n_k} \to \ell$.

Plain English. $L = \limsup a_n$ means two things simultaneously: L is exceeded infinitely often (up to any tolerance), and L is an eventual upper bound (from some point onward).

Fully quantified. $L = \limsup_{n\to\infty} a_n \iff$

(i) $\forall \varepsilon > 0\ \exists^{\infty} n,\ a_n > L - \varepsilon.$

(ii) $\forall \varepsilon > 0\ \exists N\ \forall n \geq N,\ a_n < L + \varepsilon.$

Consequence. $\limsup a_n$ is the largest subsequential limit of (a_n). There always exists a subsequence $a_{n_k} \to L$.

Plain English. The set of convergent real sequences is closed under the standard arithmetic operations, and limits commute with those operations. If $x_n \to x$ and $y_n \to y$ then $x_n \pm y_n \to x \pm y$, $cx_n \to cx$, $x_n y_n \to xy$, and $x_n/y_n \to x/y$ (provided $y \neq 0$).

Skeleton / checklist.

1. Assume $a_n \to L$.
2. Use $\varepsilon = 1$: get N s.t. $|a_n - L| < 1$ for all $n \geq N$.
3. For $n \geq N$: $|a_n| \leq |L| + 1$ (triangle inequality).
4. Bound the finite head $\{a_1, \ldots, a_{N-1}\}$ by their maximum.
5. Set $M = \max\big(|a_1|, \ldots, |a_{N-1}|, |L| + 1\big)$.

Plain English. Changing finitely many terms of a sequence does not affect convergence or its limit.

Fully quantified.

$$\text{If } (a_n) \text{ and } (b_n) \text{ differ in only finitely many terms,}$$
$$\text{then } a_n \to L \iff b_n \to L.$$

Consequence. Convergence, divergence, and the limit value are all tail properties: they are determined entirely by the eventual behavior of the sequence, not its initial segment.

Plain English. Whether a sequence converges (and to what limit) depends only on its eventual behavior finitely many initial terms are irrelevant.

Fully quantified.

$$\forall a : \mathbb{N} \to \mathbb{R}\ \forall L \in \mathbb{R}\ \forall k \in \mathbb{N},$$
$$(a_n \to L) \iff (a_{n+k} \to L).$$

Why it matters. Formally justifies every "without loss of generality, assume $n \geq N_0 \ldots$" move in analysis proofs. Also: $\limsup$, $\liminf$, Cauchy, and boundedness are all tail properties.

Skeleton / checklist.

1. Put the bounded sequence inside $[A, B]$.
2. Bisect; pick the half containing infinitely many terms.
3. Iterate: nested closed intervals with lengths $\to 0$.
4. Pick indices n_k so a_{n_k} lies in the k-th interval.
5. Nested Interval Theorem gives a point L; then $a_{n_k} \to L$.

Skeleton / checklist.

1. Assume (a_n) is Cauchy.
2. Use $\varepsilon = 1$: get N s.t. $|a_m - a_n| < 1$ for all $m, n \geq N$.
3. Fix $n = N$: for $m \geq N$, $|a_m| \leq |a_N| + 1$.
4. Bound the finite head $\{a_1, \ldots, a_{N-1}\}$ by their maximum.
5. Combine: one global bound M.

Skeleton / checklist.

1. Assume $a_n \to L$ and $a_n \to M$.
2. If $L \neq M$, set $\varepsilon = \frac{1}{2}|L - M| > 0$.
3. Get N_1, N_2 for each limit with tolerance ε.
4. For $n \geq \max(N_1, N_2)$: $|L - M| \leq |L - a_n| + |a_n - M| < 2\varepsilon = |L - M|$.
5. Contradiction $\Rightarrow L = M$.

Skeleton / checklist.

1. Define $s_n = \sup_{k \geq n} a_k$, $i_n = \inf_{k \geq n} a_k$.
2. Note $i_n \leq a_n \leq s_n$; (s_n) decreasing, (i_n) increasing.
3. ($\Rightarrow$) If $a_n \to L$: tail in $(L-\varepsilon, L+\varepsilon)$ forces $s_n \to L$ and $i_n \to L$.
4. ($\Leftarrow$) If $s_n \to L$ and $i_n \to L$: squeeze theorem gives $a_n \to L$.

Flashcards — B Deck: Formula-First (Avery 5388)

$$\forall a \in A,\ a \leq u.$$

$$\forall a \in A,\ \ell \leq a.$$

$$\exists u \in \mathbb{R}\ \forall a \in A,\ a \leq u$$
$$\wedge\ \exists \ell \in \mathbb{R}\ \forall a \in A,\ \ell \leq a.$$

$$(m \in A) \ \wedge \ \forall a \in A, \ a \leq m.$$

$$(m \in A) \ \wedge \ \forall a \in A, \ m \leq a.$$

$$\forall a \in A, \ a \leq s$$
$$\wedge \ \forall u \in \mathbb{R}, \ \big(\forall a \in A, \ a \leq u\big) \to s \leq u.$$

$$\forall a \in A,\ t \leq a$$
$$\wedge\ \forall \ell \in \mathbb{R},\ \big(\forall a \in A,\ \ell \leq a\big) \to \ell \leq t.$$

$$\exists M > 0\ \forall n \in \mathbb{N},\ |a_n| \leq M.$$

$$\forall M > 0\ \exists n \in \mathbb{N},\ |a_n| > M.$$

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

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

$$\forall \varepsilon > 0 \ \exists N \in \mathbb{N} \ \forall m, n \geq N, \ |a_m - a_n| < \varepsilon.$$

$$\exists \varepsilon > 0 \ \forall N \in \mathbb{N} \ \exists m, n \geq N, \ |a_m - a_n| \geq \varepsilon.$$

$$\begin{aligned}(A \neq \emptyset) \ \wedge \ (\exists u \in \mathbb{R} \ \forall a \in A, \ a \leq u) \\ \rightarrow \ \exists s \in \mathbb{R} \ \big[(\forall a \in A, \ a \leq s) \ \wedge \\ \forall u, \ (\forall a \in A, \ a \leq u) \rightarrow s \leq u\big].\end{aligned}$$

$$\forall \varepsilon > 0 \ \exists n \in \mathbb{N}, \ \tfrac{1}{n} < \varepsilon.$$

$$\exists \sigma : \mathbb{N} \to \mathbb{N} \left(\forall k < \ell,\ \sigma(k) < \sigma(\ell) \right) \wedge\ a_{\sigma(k)} \to L.$$

$$\forall a : \mathbb{N} \to \mathbb{R}\ \exists b \text{ subseq. of } a,$$
$$\Big(\forall k,\ b_k \leq b_{k+1} \Big) \vee \Big(\forall k,\ b_{k+1} \leq b_k \Big).$$

$$\left(a_{2n} \to L\ \wedge\ a_{2n+1} \to L \right) \to a_n \to L.$$

$\big(b \text{ subseq. of } a\big) \wedge \big(c \text{ subseq. of } b\big) \to \big(c \text{ subseq. of } a\big).$

$\forall n,\ a_n \leq a_{n+1}.$

$\forall n,\ a_{n+1} \leq a_n.$

$$\Big(\forall n,\ x_n \le y_n\Big) \wedge (x_n \to x) \wedge (y_n \to y) \to x \le y.$$

$$(x_n \to x)\ \wedge\ (y_n \to y)\ \to\ x_n y_n \to xy.$$

$$s = \sup A\ \Rightarrow\ \forall \varepsilon > 0\ \exists a \in A,\ s - \varepsilon < a.$$

$$(a_n \to L) \iff (a_{n+k} \to L)\ \forall k \in \mathbb{N}.$$

$$(a_n),\ (b_n) \text{ differ finitely } \Rightarrow\ (a_n \to L \iff b_n \to L).$$

$$\forall \varepsilon > 0\ \exists^{\infty} n,\ a_n > L - \varepsilon$$
$$\wedge\ \forall \varepsilon > 0\ \exists N\ \forall n \geq N,\ a_n < L + \varepsilon.$$

$$\forall \varepsilon > 0 \ \exists^{\infty} n, \ a_n < \ell + \varepsilon$$
$$\wedge \ \forall \varepsilon > 0 \ \exists N \ \forall n \geq N, \ a_n > \ell - \varepsilon.$$

$$\big(a_n \to L \ \wedge \ a_n \to M\big) \to L = M.$$

$$\Big(\forall \varepsilon > 0 \ \exists N \ \forall n \geq N, \ |a_n - L| < \varepsilon\Big)$$
$$\to \Big(\exists M > 0 \ \forall n, \ |a_n| \leq M\Big).$$

$$\big(a_n \to L \ \wedge \ b \text{ subseq. of } a\big) \to b_n \to L.$$

$$\Big(\exists M > 0 \ \forall n, \ |a_n| \leq M\Big)$$
$$\to$$
$$\Big(\exists b \text{ subseq. of } a, \ \exists L \in \mathbb{R}, \ b_n \to L\Big).$$

$$\Big(\forall \varepsilon > 0 \ \exists N \ \forall n \geq N, \ |a_n - L| < \varepsilon\Big)$$
$$\to$$
$$\Big(\forall \varepsilon > 0 \ \exists N \ \forall m, n \geq N, \ |a_m - a_n| < \varepsilon\Big).$$

$$\Big(\forall \varepsilon > 0\ \exists N\ \forall m, n \geq N,\ |a_m - a_n| < \varepsilon\Big)$$
$$\to \Big(\exists M > 0\ \forall n,\ |a_n| \leq M\Big).$$

$$\Big(\forall \varepsilon > 0\ \exists N\ \forall m, n \geq N,\ |a_m - a_n| < \varepsilon\Big)$$
$$\leftrightarrow$$
$$\Big(\exists L \in \mathbb{R}\ \forall \varepsilon > 0\ \exists N\ \forall n \geq N,\ |a_n - L| < \varepsilon\Big).$$

$$\Big(\forall n,\ a_n \leq a_{n+1}\Big) \wedge \Big(\exists M\ \forall n,\ a_n \leq M\Big)$$
$$\to\ a_n \to \sup\{a_n : n \in \mathbb{N}\}.$$

$$\Big(\forall n,\ a_n \le b_n \le c_n\Big) \wedge (a_n \to L) \wedge (c_n \to L)$$
$$\to\ b_n \to L.$$

$$\neg(\exists L \in \mathbb{R},\ a_n \to L) \iff$$
$$\Big(\forall M > 0\ \exists n,\ |a_n| > M\Big)$$
$$\vee\ \Big(\exists \text{ subseqs } b, c\ \exists L \neq L',\ b_n \to L\ \wedge\ c_n \to L'\Big).$$

$$\liminf_{n\to\infty} a_n\ \le\ \limsup_{n\to\infty} a_n.$$

$$\Big(\exists L \in \mathbb{R},\ a_n \to L\Big)$$
$$\leftrightarrow$$
$$\Big(\liminf_{n\to\infty} a_n = \limsup_{n\to\infty} a_n \in \mathbb{R}\Big).$$

Bounded Set. A set is bounded if it has both an upper bound and a lower bound.

Lower Bound (Set). A real number ℓ is a lower bound of $A \subseteq \mathbb{R}$ if every element of A is $\geq \ell$. Note: ℓ need not belong to A.

Upper Bound (Set). A real number u is an upper bound of $A \subseteq \mathbb{R}$ if every element of A is $\leq u$. Note: u need not belong to A.

Supremum (LUB). $s = \sup A$: s is an upper bound of A, and $s \leq$ every other upper bound.

ε-move: $\forall \varepsilon > 0\ \exists a \in A\ (s - \varepsilon < a)$.

Minimum. m is the minimum of A: it belongs to A and is a lower bound of A. Contrast with inf: min must be in A.

Maximum. m is the maximum of A: it belongs to A and is an upper bound of A. Contrast with sup: max must be in A.

NOT Bounded (Sequence). Every proposed bound is exceeded: no single M captures all terms.

Bounded Sequence. All terms lie within fixed distance M from 0. The bound M is global it works for every n at once.

Infimum (GLB). $t = \inf A$: t is a lower bound of A, and every other lower bound ℓ satisfies $\ell \leq t$.

ε-move: $\forall \varepsilon > 0\ \exists a \in A\ (a < t + \varepsilon)$.

Cauchy Sequence. Terms eventually cluster together: for large enough N, any two terms beyond N are within ε. Key distinction from convergence: no limit L appears.

NOT Convergent to L. Some tolerance $\varepsilon > 0$ is violated infinitely often. Quantifier order flips to $\exists\varepsilon\ \forall N\ \exists n$ the signature of negation.

Convergence ($a_n \to L$). For every tolerance $\varepsilon > 0$, all sufficiently late terms lie within ε of L. The quantifier order $\forall\varepsilon\ \exists N\ \forall n$ is the essential pattern.

Archimedean Property (workhorse form). For any $\varepsilon > 0$, some $1/n$ undershoots it. This justifies the "$\varepsilon = 1/n$" move ubiquitous in analysis proofs.

Equivalent form: $\forall x \in \mathbb{R}\ \exists n \in \mathbb{N},\ n > x$ ($\mathbb{N}$ is unbounded in $\mathbb{R}$).

Axiom of Completeness. Every nonempty bounded-above subset of $\mathbb{R}$ has a supremum in $\mathbb{R}$. This fails in $\mathbb{Q}$ it is the defining property of $\mathbb{R}$.

NOT Cauchy. Some fixed gap $\varepsilon > 0$ is never closed: arbitrarily far out, two terms remain at least ε apart.

Convergence via Even/Odd Subsequences. Both parity subsequences converging to L forces the full sequence to L. Special case of the Partition Convergence Principle.

Monotone Subsequence Theorem. Every real sequence has a monotone subsequence. Proved via peak indices. Combined with MCT gives an alternative proof of BW.

Subsequential Limit. L is a subsequential limit of (a_n): some subsequence converges to L. If $(a_n) \to L$ then L is the only subsequential limit.

Decreasing sequence. Each term is $\geq$ the next. Non-strict equality permitted. Strictly decreasing requires $a_{n+1} < a_n$.

Increasing sequence. Each term is $\leq$ the next. Non-strict equality permitted. Strictly increasing requires $a_n < a_{n+1}$.

Subsequence of a Subsequence. Subsequence composition is transitive. Composition of strictly increasing index maps is strictly increasing. Used whenever BW is applied twice in a single proof.

ε-Characterization of Supremum. $\sup A$ is always approachable from inside A: $s - \varepsilon$ is never an upper bound. Symmetric: $t = \inf A \Rightarrow \forall \varepsilon > 0\ \exists a \in A,\ a < t + \varepsilon$. Engine of nearly every completeness proof.

Algebra of Limits. Limits commute with $+$, $-$, scalar multiple, $\times$, and $\div$ (with $y \neq 0$).

Order Limit Theorem. Termwise inequality is preserved in the limit. Warning: strict inequality $x_n < y_n$ does not guarantee $x < y$ the limit can collapse the gap.

lim sup Characterization. $L = \limsup a_n$: exceeded infinitely often from below, eventually bounded above by L. L is the largest subsequential limit; a subsequence $a_{n_k} \to L$ always exists.

Finite Modification Theorem. Changing finitely many terms preserves convergence and the limit. Consequence of convergence being a tail property.

Convergence is a Tail Property. Convergence and its limit are unaffected by any finite initial segment. Formally justifies every "assume $n \geq N_0$" move in proofs.

Convergent $\Rightarrow$ Bounded. Every convergent sequence is bounded. Use $\varepsilon = 1$ to bound the tail, then take max with the finite head.

Uniqueness of Limits. A sequence in $\mathbb{R}$ cannot converge to two different values. Proved by contradiction using $\varepsilon = \frac{1}{2}|L - M|$.

$\liminf$ Characterization. $\ell = \liminf a_n$: undershot infinitely often from above, eventually bounded below by ℓ. ℓ is the smallest subsequential limit; a subsequence $a_{n_k} \to \ell$ always exists.

Convergent ⇒ Cauchy. Every convergent sequence is Cauchy. Use the triangle inequality: $|a_m - a_n| \le |a_m - L| + |L - a_n|$.

Bolzano–Weierstrass. Every bounded real sequence has a convergent subsequence. Proved by repeated bisection of a containing interval.

Subsequence Inherits Limit. If $(a_n) \to L$ then every subsequence also converges to L. The same N works since $n_k \ge k$.

Monotone Convergence Theorem (increasing case). A bounded increasing sequence converges to its supremum. The decreasing case is symmetric: bounded decreasing $\to$ inf.

Cauchy $\Leftrightarrow$ Convergent (in $\mathbb{R}$). Completeness of $\mathbb{R}$: a sequence converges iff it is Cauchy. ($\Rightarrow$) always holds; ($\Leftarrow$) uses Bolzano–Weierstrass + uniqueness.

Cauchy $\Rightarrow$ Bounded. Every Cauchy sequence is bounded. Use $\varepsilon = 1$ to bound the tail from a_N, then take max with the finite head.

$\liminf \leq \limsup$. The eventual lower oscillation level never exceeds the eventual upper one. Follows from $i_n \leq s_n$ for all n, taking limits.

Divergence Criterion. (a_n) diverges iff it is unbounded or has two subsequences with different limits. Practical test: exhibit two subsequences converging to distinct values.

Squeeze Theorem. A sequence trapped between two sequences sharing a common limit must converge to that limit. Key tool: $L - \varepsilon < a_n \leq b_n \leq c_n < L + \varepsilon$ for $n \geq \max(N_a, N_c)$.

Convergence $\Leftrightarrow \liminf = \limsup$. A sequence converges iff its liminf and limsup agree (and are finite). When they do, the common value is the limit. Proved via the squeeze theorem on $i_n \le a_n \le s_n$.

Chapter 12

Complex Numbers ($\mathbb{C}$)

Where You Are in the Journey

$$\mathbb{N}, \mathbb{Z}, \mathbb{Q} \;\to\; \text{Real Numbers } (\mathbb{R}) \;\to\; \text{Complex Numbers } (\mathbb{C}) \;\to\; \text{Complex Analysis} \;\to\; \cdots$$

How we got here. The reals are algebraically incomplete: $x^2 = -1$ has no real solution. The complex numbers adjoin $i = \sqrt{-1}$, yielding an algebraically closed field — every polynomial splits completely.

What this chapter will build. Construction of $\mathbb{C}$ as $\mathbb{R}^2$ with complex multiplication, the field axioms, modulus and argument, polar form, De Moivre's theorem, and the fundamental theorem of algebra.

Where this leads. Complex analysis studies differentiable functions on $\mathbb{C}$. The algebraic closure of $\mathbb{C}$ is the starting point for algebraic geometry over $\mathbb{C}$.

Status: Planned

Coming Soon

Notes, proofs, and exercises will appear here in a future revision.

12.1 Notes

To be populated.

12.2 Proofs

To be populated.

12.3 Capstone

To be populated.

Part III

Abstract Mathematics

Analysis

Chapter 13

Real Analysis

13.1 Notes

Where You Are in the Journey

Propositional Logic $\to$ Predicate Calculus $\to$ Sets & Functions $\to$ Proof Techniques $\to$ $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$ $\to$ Real Analysis $\to$ Algebraic Structures $\to$ Metric Spaces $\to$ Topology $\to$ $\cdots$

How we got here. Volume II established the real numbers as a complete ordered field and proved the axiomatic foundations: field and order axioms, bounds, suprema, and the completeness axiom. Real analysis takes these tools and develops the theory of sequences, series, and their limiting behaviour on $\mathbb{R}$.

What this chapter builds. We develop rigorous ε-N analysis on $\mathbb{R}$: convergence of sequences, the algebra of limits, monotone convergence, Cauchy sequences, subsequences and Bolzano–Weierstrass, limit superior and inferior, and infinite series with convergence tests.

Where this leads. Metric spaces generalise the distance structure of $\mathbb{R}$ to arbitrary sets. Every theorem here has a metric-space analogue. Real analysis on $\mathbb{R}$ is the prototype for all of it.

Structural Roadmap

Sequences $\longrightarrow$ Convergence $\longrightarrow$ Subsequences $\longrightarrow$ Series $\longrightarrow$ Asymptotics $\longrightarrow$ Applications

The axiomatic foundations of $\mathbb{R}$ (field axioms, intervals, bounds, completeness) live in Volume II, Chapter: Real Numbers. The progression here begins where those foundations end:

1. Sequences and boundedness
2. Convergence and the algebra of limits
3. Monotone sequences and the monotone convergence theorem
4. Cauchy sequences and completeness revisited
5. Subsequences, Bolzano–Weierstrass, and the subsequence toolkit
6. Recurrence inequalities and iterative control
7. Limit superior and limit inferior
8. Growth and asymptotics
9. Infinite series and partial sums
10. Absolute and conditional convergence
11. Convergence tests (comparison, ratio, root, alternating)
12. Rearrangements and power series

Remark 13.1 (Primary sources). Abbott, Understanding Analysis; Rudin, Principles of Mathematical Analysis; Pons, Real Analysis for the Undergraduate; Ross, Elementary Analysis.

13.1.1 Sequences

Sequences Quick Reference

Core items	Key definitions/results introduced in this file.
How to use	Read the boxed items first; proofs and consequences follow.
Dependencies	Refer back to earlier sections as needed.

13.1.1.1 Basic Definitions

Definition (Sequence)

A sequence in a set X is a function

$$x : \mathbb{N} \to X.$$

For each $n \in \mathbb{N}$, the value $x(n)$ is denoted

$$x_n := x(n),$$

and the sequence is written

$$(x_n)_{n\in\mathbb{N}} \quad \text{or simply} \quad (x_n).$$

Remark 13.2 (Logical Form).

$$\exists x \left(x : \mathbb{N} \to X \ \wedge \ \forall n \in \mathbb{N},\ x_n = x(n)\right).$$

Remark 13.3 (Common Notation). Equivalent notations:

$$(x_n)_{n=1}^{\infty}, \qquad (x_n), \qquad \{x_n\}_{n=1}^{\infty}, \qquad n \mapsto x_n.$$

All represent the same underlying function.

Remark 13.4. When $X = \mathbb{R}$, we speak of a real sequence. When $X = \mathbb{R}^m$, we speak of a vector-valued sequence.

Definition (Subsequence)

Let $(x_n)_{n\in\mathbb{N}}$ be a sequence in X, i.e., a function

$$x : \mathbb{N} \to X, \qquad n \mapsto x_n.$$

A function $y : \mathbb{N} \to X$ is a subsequence of (x_n) if and only if there exists a strictly increasing function

$$\sigma : \mathbb{N} \to \mathbb{N}, \qquad k < \ell \implies \sigma(k) < \sigma(\ell),$$

such that $y = x \circ \sigma$. Writing $n_k := \sigma(k)$, the composed function takes the form

$$(x \circ \sigma)(k) = x(\sigma(k)) = x(n_k) =: x_{n_k},$$

and the subsequence is written

$$(x_{n_k})_{k\in\mathbb{N}} \quad \text{or simply} \quad (x_{n_k}).$$

Remark 13.5 (Logical Form). A function $y : \mathbb{N} \to X$ is a subsequence of x if and only if

$$\exists\, \sigma : \mathbb{N} \to \mathbb{N} \left(\forall\, k < \ell \in \mathbb{N},\ \sigma(k) < \sigma(\ell) \ \wedge\ y = x \circ \sigma \right).$$

The quantifier $\exists\, \sigma$ asserts that a witnessing index function can be exhibited. The subsequence is the composition $y = x \circ \sigma$, not σ itself.

Remark 13.6 (Function Composition). The composition $x \circ \sigma$ makes the two-layer structure explicit:

$$\mathbb{N} \xrightarrow{\sigma} \mathbb{N} \xrightarrow{x} X.$$

σ selects an infinite subset of indices in their natural order; x maps those indices to elements of X. The subsequence is entirely determined by the choice of σ.

Remark 13.7 (Implications of Strict Monotonicity). Since σ is strictly increasing, the following hold.

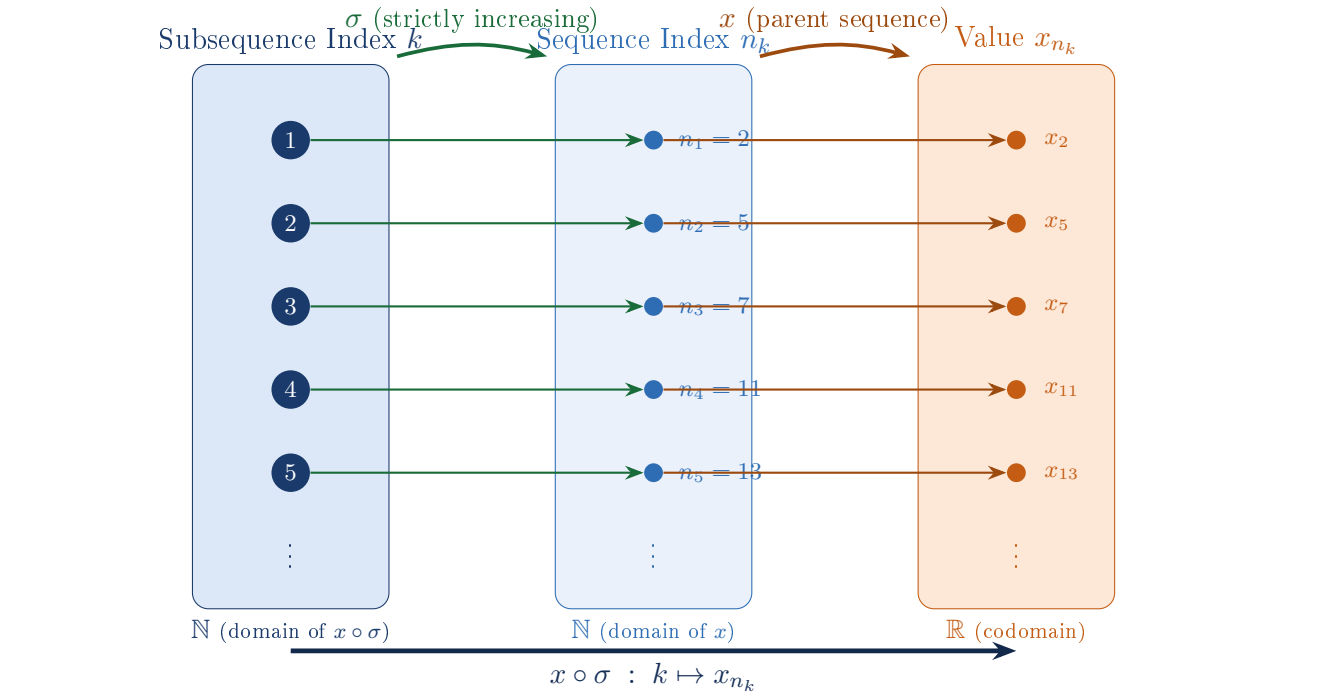
- Injectivity. Strict monotonicity implies injectivity: $k \neq \ell \implies \sigma(k) \neq \sigma(\ell)$, since either $k < \ell$, giving $\sigma(k) < \sigma(\ell)$, or $k > \ell$, giving $\sigma(k) > \sigma(\ell)$.
- Index growth. $\sigma(k) \geq k$ for all $k \in \mathbb{N}$. This follows directly from strict monotonicity and the discreteness of $\mathbb{N}$: $\sigma(1) \geq 1$, and $\sigma(k+1) \geq \sigma(k) + 1$, so $\sigma(k) \geq k$ at each step without appeal to a separate induction argument.
- Divergence of indices. Since $\sigma(k) \geq k$ for all k and $k \to \infty$, it follows immediately that $\sigma(k) \to \infty$. This is not an independent assumption it is a consequence of strict monotonicity alone. In particular, $n_k \to \infty$, so every subsequence reaches arbitrarily far into the original sequence. This fact is used in convergence proofs where n_k must be eventually large.
- Order preservation. $k < \ell \implies n_k < n_\ell$, so the relative order of selected terms is unchanged.
- Deletion only. Terms of (x_n) may be omitted but never duplicated or reinserted out of order. A subsequence is a selection, not a rearrangement.

Remark 13.8 (Common Notation). Equivalent notations for the subsequence $x \circ \sigma$:

$$(x_{n_k})_{k=1}^{\infty}, \qquad (x_{n_k}), \qquad k \mapsto x_{n_k}.$$

The index k tracks position within the subsequence; $n_k = \sigma(k)$ tracks position within the original sequence. These two indices must be kept distinct in proofs.

Remark 13.9 (Sources). The functional definition of a sequence as a map $x : \mathbb{N} \to X$ appears in ?, ?, ?, ?, and ?. The compositional formulation $x \circ \sigma$ is made explicit in ?, Chapter 6 [?] and ?, Chapter 3 [?].



Four consequences of strict monotonicity:

Order Preservation
$k < \ell \Rightarrow \sigma(k) < \sigma(\ell)$
Strict monotonicity. The defining property of σ.

Index Growth
$\sigma(k) \geq k \quad \forall k \in \mathbb{N}$
From strict monotonicity and discreteness of $\mathbb{N}$.

Injectivity
$k \neq \ell \Rightarrow \sigma(k) \neq \sigma(\ell)$
Strict monotonicity implies injectivity directly.

Unbounded Indices
$\sigma(k) \to \infty$ as $k \to \infty$
Since $\sigma(k) \geq k \to \infty$. Follows from index growth.

Key Identity $\quad (x \circ \sigma)(k) \;=\; x\big(\sigma(k)\big) \;=\; x(n_k) \;=:\; x_{n_k}$

k = position in subsequence

$n_k \;=\; \sigma(k)$ = position in original sequence

These two indices must be kept distinct in every proof.

Figure 13.1: Subsequence as function composition $\mathbb{N} \xrightarrow{\sigma} \mathbb{N} \xrightarrow{x} \mathbb{R}$.

13.1.1.2 Main Theorems

Theorem 13.10 (A Sequence Is Determined by Its Values). Let $x, y : \mathbb{N} \to X$ be sequences. If

$$\forall n \in \mathbb{N}, \quad x_n = y_n,$$

then $x = y$ as functions.

Remark 13.11. This theorem relies only on the definition of equality of functions. It implies that sequence equality is pointwise equality.

Theorem 13.12 (Every Subsequence Is a Sequence). If (x_n) is a sequence in X and (n_k) is strictly increasing in $\mathbb{N}$, then (x_{n_k}) is a sequence in X.

Remark 13.13. This theorem depends on function composition. It implies that subsequences remain within the same ambient space. Later, this becomes crucial in BolzanoWeierstrass and compactness arguments.

13.1.1.3 Consequences

The logical implication of this section is:

A sequence is not merely an ordered list, but a function with domain $\mathbb{N}$. Therefore:

$$\text{Sequence theory} = \text{Function theory on } \mathbb{N}.$$

This structural viewpoint allows us to:

- Define limits using quantifiers over $\mathbb{N}$.
- Treat subsequences as compositions.
- Apply functional equality rigorously.

Remark 13.14 (Logical Structure). The foundational flow is:

$$\text{Definition of Sequence} \Rightarrow \text{Equality Theorem} \Rightarrow \text{Subsequence Definition} \Rightarrow \text{Subsequence Theorem}.$$

Thus all later limit theory rests on the functional interpretation.

13.1.1.4 Canonical Examples

Example (Constant sequence)

Fix $c \in \mathbb{R}$. Define

$$x_n := c \quad \text{for all } n \in \mathbb{N}.$$

This is the simplest example of a sequence.

Example (Arithmetic sequence)

Given $a, d \in \mathbb{R}$, define

$$x_n := a + (n-1)d.$$

Each term differs from the previous one by the fixed increment d.

Example (Geometric sequence)

Given $a, r \in \mathbb{R}$, define

$$x_n := ar^{n-1}.$$

Each term is obtained by multiplying the previous one by the ratio r.

Example (Harmonic sequence)

$$x_n := \frac{1}{n}.$$

This sequence decreases to 0.

Example (Alternating sequence)

$$x_n := (-1)^n.$$

This sequence oscillates between 1 and -1.

Example (Polynomial growth sequence)

Let $k \in \mathbb{N}$. Define

$$x_n := n^k.$$

This sequence grows without bound.

Example (Exponential decay sequence)

$$x_n := \frac{1}{2^n}.$$

This sequence decreases rapidly toward 0.

Remark 13.15 (Purpose of Canonical Examples). These examples serve as test cases for:

- boundedness,
- monotonicity,
- convergence and divergence,
- oscillation,
- growth rates.

Nearly every theorem about sequences can be sanity-checked against this list.

13.1.1.5 Logical Classification Table

Sequence	Bounded	Monotone	Convergent	Divergent
Constant $x_n = c$	Yes	Yes	Yes	No
Arithmetic $x_n = a + (n-1)d$	If $d = 0$	If $d \geq 0$ or $d \leq 0$	If $d = 0$	If $d \neq 0$
Geometric $x_n = ar^{n-1}$	If $\lvert r\rvert \leq 1$	If $r \geq 0$	If $\lvert r\rvert < 1$	If $\lvert r\rvert > 1$
Harmonic $x_n = \frac{1}{n}$	Yes	Decreasing	Yes	No
Alternating $x_n = (-1)^n$	Yes	No	No	Yes
Polynomial $x_n = n^k$	No	Increasing	No	Yes
Exponential decay $x_n = \frac{1}{2^n}$	Yes	Decreasing	Yes	No

Remark 13.16 (Logical Dependencies). The classifications rely on later results:

- Monotone sequences converge if and only if they are bounded (Monotone Convergence Theorem).
- Convergent sequences are bounded.
- Divergence may occur via unbounded growth or oscillation.

Thus this table previews the structure of future theorems.

13.1.2 Bounded Sequences and Types of Bounds

Sequence Bounds Quick Reference

Core items	Key definitions/results introduced in this file.
How to use	Read the boxed items first; proofs and consequences follow.
Dependencies	Refer back to earlier sections as needed.

13.1.2.1 Bounded Sequences

13.1.2.2 Basic Definitions

Definition (Bounded sequence)

A sequence (x_n) in $\mathbb{R}$ is called bounded if there exists a real number $M > 0$ such that

$$\lvert x_n\rvert \leq M \quad \text{for all } n \in \mathbb{N}.$$

Remark 13.17 (Equivalent form). A sequence (x_n) is bounded if and only if there exist real numbers m and M such that

$$m \leq x_n \leq M \quad \text{for all } n \in \mathbb{N}.$$

Remark 13.18 (Logical form).

$$\exists M > 0\, \forall n \in \mathbb{N}\, \big(|x_n| \le M\big).$$

$$\exists m, M \in \mathbb{R}\, \forall n \in \mathbb{N}\, \big(m \le x_n \le M\big).$$

Remark 13.19 (Three structural viewpoints of boundedness). A sequence may be understood as bounded at three distinct but equivalent levels:

1. Function viewpoint. A sequence $\{x_n\}_{n=1}^{\infty}$ is a function $x : \mathbb{N} \to \mathbb{R}$, $x(n) = x_n$. Boundedness is the condition that this function is bounded on its domain.
2. Set (range) viewpoint. The values form the subset $\{x_n : n \in \mathbb{N}\} \subseteq \mathbb{R}$. The sequence is bounded if and only if this range is a bounded subset of $\mathbb{R}$.
3. Order-theoretic viewpoint. Boundedness is not a property of the index set $\mathbb{N}$ but of the range in $\mathbb{R}$. The sequence bounds (upper and lower) are exactly the set-theoretic bounds of the range, as defined in Section 11.1.3.

These perspectives are logically equivalent but conceptually distinct, and each becomes useful in different parts of analysis.

Definition (Upper bound for a sequence)

A real number M is an upper bound for the sequence (x_n) if

$$x_n \le M \quad \text{for all } n \in \mathbb{N}.$$

Definition (Lower bound for a sequence)

A real number m is a lower bound for the sequence (x_n) if

$$m \le x_n \quad \text{for all } n \in \mathbb{N}.$$

Definition (Bounded above / below)

A sequence (x_n) is said to be

- bounded above if it has at least one upper bound;
- bounded below if it has at least one lower bound;
- bounded if it is both bounded above and bounded below.

Remark 13.20. These are special cases of the set-theoretic definitions in Section 11.1.3, applied to the range $\{x_n : n \in \mathbb{N}\}$.

Definition (Supremum of a sequence)

If the set $\{x_n : n \in \mathbb{N}\}$ is bounded above, its least upper bound is called the supremum of the sequence, denoted

$$\sup\{x_n\} \quad \text{or} \quad \sup_{n \in \mathbb{N}} x_n.$$

Definition (Infimum of a sequence)

If the set $\{x_n : n \in \mathbb{N}\}$ is bounded below, its greatest lower bound is called the infimum of the sequence, denoted

$$\inf\{x_n\} \quad \text{or} \quad \inf_{n\in\mathbb{N}} x_n.$$

Definition (Maximum and minimum of a sequence)

A sequence (x_n) has a maximum if there exists $N \in \mathbb{N}$ such that

$$x_N \geq x_n \quad \text{for all } n \in \mathbb{N}.$$

A sequence (x_n) has a minimum if there exists $N \in \mathbb{N}$ such that

$$x_N \leq x_n \quad \text{for all } n \in \mathbb{N}.$$

Remark 13.21. If a maximum exists, then $\max\{x_n\} = \sup\{x_n\}$, and similarly for a minimum and the infimum. The supremum or infimum of a sequence need not be attained by any term.

13.1.2.3 Main Theorems

13.1.2.4 Consequences

$$\text{Bounded sequence} \iff \text{Range } \{x_n\} \text{ is bounded in } \mathbb{R}.$$

Thus sequence boundedness reduces to set boundedness, and all results about suprema and infima of sets apply directly to sequences.

Remark 13.22 (Logical Structure).

$$\text{Upper/Lower bounds} \Rightarrow \text{Boundedness} \Rightarrow \text{Supremum/Infimum} \Rightarrow \text{Maximum/Minimum}.$$

All later extremal and limit arguments depend on this hierarchy.

13.1.3 Convergence of Sequences in $\mathbb{R}$

Convergence Quick Reference

Core items	Key definitions/results introduced in this file.
How to use	Read the boxed items first; proofs and consequences follow.
Dependencies	Refer back to earlier sections as needed.

13.1.3.1 Convergence of Sequences

13.1.3.2 Basic Definitions

Definition (Convergence informal description)

A sequence (x_n) of real numbers is said to converge to a real number L if the terms x_n can be made arbitrarily close to L by taking n sufficiently large.

Remark 13.23. This description captures the intuition of convergence but is not logically precise. A rigorous definition is given below using ε-bounds.

Definition (Convergence ε definition)

Let (x_n) be a sequence in $\mathbb{R}$ and let $L \in \mathbb{R}$. We say that (x_n) converges to L, and write

$$x_n \to L \quad \text{or} \quad \lim_{n\to\infty} x_n = L,$$

if

$$\forall \varepsilon > 0\ \exists N \in \mathbb{N}\ \forall n \in \mathbb{N}\ \big(n \geq N \to |x_n - L| < \varepsilon\big).$$

Remark 13.24 (Logical form).

$$\forall \varepsilon > 0\ \exists N \in \mathbb{N}\ \forall n \in \mathbb{N}\ \big(n \geq N \Rightarrow x_n \in (L-\varepsilon,\, L+\varepsilon)\big).$$

Remark 13.25. The number N may depend on ε, but must work for all $n \geq N$.

Remark 13.26 (Source alignment). The ε–definition of convergence is standard in real analysis texts; see, e.g., ?????.

Definition (Neighborhood formulation of convergence in $\mathbb{R}$)

Let (x_n) be a sequence in $\mathbb{R}$ and let $L \in \mathbb{R}$. The sequence (x_n) converges to L if and only if

$$\forall U \subseteq \mathbb{R}\ \big(U \text{ is an open neighborhood of } L \to \exists N \in \mathbb{N}\ \forall n \geq N,\ x_n \in U\big).$$

Remark 13.27. Equivalently, (x_n) converges to L if and only if

for every open set U containing L, all but finitely many terms of the sequence lie in U.

Remark 13.28 (Source alignment). Neighborhood formulations of convergence are standard in topology/metric space treatments; see, e.g., ???.

13.1.3.3 Main Theorems

Theorem 13.29 (Every convergent sequence is bounded). Let (x_n) be a sequence in $\mathbb{R}$. If (x_n) converges, then (x_n) is bounded.

Remark 13.30. Boundedness is a necessary but not sufficient condition for convergence.

Remark 13.31. This theorem assumes the ε–definition of convergence (and standard inequalities for absolute value).

This theorem implies: whenever you prove $x_n \to L$, you automatically obtain a bound $|x_n| \leq M$ for all n.

Theorem 13.32 (Equivalence of convergence definitions in $\mathbb{R}$). For a sequence (x_n) in $\mathbb{R}$ and a point $L \in \mathbb{R}$, the following are equivalent:

1. (x_n) converges to L in the ε-sense.
2. For every $\varepsilon > 0$, eventually $x_n \in (L - \varepsilon, L + \varepsilon)$.
3. For every open neighborhood U of L, eventually $x_n \in U$.

Remark 13.33. In $\mathbb{R}$, the ε-definition and the topological definition of convergence coincide because open neighborhoods are precisely unions of open intervals.

Remark 13.34. This theorem assumes: the standard topology on $\mathbb{R}$ (equivalently, the usual metric).

This theorem implies: you may freely switch between ε–language and neighborhood language when working in $\mathbb{R}$.

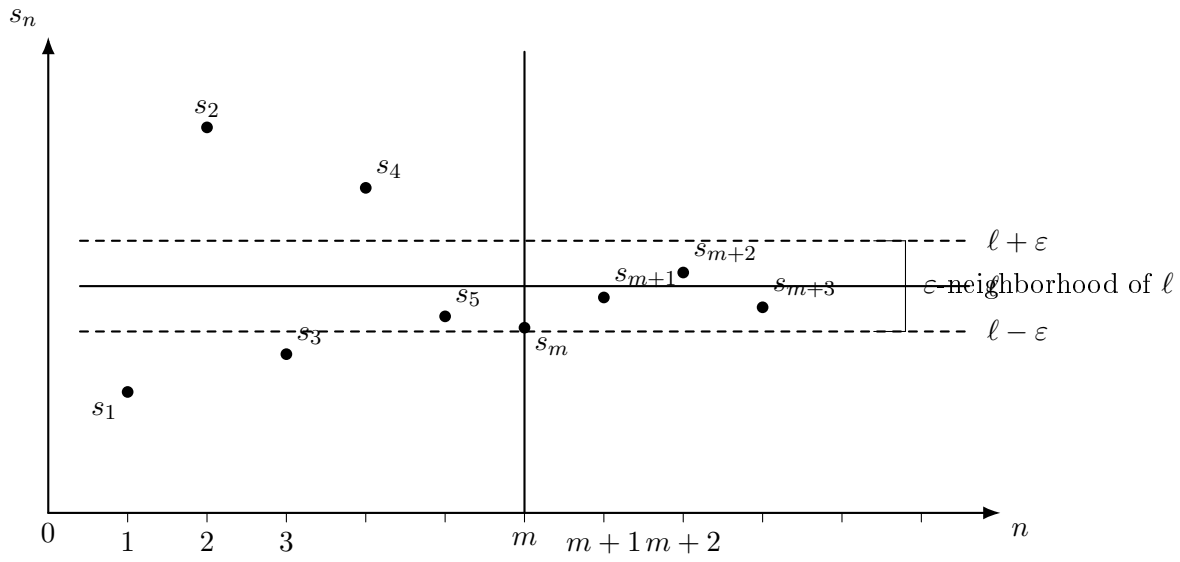


Figure 13.2: Illustration of ε–convergence of a sequence (s_n) to ℓ.

13.1.3.4 Consequences

The logical implication of this entire section is:

- The informal "terms get closer to L" intuition is made rigorous by a quantifier pattern

$$\forall \varepsilon > 0 \; \exists N \; \forall n \geq N : \; |x_n - L| < \varepsilon.$$

- Convergence forces boundedness:

$$x_n \to L \;\Rightarrow\; (x_n) \text{ is bounded.}$$

- In $\mathbb{R}$, the ε–definition and the neighborhood definition are equivalent, so either language may be used depending on context.

Remark 13.35 (Logical Structure). The major convergence statements interlock as follows:

$$\text{(Informal intuition)} \Rightarrow (\varepsilon\text{–definition}) \Rightarrow \text{Boundedness Theorem}$$

and in parallel

$$(\varepsilon\text{–definition}) \Longleftrightarrow \text{Neighborhood definition in } \mathbb{R}.$$

Proposition 13.36 (Bounded Does Not Imply Convergent). The converse of the statement

$$\text{convergent} \Rightarrow \text{bounded}$$

is false.

Remark 13.37 (Dependence on the metric). The definition of convergence depends on the metric of the space.

In $\mathbb{R}$, the metric is

$$d(x, y) = |x - y|,$$

so the familiar ε–definition of convergence uses absolute value.

However, in a general metric space (X, d), convergence is defined by

$$d(x_n, x) < \varepsilon.$$

Thus absolute value is not an intrinsic feature of convergence; it is simply the metric on $\mathbb{R}$.

13.1.4 K-Tails of Sequences

K Tails Quick Reference	
Core items	Key definitions/results introduced in this file.
How to use	Read the boxed items first; proofs and consequences follow.
Dependencies	Refer back to earlier sections as needed.

13.1.4.1 K-Tails of Sequences

The long-term behavior of a sequence is completely determined by its tails. Finite initial segments play no role in convergence, Cauchy behavior, limit superior/inferior, or subsequence extraction.

This section formalizes that principle.

13.1.4.2 Basic Definitions

Definition (k-tail of a sequence)

Let (x_n) be a sequence in $\mathbb{R}$ and let $k \in \mathbb{N}$. The k-tail of (x_n) is the sequence

$$(x_k, x_{k+1}, x_{k+2}, \dots).$$

Definition (Tail set)

Let (x_n) be a sequence and let $k \in \mathbb{N}$. The tail set beginning at k is

$$T_k := \{x_n : n \geq k\}.$$

Remark 13.38. A k-tail removes the first $k-1$ terms of the sequence. All analytic behavior of a sequence occurs in its tails.

13.1.4.3 Main Theorems

Theorem 13.39 (Convergence is a Tail Property). A sequence (x_n) converges to L if and only if every k-tail converges to L.

Theorem 13.40 (Finite Modification Theorem). If two sequences differ in only finitely many terms, then either both converge to the same limit or neither converges.

Theorem 13.41 (Cauchy is a Tail Property). A sequence (x_n) is Cauchy if and only if every k-tail is Cauchy.

Theorem 13.42 (Tail Suprema are Monotone). Let (x_n) be bounded and define

$$s_k := \sup T_k.$$

Then (s_k) is a decreasing sequence.

Theorem 13.43 (Limit Superior via Tails). If (x_n) is bounded and

$$s_k := \sup T_k,$$

then

$$\limsup x_n = \lim_{k \to \infty} s_k.$$

Remark 13.44. Limit superior captures the eventual upper behavior of shrinking tails.

Theorem 13.45 (Subsequences are Iterated Tails). Every subsequence is obtained by repeatedly passing to later tails.

13.1.4.4 Consequences

The logical implications of this section are:

- Convergence ignores finite initial segments.
- Cauchy behavior ignores finite initial segments.
- $\limsup$ and $\liminf$ are defined entirely through tail sets.
- Subsequences are constructed by iterating tails.
- All asymptotic behavior is tail behavior.

Corollary 13.46 (Finite Alteration Does Not Affect Limit). Removing or altering finitely many terms of a convergent sequence does not change its limit.

Corollary 13.47 (Tails of a Convergent Sequence Converge). If (x_n) converges, then any tail of (x_n) converges to the same limit.

Remark 13.48. The phrase "for all sufficiently large n" is precisely tail language.

13.1.4.5 Structural Principle

Remark 13.49 (Tail Principle). In real analysis, properties defined by "eventually" or "for sufficiently large indices" depend only on the tails of sequences.

Finite behavior is analytically irrelevant.

Remark 13.50 (Logical Structure).

$$\text{Tail containment} \Rightarrow \text{Convergence} \Rightarrow \text{Finite modification invariance.}$$

$$\text{Tail nesting} \Rightarrow \text{Monotonic tail suprema} \Rightarrow \limsup \text{ theory.}$$

13.1.5 Algebra of Sequences

Algebra Of Sequences Quick Reference	
Core items	Key definitions/results introduced in this file.
How to use	Read the boxed items first; proofs and consequences follow.
Dependencies	Refer back to earlier sections as needed.

13.1.5.1 Algebra of Sequences

The set of all real sequences is closed under the usual algebraic operations. Moreover, convergence is preserved under these operations, provided mild conditions are satisfied.

13.1.5.2 Basic Definitions

Definition (Algebraic operations)

Let (x_n) and (y_n) be sequences in $\mathbb{R}$, and let $c \in \mathbb{R}$. Define new sequences by:

1. Sum:
$$(x_n) + (y_n) := (x_n + y_n)$$
2. Difference:
$$(x_n) - (y_n) := (x_n - y_n)$$
3. Scalar multiple:
$$c(x_n) := (cx_n)$$
4. Product:
$$(x_n)(y_n) := (x_n y_n)$$
5. Quotient:
$$\frac{(x_n)}{(y_n)} := \left(\frac{x_n}{y_n}\right), \quad \text{provided } y_n \neq 0 \text{ for all } n.$$

13.1.5.3 Main Theorems

Theorem 13.51 (Uniqueness of Limits). If a sequence (x_n) converges to a real number L and also converges to a real number M, then $L = M$.

Theorem 13.52 (Order Limit Theorem). Let (x_n) and (y_n) be sequences of real numbers such that $x_n \leq y_n$ for all $n \in \mathbb{N}$. If $x_n \to x$ and $y_n \to y$, then $x \leq y$.

Theorem 13.53 (Squeeze Theorem). Let (x_n), (y_n), and (z_n) be sequences in $\mathbb{R}$. Assume that

$$x_n \leq y_n \leq z_n \quad \text{for all } n \in \mathbb{N}.$$

If $x_n \to L$ and $z_n \to L$, then $y_n \to L$.

Theorem 13.54 (Algebra of limits). Let (x_n) and (y_n) be sequences in $\mathbb{R}$ such that

$$x_n \to x, \qquad y_n \to y.$$

Then:

1. Sum rule
$$x_n + y_n \to x + y$$
2. Difference rule
$$x_n - y_n \to x - y$$
3. Scalar multiple rule
$$cx_n \to cx \quad \text{for any } c \in \mathbb{R}$$

4. Product rule
$$x_n y_n \to xy$$

5. Quotient rule If $y \neq 0$ and $y_n \neq 0$ for all sufficiently large n, then
$$\frac{x_n}{y_n} \to \frac{x}{y}.$$

Remark 13.55. Each part of the algebra of limits theorem is proved directly from the εdefinition of convergence using:

- the triangle inequality,
- boundedness of convergent sequences,
- basic inequalities in $\mathbb{R}$.

No appeal to continuity is required.

Remark 13.56. The quotient rule requires the additional assumption $y \neq 0$. This ensures that the sequence $(1/y_n)$ is eventually well-defined and bounded.

13.1.5.4 Consequences

The logical implication of this entire section is:

- Algebraic operations on sequences produce sequences (closure).
- When sequences converge, their limits interact predictably with addition, subtraction, scalar multiplication, multiplication, and (with mild extra hypotheses) division.
- Order hypotheses are inherited by limits (Order Limit Theorem), and sandwiching forces convergence (Squeeze Theorem).

Corollary 13.57 (Limit Respects Upper Bounds). If $x_n \leq b$ for all n and $x_n \to x$, then $x \leq b$.

Corollary 13.58 (Convergent Times Bounded Is Bounded). If (x_n) converges and (y_n) is bounded, then the product sequence $(x_n y_n)$ is bounded.

Corollary 13.59 (Convergence of Absolute Values). If (x_n) converges, then $(|x_n|)$ converges and
$$|x_n| \to |x|.$$

Remark 13.60. These algebraic rules justify the informal manipulation of limits familiar from calculus, but their validity rests entirely on the εdefinition of convergence.

Remark 13.61 (Logical Structure). The major sequence theorems interlock as follows:

$$\text{Uniqueness of Limits} \Rightarrow \text{Well-definedness of ``the'' limit} \Rightarrow \text{Algebra of limits.}$$

and

$$\text{Order Limit Theorem} \Rightarrow \text{Squeeze Theorem} \Rightarrow \text{Many convergence proofs in practice.}$$

13.1.6 Monotone Sequences and Monotone Convergence

Monotone Sequences Quick Reference

Core items	Key definitions/results introduced in this file.
How to use	Read the boxed items first; proofs and consequences follow.
Dependencies	Refer back to earlier sections as needed.

13.1.6.1 Basic Definitions

Definition (Increasing and decreasing sequences)

A sequence (x_n) is called

- increasing if $x_n \leq x_{n+1}$ for all $n \in \mathbb{N}$,
- strictly increasing if $x_n < x_{n+1}$ for all n,
- decreasing if $x_{n+1} \leq x_n$ for all n,
- strictly decreasing if $x_{n+1} < x_n$ for all n.

Definition (Monotone sequence)

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

13.1.6.2 Main Theorem

Theorem 13.62 (Monotone Convergence Theorem). Every bounded monotone sequence of real numbers converges.

More precisely:

1. If (x_n) is increasing and bounded above, then
$$\lim_{n\to\infty} x_n = \sup\{x_n : n \in \mathbb{N}\}.$$
2. If (x_n) is decreasing and bounded below, then
$$\lim_{n\to\infty} x_n = \inf\{x_n : n \in \mathbb{N}\}.$$

13.1.6.3 Consequences

The logical implication of this entire section is:

$$\text{Monotone} + \text{Bounded} \;\Rightarrow\; \text{Convergent}.$$

More precisely,

$$\text{Increasing and bounded above} \Rightarrow \lim x_n = \sup\{x_n\},$$
$$\text{Decreasing and bounded below} \Rightarrow \lim x_n = \inf\{x_n\}.$$

Remark 13.63 (Extended real limits). If (x_n) is increasing but not bounded above, then

$$x_n \to +\infty.$$

If (x_n) is decreasing but not bounded below, then

$$x_n \to -\infty.$$

Thus every monotone sequence has a limit in

$$\mathbb{R} \cup \{\pm\infty\}.$$

Remark 13.64 (Logical structure).

$$\text{Increasing/Decreasing} \Rightarrow \text{Monotone} \Rightarrow \text{Monotone Convergence}.$$

The Monotone Convergence Theorem relies on:

- Order structure,
- Boundedness,
- The Least Upper Bound Property (Completeness).

In fact, it is equivalent (in the presence of ordered field axioms) to:

- Bolzano–Weierstrass,
- The Cauchy Criterion,
- The Nested Interval Property.

13.1.7 Monotone Approximation and Completeness Equivalences

Monotone Approximation Quick Reference	
Core items	Key definitions/results introduced in this file.
How to use	Read the boxed items first; proofs and consequences follow.
Dependencies	Refer back to earlier sections as needed.

13.1.7.1 Monotone Approximation of Suprema and Infima

Proposition 13.65 (Monotone Approximation of Bounds). Let $S \subset \mathbb{R}$ be nonempty and bounded. Then there exist monotone sequences $\{x_n\}_{n\in\mathbb{N}}$ and $\{y_n\}_{n\in\mathbb{N}}$ such that $x_n, y_n \in S$ for all n and

$$\sup S = \lim_{n\to\infty} x_n, \qquad \inf S = \lim_{n\to\infty} y_n.$$

Remark 13.66 (Source note). ?Proposition 2.1.13.

13.1.7.2 Supremum Case: Proof + Dissection

Quantifier Dissection (Supremum engine). Let $\alpha := \sup S$. The core logical clause used to build a sequence is

$$(\forall \varepsilon > 0)(\exists s \in S)\ \big(\alpha - \varepsilon < s \le \alpha\big).$$

Choosing a specific ε-schedule, e.g. $\varepsilon_n = 1/n$, yields

$$(\forall n \in \mathbb{N})(\exists x_n \in S)\ \Big(\alpha - \frac{1}{n} < x_n \le \alpha\Big),$$

and monotonicity is forced by the definitional operator

$$X_n := \max\{x_1, \dots, x_n\}.$$

Logical Skeleton (Supremum case).

$$\Big(S \neq \varnothing\ \wedge\ S \text{ bounded above}\Big) \Rightarrow \exists \alpha \in \mathbb{R}\ (\alpha = \sup S) \Rightarrow \exists (X_n) \subset S\ \Big(X_n \uparrow \alpha\Big).$$

Logical Skeleton (Pure quantifier form; supremum part).

$$\forall S \subset \mathbb{R}\Big[\big((\exists s_0 \in S) \wedge (\exists M)(\forall s \in S)(s \le M)\big) \to \exists (X_n)_{n\in\mathbb{N}}$$

$$\Big((\forall n)(X_n \in S) \wedge (\forall n)(X_n \le X_{n+1}) \wedge (\forall \varepsilon > 0)(\exists N)(\forall n \ge N)(|X_n - \sup S| < \varepsilon)\Big)\Big].$$

Construction Mechanism (Supremum).

- Approximate from inside: for each n, pick $x_n \in S$ with $\sup S - \frac{1}{n} < x_n \le \sup S$.
- Monotonize: set $X_n := \max\{x_1, \dots, x_n\}$.
- Conclude: $X_n \uparrow \sup S$.

13.1.7.3 Infimum Case: Symmetric Proof + Dissection

Quantifier Dissection (Infimum engine). Let $\beta := \inf S$. The core clause is

$$(\forall \varepsilon > 0)(\exists s \in S)\ \big(\beta \le s < \beta + \varepsilon\big).$$

Choosing $\varepsilon_n = 1/n$ yields

$$(\forall n \in \mathbb{N})(\exists y_n \in S)\ \Big(\beta \le y_n < \beta + \frac{1}{n}\Big),$$

and monotonicity is enforced by

$$Y_n := \min\{y_1, \dots, y_n\}.$$

Logical Skeleton (Infimum case).

$$\Big(S \neq \varnothing\ \wedge\ S \text{ bounded below}\Big) \Rightarrow \exists \beta \in \mathbb{R}\ (\beta = \inf S) \Rightarrow \exists (Y_n) \subset S\ \Big(Y_n \downarrow \beta\Big).$$

Logical Skeleton (Pure quantifier form; infimum part).

$$\forall S \subset \mathbb{R}\Big[\big((\exists s_0 \in S) \wedge (\exists m)(\forall s \in S)(m \le s)\big) \to \exists (Y_n)_{n\in\mathbb{N}}$$

$$\Big((\forall n)(Y_n \in S) \wedge (\forall n)(Y_{n+1} \le Y_n) \wedge (\forall \varepsilon > 0)(\exists N)(\forall n \ge N)(|Y_n - \inf S| < \varepsilon)\Big)\Big].$$

Construction Mechanism (Infimum).

- Approximate from inside: for each n, pick $y_n \in S$ with $\inf S \le y_n < \inf S + \frac{1}{n}$.
- Monotonize: set $Y_n := \min\{y_1, \dots, y_n\}$.
- Conclude: $Y_n \downarrow \inf S$.

13.1.7.4 Least Upper Bound Property $\iff$ Monotone Convergence

Definition (Least Upper Bound Property (LUB))

We say $\mathbb{R}$ has the least upper bound property if every nonempty set $A \subset \mathbb{R}$ that is bounded above has a supremum in $\mathbb{R}$.

Theorem 13.67 (Monotone Convergence Theorem (MCT)). Every bounded monotone sequence of real numbers converges.

Theorem 13.68 (Equivalence: LUB $\iff$ MCT). The following are equivalent:

1. $\mathbb{R}$ has the least upper bound property.
2. Every bounded monotone sequence of real numbers converges.

13.1.7.5 (LUB ⇒ MCT): Proof + Dissection

Quantifier Dissection (LUB ⇒ MCT). The convergence proof is the same quantifier engine as supremum approximation:

$$\alpha = \sup A \;\Rightarrow\; (\forall \varepsilon > 0)\ \exists N\ (x_N > \alpha - \varepsilon),$$

and monotonicity upgrades $\exists N$ to a tail statement:

$$(\forall \varepsilon > 0)(\exists N)(\forall n \geq N)\ (\alpha - \varepsilon < x_n \leq \alpha).$$

Logical Skeleton (LUB ⇒ MCT).

$$\text{bounded increasing } (x_n) \Rightarrow A = \{x_n\} \text{ bounded above} \Rightarrow \exists \alpha = \sup A \Rightarrow x_n \to \alpha.$$

Logical Skeleton (Pure quantifier form; LUB ⇒ MCT).

$$\Big(\forall A \subset \mathbb{R}\big((\exists a_0 \in A) \wedge (\exists M)(\forall a \in A)(a \leq M) \Rightarrow \exists \sup A\big)\Big) \Rightarrow$$

$$\forall (x_n)\Big(\big((\forall n)(x_n \leq x_{n+1}) \wedge (\exists M)(\forall n)(x_n \leq M)\big) \Rightarrow \exists L\ (\forall \varepsilon > 0)(\exists N)(\forall n \geq N)(|x_n - L| < \varepsilon)\Big).$$

Construction Mechanism (LUB ⇒ MCT).

- Form the image set $A = \{x_n\}$.
- Use LUB to produce $\alpha = \sup A$.
- Use the ε-witness property of sup to get one index N with $x_N > \alpha - \varepsilon$.
- Use monotonicity to push that inequality to all $n \geq N$.

13.1.7.6 (MCT ⇒ LUB): Proof + Dissection (Bisection Construction)

Quantifier Dissection (MCT ⇒ LUB). The bisection construction is a quantifier-control machine:

- Maintain a universal invariant: $(\forall a \in A)(a \leq u_n)$ (upper-bound witness).
- Maintain a failure invariant: l_n is not an upper bound, i.e. $(\exists a \in A)(a > l_n)$ (inside-point witness).
- Force a shrinking bracket: $u_n - l_n \to 0$ (metric control).
- Use MCT to produce limits of monotone bounded sequences and collapse the bracket.
- Convert "l_N is not an upper bound" into the required ε-witness for sup:

$$(\forall \varepsilon > 0)(\exists a \in A)(s - \varepsilon < a).$$

Logical Skeleton (MCT $\Rightarrow$ LUB).

$A \neq \varnothing$, A bounded above $\Rightarrow \exists(l_n)\uparrow$, $\exists(u_n)\downarrow$, $u_n - l_n \to 0 \Rightarrow l_n \to s$, $u_n \to s \Rightarrow s$ is an upper bound and is least =

Logical Skeleton (Pure quantifier form; MCT $\Rightarrow$ LUB).

$$\Big(\forall(x_n)\big(((\forall n)(x_n \leq x_{n+1}) \wedge (\exists M)(\forall n)(x_n \leq M)) \Rightarrow \exists L(\forall \varepsilon > 0)(\exists N)(\forall n \geq N)(|x_n - L| < \varepsilon)\big)\Big) \Rightarrow$$

$$\forall A \subset \mathbb{R}\Big(((\exists a_0 \in A) \wedge (\exists U)(\forall a \in A)(a \leq U)) \Rightarrow \exists s \in \mathbb{R}\ \big(s = \sup A\big)\Big).$$

Construction Mechanism (MCT $\Rightarrow$ LUB).

- Pick $l_1 \in A$ and an upper bound u_1.
- Repeatedly bisect: $m_n = (l_n + u_n)/2$.
- If m_n is an upper bound, tighten from above: $u_{n+1} = m_n$.
- If not, tighten from below: $l_{n+1} = m_n$.
- The bracket width shrinks geometrically, and MCT provides limits.
- Use invariants to prove the limit is the least upper bound.

13.1.7.7 Logical Implications for the Journey

Remark 13.69 (Why this matters structurally). The chain of ideas in this subsection is a reusable meta-template:

$$\text{(Order axiom / extremal property)} \Longleftrightarrow \text{(Monotone sequential convergence)}.$$

It explains why so much of early analysis can be done with monotone sequences: they are a computational interface to completeness.

Remark 13.70 (Downstream reuse). The quantifier pattern

$$(\forall \varepsilon > 0)(\exists \text{witness}) \Rightarrow (\exists \text{sequence with controlled tail})$$

reappears in:

- Bolzano–Weierstrass (extracting convergent subsequences),
- $\limsup$ / $\liminf$ (tail suprema as a decreasing sequence),
- compactness (finite subcovers as "finite witnesses"),
- later approximation theorems (simple functions, step functions, etc.).

13.1.8 Cauchy Sequences

Cauchy Quick Reference	
Core items	Key definitions/results introduced in this file.
How to use	Read the boxed items first; proofs and consequences follow.
Dependencies	Refer back to earlier sections as needed.

13.1.8.1 Cauchy Sequences

13.1.8.2 Basic Definitions

Definition (Cauchy sequence)

A sequence (a_n) in $\mathbb{R}$ is Cauchy if

$$\forall \varepsilon > 0 \, \exists N \in \mathbb{N} \, \forall m, n \in \mathbb{N}, \quad (m \geq N \wedge n \geq N) \Rightarrow |a_n - a_m| < \varepsilon.$$

Remark 13.71. A Cauchy sequence is one whose terms become arbitrarily close to each other without mentioning a limit.

13.1.8.3 Main Theorems

Theorem 13.72 (Every convergent sequence is Cauchy). If (a_n) converges in $\mathbb{R}$, then (a_n) is a Cauchy sequence.

Corollary 13.73 (Every convergent series has Cauchy partial sums). If $\sum_{n=1}^{\infty} a_n$ converges and $s_n := \sum_{k=1}^{n} a_k$ denotes the sequence of partial sums, then (s_n) is a Cauchy sequence. Equivalently,

$$\forall \varepsilon > 0 \; \exists N \text{ s.t. } m > n \geq N \implies \left| \sum_{k=n+1}^{m} a_k \right| < \varepsilon.$$

Theorem 13.74 (Every convergent sequence is Cauchy). If (a_n) converges in $\mathbb{R}$, then (a_n) is a Cauchy sequence.

Corollary 13.75 (Every convergent series has Cauchy partial sums). If $\sum_{n=1}^{\infty} a_n$ converges and $s_n := \sum_{k=1}^{n} a_k$ denotes the sequence of partial sums, then (s_n) is a Cauchy sequence. Equivalently,

$$\forall \varepsilon > 0 \; \exists N \text{ s.t. } m > n \geq N \implies \left| \sum_{k=n+1}^{m} a_k \right| < \varepsilon.$$

Theorem 13.76 (Cauchy sequences are bounded). If (a_n) is a Cauchy sequence in $\mathbb{R}$, then (a_n) is bounded.

Theorem 13.77 (Bolzano–Weierstrass Theorem). Every bounded sequence in $\mathbb{R}$ has a convergent subsequence.

Remark 13.78 (Logical Position). Bolzano–Weierstrass is the bridge between boundedness and convergence. A bounded sequence need not converge, but it cannot avoid convergence entirely: some subsequence must converge.

Remark 13.79 (Dependence on Completeness). The proof relies on the Monotone Convergence Theorem, which in turn depends on the Least Upper Bound Property of $\mathbb{R}$. Thus Bolzano–Weierstrass is a manifestation of completeness.

Corollary 13.80 (Sequential Compactness of Closed Intervals). Every sequence in a closed bounded interval $[a, b] \subset \mathbb{R}$ has a convergent subsequence whose limit lies in $[a, b]$.

Remark 13.81 (Structural Consequence). Bolzano–Weierstrass supplies the critical step in the Cauchy Criterion:

$$\text{Cauchy} \Rightarrow \text{bounded} \Rightarrow \text{convergent subsequence} \Rightarrow \text{full convergence.}$$

It is therefore one of the equivalent formulations of completeness.

Theorem 13.82 (Cauchy Criterion in $\mathbb{R}$). A sequence (a_n) of real numbers converges if and only if it is Cauchy.

13.1.8.4 Consequences

The logical implication of this entire section is:

Remark 13.83 (Interdependence of the Major Theorems on Sequences). At first glance, the principal results of this section may appear to be independent facts:

- Convergent $\Rightarrow$ Cauchy
- Cauchy $\Rightarrow$ bounded
- Bolzano–Weierstrass (bounded $\Rightarrow$ convergent subsequence)
- Cauchy Criterion (Cauchy $\Leftrightarrow$ convergent)

In reality, these theorems form a tightly interlocking structure whose foundation is the completeness of $\mathbb{R}$.

Structural Dependencies.

1. Convergence $\Rightarrow$ Cauchy. This direction uses only the triangle inequality. It does not rely on completeness.
2. Cauchy $\Rightarrow$ bounded. Once the terms eventually cluster tightly, the entire sequence must lie inside some finite interval.
3. Bounded $\Rightarrow$ convergent subsequence (Bolzano–Weierstrass). Boundedness alone does not guarantee convergence, but it guarantees partial convergence.

4. Cauchy $\Rightarrow$ Convergent (Cauchy Criterion). This direction synthesizes the previous results:

$$\text{Cauchy} \Rightarrow \text{bounded} \Rightarrow \text{convergent subsequence} \Rightarrow \text{full convergence.}$$

Logical Structure.

$$\text{Convergent} \Rightarrow \text{Cauchy} \Rightarrow \text{Bounded} \Rightarrow \text{Convergent subsequence.}$$

Completeness of $\mathbb{R}$ upgrades the final step:

$$\text{Cauchy} \iff \text{Convergent.}$$

Without completeness (for example in $\mathbb{Q}$), the implication Cauchy $\Rightarrow$ Convergent fails.

Remark 13.84 (Direct Consequences of Bolzano–Weierstrass). The Bolzano–Weierstrass Theorem yields several immediate structural facts:

1. Bounded nonconvergent sequences oscillate. If a bounded sequence does not converge, then it must admit at least two subsequences converging to different limits.
2. Extremal subsequences exist. For every bounded sequence (a_n), there exist subsequences converging to $\limsup a_n$ and to $\liminf a_n$.
3. Sequential compactness of closed intervals. Every sequence contained in a closed bounded interval $[a, b]$ admits a convergent subsequence whose limit lies in $[a, b]$.

These facts show that Bolzano–Weierstrass controls the long-term structure of bounded sequences.

Extended Logical Chain.

Combining all major results of this section, we obtain:

$$\text{Convergent} \Rightarrow \text{Cauchy} \Rightarrow \text{Bounded} \Rightarrow \text{Convergent subsequence}$$

and, using completeness,

$$\text{Cauchy} \iff \text{Convergent.}$$

Thus boundedness alone does not ensure convergence, but it prevents total divergence. Some limiting behavior must emerge.

Conceptual Summary.

Convergence is an external statement about approaching a number. Cauchy is an internal statement about self-consistency of the sequence.

Completeness asserts that internal consistency is sufficient: there are no "holes" in the real line.

Thus the Cauchy Criterion is not merely a technical tool it is an equivalent formulation of completeness.

13.1.9 Subsequences

Subsequences Quick Reference

Core items	Key definitions/results introduced in this file.
How to use	Read the boxed items first; proofs and consequences follow.
Dependencies	Refer back to earlier sections as needed.

13.1.9.1 Basic Definitions

Definition (Subsequence)

Let (a_n) be a sequence in $\mathbb{R}$. A sequence (a_{n_k}) is called a subsequence of (a_n) if there exists a function

$$k \mapsto n_k$$

from $\mathbb{N}$ to $\mathbb{N}$ such that

1. $n_k < n_{k+1}$ for all $k \in \mathbb{N}$, and
2. the kth term of the new sequence is a_{n_k}.

Remark 13.85. A subsequence is obtained by selecting terms from (a_n) without duplication while preserving their original order. The index sequence (n_k) determines which terms are selected.

Lemma 13.86 (Index Growth). If (n_k) is a strictly increasing sequence in $\mathbb{N}$, then $n_k \geq k$ for all $k \in \mathbb{N}$. In particular, $n_k \to \infty$ as $k \to \infty$.

Definition (Subsequential Limit)

Let (a_n) be a sequence. A real number L is called a subsequential limit of (a_n) if there exists a subsequence (a_{n_k}) such that $a_{n_k} \to L$.
The set of all subsequential limits of (a_n) is denoted $\mathcal{L}(a_n)$.

13.1.9.2 Main Theorems

Theorem 13.87 (Subsequences Inherit Limits). Let (a_n) be a sequence of real numbers and let (a_{n_k}) be any subsequence. If $a_n \to L$, then $a_{n_k} \to L$.

Remark 13.88 (Consistency of Subsequential Limits). If $(a_n) \to L$, then every subsequence converges to L. Consequently, a convergent sequence has exactly one subsequential limit:

$$a_n \to L \implies \mathcal{L}(a_n) = \{L\}.$$

Remark 13.89 (Detecting Divergence). Contrapositive: if two subsequences of (a_n) converge to different limits, then (a_n) does not converge. This provides a practical test for divergence.

Theorem 13.90 (Convergence via Even and Odd Subsequences). Let (a_n) be a sequence. If the even-indexed subsequence (a_{2n}) and the odd-indexed subsequence (a_{2n+1}) both converge to the same limit L, then $(a_n) \to L$.

Remark 13.91. More generally, if (a_n) can be partitioned into finitely many subsequences, each converging to the same limit L, then $a_n \to L$.

Theorem 13.92 (Bolzano–Weierstrass). Every bounded sequence of real numbers has a convergent subsequence. Equivalently: if (a_n) is bounded in $\mathbb{R}$, then there exist $L \in \mathbb{R}$ and a strictly increasing sequence (n_k) in $\mathbb{N}$ such that $a_{n_k} \to L$.

Corollary 13.93 (Existence of Subsequential Limits). Every bounded sequence admits at least one subsequential limit. That is, if (a_n) is bounded, then $\mathcal{L}(a_n) \neq \varnothing$.

Corollary 13.94 (Sequential Compactness of Closed Intervals). Every sequence contained in a closed bounded interval $[a, b]$ has a convergent subsequence whose limit lies in $[a, b]$.

Theorem 13.95 (Monotone Subsequence Theorem). Every sequence in $\mathbb{R}$ has a monotone subsequence.

Remark 13.96 (Alternative Proof of Bolzano–Weierstrass). The Monotone Subsequence Theorem combined with the Monotone Convergence Theorem provides an alternative proof of Bolzano–Weierstrass:

1. Every sequence has a monotone subsequence (Monotone Subsequence Theorem).
2. A bounded monotone sequence converges (Monotone Convergence Theorem).
3. Therefore, every bounded sequence has a convergent subsequence.

13.1.9.3 Divergence Criteria

Theorem 13.97 (Characterization of Divergence). A sequence (a_n) diverges if and only if at least one of the following holds:

1. (a_n) is unbounded.
2. (a_n) has two subsequences converging to different finite limits.
3. (a_n) has a subsequence diverging to $+\infty$ or $-\infty$.

13.1.9.4 Consequences and Structural Summary

This section establishes three structural principles.

(1) Inheritance of Limits.

$$a_n \to L \implies a_{n_k} \to L.$$

Convergence is preserved under passage to subsequences.

(2) Reconstruction from Subsequences.

$$(a_{2n} \to L) \wedge (a_{2n+1} \to L) \implies a_n \to L.$$

Convergence can be established by verifying it on a finite partition.

(3) Emergence of Subsequential Limits.

$$\text{Bounded} \implies \text{Convergent subsequence exists.}$$

Boundedness alone does not ensure convergence, but it guarantees the existence of convergent subsequences.

Remark 13.98 (Connection to Limit Superior and Inferior). For a bounded sequence (a_n), the set $\mathcal{L}(a_n)$ of subsequential limits is nonempty, closed, and bounded. Moreover,

$$\limsup_{n\to\infty} a_n = \sup \mathcal{L}(a_n) = \max \mathcal{L}(a_n),$$

$$\liminf_{n\to\infty} a_n = \inf \mathcal{L}(a_n) = \min \mathcal{L}(a_n).$$

Thus lim sup and lim inf are themselves subsequential limits, and they are the largest and smallest such limits.

Remark 13.99 (Logical Structure).

$$\text{Convergent} \implies \text{All subsequences converge to the same limit,}$$

$$\text{Bounded} \xRightarrow{\text{Bolzano–Weierstrass}} \text{Existence of subsequential limits.}$$

These results form the bridge between basic convergence theory and completeness theory.

Bolzano–Weierstrass is equivalent (in an ordered field) to:

- The Monotone Convergence Theorem,
- The Cauchy Criterion,
- The Nested Interval Property,
- The Least Upper Bound Property.

Remark 13.100 (Inheritance and Reflection of Properties). Different sequence properties behave differently with respect to subsequences.

Property	Inherited by Subsequences?	Reflected by One Subsequence?	Structural Type
Convergent	Yes	No	Tail Property
Cauchy	Yes	No	Tail Property
Bounded	Yes	No	Global Property
Monotone	Yes	No	Global Structural
Eventually monotone	Yes	No	Tail Property
Every subseq. convergent	—	Yes	Universal Property
Every subseq. Cauchy	—	Yes	Universal Property

Remark 13.101. Inherited means: if (a_n) has property P, then every subsequence has P.

Reflected means: if some subsequence has property P, then (a_n) has P.

Most natural properties are inherited but not reflected. The sequence $a_n = (-1)^n$ is bounded and has convergent subsequences, but does not converge.

13.1.10 Subsequence Toolkit

Subsequence Toolkit Quick Reference

Core items	Key definitions/results introduced in this file.
How to use	Read the boxed items first; proofs and consequences follow.
Dependencies	Refer back to earlier sections as needed.

Remark 13.102 (Relationship to Subsequences Notes). The results in this section assume the foundational facts about subsequences proved in the main subsequences section, including:

- Limits are inherited by subsequences.
- Convergent sequences have unique limits.
- BolzanoWeierstrass Theorem.

This section abstracts and generalizes those structural principles.

13.1.10.1 Partition Convergence Principles

Theorem 13.103 (Finite Partition Convergence Principle). Let (a_n) be a sequence in $\mathbb{R}$. Suppose the index set $\mathbb{N}$ admits a finite partition

$$\mathbb{N} = E_1 \cup E_2 \cup \cdots \cup E_k,$$

where:

1. $E_i \cap E_j = \varnothing$ whenever $i \neq j$,

2. each E_i is infinite.

For each $i \in \{1, \dots, k\}$, let $(a_n)_{n \in E_i}$ denote the subsequence indexed by E_i.

If each subsequence $(a_n)_{n \in E_i}$ converges to the same limit L, then the full sequence (a_n) converges to L.

Remark 13.104. The theorem generalizes to metric spaces: replace $|a_n - L|$ with $d(a_n, L)$.

Corollary 13.105 (Residue-Class Convergence). Let (a_n) be a sequence in $\mathbb{R}$ and let $k \geq 1$. If for each $r \in \{0, 1, \dots, k-1\}$ the subsequence (a_{kn+r}) converges to the same limit L, then $(a_n) \to L$.

Corollary 13.106 (Even-Odd Convergence). Let (a_n) be a sequence. If $(a_{2n}) \to L$ and $(a_{2n+1}) \to L$, then $(a_n) \to L$.

13.1.10.2 Inheritance of Sequence Properties

Definition (Inherited Property)

A property $\mathcal{P}$ of sequences is inherited by subsequences if whenever (a_n) satisfies $\mathcal{P}$, every subsequence (a_{n_k}) also satisfies $\mathcal{P}$.

Definition (Reflected Property)

A property $\mathcal{P}$ of sequences is reflected by subsequences if whenever some subsequence (a_{n_k}) satisfies $\mathcal{P}$, the original sequence (a_n) also satisfies $\mathcal{P}$.

Proposition 13.107 (Convergence is Inherited). If $(a_n) \to L$, then every subsequence $(a_{n_k}) \to L$.

Proposition 13.108 (Boundedness is Inherited). If (a_n) is bounded, then every subsequence (a_{n_k}) is bounded.

Proposition 13.109 (Monotonicity is Inherited). If (a_n) is monotone increasing (resp. decreasing), then every subsequence (a_{n_k}) is monotone increasing (resp. decreasing).

Proposition 13.110 (Cauchy Property is Inherited). If (a_n) is Cauchy, then every subsequence (a_{n_k}) is Cauchy.

Example (Convergence is Not Reflected)

The sequence $a_n = (-1)^n$ has the convergent subsequence

$$(a_{2n}) = (1, 1, 1, \dots) \to 1,$$

but (a_n) itself does not converge.
Similarly, $(a_{2n+1}) = (-1, -1, -1, \dots) \to -1$.
This shows that a divergent sequence can have convergent subsequences, so convergence is inherited but not reflected.

Example (Boundedness is Not Reflected)

Let $a_n = n$ for n odd and $a_n = 0$ for n even. The subsequence $(a_{2n}) = (0, 0, 0, \ldots)$ is bounded, but (a_n) is unbounded.

Example (Monotonicity is Not Reflected)

The sequence $a_n = (-1)^n$ has the monotone (constant) subsequence $(a_{2n}) = (1, 1, 1, \ldots)$, but (a_n) is not monotone.

13.1.10.3 Additional Subsequence Tools

Proposition 13.111 (Subsequence of a Subsequence). A subsequence of a subsequence is a subsequence of the original sequence.

Theorem 13.112 (Dense Subsequence Criterion). Let (a_n) be a sequence and $L \in \mathbb{R}$. If every neighborhood of L contains infinitely many terms of (a_n), then (a_n) has a subsequence converging to L.

Remark 13.113 (Connection to Cluster Points). The Dense Subsequence Criterion characterizes subsequential limits: L is a subsequential limit of (a_n) if and only if every neighborhood of L contains infinitely many terms of (a_n). In topological language, L is a cluster point (or accumulation point) of the sequence.

Theorem 13.114 (Diagonal Subsequence Lemma). Let $(a_n^{(m)})_{n=1}^{\infty}$ be a sequence of sequences indexed by $m \in \mathbb{N}$. Suppose that for each m, the sequence $(a_n^{(m)})_{n=1}^{\infty}$ converges to a limit L_m, and suppose $L_m \to L$ as $m \to \infty$.

Then there exists a strictly increasing sequence (n_m) such that the diagonal sequence $(a_{n_m}^{(m)})$ converges to L.

Remark 13.115. The Diagonal Subsequence Lemma is used extensively in functional analysis and measure theory, particularly in proofs involving weak convergence and the construction of convergent subsequences from families of sequences.

13.1.10.4 Structural Classification of Sequence Properties

Definition (Tail Property)

A property $\mathcal{P}$ of sequences is called a tail property if whenever two sequences (a_n) and (b_n) satisfy

$$a_n = b_n \quad \text{for all } n \geq N$$

for some $N \in \mathbb{N}$, then

$$(a_n) \text{ satisfies } \mathcal{P} \iff (b_n) \text{ satisfies } \mathcal{P}.$$

In other words, altering finitely many initial terms does not affect whether the sequence has property $\mathcal{P}$.

Definition (Universal Subsequence Property)

A property $\mathcal{P}$ of sequences is called a universal subsequence property if for every sequence (a_n),

$$(a_n) \text{ satisfies } \mathcal{P} \iff \text{every subsequence of } (a_n) \text{ satisfies } \mathcal{P}.$$

Remark 13.116 (Universal Properties and Inheritance). Universal subsequence properties are necessarily inherited. If $\mathcal{P}$ is universal and (a_n) satisfies $\mathcal{P}$, then by definition every subsequence satisfies $\mathcal{P}$.

Proposition 13.117 (Reflection via Universal Properties). Let $\mathcal{P}$ be a universal subsequence property. Then:

1. $\mathcal{P}$ is inherited by subsequences.
2. $\mathcal{P}$ is reflected by the collection of all subsequences (though not necessarily by any single subsequence).

Example (Universal Property: Every Subsequence is Cauchy)

The property "every subsequence is Cauchy" is a universal subsequence property. Moreover, this property is equivalent to the sequence being Cauchy:

- If (a_n) is Cauchy, then every subsequence is Cauchy (inheritance).
- If every subsequence is Cauchy, then (a_n) is Cauchy (since (a_n) is a subsequence of itself).

In $\mathbb{R}$, this is equivalent to convergence.

Example (Universal Property: Every Subsequence Has a Convergent Subsubsequence)

The property "every subsequence has a convergent subsubsequence" is equivalent to boundedness:

- If (a_n) is bounded, then every subsequence is bounded (inheritance), and by Bolzano–Weierstrass, every bounded subsequence has a convergent subsubsequence.
- Conversely, if (a_n) is unbounded, then it has a subsequence (a_{n_k}) with $|a_{n_k}| \to \infty$. This subsequence has no bounded subsubsequence, hence no convergent subsubsequence.

This characterization connects subsequence logic to compactness: a set $K \subseteq \mathbb{R}$ is sequentially compact if and only if every sequence in K has a convergent subsequence with limit in K.

Example (Tail Properties)

The following are tail properties:

- Convergence (and convergence to a specific limit L)
- Cauchy property
- Boundedness
- Eventually monotone
- Eventually constant

Example (Non-Tail Properties)

The following are not tail properties:

- Monotonicity (changing a_1 can destroy monotonicity)
- Having $a_1 = 0$

13.1.10.5 Summary of Property Classification

Property	Inherited by Subsequences?	Reflected by One Subsequence?	Structural Type
Convergent	Yes	No	Tail Property
Cauchy	Yes	No	Tail Property
Bounded	Yes	No	Tail Property*
Monotone	Yes	No	Global Structural
Eventually monotone	Yes	No	Tail Property
Every subseq. convergent	Yes	Yes (all)	Universal Property
Every subseq. Cauchy	Yes	Yes (all)	Universal Property

*Boundedness is a tail property (changing finitely many terms preserves boundedness) and also a global magnitude condition (it constrains all terms).

Remark 13.118 (Interpretation of the Table). Inherited means: if (a_n) has property P, then every subsequence has P.

Reflected by one means: if some single subsequence has property P, then (a_n) has P. This is rare.

Reflected by all means: if every subsequence has property P, then (a_n) has P. This holds for universal properties.

Most natural properties are inherited but not reflected by any single subsequence. Universal properties are both inherited and reflected by the totality of subsequences.

Remark 13.119 (Hierarchy of Subsequence Properties). The structural types form a hierarchy:

Type	Characterization
Tail Property	Determined by sufficiently late terms
Global Property	Depends on all terms (not just tail)
Universal Property	Equivalent to holding for all subsequences

Understanding this hierarchy clarifies which proof strategies apply:

- Tail properties allow discarding finitely many "bad" initial terms.
- Global properties require controlling the entire sequence.
- Universal properties can be verified by checking all subsequences, or refuted by finding one counterexample subsequence.

13.1.11 Recurrence Inequalities and Iterative Control

Recurrence Relations Quick Reference

Core items	Key definitions/results introduced in this file.
How to use	Read the boxed items first; proofs and consequences follow.
Dependencies	Refer back to earlier sections as needed.

Many arguments involving sequences and subsequences rely on propagating an inequality across indices. This section isolates that structure.

Definition (Recurrence Inequality)

Let (u_n) be a nonnegative sequence, i.e., $u_n \geq 0$ for all $n \geq 0$. We say (u_n) satisfies a recurrence inequality if there exists a function $T : [0, \infty) \to [0, \infty)$ such that

$$u_{n+1} \leq T(u_n) \quad \text{for all } n \geq 0.$$

Remark 13.120. A recurrence inequality provides local control: each term is bounded in terms of the previous term. The main technique is to iterate the inequality to obtain global control.

Definition (Iterates of a Function)

Let $T : [0, \infty) \to [0, \infty)$ be a function. The n-fold iterate of T, denoted T^n, is defined recursively by

$$T^0(x) := x, \qquad T^{n+1}(x) := T(T^n(x)).$$

Lemma 13.121 (Iterative Propagation). Let (u_n) be a nonnegative sequence satisfying

$$u_{n+1} \leq T(u_n)$$

for all $n \geq 0$, where $T : [0, \infty) \to [0, \infty)$ is monotone increasing. Then for all $n \geq 0$,

$$u_n \leq T^n(u_0).$$

Remark 13.122 (Structural Insight). The method is:

1. Extract a recurrence inequality.
2. Iterate it (often by induction).
3. Obtain explicit asymptotic control.
4. Deduce convergence or Cauchy behavior.

This pattern underlies contraction mapping arguments, geometric decay estimates, and certain subadditivity arguments.

Definition (Multiplicative Recurrence)

A nonnegative sequence (u_n) satisfies a multiplicative recurrence if

$$u_{n+1} \le cu_n \quad \text{for some } 0 < c < 1.$$

Proposition 13.123 (Multiplicative Recurrence Bound). If (u_n) is nonnegative and satisfies $u_{n+1} \le cu_n$ with $0 < c < 1$, then

$$u_n \le c^n u_0.$$

In particular, $u_n \to 0$ as $n \to \infty$.

Definition (Affine Recurrence)

A nonnegative sequence (u_n) satisfies an affine recurrence if

$$u_{n+1} \le cu_n + b \quad \text{for some } 0 < c < 1 \text{ and } b \ge 0.$$

Proposition 13.124 (Affine Recurrence Bound). If (u_n) is nonnegative and satisfies $u_{n+1} \le cu_n + b$ with $0 < c < 1$ and $b \ge 0$, then

$$u_n \le c^n u_0 + \frac{b(1-c^n)}{1-c}.$$

In particular,

$$\limsup_{n\to\infty} u_n \le \frac{b}{1-c}.$$

Theorem 13.125 (Discrete Grönwall Inequality). Let (u_n) and (β_n) be nonnegative sequences satisfying

$$u_n \le \alpha + \sum_{k=0}^{n-1} \beta_k u_k \quad \text{for all } n \ge 0,$$

where $\alpha \ge 0$ is a constant. Then

$$u_n \le \alpha \prod_{k=0}^{n-1} (1+\beta_k).$$

In particular,

$$u_n \le \alpha \exp\left(\sum_{k=0}^{n-1} \beta_k\right).$$

Remark 13.126 (Connection to Continuous Grönwall Inequality). The discrete Grönwall inequality is the analogue of the continuous Grönwall–Bellman inequality: if $u : [a, b] \to \mathbb{R}$ satisfies

$$u(t) \leq \alpha + \int_a^t \beta(s)\, u(s)\, ds,$$

with $\alpha \geq 0$ and $\beta(s) \geq 0$, then

$$u(t) \leq \alpha \exp\left(\int_a^t \beta(s)\, ds \right).$$

This continuous version is fundamental in the theory of ordinary differential equations, particularly for proving uniqueness and continuous dependence on initial conditions.

Definition (Subsequence Recurrence)

Let (a_n) be a sequence. We say that a subsequence recurrence inequality holds along (a_{n_k}) if there exists a function T such that

$$|a_{n_{k+1}} - a_{n_k}| \leq T(|a_{n_k} - a_{n_{k-1}}|)$$

for all sufficiently large k.

Remark 13.127 (Induction on Subsequence Indices). When verifying a property $\mathcal{P}(k)$ for all $k \in \mathbb{N}$, standard mathematical induction applies to the subsequence index k, regardless of the values of the original indices n_k.

Definition (Terminology)

- Propagation: Repeated application of a recurrence inequality.
- Iterative control: Bounding u_n via repeated composition T^n.
- Geometric decay: Bounds of the form $u_n \leq Cc^n$ with $0 < c < 1$.
- Telescoping argument: Summing successive differences to control total variation.
- Block decomposition: Writing $n = qk + r$ to propagate additive bounds.

Definition (Subadditive Sequence)

A sequence (a_n) is subadditive if

$$a_{m+n} \leq a_m + a_n \quad \text{for all } m, n \in \mathbb{N}.$$

Definition (Superadditive Sequence)

A sequence (a_n) is superadditive if

$$a_{m+n} \geq a_m + a_n \quad \text{for all } m, n \in \mathbb{N}.$$

Remark 13.128. A sequence (a_n) is superadditive if and only if $(-a_n)$ is subadditive. This duality allows results about subadditive sequences to be transferred to superadditive sequences by negation.

Theorem 13.129 (Fekete's Lemma for Subadditive Sequences). If (a_n) is subadditive, then

$$\lim_{n\to\infty} \frac{a_n}{n} = \inf_{k\geq 1} \frac{a_k}{k}.$$

In particular, the limit exists (possibly equal to $-\infty$).

Theorem 13.130 (Fekete's Lemma for Superadditive Sequences). If (a_n) is superadditive, then

$$\lim_{n\to\infty} \frac{a_n}{n} = \sup_{k\geq 1} \frac{a_k}{k}.$$

In particular, the limit exists (possibly equal to $+\infty$).

Example (Contraction-Type Recurrence)

Suppose (x_n) satisfies

$$|x_{n+1} - x_n| \leq c|x_n - x_{n-1}| \quad \text{for all } n \geq 1,$$

where $0 < c < 1$.

Proposition 13.131 (Contraction Sequence Is Cauchy). If (x_n) satisfies a contraction-type recurrence with $0 < c < 1$, then (x_n) is Cauchy.

Example (Contraction Along a Subsequence)

Suppose (a_{n_k}) is a subsequence satisfying

$$|a_{n_{k+1}} - a_{n_k}| \leq c|a_{n_k} - a_{n_{k-1}}| \quad \text{for all } k \geq 1,$$

where $0 < c < 1$.

Proposition 13.132 (Contraction Subsequence Is Cauchy). If (a_{n_k}) satisfies a contraction-type recurrence with $0 < c < 1$, then (a_{n_k}) is Cauchy.

Remark 13.133 (Unifying Structure). The results in this section share a common logical structure:

$$\text{Local inequality} \implies \text{Iterated bound} \implies \text{Explicit asymptotic control} \implies \text{Tail property}.$$

Remark 13.134 (Hierarchy of Recurrence Types). The section develops the following hierarchy:

Type	Key Result
Multiplicative ($u_{n+1} \leq cu_n$)	Geometric decay: $u_n \leq c^n u_0$
Affine ($u_{n+1} \leq cu_n + b$)	Bounded limit: $\limsup u_n \leq \frac{b}{1-c}$
Subadditive ($a_{m+n} \leq a_m + a_n$)	Fekete: $\lim \frac{a_n}{n} = \inf \frac{a_k}{k}$
Discrete integral ($u_n \leq \alpha + \sum \beta_k u_k$)	Grönwall: $u_n \leq \alpha \exp(\sum \beta_k)$

13.1.12 Limit Superior and Limit Inferior

Limsup Liminf Quick Reference

Core items	Key definitions/results introduced in this file.
How to use	Read the boxed items first; proofs and consequences follow.
Dependencies	Refer back to earlier sections as needed.

13.1.12.1 Limit Superior and Limit Inferior

13.1.12.2 Basic Definitions

Definition (Tail supremum and infimum)

Let (a_n) be a bounded real sequence. For each $n \in \mathbb{N}$ define

$$s_n := \sup\{a_k : k \geq n\}, \qquad i_n := \inf\{a_k : k \geq n\}.$$

The sequences (s_n) and (i_n) are called the tail suprema and tail infima of (a_n).

Theorem 13.135 (Monotonicity of tail suprema and infima). Let (x_n) be a real sequence and define

$$s_n := \sup\{x_k : k \geq n\}, \qquad i_n := \inf\{x_k : k \geq n\}.$$

Then (s_n) is decreasing and (i_n) is increasing. In particular, each of (s_n) and (i_n) has a limit in $\mathbb{R} \cup \{\pm\infty\}$, and these limits are $\limsup x_n$ and $\liminf x_n$.

Definition (Limit superior and limit inferior)

Let (a_n) be a bounded real sequence. The limit superior and limit inferior of (a_n) are defined by

$$\limsup_{n\to\infty} a_n := \lim_{n\to\infty} s_n, \qquad \liminf_{n\to\infty} a_n := \lim_{n\to\infty} i_n,$$

where s_n and i_n are the tail supremum and tail infimum sequences.

Definition (Limit superior and limit inferior)

Let (x_n) be a real sequence. Define

$$s_n := \sup\{x_k : k \geq n\}, \qquad i_n := \inf\{x_k : k \geq n\}.$$

The limit superior and limit inferior of (x_n) are

$$\limsup_{n\to\infty} x_n := \lim_{n\to\infty} s_n, \qquad \liminf_{n\to\infty} x_n := \lim_{n\to\infty} i_n,$$

provided these limits exist in $\mathbb{R} \cup \{\pm\infty\}$.

13.1.12.3 Main Theorems

Theorem 13.136 (Equivalent formulation as inf-sup). For any bounded sequence (a_n),

$$\limsup_{n\to\infty} a_n = \inf_{n\geq 1} \sup_{k\geq n} a_k, \qquad \liminf_{n\to\infty} a_n = \sup_{n\geq 1} \inf_{k\geq n} a_k.$$

Theorem 13.137 (Equivalent characterizations). Let (a_n) be bounded and let $L = \limsup_{n\to\infty} a_n$. Then:

1. $a_n > L - \varepsilon$ for infinitely many n for every $\varepsilon > 0$.
2. For every $\varepsilon > 0$, $a_n < L + \varepsilon$ for all sufficiently large n.

Similarly, if $l = \liminf_{n\to\infty} a_n$, then:

1. $a_n < l + \varepsilon$ for infinitely many n for every $\varepsilon > 0$.
2. For every $\varepsilon > 0$, $a_n > l - \varepsilon$ for all sufficiently large n.

Theorem 13.138 (Basic inequalities). Let (a_n) be bounded. Then

$$\inf_{n\in\mathbb{N}} a_n \;\leq\; \liminf_{n\to\infty} a_n \;\leq\; \limsup_{n\to\infty} a_n \;\leq\; \sup_{n\in\mathbb{N}} a_n.$$

Theorem 13.139 (Sign symmetry). For any bounded sequence (a_n),

$$\limsup_{n\to\infty}(-a_n) = -\liminf_{n\to\infty} a_n, \qquad \liminf_{n\to\infty}(-a_n) = -\limsup_{n\to\infty} a_n.$$

Theorem 13.140 (Order preservation). Let (a_n) and (b_n) be bounded sequences. If

$$a_n \leq b_n \quad \text{for all } n,$$

then

$$\liminf_{n\to\infty} a_n \;\leq\; \liminf_{n\to\infty} b_n, \qquad \limsup_{n\to\infty} a_n \;\leq\; \limsup_{n\to\infty} b_n.$$

Theorem 13.141 (Subadditivity of lim sup). Let (a_n) and (b_n) be bounded sequences. Then

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

Theorem 13.142 (Superadditivity of lim inf). Let (a_n) and (b_n) be bounded sequences. Then

$$\liminf_{n\to\infty} a_n + \liminf_{n\to\infty} b_n \;\leq\; \liminf_{n\to\infty}(a_n + b_n).$$

Theorem 13.143 (Equality when one sequence converges). Suppose (a_n) converges to a and (b_n) is bounded. Then

$$\limsup_{n\to\infty}(a_n + b_n) = a + \limsup_{n\to\infty} b_n,$$
$$\liminf_{n\to\infty}(a_n + b_n) = a + \liminf_{n\to\infty} b_n.$$

Theorem 13.144 (Addition law for convergent sequences). Suppose (a_n) and (b_n) are bounded sequences such that $a_n \to a$ and $b_n \to b$. Then

$$\limsup(a_n + b_n) = a + b \qquad \text{and} \qquad \liminf(a_n + b_n) = a + b.$$

Theorem 13.145 (Scalar multiplication). Let (a_n) be a bounded sequence and $c \in \mathbb{R}$. Then:

1. If $c > 0$:
$$\limsup_{n\to\infty}(c \cdot a_n) = c \cdot \limsup_{n\to\infty} a_n, \qquad \liminf_{n\to\infty}(c \cdot a_n) = c \cdot \liminf_{n\to\infty} a_n.$$
2. If $c < 0$:
$$\limsup_{n\to\infty}(c \cdot a_n) = c \cdot \liminf_{n\to\infty} a_n, \qquad \liminf_{n\to\infty}(c \cdot a_n) = c \cdot \limsup_{n\to\infty} a_n.$$
3. If $c = 0$: $\limsup(c \cdot a_n) = \liminf(c \cdot a_n) = 0$.

Theorem 13.146 (Product inequality for nonnegative sequences). Let (a_n) and (b_n) be bounded sequences with $a_n \geq 0$ and $b_n \geq 0$ for all n. Then

$$\limsup_{n\to\infty}(a_n b_n) \leq \big(\limsup_{n\to\infty} a_n\big)\big(\limsup_{n\to\infty} b_n\big).$$

Theorem 13.147 (Characterization via subsequences). Let (a_n) be bounded. Then

$$\limsup_{n\to\infty} a_n = \sup\{\ell : \ell \text{ is a subsequential limit of } (a_n)\},$$

and

$$\liminf_{n\to\infty} a_n = \inf\{\ell : \ell \text{ is a subsequential limit of } (a_n)\}.$$

Theorem 13.148 (Extremal subsequences). Let (a_n) be bounded. Then there exist subsequences (a_{n_k}) and (a_{m_k}) such that

$$a_{n_k} \to \limsup_{n\to\infty} a_n, \qquad a_{m_k} \to \liminf_{n\to\infty} a_n.$$

Theorem 13.149 (Convergence criterion). A bounded sequence (a_n) converges if and only if

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

In this case,

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

Example (Strict inequality in subadditivity)

Let $a_n = (-1)^n$ and $b_n = (-1)^{n+1} = -a_n$. Then $a_n + b_n = 0$ for all n, so

$$\limsup(a_n + b_n) = 0.$$

However, $\limsup a_n = 1$ and $\limsup b_n = 1$, so

$$\limsup a_n + \limsup b_n = 2.$$

Thus the inequality $\limsup(a_n + b_n) \leq \limsup a_n + \limsup b_n$ can be strict.

Example (Strict inequality in product rule)

Let $a_n = \frac{1+(-1)^n}{2}$ (alternates $0, 1, 0, 1, \ldots$) and $b_n = \frac{1+(-1)^{n+1}}{2}$ (alternates $1, 0, 1, 0, \ldots$). Then $a_n b_n = 0$ for all n, so $\limsup(a_n b_n) = 0$. But $\limsup a_n = 1$ and $\limsup b_n = 1$, so the product of lim sup's is 1.

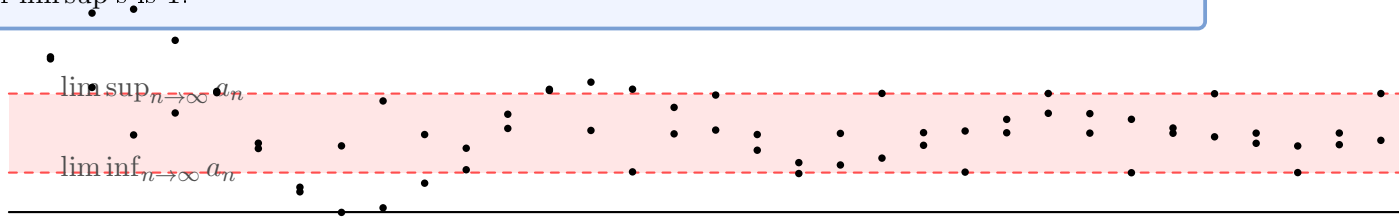


Figure 13.3: Illustration of $\liminf a_n$ and $\limsup a_n$: the oscillation persists, but the transient dies out so points eventually lie within the band.

13.1.12.4 Consequences

The logical implication of this entire section is:

- Tail suprema/infima are monotone (hence have extended real limits).
- lim sup and lim inf can be characterized in multiple equivalent ways: as limits of tail extrema, as inf–sup / sup–inf, and via ε-style "infinitely often" / "eventually" conditions.
- lim sup and lim inf behave functorially with order, addition, and scalar multiplication (with inequalities in general, and equalities under convergence hypotheses).
- A bounded sequence converges exactly when its two extreme limiting behaviors coincide:

$$\limsup a_n = \liminf a_n \iff a_n \text{ converges},$$

  and then all three values equal the common limit.

Remark 13.150. The limit superior and limit inferior capture the largest and smallest possible limiting behavior of a bounded sequence.

The number $\limsup a_n$ is the largest subsequential limit, while $\liminf a_n$ is the smallest subsequential limit. A sequence converges precisely when these two extreme behaviors coincide.

Remark 13.151 (Logical Structure). A clean dependency chain is:

Tail sup/inf $\Rightarrow$ Monotonicity of tail sup/inf $\Rightarrow$ lim sup / lim inf (definitions) $\Rightarrow$ Characterizations and algebraic lav

Remark 13.152 (From limit superior and inferior to series). The machinery of limit superior and inferior developed in this section is not merely a tool for studying individual sequences — it is the natural language for convergence tests on infinite series.

Recall that an infinite series

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

is defined as the limit of its sequence of partial sums

$$S_N := \sum_{n=1}^{N} a_n.$$

Convergence of the series is therefore convergence of the sequence (S_N), and all prior theory applies directly.

The connection to limsup and liminf is concrete:

- The Root Test determines convergence via $\limsup_{n\to\infty} |a_n|^{1/n}$, which exists for any sequence and captures the worst-case exponential growth rate of the terms.
- The Ratio Test uses $\limsup_{n\to\infty} |a_{n+1}/a_n|$ and $\liminf_{n\to\infty} |a_{n+1}/a_n|$ to give sharp convergence and divergence conditions.
- The Cauchy condensation test and comparison arguments rely on tail behavior, which is precisely what limsup and liminf measure.

In each case, the test reduces a question about a series to a question about the limiting behavior of a sequence of real numbers — exactly the setting limsup and liminf were built for.

The series section that follows should therefore be read as an application of the full sequence theory developed so far, with limsup and liminf as the central technical tool.

13.1.13 Growth and Asymptotic Behavior of Sequences

Growth And Asymptotics Quick Reference

Core items	Key definitions/results introduced in this file.
How to use	Read the boxed items first; proofs and consequences follow.
Dependencies	Refer back to earlier sections as needed.

13.1.13.1 Basic Definitions

Definition (Ratio behavior)

Let (a_n) be a sequence with $a_n \neq 0$ for large n. The ratio behavior of (a_n) is governed by the limit (if it exists)

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

Definition (Root behavior)

Let (a_n) be a sequence with $a_n \geq 0$ for large n. The root behavior of (a_n) is governed by the limit (if it exists)

$$\lim_{n\to\infty} \sqrt[n]{a_n}.$$

Definition (Big-O notation)

Let (a_n) and (b_n) be sequences with $b_n \neq 0$ eventually. We write

$$a_n = O(b_n)$$

if there exist constants $C > 0$ and N such that

$$|a_n| \leq C|b_n| \quad \text{for all } n \geq N.$$

Definition (Asymptotic comparison)

Let (a_n) and (b_n) be sequences with $b_n \neq 0$ eventually.

$$a_n = o(b_n) \quad \text{if} \quad \lim_{n\to\infty} \frac{a_n}{b_n} = 0.$$

$$a_n \sim b_n \quad \text{if} \quad \lim_{n\to\infty} \frac{a_n}{b_n} = 1.$$

Definition (Polynomial, exponential, and factorial growth)

A sequence exhibits:

- Polynomial growth if $a_n \sim n^k$ for some $k > 0$.
- Exponential growth if $a_n \sim c^n$ for some $c > 1$.
- Factorial growth if $a_n \sim n!$.

13.1.13.2 Main Theorems

Theorem 13.153 (Ratio theorem for sequences). Let (a_n) be positive and suppose

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

1. If $L < 1$, then $a_n \to 0$.
2. If $L > 1$, then $a_n \to \infty$.

Theorem 13.154 (Root theorem for sequences). Let (a_n) be positive and suppose

$$\lim_{n\to\infty} \sqrt[n]{a_n} = L.$$

1. If $L < 1$, then $a_n \to 0$.
2. If $L > 1$, then $a_n \to \infty$.

Theorem 13.155 (Properties of asymptotic equivalence). The relation $a_n \sim b_n$ is:

- reflexive,
- symmetric,
- transitive.

Hence asymptotic equivalence defines an equivalence relation on sequences that are eventually nonzero.

Theorem 13.156 (Stolz–Cesàro). Let (a_n) and (b_n) be sequences with (b_n) strictly increasing and $b_n \to \infty$. If

$$\lim_{n\to\infty} \frac{a_{n+1} - a_n}{b_{n+1} - b_n} = L,$$

then

$$\lim_{n\to\infty} \frac{a_n}{b_n} = L.$$

Theorem 13.157 (Fundamental asymptotic limits).

$$\lim_{n\to\infty} n^{1/n} = 1, \qquad \lim_{n\to\infty} \frac{\log n}{n^\alpha} = 0 \quad (\alpha > 0),$$

$$\lim_{n\to\infty} \left(1 + \frac{x}{n}\right)^n = e^x.$$

13.1.13.3 Consequences

The logical implication of this section is:

$$\text{Local growth ratios} \Rightarrow \text{Global asymptotic classification.}$$

Remark 13.158 (Connection to series tests).

$$\text{Ratio behavior} \Rightarrow \text{Ratio Test for series.}$$

$$\text{Root behavior} \Rightarrow \text{Root Test for series.}$$

Thus asymptotic growth tools directly govern convergence of infinite series.

Remark 13.159 (Hierarchy of growth). For large n:

$$\log n \ll n^k \ll c^n \ll n!$$

for any $k > 0$ and $c > 1$.

Remark 13.160 (Logical Structure).

Difference quotients ⇒ Stolz–Cesàro ⇒ Asymptotic comparison ⇒ Growth hierarchy ⇒ Series convergence behavi

13.1.14 Series

Series Quick Reference	
Core items	Key definitions/results introduced in this file.
How to use	Read the boxed items first; proofs and consequences follow.
Dependencies	Refer back to earlier sections as needed.

13.1.14.1 Basic Definitions

Definition (Series)

Let (a_n) be a sequence of real numbers. The series associated with (a_n) is the formal expression

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

Remark 13.161. A series is not itself a number, but a symbolic object whose meaning is defined via its sequence of partial sums.

Definition (Partial sums)

Given a sequence (a_n), define the sequence of partial sums (s_N) by

$$s_N := \sum_{n=1}^{N} a_n.$$

Remark 13.162. The series $\sum_{n=1}^{\infty} a_n$ is said to converge if the sequence of partial sums (s_N) converges in $\mathbb{R}$.

Remark 13.163 (Logical structure).

$$\sum_{n=1}^{\infty} a_n \text{ converges} \iff \exists L \in \mathbb{R} \left(\lim_{N\to\infty} s_N = L \right).$$

13.1.14.2 Main Theorems

Theorem 13.164 (Cauchy Condensation Test). Let (a_n) be a nonincreasing sequence of nonnegative real numbers. Then the series

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

converges if and only if the condensed series

$$\sum_{k=0}^{\infty} 2^k \, a_{2^k}$$

converges.

13.1.14.3 Consequences

The logical implication of this section is:

- A series is completely determined by its sequence of partial sums.
- Convergence of a series is therefore a special case of sequence convergence.
- Tests for series are methods for proving convergence of the associated partial-sum sequence.
- The Cauchy Condensation Test reduces certain monotone nonnegative series to a sparser dyadic subsequence.

Remark 13.165 (Structural Position).

$$\text{Series convergence} = \text{Convergence of partial sums.}$$

Thus all sequence results (Cauchy Criterion, Bolzano–Weierstrass, Algebra of Limits, etc.) apply immediately to series via the sequence (s_N).

13.1.15 Absolute and Conditional Convergence

Absolute Convergence Quick Reference

Core items	Key definitions/results introduced in this file.
How to use	Read the boxed items first; proofs and consequences follow.
Dependencies	Refer back to earlier sections as needed.

13.1.15.1 Basic Definitions

Definition (Absolute convergence)

A series

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

is said to converge absolutely if the series

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

converges.

Definition (Conditional convergence)

A series

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

is said to converge conditionally if it converges, but does not converge absolutely.

Remark 13.166 (Logical form).

$$\text{Absolute convergence} \iff \sum |a_n| \text{ converges.}$$

$$\text{Conditional convergence} \iff \sum a_n \text{ converges and } \sum |a_n| \text{ diverges.}$$

Remark 13.167. Absolute convergence is a stronger property than convergence. It imposes global control on the total variation of the series.

13.1.15.2 Main Theorems

Theorem 13.168 (Absolute convergence implies convergence). If

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

converges, then

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

converges.

Remark 13.169. The proof uses only:

- the triangle inequality,
- the Cauchy Criterion,
- completeness of $\mathbb{R}$.

Theorem 13.170 (Comparison via absolute values). If $|a_n| \leq b_n$ for all n, where $b_n \geq 0$ and

$$\sum b_n$$

converges, then

$$\sum a_n$$

converges absolutely (and hence converges).

Theorem 13.171 (Absolute convergence is rearrangement invariant). If a series converges absolutely, then every rearrangement of the series converges to the same sum.

Remark 13.172. The proof of this theorem requires additional combinatorial estimates and will be developed later in the study of rearrangements. The key idea is that absolute convergence prevents cancellation effects from altering the limit.

13.1.15.3 Canonical Examples

Example (Geometric series)

For $|r| < 1$,

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

converges absolutely since

$$\sum |r|^n$$

is geometric.

Example (Alternating harmonic series)

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

converges, but

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

diverges.
Hence it is conditionally convergent.

13.1.15.4 Consequences

The logical implication of this section is:

$$\text{Absolute convergence} \Rightarrow \text{Convergence.}$$

However,

$$\text{Convergence} \not\Rightarrow \text{Absolute convergence.}$$

Remark 13.173 (Logical Structure). The major sequence theorems interlock as follows:

$$\sum |a_n| \text{ converges} \Rightarrow \text{Cauchy partial sums} \Rightarrow \text{Convergent series.}$$

Absolute convergence therefore sits structurally between:

$$\text{Comparison tests} \quad \text{and} \quad \text{Rearrangement theory.}$$

Remark 13.174 (Philosophical interpretation). Absolute convergence eliminates the possibility that convergence is caused merely by oscillatory cancellation. It measures the total accumulated magnitude of the series.

13.1.16 Tests for Series

Series Tests Quick Reference	
Core items	Key definitions/results introduced in this file.
How to use	Read the boxed items first; proofs and consequences follow.
Dependencies	Refer back to earlier sections as needed.

13.1.16.1 Basic Logical Structure

All convergence tests reduce to properties of the partial sums

$$s_N = \sum_{n=1}^{N} a_n.$$

Thus every test ultimately proves that (s_N) is either:

- bounded and monotone, or
- Cauchy.

13.1.16.2 Main Theorems

Theorem 13.175 (Direct Comparison Test). Let $a_n, b_n \geq 0$.

1. If $a_n \leq b_n$ eventually and $\sum b_n$ converges, then $\sum a_n$ converges.
2. If $a_n \geq b_n$ eventually and $\sum b_n$ diverges, then $\sum a_n$ diverges.

Theorem 13.176 (Limit Comparison Test). Let $a_n, b_n > 0$ and suppose

$$\lim_{n\to\infty} \frac{a_n}{b_n} = L, \quad 0 < L < \infty.$$

Then

$$\sum a_n \text{ converges } \iff \sum b_n \text{ converges.}$$

Theorem 13.177 (Ratio Test). Let

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

1. If $L < 1$, the series converges absolutely.
2. If $L > 1$, the series diverges.

Theorem 13.178 (Root Test). Let

$$L = \limsup_{n\to\infty} \sqrt[n]{|a_n|}.$$

1. If $L < 1$, the series converges absolutely.
2. If $L > 1$, it diverges.

Theorem 13.179 (Integral Test). Let f be continuous, positive, decreasing on $[1, \infty)$. Let $a_n = f(n)$.

Then

$$\sum a_n \text{ converges } \iff \int_1^\infty f(x)\,dx \text{ converges.}$$

Theorem 13.180 (p-Series).

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

converges iff $p > 1$.

Theorem 13.181 (Alternating Series Test). If $b_n \geq 0$, decreasing, and $b_n \to 0$, then

$$\sum (-1)^{n+1} b_n$$

converges.

Theorem 13.182 (Dirichlet Test). If

- partial sums of $\sum a_n$ are bounded,
- b_n is monotone and $b_n \to 0$,

then $\sum a_n b_n$ converges.

Theorem 13.183 (Abel Test). If $\sum a_n$ converges and b_n is bounded monotone, then $\sum a_n b_n$ converges.

13.1.16.3 Consequences

Hierarchy of strength:

$$\text{Root} \Rightarrow \text{Ratio} \Rightarrow \text{Comparison.}$$

Absolute convergence tests imply unconditional stability.

Alternating / Dirichlet / Abel capture cancellation-driven convergence.

Remark 13.184 (Logical Core). All tests ultimately rely on:

- comparison,
- Cauchy criterion,
- bounded monotone convergence.

13.1.17 Manipulation and Rearrangement of Series

Series Rearrangements Quick Reference	
Core items	Key definitions/results introduced in this file.
How to use	Read the boxed items first; proofs and consequences follow.
Dependencies	Refer back to earlier sections as needed.

13.1.17.1 Basic Definitions

Definition (Rearrangement of a series)

Let $\sum_{n=1}^{\infty} a_n$ be a series. A rearrangement of this series is any series of the form

$$\sum_{n=1}^{\infty} a_{\sigma(n)},$$

where $\sigma : \mathbb{N} \to \mathbb{N}$ is a bijection.

Remark 13.185. A rearrangement preserves all terms of the series, but possibly changes their order.

Definition (Regrouping)

A regrouping of a series consists of inserting parentheses into the series in such a way that finitely many consecutive terms are summed together before taking limits.

13.1.17.2 Main Theorems

Theorem 13.186 (Absolute convergence is stable under rearrangement). If $\sum a_n$ converges absolutely, then every rearrangement

$$\sum a_{\sigma(n)}$$

converges and has the same sum.

Theorem 13.187 (Riemann Rearrangement Theorem). If $\sum a_n$ converges conditionally (i.e. converges but not absolutely), then for every $L \in \mathbb{R}$ there exists a rearrangement that converges to L.

Moreover, there exist rearrangements that diverge to $+\infty$ or $-\infty$.

Theorem 13.188 (Regrouping and convergence). If $\sum a_n$ converges, then any regrouping of finitely many consecutive terms converges to the same sum.

Remark 13.189. Regrouping preserves convergence. Rearrangement does not unless the series converges absolutely.

Definition (Cauchy product)

Let $\sum a_n$ and $\sum b_n$ be two series. The Cauchy product is the series

$$\sum_{n=0}^{\infty} c_n, \quad c_n := \sum_{k=0}^{n} a_k b_{n-k}.$$

Theorem 13.190 (Cauchy Product Theorem). If both $\sum a_n$ and $\sum b_n$ converge absolutely, then the Cauchy product converges absolutely and

$$\sum c_n = \left(\sum a_n\right)\left(\sum b_n\right).$$

Remark 13.191. If convergence is not absolute, the Cauchy product may fail to converge.

Similarly, rearrangements can change the value.

13.1.17.3 Consequences and Logical Structure

The hierarchy of stability is:

$$\text{Absolute convergence} \Rightarrow \text{Rearrangement stability} \Rightarrow \text{Cauchy product stability}.$$

Conditional convergence implies instability under rearrangement.

Remark 13.192 (Philosophical Summary). Absolute convergence behaves like finite sums.

Conditional convergence behaves like an infinite balancing act: order matters.

13.1.18 Power Series and Radius of Convergence

Power Series Quick Reference

Core items	Key definitions/results introduced in this file.
How to use	Read the boxed items first; proofs and consequences follow.
Dependencies	Refer back to earlier sections as needed.

13.1.18.1 Basic Definitions

Definition (Power series)

Let (a_n) be a sequence of real numbers and let $c \in \mathbb{R}$. A power series centered at c is a series of the form

$$\sum_{n=0}^{\infty} a_n (x-c)^n.$$

Definition (Radius of convergence)

The radius of convergence of a power series

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

is the number $R \in [0, \infty]$ such that:

- the series converges absolutely whenever $|x-c| < R$,
- the series diverges whenever $|x-c| > R$.

Definition (Interval of convergence)

The interval of convergence is the set of x for which the series converges. It is of the form

$$(c-R, c+R)$$

possibly including one or both endpoints.

13.1.18.2 Main Theorems

Theorem 13.193 (Radius of Convergence Theorem). For every power series

$$\sum_{n=0}^{\infty} a_n(x-c)^n,$$

there exists $R \in [0, \infty]$ such that:

1. The series converges absolutely for all $|x-c| < R$.
2. The series diverges for all $|x-c| > R$.

Theorem 13.194 (Cauchy–Hadamard Formula). Let

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

be a power series. Then the radius of convergence is

$$R = \frac{1}{\limsup_{n\to\infty} |a_n|^{1/n}}.$$

Theorem 13.195 (Term-by-term differentiation). Let

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

have radius of convergence $R > 0$. Then for all $|x-c| < R$:

1. The series converges uniformly on every closed interval
$$[c-r, c+r] \subset (c-R, c+R).$$
2. The function f is differentiable on $(c-R, c+R)$.
3. The derivative is obtained by term-by-term differentiation:
$$f'(x) = \sum_{n=1}^{\infty} n a_n (x-c)^{n-1}.$$
4. The differentiated series has the same radius of convergence R.

13.1.18.3 Consequences and Logical Structure

Remark 13.196 (Structural Position). Power series sit at the intersection of:

$$\text{Sequences} \to \text{Series} \to \text{Absolute convergence} \to \text{Root test} \to \text{limsup}.$$

The Cauchy–Hadamard formula is the culmination of the entire limsup theory.

Remark 13.197 (Uniform convergence inside the radius). On every compact subinterval of $(c-R, c+R)$, power series converge uniformly. This makes them exceptionally well-behaved:

$$\text{Inside } R: \quad \text{Uniform convergence} \Rightarrow \text{Continuous} \Rightarrow \text{Differentiable} \Rightarrow \text{Smooth}.$$

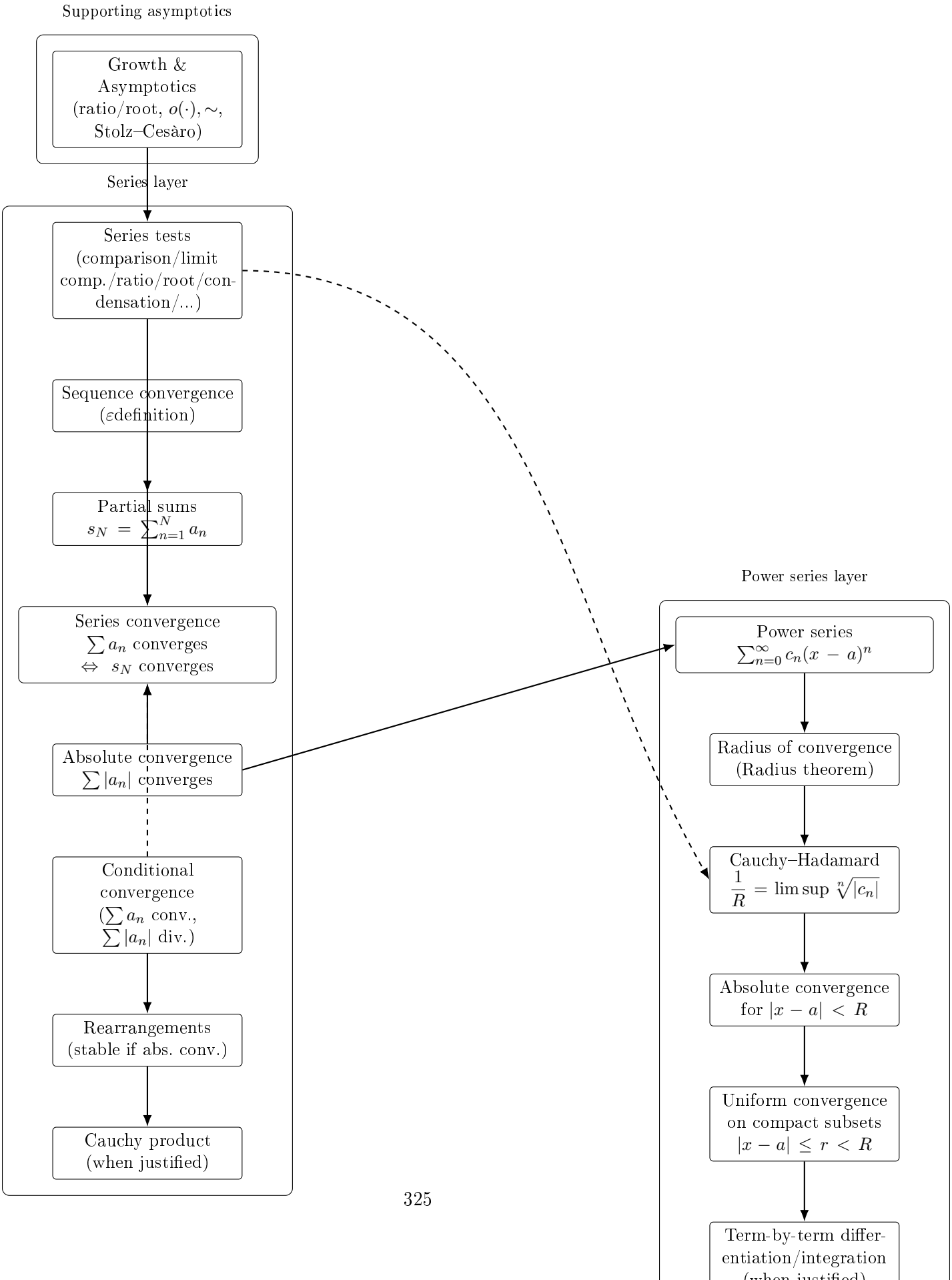
Remark 13.198 (Completeness connection). The existence of R ultimately depends on:

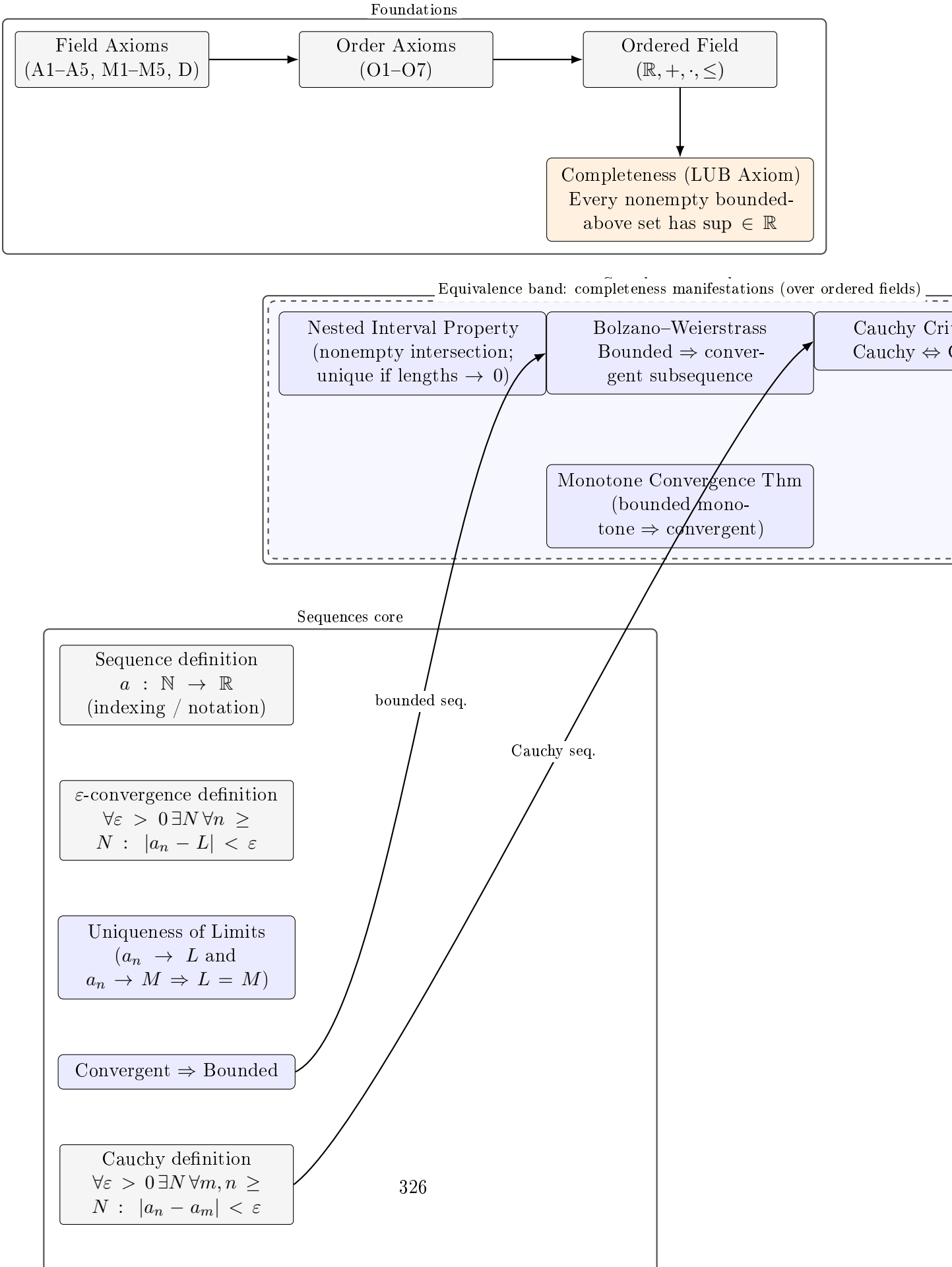
- completeness of $\mathbb{R}$,
- limsup existence,
- root test,
- absolute convergence theory.

Thus power series are a structural synthesis of the entire sequence and series development.

Sequence Applications Quick Reference	
Core items	Key definitions/results introduced in this file.
How to use	Read the boxed items first; proofs and consequences follow.
Dependencies	Refer back to earlier sections as needed.

Theorems1 Quick Reference	
Core items	Key definitions/results introduced in this file.
How to use	Read the boxed items first; proofs and consequences follow.
Dependencies	Refer back to earlier sections as needed.


Supporting asymptotics
Growth & Asymptotics (ratio/root, $o(\cdot), \sim$, Stolz–Cesàro)
Series layer
Series tests (comparison/limit comp./ratio/root/condensation/...)
Sequence convergence (εdefinition)
Partial sums $s_N = \sum_{n=1}^{N} a_n$
Series convergence $\sum a_n$ converges $\Leftrightarrow$ s_N converges
Absolute convergence $\sum |a_n|$ converges
Conditional convergence ($\sum a_n$ conv., $\sum |a_n|$ div.)
Rearrangements (stable if abs. conv.)
Cauchy product (when justified)
Power series layer
Power series $\sum_{n=0}^{\infty} c_n (x-a)^n$
Radius of convergence (Radius theorem)
Cauchy–Hadamard $\frac{1}{R} = \limsup \sqrt[n]{|c_n|}$
Absolute convergence for $|x-a| < R$
Uniform convergence on compact subsets $|x-a| \le r < R$
Term-by-term differentiation/integration (when justified)


Foundations
Field Axioms (A1–A5, M1–M5, D)
Order Axioms (O1–O7)
Ordered Field $(\mathbb{R}, +, \cdot, \leq)$
Completeness (LUB Axiom) Every nonempty bounded-above set has sup $\in \mathbb{R}$
Equivalence band: completeness manifestations (over ordered fields)
Nested Interval Property (nonempty intersection; unique if lengths $\to 0$)
Bolzano–Weierstrass Bounded $\Rightarrow$ convergent subsequence
Cauchy Cri
Cauchy $\Leftrightarrow$ C
Monotone Convergence Thm (bounded monotone $\Rightarrow$ convergent)
Sequences core
Sequence definition $a : \mathbb{N} \to \mathbb{R}$ (indexing / notation)
ε-convergence definition $\forall \varepsilon > 0 \, \exists N \, \forall n \geq N : |a_n - L| < \varepsilon$
Uniqueness of Limits ($a_n \to L$ and $a_n \to M \Rightarrow L = M$)
Convergent $\Rightarrow$ Bounded
Cauchy definition $\forall \varepsilon > 0 \, \exists N \, \forall m, n \geq N : |a_n - a_m| < \varepsilon$
bounded seq.
Cauchy seq.

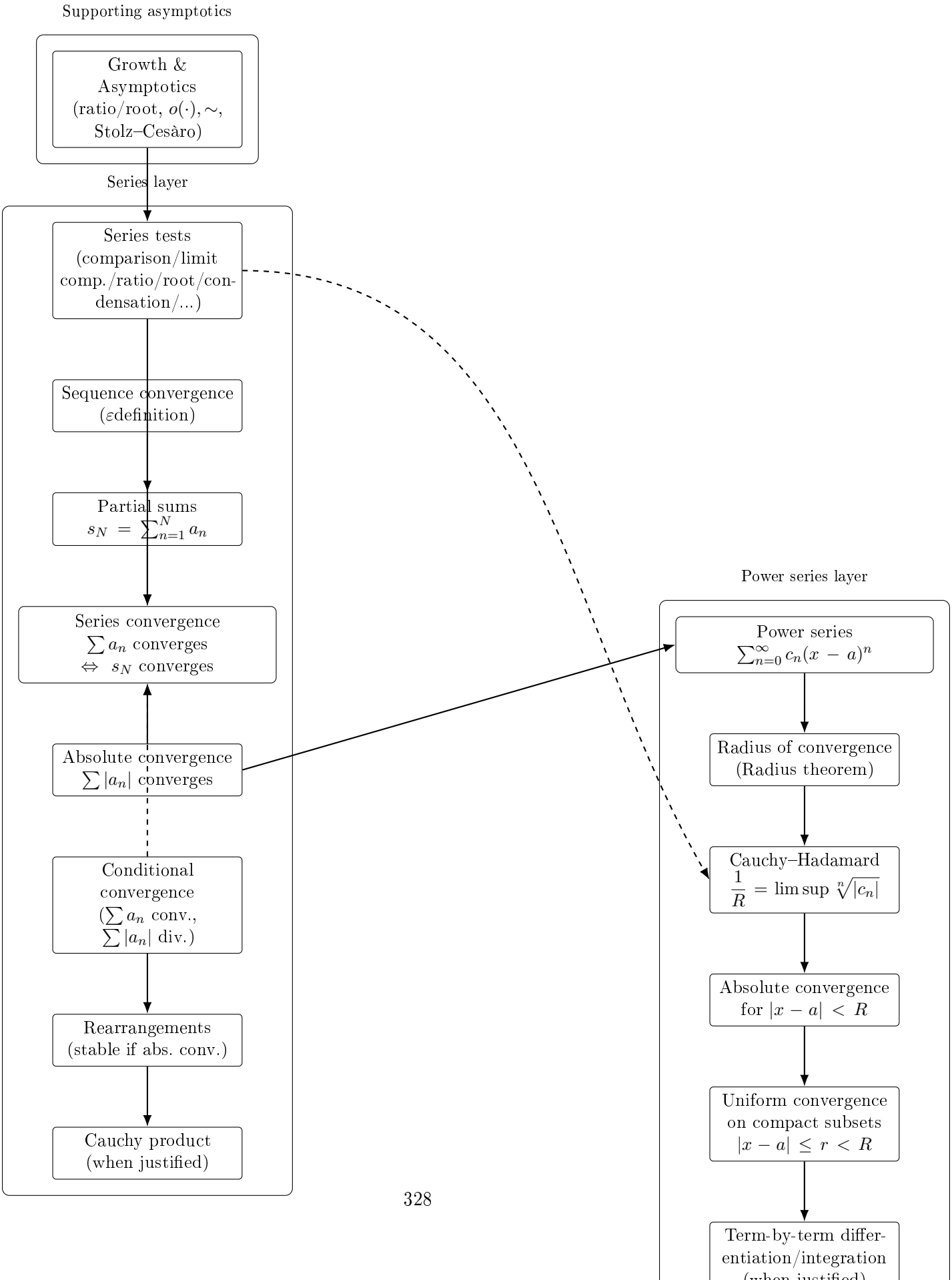
Theorems2 Quick Reference

Core items	Key definitions/results introduced in this file.
How to use	Read the boxed items first; proofs and consequences follow.
Dependencies	Refer back to earlier sections as needed.

13.1.18.4 Youtube proof example

Youtube Proof Quick Reference

Core items	Key definitions/results introduced in this file.
How to use	Read the boxed items first; proofs and consequences follow.
Dependencies	Refer back to earlier sections as needed.


Supporting asymptotics
Growth & Asymptotics (ratio/root, $o(\cdot), \sim$, Stolz–Cesàro)
Series layer
Series tests (comparison/limit comp./ratio/root/condensation/...)
Sequence convergence (εdefinition)
Partial sums $s_N = \sum_{n=1}^{N} a_n$
Series convergence $\sum a_n$ converges $\Leftrightarrow$ s_N converges
Absolute convergence $\sum |a_n|$ converges
Conditional convergence ($\sum a_n$ conv., $\sum |a_n|$ div.)
Rearrangements (stable if abs. conv.)
Cauchy product (when justified)
Power series layer
Power series $\sum_{n=0}^{\infty} c_n (x-a)^n$
Radius of convergence (Radius theorem)
Cauchy–Hadamard $\frac{1}{R} = \limsup \sqrt[n]{|c_n|}$
Absolute convergence for $|x-a| < R$
Uniform convergence on compact subsets $|x-a| \le r < R$
Term-by-term differentiation/integration

Real Analysis Sequences Completion Checklist

I. Real Number Foundations

Field & Order Structure

Path Quick Reference	
Core items	Key definitions/results introduced in this file.
How to use	Read the boxed items first; proofs and consequences follow.
Dependencies	Refer back to earlier sections as needed.

- ☐ Field axioms of $\mathbb{R}$
- ☐ Order axioms of $\mathbb{R}$
- ☐ Trichotomy Law
- ☐ Compatibility of order with addition
- ☐ Compatibility of order with multiplication

Absolute Value

- ☐ Definition of absolute value
- ☐ $|x| \geq 0$ and $|x| = 0 \iff x = 0$
- ☐ $|xy| = |x||y|$
- ☐ Triangle inequality
- ☐ Reverse triangle inequality

Suprema & Infima

- ☐ Definition of upper bound
- ☐ Definition of lower bound
- ☐ Definition of supremum
- ☐ Definition of infimum
- ☐ Supremum is unique
- ☐ Completeness (Least Upper Bound Property)

Archimedean Property

- ☐ Archimedean property statement
- ☐ Equivalent forms
- ☐ Proof that $\frac{1}{n} \to 0$

II. Basic Sequence Theory

Definitions

- ☐ Definition of sequence ($\mathbb{N} \to \mathbb{R}$)
- ☐ Definition of convergence (εN form)
- ☐ Negation of convergence

Fundamental Theorems

- ☐ Uniqueness of limits
- ☐ Convergent $\Rightarrow$ bounded

Algebra of Limits

- ☐ Sum rule
- ☐ Scalar multiple rule
- ☐ Product rule
- ☐ Quotient rule (nonzero limit)

Order Limit Theorems

- ☐ Limit preserves inequalities
- ☐ Squeeze theorem

III. Structural Sequence Theory

Monotone Sequences

- ☐ Definition of monotone increasing

- ☐ Definition of monotone decreasing
- ☐ Monotone Convergence Theorem

Subsequences

- ☐ Definition of subsequence
- ☐ n_k strictly increasing
- ☐ Subsequence of convergent sequence converges to same limit
- ☐ Subsequence of subsequence lemma
- ☐ Index growth fact: $n_k \geq k$

Bolzano–Weierstrass

- ☐ Every bounded sequence has a convergent subsequence

IV. Cauchy Theory & Completeness

Cauchy Sequences

- ☐ Definition of Cauchy sequence
- ☐ Convergent $\Rightarrow$ Cauchy
- ☐ Cauchy $\Rightarrow$ bounded
- ☐ Cauchy $\Rightarrow$ convergent (Completeness of $\mathbb{R}$)

V. Limit Superior / Limit Inferior

- ☐ Definition of tail set
- ☐ Definition of $s_n = \sup_{k \geq n} a_k$
- ☐ Definition of $\limsup a_n$
- ☐ Definition of $\liminf a_n$
- ☐ $\limsup$ always exists (possibly $\pm\infty$)
- ☐ $\liminf$ always exists
- ☐ $\liminf \leq \limsup$
- ☐ Convergence $\iff \limsup = \liminf$

- ☐ $\limsup$ equals largest subsequential limit
- ☐ $\liminf$ equals smallest subsequential limit

VI. Subsequence Toolkit (Advanced)

- ☐ Finite Partition Convergence Principle
- ☐ Residue class convergence
- ☐ Even/odd convergence principle
- ☐ Dense subsequence criterion
- ☐ Diagonal subsequence lemma
- ☐ Inherited vs. reflected properties
- ☐ Tail properties
- ☐ Universal subsequence properties

Final Mastery Check

- ☐ Prove Monotone Convergence from completeness
- ☐ Prove Bolzano–Weierstrass
- ☐ Prove Cauchy $\iff$ Convergent in $\mathbb{R}$
- ☐ Extract convergent subsequences intentionally
- ☐ Compute $\limsup$ and $\liminf$ in nontrivial examples
- ☐ Characterize convergence via $\limsup$ / $\liminf$

Chapter 14

Introduction to Metric Spaces

14.1 Notes

Where You Are in the Journey

Propositional Logic $\to$ Predicate Calculus $\to$ Sets & Functions $\to$ $\mathbb{R}$, Real Analysis $\to$ Algebraic Structures $\to$ Metric Spaces $\to$ Topology $\to \cdots$

How we got here. Real analysis developed convergence, Cauchy sequences, and completeness specifically for $\mathbb{R}$, using $|x - y|$ as the distance. Metric spaces ask: which of those theorems depend only on the properties of distance, not on the specific structure of $\mathbb{R}$?

What this chapter builds. We define metric spaces axiomatically, then port the key analytic concepts — open and closed sets, sequences and limits, completeness, compactness, and connectedness — into this abstract setting. The triangle inequality is the engine driving every result.

Where this leads. Topology abstracts further by replacing the metric with a collection of open sets. Functional analysis studies metric spaces of functions. Every normed vector space is a metric space.

14.1.1 Metric Spaces

Structural Roadmap

The development of metric space theory follows the same definition–theorem–structure architecture used in the real line foundations, but now in a fully abstract setting.

Each major topic is organized as:

$$\text{Definitions} \longrightarrow \text{Main Theorems} \longrightarrow \text{Consequences and Logical Structure}$$

The global progression is:

1. Definition of a metric space
2. Open balls
3. Convergence in metric spaces
4. Cauchy sequences and completeness
5. Compactness and sequential compactness
6. The real line as a complete metric space

Remark 14.1. This section abstracts the distance structure previously used on $\mathbb{R}$:

$$d(x, y) = |x - y|.$$

The real line now becomes a special case of a metric space.

Remark 14.2. Metric space theory isolates the notion of distance from order structure and supremum-based completeness. It allows convergence and Cauchy behavior to be studied independently of algebraic or order properties.

14.1.2 Real Numbers as a Metric Space

14.1.2.1 Basic Definitions

Definition 14.3 (Modulus / Absolute value). The modulus (absolute value) is the function

$$|\cdot| : \mathbb{R} \to \mathbb{R}$$

defined by

$$|x| := \begin{cases} x, & x \geq 0, \\ -x, & x < 0. \end{cases}$$

Definition 14.4 (Distance function / Metric). Let X be a nonempty set. A function $d : X \times X \to \mathbb{R}$ is called a metric on X if for all $x, y, z \in X$:

$$\begin{aligned} d(x, y) &\geq 0, \\ d(x, y) = 0 &\iff x = y, \\ d(x, y) &= d(y, x), \\ d(x, z) &\leq d(x, y) + d(y, z). \end{aligned}$$

The pair (X, d) is called a metric space.

Definition 14.5 (Euclidean norm on $\mathbb{R}^n$). For $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$, the Euclidean norm is defined by

$$\|x\|_2 := \left(\sum_{i=1}^{n} x_i^2 \right)^{1/2}.$$

Definition 14.6 (ℓ^p metrics). Let $1 \le p < \infty$. For $x, y \in \mathbb{R}^n$, define

$$d_p(x,y) = \left(\sum_{i=1}^{n} |x_i - y_i|^p \right)^{1/p}.$$

For $p = \infty$, define

$$d_\infty(x,y) = \max_{1 \le i \le n} |x_i - y_i|.$$

Remark 14.7. Throughout these notes, $|\cdot|$ denotes absolute value on $\mathbb{R}$, while $\|\cdot\|_2$ denotes the Euclidean norm on $\mathbb{R}^n$.

14.1.2.2 Main Theorems

Theorem 14.8 (Properties of the modulus). For all $x, y \in \mathbb{R}$:

$$\begin{aligned} |x| &\ge 0, \\ |x| = 0 &\iff x = 0, \\ |-x| &= |x|, \\ |xy| &= |x|\,|y|, \\ |x+y| &\le |x| + |y|, \\ \big||x| - |y|\big| &\le |x-y|, \\ -|x| &\le x \le |x|. \end{aligned}$$

Remark 14.9. The triangle inequality is the key property that allows $|\cdot|$ to induce a metric:

$$d(x,y) = |x-y|.$$

14.1.2.3 Canonical Examples

Example 14.10 (Discrete metric). Let X be any nonempty set. Define

$$d(x,y) := \begin{cases} 0, & x = y, \\ 1, & x \ne y. \end{cases}$$

Then d is a metric on X.

Example 14.11 (Taxicab (Manhattan) metric). On $\mathbb{R}^n$, define

$$d_1(x,y) := \sum_{i=1}^{n} |x_i - y_i|.$$

Example 14.12 (London railway metric). Fix a base point $0 \in \mathbb{R}^n$. Define

$$d_L(x,y) := \begin{cases} \|x\|_2 + \|y\|_2, & x \ne y, \\ 0, & x = y. \end{cases}$$

Then d_L is a metric.

14.1.2.4 Consequences

Remark 14.13 (Geometric dependence on metric). Let $(\mathbb{R}^2, d)$ be equipped with a norm $\|\cdot\|$. The unit sphere is

$$S = \{x \in \mathbb{R}^2 : \|x\| = 1\}.$$

Changing the norm changes the geometry:

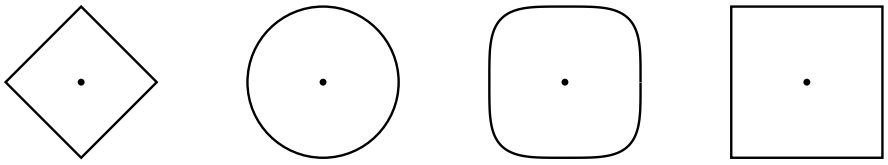
- ℓ^2 circle
- ℓ^1 diamond
- ℓ^∞ square
- $1 < p < \infty$ smooth interpolation between diamond and square

Although geometry changes, the logical structure of convergence in metric spaces depends only on the metric axioms.

Remark 14.14 (Logical Structure).

$$\text{Absolute Value} \Rightarrow \text{Triangle Inequality} \Rightarrow \text{Metric} \Rightarrow \text{Open Balls} \Rightarrow \text{Convergence Definition}$$

14.1.2.5 Geometric Illustration



Taxicab ℓ^1 Euclidean ℓ^2 ℓ^4 Max ℓ^∞

Remark 14.15 (Structural Position). Metric spaces are introduced only after sequences, convergence, and completeness have been fully understood in $\mathbb{R}$.

This ensures that abstraction follows mastery rather than preceding it.

Chapter 15

Introduction to Topology

15.1 Notes

Where You Are in the Journey

Propositional Logic $\to$ Sets & Functions $\to$ $\mathbb{R}$, Real Analysis $\to$ Algebraic Structures $\to$ Metric Spaces $\to$ Point-Set Topology $\to$ Measure Theory $\to$ $\cdots$

How we got here. Metric spaces showed us that convergence and continuity can be defined without coordinates — only distance matters. Point-set topology takes the final abstraction step: it discards distance altogether, retaining only the family of "open sets" as the primitive notion.

What this chapter builds. We define topological spaces axiomatically, study the open/closed set framework, and develop continuity, compactness, and connectedness in full generality. The Hausdorff, second-countable, and metric conditions are studied as special cases that recover familiar properties.

Where this leads. Measure theory requires a σ-algebra of measurable sets, which is a cousin of the topology. Differential geometry and algebraic topology build on the topological foundations established here.

15.1.1 Introduction to Topology

Structural Roadmap

The development of topology follows the same definition–theorem–structure architecture used throughout this project, but now abstracts the notion of neighborhood structure independently of distance.

Each major topic is organized as:

Definitions $\longrightarrow$ Main Theorems $\longrightarrow$ Consequences and Logical Structure

The global progression is:

1. Definition of a topological space
2. Open and closed sets
3. Interior, closure, and boundary
4. Accumulation (limit) points
5. Subspace topology
6. Compactness via open covers
7. Sequential vs. open-cover compactness

Remark 15.1. Topology isolates the concept of open sets and neighborhood structure without reference to algebra, order, or distance.

Remark 15.2. Many notions previously defined using sequences (convergence, compactness, limit points) are reinterpreted here in terms of open sets.

15.1.2 Topology of the Real Metric Space

15.1.2.1 Basic Definitions

Remark 15.3 (Why topology appears here). The metric space and topological definitions introduced in this section — open sets, closed sets, neighborhoods, limit points, and compactness — are stated in their natural generality, but their real force only becomes visible once sequences and limits are in hand.

In particular:

- A point x is a limit point of a set A if and only if some sequence in $A \setminus \{x\}$ converges to x.
- A set F is closed if and only if it contains the limits of all convergent sequences in F.
- A set K is sequentially compact if and only if every sequence in K has a subsequence converging to a point in K. By the Heine–Borel theorem, this coincides with compactness for subsets of $\mathbb{R}^n$.

These equivalences are not incidental — they reveal that the topological language and the sequential language are two dialects describing the same structure. The sequence-based characterizations are generally easier to work with in proofs, while the open-set definitions generalize more cleanly to spaces where sequences are insufficient (e.g. uncountable products).

The results that follow — convergence, Cauchy sequences, Bolzano–Weierstrass, and subsequence theory — should be read as building the sequential side of this correspondence. The topological interpretations will be noted where they arise.

Definition 15.4 (ε-neighborhood). Let $x_0 \in \mathbb{R}$ and $\varepsilon > 0$.

$$N_\varepsilon(x_0) := \{x \in \mathbb{R} : |x - x_0| < \varepsilon\}.$$

Definition 15.5 (Open ball). Let (X, d) be a metric space.

$$B_\varepsilon(x_0) := \{x \in X : d(x, x_0) < \varepsilon\}.$$

Definition 15.6 (Closed ball).

$$\overline{B}_\varepsilon(x_0) := \{x \in X : d(x, x_0) \leq \varepsilon\}.$$

Definition 15.7 (Open set). A set $U \subseteq X$ is open if

$$\forall x \in U \; \exists \varepsilon > 0 \quad \text{such that} \quad B_\varepsilon(x) \subseteq U.$$

Definition 15.8 (Closed set). A set $F \subseteq X$ is closed if $X \setminus F$ is open.

Definition 15.9 (Open cover). A family $\{U_\alpha\}_{\alpha \in I}$ of open sets is an open cover of $A \subseteq X$ if

$$A \subseteq \bigcup_{\alpha \in I} U_\alpha.$$

Definition 15.10 (Closure).

$$\overline{A} = \bigcap \{F \subseteq X : F \text{ closed and } A \subseteq F\}.$$

Definition 15.11 (Interior).

$$A^\circ = \bigcup \{U \subseteq X : U \text{ open and } U \subseteq A\}.$$

Definition 15.12 (Limit point).

$$\forall \varepsilon > 0, \quad (B_\varepsilon(x) \setminus \{x\}) \cap A \neq \varnothing.$$

Definition 15.13 (Compact set). Every open cover admits a finite subcover.

Definition 15.14 (Sequential compactness). Every sequence in K has a convergent subsequence whose limit lies in K.

Definition 15.15 (Bounded set).

$$A \subseteq B_R(x_0) \quad \text{for some } x_0, R.$$

15.1.2.2 Main Theorems

Theorem 15.16 (Neighborhood = ball in $\mathbb{R}$). With $d(x, y) = |x - y|$,

$$N_\varepsilon(x_0) = B_\varepsilon(x_0).$$

Theorem 15.17 (Ball characterization of closure).

$$x \in \overline{A} \iff \forall \varepsilon > 0, \; B_\varepsilon(x) \cap A \neq \varnothing.$$

Theorem 15.18 (Ball characterization of interior).

$$x \in A^\circ \iff \exists \varepsilon > 0,\ B_\varepsilon(x) \subseteq A.$$

Theorem 15.19 (Heine–Borel Theorem on $\mathbb{R}^n$). For $K \subseteq \mathbb{R}^n$, the following are equivalent:

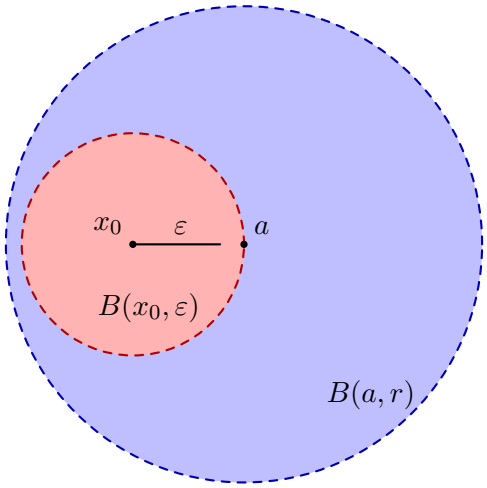
1. K compact
2. K sequentially compact
3. Every infinite subset has a limit point in K
4. K closed and bounded

Implication cycle:

$$(1) \Rightarrow (2) \Rightarrow (3) \Rightarrow (4) \Rightarrow (1).$$

Remark 15.20. Closed and bounded implies compact only in finite-dimensional spaces.

15.1.2.3 Geometric Illustration



15.1.2.4 Consequences

Remark 15.21 (Logical Structure).

$$\text{Metric} \Rightarrow \text{Open Balls} \Rightarrow \text{Open Sets} \Rightarrow \text{Closure / Interior} \Rightarrow \text{Compactness}$$

Remark 15.22. In $\mathbb{R}^n$, compactness is equivalent to closed and bounded. This depends critically on completeness.

Remark 15.23 (Structural Position). Topology is introduced after metric spaces to separate the idea of neighborhood structure from the notion of distance.

This abstraction allows compactness, continuity, and convergence to be studied in their most general setting.

15.2 Proofs

15.3 Capstone

Chapter 16

Measure Theory

Where You Are in the Journey

$\mathbb{R}$, Real Analysis $\to$ Set Algebras $\to$ Metric Spaces $\to$ Topology $\to$ Measure Theory $\to$ Functional Analysis $\to \cdots$

How we got here. Real analysis defined the Riemann integral for nice functions, but many naturally arising functions are not Riemann integrable. Measure theory builds a more powerful integration theory by replacing the interval-based approach with an abstract notion of "size" (measure) on arbitrary sets.

What this chapter will build. σ-algebras, measures (including Lebesgue measure), measurable functions, the Lebesgue integral, and the convergence theorems (monotone convergence, dominated convergence, Fatou's lemma).

Where this leads. L^p spaces are the natural setting for functional analysis. Probability theory is measure theory applied to probability spaces.

Status: Planned

Coming Soon

Notes, proofs, and exercises will appear here in a future revision.

16.1 Notes

To be populated.

16.2 Proofs

To be populated.

16.3 Capstone

To be populated.

Algebra

Chapter 17

Introduction to Algebraic Structures

17.1 Notes

Where You Are in the Journey

Propositional Logic → Predicate Calculus → Sets & Functions → Proof Techniques → Real Analysis → Intro to Algebraic Structures → Linear Algebra → Topology → ⋯

How we got here. Sets and functions gave us the language of mathematical structure, and proof techniques gave us the tools to reason about it. Real analysis showed what rigorous reasoning looks like on a concrete number system. Algebraic structures now ask the deeper question: what is the minimal set of axioms needed for the most important algebraic phenomena — identity, inverses, and arithmetic — to occur?

What this chapter builds. We construct the hierarchy of algebraic structures: groups, rings, and fields. Each is defined by progressively richer axiom systems. Fields are the scalars over which vector spaces are built, and groups are the additive backbone of every vector space.

Where this leads. Linear algebra inherits the field and group axioms directly: a vector space is an abelian group equipped with a scalar multiplication by a field, and every proof about vector spaces draws on both simultaneously.

Structural Roadmap

The development of algebraic structures in this project follows the definition–theorem–structure architecture used throughout the analysis volumes.

The primary driver is Contemporary Abstract Algebra by Joseph Gallian. The emphasis is on axiom systems, structural consequences, and the hierarchy of algebraic objects rather than computational techniques.

Each major topic is organized as:

Definitions $\longrightarrow$ Main Theorems $\longrightarrow$ Consequences and Structural Insight

The global progression is:

1. Groups and abelian groups
2. Rings and integral domains
3. Fields and their properties
4. Bridge to linear algebra

Remark 17.1 (Structural Position). The structures developed here are not studied for their own sake alone. Groups provide the additive structure of every vector space. Fields provide the scalars. The interaction between them is the subject of linear algebra.

Remark 17.2 (Dependency Note). The uniqueness theorems proved here — uniqueness of identity, uniqueness of inverses — are the same theorems invoked in vector space proofs. They are proved once here and cited by theorem number thereafter.

17.1.0.1 Groups

Binary Operations

Definition 17.3 (Binary Operation). Let G be a set. A binary operation on G is a function

$$\star : G \times G \to G.$$

For $a, b \in G$, we write $a \star b$ for the image of (a, b) under $\star$.

Remark 17.4. The codomain of $\star$ is G itself. This means that for any $a, b \in G$, the result $a \star b$ is again an element of G. This property is called closure and is built into the definition of a binary operation.

Intuitively: a binary operation takes two elements of a set and produces a third element of the same set.

Definition 17.5 (Associativity). A binary operation $\star$ on G is associative if

$$(a \star b) \star c = a \star (b \star c) \quad \text{for all } a, b, c \in G.$$

Definition 17.6 (Commutativity). A binary operation $\star$ on G is commutative if

$$a \star b = b \star a \quad \text{for all } a, b \in G.$$

Remark 17.7. Associativity and commutativity are independent properties. Addition on $\mathbb{Z}$ is both; matrix multiplication is associative but not commutative; subtraction on $\mathbb{Z}$ is neither.

Definition of a Group

Definition 17.8 (Group). A group is a pair $(G, \star)$ where G is a set and $\star$ is a binary operation on G satisfying the following axioms:

G1. Associativity. $(a \star b) \star c = a \star (b \star c)$ for all $a, b, c \in G$.

G2. Identity. There exists an element $e \in G$ such that

$$e \star a = a \star e = a \quad \text{for all } a \in G.$$

The element e is called the identity element of G.

G3. Inverses. For each $a \in G$, there exists an element $a^{-1} \in G$ such that

$$a \star a^{-1} = a^{-1} \star a = e.$$

The element a^{-1} is called the inverse of a.

Remark 17.9. Closure is not listed as a separate axiom because it is already encoded in the requirement that $\star : G \times G \to G$ is a binary operation — the codomain forces the result to stay in G.

Intuitively: a group is a set where you can combine elements, undo combinations, and the order of grouping never matters.

Remark 17.10 (Axiom Count). Some treatments list four axioms (closure, associativity, identity, inverses). Here closure is absorbed into the definition of binary operation, leaving three axioms. Both presentations define the same object.

Remark 17.11 (Notation). When the operation is understood from context, we write ab instead of $a \star b$, and call G itself a group rather than the pair $(G, \star)$. For groups whose operation is addition, we write $a + b$, use 0 for the identity, and $-a$ for the inverse of a.

Definition 17.12 (Order of a Group). The order of a group G, denoted $|G|$, is the cardinality of G as a set. If $|G|$ is finite, G is called a finite group; otherwise it is an infinite group.

Example 17.13 (Canonical Examples of Groups). (i) $(\mathbb{Z}, +)$: the integers under addition. Identity: 0. Inverse of n: $-n$. Infinite group.

(ii) $(\mathbb{Q} \setminus \{0\}, \cdot)$: nonzero rationals under multiplication. Identity: 1. Inverse of q: $1/q$. Infinite group.

(iii) $(\mathbb{Z}/n\mathbb{Z}, +)$: integers modulo n under addition. Identity: $[0]$. Inverse of $[k]$: $[n - k]$. Finite group of order n.

(iv) $(GL_n(\mathbb{R}), \cdot)$: invertible $n \times n$ real matrices under multiplication. Identity: I_n. Inverse: matrix inverse. Infinite group.

Remark 17.14. Note what fails to be a group: $(\mathbb{Z}, \cdot)$ is not a group because 2 has no multiplicative inverse in $\mathbb{Z}$. $(\mathbb{N}, +)$ is not a group because positive integers have no additive inverse in $\mathbb{N}$. These failures illustrate why each axiom is necessary.

Basic Theorems

Remark 17.15 (Why These Theorems Matter). The group axioms guarantee existence of an identity and inverses, but say nothing about uniqueness. The following theorems establish that both are unique. This is essential: without uniqueness, we cannot speak of the identity or the inverse of an element, and proofs that equate two objects via the identity or inverse would be invalid.

These same uniqueness theorems reappear in vector space proofs, where they are cited by name. They are proved once here.

Proposition 17.16 (Uniqueness of the Identity). Let G be a group. The identity element of G is unique.

Remark 17.17. The proof strategy is standard for uniqueness arguments: assume two identities exist, then show they must be equal. This pattern recurs throughout algebra whenever a definition asserts existence of a distinguished element.

Intuitively: if two elements both act as an identity, applying one to the other forces them to coincide.

Proposition 17.18 (Uniqueness of Inverses). Let G be a group. For each $a \in G$, the inverse of a is unique.

Remark 17.19. Intuitively: if two elements both undo a, then they must be the same element — because each can be obtained from the other by cancellation.

Proposition 17.20 (Cancellation Laws). Let G be a group and let $a, b, c \in G$. Then:

(i) Left cancellation: $ab = ac \implies b = c$.

(ii) Right cancellation: $ba = ca \implies b = c$.

Remark 17.21. Cancellation is what makes group equations solvable. It does not hold in general for rings or monoids without inverses.

Intuitively: multiply both sides by a^{-1} and the common factor disappears.

Proposition 17.22 (Socks-Shoes Property). Let G be a group and let $a, b \in G$. Then

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

Remark 17.23. The name comes from the observation that to undo putting on socks then shoes, you must first remove the shoes, then the socks — in reverse order.

This reversal of order is characteristic of non-abelian groups and becomes important in the theory of group homomorphisms and in matrix algebra, where $(AB)^{-1} = B^{-1}A^{-1}$.

Abelian Groups

Definition 17.24 (Abelian Group). A group $(G, \star)$ is called abelian (or commutative) if

$$a \star b = b \star a \quad \text{for all } a, b \in G.$$

Remark 17.25. An abelian group is a group with one additional axiom: commutativity of the operation. Every abelian group is a group, but not every group is abelian.

Intuitively: in an abelian group, the order in which you combine elements is irrelevant. This is the familiar arithmetic of addition on $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$, and $\mathbb{C}$.

Remark 17.26 (Structural Position). Abelian groups are the additive backbone of every vector space. The four vector space axioms governing addition — associativity, commutativity, existence of zero, existence of additive inverses — are precisely the axioms that make $(V, +)$ an abelian group.

This is why the vector space definition can be stated compactly as: a vector space over $\mathbb{F}$ is an abelian group $(V, +)$ equipped with a scalar multiplication by $\mathbb{F}$. The abelian group structure is not an analogy; it is the literal algebraic content of the first four vector space axioms.

Example 17.27 (Abelian and Non-Abelian Groups). (i) $(\mathbb{Z}, +)$, $(\mathbb{R}, +)$, $(\mathbb{C}, +)$: all abelian. These are the additive groups underlying the standard vector spaces.

(ii) $(\mathbb{Z}/n\mathbb{Z}, +)$: abelian for all $n \geq 1$.

(iii) $(GL_n(\mathbb{R}), \cdot)$ for $n \geq 2$: not abelian, since matrix multiplication does not commute in general.

(iv) $(S_n, \circ)$ for $n \geq 3$: the symmetric group on n symbols under composition is not abelian.

Remark 17.28 (Additive Notation Convention). For abelian groups, it is standard to write the operation as $+$, the identity as 0, and the inverse of a as $-a$. This additive notation is used throughout linear algebra, where $(V, +)$ is always an abelian group.

17.1.0.2 Rings

Definition of a Ring

Remark 17.29 (Motivation). A group has one binary operation. A ring has two: addition and multiplication. The addition makes the ring an abelian group. Multiplication is layered on top, connected to addition through the distributive laws.

Rings are the natural algebraic home of arithmetic. The integers $\mathbb{Z}$, polynomials $\mathbb{Z}[x]$, and square matrices $M_n(\mathbb{R})$ are all rings. What they share is not the specific objects but the axiom structure — and every theorem proved here holds simultaneously in all of them.

Definition 17.30 (Ring). A ring is a triple $(R, +, \cdot)$ where R is a set and $+$ and $\cdot$ are binary operations on R satisfying:

R1. Additive abelian group. $(R, +)$ is an abelian group:

- $(a + b) + c = a + (b + c)$ for all $a, b, c \in R$.
- $a + b = b + a$ for all $a, b \in R$.
- There exists $0 \in R$ such that $a + 0 = a$ for all $a \in R$.
- For each $a \in R$, there exists $-a \in R$ such that $a + (-a) = 0$.

R2. Multiplicative associativity. $(a \cdot b) \cdot c = a \cdot (b \cdot c)$ for all $a, b, c \in R$.

R3. Distributivity.

- $a \cdot (b + c) = a \cdot b + a \cdot c$ for all $a, b, c \in R$.
- $(a + b) \cdot c = a \cdot c + b \cdot c$ for all $a, b, c \in R$.

Remark 17.31 (What a Ring Does and Does Not Require). A ring does not require:

- commutativity of multiplication ($ab = ba$ need not hold),
- a multiplicative identity (1 need not exist),
- multiplicative inverses (a^{-1} need not exist).

Each additional requirement produces a richer structure. A ring with a multiplicative identity is a ring with unity. A commutative ring with unity where every nonzero element has a multiplicative inverse is a field — covered in the next section.

Intuitively: a ring is the minimal structure needed for addition, subtraction, and multiplication to coexist and interact sensibly. Division is not guaranteed.

Definition 17.32 (Commutative Ring). A ring $(R, +, \cdot)$ is commutative if

$$a \cdot b = b \cdot a \quad \text{for all } a, b \in R.$$

Definition 17.33 (Ring with Unity). A ring $(R, +, \cdot)$ is a ring with unity if there exists $1 \in R$ such that

$$1 \cdot a = a \cdot 1 = a \quad \text{for all } a \in R.$$

The element 1 is called the multiplicative identity or unity. When it exists, it is unique (proof identical to Proposition 17.16 applied to $(R, \cdot)$).

Example 17.34 (Canonical Examples of Rings). (i) $(\mathbb{Z}, +, \cdot)$: integers. Commutative ring with unity 1. No multiplicative inverses for $|n| \neq 1$.

(ii) $(\mathbb{Z}/n\mathbb{Z}, +, \cdot)$: integers modulo n. Commutative ring with unity [1].

(iii) $(M_n(\mathbb{R}), +, \cdot)$: $n \times n$ real matrices. Ring with unity I_n. Not commutative for $n \geq 2$.

(iv) $(\mathbb{Z}[x], +, \cdot)$: polynomials with integer coefficients. Commutative ring with unity 1.

(v) The trivial ring $\{0\}$, where $0 = 1$, is the only ring in which the additive and multiplicative identities coincide.

Basic Theorems

Remark 17.35 (What Needs Proving). The ring axioms say nothing explicitly about how multiplication interacts with the additive identity 0 or with additive inverses. These must be derived. The key tool in both proofs is distributivity — the bridge between the two operations.

Notice that these proofs make no assumption about which ring you are in. They work in $\mathbb{Z}$, in $M_n(\mathbb{R})$, in $\mathbb{Z}[x]$, in any ring simultaneously.

Proposition 17.36 (Multiplication by Zero). Let R be a ring. For all $a \in R$,

$$a \cdot 0 = 0 \cdot a = 0.$$

Remark 17.37. The proof uses distributivity to produce the equation $a \cdot 0 + a \cdot 0 = a \cdot 0$, then applies additive cancellation to conclude $a \cdot 0 = 0$.

This is not circular: 0 on the left side is the additive identity of the ring; the 0 on the right is the same element being derived as a consequence of cancellation. The proof works because additive cancellation was already proved for all abelian groups (Proposition 17.20).

Proposition 17.38 (Multiplication by Additive Inverse). Let R be a ring. For all $a, b \in R$,

$$a \cdot (-b) = -(a \cdot b) \quad \text{and} \quad (-a) \cdot b = -(a \cdot b).$$

In any ring with unity, $(-1) \cdot a = -a$.

Remark 17.39. The proof strategy is: show $a \cdot (-b)$ satisfies the defining property of the additive inverse of $a \cdot b$, then invoke uniqueness of additive inverses (Proposition 17.18).

This is the same strategy used in the vector space proof that $(-1)v = -v$ — because that proof is just this theorem applied to the scalar field $\mathbb{F}$ acting on V.

Integral Domains

Remark 17.40 (Motivation). In $\mathbb{Z}$, if $ab = 0$ then $a = 0$ or $b = 0$. This feels obvious but it is not an axiom of rings — it fails in $\mathbb{Z}/6\mathbb{Z}$, where $[2] \cdot [3] = [0]$ even though neither $[2]$ nor $[3]$ is zero. Elements that behave like $[2]$ and $[3]$ are called zero divisors, and rings without them are integral domains.

This property matters because it is exactly what is needed for multiplicative cancellation — and for the zero product argument that appears throughout linear algebra.

Definition 17.41 (Zero Divisor). Let R be a commutative ring with unity. A nonzero element $a \in R$ is a zero divisor if there exists a nonzero $b \in R$ such that $a \cdot b = 0$.

Definition 17.42 (Integral Domain). A commutative ring with unity R is an integral domain if R has no zero divisors. Equivalently,

$$a \cdot b = 0 \implies a = 0 \text{ or } b = 0 \quad \text{for all } a, b \in R.$$

Remark 17.43 (Connection to Linear Algebra). The zero product property is the theorem cited in the vector space proof that $av = \mathbf{0} \Rightarrow a = 0$ or $v = \mathbf{0}$. Specifically, the case $a \neq 0$ uses the fact that the scalar field $\mathbb{F}$ has no zero divisors — which holds because every field is an integral domain (proved in the next section).

The proof is not about vectors at all. It is about the scalar field.

Proposition 17.44 (Cancellation in Integral Domains). Let R be an integral domain and $a, b, c \in R$ with $a \neq 0$. Then

$$ab = ac \implies b = c.$$

Example 17.45 (Integral Domains and Non-Examples). (i) $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$: all integral domains.

(ii) $\mathbb{Z}[x]$: integral domain.

(iii) $\mathbb{Z}/p\mathbb{Z}$ for prime p: integral domain (in fact a field, as shown in the next section).

(iv) $\mathbb{Z}/6\mathbb{Z}$: not an integral domain, since $[2][3] = [0]$ with $[2], [3] \neq [0]$.

(v) $M_2(\mathbb{R})$: not an integral domain — not commutative, and admits zero divisors.

17.1.0.3 Fields

Definition of a Field

Remark 17.46 (Motivation). A ring allows addition, subtraction, and multiplication. A field adds division: every nonzero element has a multiplicative inverse. This is the structure of $\mathbb{Q}$, $\mathbb{R}$, and $\mathbb{C}$ — the number systems where you can always solve $ax = b$ for $a \neq 0$.

Fields are the scalars of linear algebra. Axler's $\mathbb{F}$ denotes an arbitrary field, meaning every theorem in linear algebra holds simultaneously over $\mathbb{R}$, $\mathbb{C}$, $\mathbb{Q}$, and any other field — because the proofs use only field axioms, not properties specific to real or complex numbers.

Definition 17.47 (Field). A field is a commutative ring with unity $(F, +, \cdot)$ satisfying:

F1. Additive abelian group. $(F, +)$ is an abelian group with identity 0.

F2. Multiplicative abelian group on nonzero elements. $(F \setminus \{0\}, \cdot)$ is an abelian group with identity 1. Explicitly: for each $a \in F$ with $a \neq 0$, there exists $a^{-1} \in F$ such that $a \cdot a^{-1} = 1$.

F3. Distributivity. $a \cdot (b + c) = a \cdot b + a \cdot c$ for all $a, b, c \in F$.

F4. Non-triviality. $0 \neq 1$.

Remark 17.48 (Unpacking the Definition). A field is simultaneously:

- $(F, +)$: an abelian group (additive structure),
- $(F \setminus \{0\}, \cdot)$: an abelian group (multiplicative structure),
- connected by distributivity.

The non-triviality axiom $0 \neq 1$ rules out the trivial ring $\{0\}$, which would otherwise technically satisfy the other axioms.

Example 17.49 (Fields). (i) $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$: the standard fields.

(ii) $\mathbb{Z}/p\mathbb{Z}$ for any prime p: a finite field with p elements, denoted $\mathbb{F}_p$.

(iii) $\mathbb{Q}(\sqrt{2}) = \{a + b\sqrt{2} : a, b \in \mathbb{Q}\}$: a field extending $\mathbb{Q}$.

Example 17.50 (Non-Fields). (i) $\mathbb{Z}$: not a field. The element 2 has no multiplicative inverse in $\mathbb{Z}$.

(ii) $\mathbb{Z}/6\mathbb{Z}$: not a field. $[2]$ has no inverse since $\gcd(2, 6) \neq 1$. (Also not an integral domain.)

(iii) $M_n(\mathbb{R})$ for $n \geq 2$: not a field. Not commutative, and singular matrices have no inverse.

Basic Theorems

Remark 17.51 (Why These Theorems Matter). These are the theorems cited in every vector space proof that crosses from the scalar side to the vector side. They live here, in the field, and are cited by number when needed in linear algebra.

Proposition 17.52 (Every Field is an Integral Domain). Every field is an integral domain.

Remark 17.53. This is the theorem behind the linear algebra argument: if $av = \mathbf{0}$ and $a \neq 0$, then $v = \mathbf{0}$. The scalar a lives in a field $\mathbb{F}$, and the key step is that $\mathbb{F}$ has no zero divisors.

Proposition 17.54 (Zero Product Property in a Field). Let $\mathbb{F}$ be a field and $a, b \in \mathbb{F}$. Then

$$ab = 0 \implies a = 0 \text{ or } b = 0.$$

Proposition 17.55 (Nonzero Scalars Have Inverses). Let $\mathbb{F}$ be a field and $a \in \mathbb{F}$ with $a \neq 0$. Then there exists a unique $a^{-1} \in \mathbb{F}$ such that $a \cdot a^{-1} = 1$.

Remark 17.56. This is the theorem cited in the vector space proof that $av = \mathbf{0}$ with $a \neq 0$ implies $v = \mathbf{0}$: the step "multiply both sides by a^{-1}" is valid precisely because a^{-1} exists and is unique.

Proposition 17.57 (Characteristic of a Field). Let $\mathbb{F}$ be a field. The characteristic of $\mathbb{F}$ is the smallest positive integer n such that

$$\underbrace{1 + 1 + \cdots + 1}_{n} = 0,$$

or 0 if no such n exists. The characteristic of a field is either 0 or a prime p.

Remark 17.58. $\mathbb{Q}$, $\mathbb{R}$, and $\mathbb{C}$ all have characteristic 0. $\mathbb{Z}/p\mathbb{Z}$ has characteristic p.

In Axler, $\mathbb{F}$ denotes $\mathbb{R}$ or $\mathbb{C}$, both of characteristic 0. Some results (such as the existence of eigenvalues) require characteristic 0 or algebraic closure and would fail over $\mathbb{F}_p$.

Bridge to Linear Algebra

Remark 17.59 (Why This Chapter Exists). Every proof in Axler's Chapter 1 draws on exactly two sources: properties of the field $\mathbb{F}$, and properties of the vector space V. This chapter is the permanent home of the field side. The table below maps each type of linear algebra proof step to the theorem it cites from this chapter.

Proof step in linear algebra	Structure used	Theorem
"since $a \neq 0$, a^{-1} exists"	Field $\mathbb{F}$	Prop. 17.55
"$ab = 0$ and $a \neq 0$ implies $b = 0$"	Field $\mathbb{F}$	Prop. 17.54
"$(-1)v = -v$"	Ring theorem on $\mathbb{F}$	Prop. 17.38
"$0 \cdot v = \mathbf{0}$"	Ring theorem on $\mathbb{F}$ acting on V	Prop. 17.36
"the additive identity is unique"	Abelian group $(V, +)$	Prop. 17.16
"additive inverses are unique"	Abelian group $(V, +)$	Prop. 17.18
"$-(-v) = v$"	Abelian group $(V, +)$	Prop. 17.18

Remark 17.60 (The Interaction Layer). The table above separates into two kinds of steps: those that use the structure of the field $\mathbb{F}$ alone (rows 1–3), and those that use the abelian group structure of V alone (rows 5–7). Row 4 is the interaction: it uses a field theorem applied across scalar multiplication to a vector.

This separation is the heart of the vector space structure. The field $\mathbb{F}$ and the abelian group V are independent objects connected by the scalar multiplication axioms. When a proof crosses that connection, it is using the interaction layer. When it stays on one side, it is using either field theory or group theory alone. Knowing which side you are on at each step is what the DU/TA/AM annotation system makes explicit.

Model-Theoretic Foundations

Model theory provides a precise language for describing algebraic structures: a signature names the operations and constants a structure carries, and a model (or structure) interprets those names on a concrete set. Every algebraic structure in these blueprints is a model of some first-order signature.

Definition 17.61 (First-Order Signature). A first-order signature $\mathcal{L}$ consists of:

- Constant symbols $c_1, c_2, \ldots$ naming distinguished elements,
- Function symbols $f_1, f_2, \ldots$, each equipped with a specified arity $n \in \mathbb{N}$ naming n-ary operations,
- Relation symbols $R_1, R_2, \ldots$ (optional), each with a specified arity naming predicates.

Definition 17.62 ($\mathcal{L}$-Structure). Given a signature $\mathcal{L}$, an $\mathcal{L}$-structure $\mathcal{A}$ consists of:

- A nonempty set A, called the universe (or carrier set),
- For each constant symbol c: an element $c^{\mathcal{A}} \in A$,
- For each n-ary function symbol f: a function $f^{\mathcal{A}} : A^n \to A$,
- For each n-ary relation symbol R: a subset $R^{\mathcal{A}} \subseteq A^n$.

We write $\mathcal{A} = (A, \ldots)$, listing the universe and interpretations.

Remark 17.63. The signature is the type; the structure is the instance. A group, for example, has signature $\mathcal{L}_{\text{grp}} = \{*, e, {}^{-1}\}$ where $*$ has arity 2, ${}^{-1}$ has arity 1, and e is a constant. Any group $(G, *, e, {}^{-1})$ is then an $\mathcal{L}_{\text{grp}}$-structure satisfying the group axioms. The axioms themselves are first-order sentences in $\mathcal{L}_{\text{grp}}$, and a model of those sentences is precisely a group. This framework unifies all algebraic structures: rings, fields, vector spaces, and so on are models of their respective signatures.

Functions

A function is the primitive notion underlying every algebraic operation. Before studying operations on sets, we record the full structure of functions and their properties.

Definition 17.64 (Function). Let A and B be sets. A function $f : A \to B$ is a rule such that

$$(\forall\, a \in A)\ (\exists!\, b \in B)\ (f(a) = b).$$

The components of f are:

- Domain: $\mathrm{dom}(f) = A$,
- Codomain: $\mathrm{cod}(f) = B$,
- Graph: $\Gamma_f = \{(a,\, f(a)) : a \in A\} \subseteq A \times B$.

Two functions are equal iff they have the same domain, codomain, and graph.

Definition 17.65 (Image and Preimage). Let $f : A \to B$, $S \subseteq A$, $T \subseteq B$.

- Image of f: $\mathrm{Im}(f) = f(A) = \{\, f(a) : a \in A \,\} \subseteq B$.
- Image of a subset S: $f(S) = \{\, f(s) : s \in S \,\} \subseteq B$.
- Preimage of a subset T: $f^{-1}(T) = \{\, a \in A : f(a) \in T \,\} \subseteq A$.

Note: $f^{-1}(T)$ is defined for every function f; it does not require f to be invertible.

Definition 17.66 (Injective, Surjective, Bijective). Let $f : A \to B$.

- Injective (one-to-one):
$$(\forall\, a_1, a_2 \in A)\ \big(f(a_1) = f(a_2) \ \Rightarrow\ a_1 = a_2\big).$$
Equivalently, distinct inputs give distinct outputs.
- Surjective (onto):
$$(\forall\, b \in B)\ (\exists\, a \in A)\ (f(a) = b).$$
Equivalently, $\mathrm{Im}(f) = B$.
- Bijective: f is both injective and surjective. Equivalently, f establishes a one-to-one correspondence between A and B.

Definition 17.67 (Inverse Function). If $f : A \to B$ is bijective, then its inverse $f^{-1} : B \to A$ is the unique function satisfying

$$(\forall\, a \in A) \quad f^{-1}(f(a)) = a, \qquad (\forall\, b \in B) \quad f(f^{-1}(b)) = b.$$

That is, $f^{-1} \circ f = \mathrm{id}_A$ and $f \circ f^{-1} = \mathrm{id}_B$. A function has an inverse if and only if it is bijective.

Remark 17.68. Injectivity, surjectivity, and bijectivity measure how faithfully f maps A into B. Injective: no collisions. Surjective: full coverage. Bijective: perfect pairing. In algebra, the morphisms that preserve structure and are bijective (isomorphisms) are the ones that identify structures as the same up to relabelling. Note also the distinction between $\mathrm{Im}(f) \subseteq B$ (a subset) and $f^{-1}(T) \subseteq A$ (a preimage): the former goes forward, the latter pulls back, and neither requires the function to be invertible.

Unary (Singular) Operation

Before defining binary operations (arity 2), we fix the simpler case of arity 1. Unary operations appear throughout algebra: negation $a \mapsto -a$, multiplicative inverse $a \mapsto a^{-1}$, complex conjugation $z \mapsto \bar{z}$, and set complementation.

Definition 17.69 (Unary Operation). Let S be a set. A unary operation on S is a function

$$f : S \to S.$$

Components:

- Underlying set: S (domain and codomain coincide),
- Function: f,
- Closure: $(\forall\, x \in S)\ (\exists!\, y \in S)\ (y = f(x))$.

Model-theoretic signature:

$$\mathcal{L}_{\text{unary}} = \{\, f \,\}, \qquad f \text{ a function symbol of arity 1.}$$

The structure (S, f) is then an $\mathcal{L}_{\text{unary}}$-structure with no axioms beyond closure.

Remark 17.70. A unary operation is to a binary operation what a button is to a lever: it takes one input and produces one output, always within the same set. In group theory the inverse map $^{-1} : G \to G$ is a unary operation built into the signature; the axiom $a \cdot a^{-1} = e$ then constrains it. Similarly, in a ring, negation $- : A \to A$ is unary. Naming unary operations explicitly in the signature (rather than deriving them) is the model-theoretic convention that keeps axioms first-order.

New at Set

$\exists\ a, b, c \in S$

Definition 17.71 (Set). A set S is a well-defined collection of distinct objects called elements. We write $a \in S$ to denote that a is an element of S, and $a \notin S$ otherwise. Two sets are equal if and only if they have the same elements:

$$S = T \iff \forall x : (x \in S \iff x \in T).$$

Remark 17.72. A set carries no additional structure no operations, no ordering, no notion of distance. It is the blank slate from which all algebraic structures are built by adding operations and axioms. When we define a Magma $(A, *)$, the A is just a set; the operation $*$ is layered on top.

Extended by: Binary Operation, Magma (add a closed map)

Note: No operations or axioms at this level. The carrier set is the only ingredient.

Binary Operation $* : A \times A \to A$ — Structure Blueprint

New at Binary Operation

f $* : A \times A \to A$	[the binary map]
$\forall\ \forall a, b \in A :\ a * b \in A$	[closure]

Inherited from Set

$\exists\ a,\ b,\ c \in A$	[elements of the carrier set]

Definition 17.73 (Binary Operation). Let A be a set. A binary operation on A is a function

$$* : A \times A \to A$$

that assigns to each ordered pair $(a, b) \in A \times A$ a unique element $a * b \in A$. The defining property is closure: $\forall a, b \in A,\ a * b \in A$, i.e. the operation never escapes the set.

Model-theoretic signature:

$$\mathcal{L}_{\text{bin}} = \{*\}, \qquad * \text{ a function symbol of arity } 2.$$

Remark 17.74. A binary operation is a concept, not a structure in its own right it is the interface that all group-like structures implement. Closure is the only requirement: there are no demands on how $*$ behaves (no associativity, no identity, no inverses). In OOP terms, this is a pure abstract interface; Magma is the first concrete class that implements it.

IS-A: Set (carrier A)
HAS-A: implemented by Magma and all descendants

Note: Addition $+$ on $\mathbb{Z}$, multiplication $\cdot$ on $\mathbb{R}$, and function composition are all binary operations.

Magma $(A, *)$ Structure Blueprint

New at Magma

f $* : A \times A \to A$	[closed binary operation]
$\forall\ \forall a, b \in A :\ a * b \in A$	[closure]

Inherited from Set

$\exists\ a,\ b,\ c \in A$	[elements of the carrier set]

Definition 17.75 (Magma). A magma is a pair $(A, *)$ where A is a nonempty set and $* : A \times A \to A$ is a binary operation satisfying:

$$\forall\, a, b \in A :\ a * b \in A. \qquad \text{(Closure)}$$

No further axioms are imposed on $*$.

Model-theoretic signature:

$$\mathcal{L}_{\text{magma}} = \{*\}, \qquad * \text{ a function symbol of arity 2.}$$

Remark 17.76. A magma is the weakest non-trivial algebraic structure the bare minimum for calling something an algebra. The single axiom, closure, merely says that applying $*$ to elements of A stays inside A. This rules out partial operations but imposes no coherence (no associativity, no identity, no inverses). Most naturally occurring operations satisfy much more, which is why the Magma is usually the invisible foundation rather than the object of study itself.

IS-A: Set HAS-A: Binary Operation
Extended by: Semigroup (adds associativity)

Example: $(\mathbb{N}$, rock-paper-scissors) is a magma closed, but not associative.

Semigroup (A, *) — Structure Blueprint

New at Semigroup

$\forall\ \forall a, b, c \in A:\ (a * b) * c = a * (b * c)$ [associativity]

Inherited from Magma

f $* : A \times A \to A$ [closed binary operation]

$\forall\ \forall a, b \in A:\ a * b \in A$ [closure]

Inherited from Set

$\exists\ a,\ b,\ c \in A$ [elements of the carrier set]

Definition 17.77 (Semigroup). A semigroup is a pair $(A, *)$ where A is a nonempty set and $* : A \times A \to A$ is a binary operation satisfying:

$$\forall a, b \in A : a * b \in A \qquad \text{(Closure)}$$

$$\forall a, b, c \in A : (a * b) * c = a * (b * c) \qquad \text{(Associativity)}$$

Model-theoretic signature:

$$\mathcal{L}_{\text{semi}} = \{*\}, \qquad * \text{ a function symbol of arity 2.}$$

Remark 17.78. Associativity is what allows us to write $a * b * c$ without parentheses the result is the same regardless of how we bracket a chain of operations. This seemingly minor property has profound consequences: it enables induction arguments over products, underpins the theory of free monoids (strings), and is what makes the integers under addition well-behaved. The semigroup is the first structure where the operation itself has meaningful algebraic content.

IS-A: Magma HAS-A: Binary Operation
Extended by: Monoid (adds identity element)

Examples: $(\mathbb{Z}^+, +)$, $(\mathbb{N}, \cdot)$, non-empty strings under concatenation.
Non-example: $(\mathbb{Z}, -)$ subtraction is not associative.

Monoid (A, *, e) — Structure Blueprint

New at Monoid

$\exists\ e \in A$	[identity element]
$\forall\ \forall a \in A:\ a * e = e * a = a$	[identity law]

Inherited from Semigroup

$\forall\ \forall a, b, c \in A:\ (a * b) * c = a * (b * c)$	[associativity]

Inherited from Magma

f $* : A \times A \to A$	[closed binary operation]
$\forall\ \forall a, b \in A:\ a * b \in A$	[closure]

Definition 17.79 (Monoid). A monoid is a triple $(A, *, e)$ where A is a nonempty set, $* : A \times A \to A$ is a binary operation, and $e \in A$, satisfying:

$$\forall a, b \in A : a * b \in A \qquad \text{(Closure)}$$
$$\forall a, b, c \in A : (a * b) * c = a * (b * c) \qquad \text{(Associativity)}$$
$$\forall a \in A : a * e = e * a = a \qquad \text{(Identity)}$$

The element e is called the identity (or neutral element) of $*$. It is unique: if e' also satisfies the identity law, then $e = e * e' = e'$.

Model-theoretic signature:

$$\mathcal{L}_{\text{mon}} = \{ *, e \}, \quad * \text{ arity } 2, \quad e \text{ a constant symbol.}$$

Remark 17.80. The identity element gives the monoid a "do nothing" operation a baseline from which all elements are measured. This is what makes monoids so prevalent in computing: the empty string is the identity under concatenation, 0 is the identity under addition, 1 under multiplication, and the empty list under append. Monoids are the algebraic abstraction of combining things with a neutral default.

IS-A: Semigroup HAS-A: Binary Operation
Extended by: Group (adds inverses); Ring (HAS-A Monoid under $\cdot$)

Examples: $(\mathbb{N}, +, 0)$, $(\mathbb{N}, \cdot, 1)$, $(\Sigma^*, \|, \varepsilon)$ strings under concatenation.
Identity is unique: if $a * e' = e' * a = a$ for all a, then $e' = e$.

Group $(A, *, e, a^{-1})$ — Structure Blueprint

New at Group

$\exists\ \forall a \in A,\ \exists a^{-1} \in A$	[inverse element]
$\forall\ \forall a \in A:\ a * a^{-1} = a^{-1} * a = e$	[inverse law]

Inherited from Monoid

$\exists\ e \in A$	[identity element]
$\forall\ \forall a \in A:\ a * e = e * a = a$	[identity law]

Inherited from Semigroup

$\forall\ \forall a, b, c \in A:\ (a * b) * c = a * (b * c)$	[associativity]

Inherited from Magma

f $* : A \times A \to A$	[closed binary operation]
$\forall\ \forall a, b \in A:\ a * b \in A$	[closure]

Definition 17.81 (Group). A group is a quadruple $(A, *, e, {}^{-1})$ where A is a nonempty set, $* : A \times A \to A$, $e \in A$, and ${}^{-1} : A \to A$, satisfying:

$$\forall a, b \in A : a * b \in A \qquad \text{(Closure)}$$
$$\forall a, b, c \in A : (a * b) * c = a * (b * c) \qquad \text{(Associativity)}$$
$$\forall a \in A : a * e = e * a = a \qquad \text{(Identity)}$$
$$\forall a \in A : a * a^{-1} = a^{-1} * a = e \qquad \text{(Inverses)}$$

The inverse a^{-1} is unique for each a, and $(a^{-1})^{-1} = a$.

Model-theoretic signature:

$$\mathcal{L}_{\text{grp}} = \{ *,\ e,\ {}^{-1} \}, \quad * \text{ arity } 2, \quad {}^{-1} \text{ arity } 1, \quad e \text{ a constant.}$$

Remark 17.82. Inverses are what turn a monoid into a group, and the conceptual leap is significant: every element now has an "undo." This is the algebraic abstraction of reversible processes rotations of a shape, permutations of a set, invertible linear maps. Group theory is arguably the most central structure in all of algebra, connecting symmetry in geometry, number theory (e.g. $(\mathbb{Z}/n\mathbb{Z})^*$), and the classification of elementary particles in physics. Note that commutativity is not required: $a * b$ need not equal $b * a$.

IS-A: Monoid HAS-A: Binary Operation
Extended by: Abelian Group (adds commutativity)

Examples: $(\mathbb{Z}, +, 0)$, $(S_n, \circ, \mathrm{id})$ permutations, $(GL_n(\mathbb{R}), \cdot, I)$ invertible matrices.
Non-example: $(\mathbb{N}, +)$ no additive inverses.

Abelian Group $(A, +, 0, -a)$ — Structure Blueprint

New at Abelian Group

$\forall\ \forall a, b \in A:\ a + b = b + a$	[commutativity]
$\exists$ rename: $e \to 0, \quad a^{-1} \to -a, \quad * \to +$	[notational convention]

Inherited from Group

$\exists\ \forall a \in A,\ \exists -a \in A$	[additive inverse]
$\forall\ \forall a \in A:\ a + (-a) = (-a) + a = 0$	[inverse law]
$\exists\ 0 \in A$	[additive identity]
$\forall\ \forall a \in A:\ a + 0 = 0 + a = a$	[identity law]

Inherited from Semigroup

$\forall\ \forall a, b, c \in A:\ (a + b) + c = a + (b + c)$	[associativity]

Definition 17.83 (Abelian Group). A group $(A, +, 0, -)$ is called abelian (or commutative) if it additionally satisfies:

$$\forall a, b \in A:\ a + b = b + a. \qquad \text{(Commutativity)}$$

By convention, the operation is written $+$, the identity as 0, and the inverse of a as $-a$. An abelian group is thus a quintuple $(A, +, 0, -)$ satisfying closure, associativity, identity, inverses, and commutativity.

Model-theoretic signature:

$$\mathcal{L}_{\text{ab}} = \{+, 0, -\}, \quad + \text{ arity } 2, \quad - \text{ arity } 1, \quad 0 \text{ a constant.}$$

Remark 17.84. Commutativity is one axiom, but it changes the character of the structure profoundly. Abelian groups are far more tractable than general groups: their subgroup structure is simpler, they admit a full classification theorem (every finitely generated abelian group is a product of cyclic groups), and they form the additive backbone of every ring and field. The additive notation $+$ signals commutativity by convention throughout algebra.

IS-A: Group HAS-A: Binary Operation
Extended by: Ring (HAS-A Ab. Group under $+$); Vector Space (HAS-A Ab. Group on vectors)
Examples: $(\mathbb{Z}, +, 0)$, $(\mathbb{R}, +, 0)$, $(\mathbb{Z}/n\mathbb{Z}, +, \bar{0})$.
Non-example: $(S_3, \circ)$ non-abelian group of order 6.

Ring $(A, +, \cdot, 0, 1)$ — Structure Blueprint

New at Ring

f $\cdot : A \times A \to A$	[multiplicative binary operation]
$\exists\ 1 \in A$	[multiplicative identity]
$\forall\ \forall a \in A :\ a \cdot 1 = 1 \cdot a = a$	[identity law for $\cdot$]
$\forall\ \forall a, b, c \in A :\ (a \cdot b) \cdot c = a \cdot (b \cdot c)$	[associativity of $\cdot$]
$\forall\ \forall a, b, c \in A :\ a \cdot (b + c) = a \cdot b + a \cdot c$	[left distributivity]
$\forall\ \forall a, b, c \in A :\ (a + b) \cdot c = a \cdot c + b \cdot c$	[right distributivity]

HAS-A: Abelian Group under addition

f $+ : A \times A \to A$	[addition]
$\exists\ 0 \in A$	[additive identity]
$\exists\ \forall a,\ \exists -a \in A$	[additive inverse]
$\forall\ \forall a, b :\ a + b = b + a$	[commutativity of $+$]
$\forall\ \forall a, b, c :\ (a + b) + c = a + (b + c)$	[associativity of $+$]

HAS-A: Monoid under multiplication

f $\cdot : A \times A \to A$	[multiplication]
$\exists\ 1 \in A$	[multiplicative identity]
$\forall\ \forall a, b, c :\ (a \cdot b) \cdot c = a \cdot (b \cdot c)$	[associativity of $\cdot$]

Definition 17.85 (Ring). A ring is a tuple $(A, +, \cdot, 0, 1)$ where A is a nonempty set satisfying:

$$(A, +, 0, -) \text{ is an abelian group} \qquad \text{(Ab. Group under } +\text{)}$$
$$(A, \cdot, 1) \text{ is a monoid} \qquad \text{(Monoid under } \cdot\text{)}$$
$$\forall a, b, c \in A : a \cdot (b + c) = a \cdot b + a \cdot c \qquad \text{(Left Distributivity)}$$
$$\forall a, b, c \in A : (a + b) \cdot c = a \cdot c + b \cdot c \qquad \text{(Right Distributivity)}$$

Note: $\cdot$ need not be commutative, and multiplicative inverses need not exist.

Model-theoretic signature:

$$\mathcal{L}_{\text{ring}} = \{+, \cdot, 0, 1, -\}, \quad +, \cdot \text{ arity } 2, \quad - \text{ arity } 1, \quad 0, 1 \text{ constants.}$$

Remark 17.86. A ring is the algebraic formalization of arithmetic: two operations linked by distributivity. The integers $\mathbb{Z}$ are the canonical example. Distributivity is the bridge between addition and multiplication it is what makes the two operations interact coherently rather than live as independent structures on the same set. Rings with commutative multiplication are called commutative rings; those where every nonzero element has a multiplicative inverse are division rings; both together give a field.

HAS-A: Abelian Group $(A, +, 0, -)$; Monoid $(A, \cdot, 1)$
Extended by: Field (adds mult. inverses + commutativity)

Examples: $(\mathbb{Z}, +, \cdot)$, $M_n(\mathbb{R})$ matrices, $\mathbb{Z}[x]$ polynomials.
Non-example: $(\mathbb{N}, +, \cdot)$ no additive inverses.

Field $(A, +, \cdot, 0, 1, a^{-1})$ — Structure Blueprint

New at Field

$\exists\ \forall a \in A \setminus \{0\},\ \exists a^{-1} \in A$	[multiplicative inverse]
$\forall\ \forall a \in A \setminus \{0\} :\ a \cdot a^{-1} = a^{-1} \cdot a = 1$	[inverse law]
$\forall\ \forall a, b \in A :\ a \cdot b = b \cdot a$	[commutativity of $\cdot$]

Inherited from Ring

$\forall\ \forall a, b, c \in A :\ a \cdot (b + c) = a \cdot b + a \cdot c$	[left distributivity]
$\forall\ \forall a, b, c \in A :\ (a + b) \cdot c = a \cdot c + b \cdot c$	[right distributivity]
$\exists\ 1 \in A$	[multiplicative identity]
$\forall\ \forall a, b, c \in A :\ (a \cdot b) \cdot c = a \cdot (b \cdot c)$	[associativity of $\cdot$]

Inherited from Abelian Group (under addition)

f $+ : A \times A \to A$	[addition]
$\exists\ 0 \in A$	[additive identity]
$\exists\ \forall a,\ \exists -a \in A$	[additive inverse]
$\forall\ \forall a, b \in A :\ a + b = b + a$	[commutativity of $+$]
$\forall\ \forall a, b, c \in A :\ (a + b) + c = a + (b + c)$	[associativity of $+$]

Definition 17.87 (Field). A field is a tuple $(A, +, \cdot, 0, 1)$ where A is a set with $|A| \geq 2$, satisfying all ring axioms plus:

$$\forall a, b \in A : a \cdot b = b \cdot a \qquad \text{(Commutativity of } \cdot\text{)}$$

$$\forall a \in A \setminus \{0\},\ \exists a^{-1} \in A : a \cdot a^{-1} = a^{-1} \cdot a = 1 \qquad \text{(Multiplicative Inverses)}$$

Equivalently, $(A \setminus \{0\}, \cdot, 1, {}^{-1})$ is an abelian group. The condition $0 \neq 1$ (i.e. $|A| \geq 2$) excludes the trivial ring.

Model-theoretic signature:

$$\mathcal{L}_{\text{fld}} = \{ +, \cdot, 0, 1, -, {}^{-1} \}, \quad +, \cdot \text{ arity } 2, \quad -, {}^{-1} \text{ arity } 1, \quad 0, 1 \text{ constants.}$$

Remark 17.88. A field is a ring where multiplication is as strong as addition: every nonzero element can be divided by. This is the algebraic setting for linear algebra ($\mathbb{R}$, $\mathbb{C}$, $\mathbb{Q}$) and number theory (finite fields $\mathbb{F}_p$). Division — the operation conspicuously absent from rings — is what makes solving linear equations ($ax = b \Rightarrow x = a^{-1}b$) always possible for $a \neq 0$. The insistence on commutativity of $\cdot$ distinguishes fields from division rings (skew fields), where inverses exist but $ab \neq ba$ generally.

IS-A: Ring
Used by: Vector Space (USES-A Field for scalars)

Examples: $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$, $\mathbb{F}_p = \mathbb{Z}/p\mathbb{Z}$ (p prime).
Non-example: $\mathbb{Z}$ — no multiplicative inverse for 2.

Vector Space $(V, +, 0, -v, F, \cdot)$ — Structure Blueprint

New at Vector Space

f $\cdot : F \times V \to V$	[scalar multiplication]
$\forall\ \forall a \in F, \mathbf{u}, \mathbf{v} \in V :\ a \cdot (\mathbf{u} + \mathbf{v}) = a \cdot \mathbf{u} + a \cdot \mathbf{v}$	[distributivity over vector addition]
$\forall\ \forall a, b \in F, \mathbf{v} \in V :\ (a + b) \cdot \mathbf{v} = a \cdot \mathbf{v} + b \cdot \mathbf{v}$	[distributivity over scalar addition]
$\forall\ \forall a, b \in F, \mathbf{v} \in V :\ (ab) \cdot \mathbf{v} = a \cdot (b \cdot \mathbf{v})$	[compatibility of scalar mult.]
$\forall\ \forall \mathbf{v} \in V :\ 1_F \cdot \mathbf{v} = \mathbf{v}$	[identity scalar]

HAS-A: Abelian Group on V

f $+ : V \times V \to V$	[vector addition]
$\exists\ \mathbf{0} \in V$	[zero vector]
$\exists\ \forall \mathbf{v},\ \exists -\mathbf{v} \in V$	[additive inverse]
$\forall\ \forall \mathbf{u}, \mathbf{v} \in V :\ \mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$	[commutativity]
$\forall\ \forall \mathbf{u}, \mathbf{v}, \mathbf{w} :\ (\mathbf{u} + \mathbf{v}) + \mathbf{w} = \mathbf{u} + (\mathbf{v} + \mathbf{w})$	[associativity]

USES-A: Field F for scalars

f $+, \cdot : F \times F \to F$	[field operations]
$\exists\ 0_F, 1_F \in F$	[additive & mult. identity]
$\exists\ \forall a \in F \setminus \{0\},\ \exists a^{-1}$	[mult. inverse]
$\forall\ \forall a, b \in F :\ ab = ba$	[commutativity of $\cdot$]

Definition 17.89 (Vector Space). Let F be a field. A vector space over F is a tuple $(V, +, \mathbf{0}, -, F, \cdot)$ where $(V, +, \mathbf{0}, -)$ is an abelian group and $\cdot : F \times V \to V$ satisfies, for all $a, b \in F$ and $\mathbf{u}, \mathbf{v} \in V$:

$$a \cdot (\mathbf{u} + \mathbf{v}) = a \cdot \mathbf{u} + a \cdot \mathbf{v} \qquad \text{(Distributivity over vector } +\text{)}$$
$$(a + b) \cdot \mathbf{v} = a \cdot \mathbf{v} + b \cdot \mathbf{v} \qquad \text{(Distributivity over scalar } +\text{)}$$
$$(ab) \cdot \mathbf{v} = a \cdot (b \cdot \mathbf{v}) \qquad \text{(Compatibility)}$$
$$1_F \cdot \mathbf{v} = \mathbf{v} \qquad \text{(Identity scalar)}$$

Elements of V are called vectors; elements of F are called scalars.

Model-theoretic signature:

$$\mathcal{L}_{\mathrm{VS}} = \{ +_V,\ \mathbf{0},\ -_V,\ \cdot \} \cup \mathcal{L}_{\mathrm{fld}},$$

where $\cdot$ has arity 2 (one scalar, one vector) and $\mathcal{L}_{\mathrm{fld}}$ names the field operations on F.

Remark 17.90. A vector space decouples two roles: vectors are the objects being combined (via $+$), and scalars are the coefficients that stretch or shrink them (via $\cdot$). The scalars must form a field so that scalar division is always possible — this is what makes Gaussian elimination work. The four scalar-multiplication axioms together say that scaling behaves consistently with both kinds of addition and with the field's own multiplication. Linear algebra — spanning sets, bases, dimension, linear maps — is entirely the theory of vector spaces over a field.

HAS-A: Abelian Group $(V, +, \mathbf{0})$ on vectors
USES-A: Field F for scalars (external dependency)

Examples: $\mathbb{R}^n$ over $\mathbb{R}$, $\mathbb{C}^n$ over $\mathbb{C}$, polynomials $F[x]$ over F.
Note: Changing the field changes the space: $\mathbb{C}$ is dim-2 over $\mathbb{R}$, dim-1 over $\mathbb{C}$.

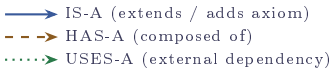
Group-Like Algebraic Structures

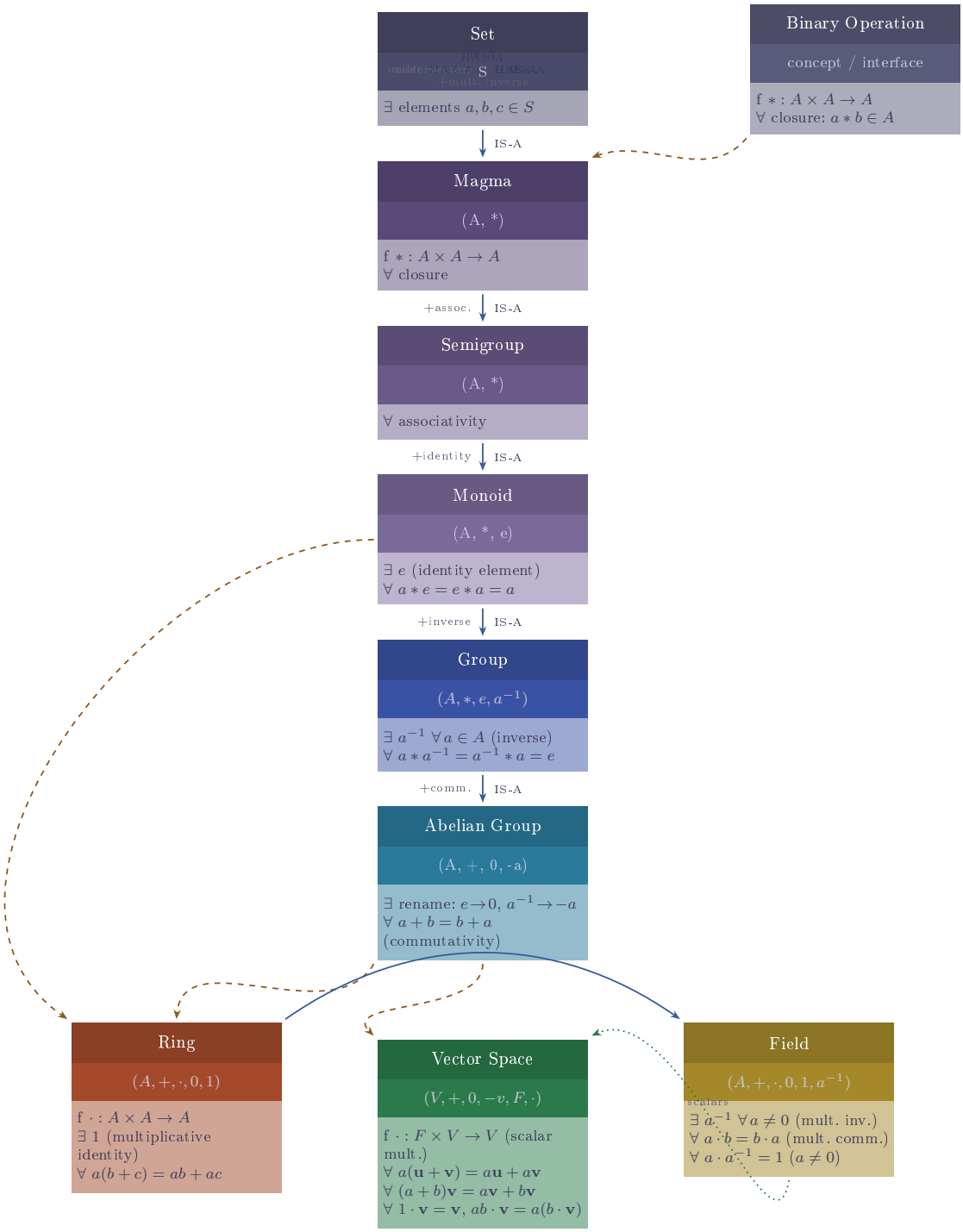
DAG OVERVIEW ů IS-A / HAS-A / USES-A RELATIONSHIPS

∃ Distinguished element (existentially quantified)

f Operation (binary map, closed signature)

∀ Axiom / property (universally quantified)





is-a Strict extension—child satisfies all parent axioms plus at least one new constraint.

has-a Composition—structure contains another as a component (e.g. Ring has-a Ab. Group under + and Monoid under ·).

uses-a External dependency—one structure acts on or requires another without inheriting (e.g. Vector Space uses-a Field for scalars).

Chapter 18

Set Algebras

18.1 Notes

Where You Are in the Journey

Propositional Logic $\rightarrow$ Sets & Functions $\rightarrow$ $\mathbb{R}$ $\rightarrow$ Algebraic Structures $\rightarrow$ Set Algebras $\rightarrow$ Measure Theory $\rightarrow$ $\cdots$

How we got here. Set theory gave us the basic operations — union, intersection, complement, and power set — and the language of subsets of a fixed space X. Algebraic structures showed us how to classify mathematical objects by their closure properties under operations. Set algebras bring these two threads together: we ask which collections of subsets are closed under which set-theoretic operations, and what algebraic laws govern them.

What this chapter builds. We develop the hierarchy of closure systems on 2^X: rings of sets, algebras (fields) of sets, σ-rings, and σ-algebras. Alongside this, we study characteristic functions as a bridge between subsets and arithmetic, revealing that 2^X carries Boolean ring and $\mathbb{F}_2$-vector space structure. The power set itself is the ambient σ-algebra and the universal example.

Where this leads. Measure theory requires a σ-algebra as its foundational input: a measure is defined on a σ-algebra, and measurability is a membership condition in one. The Boolean algebra structure of 2^X connects back to propositional logic. Characteristic functions reappear as simple functions in Lebesgue integration theory.

Structural Roadmap

Each major topic is organised as:

$$\text{Definitions} \longrightarrow \text{Main Theorems} \longrightarrow \text{Consequences and Structural Insight}$$

The global progression is:

1. Families of sets and closure operations
2. Finite closure structures: rings of sets and algebras of sets
3. Countable closure structures: σ-rings and σ-algebras
4. The power set as a Boolean ring and $\mathbb{F}_2$-vector space
5. Characteristic functions and the identification $2^X \cong \{0,1\}^X$

Remark 18.1 (Primary sources). The treatment follows Kolmogorov and Fomin Introductory Real Analysis, Chapter 1, and standard accounts in measure theory (Rudin, Royden). Characteristic functions and the Boolean ring structure appear in Halmos Measure Theory.

18.1.1 Systems of Sets

Systems of Sets — Quick Reference

Structure	Closure properties	Detail
Family of sets	Any $\mathcal{F} \subseteq 2^X$	Def
Closed under $*$	Inputs in $\mathcal{F}$ implies output in $\mathcal{F}$	Def
Ring of sets	Closed under $\cup$ and $\setminus$	Def
Algebra of sets	Contains X; closed under $(\cdot)^c$ and $\cup$	Def
σ-ring	Ring closed under countable $\cup$	Def
σ-algebra	Algebra closed under countable $\cup$	Def
Key results:		
Algebra = ring + X	Equivalent characterisation	Prop
σ-algebra closure	Also closed under countable $\cap$, $\setminus$, finite Boolean ops	Prop

Let X be a set. A family of sets on X is a collection $\mathcal{F} \subseteq 2^X$. The objects of study are not elements of X but collections of subsets, classified by their stability under set-theoretic operations.

Definition (Family of Sets)

Let X be a set. A family of sets on X is any collection $\mathcal{F} \subseteq 2^X$.

Definition (Closure Under an Operation)

Let $\mathcal{F} \subseteq 2^X$ and let $*$ be an operation on subsets of X. We say $\mathcal{F}$ is closed under $*$ if whenever the inputs to $*$ belong to $\mathcal{F}$, the output also belongs to $\mathcal{F}$.

Remark 18.2 (Typical operations). The operations of interest are $\cup$, $\cap$, $(\cdot)^c$, $\setminus$, and $\triangle$ (symmetric difference). Different combinations of closure under these operations produce the distinct algebraic structures defined below.

Definition (Ring of Sets)

A nonempty collection $\mathcal{R} \subseteq 2^X$ is a ring of sets if it is closed under finite union and set difference:

(i) $A, B \in \mathcal{R} \Rightarrow A \cup B \in \mathcal{R}$,

(ii) $A, B \in \mathcal{R} \Rightarrow A \setminus B \in \mathcal{R}$.

Remark 18.3 (Intersection and symmetric difference). A ring of sets is automatically closed under $\cap$ and $\triangle$, since $A \cap B = A \setminus (A \setminus B)$ and $A \triangle B = (A \setminus B) \cup (B \setminus A)$. Under symmetric difference and intersection, a ring of sets is an abelian group under $\triangle$ and a commutative monoid under $\cap$, making it a Boolean ring in the algebraic sense.

Definition (Algebra of Sets)

A collection $\mathcal{A} \subseteq 2^X$ is an algebra of sets (also called a field of sets) if:

(i) $X \in \mathcal{A}$,

(ii) $A \in \mathcal{A} \Rightarrow A^c \in \mathcal{A}$,

(iii) $A, B \in \mathcal{A} \Rightarrow A \cup B \in \mathcal{A}$.

Proposition 18.4 (Algebra = Ring + X). An algebra of sets is precisely a ring of sets that contains X.

Remark 18.5 (Finite Boolean stability). An algebra of sets is closed under all finite Boolean combinations: finite unions, finite intersections, complements, differences, and symmetric differences. The qualifier "finite" is essential; the next step is to add closure under countably infinite operations.

Definition (σ-Ring)

A ring of sets $\mathcal{R}$ is a σ-ring if it is also closed under countable unions:

$$A_1, A_2, \cdots \in \mathcal{R} \Rightarrow \bigcup_{n=1}^{\infty} A_n \in \mathcal{R}.$$

Definition (σ-Algebra)

A collection $\mathcal{F} \subseteq 2^X$ is a σ-algebra if:

(i) $X \in \mathcal{F}$,

(ii) $A \in \mathcal{F} \Rightarrow A^c \in \mathcal{F}$,

(iii) $A_1, A_2, \cdots \in \mathcal{F} \Rightarrow \bigcup_{n=1}^{\infty} A_n \in \mathcal{F}$.

Proposition 18.6 (Closure consequences of a σ-algebra). Every σ-algebra $\mathcal{F}$ is also closed under countable intersections, set differences, and all finite Boolean combinations.

Remark 18.7 (Finite vs. countable closure). The passage from an algebra to a σ-algebra is precisely the step from finite to countable closure. This is what makes σ-algebras compatible with limit

processes: if measurable sets are approximated by sequences of other measurable sets, the limit set remains measurable.

Structure	Contains X	Closed under $(\cdot)^c$	Closed under $\bigcup$
Ring of sets	No	No	Finite only
Algebra	Yes	Yes	Finite only
σ-Ring	No	No	Countable
σ-Algebra	Yes	Yes	Countable

Remark 18.8 (Reading the hierarchy). Each row is strictly stronger than the one above it: every algebra is a ring, and every σ-algebra is both an algebra and a σ-ring. The top of the hierarchy, the σ-algebra, is the natural domain for measure theory because it combines the Boolean stability of an algebra with the limit-compatibility of countable closure.

18.1.2 The Power Set and Characteristic Functions

Power Set and Characteristic Functions — Quick Reference

Concept	Meaning	Detail
Power set 2^X	All subsets of X; $\lvert 2^X\rvert = 2^{\lvert X\rvert}$	Def
Characteristic function	$\chi_A : X \to \{0,1\}$; encodes membership in A	Def
$2^X \cong \{0,1\}^X$	Bijection $A \mapsto \chi_A$	Prop
Boolean ring structure	$(2^X, \triangle, \cap)$ is a Boolean ring	Prop
$\mathbb{F}_2$-vector space	$(2^X, \triangle)$ is a vector space over $\mathbb{F}_2$	Prop

Definition (Power Set)

The power set of X is

$$2^X := \{\, A \mid A \subseteq X \,\}.$$

Remark 18.9 (Cardinality). The notation 2^X reflects the identity $\lvert 2^X\rvert = 2^{\lvert X\rvert}$ for finite X: each element of X independently either belongs to a subset (1) or does not (0), giving $2^{\lvert X\rvert}$ binary choices in total. For infinite X, the same exponential notation is retained and 2^X denotes the cardinal power.

Definition (Characteristic Function)

Let $A \subseteq X$. The characteristic function (or indicator function) of A is the function $\chi_A : X \to \{0,1\}$ defined by

$$\chi_A(x) := \begin{cases} 1 & \text{if } x \in A, \\ 0 & \text{if } x \notin A. \end{cases}$$

Proposition 18.10 (Identification of subsets with functions). The map $A \mapsto \chi_A$ is a bijection

$$2^X \xrightarrow{\sim} \{0,1\}^X.$$

Consequently, subsets of X and functions from X to $\{0,1\}$ are in exact correspondence.

Remark 18.11 (Functional perspective). This bijection means we can study subsets of X as functions, and vice versa. The set-theoretic operations then correspond to pointwise arithmetic: $\chi_{A\cup B} = \max(\chi_A, \chi_B)$, $\chi_{A\cap B} = \min(\chi_A, \chi_B) = \chi_A \cdot \chi_B$, and $\chi_{A^c} = 1 - \chi_A$. The symmetric difference corresponds to addition modulo 2: $\chi_{A\triangle B} = \chi_A + \chi_B \pmod 2$.

Proposition 18.12 (Boolean ring structure on 2^X). The power set $(2^X, \triangle, \cap)$, with symmetric difference as addition and intersection as multiplication, is a Boolean ring: a commutative ring in which every element is idempotent ($A \cap A = A$).

Proposition 18.13 ($\mathbb{F}_2$-vector space structure on 2^X). The power set $(2^X, \triangle)$, with scalar multiplication defined by $0 \cdot A = \varnothing$ and $1 \cdot A = A$, is a vector space over the two-element field $\mathbb{F}_2 = \{0,1\}$.

Remark 18.14 (Consequence). This vector space structure means any family of subsets closed under $\triangle$ and containing $\varnothing$ is a subspace of 2^X over $\mathbb{F}_2$. Characteristic functions are then the natural coordinate vectors, and the Boolean operations become linear-algebraic operations. This bridge between set theory and linear algebra reappears whenever one needs to decompose or represent families of measurable sets.

Remark 18.15 (Transition). The full power set 2^X is itself a σ-algebra — the largest one on X. Every σ-algebra on X is a sub-σ-algebra of 2^X, just as every subspace of a vector space sits inside the ambient space. This ambient structure makes 2^X the natural starting point for constructing the Borel σ-algebra in measure theory.

Chapter 19

Linear Algebra

19.1 Notes

Where You Are in the Journey

Propositional Logic $\to$ Predicate Calculus $\to$ Sets & Functions $\to$ Proof Techniques $\to$ $\mathbb{R}$ $\to$ Algebraic Structures $\to$ Linear Algebra $\to$ Topology $\to$ $\cdots$

How we got here. Algebraic structures gave us the hierarchy of groups, rings, and fields. A field is the scalar system: a set of numbers with addition and multiplication. An abelian group is the additive backbone. Linear algebra asks: what happens when you combine an abelian group of "vectors" with scalar multiplication from a field?

What this chapter builds. We develop vector spaces from the axioms, then build the theory of subspaces, linear combinations, span, linear independence, bases, and dimension. Linear maps are the central objects of study — not matrices, which are only representations relative to a basis.

Where this leads. Every proof about eigenvalues, the spectral theorem, and inner product spaces relies directly on the axioms established here. Functional analysis generalises vector spaces to infinite dimensions; topology provides the analytic tools for that generalisation.

19.1.1 Linear Algebra

Structural Roadmap

The development of linear algebra in this project follows the definition–theorem–structure architecture used throughout the analysis volumes.

The primary driver is Linear Algebra Done Right by Sheldon Axler. The emphasis is on linear maps, invariant structure, and conceptual clarity rather than computational techniques.

Each major topic is organized as:

$$\text{Definitions} \longrightarrow \text{Main Theorems} \longrightarrow \text{Consequences and Structural Insight}$$

The global progression is:

1. Vector spaces and subspaces
2. Linear combinations, span, and linear independence
3. Bases and dimension
4. Linear maps
5. Null space and range
6. Matrix representations of linear maps
7. Eigenvalues and eigenvectors
8. Invariant subspaces
9. Triangularization and diagonalization
10. Inner product spaces
11. Orthogonality and projections
12. Self-adjoint and normal operators
13. Spectral theorem (finite-dimensional case)

Remark 19.1. This treatment prioritizes structural understanding over algorithmic manipulation. Determinants are introduced only after linear maps and eigenvalues have been conceptually understood.

Remark 19.2. The central object of study is not matrices, but linear maps between vector spaces. Matrices serve only as representations relative to a basis.

Remark 19.3 (Structural Position). Linear algebra is developed independently of analysis. Its structural viewpoint will later support operator theory, Hilbert spaces, and functional analysis.

19.1.1.1 Preliminary Definitions

Definition 19.4 (Abelian group). A set G with a binary operation $* : G \times G \to G$ is an abelian group if:

1. Associativity:
$$(a * b) * c = a * (b * c) \quad \text{for all } a, b, c \in G.$$

2. Identity element: There exists $e \in G$ such that

$$a * e = e * a = a \quad \text{for all } a \in G.$$

3. Inverse element: For each $a \in G$ there exists $b \in G$ such that

$$a * b = b * a = e.$$

4. Commutativity:

$$a * b = b * a \quad \text{for all } a, b \in G.$$

Remark 19.5. An abelian group is simply a group whose operation is commutative.

Definition 19.6 (Field — structural form). A field is a set $\mathbb{F}$ equipped with two operations $+$ and $\cdot$ such that:

- $(\mathbb{F}, +)$ is an abelian group,
- $(\mathbb{F} \setminus \{0\}, \cdot)$ is an abelian group,
- Multiplication distributes over addition.

Remark 19.7 (Expanded Axioms). Expanding the abelian group axioms yields the familiar list:

- Associativity, commutativity, identity, and inverse for addition.
- Associativity, commutativity, identity, and inverse (for nonzero elements) for multiplication.
- The distributive law.

Thus the usual field axioms are not independent assumptions, but consequences of the two abelian group structures together with distributivity.

Remark 19.8 (Structural Summary). A field consists of two compatible abelian group structures:

Structure	Additive Group	Multiplicative Structure
Field	Abelian	Abelian (on nonzero elements)

Distributivity links the two operations.

19.1.1.2 The Complex Numbers

Definition 19.9 (Complex Numbers). The set of complex numbers, denoted $\mathbb{C}$, is defined as

$$\mathbb{C} := \{(a, b) : a, b \in \mathbb{R}\},$$

with operations defined by

$$(a, b) + (c, d) := (a + c,\ b + d),$$
$$(a, b) \cdot (c, d) := (ac - bd,\ ad + bc).$$

Remark 19.10. We identify (a, b) with the symbol

$$a + bi,$$

where $i := (0, 1)$ satisfies $i^2 = -1$.

Remark 19.11 (Field structure of $\mathbb{C}$). With the operations defined above, $\mathbb{C}$ satisfies the field axioms:

- $(\mathbb{C}, +)$ is an abelian group,
- $(\mathbb{C} \setminus \{0\}, \cdot)$ is an abelian group,
- multiplication distributes over addition.

Hence $\mathbb{C}$ is a field containing $\mathbb{R}$ as the subset $\{(a, 0) : a \in \mathbb{R}\}$.

19.1.1.3 Basic Definitions

Definition 19.12 (Vector Space — informal description). A vector space is a set equipped with addition and scalar multiplication that behave like familiar vector arithmetic.

Remark 19.13. The informal description suggests structure, but precision requires axioms.

Definition 19.14 (Vector Space). Let $\mathbb{F}$ be a field. A vector space over $\mathbb{F}$ is a triple $(V, +, \cdot)$ where V is a set, $+ : V \times V \to V$ is a binary operation called vector addition, and $\cdot : \mathbb{F} \times V \to V$ is a binary operation called scalar multiplication, such that:

- $(V, +)$ is an abelian group,
- $\alpha(u + v) = \alpha u + \alpha v$ for all $\alpha \in \mathbb{F}$ and $u, v \in V$,
- $(\alpha + \beta)v = \alpha v + \beta v$ for all $\alpha, \beta \in \mathbb{F}$ and $v \in V$,
- $(\alpha\beta)v = \alpha(\beta v)$ for all $\alpha, \beta \in \mathbb{F}$ and $v \in V$,
- $1_{\mathbb{F}} v = v$ for all $v \in V$.

The elements of V are called vectors and the elements of $\mathbb{F}$ are called scalars.

Remark 19.15 (Expanded Axioms). Expanding the abelian group structure of $(V, +)$ yields the full list of vector space axioms:

- Associativity of addition: $(u + v) + w = u + (v + w)$ for all $u, v, w \in V$.
- Commutativity of addition: $u + v = v + u$ for all $u, v \in V$.
- Additive identity: There exists $\mathbf{0} \in V$ such that $v + \mathbf{0} = v$ for all $v \in V$.
- Additive inverse: For each $v \in V$ there exists $-v \in V$ such that $v + (-v) = \mathbf{0}$.

- Distributivity over vector addition: $\alpha(u+v) = \alpha u + \alpha v$ for all $\alpha \in \mathbb{F}$, $u, v \in V$.
- Distributivity over scalar addition: $(\alpha+\beta)v = \alpha v + \beta v$ for all $\alpha, \beta \in \mathbb{F}$, $v \in V$.
- Compatibility: $(\alpha\beta)v = \alpha(\beta v)$ for all $\alpha, \beta \in \mathbb{F}$, $v \in V$.
- Unit law: $1_{\mathbb{F}} v = v$ for all $v \in V$.

Thus the structural definition compresses eight axioms into one abelian group condition plus four scalar multiplication laws.

Remark 19.16 (Structural Summary). A vector space combines an abelian group with a compatible scalar action:

Structure	Additive Group	Scalar Action
Vector Space	Abelian	$\mathbb{F}$-linear

The scalar multiplication axioms express compatibility between the field $\mathbb{F}$ and the group $(V, +)$.

Definition 19.17 (Zero vector). Let V be a vector space over a field $\mathbb{F}$. The zero vector of V, denoted $\mathbf{0}$, is the additive identity of the abelian group $(V, +)$; that is, the unique element $\mathbf{0} \in V$ such that

$$v + \mathbf{0} = v \quad \text{for all } v \in V.$$

19.1.1.4 Tuples and Lists

Definition 19.18 (n-tuple). Let $n \in \mathbb{N}$ and let X be a set. An n-tuple of elements of X is a function

$$f : \{1, 2, \ldots, n\} \to X.$$

We write such a function as $(x_1, \ldots, x_n)$, where $x_i := f(i)$ for each i.

Definition 19.19 (Finite list). Let X be a set. A finite list of elements of X is a function $f : \{1, 2, \ldots, n\} \to X$ for some $n \in \mathbb{N}$.

Definition 19.20 (Set). A set is a collection of distinct elements with no inherent ordering.

Example 19.21. Consider the elements 1 and 2. The set $\{1, 2, 2\}$ equals $\{1, 2\}$ because sets do not record duplicates. However, the list $(1, 2, 2)$ differs from $(1, 2)$ because lists record both order and repetition.

Definition 19.22 (Coordinate). Let $\mathbb{F}$ be a field and let $v = (v_1, \ldots, v_n) \in \mathbb{F}^n$. The element v_i is called the ith coordinate of v.

19.1.1.5 The Function-Space Viewpoint

Definition 19.23 (Set Exponentiation). Let X and Y be sets. Define

$$Y^X := \{f : X \to Y\},$$

the set of all functions from X to Y.

Remark 19.24. If $|X| = n$ and $|Y| = m$ are finite, then $|Y^X| = m^n$. Thus the notation Y^X agrees with the rules of cardinal arithmetic.

Definition 19.25 ($\mathbb{F}^S$ — the function space over a set). Let $\mathbb{F}$ be a field and let S be a set. Define

$$\mathbb{F}^S := \{f : S \to \mathbb{F}\},$$

the set of all functions from S to $\mathbb{F}$, with pointwise operations:

$$(f+g)(x) := f(x) + g(x), \qquad (\lambda f)(x) := \lambda f(x),$$

for all $f, g \in \mathbb{F}^S$, $\lambda \in \mathbb{F}$, and $x \in S$.

Example 19.26 ($\mathbb{F}^S$ is a vector space). With the pointwise operations above, $\mathbb{F}^S$ is a vector space over $\mathbb{F}$. The zero vector is $\mathbf{0}(x) = 0$ for all $x \in S$, and the additive inverse of f is $(-f)(x) = -f(x)$ for all $x \in S$. All vector space axioms follow from the field axioms applied pointwise; for example,

$$((f+g)+h)(x) = f(x) + g(x) + h(x) = (f + (g+h))(x),$$

so associativity holds because addition in $\mathbb{F}$ is associative. Every other axiom follows similarly.

Remark 19.27 (Zero vector in $\mathbb{F}^S$). In the function space $\mathbb{F}^S$, the zero vector is the function

$$\mathbf{0} : S \to \mathbb{F} \quad \text{defined by} \quad \mathbf{0}(x) = 0 \text{ for all } x \in S.$$

19.1.1.6 Coordinate Spaces as Function Spaces

Definition 19.28 (Coordinate space as a function space). Let $\mathbb{F}$ be a field and let $[n] := \{1, \dots, n\}$. Then

$$\mathbb{F}^n \;=\; \mathbb{F}^{[n]} \;=\; \{f : [n] \to \mathbb{F}\}.$$

An element $v \in \mathbb{F}^n$ is a function $v : [n] \to \mathbb{F}$, written $v = (v_1, \dots, v_n)$ with $v_i := v(i)$. Thus an n-tuple is simply a function from a finite index set, and $\mathbb{F}^n$ inherits its vector space structure as a special case of $\mathbb{F}^S$.

Remark 19.29. With the componentwise operations defined above, $\mathbb{F}^n$ is a vector space over $\mathbb{F}$ for any field $\mathbb{F}$ and any $n \in \mathbb{N}$.

19.1.1.7 The Space $\mathbb{F}^\infty$

Definition 19.30 ($\mathbb{F}^\infty$ — the space of sequences). Let $\mathbb{F}$ be a field. Define

$$\mathbb{F}^\infty := \mathbb{F}^{\mathbb{N}} = \{f : \mathbb{N} \to \mathbb{F}\} = \{(x_1, x_2, \dots) : x_k \in \mathbb{F} \text{ for all } k \in \mathbb{N}\}.$$

An element of $\mathbb{F}^\infty$ is called a sequence of elements of $\mathbb{F}$. With pointwise operations inherited from $\mathbb{F}^S$, $\mathbb{F}^\infty$ is a vector space over $\mathbb{F}$.

Remark 19.31 (Conceptual Unification). The passage

$$\mathbb{F}^n \quad \longrightarrow \quad \mathbb{F}^{\mathbb{N}} \quad \longrightarrow \quad \mathbb{F}^S$$

shows that finite tuples, sequences, and general function spaces are all instances of the same construction. Finite-dimensional linear algebra is the special case where the index set is finite, and this viewpoint connects directly to infinite-dimensional spaces and functional analysis.

19.1.1.8 Basic Propositions

Proposition 19.32 (Uniqueness of the Additive Identity). Let V be a vector space over a field $\mathbb{F}$. The additive identity of V is unique.

Proposition 19.33 (Uniqueness of the Additive Inverse). Let V be a vector space over a field $\mathbb{F}$. For each $v \in V$, the additive inverse of v is unique.

Proposition 19.34 (Scalar Multiplication by Zero and Negation). Let V be a vector space over a field $\mathbb{F}$, let $v \in V$, and let $\alpha \in \mathbb{F}$. Then:

1. $0_{\mathbb{F}} v = \mathbf{0}$,
2. $\alpha \mathbf{0} = \mathbf{0}$,
3. $(-1_{\mathbb{F}})v = -v$.

Chapter 20

Abstract Algebra

20.1 Notes

Where You Are in the Journey

Propositional Logic $\to$ Sets & Functions $\to$ $\mathbb{Z}$, $\mathbb{R}$ $\to$ Algebraic Structures $\to$ Linear Algebra $\to$ Abstract Algebra $\to$ Algebraic Geometry $\to$ $\cdots$

How we got here. Algebraic structures introduced groups, rings, and fields via their axioms. Abstract algebra deepens this study: it asks what properties follow from the axioms alone, develops the theory of homomorphisms and isomorphisms, and classifies structures up to structural equivalence.

What this chapter builds. We develop group theory (subgroups, cosets, Lagrange's theorem, normal subgroups, quotient groups, homomorphisms, isomorphisms, and the isomorphism theorems), ring theory (ideals, quotient rings, polynomial rings, and factorisation), and field extensions.

Where this leads. Galois theory uses group theory to answer questions about polynomial equations. Algebraic geometry studies polynomial ideals and their geometric solutions. Number theory uses ring-theoretic tools throughout.

20.1.1 Abstract Algebra

Structural Roadmap

The development of abstract algebra in this project follows the definition–theorem–structure architecture used throughout the analysis volumes.

The primary driver is Contemporary Abstract Algebra by Joseph A. Gallian (7th ed.).

The emphasis is on algebraic structure as an abstraction of symmetry, arithmetic, and linear

operations. Each major topic is organized as:

$$\text{Definitions} \longrightarrow \text{Structural Theorems} \longrightarrow \text{Classification and Applications}$$

The global progression follows Gallian in five structural stages:

1. Foundations: Integers and Equivalence Relations

   (a) Modular arithmetic
   (b) Equivalence relations and partitions
   (c) Functions and mappings

   Structural Theme: Congruence and equivalence encode algebraic structure via partitions. This stage formalizes symmetry at the level of arithmetic.

2. Groups

   (a) Definition and basic properties
   (b) Subgroups
   (c) Cyclic groups
   (d) Permutation groups
   (e) Isomorphisms
   (f) Cosets and Lagranges Theorem
   (g) Normal subgroups and factor groups
   (h) Homomorphisms
   (i) Fundamental Theorem of Finite Abelian Groups

   Structural Theme: Groups formalize symmetry. Cosets measure deviation from subgroup structure. Normality permits quotient construction. Homomorphisms reveal structural preservation. Classification emerges in the finite abelian case.

3. Rings

   (a) Definition and examples
   (b) Integral domains
   (c) Ideals and factor rings
   (d) Ring homomorphisms
   (e) Polynomial rings
   (f) Unique factorization domains
   (g) Euclidean domains

   Structural Theme: Rings generalize arithmetic. Ideals control structure via quotients. Factorization encodes algebraic rigidity. Euclidean domains permit algorithmic structure.

4. Fields and Extensions

   (a) Vector spaces
   (b) Field extensions
   (c) Algebraic extensions
   (d) Finite fields
   (e) Geometric constructions

   Structural Theme: Fields unify arithmetic and linear algebra. Extensions enlarge solvability. Finite fields exhibit deep classification. Constructibility links algebra and geometry.

5. Special Topics

   (a) Sylow Theorems
   (b) Finite simple groups
   (c) Generators and relations
   (d) Symmetry groups
   (e) Group actions and counting (Burnside)
   (f) Coding theory
   (g) Galois theory
   (h) Cyclotomic extensions

   Structural Theme: Group actions connect algebra to combinatorics. Sylow theory controls finite structure. Galois theory connects symmetry and solvability. Cyclotomic extensions bridge algebra and number theory.

Remark 20.1. Abstract algebra studies algebraic structures defined by operations subject to axioms. Each structure (group, ring, field) is a package consisting of:

$$\text{Set} + \text{Operations} + \text{Axioms}.$$

Remark 20.2. The central construction principle is:

$$\text{Substructure} \longrightarrow \text{Quotient Structure}.$$

Normal subgroups and ideals make quotient constructions possible.

Remark 20.3 (Structural Position). Abstract algebra in this project builds on:

- Set theory and logic (Phase 0)
- Linear algebra (vector space structure)
- Real analysis (structural proof discipline)

Group theory forms the conceptual backbone. Ring theory extends arithmetic. Field theory culminates in symmetry of roots (Galois theory).

20.1.2 Integers

20.1.2.1 Basic Definitions and Theorems

Axiom 20.4 (Well Ordering Principle). Every nonempty set of positive integers contains a smallest member.

$$\forall A \Big(\big(A \subseteq \mathbb{N} \wedge \exists a \, (a \in A) \big) \Rightarrow \exists m \, \big(m \in A \wedge \forall a \, (a \in A \Rightarrow m \leq a) \big) \Big).$$

Remark 20.5. The well-ordering principle asserts that every nonempty subset, A, of $\mathbb{N}$ possesses a least element. Its quantifier structure is:

$$\forall A \, (\exists a \in A \Rightarrow \exists m \in A \, \forall a \in A).$$

This principle is equivalent to the principle of mathematical induction.

Remark 20.6 (Axiomatic Status of the Well-Ordering Principle). The Well-Ordering Principle is not a theorem of elementary arithmetic. It cannot be proved using only the usual algebraic laws of addition and multiplication on $\mathbb{N}$.

Rather, it is taken as an axiom describing the order structure of the natural numbers. In standard foundations, it is equivalent to the Principle of Mathematical Induction and the Least Element Principle. Thus one must assume one of these principles in order to derive the others.

In particular, arithmetic identities alone do not imply that every nonempty subset of $\mathbb{N}$ has a smallest element; this property is part of the defining structure of the natural numbers.

Remark 20.7 (Intuition). Intuitively: the Well-Ordering Principle says you can always find a "smallest" object in any nonempty collection of positive integers. This gives us a foothold once we have a minimal element, we can derive divisibility and gcd properties by contradiction.

Definition 20.8 (Numerator and Denominator). Let $a, b \in \mathbb{Z}$ with $b \neq 0$. In the expression $\dfrac{a}{b}$, we call a the numerator and b the denominator.

Definition 20.9 (Multiple). Let $a, b \in \mathbb{Z}$. We say a is a multiple of b if $a = bk$ for some $k \in \mathbb{Z}$.

Definition 20.10 (Divisibility). Let $a, b \in \mathbb{Z}$ with $b \neq 0$. We say b divides a, written $b \mid a$, if there exists an integer k such that

$$a = bk.$$

If b divides a, we call b a divisor (or factor) of a, and a a multiple of b. If b does not divide a, we write $b \nmid a$.

Remark 20.11 (Unpacking Divisibility). The statement $b \mid a$ is a claim about the existence of an integer k such that $a = bk$. It says nothing about a remainder — divisibility means the remainder is exactly zero.

For example:

- $3 \mid 12$ because $12 = 3 \cdot 4$.
- $5 \nmid 13$ because $13 = 5 \cdot 2 + 3$ and the remainder $3 \neq 0$.

This is the foundation for gcd, lcm, and Bézout's Identity.

Remark 20.12 (Intuition). Intuitively: $b \mid a$ means b fits into a a whole number of times, with nothing left over. Division is exact.

Lemma 20.13 (Divisibility of Linear Combinations). Let $a, b \in \mathbb{Z}$. If $t \mid a$ and $t \mid b$, then t divides every integer linear combination of a and b. That is,

$$t \mid (ua + vb) \quad \text{for all } u, v \in \mathbb{Z}.$$

Remark 20.14. This lemma is used constantly in number theory. Its most important instance is: since $\gcd(a, b)$ divides both a and b, it divides every linear combination $ua + vb$. In particular it divides $d = as + bt$ from Bézout's Identity, which is how we verify the greatest-ness of the gcd.

Remark 20.15 (Intuition). Intuitively: if t fits into both a and b individually with no remainder, then t must fit into any way you scale and combine them — the remainders simply cannot accumulate.

Theorem 20.16 (Division Algorithm). Let a and b be integers with $b > 0$. Then there exist unique integers q and r such that

$$a = bq + r, \qquad 0 \leq r < b.$$

Remark 20.17. In the Division Algorithm, $a = bq + r$, we say a is the dividend and b is the divisor. This corresponds to the division a/b, yielding quotient q and remainder r. When $r = 0$, we write $b \mid a$ ("b divides a"), meaning $a = bq$ for some integer q.

?

Definition 20.18 (Quotient and Remainder). Let $a, b \in \mathbb{Z}$ with $b > 0$. If integers q and r satisfy

$$a = bq + r, \qquad 0 \leq r < b,$$

then:

- q is called the quotient of a divided by b,
- r is called the remainder of a divided by b.

Remark 20.19. The conditions $0 \leq r < b$ ensure uniqueness. Without this bound on r, the representation

$$a = bq + r$$

would not be unique, since one could replace (q, r) with

$$(q + 1, r - b), \quad (q - 1, r + b), \quad \text{etc.}$$

The remainder condition selects exactly one representative.

Remark 20.20 (Interpretation of quotient and remainder). The quotient measures how many full copies of b fit into a, while the remainder measures the leftover part that is strictly smaller than b.

Definition 20.21 (Prime Number). An integer $p > 1$ is called prime if its only positive divisors are 1 and p.

Definition 20.22 (Greatest Common Divisor). Let a and b be integers, not both zero. The greatest common divisor of a and b is the largest of all common divisors of a and b.

A positive integer d is called the greatest common divisor of a and b if:

1. $d \mid a$ and $d \mid b$ (that is, d divides both a and b), and
2. Whenever c is an integer such that $c \mid a$ and $c \mid b$, then $c \mid d$.

The greatest common divisor of a and b is denoted

$$\gcd(a, b).$$

Remark 20.23 (Intuition). Intuitively: the gcd is the largest ruler that measures both a and b exactly, with no remainder. Any other common divisor must itself be measured exactly by the gcd.

Definition 20.24 (Relatively Prime Integers). Two integers a and b are said to be relatively prime (or coprime) if

$$\gcd(a, b) = 1.$$

Remark 20.25. If $\gcd(a, b) = 1$, then the only positive integer dividing both a and b is 1. In this case, a and b share no common prime factors.

Theorem 20.26 (Bézout's Identity). Let a and b be integers, not both zero. Then there exist integers s and t such that

$$\gcd(a, b) = as + bt.$$

Moreover, $\gcd(a, b)$ is the smallest positive integer of the form

$$as + bt, \quad s, t \in \mathbb{Z}.$$

Remark 20.27. The set

$$\{\, as + bt : s, t \in \mathbb{Z} \,\}$$

is called the set of integer linear combinations of a and b. Bézout's Identity asserts that the greatest common divisor is the minimal positive element of this set.

Remark 20.28 (Structural Map of the Proof). This proof has three distinct logical engines.

1. Well-Ordering Principle (existence of a minimal positive combination). We define

   $$S = \{\, am + bn : m, n \in \mathbb{Z},\ am + bn > 0 \,\}.$$

   The Well-Ordering Principle is invoked exactly here to guarantee that S has a smallest element d. This is the sole point where order-theoretic structure on $\mathbb{N}$ enters.

2. Division Algorithm (turning minimality into divisibility). After choosing the minimal $d \in S$, we apply the Division Algorithm to write

   $$a = dq + r, \qquad 0 \leq r < d.$$

   The key move is: if $r > 0$, then r can be rewritten as another positive integer linear combination of a and b, hence $r \in S$, contradicting the minimality of d. Therefore $r = 0$ and $d \mid a$ (and similarly $d \mid b$). So the Division Algorithm is used exactly here to convert the order statement "d is smallest" into the algebraic statement "d divides."

3. Greatest-ness (universal property among common divisors). Finally, for any common divisor d' of a and b, writing $a = d'h$ and $b = d'k$ shows

$$d = as + bt = d'(hs + kt),$$

so $d' \mid d$. This establishes that d is the greatest common divisor in the divisibility order.

Prototype Euclidean-domain argument. The overall pattern is the Euclidean-domain template:

(nonempty set of "sizes") $\xRightarrow{\text{well-ordering}}$ minimal element $\xRightarrow{\text{division algorithm}}$ divisibility + gcd characterization.

In $\mathbb{Z}$, the "size" is the usual order on positive integers; in a general Euclidean domain, the "size" is a Euclidean function $\delta : R \setminus \{0\} \to \mathbb{N}$, and the same minimality-plus-division strategy produces gcds and Bézout-type identities.

Corollary 20.29 (Bezout's Identity — Coprime Case). Let a and b be relatively prime integers. Then there exist integers s and t such that

$$as + bt = 1.$$

Remark 20.30 (Bézout in Both Directions). Bézout's Identity is often applied in two distinct directions, and it is important to keep them separate.

Forward direction. If $\gcd(a, b) = d$, then there exist integers s, t such that

$$as + bt = d.$$

This is the content of the theorem itself.

Reverse direction. If there exist integers s, t such that $as + bt = 1$, then $\gcd(a, b) = 1$.

Proof of reverse direction. Let $d = \gcd(a, b)$. By the forward direction, $d \mid a$ and $d \mid b$, so by Lemma 20.13, $d \mid (as + bt) = 1$. Since $d > 0$ and $d \mid 1$, we conclude $d = 1$.

More generally, if $as + bt = c$ for some integers s, t, then $\gcd(a, b) \mid c$, since $\gcd(a, b)$ divides any linear combination of a and b.

Remark 20.31 (Intuition). Intuitively: a linear combination equaling 1 leaves no room for any common factor greater than 1 — if $d > 1$ divided both a and b, it would divide the combination too, but nothing greater than 1 divides 1.

Theorem 20.32 (Equivalent Characterizations of Relatively Prime Integers). For integers a and b, the following are equivalent:

1. $\gcd(a, b) = 1$.
2. There exist integers s, t such that $as + bt = 1$.
3. a and b share no common prime factor.

Theorem 20.33 (Euclid's Lemma). Let p be a prime integer. If

$$p \mid ab,$$

then

$$p \mid a \quad \text{or} \quad p \mid b.$$

Remark 20.34 (Intuition). Intuitively: primes cannot be fooled by multiplication — if a prime divides a product, it must have divided one of the factors going in. This rigidity is what makes unique prime factorization possible.

Definition 20.35 (Least Common Multiple). Let a and b be nonzero integers. The least common multiple of a and b is the unique positive integer m such that:

1. $a \mid m$ and $b \mid m$ (so m is a common multiple), and
2. If c is any positive integer with $a \mid c$ and $b \mid c$, then $m \mid c$.

The least common multiple of a and b is denoted

$$\operatorname{lcm}(a, b).$$

Remark 20.36 (Intuition). Intuitively: the lcm is the smallest number that both a and b divide into evenly. Any other common multiple must itself be a multiple of the lcm.

Remark 20.37 (Relation with the Greatest Common Divisor). For nonzero integers a and b,

$$\gcd(a, b)\ \operatorname{lcm}(a, b) = |ab|.$$

Thus the greatest common divisor measures the shared divisibility of a and b, while the least common multiple measures their combined multiplicative content.

In prime factorization terms, gcd takes the minimum exponent of each prime, whereas lcm takes the maximum exponent.

Theorem 20.38 (Fundamental Theorem of Arithmetic). Every integer $n > 1$ is either prime or can be written as a product of primes. Moreover, this factorization is unique up to the order of the factors.

That is, if

$$n = p_1 p_2 \cdots p_r \quad \text{and} \quad n = q_1 q_2 \cdots q_s,$$

where each p_i and q_j is prime, then

$$r = s,$$

and after a reordering of the q_j,

$$p_i = q_i \quad \text{for all } i.$$

Theorem 20.39 (Structural Relation Between gcd and lcm). Let a and b be nonzero integers. Then

$$\gcd(a, b)\ \operatorname{lcm}(a, b) = |ab|.$$

Remark 20.40 (Intuition). Intuitively: a and b together contain a fixed amount of prime material. The gcd captures what they share; the lcm captures everything between them. Together they account for the full prime content of ab.

20.1.2.2 Consequences and Logical Implications

Remark 20.41 (Logical Dependency Chain). The development of the integers in this section follows the chain

$$\begin{aligned}\text{Well-Ordering Principle} &\Rightarrow \text{Division Algorithm} \Rightarrow \text{Bézout's Identity} \\ &\Rightarrow \text{Euclid's Lemma} \Rightarrow \text{Fundamental Theorem of Arithmetic.}\end{aligned}$$

Thus the order structure of $\mathbb{N}$ ultimately governs prime factorization in $\mathbb{Z}$.

Remark 20.42 (Equivalences). The following principles are logically equivalent:

- The Well-Ordering Principle.
- The Principle of Mathematical Induction.
- The Least Element Principle.

Each encodes the same structural property: the integers admit no infinite strictly descending chains.

Remark 20.43 (Divisibility Structure). Bézout's Identity upgrades divisibility into a linear-combination statement:

$$\gcd(a,b) = \min\{\, as + bt > 0 : s,t \in \mathbb{Z} \,\}.$$

Thus the gcd is characterized by a universal property: it is the greatest element (under divisibility) among common divisors of a and b.

Remark 20.44 (Prime Structure). Euclid's Lemma implies that primes behave rigidly under multiplication:

$$p \mid ab \Rightarrow p \mid a \text{ or } p \mid b.$$

This rigidity is exactly what makes unique prime factorization possible. Without Euclid's Lemma, uniqueness would fail.

Remark 20.45 (Arithmetic Decomposition). The Fundamental Theorem of Arithmetic shows that every integer $n > 1$ decomposes uniquely into prime powers. Thus $\mathbb{Z}$ is a Unique Factorization Domain (UFD).

Remark 20.46 (Duality of gcd and lcm). The identity

$$\gcd(a,b)\ \operatorname{lcm}(a,b) = |ab|$$

reveals a structural duality:

- gcd measures shared prime factors (intersection).
- lcm measures total prime coverage (union).

Together they partition the prime-power structure of ab.

Remark 20.47 (Structural Position). This section establishes that:

- $\mathbb{Z}$ has a well-ordered positive part.
- Divisibility can be analyzed via minimal elements.
- Primes control multiplicative structure.
- Linear combinations control common divisors.

These ideas generalize later to:

- Euclidean domains,
- Principal ideal domains,
- Unique factorization domains.

20.1.3 Modular Arithmetic

20.1.3.1 Basic Definitions and Theorems

Definition 20.48 (Remainder (Modular Reduction)). Let $a \in \mathbb{Z}$ and $n \in \mathbb{Z}^+$. We say $a \bmod n$ is the unique integer r satisfying

$$a = qn + r, \quad 0 \leq r < n,$$

for some integer q. The value of $a \bmod n$ is r. We call r the remainder, q the quotient, and n the modulus of a upon division by n.

Remark 20.49 (Computing $a \bmod n$). To compute $a \bmod n$, apply the Division Algorithm: divide a by n, and the remainder r is the value of $a \bmod n$.

Examples.

- $17 \bmod 5 = 2$, since $17 = 3 \cdot 5 + 2$.
- $20 \bmod 4 = 0$, since $20 = 5 \cdot 4 + 0$.
- $3 \bmod 7 = 3$, since $3 = 0 \cdot 7 + 3$.
- $-1 \bmod 5 = 4$, since $-1 = (-1) \cdot 5 + 4$ and $0 \leq 4 < 5$.

Note the last example carefully: for negative integers, the remainder is still required to satisfy $0 \leq r < n$, so the quotient q may be negative.

Equivalently, using the floor function,

$$a \bmod n = a - n \left\lfloor \frac{a}{n} \right\rfloor,$$

where $\lfloor x \rfloor$ denotes the greatest integer less than or equal to x. This formula computes $q = \lfloor a/n \rfloor$ and then recovers $r = a - qn$.

Remark 20.50 (Intuition). Intuitively: $a \bmod n$ asks "after fitting as many full copies of n into a as possible, what is left over?" The remainder is always in $[0, n)$ regardless of the sign of a.

Lemma 20.51 (Idempotence of mod). For any $a \in \mathbb{Z}$ and $n \in \mathbb{Z}^+$,

$$(a \bmod n) \bmod n = a \bmod n.$$

Remark 20.52 (Intuition). Intuitively: once a value has been reduced modulo n, it is already in the range $[0, n)$, so reducing again changes nothing.

Remark 20.53 (Applying mod$n$ to an Equation Term by Term). A key technique in modular arithmetic is reducing an entire equation modulo n by applying mod$n$ to each term. This works because multiples of n vanish under mod$n$:

$$kn \bmod n = 0 \quad \text{for any } k \in \mathbb{Z}.$$

General procedure. Write each integer using the Division Algorithm:

$$a = q_1 n + r_1, \qquad b = q_2 n + r_2.$$

Then any expression built from a and b using addition and multiplication can be reduced modulo n by discarding all multiples of n and retaining only the remainders.

Example: addition.

$$a + b = (q_1 + q_2)n + (r_1 + r_2),$$

so

$$(a + b) \bmod n = (r_1 + r_2) \bmod n = (a \bmod n + b \bmod n) \bmod n.$$

The extra outer mod$n$ is needed because $r_1 + r_2$ may exceed n.

Example: multiplication.

$$ab = (q_1 n + r_1)(q_2 n + r_2) = q_1 q_2 n^2 + q_1 r_2 n + q_2 r_1 n + r_1 r_2.$$

Every term except $r_1 r_2$ is a multiple of n, so

$$(ab) \bmod n = (r_1 r_2) \bmod n = (a \bmod n)(b \bmod n) \bmod n.$$

Concrete example. Compute $(17 \cdot 13) \bmod 5$.

$$17 \bmod 5 = 2, \qquad 13 \bmod 5 = 3, \qquad 2 \cdot 3 = 6, \qquad 6 \bmod 5 = 1.$$

Check: $17 \cdot 13 = 221$ and $221 = 44 \cdot 5 + 1$, so $221 \bmod 5 = 1$. ✓

Why this works. Multiples of n contribute 0 under mod$n$. So when expanding any polynomial expression in a and b, all cross-terms involving n vanish, leaving only the product of remainders. This is formalized in the Compatibility with Arithmetic theorem below.

Remark 20.54 (Intuition). Intuitively: reducing modulo n is like working on a clock face with n positions — you only ever care about where you land, not how many full laps you made to get there. Addition and multiplication just move you around the clock, and multiples of n bring you back to 0.

Definition 20.55 (Congruence Modulo n). Let n be a positive integer. For integers a and b, we say that

$$a \equiv b \pmod{n}$$

if n divides $a - b$; that is,

$$n \mid (a - b).$$

This relation is called congruence modulo n.

Remark 20.56 (Unpacking the Definition). The statement

$$a \equiv b \pmod{n}$$

means that a and b leave the same remainder upon division by n.

Equivalently, there exists an integer k such that

$$a = b + kn.$$

Thus two integers are congruent modulo n precisely when they differ by a multiple of n.

Remark 20.57 (Intuition). Intuitively: two integers are congruent modulo n if they land on the same position when you wrap the number line around a circle of circumference n. The difference between them is always an exact number of full laps.

Definition 20.58 (Congruence Class). The congruence class of a modulo n is the set of all integers congruent to a modulo n:

$$[a]_n = \{a + kn : k \in \mathbb{Z}\}.$$

Remark 20.59. The remainder r in the Division Algorithm $a = bq + r$ is also called the residue of a modulo b, and $[a]_n$ is also called a residue class. The collection of all congruence classes modulo n forms the set

$$\mathbb{Z}/n\mathbb{Z} = \big\{\, [0]_n,\, [1]_n,\, \ldots,\, [n-1]_n \,\big\}.$$

$|\mathbb{Z}/n\mathbb{Z}| = n$, meaning $\mathbb{Z}/n\mathbb{Z}$ has exactly n elements, corresponding to the n possible remainders upon division by n.

Remark 20.60 (Intuition). Intuitively: congruence classes partition all of $\mathbb{Z}$ into n equally spaced infinite families. Every integer belongs to exactly one class, determined by its remainder upon division by n.

Theorem 20.61 (Congruence is an Equivalence Relation). Let n be a positive integer. Congruence modulo n is an equivalence relation on $\mathbb{Z}$. That is, for all $a, b, c \in \mathbb{Z}$:

1. Reflexivity: $a \equiv a \pmod{n}$.
2. Symmetry: If $a \equiv b \pmod{n}$, then $b \equiv a \pmod{n}$.
3. Transitivity: If $a \equiv b \pmod{n}$ and $b \equiv c \pmod{n}$, then $a \equiv c \pmod{n}$.

Remark 20.62 (Intuition). Intuitively: an equivalence relation is just a precise way of saying "these things behave the same for our purposes." Congruence mod n groups integers by their remainder — same remainder, same class, same behavior under arithmetic mod n.

Theorem 20.63 (Compatibility with Arithmetic). Let n be a positive integer and suppose

$$a \equiv b \pmod{n} \quad \text{and} \quad c \equiv d \pmod{n}.$$

Then:

1. $a + c \equiv b + d \pmod{n}$,
2. $a - c \equiv b - d \pmod{n}$,
3. $ac \equiv bd \pmod{n}$.

Remark 20.64 (Intuition). Intuitively: you can swap any integer for a congruent one before doing arithmetic, and the result will still be congruent to what you would have gotten. This is what makes working with representatives of classes — rather than the classes themselves — legitimate.

Remark 20.65 (Consequence for $\mathbb{Z}/n\mathbb{Z}$). The compatibility theorem means that addition and multiplication of congruence classes are well-defined:

$$[a]_n + [b]_n := [a + b]_n, \qquad [a]_n \cdot [b]_n := [ab]_n.$$

The choice of representative within a class does not affect the result. Under these operations, $\mathbb{Z}/n\mathbb{Z}$ is a commutative ring.

20.1.3.2 Consequences and Logical Implications

Remark 20.66 (Connection to the Integers). Congruence modulo n is the first instance of a general construction: forming a quotient of $\mathbb{Z}$ by identifying elements that differ by a multiple of n. The set $\mathbb{Z}/n\mathbb{Z}$ is the quotient of $\mathbb{Z}$ by the subgroup $n\mathbb{Z} = \{\ldots, -2n, -n, 0, n, 2n, \ldots\}$.

This construction depends directly on the Division Algorithm: every integer belongs to exactly one congruence class modulo n, corresponding to its remainder upon division by n.

Remark 20.67 (When $\mathbb{Z}/n\mathbb{Z}$ is a Field). The ring $\mathbb{Z}/n\mathbb{Z}$ is a field if and only if n is prime.

When $n = p$ is prime, every nonzero element $[a]_p$ satisfies $\gcd(a, p) = 1$, so by Bézout's Identity there exist integers s, t with $as + pt = 1$, giving $[a]_p \cdot [s]_p = [1]_p$. Thus every nonzero element has a multiplicative inverse.

When n is composite, say $n = ab$ with $1 < a, b < n$, the element $[a]_n$ is a zero divisor: $[a]_n \cdot [b]_n = [0]_n$ with neither factor zero, so $\mathbb{Z}/n\mathbb{Z}$ is not an integral domain and hence not a field.

Remark 20.68 (Intuition). Intuitively: a field requires every nonzero element to have a multiplicative inverse. In $\mathbb{Z}/n\mathbb{Z}$, an element $[a]_n$ has an inverse exactly when a and n share no common factor — and when n is prime, this is true for every nonzero element.

Remark 20.69 (Structural Position). Modular arithmetic occupies the boundary between the integers and abstract algebra:

- It depends on the Division Algorithm and gcd theory developed in the integers section.
- It is the prototype for quotient ring constructions, which generalize to arbitrary ideals in commutative ring theory.
- The fields $\mathbb{Z}/p\mathbb{Z}$ are the simplest finite fields, foundational in algebra, number theory, and cryptography.

20.1.4 Induction

20.1.4.1 Definitions and Theorems

Definition 20.70 (Principle of Mathematical Induction). Let $a \in \mathbb{Z}$ and let $S \subseteq \mathbb{Z}$ be a set of integers such that $a \in S$.

Suppose S has the property that whenever an integer $n \geq a$ belongs to S, then $n+1$ also belongs to S. That is,

$$n \in S \;\Rightarrow\; n+1 \in S \quad \text{for all } n \geq a.$$

Then S contains every integer greater than or equal to a. In other words,

$$\{n \in \mathbb{Z} : n \geq a\} \subseteq S.$$

Definition 20.71 (Mathematical Induction Logical Form). Let $a \in \mathbb{Z}$ and let $P(n)$ be a predicate defined for integers $n \geq a$.

Suppose:

1. Base case:
$$P(a) \text{ is true.}$$
2. Inductive step:
$$\forall n \geq a \, \big(P(n) \Rightarrow P(n+1)\big).$$

Then

$$\forall n \geq a, \; P(n) \text{ is true.}$$

Remark 20.72 (Equivalence with the Well-Ordering Principle). The Principle of Mathematical Induction is logically equivalent to the Well-Ordering Principle.

Equivalence idea:

- Well-Ordering $\Rightarrow$ Induction: If a statement $P(n)$ were false for some $n \geq a$, then the set of counterexamples would be nonempty. By the Well-Ordering Principle, it would contain a smallest element. This smallest counterexample contradicts the inductive hypothesis.
- Induction $\Rightarrow$ Well-Ordering: If a nonempty set of integers had no smallest element, one can use induction to show that no integer belongs to it, yielding a contradiction.

Thus induction and well-ordering are two formulations of the same fundamental structural property of the integers.

Remark 20.73 (Logical Structure of Induction). The Principle of Mathematical Induction has the quantifier pattern

$$(\text{initial truth}) \;\wedge\; (\text{truth propagates}) \;\Rightarrow\; (\text{universal truth}).$$

More precisely,

$$\big(P(a) \ \wedge \ \forall n \geq a\,(P(n) \Rightarrow P(n+1))\big) \ \Rightarrow \ \forall n \geq a\,P(n).$$

Thus induction is a mechanism for converting a local implication ($P(n) \Rightarrow P(n+1)$) into a global conclusion ($P(n)$ holds for all $n \geq a$).

It upgrades step-by-step propagation into a universal statement.

Theorem 20.74 (Second Principle of Mathematical Induction (Strong Induction)). Let $a \in \mathbb{Z}$ and let $S \subseteq \mathbb{Z}$ be a set of integers such that $a \in S$.

Suppose S has the following property: for every integer $n \geq a$,

$$\big(\forall k \in \mathbb{Z}\ (a \leq k < n \Rightarrow k \in S)\big) \ \Rightarrow \ n \in S.$$

Then

$$\{\, n \in \mathbb{Z} : n \geq a \,\} \subseteq S,$$

i.e., S contains every integer greater than or equal to a.

Remark 20.75 (Weak vs. Strong Induction). There are two common formulations of induction:

Weak (ordinary) induction:

$$P(a) \ \wedge \ \forall n \geq a\,(P(n) \Rightarrow P(n+1)) \ \Rightarrow \ \forall n \geq a\,P(n).$$

Strong induction:

$$P(a) \ \wedge \ \forall n > a \left(\big(\forall k\ (a \leq k < n \Rightarrow P(k))\big) \Rightarrow P(n)\right) \ \Rightarrow \ \forall n \geq a\,P(n).$$

Difference: Weak induction assumes only the immediately preceding case $P(n)$ to prove $P(n+1)$. Strong induction assumes all earlier cases $P(a), \ldots, P(n-1)$ to prove $P(n)$.

Use cases:

- Weak induction is natural when the proof of $P(n+1)$ depends only on $P(n)$.
- Strong induction is appropriate when the proof of $P(n)$ requires information about several earlier values, such as in factorization arguments (e.g., the Fundamental Theorem of Arithmetic) or recursive definitions.

Although they appear different, weak and strong induction are logically equivalent; each can be derived from the other. They are two formulations of the same structural property of the integers.

20.1.4.2 Consequences

Remark 20.76 (Logical Equivalence). The following principles are logically equivalent:

- The Well-Ordering Principle.
- The Principle of Mathematical Induction.
- The Second (Strong) Principle of Mathematical Induction.
- The Least Element Principle.

Each encodes the same structural property of the integers: every nonempty subset of $\mathbb{N}$ has a minimal element, and there are no infinite strictly descending chains.

Remark 20.77 (Induction as Minimal Counterexample Argument). Induction may be reformulated as a minimal counterexample principle.

To prove a statement $P(n)$ for all $n \geq a$, it suffices to assume that a counterexample exists, choose the smallest such counterexample using well-ordering, and derive a contradiction.

Thus induction and minimal-counterexample arguments are two views of the same logical engine.

Remark 20.78 (Recursive Definitions). Induction justifies recursive constructions.

If a function or object is defined by:

- specifying its value at a, and
- specifying how to construct its value at $n+1$ from its value at n,

then induction guarantees that the definition extends uniquely to all $n \geq a$.

Thus induction underlies the construction of:

- exponentiation,
- factorials,
- recursively defined sequences,
- algorithms on the integers.

Remark 20.79 (Structural Position). Induction is not merely a proof technique; it characterizes the order structure of the integers.

It upgrades a local propagation rule

$$P(n) \Rightarrow P(n+1)$$

into a global universal conclusion

$$\forall n \geq a \, P(n).$$

This mechanism is the prototype for:

- Euclidean-domain arguments (minimal element methods),
- termination proofs in algorithms,
- structural recursion in algebra and logic.

Remark 20.80 (Failure Outside Well-Ordered Sets). Induction depends essentially on the well-ordering of $\mathbb{N}$. In sets that admit infinite descending chains (such as $\mathbb{Z}$ under the usual order), induction in this form fails.

Thus induction is a structural consequence of well-ordering, not a purely algebraic property.

Remark 20.81 (Logical Flow).

$$\text{Well-Ordering} \iff \text{Induction} \iff \text{Strong Induction}$$

These are different formulations of the same foundational order-theoretic principle governing the integers.

20.1.4.3 Consequences

Corollary 20.82. <Immediate consequence.>

Remark 20.83 (Structural Insight). <Explain what this section reveals about the structure.>

Remark 20.84 (Logical Structure).

$$\text{Local Definitions} + \text{Theorems} \Rightarrow \text{Structural Consequences.}$$

20.1.5 Relations and Functions

20.1.5.1 Definitions and Theorems

Definition 20.85 (Binary Relation). Let A and B be sets. A binary relation from A to B is a subset

$$R \subseteq A \times B.$$

If $(a, b) \in R$, we write

$$a \, R \, b.$$

If $A = B$, we call R a relation on A.

Definition 20.86 (Properties of Relations). Let R be a relation on a set S.

- Reflexive:

$$\forall a \in S, \ (a, a) \in R.$$

- Symmetric:
$$(a, b) \in R \Rightarrow (b, a) \in R.$$
- Transitive:
$$(a, b) \in R \wedge (b, c) \in R \Rightarrow (a, c) \in R.$$
- Antisymmetric:
$$(a, b) \in R \wedge (b, a) \in R \Rightarrow a = b.$$

Definition 20.87 (Equivalence Relation). A relation R on S is an equivalence relation if it is reflexive, symmetric, and transitive.

Definition 20.88 (Equivalence Class). Let R be an equivalence relation on S. For $a \in S$, the equivalence class of a is

$$[a] := \{\, x \in S : (x, a) \in R\} = \{\, x \in S : xRa \,\}.$$

Theorem 20.89 (Equivalence Relations and Partitions). Let R be an equivalence relation on a set S. Then the set of equivalence classes of R forms a partition of S.

Conversely, every partition of S determines an equivalence relation on S.

Definition 20.90 (Function (Mapping)). Let A and B be sets. A function from A to B is a relation $f \subseteq A \times B$ such that:

1. For every $a \in A$, there exists $b \in B$ with $(a, b) \in f$.
2. If $(a, b_1) \in f$ and $(a, b_2) \in f$, then $b_1 = b_2$.

We write $f : A \to B$.

Definition 20.91 (Injective, Surjective, Bijective). Let $f : A \to B$.

- Injective:
$$f(a_1) = f(a_2) \Rightarrow a_1 = a_2.$$
- Surjective:
$$\forall b \in B,\ \exists a \in A \text{ such that } f(a) = b.$$
- Bijective: both injective and surjective.

Definition 20.92 (Image and Preimage). Let $f : A \to B$.

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

For $C \subseteq B$,
$$f^{-1}(C) = \{a \in A : f(a) \in C\}.$$

Definition 20.93 (Composition of Functions). Let $f : A \to B$ and $g : B \to C$. Define

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

Theorem 20.94 (Basic Properties of Function Composition). Let

$$a : A \to B, \quad b : B \to C, \quad g : C \to D.$$

1. Associativity:
$$g \circ (b \circ a) = (g \circ b) \circ a.$$
2. If a and b are injective, then $b \circ a$ is injective.
3. If a and b are surjective, then $b \circ a$ is surjective.
4. If a is bijective, then there exists $a^{-1} : B \to A$ such that
$$a^{-1} \circ a = \mathrm{id}_A, \quad a \circ a^{-1} = \mathrm{id}_B .$$

20.1.5.2 Consequences

Remark 20.95 (Structural Map).

$$\text{Relation} \supset \text{Function} \supset \text{Bijective Function}.$$

Remark 20.96 (Equivalence Relations Produce Quotients). If R is an equivalence relation on S, then

$$S/R = \{[a] : a \in S\}$$

is the associated quotient set.

Remark 20.97 (Partitions vs. Structure). Equivalence relations encode structural indistinguishability. Partitions encode decomposition. These viewpoints are equivalent.

Remark 20.98 (Upgrade Path). These concepts reappear as:

- Congruence modulo n,
- Cosets and quotient groups,
- Kernels of homomorphisms,
- Quotient spaces in topology.

Example Proof

$ab = \text{lcm}(a, b)\ \gcd(a, b)$

(Three-column format: tag / step / justification)

DU = Definition Unpacked TA = Theorem Applied AM = Algebraic Manipulation

Theorem 20.99 ($ab = \text{lcm}(a, b)\ \gcd(a, b)$). Let a and b be positive integers. Then

$$ab \;=\; \text{lcm}(a, b)\ \gcd(a, b).$$

Proof.

Part 1: Setup — write a and b in terms of their gcd.

Tag	Step	Justification
DU	Let $d = \gcd(a, b)$.	Definition of gcd; d is the greatest common divisor of a and b.
DU	Write $a = da_1$ and $b = db_1$ for positive integers a_1, b_1.	Since $d \mid a$ and $d \mid b$, we can factor d out of each.
TA	$\gcd(a_1, b_1) = 1$.	(Sub-argument by contradiction — see box below.)

Sub-argument (contradiction): Why must $\gcd(a_1, b_1) = 1$?

Tag	Step	Justification
DU	Suppose $\gcd(a_1, b_1) = c > 1$.	Assume for contradiction that a_1 and b_1 share a common factor.
DU	Then $c \mid a_1$ and $c \mid b_1$.	Definition of common divisor.
AM	So $dc \mid da_1 = a$ and $dc \mid db_1 = b$.	Multiplying both sides of each divisibility by d.
DU	Thus dc is a common divisor of a and b.	It divides both a and b.
AM	$dc > d$.	Since $c > 1$.
DU	This contradicts $d = \gcd(a, b)$.	d is the greatest common divisor, so no common divisor can exceed d. Contradiction. $\square$

Therefore $\gcd(a_1, b_1) = 1$; that is, a_1 and b_1 are coprime.

Part 2: Identify $\text{lcm}(a, b)$.

Tag	Step	Justification
DU	Claim: $\operatorname{lcm}(a,b) = da_1b_1$.	We verify both conditions in the definition of lcm.
DU	$a \mid da_1b_1$ since $da_1b_1 = a \cdot b_1$.	$a = da_1$, so $da_1b_1 = (da_1)b_1 = ab_1$. Hence $a \mid da_1b_1$.
DU	$b \mid da_1b_1$ since $da_1b_1 = b \cdot a_1$.	$b = db_1$, so $da_1b_1 = a_1(db_1) = a_1b$. Hence $b \mid da_1b_1$.
TA	If $a \mid c$ and $b \mid c$, then $da_1b_1 \mid c$.	Write $c = au = da_1u$ and $c = bv = db_1v$. Then $a_1u = b_1v$. Since $\gcd(a_1,b_1) = 1$, we get $b_1 \mid u$, say $u = b_1w$. Then $c = da_1b_1w$, so $da_1b_1 \mid c$.
DU	$\therefore \operatorname{lcm}(a,b) = da_1b_1$.	Both conditions of the lcm definition are satisfied.

Part 3: Compute $\operatorname{lcm}(a,b)\ \gcd(a,b)$.

Tag	Step	Justification
AM	$\operatorname{lcm}(a,b)\ \gcd(a,b) = (da_1b_1) \cdot d$	Substituting $\operatorname{lcm}(a,b) = da_1b_1$ and $\gcd(a,b) = d$.
AM	$= d^2a_1b_1$	Collecting the two factors of d.
AM	$= (da_1)(db_1)$	Regrouping.
AM	$= ab$	Since $a = da_1$ and $b = db_1$.

Therefore $ab = \operatorname{lcm}(a,b)\ \gcd(a,b)$. ■

Study Notes. Where did each tool appear?

Tool	Role in this proof
Definition of gcd	Gave us $d \mid a$ and $d \mid b$, allowing us to write $a = da_1$, $b = db_1$.
Proof by contradiction	Embedded sub-argument showing $\gcd(a_1,b_1) = 1$. The key: if they shared a factor $c > 1$, then $dc > d$ would be a common divisor, contradicting d being greatest.
Definition of lcm (universal property)	Used to verify the candidate da_1b_1 rather than just assert it. Required checking both (i) it is a common multiple, and (ii) it divides every common multiple.
Coprimality of a_1, b_1	The essential ingredient in Part 2 that makes the minimality argument work: $\gcd(a_1,b_1) = 1$ forced $b_1 \mid u$.
Algebraic regrouping	The final computation is just $d^2a_1b_1 = (da_1)(db_1)$. All the work was in setting up the right objects.

Proof GCD is a linear combination

$\gcd(a, b) = as + bt$ for some $s, t \in \mathbb{Z}$

(Three-column format: tag / step / justification)

DU = Definition Unpacked TA = Theorem Applied AM = Algebraic Manipulation

Theorem 20.100 (GCD as Smallest Positive Linear Combination). Let a and b be positive integers. Then $\gcd(a, b)$ is the smallest positive element of the set

$$S = \{am + bn \mid m, n \in \mathbb{Z},\ am + bn > 0\}.$$

Proof.

Part 1: S is nonempty and has a smallest member.

Tag	Step	Justification
DU	$S = \{am + bn \mid m, n \in \mathbb{Z},\ am + bn > 0\}$.	Define S as the set of all positive integer linear combinations of a and b.
DU	$S \neq \emptyset$.	If some m, n give $am + bn < 0$, replace m, n with $-m, -n$ to get a positive value.
TA	Let $d = as + bt$ be the smallest member of S.	By the Well-Ordering Principle, every nonempty set of positive integers has a least element.

Part 2: d divides both a and b.

Tag	Step	Justification
TA	Write $a = dq + r, \quad 0 \leq r < d$.	Division Algorithm applied to a and d.
AM	$r = a - dq = a - (as + bt)q = a(1 - sq) + b(-tq)$.	Substituting $d = as + bt$ and expanding.
DU	If $r > 0$, then $r \in S$.	r is a positive integer linear combination of a and b.
TA	$r < d$, contradicting $d = \min S$.	But $r < d$ by the Division Algorithm, contradicting minimality of d.
DU	$\therefore\ r = 0$, so $d \mid a$.	The remainder must be zero; hence d divides a.
DU	$d \mid b$.	By the same argument applied symmetrically to b.
DU	$\therefore\ d$ is a common divisor of a and b.	d divides both a and b.

Part 3: d is the greatest common divisor.

Tag	Step	Justification
DU	Let d' be any common divisor of a and b.	Suppose d' is an arbitrary common divisor; write $a = d'h$ and $b = d'k$.
AM	$d = as + bt = (d'h)s + (d'k)t = d'(hs + kt)$.	Substituting $a = d'h$ and $b = d'k$ into $d = as + bt$.
DU	$\therefore\ d' \mid d$.	d is an integer multiple of d', so $d' \leq d$.
DU	$\therefore\ d = \gcd(a, b)$.	d is a common divisor of a and b, and every other common divisor divides d, so d is greatest. ■

Study Notes. Where did each tool appear?

Tool	Role in this proof
Well-Ordering Principle	Guaranteed that S, being a nonempty set of positive integers, has a least element d.
Division Algorithm	Used to write $a = dq + r$ and derive that $r \in S$ if $r > 0$, forcing $r = 0$.
Proof by contradiction	If $r > 0$ then $r \in S$ with $r < d$, contradicting minimality of d.
Linear combination structure	The form $d = as + bt$ was essential in showing every common divisor d' satisfies $d' \mid d$.
Symmetry argument	After showing $d \mid a$, the identical argument applies to b without repeating the full proof.

Chapter 21

Algebraic Geometry

21.1 Notes

Where You Are in the Journey

Sets & Functions $\to$ $\mathbb{R}$ $\to$ Algebraic Structures $\to$ Abstract Algebra $\to$ Algebraic Geometry $\to$ $\cdots$

How we got here. Abstract algebra gave us the theory of polynomial rings and ideals. Algebraic geometry asks the geometric question: what do the solutions of polynomial equations look like? Every polynomial ideal determines a geometric object; every geometric object corresponds to an ideal.

What this chapter builds. We study affine varieties (solution sets of polynomial systems), the Nullstellensatz (the fundamental correspondence between ideals and varieties), and Gröbner bases as a computational tool for working with polynomial ideals.

Where this leads. Algebraic geometry in its modern form uses scheme theory and cohomology. This chapter provides the classical foundations.

21.1.1 Algebraic Geometry

Structural Roadmap

The development of algebraic geometry in this project follows the definition–theorem–structure architecture used throughout the analysis volumes. The primary driver is Beginning in Algebraic Geometry by Emily Clader and Dustin Ross. The emphasis is on the interplay between polynomial algebra and geometric structure, building from affine varieties through to projective geometry and culminating topics in modern algebraic geometry. Each major topic is organized as:

$$\text{Definitions} \longrightarrow \text{Main Theorems} \longrightarrow \text{Consequences and Structural Insight}$$

The global progression follows Clader–Ross in three stages:

1. Algebraic Foundations
   (a) Polynomial rings
2. Affine Algebraic Geometry
   (a) Varieties and ideals
   (b) Irreducibility of affine varieties
   (c) Coordinate rings
   (d) Polynomial maps
   (e) Proof of the Nullstellensatz
   (f) Dimension
   (g) Smoothness
   (h) Products
3. Projective Algebraic Geometry
   (a) Projective varieties
   (b) Maps of projective varieties
   (c) Quasiprojective varieties
   (d) Culminating topics

Remark 21.1. This treatment prioritizes geometric intuition grounded in rigorous commutative algebra. The book opens with polynomial rings as the algebraic foundation before any geometric objects are introduced, reflecting the philosophy that the algebra must be solid before the geometry can be understood.

Remark 21.2. The central objects of study are varieties and the maps between them. The Nullstellensatz is the pivotal theorem of the affine theory, establishing the precise correspondence between geometry and algebra that underlies all subsequent development.

Remark 21.3 (Structural Position). Algebraic geometry is developed building on the commutative algebra and linear algebra established in earlier chapters. The affine theory is developed completely before projective geometry is introduced, so that projective space can be understood as a natural compactification and generalization of affine space.

21.1.1.1 Preliminary Definitions

Definition 21.4 (Ring). A set R with two binary operations $+ : R \times R \to R$ and $\cdot : R \times R \to R$ is a ring if:

1. Additive structure: $(R, +)$ is an abelian group.

2. Associativity of multiplication:

$$(a \cdot b) \cdot c = a \cdot (b \cdot c) \quad \text{for all } a, b, c \in R.$$

3. Multiplicative identity: There exists $1 \in R$ such that

$$1 \cdot a = a \cdot 1 = a \quad \text{for all } a \in R.$$

4. Distributivity:

$$a \cdot (b + c) = a \cdot b + a \cdot c \quad \text{and} \quad (a + b) \cdot c = a \cdot c + b \cdot c \quad \text{for all } a, b, c \in R.$$

Remark 21.5. A ring generalizes a field by dropping the requirement that nonzero elements have multiplicative inverses, and by not requiring multiplication to be commutative. When multiplication is commutative, we call R a commutative ring.

Definition 21.6 (Commutative Ring). A ring R is commutative if

$$a \cdot b = b \cdot a \quad \text{for all } a, b \in R.$$

Remark 21.7 (Structural Summary). The hierarchy of algebraic structures encountered so far is:

Structure	Additive Group	Multiplicative Structure
Ring	Abelian	Associative, with identity
Commutative Ring	Abelian	Associative, commutative, with identity
Field	Abelian	Abelian on nonzero elements

Every field is a commutative ring, but not every commutative ring is a field.

21.1.1.2 Polynomial Rings

Definition 21.8 (Polynomial). Let R be a commutative ring. A polynomial in one variable over R is a formal expression

$$f = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0,$$

where $n \in \mathbb{N}$, the coefficients $a_0, a_1, \ldots, a_n \in R$, and x is a formal symbol called an indeterminate.

Remark 21.9. The indeterminate x is not a variable ranging over values in R. It is a formal placeholder that encodes the coefficient sequence. Two polynomials are equal if and only if all their coefficients are equal.

Definition 21.10 (Degree). Let $f = a_n x^n + \cdots + a_0$ be a polynomial over R. If $a_n \neq 0$, then the degree of f is

$$\deg(f) := n.$$

The coefficient a_n is called the leading coefficient of f. A polynomial with leading coefficient 1 is called monic.

Remark 21.11. The zero polynomial $f = 0$ has no leading coefficient. Its degree is left undefined, or assigned $-\infty$ by convention to preserve the identity $\deg(fg) = \deg(f) + \deg(g)$.

Definition 21.12 (Polynomial Ring). Let R be a commutative ring. The polynomial ring over R in one indeterminate x, denoted $R[x]$, is the set of all polynomials in x with coefficients in R, equipped with addition and multiplication defined by:

$$\left(\sum_i a_i x^i\right) + \left(\sum_i b_i x^i\right) := \sum_i (a_i + b_i) x^i,$$

$$\left(\sum_i a_i x^i\right) \cdot \left(\sum_j b_j x^j\right) := \sum_k \left(\sum_{i+j=k} a_i b_j\right) x^k.$$

Example 21.13. Let $R = \mathbb{Z}$ and consider

$$f = 2x^2 + 3x + 1, \qquad g = x + 4 \quad \in \mathbb{Z}[x].$$

Then:

$$f + g = 2x^2 + 4x + 5,$$
$$f \cdot g = 2x^3 + 11x^2 + 13x + 4.$$

Remark 21.14 (Structural Summary). With these operations, $R[x]$ is itself a commutative ring. The original ring R embeds into $R[x]$ as the constant polynomials.

Property	Holds in $R[x]$?
Commutative ring	Always
Field	Only in degenerate cases
R embeds in $R[x]$	Always

21.1.1.3 Polynomial Rings in Several Variables

Definition 21.15 (Polynomial Ring in n Variables). Let R be a commutative ring and let $n \in \mathbb{N}$. The polynomial ring in n variables over R, denoted

$$R[x_1, x_2, \ldots, x_n],$$

is defined inductively by

$$R[x_1, \ldots, x_n] := R[x_1, \ldots, x_{n-1}][x_n].$$

Its elements are finite sums of the form

$$f = \sum_{\alpha} a_\alpha \, x^\alpha,$$

where the sum ranges over multi-indices $\alpha = (\alpha_1, \ldots, \alpha_n) \in \mathbb{N}^n$, the coefficients $a_\alpha \in R$, and $x^\alpha := x_1^{\alpha_1} \cdots x_n^{\alpha_n}$.

Definition 21.16 (Monomial). A monomial in $R[x_1, \ldots, x_n]$ is a polynomial of the form

$$x^\alpha = x_1^{\alpha_1} \cdots x_n^{\alpha_n}$$

for some multi-index $\alpha \in \mathbb{N}^n$.

Definition 21.17 (Total Degree). The total degree of the monomial x^α is

$$|\alpha| := \alpha_1 + \alpha_2 + \cdots + \alpha_n.$$

The degree of a polynomial $f \in R[x_1, \ldots, x_n]$ is the maximum total degree among all monomials with nonzero coefficient.

Example 21.18. In $\mathbb{R}[x, y, z]$, the polynomial

$$f = 3x^2y + xy^2z - 5z^3$$

has three terms with total degrees 3, 4, and 3 respectively. Hence $\deg(f) = 4$.

Remark 21.19 (Relevance to Algebraic Geometry). Polynomial rings in several variables are the foundational algebraic object of algebraic geometry. The geometric objects studied in subsequent sections — affine varieties, ideals, coordinate rings — are all defined in terms of $k[x_1, \ldots, x_n]$ for a field k. The interplay between the algebra of $k[x_1, \ldots, x_n]$ and the geometry of its zero sets is the central theme of Clader–Ross.

21.1.1.4 Ideals

Definition 21.20 (Ideal). Let R be a commutative ring. A subset $I \subseteq R$ is an ideal of R if:

1. $0 \in I$,
2. $a + b \in I$ for all $a, b \in I$,
3. $r \cdot a \in I$ for all $r \in R$ and $a \in I$.

Definition 21.21 (Generated Ideal). Let R be a commutative ring and let $f_1, \ldots, f_m \in R$. The ideal generated by $f_1, \ldots, f_m$ is

$$\langle f_1, \ldots, f_m \rangle := \left\{ \sum_{i=1}^{m} r_i f_i : r_i \in R \right\}.$$

An ideal of this form is called finitely generated.

Definition 21.22 (Finitely Generated Ideal). An ideal $I \subseteq R$ is finitely generated if there exist $f_1, \ldots, f_m \in R$ such that $I = \langle f_1, \ldots, f_m \rangle$.

Definition 21.23 (Noetherian Ring). A commutative ring R is Noetherian if every ideal of R is finitely generated.

Theorem 21.24 (Hilbert Basis Theorem). If R is a Noetherian ring, then the polynomial ring $R[x]$ is also Noetherian.

In particular, $k[x_1, \ldots, x_n]$ is Noetherian for any field k.

Remark 21.25. The Hilbert Basis Theorem guarantees that every ideal in $k[x_1, \ldots, x_n]$ is finitely generated. This is a foundational finiteness result: it means every algebraic variety can be cut out by finitely many polynomial equations.

21.1.1.5 Quotient Rings

Definition 21.26 (Quotient Ring). Let R be a commutative ring and let $I \subseteq R$ be an ideal. The quotient ring R/I is the set of cosets

$$R/I := \{a + I : a \in R\},$$

equipped with addition and multiplication defined by

$$(a + I) + (b + I) := (a + b) + I,$$

$$(a + I) \cdot (b + I) := (a \cdot b) + I.$$

Remark 21.27. The quotient ring R/I is a commutative ring. Elements of R/I are equivalence classes under the relation $a \sim b \iff a - b \in I$.

Example 21.28. In $\mathbb{R}[x]$, consider the ideal $I = \langle x^2 + 1 \rangle$. The quotient ring

$$\mathbb{R}[x]/\langle x^2 + 1 \rangle \cong \mathbb{C},$$

since in the quotient, x satisfies $x^2 = -1$, which is precisely the defining relation of $i \in \mathbb{C}$.

Remark 21.29 (Structural Position). Quotient rings of polynomial rings are the coordinate rings of affine varieties, which are studied in depth in the next chapter. The passage

$$k[x_1, \ldots, x_n] \longrightarrow k[x_1, \ldots, x_n]/I$$

is the algebraic encoding of restricting from all of $\mathbb{A}^n$ to the variety $V(I)$.

Proof Completion Metrics

Chapter	Completed
Propositional Logic (Bjørndahl, Suppes)	24
Predicate Logic (Suppes Ch. 6–8)	12
Real Analysis (Abbott, Johar, Pons, Ross)	12
Algebraic Structures (groups, rings, fields)	8
Abstract Algebra (Gallian)	9
Axiom Systems / Tao Ch. 2	8
Linear Algebra (Axler)	5
Total	78

Glossary of Foundational Terms

Definition 21.30 (Logic). Logic is the formal study of valid reasoning, including the structure of statements and the rules by which conclusions may be derived from assumptions.

Definition 21.31 (Deduction). Deduction is the process of deriving conclusions from given assumptions by the application of accepted rules of reasoning.

Definition 21.32 (Inference). An inference is a single application of a rule of reasoning that produces a new statement from one or more given statements.

Definition 21.33 (Mathematical Axiom). A mathematical axiom is a statement assumed to be true without proof and used as a foundational starting point for logical deduction within a given mathematical theory.

Definition 21.34 (Proof). A proof is a finite sequence of statements, each of which is either a premise, an assumption, or follows from earlier statements by a permitted rule of inference, such that the final statement is the desired conclusion.

Definition 21.35 (Derivation). A derivation is an ordered sequence of formulas constructed according to the rules of a formal system, where each formula is either a premise, an assumption, or is obtained from previous formulas by a rule of inference.

Definition 21.36 (Assumption). An assumption is a statement temporarily introduced in a proof for the purpose of deriving further statements. Assumptions must be properly discharged before a proof is completed.

Definition 21.37 (Proposition). A proposition is a declarative statement that has a definite truth value, namely true or false.

Definition 21.38 (Truth Value). The truth value of a proposition is the value true or false assigned to it under a given interpretation.

Definition 21.39 (Truth-Functional). A connective is truth-functional if the truth value of the compound proposition it forms depends only on the truth values of its components.

Definition 21.40 (Tautology). A tautology is a proposition that is true under every possible interpretation of its sentential variables.

Definition 21.41 (Formula). A formula is a finite symbolic expression constructed from atomic formulas using logical connectives according to the formation rules of a logical language.

Definition 21.42 (Well-Formed Formula). A well-formed formula (WFF) is a formula constructed in strict accordance with the formation rules of a logical language.

Definition 21.43 (Sentential Variable). A sentential variable is a symbol that stands for an unspecified proposition.

Definition 21.44 (Sentential Form). A sentential form is a symbolic expression composed of sentential variables and logical connectives.

Definition 21.45 (Connective). A connective is a logical operator that combines propositions to form compound propositions.

Definition 21.46 (Negation). A negation is a connective that reverses the truth value of a proposition.

Definition 21.47 (Disjunction). A disjunction is a connective that forms a proposition true if at least one of its component propositions is true.

Definition 21.48 (Conditional Proposition). A conditional proposition is a compound proposition of the form "if P, then Q".

Definition 21.49 (Material Implication). Material implication is the truth-functional connective corresponding to conditional propositions, false only when the antecedent is true and the consequent is false.

Definition 21.50 (Antecedent). The antecedent is the proposition following "if" in a conditional.

Definition 21.51 (Consequent). The consequent is the proposition following "then" in a conditional.

Definition 21.52 (Converse). The converse of $P \to Q$ is $Q \to P$.

Definition 21.53 (Inverse). The inverse of $P \to Q$ is $\neg P \to \neg Q$.

Definition 21.54 (Contrapositive). The contrapositive of $P \to Q$ is $\neg Q \to \neg P$, and is logically equivalent to the original conditional.

Definition 21.55 (Biconditional Proposition). A biconditional proposition asserts that two propositions have the same truth value.

Definition 21.56 (Material Equivalence). Material equivalence is the truth-functional connective corresponding to biconditional propositions.

Definition 21.57 (Equivalent Propositions). Two propositions are equivalent if they have the same truth value under all interpretations.

Definition 21.58 (Argument). An argument consists of premises together with a conclusion claimed to follow from them.

Definition 21.59 (Valid Argument). An argument is valid if there is no interpretation under which all premises are true and the conclusion is false.

Definition 21.60 (Formal Derivation). A formal derivation is a syntactic sequence of formulas produced by rules of inference from premises and assumptions.

Definition 21.61 (Modus Ponens). (Modus Ponens; also called the Law of Detachment) From $P \to Q$ and P, infer Q.

Definition 21.62 (Modus Tollens). From $P \to Q$ and $\neg Q$, infer $\neg P$.

Definition 21.63 (Hypothetical Syllogism). From $P \to Q$ and $Q \to R$, infer $P \to R$.

Definition 21.64 (Disjunctive Syllogism). (Also called modus tollendo ponens) From $P \vee Q$ and $\neg P$, infer Q (and symmetrically).

Definition 21.65 (Addition). (Disjunction Introduction) From P, infer $P \vee Q$.

Definition 21.66 (Simplification). (Conjunction Elimination) From $P \wedge Q$, infer P (or Q).

Definition 21.67 (Adjunction). (Conjunction Introduction) From P and Q, infer $P \wedge Q$.

Definition 21.68 (Double Negation). From $\neg\neg P$, infer P, and conversely.

Definition 21.69 (Disjunction Elimination). (Proof by Cases) If $P \vee Q$ and R follows from each of P and Q, infer R.

Definition 21.70 (Conditional Proof). If assuming P leads to Q, infer $P \to Q$.

Definition 21.71 (Indirect Proof). (Reductio ad Absurdum) If assuming $\neg P$ leads to a contradiction, infer P.