Proof Writing

Last modified: 2026-11-11 23:25:19 (PST)

This page collects general advice on how to write proofs and derivations. The goal of a proof is not just to convince yourself that a result is true; it is to convince a *reader*, and to show them *why* it is true. Each principle on this page serves that goal. The symbols for logical entailment are listed under Definitions for proofs below.

Definitions for proofs

Definition 1 (Proposition, statement)**.** A **proposition**, or **statement**, is a declarative sentence that is either true or false (Wikipedia contributors 2026z).

Example 1 (Propositions, and a sentence that is not one)**.**

- "$2 + 3 = 5$" is a proposition, and it is true.
- "7 is even" is a proposition, and it is false.
- "$x > 2$" is not a proposition on its own: it is true when $x = 3$ and false when $x = 1$, so it has no truth value until x is given a value.

Definition 2 (Predicate)**.** A **predicate** (Wikipedia contributors 2026w) $P(x)$ is a declarative sentence about a variable x that becomes a proposition (Definition 1) whenever a value is substituted for x.

Example 2 (A predicate)**.** Let $P(x)$ be the predicate "$x > 2$", for real numbers x.

- $P(3)$ is the proposition "$3 > 2$", which is true.
- $P(1)$ is the proposition "$1 > 2$", which is false.

Definition 3 (Domain of discourse, universe of discourse)**.** The **domain of discourse** (or **universe of discourse**) of a predicate $P(x)$ is the set of values its variable x may take. A predicate in several variables has a domain of discourse for each variable.

Example 3 (The domain changes the truth value)**.**
Let $P(x)$ be the predicate "$x^2 = 2$". If its domain of discourse is the real numbers, then $P(x)$ is true for $x = \sqrt{2}$. If its domain of discourse is the integers, then $P(x)$ is false for every x, since the squares of the integers are $0, 1, 4, 9, ...$, and none of them is 2.

Definition 4 (Logical connective)**.** A **logical connective** (Wikipedia contributors 2026n) builds a new proposition (Definition 1) from one or two propositions P and Q, with a truth value that depends only on the truth values of P and Q. These notes use five, each defined next: not, and, or, implies, and if and only if.

Definition 5 (Not)**.** The logical connective (Definition 4) **not** (Wikipedia contributors 2026v) is written $\neg P$ ("not P"); it is true when P is false, and false when P is true.

Example 4 (Negating a true and a false proposition)**.** Let P be "$5 > 3$", which is true, and let Q be "$2 > 3$", which is false.

- $\neg P$, "it is not true that $5 > 3$", is false, because P is true.
- $\neg Q$, "it is not true that $2 > 3$", is true, because Q is false.

Definition 6 (And)**.** The logical connective (Definition 4) **and** is written (Wikipedia contributors 2026m) $P \wedge Q$ ("P and Q"); it is true when P and Q are both true, and false otherwise.

Example 5 (Two propositions joined by "and")**.** Let P be "$3 < 5$", which is true, let Q be "$5 < 4$", which is false, and let R be "$5 < 7$", which is true.

- $P \wedge R$ is true, because P and R are both true.
- $P \wedge Q$ is false, because Q is false.

Definition 7 (Or)**.** The logical connective (Definition 4) **or** is written (Wikipedia contributors 2026p) $P \vee Q$ ("P or Q"); it is true when at least one of P and Q is true, and false when both are false.

Example 6 (Two propositions joined by "or")**.** Let P be "$3 < 2$", which is false, let Q be "$5 < 7$", which is true, and let R be "$5 < 4$", which is false.

- $P \vee Q$ is true, because Q is true.
- $P \vee R$ is false, because P and R are both false.

Definition 8 (Implies)**.** The logical connective (Definition 4) **implies** is written (Wikipedia contributors 2026r) $P \Rightarrow Q$ ("if P, then Q", or "P implies Q"); it is false when P is true and Q is false, and true otherwise.

Example 7 (Implications with a true and a false premise)**.** Take $x = 3$.

- $x > 2 \Rightarrow x > 1$ is true: the premise $3 > 2$ is true and the conclusion $3 > 1$ is true.
- $x > 2 \Rightarrow x > 4$ is false: the premise $3 > 2$ is true, but the conclusion $3 > 4$ is false.
- $x > 5 \Rightarrow x > 4$ is true: the premise $3 > 5$ is false, so the implication is true whatever the conclusion.

Definition 9 (If and only if)**.** The logical connective (Definition 4) **if and only if** is written (Wikipedia contributors 2026l) $P \Leftrightarrow Q$ ("P if and only if Q"); it is true when P and Q have the same truth value, and false otherwise.

Example 8 ("If and only if" with equal and different truth values)**.** Take $x = 4$.

- "x is even $\Leftrightarrow x > 3$" is true: 4 is even and $4 > 3$, so both sides are true.
- "$x > 3 \Leftrightarrow x > 5$" is false: $4 > 3$ is true, but $4 > 5$ is false.

Example 9 (Truth values of connectives)**.** Let P be "$2 + 3 = 5$", which is true, and let Q be "7 is even", which is false.

- $\neg Q$, "7 is not even", is true, because Q is false.
- $P \wedge Q$ is false, because Q is false.
- $P \vee Q$ is true, because P is true.
- $P \Rightarrow Q$ is false, because P is true and Q is false.
- $Q \Rightarrow P$ is true, because Q is false.
- $P \Leftrightarrow Q$ is false, because P and Q have different truth values.

Definition 10 (Truth table)**.** A **truth table** (Wikipedia contributors 2026ad) for a proposition built from propositions $P_1, \ldots, P_n$ with logical connectives has one row for each of the 2^n combinations of truth values of $P_1, \ldots, P_n$, and gives the truth value of the proposition in each row.

Example 10 (A truth table for $P \Rightarrow Q$)**.** The proposition $P \Rightarrow Q$ is built from two propositions, so its truth table has $2^2 = 4$ rows. By the definition of implies, $P \Rightarrow Q$ is false only in the row where P is true and Q is false:

P	Q	$P \Rightarrow Q$
T	T	T
T	F	F
F	T	T
F	F	T

Definition 11 (Logically equivalent)**.** Two propositions A and B, each built from some or all of the propositions $P_1, \ldots, P_n$ with logical connectives, are **logically equivalent** (Wikipedia contributors 2026q), written $A \equiv B$, when they have the same truth value in every row of a truth table on $P_1, \ldots, P_n$.

Example 11 ($P \Rightarrow Q$ is logically equivalent to $\neg P \vee Q$)**.** Adding a column for $\neg P \vee Q$ to the truth table of Example 10 gives:

P	Q	$\neg P$	$\neg P \vee Q$	$P \Rightarrow Q$
T	T	F	T	T
T	F	F	F	F
F	T	T	T	T
F	F	T	T	T

The last two columns agree in every row, so $(P \Rightarrow Q) \equiv (\neg P \vee Q)$.

Theorem 1 (A biconditional is two conditionals)**.** *For propositions P and Q,*

$$(P \Leftrightarrow Q) \equiv ((P \Rightarrow Q) \wedge (Q \Rightarrow P)).$$

Proof

Proof. By the definitions of if and only if, implies and and, the truth table is:

P	Q	$P \Rightarrow Q$	$Q \Rightarrow P$	$(P \Rightarrow Q) \wedge (Q \Rightarrow P)$	$P \Leftrightarrow Q$
T	T	T	T	T	T
T	F	F	T	F	F
F	T	T	F	F	F
F	F	T	T	T	T

The last two columns match in every row, so the two propositions are logically equivalent. □

Example 12 (Proving an "if and only if" in two parts)**.**
Let n be an integer. By a biconditional is two conditionals, to prove "n is even if and only if $n+2$ is even", it is enough to prove both conditionals:

- If $n = 2m$, then $n + 2 = 2(m + 1)$, which is even.
- If $n + 2 = 2m$, then $n = 2(m - 1)$, which is even.

Theorem 2 (A cycle of conditionals)**.** *Let $P_1, P_2, \ldots, P_k$ be propositions. If $P_1 \Rightarrow P_2$, $P_2 \Rightarrow P_3$, ..., $P_{k-1} \Rightarrow P_k$ and $P_k \Rightarrow P_1$ are all true, then $P_1, \ldots, P_k$ all have the same truth value.*

Proof

Proof. There are two cases.

- None of $P_1, \ldots, P_k$ is true. Then all of them are false, so they all have the same truth value.
- Some P_i is true. By the definition of implies, for any P and Q, if $P \Rightarrow Q$ and P are both true, then Q is true. So $P_i \Rightarrow P_{i+1}$ makes P_{i+1} true, then $P_{i+1} \Rightarrow P_{i+2}$ makes P_{i+2} true, and so on around the cycle, with $P_k \Rightarrow P_1$ carrying the truth of P_k back to P_1. After at most $k - 1$ steps, every proposition in the list is true.

□

Example 13 (Three equivalent statements)**.**
Let n be an integer, and let:

- P_1 be "n is even";
- P_2 be "$n + 1$ is odd";
- P_3 be "$n + 2$ is even".

If $n = 2m$, then $n + 1 = 2m + 1$, so $P_1 \Rightarrow P_2$. If $n + 1 = 2m + 1$, then $n + 2 = 2(m + 1)$, so $P_2 \Rightarrow P_3$. If $n + 2 = 2m$, then $n = 2(m - 1)$, so $P_3 \Rightarrow P_1$. By a cycle of conditionals, the three propositions are all true or all false. Three conditionals were enough, where checking each pair both ways would take six.

Theorem 3 (De Morgan's law for a conjunction)**.** *For any propositions P and Q (Wikipedia contributors 2026f):*

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

Proof

Proof. Each side is built from P and Q, so by Definition 11 it is enough to compare the two sides in each of the four rows of a truth table:

P	Q	$\neg(P \wedge Q)$	$\neg P \vee \neg Q$
T	T	F	F
T	F	T	T
F	T	T	T
F	F	T	T

The third and fourth columns agree in every row. □

Theorem 4 (De Morgan's law for a disjunction)**.** *For any propositions P and Q (Wikipedia contributors 2026f):*

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

Proof

Proof. Each side is built from P and Q, so by Definition 11 it is enough to compare the two sides in each of the four rows of a truth table:

P	Q	$\neg(P \vee Q)$	$\neg P \wedge \neg Q$
T	T	F	F
T	F	F	F
F	T	F	F
F	F	T	T

The third and fourth columns agree in every row. □

Definition 12 (Tautology)**.** A **tautology** is a proposition that is true in every row of its truth table.

Example 14 ($P \vee \neg P$ is a tautology)**.**
By the definitions of or and not, the truth table of $P \vee \neg P$ is:

P	$\neg P$	$P \vee \neg P$
T	F	T
F	T	T

The last column is true in every row, so $P \vee \neg P$ is a tautology.

Definition 13 (Contradiction)**.** A **contradiction** is a proposition that is false in every row of its truth table.

Example 15 ($P \wedge \neg P$ is a contradiction)**.**
By the definitions of and and not, the truth table of $P \wedge \neg P$ is:

P	$\neg P$	$P \wedge \neg P$
T	F	F
F	T	F

The last column is false in every row, so $P \wedge \neg P$ is a contradiction.

Theorem 5 (Equivalence is a tautological biconditional)**.** *For propositions A and B, A and B are logically equivalent exactly when $A \Leftrightarrow B$ is a tautology.*

Proof

Proof. By the definition of if and only if, $A \Leftrightarrow B$ is true in a row of the truth table exactly when A and B have the same truth value in that row. So $A \Leftrightarrow B$ is true in every row exactly when A and B have the same truth value in every row, which is the definition of $A \equiv B$. □

Example 16 (A tautological biconditional)**.**
Example 11 showed that $P \Rightarrow Q$ and $\neg P \vee Q$ are logically equivalent. Adding a column for the biconditional to its truth table gives:

P	Q	$\neg P \vee Q$	$P \Rightarrow Q$	$(P \Rightarrow Q) \Leftrightarrow (\neg P \vee Q)$
T	T	T	T	T
T	F	F	F	T
F	T	T	T	T
F	F	T	T	T

The last column is true in every row, so the biconditional is a tautology, as equivalence as a tautological biconditional says it must be.

Theorem 6 (Double negation)**.** *For a proposition P,*

$$\neg\neg P \equiv P.$$

Proof

Proof. By the definition of not, the truth table is:

P	$\neg P$	$\neg\neg P$
T	F	T
F	T	F

The first and last columns match in both rows, so the two sides are logically equivalent. □

Example 17 (Removing a double negation)**.**
Let P be the proposition "7 is odd". Then $\neg\neg P$ is "it is not the case that 7 is not odd". By double negation, $\neg\neg P$ says the same thing as P: 7 is odd.

Definition 14 (Quantifier)**.** A **quantifier** (Wikipedia contributors 2026ab) turns a predicate $P(x)$ (Definition 2) about the elements x of a set S into a proposition (Definition 1). There are two, each defined next: the universal quantifier and the existential quantifier.

Definition 15 (Universal quantifier)**.** The **universal quantifier** (Wikipedia contributors 2026ae) "for all" is a quantifier (Definition 14): $\forall x \in S, P(x)$ ("for every x in S, $P(x)$") is true when $P(x)$ is true for every element x of S, and false otherwise.

Example 18 (A universally quantified proposition)**.** Let $P(x)$ be the predicate "x is even".

- "$\forall x \in \{2, 4, 6\}, P(x)$" is true, because 2, 4, and 6 are all even.
- "$\forall x \in \{2, 3, 4\}, P(x)$" is false, because $P(3)$ is false.

Definition 16 (Existential quantifier)**.** The **existential quantifier** (Wikipedia contributors 2026h) "there exists" is a quantifier (Definition 14): $\exists x \in S, P(x)$ ("there is an x in S such that $P(x)$") is true when $P(x)$ is true for at least one element x of S, and false otherwise.

Example 19 (An existentially quantified proposition)**.** Let $P(x)$ be the predicate "x is even".

- "$\exists x \in \{1, 2, 3\}, P(x)$" is true, because $P(2)$ is true.
- "$\exists x \in \{1, 3, 5\}, P(x)$" is false, because none of 1, 3, and 5 is even.

Example 20 (Quantified propositions)**.** Let $P(x)$ be the predicate "$x > 2$", for integers x.

- "$\exists x \in \{1, 2, 3\}, P(x)$" is true, because $P(3)$ is true.
- "$\forall x \in \{1, 2, 3\}, P(x)$" is false, because $P(1)$ is false.
- "$\forall x \in \{3, 4, 5\}, P(x)$" is true, because $3 > 2$, $4 > 2$, and $5 > 2$ are all true.

Theorem 7 (De Morgan's laws for quantifiers)**.** *For a set S and a predicate $P(x)$ whose domain of discourse is S:*

$$\neg(\forall x \in S, P(x)) \equiv \exists x \in S, \neg P(x)$$
$$\neg(\exists x \in S, P(x)) \equiv \forall x \in S, \neg P(x)$$

Proof

Proof.

1. By the definition of the universal quantifier, $\forall x \in S, P(x)$ is false exactly when $P(x)$ is false for at least one x in S. By the definition of the existential quantifier, that is exactly when $\exists x \in S, \neg P(x)$ is true.
2. By the definition of the existential quantifier, $\exists x \in S, P(x)$ is false exactly when $P(x)$ is false for every x in S. By the definition of the universal quantifier, that is exactly when $\forall x \in S, \neg P(x)$ is true.

□

Remark 1 (Why the two laws share a theorem). The two laws in De Morgan's laws for quantifiers are one rule seen from each side: negating a quantifier swaps $\forall$ and $\exists$ and moves the negation onto the predicate. De Morgan's laws for a conjunction and a disjunction of two propositions pair up the same way, and these laws extend them from two propositions to a whole set of them.

Example 21 (Negating "every integer is even")**.**
Let S be the integers, and let $P(x)$ be "x is even". By the first of De Morgan's laws for quantifiers, $\neg(\forall x \in S, P(x)) \equiv \exists x \in S, \neg P(x)$: the negation of "every integer is even" is "some integer is not even". By the second law, $\neg(\exists x \in S, P(x)) \equiv \forall x \in S, \neg P(x)$: the negation of "some integer is even" is "no integer is even".

Definition 17 (Inequality)**.** An **inequality** is a proposition (Definition 1) or a predicate (Definition 2) that compares two expressions a and b whose values are real numbers, in one of the forms:

- $a < b$ ("a is less than b");
- $a \leq b$ ("a is less than or equal to b");
- $a > b$ ("a is greater than b");
- $a \geq b$ ("a is greater than or equal to b").

Example 22 (Inequalities)**.**

- "$3 < 5$" is an inequality, and it is true.
- "$5 \leq 3$" is an inequality, and it is false.
- "$x \geq 2$" is an inequality and a predicate: it is true when $x = 2$ and false when $x = 1$.
- "$x = 2$" is not an inequality: it is an equality.

Definition 18 (Logical entailment, logical consequence, deductive consequence)**.** Propositions or predicates $P_1, \ldots, P_n$ (Definition 1, Definition 2) **logically entail** (Wikipedia contributors 2026o) a proposition or predicate Q when every choice of values for their variables that makes all of $P_1, \ldots, P_n$ true also makes Q true. Then Q is a **logical consequence**, or **deductive consequence**, of $P_1, \ldots, P_n$.

Example 23 (Entailment, and its failure)**.** Let x be an integer.

- "$x = 2$" entails "$x^2 = 4$": the only value that makes "$x = 2$" true is $x = 2$, and $2^2 = 4$.
- "$x > 3$" and "$x < 5$" together entail "$x = 4$": 4 is the only integer greater than 3 and less than 5.
- "$x^2 = 4$" does not entail "$x = 2$": the value $x = -2$ makes "$x^2 = 4$" true, since $(-2)^2 = 4$, but makes "$x = 2$" false.

We can use any of:

- $\therefore$ (`\therefore` in LaTeX),
- $\Rightarrow$ (`\Rightarrow`),
- $\models$ (`\models`)

to denote logical entailment. $\Rightarrow$ also writes the connective "implies"; P entails Q exactly when $P \Rightarrow Q$ is true for every choice of values for the variables in P and Q.

Let's save $\to$ (`\rightarrow`) for convergence results.

Definition 19 (Logical argument)**.** A **logical argument** is a list of propositions $P_1, \dots, P_n$ together with one more proposition Q. It is written with "so" or $\therefore$ before Q.

Definition 20 (Premise)**.** The **premises** of a logical argument $P_1, \dots, P_n \therefore Q$ are the propositions $P_1, \dots, P_n$.

Definition 21 (Conclusion)**.** The **conclusion** of a logical argument $P_1, \dots, P_n \therefore Q$ is the proposition Q.

Example 24 (A logical argument)**.**
"If n is even, then n^2 is even. n is even. So n^2 is even." is a logical argument. Its premises are "if n is even, then n^2 is even" and "n is even". Its conclusion is "n^2 is even".

Definition 22 (Rule of inference)**.** A **rule of inference** is a pattern of logical argument, written with propositions $P, Q, \dots$ as placeholders, whose premises logically entail its conclusion whatever propositions are put in place of $P, Q, \dots$. It is written with its premises above a line and its conclusion below:

$$\frac{P_1, \dots, P_n}{Q}.$$

Example 25 (A pattern that is not a rule of inference)**.**
The pattern

$$\frac{P \Rightarrow Q, \quad Q}{P}$$

is not a rule of inference. In the row of the truth table where P is false and Q is true, both premises are true, by the definition of implies, but the conclusion P is false. For example, "if $n = 2$, then n is even" and "n is even" are both true for $n = 4$, but "$n = 2$" is false.

Theorem 8 (Modus ponens)**.** *For propositions P and Q, this is a rule of inference:*

$$\frac{P \Rightarrow Q, \quad P}{Q}$$

Proof

Proof. By the definition of implies, $P \Rightarrow Q$ is false only when P is true and Q is false. Both premises are true, so P is true and $P \Rightarrow Q$ is true. So Q is not false, and the conclusion Q is true. □

Theorem 9 (Modus tollens)**.** *For propositions P and Q, this is a rule of inference:*

$$\frac{P \Rightarrow Q, \quad \neg Q}{\neg P}$$

Proof

Proof. Suppose both premises are true. By the definition of not, Q is false. By the definition of implies, a true $P \Rightarrow Q$ with a false Q needs a false P. So P is false, and the conclusion $\neg P$ is true. □

Theorem 10 (Disjunctive syllogism)**.** *For propositions P and Q, this is a rule of inference:*

$$\frac{P \vee Q, \quad \neg P}{Q}$$

Proof

Proof. Suppose both premises are true. By the definition of not, P is false. By the definition of or, $P \vee Q$ is true only if P is true or Q is true. So the conclusion Q is true. □

Theorem 11 (Addition)**.** *For propositions P and Q, this is a rule of inference:*

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

Proof

Proof. Suppose the premise is true. By the definition of or, $P \vee Q$ is true when P is true. So the conclusion $P \vee Q$ is true. □

Theorem 12 (Simplification)**.** *For propositions P and Q, this is a rule of inference:*

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

Proof

Proof. Suppose the premise is true. By the definition of and, $P \wedge Q$ is true only if P and Q are both true. So the conclusion P is true. □

Theorem 13 (Conjunction)**.** *For propositions P and Q, this is a rule of inference:*

$$\frac{P, \quad Q}{P \wedge Q}$$

Proof

Proof. Suppose both premises are true. By the definition of and, $P \wedge Q$ is true when P and Q are both true. So the conclusion $P \wedge Q$ is true. □

Example 26 (Chaining rules of inference)**.**
Suppose these three premises are true:

- "if it rains, then the game is cancelled";
- "if the game is cancelled, then we go to the museum";
- "we do not go to the museum".

By modus tollens on the second and third premises, "the game is not cancelled" is true. By modus tollens again, on the first premise and the fact derived from them, "it does not rain" is true. So the conclusion "it does not rain" follows from the premises.

Definition 23 (Proof)**.** A **proof** (Wikipedia contributors 2026t) of a proposition (Definition 1) is a finite list of statements that ends with that proposition, in which each statement is

- an assumption,
- a definition,
- a previously proved result, or
- a consequence of earlier statements in the list, by one rule of logic or algebra.

Example 27 (The sum of two even integers is even)**.** Let a and b be even integers (definition). Then $a = 2j$ and $b = 2k$ for some integers j and k, and

$$\begin{aligned} a + b &= 2j + 2k && \text{(substitute } a = 2j \text{ and } b = 2k\text{)} \\ &= 2(j + k) && \text{(distributive law)} \end{aligned}$$

The second line uses the distributive law. Since $j+k$ is an integer, $a+b$ is even (definition). For example,

$$\begin{aligned} 4 + 10 &= 2 \cdot 2 + 2 \cdot 5 \\ &= 2 \cdot 7 \\ &= 14. \end{aligned}$$

Definition 24 (Derivation)**.** A **derivation** is a proof (Definition 23), or a part of one, written as a chain of expressions, one per line, in which each line is joined to the line before it by $=$ or by an inequality sign (Definition 17), and carries an annotation that names the reason that step holds: an assumption, a definition, a previously proved result, or a rule of algebra.

Example 28 (A derivation with equalities and inequalities)**.** For every real number x:

$$\begin{aligned} (x+1)^2 &= x^2 + 2x + 1 && \text{(square of a sum)} \\ &\geq 2x + 1 && (x^2 \geq 0) \\ &> 2x && (1 > 0) \end{aligned}$$

The first step uses the square of a sum. The steps are $=$, $\geq$, and $>$, so the derivation shows $(x+1)^2 > 2x$. For example, with $x = 3$: $(3+1)^2 = 16$, and $16 > 6 = 2 \cdot 3$.

Theorem 14 (What a derivation shows)**.** *Consider a derivation (Definition 24) with expressions $E_0, E_1, \ldots, E_n$, where each line E_j is joined to the line E_{j-1} before it by $=$, $\leq$ or $<$.*

1. *If every step is $=$, then $E_0 = E_n$.*
2. *If every step is $=$ or $\leq$, then $E_0 \leq E_n$.*
3. *If every step is $=$, $\leq$, or $<$, and at least one step is $<$, then $E_0 < E_n$.*

The same holds with $\geq$ and $>$ in place of $\leq$ and $<$.

Proof

Proof. First, a fact about two steps in a row. For numbers a, b and c, if $a\ R\ b$ and $b\ S\ c$, where R and S are each one of $=$, $\leq$ and $<$, then $a\ T\ c$, where T is

- $<$ if R or S is $<$;
- otherwise $\leq$ if R or S is $\leq$;
- otherwise $=$.

This holds in each case:

- If R and S are both $=$, then $a = c$ (equalities are transitive).
- If exactly one of R and S is $=$, replace b by the equal expression in the other relation (substitution). This gives the other relation between a and c.
- If neither is $=$, then each of $a\ R\ b$ and $b\ S\ c$ is $\leq$ or $<$, and $\leq$ means $<$ or $=$ (strict and non-strict inequalities). A case with an $=$ is covered by the previous bullet,

and the remaining case is $a < b$ and $b < c$, which gives $a < c$ because $<$ is transitive on the real numbers. So two steps that are each $\leq$ or $<$ give $<$ if either is $<$, and otherwise $\leq$.

Now let T_j be the relation named by the same rule applied to the first j steps of the chain, with T_0 being $=$. We show $E_0 \; T_j \; E_j$ by induction on j. For $j = 0$, $E_0 = E_0$. If $E_0 \; T_j \; E_j$, then the step from E_j to E_{j+1} and the fact above give $E_0 \; T_{j+1} \; E_{j+1}$. At $j = n$, T_n is $=$ when every step is $=$, which is part 1. It is $\leq$ or $=$ when every step is $=$ or $\leq$, and $=$ implies $\leq$, which is part 2. It is $<$ when some step is $<$, which is part 3.
The argument for $\geq$ and $>$ is the same with the inequalities reversed. □

Example 29 (Checking a derivation line by line)**.** The derivation in Example 28 has the steps $=$, $\geq$ and $>$, so by Theorem 14 it shows $(x+1)^2 > 2x$. The code evaluates the four lines of that derivation at several values of x, and checks each step and the comparison of the first line with the last.

```r
x <- c(-3, -1, 0, 2, 3)
lines <- data.frame(
  x = x,
  line_0 = (x + 1)^2,
  line_1 = x^2 + 2 * x + 1,
  line_2 = 2 * x + 1,
  line_3 = 2 * x
)
steps <- with(
  lines,
  data.frame(
    x = x,
    step_1_equal = line_0 == line_1,
    step_2_geq = line_1 >= line_2,
    step_3_greater = line_2 > line_3,
    first_greater_than_last = line_0 > line_3
  )
)
lines
#>    x line_0 line_1 line_2 line_3
#> 1 -3      4      4     -5     -6
#> 2 -1      0      0     -1     -2
#> 3  0      1      1      1      0
#> 4  2      9      9      5      4
#> 5  3     16     16      7      6
steps
#>    x step_1_equal step_2_geq step_3_greater first_greater_than_last
#> 1 -3         TRUE       TRUE           TRUE                    TRUE
#> 2 -1         TRUE       TRUE           TRUE                    TRUE
#> 3  0         TRUE       TRUE           TRUE                    TRUE
#> 4  2         TRUE       TRUE           TRUE                    TRUE
#> 5  3         TRUE       TRUE           TRUE                    TRUE
```

At all 5 values of x, the three steps hold in 5, 5 and 5 cases, and the first line exceeds the last in 5 cases, as Theorem 14 says. At $x = 0$, for example, the first line is 1 and the last is 0.

Definition 25 (Result)**.** A **result** is a proposition (Definition 1) that has a proof (Definition 23). These notes label each result by its role: as a theorem, a lemma, or a corollary, each defined next.

Example 30 (A proposition that is a result, and one that is not)**.** "$2 + 3 = 5$" is a proposition (Definition 1) that has a proof, the computation of $2 + 3$, so it is a result. "7 is even" is a proposition that is false, so it has no proof and is not a result.

Remark 2 ("Proposition" as a kind of result)*.* Some sources also call a result of moderate importance a "proposition". In these notes, "proposition" always means a sentence that is true or false (Definition 1), whether or not it has been proved.

Definition 26 (Theorem)**.** A **theorem** (Wikipedia contributors 2026ac) is a result (Definition 25) of interest in its own right.

Example 31 (A theorem)**.** "The sum of two even integers is even" is a result: Example 27 proves it. It is of interest in its own right, since it applies whenever two even integers are added, as in $2 + 4 = 6$, so it is a theorem.

Definition 27 (Lemma)**.** A **lemma** (Wikipedia contributors 2026k) is a result (Definition 25) proved mainly as a step in the proof of another result.

Example 32 (A lemma in the proof of a larger result)**.** To show that the sum of four even integers a, b, c, and d is even, write $a+b+c+d = (a+b)+(c+d)$. By the result of Example 27, $a + b$ and $c + d$ are even, and applying that result again to these two sums shows that their sum is even. Here that result is proved mainly as a step in the proof of another result, so it serves as a lemma. For example, $2 + 4 + 6 + 8 = (2 + 4) + (6 + 8) = 6 + 14 = 20$, which is even.

Definition 28 (Corollary)**.** A **corollary** (Wikipedia contributors 2026d) is a result (Definition 25) that follows from a theorem with little extra work.

Example 33 (A corollary of a result)**.** Example 27 proves the result "the sum of two even integers is even". A corollary follows with little extra work: the sum of three even integers a, b, and c is even, because $a+b+c=(a+b)+c$, $a+b$ is even by that result, and applying the result again to the even integers $a+b$ and c shows that $(a+b)+c$ is even. For example,

$$\begin{aligned} 2+4+6 &= (2+4)+6 \\ &= 6+6 \\ &= 12, \end{aligned}$$

which is even.

Definition 29 (Counterexample)**.** A **counterexample** (Wikipedia contributors 2026e) to a proposition of the form "for every x in A, $P(x)$", where $P(x)$ is a predicate (Definition 2), is an element x of A for which $P(x)$ is false. A single counterexample shows that the proposition is false.

Example 34 (A counterexample)**.** The proposition "for every integer n, $n+n>n$" is false. The integer $n=0$ is a counterexample: $0+0=0$, and $0>0$ is false. The proposition is true for some integers, such as $n=3$, since $3+3=6>3$, but one counterexample is enough to make it false.

Definition 30 (Disproof)**.** A **disproof** of a proposition P is a proof of its negation, $\neg P$.

Example 35 (A disproof)**.** Claim: every prime is odd.
Disproof: the integer 2 is prime, and 2 is not odd, since $2=2k+1$ would need $k=1/2$, which is not an integer. So there is a prime that is not odd, which is the negation of the claim.
This disproof gives a counterexample, 2.

Definition 31 (Constructive proof)**.** A **constructive proof** (Wikipedia contributors 2026a) of a proposition $\exists x \in S,\ P(x)$ with an existential quantifier names a particular element of S and shows that P is true of it.

Remark 3 (Several existential quantifiers)*.* A proposition can open with several existential quantifiers. A constructive proof of it names one element for each quantifier. For example, a constructive proof of "there are integers m and n with $m + n = 1$" names one value of m and one of n.

Example 36 (A constructive proof)**.** Claim: there are integers m and n with $m+n = 1$.
Proof: take $m = 0$ and $n = 1$. Then $m + n = 0 + 1 = 1$.
This proof is a constructive proof, since it names an integer for each of the two existential quantifiers.

Definition 32 (Non-constructive proof)**.** A **non-constructive proof** (Wikipedia contributors 2026a) of a proposition $\exists x \in S,\ P(x)$ shows that it is true without naming any element x of S for which $P(x)$ is true. It is a proof that is not a constructive proof.

Example 37 (A non-constructive proof)**.** Claim: among any three integers, there are two that are both even or both odd.
Proof: each integer is even or odd. If at most one of the three integers were even, and at most one were odd, there would be at most two integers in all. So at least two of the three are even, or at least two are odd.
This proof is a non-constructive proof, since it does not say which two of the integers share their parity.

Definition 33 (Converse)**.** The **converse** (Wikipedia contributors 2026c) of the proposition "if P, then Q" is the proposition "if Q, then P".

Example 38 (A true proposition with a false converse)**.** For an integer n, call n a multiple of 4 if $n = 4k$ for some integer k.

- "If n is a multiple of 4, then n is even" is true:

$$\begin{aligned} n &= 4k \\ &= 2 \cdot (2k), \end{aligned}$$

  and $2k$ is an integer.

- Its converse, "if n is even, then n is a multiple of 4", is false. The integer $n = 2$ is a counterexample (Definition 29): 2 is even, but $2 = 4k$ only for $k = \frac{1}{2}$, which is not an integer.

So a proposition can be true while its converse is false.

Definition 34 (Contrapositive)**.** The **contrapositive** (Wikipedia contributors 2026b) of the proposition "if P, then Q" is the proposition "if $\neg Q$, then $\neg P$".

Theorem 15 (A conditional is equivalent to its contrapositive)**.** *For any propositions P and Q,*

$$(P \Rightarrow Q) \equiv (\neg Q \Rightarrow \neg P).$$

Proof

Proof. By the definition of implies, $P \Rightarrow Q$ is false exactly when P is true and Q is false. By the same definition, $\neg Q \Rightarrow \neg P$ is false exactly when $\neg Q$ is true and $\neg P$ is false, that is, exactly when Q is false and P is true. So the two propositions are false in the same row of a truth table and true in the other three, which is Definition 11. □

Example 39 (The contrapositive versus the converse)**.** The proposition "if n is a multiple of 4, then n is even" from Example 38 has:

- converse (Definition 33): "if n is even, then n is a multiple of 4", which is false;
- contrapositive (Definition 34): "if n is not even, then n is not a multiple of 4", which is true, by Theorem 15, because the original proposition is true.

Definition 35 (Sufficient condition)**.** When the proposition "if P, then Q" is true, P is a **sufficient condition** (Wikipedia contributors 2026u) for Q.

Example 40 (A sufficient condition)**.** For a real number x, "if $x > 5$, then $x > 2$" is true, because a number greater than 5 is greater than 2. So $x > 5$ is a sufficient condition for $x > 2$: knowing that $x = 6$ satisfies $x > 5$ is enough to conclude that $6 > 2$.

Definition 36 (Necessary condition)**.** When the proposition "if P, then Q" is true, Q is a **necessary condition** (Wikipedia contributors 2026u) for P.

Example 41 (A necessary condition)**.** For an integer n, "if n is a multiple of 6, then n is a multiple of 3" is true, because $6k = 3 \cdot (2k)$ for every integer k. So being a multiple of 3 is a necessary condition for being a multiple of 6: $12 = 6 \cdot 2$ is a multiple of 6, and it is also a multiple of 3, since $12 = 3 \cdot 4$.

Definition 37 (Necessary and sufficient condition)**.** When both "if P, then Q" and its converse (Definition 33) are true, P is a **necessary and sufficient condition** (Wikipedia contributors 2026u) for Q (Definition 35, Definition 36), written "P if and only if Q".

Example 42 (Necessary but not sufficient)**.** In Example 38, "if n is a multiple of 4, then n is even" is true, so:

- being a multiple of 4 is a sufficient condition for n to be even;
- being even is a necessary condition for n to be a multiple of 4.

Being even is not a sufficient condition for being a multiple of 4: $n = 2$ is even, but not a multiple of 4.

Definition 38 (Proof by contradiction)**.** A **proof by contradiction** (Wikipedia contributors 2026x) of a proposition P assumes that P is false, and derives from that assumption a proposition that is known to be false.

Example 43 (No integer is both even and odd)**.** Suppose, for contradiction, that some integer n is both even and odd (definition). Then $n = 2j$ and $n = 2k+1$ for some integers j and k, and

$$\begin{aligned} 2j &= 2k+1 && \text{(both equal } n\text{)} \\ 2j-2k &= 1 && \text{(subtract } 2k \text{ from both sides)} \\ 2(j-k) &= 1 && \text{(distributive law)} \\ j-k &= \tfrac{1}{2} && \text{(divide both sides by 2)} \end{aligned}$$

The third line uses the distributive law. But $j-k$ is an integer, and $\frac{1}{2}$ is not, so the last line is false. So no integer is both even and odd. For example, $6 = 2 \cdot 3$ is even, and $6 = 2k+1$ would need $k = 2.5$, which is not an integer.

Theorem 16 (A proof by contradiction proves its proposition)**.** *Let P and Q be propositions (Definition 1). If $\neg P \Rightarrow Q$ is true and Q is false, then P is true.*

Proof

Proof. By Definition 8, $\neg P \Rightarrow Q$ is false only when $\neg P$ is true and Q is false. Here Q is false, and $\neg P \Rightarrow Q$ is true, so $\neg P$ is not true. Then $\neg P$ is false. Since $\neg P$ is true whenever P is false, P must be true. □

Theorem 17 (Why a proof by contradiction works)**.** *For propositions P and C,*

$$P \equiv (\neg P \Rightarrow (C \wedge \neg C)).$$

Proof

Proof. By the definitions of and and not, $C \wedge \neg C$ is false in every row, so it is a contradiction. So, by the definition of implies, $\neg P \Rightarrow (C \wedge \neg C)$ is true exactly in the rows where $\neg P$ is false, which are the rows where P is true:

P	C	$\neg P$	$C \wedge \neg C$	$\neg P \Rightarrow (C \wedge \neg C)$
T	T	F	F	T
T	F	F	F	T
F	T	T	F	F
F	F	T	F	F

The first and last columns match in every row, so the two are logically equivalent. □

Example 44 (A truth table for a proof by contradiction)**.** There are only four ways to assign truth values to P and Q. The code lists them, with the truth value of $\neg P \Rightarrow Q$, and marks the rows where $\neg P \Rightarrow Q$ is true and Q is false.

```
tt <- expand.grid(Q = c(TRUE, FALSE), P = c(TRUE, FALSE))
tt$not_P_implies_Q <- !(!tt$P) | tt$Q
tt$hypotheses_hold <- tt$not_P_implies_Q & !tt$Q
tt
#>       Q     P not_P_implies_Q hypotheses_hold
#> 1  TRUE  TRUE            TRUE           FALSE
#> 2 FALSE  TRUE            TRUE            TRUE
#> 3  TRUE FALSE            TRUE           FALSE
#> 4 FALSE FALSE           FALSE           FALSE
```

The hypotheses of Theorem 16 hold in 1 of the 4 rows, and in 1 of them P is true. So whenever the hypotheses hold, P is true.

Don't skip steps

Show every step of the derivation. Each line should follow from the previous line by a single, identifiable move:

- applying a definition,
- substituting a known result, or
- performing one algebraic operation.

When you skip steps, you force the reader to reconstruct your reasoning, which is exactly the work the proof was supposed to do *for* them. A step that feels "obvious" while you are writing it is often the step a reader gets stuck on. Phrases like "it can be shown that" or "clearly" usually mark a skipped step; show the step instead.

Annotate each step

For each step, note the definition, theorem, or algebraic rule that justifies it. A brief annotation to the right of the line is usually enough. Annotations make the proof checkable, and they teach the reader which tool to reach for next time.

For example, here is a derivation that the hat matrix $\mathbf{H} = \mathbf{X}(\mathbf{X}^\top\mathbf{X})^{-1}\mathbf{X}^\top$ is idempotent, $\mathbf{H}^2 = \mathbf{H}$, with each step annotated:

$$
\begin{aligned}
\mathbf{H}^2 &= \mathbf{H}\mathbf{H} && \text{(definition of a matrix power)} \\
&= \mathbf{X}(\mathbf{X}^\top\mathbf{X})^{-1}\mathbf{X}^\top\,\mathbf{X}(\mathbf{X}^\top\mathbf{X})^{-1}\mathbf{X}^\top && \text{(definition of } \mathbf{H}\text{)} \\
&= \mathbf{X}(\mathbf{X}^\top\mathbf{X})^{-1}(\mathbf{X}^\top\mathbf{X})(\mathbf{X}^\top\mathbf{X})^{-1}\mathbf{X}^\top && \text{(regroup; matrix multiplication is associative)} \\
&= \mathbf{X}\,\mathbf{I}_p\,(\mathbf{X}^\top\mathbf{X})^{-1}\mathbf{X}^\top && \text{(a matrix inverse times the matrix is } \mathbf{I}_p\text{)} \\
&= \mathbf{X}(\mathbf{X}^\top\mathbf{X})^{-1}\mathbf{X}^\top && \text{(the identity matrix changes nothing)} \\
&= \mathbf{H} && \text{(definition of } \mathbf{H}\text{)}
\end{aligned}
$$

The Linear Algebra page uses this derivation to show that $\mathbf{H}$ is a projection matrix. The probability notes derive the variance identity $\mathrm{Var}(X) = \mathbb{E}[X^2] - (\mathbb{E}[X])^2$ in the same annotated style.

Direct proof

Definition 39 (Direct proof)**.** A **direct proof** (Wikipedia contributors 2026g) of a proposition Q starts from the assumptions that Q is stated under, if any, and reaches Q through a chain of steps, each of which applies a definition, a previously proved result, or one rule of logic or algebra to the assumptions or to earlier steps. A direct proof never assumes that Q is false; compare proof by contradiction, which does.

Example 45 (The square of an odd integer is odd)**.** Claim: if n is an odd integer, then n^2 is odd.
Since n is odd, there is an integer k with $n = 2k + 1$. Then

$$\begin{aligned} n^2 &= (2k+1)^2 && \text{(substitute } n = 2k+1\text{)} \\ &= (2k)^2 + 2 \cdot 2k \cdot 1 + 1^2 && \text{(square of a sum)} \\ &= 4k^2 + 4k + 1 && \text{(arithmetic)} \\ &= 2(2k^2 + 2k) + 1 && \text{(distributive law, read from right to left)} \end{aligned}$$

and $2k^2 + 2k$ is an integer, because adding or multiplying integers gives an integer. So $n^2 = 2 \cdot (\text{an integer}) + 1$, which is the definition of odd. For example,

$$\begin{aligned} n &= 7 \\ &= 2 \cdot 3 + 1 \end{aligned}$$

gives

$$\begin{aligned} n^2 &= 49 \\ &= 2 \cdot 24 + 1, \end{aligned}$$

and $2 \cdot 3^2 + 2 \cdot 3 = 24$.
The proof above that the sum of two even integers is even is also a direct proof.

Proof by contrapositive

Definition 40 (Proof by contrapositive)**.** A **proof by contrapositive** (Wikipedia contributors 2026b) of a proposition "if P, then Q" is a direct proof of its contrapositive, "if $\neg Q$, then $\neg P$": it assumes that Q is false and reaches the conclusion that P is false. By Theorem 15, proving the contrapositive proves the original proposition.

Example 46 (If $x^2+3x<0$, then $x<0$)**.** Claim: for every real number x, if $x^2+3x<0$, then $x<0$.
A direct proof would have to start from the inequality $x^2+3x<0$, which says little about x on its own. The contrapositive (Definition 34) is "if $x \geq 0$, then $x^2+3x \geq 0$", and its assumption is easier to use.
Suppose $x \geq 0$. Then $x^2 = x \cdot x \geq 0$, because a product of two nonnegative numbers is nonnegative, and $3x \geq 0$, because $3>0$ and $x \geq 0$. So

$$\begin{aligned} x^2+3x &\geq 0+0 && \text{(add the inequalities } x^2 \geq 0 \text{ and } 3x \geq 0\text{)} \\ &= 0. && \text{(arithmetic)} \end{aligned}$$

That proves the contrapositive, so by Theorem 15 the claim holds. For example, $x=-1$ gives $x^2+3x = 1-3 = -2 < 0$, and indeed $x<0$.

Proof by cases

Definition 41 (Proof by cases, proof by exhaustion)**.** A **proof by cases** (Wikipedia contributors 2026y), also called a **proof by exhaustion**, of a proposition Q lists propositions $C_1, \ldots, C_k$, the *cases*, shows that at least one of them is true, and proves $C_i \Rightarrow Q$ for each i.

Example 47 (The square of a real number is nonnegative)**.** Claim: for every real number x, $x^2 \geq 0$.
Every real number x satisfies $x \geq 0$ or $x<0$, so these two cases cover every x.
Case 1: $x \geq 0$**.** Then $x^2 = x \cdot x$ is a product of two nonnegative numbers, so $x^2 \geq 0$.
Case 2: $x<0$**.** Then $-x>0$, and

$$\begin{aligned} x^2 &= x \cdot x && \text{(definition of } x^2\text{)} \\ &= (-x)\cdot(-x) && \text{(the two minus signs cancel)} \end{aligned}$$

is a product of two positive numbers, so $x^2>0$, and in particular $x^2 \geq 0$.
The claim holds in both cases, so it holds for every real number x.

Follow the golden rule

In general, follow the golden rule: treat your readers the way you want to be treated as a reader.

When you read someone else's proof, you want to be able to follow every step without guessing, to know which result is being used at each line, and to never be left wondering where a quantity came from. Write your own proofs to meet that same standard.

Proof by induction

Many proofs in these notes show that a statement holds for every natural number n, or for every integer n from some starting value on. Checking the statement one n at a time would never finish; induction proves it for all of them with two arguments.

Definition 42 (Proof by induction)**.** Let n_0 be an integer, and let $P(n)$ be a predicate about the integers $n \geq n_0$. A **proof by induction** (Wikipedia contributors 2026s) on n shows that $P(n)$ holds for every integer $n \geq n_0$ by showing two things:

1. **Base case:** $P(n_0)$ holds.
2. **Inductive step:** for every integer $n > n_0$, if $P(n-1)$ holds, then $P(n)$ holds.

In the inductive step, the assumption that $P(n-1)$ holds is called the **induction hypothesis**.

Remark 4 (Why the two steps suffice). The base case gives $P(n_0)$. The inductive step with $n = n_0 + 1$ then gives $P(n_0 + 1)$, with $n = n_0 + 2$ it gives $P(n_0 + 2)$, and so on: each integer $n \geq n_0$ is reached after $n - n_0$ uses of the inductive step.
When $P(n)$ is needed only for $n_0 \leq n \leq r$, it is enough to show the inductive step for $n_0 < n \leq r$; the same chain then stops at $P(r)$.

The following example uses summation notation.

Example 48 (Sum of the first n natural numbers)**.** For every $n \in \mathbb{N}$,

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

For example, with $n = 3$, $1 + 2 + 3 = 6$ and $\frac{3 \cdot 4}{2} = 6$. To prove the equation for every n, let $P(n)$ be the equation and use Definition 42 with $n_0 = 1$.
Base case, $n = 1$**.** The sum has one term:

$$\sum_{i=1}^{1} i = 1 \quad \text{(definition of a sum)}$$

and the right side is

$$\begin{aligned}
\frac{1\,(1+1)}{2} &= \frac{1\cdot 2}{2} && (1+1=2)\\
&= \frac{2}{2} && (1\cdot 2=2)\\
&= 1. && (\tfrac{2}{2}=1)
\end{aligned}$$

So $P(1)$ holds.

Inductive step. Let $n \geq 2$, and suppose $P(n-1)$ holds:

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

Then

$$\begin{aligned}
\sum_{i=1}^{n} i &= 1+2+\cdots+(n-1)+n && \text{(definition of a sum)}\\
&= (1+2+\cdots+(n-1))+n && \text{(associative law)}\\
&= \sum_{i=1}^{n-1} i + n && \text{(definition of a sum)}\\
&= \frac{(n-1)\,((n-1)+1)}{2} + n && \text{(induction hypothesis)}\\
&= \frac{(n-1)\,n}{2} + n && ((n-1)+1=n)\\
&= \frac{(n-1)\,n}{2} + \frac{2n}{2} && (n=\tfrac{2n}{2})\\
&= \frac{(n-1)\,n+2n}{2} && \text{(common denominator)}\\
&= \frac{((n-1)+2)\,n}{2} && \text{(distributive law)}\\
&= \frac{(n+1)\,n}{2} && ((n-1)+2=n+1)\\
&= \frac{n(n+1)}{2}, && \text{(commutative law)}
\end{aligned}$$

which is $P(n)$. By Definition 42, the equation holds for every $n \in \mathbb{N}$.

Formal logic

Definition 43 (Propositional logic)**. Propositional logic** (Wikipedia contributors 2026aa) is the study of logical entailment among propositions built from basic propositions $P, Q, R, ...$ using only the logical connectives. Whether such propositions entail one another depends only on the truth values of the basic propositions, not on what the basic propositions say.

Example 49 (P and $P \Rightarrow Q$ entail Q)**.** Suppose P is true and $P \Rightarrow Q$ is true. If Q were false, then $P \Rightarrow Q$ would have a true P and a false Q, so $P \Rightarrow Q$ would be false, which contradicts the assumption. So Q is true. This holds whatever P and Q say, so P and $P \Rightarrow Q$ logically entail Q.

Definition 44 (First-order logic)**. First-order logic** (Wikipedia contributors 2026i) extends propositional logic (Definition 43) by also allowing:

- variables $x, y, ...$ that stand for elements of a set;
- predicates of one or more of those variables, such as $P(x)$ or $x < y$;
- functions of them, such as $x + y$;
- equality between them;
- quantifiers $\forall$ and $\exists$ over them.

Example 50 (From "for all" to one element)**.** Let $P(x)$ be a predicate about the elements of a set S, and let $a \in S$. The propositions $\forall x \in S, P(x)$ and $a \in S$ logically entail $P(a)$: if $P(x)$ is true for every element x of S, then it is true for the element a. For example, $\forall x \in \mathbb{Z}, x^2 \geq 0$ entails $3^2 \geq 0$. Propositional logic alone cannot express this entailment, because it has no way to refer to the elements of S.

Definition 45 (Formal proof)**.** A **formal proof** (Wikipedia contributors 2026j) is a proof written in propositional logic (Definition 43) or first-order logic (Definition 44) in which each step is an assumption or follows from earlier steps by one rule taken from a fixed list of allowed rules, so that checking the proof needs no understanding of what its propositions say.

Example 51 (A three-step formal proof)**.** Suppose the allowed rules include "from P and $P \Rightarrow Q$, conclude Q". Then this is a formal proof of Q from the assumptions P and $P \Rightarrow Q$:

1. P (assumption)
2. $P \Rightarrow Q$ (assumption)
3. Q (from steps 1 and 2, by the rule above)

Further reading

- Velleman (2019) is a structured introduction to proof techniques such as direct proof, proof by contradiction, and induction.
- Barker-Plummer et al. (2011) treats proofs as formal proofs in propositional logic and first-order logic.

Free course materials on formal logic:

- Thomas Icard's Phil 150/250, Mathematical Logic course notes (Stanford) review sets, relations and functions, then cover propositional, modal and first-order logic.
- John Wilcox's logic handouts supplement the same Stanford course, with a handout per week on the syntax and semantics of propositional and first-order logic, natural deduction and metalogic.
- Eric Pacuit's Phil 151: First Order Logic (Stanford, Winter 2009) posts a syllabus, a modal logic study guide and problem sets.

References

Barker-Plummer, Dave, Jon Barwise, and John Etchemendy. 2011. *Language, Proof and Logic*. 2nd ed. CSLI Publications. https://www.amazon.com/dp/1575866323.

Velleman, Daniel J. 2019. *How to Prove It: A Structured Approach*. 3rd ed. Cambridge University Press. https://doi.org/10.1017/9781108539890.

Wikipedia contributors. 2026a. *Constructive Proof — Wikipedia, the Free Encyclopedia*. https://en.wikipedia.org/wiki/Constructive_proof.

Wikipedia contributors. 2026b. *Contraposition — Wikipedia, the Free Encyclopedia*. https://en.wikipedia.org/wiki/Contraposition.

Wikipedia contributors. 2026c. *Converse (Logic) — Wikipedia, the Free Encyclopedia*. https://en.wikipedia.org/wiki/Converse_(logic).

Wikipedia contributors. 2026d. *Corollary — Wikipedia, the Free Encyclopedia*. https://en.wikipedia.org/wiki/Corollary.

Wikipedia contributors. 2026e. *Counterexample — Wikipedia, the Free Encyclopedia*. https://en.wikipedia.org/wiki/Counterexample.

Wikipedia contributors. 2026f. *De Morgan's Laws — Wikipedia, the Free Encyclopedia*. https://en.wikipedia.org/wiki/De_Morgan%27s_laws.

Wikipedia contributors. 2026g. *Direct Proof — Wikipedia, the Free Encyclopedia*. https://en.wikipedia.org/wiki/Direct_proof.

Wikipedia contributors. 2026h. *Existential Quantification — Wikipedia, the Free Encyclopedia*. https://en.wikipedia.org/wiki/Existential_quantification.

Wikipedia contributors. 2026i. *First-Order Logic — Wikipedia, the Free Encyclopedia*. https://en.wikipedia.org/wiki/First-order_logic.

Wikipedia contributors. 2026j. *Formal Proof — Wikipedia, the Free Encyclopedia*. https://en.wikipedia.org/wiki/Formal_proof.

Wikipedia contributors. 2026k. *Lemma (Mathematics) — Wikipedia, the Free Encyclopedia*. https://en.wikipedia.org/wiki/Lemma_(mathematics).

Wikipedia contributors. 2026l. *Logical Biconditional — Wikipedia, the Free Encyclopedia*. https://en.wikipedia.org/wiki/Logical_biconditional.

Wikipedia contributors. 2026m. *Logical Conjunction — Wikipedia, the Free Encyclopedia*. https://en.wikipedia.org/wiki/Logical_conjunction.

Wikipedia contributors. 2026n. *Logical Connective — Wikipedia, the Free Encyclopedia*. https://en.wikipedia.org/wiki/Logical_connective.

Wikipedia contributors. 2026o. *Logical Consequence — Wikipedia, the Free Encyclopedia*. https://en.wikipedia.org/wiki/Logical_consequence.

Wikipedia contributors. 2026p. *Logical Disjunction — Wikipedia, the Free Encyclopedia*. https://en.wikipedia.org/wiki/Logical_disjunction.

Wikipedia contributors. 2026q. *Logical Equivalence — Wikipedia, the Free Encyclopedia*.

https://en.wikipedia.org/wiki/Logical_equivalence.

Wikipedia contributors. 2026r. *Material Conditional — Wikipedia, the Free Encyclopedia*. https://en.wikipedia.org/wiki/Material_conditional.

Wikipedia contributors. 2026s. *Mathematical Induction — Wikipedia, the Free Encyclopedia*. https://en.wikipedia.org/wiki/Mathematical_induction.

Wikipedia contributors. 2026t. *Mathematical Proof — Wikipedia, the Free Encyclopedia*. https://en.wikipedia.org/wiki/Mathematical_proof.

Wikipedia contributors. 2026u. *Necessity and Sufficiency — Wikipedia, the Free Encyclopedia*. https://en.wikipedia.org/wiki/Necessity_and_sufficiency.

Wikipedia contributors. 2026v. *Negation — Wikipedia, the Free Encyclopedia*. https://en.wikipedia.org/wiki/Negation.

Wikipedia contributors. 2026w. *Predicate (Mathematical Logic) — Wikipedia, the Free Encyclopedia*. https://en.wikipedia.org/wiki/Predicate_(mathematical_logic).

Wikipedia contributors. 2026x. *Proof by Contradiction — Wikipedia, the Free Encyclopedia*. https://en.wikipedia.org/wiki/Proof_by_contradiction.

Wikipedia contributors. 2026y. *Proof by Exhaustion — Wikipedia, the Free Encyclopedia*. https://en.wikipedia.org/wiki/Proof_by_exhaustion.

Wikipedia contributors. 2026z. *Proposition — Wikipedia, the Free Encyclopedia*. https://en.wikipedia.org/wiki/Proposition.

Wikipedia contributors. 2026aa. *Propositional Calculus — Wikipedia, the Free Encyclopedia*. https://en.wikipedia.org/wiki/Propositional_calculus.

Wikipedia contributors. 2026ab. *Quantifier (Logic) — Wikipedia, the Free Encyclopedia*. https://en.wikipedia.org/wiki/Quantifier_(logic).

Wikipedia contributors. 2026ac. *Theorem — Wikipedia, the Free Encyclopedia*. https://en.wikipedia.org/wiki/Theorem.

Wikipedia contributors. 2026ad. *Truth Table — Wikipedia, the Free Encyclopedia*. https://en.wikipedia.org/wiki/Truth_table.

Wikipedia contributors. 2026ae. *Universal Quantification — Wikipedia, the Free Encyclopedia*. https://en.wikipedia.org/wiki/Universal_quantification.