Appendix

Mathematics is made of arguments (reasoned discourse that is, not crockery-throwing). This section sketches the background material and argument techniques that we use in the book.

This section informally outlines the topics, skipping proofs. For more, [Velleman2] is excellent. Two other sources, available online, are [Hefferon] and [Beck].

Statements

Formal mathematical statements come labelled as a *Theorem* for major points, a *Corollary* for results that follow immediately from a prior one, or a *Lemma* for results chiefly used to prove others.

Statements can be complex and have many parts. The truth or falsity of the entire statement depends both on the truth value of the parts and on how the statement is put together.

Not Where P is a proposition, 'it is not the case that P' is true provided that P is false. For instance, 'n is not prime' is true only when n is the product of smaller integers.

To prove that a 'not P' statement holds, show that P is false.

And For a statement of the form 'P and Q' to be true both halves must hold: '7 is prime and so is 3' is true, while '7 is prime and 3 is not' is false.

To prove a 'P and Q', prove each half.

Or A 'P or Q' statement is true when either half holds: '7 is prime or 4 is prime' is true, while '8 is prime or 4 is prime' is false. In the case that both clauses of the statement are true, as in '7 is prime or 3 is prime', we take the statement as a whole to be true. (In everyday speech people occasionally use 'or' in an exclusive way — "Live free or die" does not intend both halves to hold — but we will not use 'or' in that way.)

To prove 'P or Q', show that in all cases at least one half holds (perhaps sometimes one half and sometimes the other, but always at least one).

If-then An 'if P then Q' statement may also appear as 'P implies Q' or 'P $\implies$ Q' or 'P is sufficient to give Q' or 'Q if P'. It is true unless P is true while Q is false. Thus 'if 7 is prime then 4 is not' is true while 'if 7 is prime then 4 is also prime' is false. (Contrary to its use in casual speech, in mathematics 'if P then Q' does not connote that P precedes Q or causes Q.)

Note this consequence of the prior paragraph: if P is false then 'if P then Q' is true irrespective of the value of Q: 'if 4 is prime then 7 is prime' and 'if 4 is prime then 7 is not' are both true statements. (They are *vacuously true*.) Also observe that 'if P then Q' is true when Q is true: 'if 4 is prime then 7 is prime' and 'if 4 is not prime then 7 is prime' are both true.

There are two main ways to establish an implication. The first way is direct: assume that P is true and use that assumption to prove Q. For instance, to show 'if a number is divisible by 5 then twice that number is divisible by 10' we can assume that the number is $5n$ and deduce that $2(5n) = 10n$. The indirect way is to prove the *contrapositive* statement: 'if Q is false then P is false' (rephrased, 'Q can only be false when P is also false'). Thus to show 'if a natural number is prime then it is not a perfect square' we can argue that if it were a square $p = n^2$ then it could be factored $p = n \cdot n$ where $n < p$ and so wouldn't be prime ($p = 0$ or $p = 1$ don't satisfy $n < p$ but they are nonprime).

Equivalent statements Sometimes, not only does P imply Q but also Q implies P. Some ways to say this are: 'P if and only if Q', 'P iff Q', 'P and Q are logically equivalent', 'P is necessary and sufficient to give Q', 'P $\iff$ Q'. An example is 'an integer is divisible by ten if and only if that number ends in 0'.

Although in simple arguments a chain like "P if and only if R, which holds if and only if S ..." may be practical, to prove that statements are equivalent we more often prove the two halves 'if P then Q' and 'if Q then P' separately.

Quantifiers

Compare these statements about natural numbers: 'there is a natural number x such that x is divisible by x^2' is true, while 'for all natural numbers x, that x is divisible by x^2' is false. The prefixes 'there is' and 'for all' are *quantifiers*.

For all The 'for all' prefix is the *universal quantifier*, symbolized $\forall$.

The most straightforward way to prove that a statement holds in all cases is to prove that it holds in each case. Thus to show that 'every number divisible by p has its square divisible by p^2', take a single number of the form pn and square it $(pn)^2 = p^2n^2$. This is a *typical element* proof. (In this kind of argument be careful not to assume properties for that element other than the ones in the

hypothesis. This argument is wrong: "If n is divisible by a prime, say 2, so that $n = 2k$ for some natural number k, then $n^2 = (2k)^2 = 4k^2$ and the square of n is divisible by the square of the prime." That is a proof for the special case $p = 2$ but it isn't a proof for all p. Contrast it with a correct one: "If n is divisible by a prime so that $n = pk$ for some natural number k then $n^2 = (pk)^2 = p^2k^2$ and so the square of n is divisible by the square of the prime.")

There exists The 'there exists' prefix is the *existential quantifier*, symbolized $\exists$.

We can prove an existence proposition by producing something satisfying the property: for instance, to settle the question of primality of $2^{2^5} + 1$, Euler exhibited the divisor 641[Sandifer]. But there are proofs showing that something exists without saying how to find it; Euclid's argument given in the next subsection shows there are infinitely many primes without giving a formula naming them.

Finally, after answering "Are there any?" affirmatively we often ask "How many?" That is, the question of uniqueness often arises in conjunction with the question of existence. Sometimes the two arguments are simpler if separated so note that just as proving something exists does not show it is unique, neither does proving something is unique show that it exists. (For instance, we can easily show that the natural number halfway between three and four is unique, even thouge no such number exists.)

Techniques of Proof

Induction Many proofs are iterative, "Here's why the statement is true for the number 0, it then follows for 1 and from there to 2 ...". These are proofs by *mathematical induction*. This technique is often not obvious to a person who has not seen it before, even to a person with a mathematical turn of mind. So we will see two examples.

We will first prove that $1 + 2 + 3 + \cdots + n = n(n+1)/2$. That formula has a natural number variable n that is free, meaning that setting n to be 1, or 2, etc., gives a family of cases of the statement: first that $1 = 1(2)/2$, second that $1 + 2 = 2(3)/2$, etc. Our induction proofs involve statements with one free natural number variable.

Each proof has two steps. In the *base step* we show that the statement holds for some intial number $i \in \mathbb{N}$. Often this step is a routine, and short, verification. The second step, the *inductive step*, is more subtle; we will show that this implication holds:

> If the statement holds from $n = i$ up to and including $n = k$ then the statement holds also in the $n = k + 1$ case (*)

(the first line is the *inductive hypothesis*). The Principle of Mathematical Induction is that completing both steps proves that the statement is true for all natural numbers greater than or equal to i.

For the sum of the initial n numbers statement the intuition behind the principle is that first, the base step directly verifies the statement for the case of the initial number $n = 1$. Then, because the inductive step verifies the implication $(*)$ for all k, that implication applied to $k = 1$ gives that the statement is true for the case of the number $n = 2$. Now, with the statement established for both 1 and 2, apply $(*)$ again to conclude that the statement is true for the number $n = 3$. In this way, we bootstrap to all numbers $n \geqslant 1$.

Here is a proof of $1 + 2 + 3 + \cdots + n = n(n+1)/2$, with separate paragraphs for the base step and the inductive step.

> For the base step we show that the formula holds when $n = 1$. That's easy; the sum of the first 1 natural number equals $1(1+1)/2$.
>
> For the inductive step, assume the inductive hypothesis that the formula holds for the numbers $n = 1$, $n = 2$, ..., $n = k$ with $k \geqslant 1$. That is, assume $1 = 1(1)/2$, and $1 + 2 = 2(3)/2$, and $1 + 2 + 3 = 3(4)/2$, through $1 + 2 + \cdots + k = k(k+1)/2$. With that, the formula holds also in the $n = k + 1$ case:
>
> $$1 + 2 + \cdots + k + (k+1) = \frac{k(k+1)}{2} + (k+1) = \frac{(k+1)(k+2)}{2}$$
>
> (the first equality follows from the inductive hypothesis).

Here is another example, proving that every integer greater than or equal to 2 is a product of primes.

> The base step is easy: 2 is the product of a single prime.
>
> For the inductive step assume that each of $2, 3, \ldots, k$ is a product of primes, aiming to show $k+1$ is also a product of primes. There are two possibilities. First, if $k + 1$ is not divisible by a number smaller than itself then it is a prime and so is the product of primes. The second possibility is that $k + 1$ is divisible by a number smaller than itself, and then by the inductive hypothesis its factors can be written as a product of primes. In either case $k + 1$ can be rewritten as a product of primes.

Contradiction Another technique of proof is to show that something is true by showing that it cannot be false. A proof by contradiction assumes that the proposition is false and derives some contradiction to known facts.

The classic example of proof by contradiction is Euclid's argument that there are infinitely many primes.

Suppose that there are only finitely many primes $p_1, \ldots, p_k$. Consider the number $p_1 \cdot p_2 \ldots p_k + 1$. None of the primes on the supposedly exhaustive list divides this number evenly since each leaves a remainder of 1. But every number is a product of primes so this can't be. Therefore there cannot be only finitely many primes.

Another example is this proof that $\sqrt{2}$ is not a rational number.

Suppose that $\sqrt{2} = m/n$, so that $2n^2 = m^2$. Factor out any 2's, giving $n = 2^{k_n} \cdot \hat{n}$ and $m = 2^{k_m} \cdot \hat{m}$. Rewrite.

$$2 \cdot (2^{k_n} \cdot \hat{n})^2 = (2^{k_m} \cdot \hat{m})^2$$

The Prime Factorization Theorem says that there must be the same number of factors of 2 on both sides, but there are an odd number of them $1 + 2k_n$ on the left and an even number $2k_m$ on the right. That's a contradiction, so a rational number with a square of 2 is impossible.

Sets, Functions, and Relations

Sets Mathematicians often work with collections. The most commonly-used kind of collection is a *set*. Sets are characterized by the Principle of Extensionality: two sets with the same elements are equal. Because of this, the order of the elements does not matter $\{2, \pi\} = \{\pi, 2\}$, and repeats collapse $\{7, 7\} = \{7\}$.

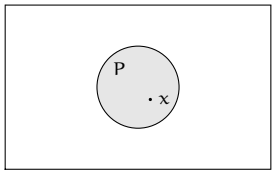
We can describe a set using a listing between curly braces $\{1, 4, 9, 16\}$ (as in the prior paragraph), or by using set-builder notation $\{x \mid x^5 - 3x^3 + 2 = 0\}$ (read "the set of all x such that ..."). We name sets with capital roman letters; for instance the set of primes is $P = \{2, 3, 5, 7, 11, \ldots\}$ (except that a few sets are so important that their names are reserved, such as the real numbers $\mathbb{R}$ and the complex numbers $\mathbb{C}$). To denote that something is an *element*, or *member*,) of a set we use '$\in$', so that $7 \in \{3, 5, 7\}$ while $8 \notin \{3, 5, 7\}$.

We say that A is a *subset* of B, written $A \subseteq B$, when $x \in A$ implies that $x \in B$. In this book we use '$\subset$' for the *proper subset* relationship that A is a subset of B but $A \neq B$ (some authors use this symbol for any kind of subset, proper or not). An example is $\{2, \pi\} \subset \{2, \pi, 7\}$. These symbols may be flipped, for instance $\{2, \pi, 5\} \supset \{2, 5\}$.

Because of Extensionality, to prove that two sets are equal $A = B$ show that they have the same members. Often we do this by showing mutual inclusion, that both $A \subseteq B$ and $A \supseteq B$. Such a proof will have a part showing that if $x \in A$ then $x \in B$, and a second part showing that if $x \in B$ then $x \in A$.

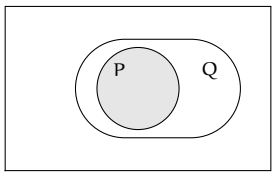
When a set has no members then it is the *empty set* $\{\}$, symbolized $\varnothing$. Any set has the empty set for a subset by the 'vacuously true' property of the definition of implication.

Diagrams We picture basic set operations with a *Venn diagram*. This shows $x \in P$.

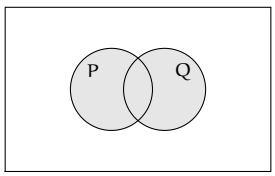


The outer rectangle contains the universe Ω of all objects under discussion. For instance, in a statement about real numbers, the rectangle encloses all members of $\mathbb{R}$. The set is pictured as a circle, enclosing its members.

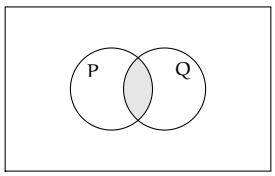
Here is the diagram for $P \subseteq Q$. It shows that if $x \in P$ then $x \in Q$.



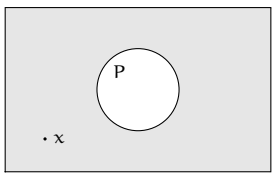
Set Operations The *union* of two sets is $P \cup Q = \{x \mid (x \in P) \text{ or } (x \in Q)\}$. The diagram shows that an element is in the union if it is in either of the sets.



The *intersection* is $P \cap Q = \{x \mid (x \in P) \text{ and } (x \in Q)\}$.



The *complement* of a set P is $P^{\text{comp}} = \{x \in \Omega \mid x \notin P\}$



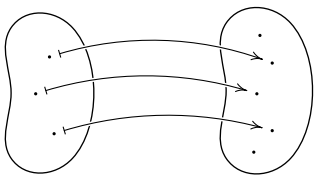
Multisets A *multiset* is a collection that is like a set in that order does not matter, but in which, unlike a set, repeats do not collapse. Thus the multiset $\{2,1,2\}$ is the same as the multiset $\{1,2,2\}$ but differs from the multiset $\{1,2\}$. Note that we use the same $\{\ldots\}$ curly brackets notation as for sets. Also as with sets, we say A is a *multiset subset* if A is a subset of B and A is a multiset.

Sequences In addition to sets and multisets, we also use collections where order matters and where repeats do not collapse. These are *sequences*, denoted with angle brackets: $\langle 2,3,7\rangle \neq \langle 2,7,3\rangle$. A sequence of length 2 is an *ordered pair*, and is often written with parentheses: $(\pi,3)$. We also sometimes say 'ordered triple', 'ordered 4-tuple', etc. The set of ordered n-tuples of elements of a set A is denoted A^n. Thus $\mathbb{R}^2$ is the set of pairs of reals.

Functions A *function* or *map* $f\colon D \to C$ is is an association between input *arguments* $x \in D$ and output *values* $f(x) \in C$ subject to the the requirement that the function must be *well-defined*, that x suffices to determine $f(x)$. Restated, the condition is that if $x_1 = x_2$ then $f(x_1) = f(x_2)$.

The set of all arguments D is f's *domain* and the set of output values is its *range* $\mathscr{R}(f)$. Often we don't work with the range and instead work with a convenient superset, the *codomain* C. For instance, we might describe the squaring function with $s\colon \mathbb{R} \to \mathbb{R}$ instead of $s\colon \mathbb{R} \to \mathbb{R}^+ \cup \{0\}$.

We picture functions with a *bean diagram*.



The blob on the left is the domain while on the right is the codomain. The function associates the three points of the domain with three in the codomain. Note that by the definition of a function every point in the domain is associated with a unique point in the codomain, but the converse needn't be true.

The association is arbitrary; no formula or algorithm is required, although in this book there typically is one. We often use y to denote $f(x)$. We also use the notation $x \overset{f}{\longmapsto} 16x^2 - 100$, read '$x$ maps under f to $16x^2 - 100$' or '$16x^2 - 100$ is the *image* of x'.

A map such as $x \mapsto \sin(1/x)$ is a combinations of simpler maps, here $g(y) = \sin(y)$ applied to the image of $f(x) = 1/x$. The *composition* of $g\colon Y \to Z$ with $f\colon X \to Y$, is the map sending $x \in X$ to $g(\,f(x)\,) \in Z$. It is denoted $g \circ f\colon X \to Z$. This definition only makes sense if the range of f is a subset of the domain of g.

An *identity map* $\mathrm{id}\colon Y \to Y$ defined by $\mathrm{id}(y) = y$ has the property that for any $f\colon X \to Y$, the composition $\mathrm{id} \circ f$ is equal to f. So an identity map plays the

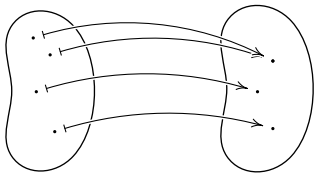
same role with respect to function composition that the number 0 plays in real number addition or that 1 plays in multiplication.

In line with that analogy, we define a *left inverse* of a map $f\colon X \to Y$ to be a function $g\colon \text{range}(f) \to X$ such that $g \circ f$ is the identity map on X. A *right inverse* of f is a $h\colon Y \to X$ such that $f \circ h$ is the identity.

For some f's there is a map that is both a left and right inverse of f. If such a map exists then it is unique because if both g_1 and g_2 have this property then $g_1(x) = g_1 \circ (f \circ g_2)(x) = (g_1 \circ f) \circ g_2(x) = g_2(x)$ (the middle equality comes from the associativity of function composition) so we call it a *two-sided inverse* or just *"the" inverse*, and denote it f^{-1}. For instance, the inverse of the function $f\colon \mathbb{R} \to \mathbb{R}$ given by $f(x) = 2x - 3$ is the function $f^{-1}\colon \mathbb{R} \to \mathbb{R}$ given by $f^{-1}(x) = (x+3)/2$.

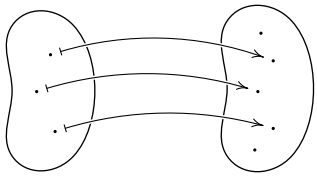
The superscript notation for function inverse 'f^{-1}' fits into a larger scheme. Functions with the same codomain as domain $f\colon X \to X$ can be iterated, so that we can consider the composition of f with itself: $f \circ f$, and $f \circ f \circ f$, etc. We write $f \circ f$ as f^2 and $f \circ f \circ f$ as f^3, etc. Note that the familiar exponent rules for real numbers hold: $f^i \circ f^j = f^{i+j}$ and $(f^i)^j = f^{i \cdot j}$. Then where f is invertible, writing f^{-1} for the inverse and f^{-2} for the inverse of f^2, etc., gives that these familiar exponent rules continue to hold, since we define f^0 to be the identity map.

The definition of function requires that for every input there is one and only one associated output value. If a function $f\colon D \to C$ has the additional property that for every output value there is at least one associated input value — that is, the additional property that f's codomain equals its range $C = \mathscr{R}(f)$ — then the function is *onto*.



A function has a right inverse if and only if it is onto. (The f pictured above has a right inverse $g\colon C \to D$ given by following the arrows backwards, from right to left. For the codomain point on the top, choose either one of the arrows to follow. With that, applying g first followed by f takes elements $y \in C$ to themselves, and so is the identity function.)

If a function $f\colon D \to C$ has the property that for every output value there is at most one associated input value — that is, if no two arguments share an image so that $f(x_1) = f(x_2)$ implies that $x_1 = x_2$ — then the function is *one-to-one*. The bean diagram from earlier illustrates.



A function has a left inverse if and only if it is one-to-one. (In the picture define $g\colon C \to D$ to follow the arrows backwards for those $y \in C$ that are at the end of an arrow, and to send the point to an arbitrary element in D otherwise. Then applying f followed by g to elements of D will act as the identity.)

By the prior paragraphs, a map has a two-sided inverse if and only if that map is both onto and one-to-one. Such a function is a *correspondence*. It associates one and only one element of the domain with each element of the codomain. Because a composition of one-to-one maps is one-to-one, and a composition of onto maps is onto, a composition of correspondences is a correspondence.

We sometimes want to shrink the domain of a function. For instance, we may take the function $f\colon \mathbb{R} \to \mathbb{R}$ given by $f(x) = x^2$ and, in order to have an inverse, limit input arguments to nonnegative reals $\hat{f}\colon \mathbb{R}^+ \cup \{0\} \to \mathbb{R}$. Then $\hat{f}$ is the *restriction* of f to the smaller domain.

Relations Some familiar mathematical things, such as '$<$' or '$=$', are most naturally understood as relations between things. A *binary relation* on a set A is a set of ordered pairs of elements of A. For example, some elements of the set that is the relation '$<$' are $(3,5)$, $(3,7)$, and $(1,100)$. Another binary relation on the natural numbers is equality; this relation is the set $\{\ldots,(-1,-1),(0,0),(1,1),\ldots\}$. Still another example is 'closer than 10', the set $\{(x,y) \mid |x-y|<10\}$. Some members of this relation are $(1,10)$, $(10,1)$, and $(42,44)$. Neither $(11,1)$ nor $(1,11)$ is a member.

Those examples illustrate the generality of the definition. All kinds of relationships (e.g., 'both numbers even' or 'first number is the second with the digits reversed') are covered.

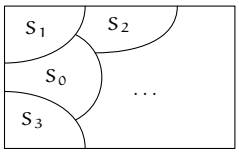
Equivalence Relations We shall need to express that two objects are alike in some way. They aren't identical, but they are related (e.g., two integers that give the same remainder when divided by 2).

A binary relation $\{(a,b),\ldots\}$ is an *equivalence relation* when it satisfies (1) *reflexivity*: any object is related to itself, (2) *symmetry*: if a is related to b then b is related to a, and (3) *transitivity*: if a is related to b and b is related to c then a is related to c. Some examples (on the integers): '$=$' is an equivalence relation, '$<$' does not satisfy symmetry, 'same sign' is a equivalence, while 'nearer than 10' fails transitivity.

Partitions In the 'same sign' relation $\{(1,3),(-5,-7),(0,0),\ldots\}$ there are three

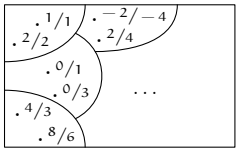
kinds of pairs, pairs with both numbers positive, pairs with both negative, and the one pair with both zero. So integers fall into exactly one of three classes, positive, or negative, or zero.

A *partition* of a set Ω is a collection of subsets $\{S_0, S_1, S_2, \ldots\}$ such that every element of S is an element of a subset $S_1 \cup S_2 \cup \cdots = S$, and overlapping parts are equal: if $S_i \cap S_j \neq \varnothing$ then $S_i = S_j$. Picture that Ω is decomposed into non-overlapping parts.



Thus the prior paragraph says that 'same sign' partitions the integers into the positives, and the negatives, and zero. Similarly, the equivalence relation '=' partitions the integers into one-element sets.

Another example is the set of strings consisting of a number, followed by a slash, followed by a nonzero number $\Omega = \{n/d \mid n, d \in \mathbb{Z} \text{ and } d \neq 0\}$. Define $S_{n,d}$ by: $\hat{n}/\hat{d} \in S_{n,d}$ if $\hat{n}d = n\hat{d}$. Checking that this is a partition of Ω is routine (observe for instance that $S_{4,3} = S_{8,6}$). This shows some parts, listing in each a couple of its infinitely many members.



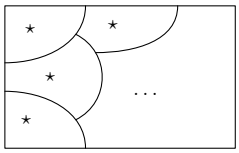
Every equivalence relation induces a partition, and every partition is induced by an equivalence. (This is routine to check.) Below are two examples.

Consider the equivalence relationship between two integers of 'gives the same remainder when divided by 2', the set $P = \{(-1,3), (2,4), (0,0), \ldots\}$. In the set P are two kinds of pairs, the pairs with both members even and the pairs with both members odd. This equivalence induces a partition where the parts are found by: for each x we define the set of numbers related to it $S_x = \{y \mid (x,y) \in P\}$. The parts are $\{\ldots, -3, -1, 1, 3, \ldots\}$ and $\{\ldots, -2, 0, 2, 4, \ldots\}$. Each part can be named in many ways; for instance, $\{\ldots, -3, -1, 1, 3, \ldots\}$ is S_1 and also is S_{-3}.

Now consider the partition of the natural numbers where two numbers are in the same part if they leave the same remainder when divided by 10, that is, if they have the same least significant digit. This partition is induced by the equivalence relation R defined by: two numbers n, m are related if they are together in the same part. For example, 3 is related to 33, but 3 is not

related to 102. Verifying the three conditions in the definition of equivalence are straightforward.

We call each part of a partition an *equivalence class*. We sometimes pick a single element of each equivalence class to be the *class representative*.



Usually when we pick representatives we have some natural scheme in mind. In that case we call them the *canonical* representatives. An example is that two fractions $3/5$ and $9/15$ are equivalent. In everyday work we often prefer to use the 'simplest form' or 'reduced form' fraction $3/5$ as the class representative.