Appendix A

Sets, Numbers, and Logic

The first section establishes some of the set notation used in the book and briefly reviews the number systems we need; the rest deals with logical implication and proofs—if-then statements, indirect proof, proof by contraposition, and what a counterexample is are all discussed, followed by a section on mathematical induction. Read this appendix alongside Chapter 1 if these are unfamiliar ideas; or use it as a reference when puzzled by something in the main text, which in the first few chapters cites it frequently.

A.0 Sets and Numbers.

First a quick mention about how sets are described; then we will comment on the numbers we will use.

Set notation.

Sets in general will be denoted here by capital letters: $S, T, \ldots$; in this book they will almost always be sets of numbers, or toward the end, of points in the plane, i.e., ordered pairs of numbers (x, y).

The notation $a \in S$ means "the number a is an element of the set S", or said more simply, "a is in S".

Sets can be defined either by listing their elements between braces:

$$S = \{1,\ 2,\ 3,\ 4,\ 5\}$$

or by describing the criteria for membership in the set:

$$S = \{x : x \text{ is an integer, } 1 \le x \le 5\},$$

read, "S is the set of all x such that x is an integer and..."

Set equality and inclusion. The notation is

$S = T$: the two sets have the same elements;

$S \subseteq T,\ T \supseteq S$: S is a subset of T (every $x \in S$ is also in T);

$S \subset T,\ T \supset S$: S is a proper subset of T ($S \subseteq T$ but $S \neq T$).

You can prove $S = T$ by proving $S \subseteq T$ and $T \subseteq S$; here is a simple example.

Example A.0A. Let $S = \{x : x^3 = x\}$. Prove that $S = \{-1, 0, 1\}$.

Solution. $\{-1,0,1\} \subseteq S$: $0^3 = 0,\ 1^3 = 1,\ (-1)^3 = -1$.

$S \subseteq \{-1,0,1\}$: $x^3 = x \ \Rightarrow\ x^3 - x = x(x+1)(x-1) = 0$

$\Rightarrow\ x = 0, -1,$ or $1.$ $\square$

Number systems. The evolution of our number system can be summarized roughly as the series of set inclusions

$$\emptyset \subset \mathbb{N} \subset \mathbb{N}_0 \subset \mathbb{Z} \subset \mathbb{Q} \subset \mathbb{R} \subset \mathbb{C}.$$

Let's talk briefly about each of these in turn.

In the beginning there was

$$\emptyset = \text{the empty set : the set with no elements.}$$

The empty set is a subset of every other set, but there is only one empty set — the set with no integers is the same as the one with no apples. Out of the empty set, man learned (only in the 20th century, actually) to construct

$$\mathbb{N} = \{1, 2, 3, \ldots, n, \ldots\} : \text{ the natural numbers.}$$

These have been around a long time, since they are needed for counting and describing the size of finite sets; it's just that they weren't defined until recently.

The set $\mathbb{N}$ was expanded a thousand years ago to

$$\mathbb{N}_0 = \{0, 1, 2, 3, \ldots\} : \text{ the non-negative integers;}$$

with the addition of zero, now the empty set also had a size. Zero made decimal notation and therefore effective calculation possible. (Try calculating with Roman numerals!)

But $\mathbb{N}_0$ is defective: you can add and multiply without leaving the system, but not always subtract: to solve $x + 5 = 3$, the system must be expanded to

$$\mathbb{Z} = \{\ldots, -2, -1, 0, 1, 2, \ldots\} : \text{ the integers.}$$

This causes some problems: negative numbers are traumatic — they don't count anything, which is what numbers are for, and the law $(-1)(-1) = 1$, passed so that $a(b+c) = ab+ac$ would always be true, has led one generation after another over the years to decide that mathematics is gibberish.

Even $\mathbb{Z}$ is defective: it doesn't contain $1/n$, a number that n-person families with pie for dessert find indispensable. In fact, the ancient Egyptians managed to do all their needed arithmetic using just the integers and the numbers $1/n$. To include them, one has to include their multiples, i.e., expand the system to

$$\mathbb{Q} = \{\tfrac{m}{n} : m, n \in \mathbb{Z}, \quad n \neq 0\} : \text{ the rational numbers.}$$

With the rational numbers we have finally what is called a *field*: a system in which one can add, subtract, multiply, and divide without ever leaving the system (and where these operations satisfy some well-known laws like $ab = ba$ and the distributive law alluded to above).

Unfortunately, many m/n can represent the same rational number. We can make the quotient m/n unique by specifying that $n > 0$ and that m and n have no common factors, i.e., that the representation is in *lowest terms*. You can't always write every fraction in lowest terms, however: it makes calculation too hard. For instance, any two rational numbers can be added only because they are "commensurable", that is, can be written so they have a common denominator; but when you do this, they won't be in lowest terms any more: $\frac{1}{2} + \frac{2}{3} = \frac{3}{6} + \frac{4}{6} = \frac{7}{6}$.

Reals and irrationals As we discuss at the end of Chapter 1, the system $\mathbb{Q}$ is still too small — it's not complete. Pythagoras discovered that the diagonal of the unit square was incommensurable with its side: in other words, that $\sqrt{2}$ was not rational. (Section A.2 gives his proof.) Suddenly there were a lot of new numbers, *irrational* numbers. There was no way of representing them except as lengths, that is, as points on a line, a representation not well-suited to calculation. But then, no one really needed them. (In a sense, it is only mathematicians who do.) At any rate, to include them, the number system had to be expanded to

$$\mathbb{R} \;=\; \text{the real numbers,}$$

thought of first as the points on a line, then many centuries later, after decimal notation had been invented, also as infinite decimals.

Like the smaller set of rational numbers, the real numbers also form a field: arithmetic operations on real numbers always lead to real numbers. They were constructed rigorously for the first time in the 19th century, in two different ways. The proofs that the so-constructed numbers have the right properties (including the Completeness Property of Chapter 1) take time and effort.

The *irrational* numbers are, in set notation,

$$\mathbb{R} \setminus \mathbb{Q}: \text{ everything in } \mathbb{R} \text{ that's not in } \mathbb{Q}.$$

They most emphatically do *not* form a field, since the arithmetic operations on irrationals do not necessarily lead to irrationals: for instance $\sqrt{2} \cdot \sqrt{2} = 2$. An irrational is either *algebraic* — like $\sqrt{2}$, a zero of a polynomial with coefficients in $\mathbb{Q}$ — or if not, *transcendental*. Proving a number is transcendental is hard.

Though we shall not make use of them in this book, one should mention the still bigger field $\mathbb{C}$ of *complex numbers* in which all polynomial equations $p(x) = 0$ have solutions; full recognition of these as a valid number system took hundreds of years, and many are still a little uncomfortable with them, since they don't measure lengths on a line. They bear the stigma of being called "imaginary", but of course in a sense they are.

Rational numbers and infinite decimals The beginning of Chapter 1 gives a brief sketch of the real numbers, thought of as infinite decimals: how you add and multiply them, and why the Completeness Property holds for them. Where do the terminating decimals fit in, and what's their relation to rational numbers?

(i) *A terminating decimal represents a rational number.*

Namely, it can be represented in the form $m/10^n$, if it has n decimal places:

$$3.141 = \tfrac{3141}{1000}, \qquad\qquad -1.42 = -\tfrac{142}{100}.$$

The rationals represented by terminating decimals are very special, however, since the above shows they are the ones writable (not in lowest terms, of course) with only powers of 10 in the denominator. Most rationals are not of this form.

(ii) *An infinite decimal is a rational number* $\Leftrightarrow$ *it is a repeating decimal.*

That is, after some point, it contains a group of digits which repeats for the rest of the decimal expansion, like $2.1333\ldots, -6.366014014014\ldots$.

Question 1 below illustrates why statement (ii) is true: briefly, a repeating decimal represents a geometric series, which can be summed to a rational number. Going the other way, in long division of n into m there are only a finite number of possibilities for the intermediate steps, so at some point, the process will start repeating.

Fact (i) above is the basis for the engineer's claim (very irritating to mathematicians) that "all numbers are rational": for in fact all experimental work and all calculations, whether done by hand or computer, of necessity use only terminating decimals.

Operations on sets. After this somewhat breezy account, let's return to more general facts about sets. Four operations produce new sets from old ones:

$$\begin{aligned} S \cup T &= \{x : x \in S \text{ or } x \in T\} && \textit{(union)};\\ S \cap T &= \{x : x \in S \text{ and } x \in T\} && \textit{(intersection)};\\ S \times T &= \{(x, y) : x \in S \text{ and } y \in T\} && \textit{(product)};\\ S \setminus T &= \{x : x \in S \text{ but } x \notin T\} && \textit{(difference)}. \end{aligned}$$

The word "or" in mathematics always has the inclusive sense: that is, "A or B" is rendered in ordinary speech by "A or B or both"; if the other (exclusive) sense is wanted, it would be written in mathematics as "A or B, but not both".

The notation (x, y) in the definition of $S \times T$ stands for the ordered pair of numbers x and y; the same notation is used for the open interval, but the context will always make it clear which is meant.

The definition of the union and intersection (and product as well, though we shall not need it) can be extended from two sets to a finite or infinite collection of sets $\{S_i\}$, where i runs over some set I of indices: the notation for this is

$$\begin{aligned} \bigcup_{i\in I} S_i &= \{x : x \in S_i \text{ for at least one } i\} && \textit{(union)}\\ \bigcap_{i\in I} S_i &= \{x : x \in S_i \text{ for every } i\} && \textit{(intersection)} \end{aligned}$$

If the collection is finite, the set of indices I is usually $\{1, 2, \ldots, n\}$, and if infinite, it is usually the natural numbers $\mathbb{N}$, although there are occasional exceptions. As examples, using the usual notation for intervals:

$$[a, b] = \{x \in \mathbb{R} : a \le x \le b\}, \qquad (a, b) = \{x \in \mathbb{R} : a < x < b\},$$

we have

$$\bigcap_{n\in\mathbb{N}} (0, 1/n) = \emptyset; \qquad \bigcup_{a\in\mathbb{R}} (0, a^2) = \mathbb{R}^+ .$$

Functions on sets.

Functions of one and two real variables are discussed in detail in Chapters 9 and 24 respectively. The words introduced below to describe functions are used very briefly in Chapter 9 and in Chapter 23, so there's no urgency in reading this quick account now. It's put here just to have as a reference.

A function f from a set S to a set T is given by a rule associating with each element $s \in S$ a corresponding element of T, denoted $f(s)$; in notation:

$$f : S \to T\ , \qquad s \to f(s)\ .$$

It is called

injective, if it sends distinct elements of S into distinct elements of T:

$$s_1 \neq s_2 \quad \text{implies} \quad f(s_1) \neq f(s_2)\ ;$$

surjective, if for each $t \in T$ there is an $s \in S$ such that $f(s) = t$;

bijective, if it is both injective and surjective.

Examples A.0B. Let $f_1(x) = x^2$ and $f_2(x) = 2x$; classify them as injective, surjective, or bijective on the following sets (here $\mathbb{R}^+ = \{x \in \mathbb{R} : x > 0\}$) :

(a) $f_1 : \mathbb{R}^+ \to \mathbb{R}$ (b) $f_1 : \mathbb{R} \to \mathbb{R}^+ \cup \{0\}$ (c) $f_1 : \mathbb{R}^+ \to \mathbb{R}^+$

(d) $f_2 : \mathbb{Z} \to \mathbb{Z}$ (e) $f_2 : \mathbb{Q} \to \mathbb{Q}$

Solution. (a) injective, since if $a > b > 0$, then $a^2 > b^2$; (b) surjective, since every non-negative real number has a square root; (c) bijective, since every positive real has a unique positive square root; (d) injective; (e) bijective.

If $f : S \to T$ is bijective, it has a unique *inverse* $g : T \to S$ defined by

$$g(t) = s \quad \text{if} \quad f(s) = t\ ;$$

given any t, such an s exists since f is surjective; the s is unique since f is injective. The association $s \to f(s)$ gives what is called a *one-one correspondence* between the elements of S and T: each element of S is paired with one and only one element of T, and vice-versa.

In the examples above, the inverse of $f_1 : \mathbb{R}^+ \to \mathbb{R}^+$ is the function g_1 given by $g_1(x) = \sqrt{x}$; the inverse of $f_2 : \mathbb{Q} \to \mathbb{Q}$ is the function g_2, where $g_2(x) = \frac{1}{2}x$.

The rest of this appendix is devoted to reviewing some points of logic.

Questions A.0 (Answers at end of Appendix A.)

1. (a) Show that $1.1232323\ldots$ is a rational number. (Use Section 4.2 (2).)
   (b) Write $\frac{3}{7}$ as an infinite decimal, using division; see why it repeats.

2. Express these sets in terms of the standard sets and intervals, and the four finite operations.

(a) $\{x \in \mathbb{R} : x \geq 0\}$ (b) $\{x \in \mathbb{Q} : a \leq x \leq b\}$

(c) $\{\text{pairs } (x, y) : a < x < b,\ c < y < d\}$ (d) $\{x \in \mathbb{Q} : x > 0\}$

(e) $\bigcup_{a \in \mathbb{N}_0} [a, a+1]$ (f) $\bigcup_{a \in \mathbb{Z}} (a, a+1)$

3. Let $S = \{x \in \mathbb{R} : x(1-x) > 0\}$. Prove $S = (0, 1)$, by proving $\subseteq$ and $\supseteq$.

A.1 If-then statements

In mathematics, statements to be proved can often be put in the form

(1) if A, then B; $A \Rightarrow B$ (read: "A implies B").

The two forms say the same thing; the second form uses a symbol called the "forward implication arrow". The A and B represent simpler statements:

A is the **hypothesis**: "what's given", "what's known";

B is the **conclusion**: "what follows", "what's to be proved".

Here are some examples. Note that often a preliminary statement must be made, explaining what the symbols in the $A \Rightarrow B$ statement stand for.

Examples A.1A

(i) A, B, C are the vertices of a triangle; a, b, c are the non-zero lengths of the opposite sides, respectively.

$$ACB \text{ is a right angle } \Rightarrow\ a^2 + b^2 = c^2\ . \qquad \textit{(true)}$$

(ii) Let a be a real number.

$$2a^6 + a^4 + 3a^2 = 0\ \Rightarrow\ a = 0\ . \qquad \textit{(true)}$$

(iii) $f(x)$ is differentiable $\Rightarrow$ $f(x)$ is continuous. *(true)*

(iv) Let a, b, n be positive integers.

$$n \text{ divides } ab\ \Rightarrow\ n \text{ divides } a \text{ or } b\ . \qquad \textit{(false)}$$

We shall use where possible the arrow notation, since it allows the hypothesis and conclusion to stand out clearly. But with the if-then form one can avoid a preliminary sentence:

"If a is a real number such that $2a^6 + a^4 + 3a^2 = 0$, then $a = 0$."

However, the problem with all of this is that in ordinary mathematical writing, the hypothesis and conclusion may not be spelled out so clearly; it is you that has to extract them from the prose sentence. For instance, (ii) and (iii) would probably appear in the form:

(ii) 0 is the only real root of $2x^6 + x^4 + 3x^2 = 0$;

(iii) a differentiable function is continuous.

Thus, if a statement is given in the form $A \Rightarrow B$, some of the work has already been done for you.

In bad mathematical writing, the ambiguities of English may make it impossible to decide what the implication is; for example,

"$f(x)$ has a relative maximum or minimum at a point a where $f'(a) = 0$."

Does this say (in rough outline) "relative max/min at a $\Rightarrow$ $f'(a) = 0$" or "$f'(a) = 0$ $\Rightarrow$ relative max/min at a"? Your guess is as good as mine.

Converse. If we interchange hypothesis and conclusion in $A \Rightarrow B$, we get

(2) $$B \Rightarrow A \qquad (\text{ or } A \Leftarrow B\ ),$$

which is called the *converse* to the statement (1).

Examples A.1B The converses to A.1A are (omitting the preliminaries):

(i) $a^2 + b^2 = c^2 \ \Rightarrow\ ACB$ is a right angle. *(true)*

(ii) $a = 0 \ \Rightarrow\ 2a^6 + a^4 + 3a^2 = 0.$ *(true)*

(iii) $f(x)$ continuous $\Rightarrow\ f(x)$ differentiable. *(false)*

(iv) n divides a or $b \ \Rightarrow\ n$ divides ab . *(true)*

As one can see from these examples, the truth or falsity of the converse is unrelated to the truth or falsity of the original statement.

Equivalent statements.

We can combine the two implication arrows into one double-ended arrow:

(3) $$A \quad \Leftrightarrow \quad B\ ,$$

which is a true statement if both $A \Rightarrow B$ and $A \Leftarrow B$ are true. If this is so, we say A and B are *equivalent* statements. To give our examples one last time:

Examples A.1C

(i) $a^2 + b^2 = c^2 \quad \Leftrightarrow \quad ACB$ is a right angle. *(true)*

(ii) $a = 0 \quad \Leftrightarrow \quad 2a^6 + a^4 + 3a^2 = 0.$ *(true)*

(iii) $f(x)$ is differentiable $\quad \Leftrightarrow \quad f(x)$ is continuous *(false)*

(iv) n divides a or $b \quad \Leftrightarrow \quad n$ divides ab . *(false)*

Necessary and sufficient. There are two verbal forms of $\Leftrightarrow$ which are in common use. We will mostly avoid them, but others do not, so you should know them. They are:

$$A \text{ if and only if B} \qquad (\text{abbreviated: } A \text{ iff } B)\ ;$$

A is a necessary and sufficient condition for B (abbreviated: nasc).

Occasionally these are separated into their component parts:

$A \Rightarrow B$: A is a *sufficient* condition for B (if A is true, B follows);

$B \Rightarrow A$: A is a *necessary* condition for B (i.e., B can't be true unless A is also true, since B implies A).

The "if and only if" is also separated:

"A, if B" : $B \Rightarrow A$; "A, only if B": $A \Rightarrow B$.

This last is the worst, since ordinary English usage is different—"only if" is considered the same as "if and only if":

"You can go only if you are invited" (..."but if you are invited, you can go" is automatically implied, or one had better be prepared for insurrection).

Stronger and weaker. If $A \Rightarrow B$ is true, but $B \Rightarrow A$ is false, we say:

A is a *stronger* statement than B; B is *weaker* than A.

Example A.1E

"ΔABC is equilateral" is stronger than "ΔABC is isosceles", since

$$\Delta ABC \text{ is equilateral } \Rightarrow \Delta ABC \text{ is isosceles.}$$

The same terminology applies to entire "if-then" statements (theorems):

Example A.1F The if-then statement

(4) ΔABC is equilateral $\Rightarrow$ ΔABC has two equal angles

can be made stronger in two different ways: make the hypothesis weaker:

(5) ΔABC is isosceles $\Rightarrow$ ΔABC has two equal angles;

or make the conclusion stronger:

(6) ΔABC is equilateral $\Rightarrow$ ΔABC has three equal angles.

Both (5) and (6) are stronger than (4) since they both imply (4): if you know (5) or (6) is true, then (4) follows, but not vice-versa.

Strengthen $A \Rightarrow B$ by making B stronger, or A weaker.

Questions A.1 (Answers at end of Appendix A)

1. Write in the form $A \Rightarrow B$, with a preliminary sentence if appropriate; write the converse without using $\Rightarrow$; mark the original statement and its converse as true or false.
   (a) An integer n is divisible by 6, provided it is divisible by 2 and 3.
   (b) The derivative of x^2 is $2x$.
   (c) A quadrilateral whose diagonals are equal is a rectangle.
   (d) Let $\{a_n\}$ be an increasing sequence. If $\{a_n\}$ is bounded, it has a limit.
   (e) Two parallel lines make equal angles with a line intersecting them.
2. Which statement is stronger, which weaker?
   (a) $a \geq 0$; $a > 0$ (b) $\{a_n\}$ is bounded above; $\{a_n\}$ is bounded
3. Form all stronger-weaker pairs from the following statements; give the pairs by using the $\Rightarrow$ symbol. (If this gets confusing, the boxed statement above will be helpful.)
   (a) An increasing sequence which is bounded above has a limit.
   (b) A bounded increasing sequence has a limit.
   (c) A bounded monotone sequence has a limit.

A.2 Contraposition and indirect proof.

We turn now to discussing a style of mathematical proof which involves forming the negatives of statements.

Negation. In general, if A is a statement, we will use either not–A or $\sim A$ to denote its negation. Often the word "not" doesn't appear explicitly in the negation. Here are three examples (in the first, a is a positive integer).

A	not–A
a is prime	a is composite or $a = 1$
$a > 2$	$a \leq 2$
$4a^2 + 2 = 3b$	$4a^2 + 2 \neq 3b$

Contraposition. In proving $A \Rightarrow B$, sometimes it is more convenient to use *contraposition*, i.e., prove the statement in its contrapositive form:

(7) not–$B \Rightarrow$ not–A (*contrapositive* of $A \Rightarrow B$).

This means exactly the same thing as $A \Rightarrow B$: if you prove one, you've proved the other. We will give a little argument for this later; however you will probably be even more convinced by looking at examples. Below, the original statement is on the left, the contrapositive is on the right; they say the same thing.

$$f(x) = x^2 \;\Rightarrow\; f'(x) = 2x \qquad f'(x) \neq 2x \;\Rightarrow\; f(x) \neq x^2$$
$$a \geq 0 \;\Rightarrow\; \sqrt{a} \text{ real} \qquad \sqrt{a} \text{ not real} \;\Rightarrow\; a < 0$$

Example A.2A Prove $2a^6 + a^4 + 3a^2 = 0 \;\Rightarrow\; a = 0$.

Solution. We use contraposition (the last line is overkill):

$$\begin{aligned} a \neq 0 \;&\Rightarrow\; a^2 > 0,\ a^4 > 0,\ a^6 > 0; \\ &\Rightarrow\; 2a^6 + a^4 + 3a^2 \;>\; 0; \\ &\Rightarrow\; 2a^6 + a^4 + 3a^2 \;\neq 0. \end{aligned}$$

□

Indirect proof. This has the same style as contraposition but is more general. To give an indirect proof that a statement S is true, we assume it is not true and derive a contradiction, i.e., show some statement C is both true and false. C can be anything.

Example A.2B Prove that $\sqrt{2}$ is irrational.

Solution. (Indirect proof). Suppose it were rational, that is,

$$\sqrt{2} \;=\; \frac{a}{b}\;;$$

we may assume the fraction on the right is in lowest terms, i.e., a and b are integers with no common factor. (Call this last clause "statement C".)

If we cross-multiply the above and square both sides, we get

$$2b^2 = \; a^2\;;$$

the left side is even, so the right side is even, which means a itself is even (since the square of an odd number is easily seen to be odd). Thus we can write $a = 2a'$,

where a' is an integer. If we substitute this into the above equation and divide both sides by 2, we get

$$b^2 = 2a'^2 ;$$

by the same reasoning as before, b is even. Since we have shown both a and b are even, they have 2 as a common factor; but this contradicts the statement C. □

The above (attributed to Pythagoras) is probably the oldest recorded indirect proof. Note how the statement C to be contradicted just appears in the course of the proof; it's not part of the statement of the theorem.

The above example is a little atypical for us, in that almost always in this book the statement S to be proved will be an if-then statement $A \Rightarrow B$. To prove it indirectly, we have to derive a contradiction from the assumption that $A \Rightarrow B$ is false, i.e., that A does not imply B: in other words, A can be true, yet B be false. So we can now formulate

Indirect proof for if-then statements.

To prove $A \Rightarrow B$ indirectly, assume A true but B false, and derive a contradiction: C and not-C are both true.

Our earlier *proof by contraposition* is just the special type of indirect proof where $C = A$. Namely, to prove $A \Rightarrow B$ by contraposition, we

(a) assume A true and B false (i.e., not-B true);

(b) prove not-$B \Rightarrow$ not-A (the contrapositive).

It follows that not-A is true, which contradicts our assumption that A is true.

To confuse you a little further, we illustrate the difference between the two styles of proof by giving two proofs of a simple proposition.

Proposition. $a^2 = 0 \;\Rightarrow\; a = 0$.

Proof by contraposition.
$$\begin{aligned} a \neq 0 &\;\Rightarrow\; a > 0 \text{ or } a < 0 ; \\ &\;\Rightarrow\; a^2 > 0 ; \\ &\;\Rightarrow\; a^2 \neq 0. \end{aligned}$$
□

Indirect proof. Suppose the conclusion is false, that is,

$$a^2 = 0, \qquad \text{but} \quad a \neq 0 .$$

Since $a \neq 0$, we can divide both sides of the above equation by a; this gives

$$a = 0 , \tag{8}$$

which contradicts our supposition that $a \neq 0$. □

Why not just stop the proof at line (8) — it says $a = 0$ and isn't that what we were supposed to prove?

This would be wrong; the last line of the proof is absolutely essential. We only got to line (8) by making a false supposition: that $a \neq 0$. Therefore (8) has no validity in itself; it is only a line in a bigger argument whose ultimate goal is to produce a contradiction.

The advantage of contraposition over the more general type of indirect proof is that since we know at the outset the statement A that is going to be contradicted, what has to to be proved (not–$B \Rightarrow$ not–A) becomes a direct statement that we hope can be proved by a direct argument.

The general argument against all indirect proofs is that they require you to focus on a false statement (A is true but B is not), and derive from it other false statements until finally you get a contradiction. Such a proof requires you to read one wrong thing after another, until a moment arrives when the author proudly announces, "This contradicts statement C !" but you can't remember that C was ever mentioned.

The other problem with indirect proofs (this applies to contraposition also) is that they require you to form the negation of statements, which is not always so easy to do in analysis, since many of the common statements are linguistically rather complicated. Directions for negating a statement are given in Appendix B, but students are well-advised not to read it until they have read most of the book; those who eat the apple of negation tend to fall into the habit of trying to prove even the simplest statements indirectly (and generally incorrectly).

In this book, we will avoid indirect proofs—even proofs by contraposition—whenever possible; if an indirect proof must be given, negation will be treated informally, which is how most professional mathematicians handle it.

Questions A.2

1. Formulate the negative statement without using "not":
   (a) In the plane, lines L and M are parallel.
   (b) Triangle ABC is isosceles.
   (c) There are infinitely many prime numbers.
2. Write the converse and contrapositive, using $\Rightarrow$, and mark T or F:

   The square of an odd integer a is odd.
3. Prove the following by contraposition (the a_i are real numbers):
   (a) if $a_1 a_2 < 0$, exactly one of the $a_i < 0$.
   (b) if $a_1 + \ldots + a_n = n$, at least one $a_i \geq 1$.

A.3 Counterexamples.

Some statements in mathematics are particular, i.e., they assert that something is true for some definite numbers, or other objects. For example.

$3^2 + 4^2 = 5^2$; ΔABC is isosceles; there is a number ≥ 22.

Other statements are general; they assert something about a whole class of numbers or other objects. For example:

(i) if a and b are numbers satisfying $a^2 = b^2$, then $a = b$;

(ii) a triangle with three equal sides has three equal angles;

(iii) every positive integer n is the sum of four squares of integers:

$$n = a_1^2 + a_2^2 + a_3^2 + a_4^2 \ ;$$

(iv) if a, b, c are numbers satisfying $ab = ac$, then $b = c$.

These respectively assert that something is true about any numbers satisfying $a^2 = b^2$, all equilateral triangles, all positive integers, any numbers satisfying $ab = ac$.

As it happens, statements (ii) and (iii) are true — (iii) is hard to prove — while (i) and (iv) are false. The problem we consider is:

How does one show a general statement like (i) *or* (iv) *is false?*

Since a general statement claims something is true for every member of some class of objects, to show it is false we only have to produce a single object in that class for which the general statement fails to hold. Such an object is called a **counterexample** to the general statement. For example a counterexample to (i) would be the pair $a = 3,\ b = -3$. (What would be a counterexample to (iv)?)

Example A.3 (uses Prop. 2.4)

(a) Prove: if the sequence $\{a_n\}$ is bounded, then $\{a_n^2\}$ is bounded.

(b) In part (a), can "bounded" be replaced by "bounded above"?

Solution.

Part (a). By hypothesis, there is a B such that (see Prop. 2.4)

$$|a_n| < B \quad \text{for all } n.$$

Squaring both sides, and using the law $|ab| = |a||b|$,

$$|a_n^2| < B^2 \quad \text{for all } n.$$

Therefore $\{a_n^2\}$ is bounded, again by Prop. 2.4.

Part (b). Solution No. 1 (by our good student)

Yes, "bounded" can be replaced by "bounded above"; to prove it, just drop the absolute value signs from the above proof:

$$\begin{aligned} \{a_n\} \text{ is bounded above } &\Rightarrow\ a_n < B \quad \text{for all } n; \\ &\Rightarrow\ a_n^2 < B^2 \quad \text{for all } n; \\ &\Rightarrow\ \{a_n^2\} \text{ is bounded above .} \quad \square?? \end{aligned} \tag{9}$$

(STOP; don't continue until you have spotted the error!)

Solution No. 2 (by our better student)

No, "bounded" cannot be replaced by "bounded above", since in the above argument which tries to prove the amended statement, (9) can fail if $a_n < 0$; thus the proof doesn't work, so the amended statement must be false. □??

Solution No. 3 (by our best student, and correct)

The amended statement is false; the sequence $\{-n\}$ is a counterexample, since it is bounded above by 0, but $\{n^2\}$ is not bounded above. □

Students in general seem to dislike counterexamples and are reluctant to produce them. Asked to show some general statement S is false, like our "better" student above they usually produce some reasoning which tries to prove S, but fails. Then they point out that the proof doesn't work, and conclude the statement must be false.

One can sympathize with this psychologically, but mathematically it's nonsense.

> *The failure of your attempted proof of S doesn't show S is false, since someone else might come along with a different argument which succeeds in proving it.*

Produce a counterexample to S instead; then no one can ever prove it.

Like students, mathematicians dislike counterexamples (particularly to their own theorems!) and producing them is a sure way to lose friends and alienate people. There seems to be something unsporting about demolishing a whole edifice of theorems by a single counterexample. Alas, it happens; worst of all in Ph.D. orals: a committee member wonders aloud how the candidate's theorem would apply to a favorite example, and a minute later the example has turned into a counterexample.

Of course you don't want to see a counterexample to your theorem: you want to know where your reasoning went astray. But even your best friend won't tell you; you have to figure it out yourself.

Questions A.3

1. Decide which of these statements is false; for each such, prove that it is false.

(a) Every positive integer is the sum of three squared integers (some of which can be zero).

(b) Three lines in the plane, no two of which are parallel, determine a unique triangle whose sides lie on the lines.

(c) It is conjectured there are an infinity of "twin primes": that is, pairs $(n, n+2)$ where both numbers are prime, like $(5, 7)$, $(11, 13)$, and $(17, 19)$. Show that in a twin prime pair, the number between the two primes is divisible by 3.

A.4 Mathematical induction.

This is a way of proving a proposition whose statement involves all positive integers $n \geq$ some n_0. Some examples (note the different values of n_0 used):

(a) $1 + x + \ldots + x^n = \dfrac{1 - x^{n+1}}{1 - x}, \quad n \geq 0.$

(b) A positive integer n is the product of one or more primes, if $n \geq 2$.

(c) The sum of the interior angles of an n-sided polygon is $(n-2)\pi, \quad n \geq 3.$

(d) $\displaystyle\int \frac{dx}{x^n} = \frac{x^{1-n}}{1-n}, \quad n \geq 2.$

We will denote the proposition by $P(n)$, to show its dependency on n. Though most think of it as a single proposition, proof by induction depends on thinking of it as a whole sequence of propositions, one for each value of n.

Proof by Mathematical Induction To prove $P(n)$, $n \geq n_0$,

(a) *prove* $P(n_0)$ *(the basis step);*

(b) *prove* $P(n+1)$ *; in the proof you are allowed to use* $P(n)$, *and if necessary,* $P(k)$ *for any lower values,* $n_0 \leq k \leq n$, *as well (the induction step).*

The best way to approach proof by induction is to see a variety of examples. As you will see, the boxed statement above has to adapted in various small ways for the different examples, but it is a good place to start.

Example A.4A Prove: $1^2 + 2^2 + 3^2 + \ldots + n^2 = \dfrac{n(n+1)(2n+1)}{6}, \quad n \geq 1.$

Solution. By induction. The basis step $P(1)$ says $1^2 = (1 \cdot 2 \cdot 3)/6$, which is true. Here is the induction step.

$$\begin{aligned} 1^2 + 2^2 + \ldots + (n+1)^2 &= (1^2 + 2^2 + \ldots + n^2) + (n+1)^2 \\ &= \frac{n(n+1)(2n+1)}{6} + (n+1)^2, && \text{using } P(n); \\ &= \frac{(n+1)\big[n(2n+1) + 6(n+1)\big]}{6}, && \text{factoring out } n+1; \\ &= \frac{(n+1)(n+2)(2n+3)}{6}, && \text{after some algebra;} \end{aligned}$$

and this last is the right side of $P(n+1)$. □

The above shows that if $P(n)$ is true, then $P(n+1)$ follows: $P(n) \Rightarrow P(n+1)$. Since we know $P(1)$ is true, this shows $P(2)$ is true; this in turn shows $P(3)$ is true, and continuing in this way, $P(n)$ is shown to be true for any value of n — it's been likened to a row of falling dominos: you tip over the first one, which then tips over the second, and ultimately the n-th is reached.

Example A.4B Prove $n! \geq 2^n$, for $n \geq$ some n_0.

Solution. The induction step runs:

$$\begin{aligned} (n+1)! &= n!\,(n+1) \\ &\geq 2^n(n+1), && \text{using } P(n); \\ &\geq 2^{n+1}, && \text{if } n+1 \geq 2\ . \end{aligned}$$

So we have proved $P(n) \Rightarrow P(n+1)$, if $n \geq 1$, i.e., for all $n \in \mathbb{N}$. □

But take $n = 2$: $2! \not\geq 2^2$. $P(2)$ is false!

Didn't we prove that $P(1) \Rightarrow P(2)$? Yes, but $P(1)$ is also false. So is $P(3)$.

The basis step is $P(4)$: $4! \geq 2^4$; the proposition is true only for $n \geq 4$. □□

The basis step is trivial or obvious most of the time, so students tend to skip it or ignore it, focussing on the harder induction step. But this can get you into trouble, as illustrated above.

Sometimes people argue over what the basis step should be. In mathematics, empty sums are assigned the value 0, and empty products are assigned the value 1 (as for example: $a^0 = 1,\ 0! = 1$). Thus the basis step for Example A.4A could also be taken as $n = 0$.

In addition to proof by induction, there is also *inductive definition* or as it is also called, *recursive definition*, in which the terms of a sequence $\{a_n\}$, $n \geq n_0$, are defined by expressing them in terms of lower values of n; as the basis, a starting value a_{n_0} must also be given.

Example A.4C Let $a_n = a_{n-1} + \dfrac{1}{n(n+1)}$, $a_0 = 0$. Find a formula for a_n.

Solution. We have

$$\begin{aligned} a_1 &= a_0 + 1/(1 \cdot 2) = 1/2\ , \\ a_2 &= a_1 + 1/(2 \cdot 3) = 2/3\ , \\ a_3 &= a_2 + 1/(3 \cdot 4) = 3/4\ , \end{aligned}$$

so we guess $\quad a_n = \dfrac{n}{n+1}\ .$

Taking this last statement as $P(n)$, we prove it by induction. It is true for a_0; as the induction step, we get

$$\begin{aligned} a_n &= a_{n-1} + \frac{1}{n(n+1)} \\ &= \frac{n-1}{n} + \frac{1}{n(n+1)}\ , && \text{using } P(n-1), \\ &= \frac{n}{n+1}\ , && \text{by algebra,} \end{aligned}$$

which completes the proof by induction. □

Notice that here we proved $P(n)$, using $P(n-1)$. This just amounts to a change of variable in the boxed method on the previous page, and is therefore equally valid.

Example A.4D Let $a_0 = 1,\ a_1 = 2,$ and $a_{n+2} = 2a_{n+1} - a_n, \quad n \geq 0.$
Prove that $a_n = n + 1$ for $n \in \mathbb{N}$.

Solution. We use as the induction step the proof of $P(n+2)$:

$$\begin{aligned} a_{n+2} &= 2a_{n+1} - a_n, && n \geq 0; \\ &= 2(n+2) - (n+1), && \text{using } P(n+1) \text{ and } P(n); \\ &= n + 3\ . \end{aligned}$$ □

Here we proved $P(n+2)$, using in the proof not just $P(n+1)$, but $P(n)$ as well. When one uses in the proof of $P(n)$ not just the preceding value but lower values of n as well, the proof method is generally referred to as **strong** or **complete** induction; in this style of induction, often more than one value of n is needed for the basis step.

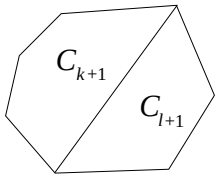
> *In strong induction, the basis step consists of all $P(n)$ not covered by the argument in the induction step, i.e., for which there are no lower $P(k)$ available to imply $P(n)$.*

In the example above, the proof of $P(n)$ uses the two previous values of n; therefore $P(1)$ and $P(0)$ are not covered by this proof, since these cases do not have two previous values — they must be proved separately, as the basis step. (In this case, both are trivial, since we are given $a_0 = 1$ and $a_1 = 2$.)

Example A.4E Prove that the sum of the interior angles of a convex polygon with n sides is $(n-2)\pi$.

Solution. A convex polygon is one where any line segment joining two vertices lies inside the polygon.

We give a proof by strong induction. To prove $P(n+1)$ we are given a convex polygon C_{n+1} with $n+1$ sides. Draw a line connecting two non-adjacent vertices. It divides C into two convex polygons C_{k+1} and C_{l+1} as shown, having respectively $k+1$ and $l+1$ sides. We have



$$\begin{aligned} \text{angle sum of } C_{n+1} &= \text{angle sum of } C_{k+1} + \text{angle sum of } C_{l+1} \\ &= (k-1)\pi + (l-1)\pi\ , && \text{by strong induction;} \\ &= (k+l-2)\pi = (n-1)\pi, && \text{since } k+l = n+1. \end{aligned}$$

The case not covered by the above is when no such dividing line segment exists; in this case C must be a triangle, and therefore $P(3)$ is the basis step, which is true: the sum of the angles of a triangle is π. □

Here one could regard the basis step as the hardest step, since probably more students could give the above induction argument than remember how to prove the statement about a triangle. The induction in this example is strong induction, since we don't know how many sides each of the smaller polygons has — just that it's less than $n+1$. (One could also give a proof by regular induction: see Question 3, but read the comments in the Answers since this work is often mishandled.)

Example A.4F Prove that every integer $n \geq 2$ is the product of primes.

Solution. Here is a case where you can only use strong induction, since there is no relation between the prime factorizations of n and $n+1$.

> Keeping in mind the freshman in one class, who on seeing the above proposition on the board, yelled triumphantly, "False!! Five is not a product of primes!" I hasten to add that in higher mathematics, a sum is allowed to have just one term, and a product just one factor. They can even have none, if you believe in the empty sums and products we mentioned earlier.

If n is prime, we are done. If not, it factors into the product of two smaller positive integers, both ≥ 2 (since the factorization is not the trivial one $n = n \cdot 1$):

$$\begin{aligned} n &= n_1 \cdot n_2, \qquad 2 \leq n_1, n_2 < n; \\ &= (p_1 p_2 \cdots p_k)(q_1 q_2 \cdots q_l), \qquad p_i, q_j \text{ primes}, \end{aligned}$$

since by strong induction, we can assume the smaller numbers n_1 and n_2 factor into the product of primes. □

Remarks. Question 3 asks for a proof that the method of induction works; it is a good example of indirect proof.

The problem many beginners have with proof by induction is, of course, the apparent circularity: "How can you assume and use $P(n)$ in the proof, since that's what you're trying to prove?" The answer is, it's $P(n+1)$ that you're trying to prove.

The same problem appears in recursive definitions: in the briefest and most efficient form, the definition of $n!$ is

$$n! = n \cdot (n-1)!, \qquad 0! = 1.$$

The definition looks circular, but because the factorial on the right is for a smaller n, the definition makes sense.

In many mathematical arguments, induction is concealed by such phrases as "similarly" or "continuing in the same way, we see that". This is a genial tradition that makes for easier reading; the preference in this book is for such informal stratagems. Induction is used explicitly only for arguments that can't be made without it. Example A.4A is a good one for that: no calculation proving the formula for $\sum_1^n i^2$ for just the first few values of n will convince anyone that the general formula for the sum is correct; only induction will do that.

Questions A.4

1. In Example A.4B, proving $n! \geq 2^n$, show that $P(0)$ is true, according to the conventions about empty products described in the subsequent remarks. If $P(0)$ is true, why can't we take it as the basis step?

2. Prove that $n < 2^n$ for all $n \in \mathbb{N}_0$.

3. Prove the method of regular induction works: that is, if $P(n_0)$ is true and $P(n+1)$ is true whenever $P(n)$ is, for $n \geq n_0$, then $P(n)$ is true for all $n \geq n_0$.

(Hint: consider the set $S = \{n \geq n_0 : P(n) \text{ is false}\}$; it has a least element.)

4. Prove Example A.4E using regular induction.

5. In Example A.4F, what is the basis step?

Exercises

A.4

1. Prove by induction that $1 + x + \ldots + x^n = \dfrac{1 - x^{n+1}}{1 - x}, \quad n \geq 0$.

2. Find a formula for $1 + 3 + 5 + \ldots + (2n - 1)$ and prove it by induction.

3. Let D denotes differentiation with respect to x. From the product rule for differentiation, $D(uv) = uDv + vDu$, and the fact that $Du = 0$ if $u(x)$ is a constant function, prove by induction that $D(x^n) = nx^{n-1}$, if $n \in \mathbb{N}_0$.

4. The coefficients of a series $\displaystyle\sum_0^\infty a_n x^n$ are given by $a_{n+2} = \dfrac{n+1}{n+3}\, a_n$.

(a) Find the power series if $a_0 = 1$ and $a_1 = 0$; prove it by induction.

(b) Find the power series if $a_0 = 0$ and $a_1 = 1$; prove it by induction.

5. The terms of a sequence $a_0, a_1, a_2, \ldots$ are given by the recursive relation

$$a_{n+1} = 2a_n - a_{n-1} + 2, \qquad a_0 = 0,\ a_1 = 1.$$

Find a formula for a_n, and prove it.

6. Fermat's Little Theorem is the basis of RSA encryption, widely used to guarantee website security. The theorem says:

if p is prime, then $n^p - n$ is an integer multiple of p, for all $n \in \mathbb{N}$.

(a) Prove this by induction if $p = 3$.

(b) Prove it if $p = 5$.

(c) Show it is false if $p = 4$ (cf. Section A.3).

7. With an unlimited supply of three-cent and seven-cent stamps, it is possible to make any integral postage n, when $n \geq n_0$. Find the smallest n_0 for which this is true, and prove it by strong induction. What is the basis step?

8. (a) Prove that $1^3 + 2^3 + \ldots + n^3 < n^4$, if $n \in \mathbb{N}$, $n > 1$.

(b) Prove the sum in part (a) is $< n^4/2$, if $n > 2$ (a bit harder).

9. Fix an $a \in \mathbb{N}$. Prove by strong induction (regular induction is clumsier here) that any $n \in \mathbb{N}_0$ can be written in the form below; what is the basis step?

$$n = qa + r, \qquad \text{where } q \in \mathbb{N}_0, \quad 0 \leq r < a\ .$$

10. Prove $n^2 < 2^n$ for $n \geq n_0$. (If stuck, Question 2 might be helpful.)

Answers to Questions

A.0

1. (a) $1.12323\ldots = 1.1 + (.1)(.2323\ldots)$;
$.2323\ldots = .23(1 + 10^{-2} + 10^{-4} + \ldots) = .23(100/99)$, by 4.2, (2).

(b) $3/7 = .428571428571\ldots$; repetition is inevitable.

2. (a) $\mathbb{R}^+ \cup \{0\}$, or $[0, \infty)$ (b) $\mathbb{Q} \cap [a, b]$ (c) $(a, b) \times (c, d)$

(d) $\mathbb{Q} \cap \mathbb{R}^+$ (e) $[0, \infty)$ (f) $\mathbb{R} \setminus \mathbb{Z}$

3. $(0, 1) \subseteq S$: if $x \in (0, 1)$, then $x > 0$, $1 - x > 0$, so $x(1 - x) > 0$.

$S \subseteq (0, 1)$: if $x(1-x) > 0$, either both factors are positive or both factors are negative. If both are positive, $x > 0$ and $1 - x > 0$, so $x \in (0, 1)$. If both are negative, $x < 0$ and $1 - x < 0$, i.e., $x < 0$ and $x > 1$, which is impossible.

A.1

1. (a) n is divisible by 2 and 3 $\Rightarrow$ n is divisible by 6. (T)
Converse: A number divisible by 6 is divisible by 2 and 3. (T)

(b) $f(x) = x^2 \Rightarrow f'(x) = 2x$. (T)
Converse: If $f'(x) = 2x$, then $f(x) = x^2$. (F)

(c) Q is a quadrilateral with equal diagonals $\Rightarrow$ Q is a rectangle. (F)
Converse: The diagonals of a rectangle are equal. (T)

(d) Let $\{a_n\}$ be increasing; $\{a_n\}$ bounded $\Rightarrow$ $\{a_n\}$ has a limit. (T)
Converse: if the increasing sequence $\{a_n\}$ has a limit, it is bounded. (T)

(e) L, M are two lines, N a third line intersecting them.
L, M parallel $\Rightarrow$ angle LN = angle MN. (T)
Converse: If two lines make equal angles with a third line intersecting them, they are parallel. (T)

2. (a) $a > 0$ is stronger (b) $\{a_n\}$ bounded is stronger

3. (a) $\Rightarrow$ (b), (c) $\Rightarrow$ (b) are the stronger-weaker pairs.

A.2

1. (a) L and M intersect in one point. (Why specify "one point"?)

(b) Triangle ABC has sides of three different lengths.

(c) There are only a finite number of primes.

2. Converse: a^2 odd $\Rightarrow$ a odd. (T)

Contrapositive: a^2 even $\Rightarrow$ a even. (T)

3. (a) $a_1 \geq 0$ and $a_2 \geq 0 \Rightarrow a_1 a_2 \geq 0$; $a_1 < 0$ and $a_2 < 0 \Rightarrow a_1 a_2 > 0$.

(b) All $a_i < 1 \Rightarrow a_1 + a_2 + \ldots + a_n < n$.

A.3

1. All are false.

(a) 7 is a counterexample;

(b) any three lines meeting in a point are a counterexample;

(c) the pair $3, 5$ is a counterexample (the only one, actually).

A.4

1. We have $0! = 2^0$, since both sides are 1; thus $P(0)$ is true. This is not the basis step, however, since the proof of $P(n+1)$, assuming $P(n)$, only works when $n \geq 1$. So the basis integer n_0 must be at least 1 to get the induction going, and as we have seen, the basis step is actually $P(4)$.

2.
$$\begin{aligned} n+1 &< 2^n + 1, && \text{using } P(n); \\ &< 2^n + 2^n, && \text{if } n \geq 1; \\ &= 2^{n+1}. \end{aligned}$$

The basis step is $P(1)$, which is true; we cannot use $P(0)$ as the basis, since the above proof doesn't show $P(0) \Rightarrow P(1)$: it requires $n \geq 1$. So we have shown $P(n)$ is true for $n \geq 1$. Nonetheless, $P(0)$ is also true (by a separate argument), so $P(n)$ is true for all $n \geq 0$.

3. We prove by indirect argument that S is empty, i.e., that $P(n)$ is true for all $n \geq n_0$.

If S is non-empty, then it contains a smallest integer $m \geq n_0$, and $P(m)$ is false. Look at the number just before m:

$$\begin{aligned} m - 1 &\geq n_0, && \text{since } m \geq n_0, \text{ and } m \neq n_0 \text{ (for } P(n_0) \text{ is true)}; \\ m - 1 &\notin S, && \text{since } m \text{ is the smallest number in } S. \end{aligned}$$

Therefore $P(m-1)$ is true; but since $P(n) \Rightarrow P(n+1)$ for $n \geq n_0$, it follows that $P(m)$ is true, contradiction. □

(Note: The self-evident fact we used:

a non-empty set of positive integers has a smallest element

is known as the *well-ordering property* of $\mathbb{N}$.)

4. The proof is almost the same: the basis step is still $P(3)$, and the only difference is that the line segment must be drawn so it divides C_{n+1} into a triangle C_3 and a C_n. Then

$$\begin{aligned} \text{angle sum of } C_{n+1} &= \text{angle sum of } C_n + \text{angle sum of } C_3 \\ &= (n-2)\pi + \pi, && \text{using } P(n) \text{ and } P(3); \\ &= (n-1)\pi\ . \end{aligned}$$

□

.

Students who think of the induction step as $P(n) \Rightarrow P(n+1)$ (rather than "prove $P(n+1)$ by using $P(n)$") will often start by drawing a C_n, then adding a vertex to make it a C_{n+1}. This is illegal — if you're proving $P(n+1)$, you must start with whatever C_{n+1} is given to you — you can't make up your own! In many areas of discrete mathematics, graph theory for instance, the $P(n) \Rightarrow P(n+1)$ approach to induction proofs produces immediate disaster. Remember: start with $P(n+1)$, if that's what you are supposed to prove.

5. The statements $P(p)$ for p prime are the basis step, though this is concealed the way the proof is given; it is worded so that the basis step is included in the proof of $P(n)$.