

Chapter 4

Mathematical Proofs

4.1 Different Types of Proofs

Now that we have looked at the basics of propositional and quantifier logic, we are ready to study mathematical proofs in depth. Recall, from Section 1.2, that the primary means for establishing new results in mathematics is called the **axiomatic method**. Certain statements are taken as axioms; this means that they are accepted without proof. A statement may be included as an axiom because it is obviously true and quite simple or (in more specialized studies) because one or more mathematicians think it may have interesting consequences.

In this book we do not use any specialized axioms. Rather, we confine ourselves to standard, “obviously true” ones, that is, axioms almost all mathematicians and you would agree are correct. Axioms fall into two categories. Those based on logic are called **logical axioms**. Axioms based on properties of a particular type of mathematical object (integers, real numbers, sets, and so on) rather than on logic are called **proper axioms**; in Euclid’s day they were called **postulates**.

Recall also that logical deduction is the only acceptable way to prove new things in mathematics. Traditionally, mathematicians have not spelled out exactly what types of reasoning should be allowed in proofs. It was generally assumed that everyone doing mathematics would have a feel for this that would be consistent with everyone else’s. But in the past only a very small percentage of the population came into contact with advanced mathematics. In today’s technological society, it’s considered valuable to make mathematics comprehensible to a wider audience than ever before. Furthermore, for various kinds of theoretical studies, including those involving the use of computers to do mathematics, it’s important to have an exact definition of what constitutes a proof.

Accordingly, this book not only presents a set of axioms but also explains numerous rules of inference that are commonly used in mathematics. A **rule of inference** is a precise rule that describes how a new statement may be asserted in a proof on the basis of its relationship to previous statements in the proof. Usually, when mathematicians refer to a proof method, they mean a rule of inference.

Example 1: Here is a rule of inference that we call the **conjunction rule**: if two statements P and Q appear as separate steps in a proof, then it’s allowable to conclude

the single statement $P \wedge Q$ in the proof. This is a particularly simple and obvious rule of inference, but many of the ones we introduce are not much more complex than this.

Formal Proofs

Definition: An **axiom system** consists of two parts: a list of statements that are to be considered axioms, and a list of rules of inference.

The lists mentioned in this definition may be finite or infinite. But in either case, the axioms and rules of inference must be clearly and unambiguously defined, so that it's always possible to determine whether any given statement is an axiom or follows from certain other statements by a rule of inference.

In the next three definitions, we assume that we have a particular axiom system in mind.

Definition: A **formal proof** is a finite sequence of statements in which every statement (or **step**) is either (1) an axiom, (2) a previously proven statement, (3) a definition, or (4) the result of applying a rule of inference to previous steps in the proof.

Definition: A **theorem** is a statement that can be formally proved. That is, it's a statement for which there's a formal proof whose last step is that statement.

Remarks: (1) There are several other words with more or less the same meaning. A relatively simple theorem may be called a **proposition**. (This usage of this word is clearly quite different from the meaning we defined for it in Chapter 2.) A theorem that is not considered very important on its own but is useful for proving a more important result is usually called a **lemma**. And a theorem that is easily proved from another theorem is usually called a **corollary** to the other theorem. There are no hard-and-fast rules for which of these words to apply to a given result. Some important results in mathematics have been labeled propositions or lemmas, perhaps because their authors were on the modest side.

(2) Sometimes it is appropriate to begin a proof with one or more **assumptions** (also called **hypotheses** or **premises** or **givens**. This usage of the word "hypotheses" is very different from its usage in the sciences, as described in Chapter 1). It is important to understand the distinction between an axiom and an assumption. An axiom is a statement that is agreed on and available for use in proofs permanently, at least within a particular subject. An assumption is a statement that is available for use only in the proof being attempted. Assumptions were made in several of the proof previews in Chapters 2 and 3, in which the goal was to prove an implication. In this chapter, we see that there are only a handful of situations in which it's legitimate to make assumptions in proofs. In fact, proving implications is essentially the *only* situation that justifies assumptions in mathematical proofs. It is *extremely important* to learn when it is appropriate to use the word "assume" and when it is not.

Example 2: In Section 2.3, whenever we verified that an argument was valid, we were essentially doing proofs from premises in an axiom system. This axiom system is very simple, having just one rule of inference: you may assert any statement that is a propositional consequence of the previous statements in the proof. Technically, this axiom system has no axioms, but practically speaking all tautologies are axioms, since every tautology is a propositional consequence of *any* set of statements.

Remarks: (1) Our definition of a formal proof allows four types of steps. Two of those—previously proven statements and definitions—are never needed to prove anything. Quoting a previously proven theorem in a proof just saves the trouble of reproving it, so while it can be a substantial time-saver, it never allows you to prove anything that you couldn't prove without quoting it. Unless specifically disallowed (which could occur on test or homework problems), it's acceptable to save time in this way when doing proofs in mathematics.

The role of definitions is more subtle, but a definition just introduces a shorthand or abbreviated way of saying something. So definitions also save time and can be very enlightening, but they don't allow you to prove anything you couldn't prove without them. You can use two sorts of definitions in proofs. You can make your own definitions, which create abbreviations (temporary or permanent) for your own convenience. But you can also quote any definition that has been given (for example, in whatever text you are using), as if it were an axiom. So if we wanted to give a bare-bones definition of what a formal proof is, we could have limited the possibilities to parts 1 and 4 of the definition.

(2) It is important to understand the difference between an axiom and a rule of inference. An axiom is a *single* statement that we agree to accept without proof and therefore may be asserted at any step in a proof. A rule of inference is never a single statement; rather, it describes some procedure for going from old statements to new ones. However, in the less formal proofs that mathematicians normally write, this distinction often gets blurred, as we soon see.

Except for very specialized and unusual systems, rules of inference are always based on logic. The ones in this book are all based on logic and, like our axioms, are quite standard.

A General-Purpose Axiom System for Mathematics

Appendix 1 consists of a detailed axiom system that we refer to throughout this book. The first part of it (all the rules of inference and groups I, II, and III of axioms) is based on logic and is **logically complete** and **sound**, meaning that its power to prove statements corresponds exactly to logical consequence. The rest of the axiom system consists of the generally accepted axioms about sets, the real numbers, and the natural numbers. Taken as a whole, our axiom system is powerful enough to derive all currently accepted theorems of mathematics, even in the most advanced subjects.

You might be surprised that it's possible to write an axiom system that encompasses all of mathematics but that is only three or four pages long. In fact, the

axioms for logic and set theory alone are sufficient for the development of all of mathematics. So if I wanted to be very economical, I could have omitted the entire sections of real number axioms and natural number axioms, and Appendix 1 would still encompass all of mathematics. These extra axioms have been included to make the system easier to use, since it is rather difficult to develop the theory of these number systems from logic and set theory alone.

You might also wonder whether this system is at all standard or if it's just a personal creation of one author. As an experiment, you might ask your favorite mathematics professor what axioms she uses in her work. Her response will probably be a smile or a puzzled expression, with words to the effect, "I don't use axioms when I do mathematics. I just use intuition and deductive reasoning, plus a few well-known principles." This type of response is probably honest, but don't let it mislead you. By the time someone becomes a professional mathematician, she has had so much experience with the usual axioms and rules of inference of mathematics that they have become second nature to her. Even if she has never had a course like this one, she is just about as comfortable using the principles discussed in this book as most people are driving a car or reciting the alphabet. Furthermore, just as these basic skills become unconscious for all of us, an experienced mathematician may not even be aware that she uses a particular set of principles that has been subtly taught to her and is quite universal. But if you were to go beyond that first response and ask her to think in detail about how she does proofs, it would almost certainly turn out that she uses a combination of principles that are exactly equivalent to the list presented here.

Informal Proofs

One major reason why this book discusses formal proofs is to help you understand that there is nothing mysterious or magical about what constitutes a proof in mathematics. It may take a great deal of ingenuity to *find* a formal proof of a given statement, but once one is produced and written down, there would normally not be a controversy about whether it's correct. Any reasonable person who is willing to be very careful and take enough time ought to be able to check a formal proof for correctness. Better yet, computer programs can be written to check them. When mathematics is done formally, it becomes a sort of a game, with exact rules like chess or tic-tac-toe. However, although the rules of a game like chess are arbitrary and so must be learned specifically, the rules of mathematics are directly based on logic and common sense, so that it should be a natural process to become fluent with them.

Now here comes the catch. In spite of the order and precision that could be brought to mathematics by sticking to formal proofs, this type of proof is almost *never* used by mathematicians. If you randomly went through a dozen mathematics books, you would probably not find a single formal proof. We'll soon see some formal proofs, and you'll easily see why they are avoided. They are often extremely long and tedious to write and even worse to read. A complete formal proof usually consists of pages of symbols, even if it is based on just one or two simple ideas.

So if mathematicians don't write formal proofs, what do they write? Naturally, we may say that they write **informal proofs**. Unfortunately, it's not possible to say exactly

what is meant by an informal proof. Furthermore, it's inaccurate to think of formal proofs and informal proofs as two, clearly separate categories. The true situation is more like a whole spectrum. On one extreme are strictly formal proofs. On the other extreme are completely informal proofs. Informal proofs are not based on any specific axiom system, generally use English more than mathematical notation, skip and/or lump together many steps, and base most of their logical assertions on commonsense reasoning. They often are laced with words like "obviously," "clearly," "it is easily seen that ...," and in the case of one well-known mathematician, "it is intuitively obvious to the most casual observer that"

Informal proofs are much easier to write and read than formal ones, and a well-written informal proof conveys information better than any formal one can. The problem with completely informal proofs, especially when used by less experienced proof writers, is that they open the door to sloppy thinking and errors.


To make all this clearer, here is a list of some of the ways in which most of the proofs written by mathematicians do not fit the definition of a formal proof:

(1) *Use of English:* Normally, when an axiom system is precisely defined, the axioms and rules of inference are stated in symbols (that is, mathematical and logical notation). It would then follow that a formal proof in that system would consist of symbolic statements, not English ones. But most mathematics proofs flow better if there are words as well as symbols, and so most mathematicians use a liberal mixture of words and symbols in their proofs. As long as these words strictly correspond to the axioms and valid logical principles, the use of English in a proof does not make the proof informal. But often, words are used to gloss over gaps in a proof, and in that case the proof must be considered informal.

(2) *Lack of an Axiom System:* We have already mentioned that most mathematicians do not consciously have an axiom system in mind when they write proofs; but unconsciously, almost all of them do follow a system that is equivalent to the one given in Appendix 1. However, a mathematician may occasionally write a proof that is not based, even unconsciously, on a clear-cut list of axioms. Such a proof would have to be called informal. On the other hand, many mathematicians would say that a nonaxiomatic proof cannot be a correct mathematical proof.

(3) *Skipping Steps:* Almost all mathematicians simply skip whatever steps they deem to be obvious. This is acceptable if the skipped steps really *are* obvious to whomever reads the proof. But this gets tricky: when you write a proof, how do you know who will be reading it and what will be obvious to whom? Something that would be obvious to most professional mathematicians would not necessarily be obvious to others. In practice, proofs are written differently for different audiences.

How should you handle this subtle point? Under what circumstances and to what extent should you leave out obvious steps? There is no pat answer to this question. Your instructor and the remainder of this text will constantly try to give you a feel for this. In the meantime, here is a good rule of thumb:

 *Do not omit any steps in a proof unless you can see clearly how to fill in all the gaps completely.* Nothing gets you into trouble more surely than skipping steps and calling them obvious, without knowing precisely how to carry out all the omitted steps.

(4) *Combining Steps:* This is a variant of skipping steps. Mathematicians often lump several easy steps into a single sentence, glossing over them rather than leaving them out entirely. The same guidelines that were described for skipping steps also apply to this practice.

(5) *Reverse Proofs:* This is a fairly common practice among mathematicians and one that can confuse the inexperienced reader. Technically speaking, when you prove a statement in mathematics, the statement you are proving should be the last step in the proof. That is, you start with things you are given (axioms and/or assumptions) and try to get to what you want. But sometimes the easiest way to figure out how to prove something is to do it in reverse, starting with what you want to prove, then looking for some statement that implies what you want, then looking for some statement that implies that statement, and finally reaching a known statement. A correct reverse proof can always be turned into a formal proof by writing the steps in the standard, forward order. But often a mathematician decides that a proof reads better in reverse and so keeps the final version that way. There is no harm in this informality if done properly. But whenever you do a proof in reverse, make sure it works forward; otherwise, you're probably proving the converse of what you should be proving.

This technique is discussed further in Section 4.6 under the heading "Reverse Reasoning," and we will see many examples of this important idea.

Good Proofs

How formal or informal should your proofs be? There is no pat answer to this. The dangers of both extremes have already been pointed out. While you are learning to write proofs, it is probably better to play it safe by keeping your proofs relatively complete. As you start to gain confidence, you can start to write more informally and skip a bit more. You should feel free to ask for guidance from your instructor and other experienced mathematicians regarding these issues, since it's quite hard to learn good proof writing without frequent feedback.

Here's another rule of thumb: *A good proof should be a clearly written outline or summary of a formal proof.* This means that when you write a proof, each statement you write (especially if it's an English sentence) should describe or indicate one or more steps that you would include in a formal proof if you were writing a formal proof. You can't do this unless you see how the formal proof should go. Once you've done that, you need to outline the formal proof in such a way that any reasonably intelligent reader, *including yourself*, should be able to understand the outline well enough to reconstruct the formal proof from it. This outlining process can require considerable thought and is what is meant by *style* in mathematics.

We have already encountered two very different styles of proofs. In the “Alternate Solutions” to Example 1 of Section 2.3, formal proofs were given for parts (a) and (c). For contrast, a less formal proof of part (c) was also included. All the proof previews in Chapters 2 and 3 are also written in good, nonformal style, as are the great majority of proofs from this point on. Occasionally we give a formal proof in the text, and there are several of them in Appendix 2. When you encounter these formal proofs, you should find it easy to understand their main advantage (that they encourage clear, correct step-by-step thinking) and their main disadvantage (that they are unwieldy, both to write and to read).

We occasionally use the term **semiformal** to describe a proof that directly parallels and summarizes a particular formal proof.

4.2 The Use of Propositional Logic in Proofs

Example 2 of Section 4.1 explained that the notion of propositional consequence introduced in Section 2.3 provides the basis for an important rule of inference. In fact, this one rule of inference is completely general with respect to propositional logic, in the sense that it includes every valid proof method based on propositional logic. Therefore, we discuss this rule of inference first and view it as the basis for all the material in this section.

Rule of Inference: Propositional Consequence

In a proof, you may assert any statement that is a propositional consequence of previous steps in the proof.

We sometimes shorten “propositional consequence” to “**prop. cons.**” or simply “**PC.**” Equivalently, we often just say that a step in a proof follows from previous steps “by propositional logic.”

Example 1: Suppose we are talking about a real number x . We know (from axiom V-15) that either $x > 0$, $x = 0$, or $x < 0$. Suppose we also know, somehow, that x is nonzero. Then PC allows us to conclude that $x > 0$ or $x < 0$. This use of PC is based on the tautology $[(P \vee Q \vee R) \wedge \sim Q] \rightarrow (P \vee R)$. This exact tautology does not appear in Appendix 3, but it is quite close to tautology 11. At any rate, you should try to reach a point where you don’t need to refer to Appendix 3 very often to check conclusions of this sort, because your own feel for logic makes it unnecessary.

Example 2: It is a theorem of calculus that if a function is differentiable, it is continuous. Suppose that we know this result, and we want to assert its contrapositive in a proof; that is, if a function is not continuous, then it is not differentiable. The rule PC allows us to do this, using the simple tautology $(P \rightarrow Q) \leftrightarrow (\sim Q \rightarrow \sim P)$.

For some more substantial examples of proofs based on propositional consequence, refer back to the examples and proof previews in Section 2.3.

As we mentioned in the previous section, having the rule of inference PC essentially makes all tautologies axioms. We now make this explicit.

All tautologies are axioms.

Note that a tautology is, by definition, always true and it's also straightforward to determine whether a given statement is a tautology. These are ideal characteristics of axioms.

Example 3: Suppose that we are trying to prove something about a real number x . If we want to, we can assert the statement that either $x = 0$ or $x \neq 0$, since this is of the form $P \vee \sim P$ (tautology 1, the law of the excluded middle). It might seem pointless to assert this disjunction, but this step might be used to set up a proof by cases, which could substantially simplify the proof.

On the basis of what we said at the beginning of this section—that propositional consequence includes all valid proof methods based on propositional logic—we could technically end this section at this point. However, PC is too general to be very convenient in most situations. Instead, mathematicians commonly use at least a half dozen more specific rules of inference. So let's now examine some of the most important of these so-called **derived rules of inference**.

Rule of Inference: Modus Ponens

If you have a step P and another step of the form $P \rightarrow Q$, you may then conclude the statement Q .

This rule of inference can be diagrammed (in the style of Section 2.3) as follows:

$$\begin{array}{l}
 P \\
 P \rightarrow Q \\
 \hline
 \therefore Q
 \end{array}$$

Despite the Latin name (which means “method of affirming”), this is a very simple rule of inference. Hopefully, you can see that modus ponens is logically correct: if you know P and also that P implies Q , then Q must follow. This is more or less the definition of implication.

A bit more formally, it is tautology 9 that justifies modus ponens. Exercise 8 asks you to derive modus ponens from propositional consequence.

Example 4: One important theorem of calculus is that if a function is differentiable, it must be continuous. Another basic result is the derivative formula for polynomials, which guarantees that polynomial functions are differentiable. Applying modus ponens to these steps yields that any given polynomial function, such as $3x^2 - 6x + 2$, must be continuous.

Starting with the next example, we occasionally illustrate a method of proof by referring to a proof of a theorem in Appendix 2 at the end of this book. Even though we have not discussed Appendix 2, you need not be intimidated by these references, because the proofs in Appendix 2 are based on the standard properties of the real number system, which are quite familiar to you from high school algebra.

Example 5: Every proof in Appendix 2 contains uses of modus ponens. A typical instance, although it’s not specifically mentioned, occurs in the proof of Theorem A-7. In that proof, we have the step $z \neq 0$, since it’s an assumption in the proof. We also have, from axiom V-11, the implication $z \neq 0 \rightarrow zz^{-1} = 1$. Modus ponens applied to these two statements yields the step $zz^{-1} = 1$.

Example 6: Let’s redo Example 1(a) of Section 2.3, using a formal proof from hypotheses, with tautologies as our only axioms and modus ponens as our only rule of inference:

(1)	$P \rightarrow Q$	Premise
(2)	$\sim R \rightarrow \sim Q$	Premise
(3)	$\sim R$	Premise
(4)	$\sim Q$	Modus ponens applied to steps 3 and 2
(5)	$(P \rightarrow Q) \rightarrow (\sim Q \rightarrow \sim P)$	Tautology
(6)	$\sim Q \rightarrow \sim P$	Modus ponens applied to steps 1 and 5
(7)	$\sim P$	Modus ponens applied to steps 4 and 6

Rule of Inference: Conditional Proof

If you can produce a proof of Q from the assumption P , you may conclude the single statement $P \rightarrow Q$ (without considering P an assumption of the proof!).

We have been using conditional proof, without calling it that, ever since our first proof preview in Chapter 2. As we mentioned then, this is by far the most common and natural way to prove implications: to prove any statement of the form $P \rightarrow Q$, start by assuming P and try to prove Q . If you succeed, you've also succeeded in proving the conditional.

Even though conditional proof is derivable from propositional consequence, it is so important that we have included it separately in our axiom system. Conditional proofs are often referred to as **direct proofs** of implications. The diagram for conditional proof looks like this:

Assume P

[Some correct intermediate steps]

Q

$\therefore P \rightarrow Q$

Remarks: (1) Technically speaking, conditional proof is not a rule of inference as we have defined it. A rule of inference is supposed to say that if you have certain *steps* in a proof, you can conclude some statement. Conditional proof says that if you can produce a certain *proof*, you can conclude some statement. This may sound like a minor distinction, but you should be aware that conditional proof has a very different flavor from normal rules of inference.

(2) Many students initially confuse modus ponens with conditional proof. The best way to keep them apart in your mind is to realize that they are essentially *opposites*. Modus ponens gives you a way of *using* an implication: it says that if you *know* $P \rightarrow Q$, then you can go from P to Q . Conditional proof gives you a way of *proving* an implication: it says if you can show how to go from P to Q , you can *conclude* $P \rightarrow Q$.

(3) When writing a conditional proof, you need to be clear about what assumption(s) are in effect at each point in your proof. If you decide, in the middle of a proof, to prove an implication by conditional proof, you have to do a certain proof from an assumption, which may be regarded as a subproof of the main proof. As long as you're within the subproof, there's an assumption being made, but when the subproof

is finished, you assert the implication and go back to the main proof without the assumption. This is not particularly confusing as long as there's only one use of conditional proof at a time in the proof, but it gets more involved if there are nested uses of conditional proof, which can occur.

Example 7: Here is an example of nested uses of conditional proof. Suppose we want to prove that if a certain statement P holds, then some set A is a subset of some set B . So we would begin by assuming P , since P is the hypothesis of the implication we are trying to prove. From this assumption, we then need to prove that $A \subseteq B$. But to prove this, we need to show that any member of A must also be a member of B . Therefore, the second line of our proof would probably be "Assume x is any member of A ." This begins another conditional proof within the outer one. The goal of the whole proof is then to prove that x is a member of B , using both assumptions.

Rule of Inference: Indirect Proof

If you can produce a proof of *any* contradiction from the assumption
 $\sim P$, you may conclude P .

The diagram for this rule of inference is

Assume $\sim P$
 [Some correct intermediate steps]
 Any contradiction

$\therefore P$

Indirect proof is sometimes called **proof by contradiction** or **reductio ad absurdum** ("reduction to the impossible").

Remarks: (1) Section 4.1 cautioned you against overusing the word "assume" in proofs. Conditional proof and indirect proof are among the few proof methods in mathematics in which it's appropriate to assume something. We see below that indirect proof is really an offshoot of conditional proof, and it can be argued that the *only* situations in mathematics in which assumptions are permitted are those using conditional proof and its offshoots.

(2) You can try to prove any statement by indirect proof, no matter what its logical form is. But indirect proof is a particularly good proof method to try when the statement you're attempting to prove is a *negation*. If you want to prove $\sim P$ by indirect proof, you assume P and try to derive a contradiction.

(3) The most common form of contradiction is a statement of the form $Q \wedge \sim Q$, but any contradiction will do. The contradiction derived need not involve the original statement P . On the other hand, you may find that from the assumption of $\sim P$, you are able to prove P . This constitutes a successful indirect proof of P , because the next step can be the contradiction $P \wedge \sim P$.

Example 8: Theorem A-8 in Appendix 2 is a typical example of an indirect proof. We want to prove the statement $\sim (x < y \wedge y < x)$, which is in the form of a negation. This is the ideal type of statement to prove indirectly. So we assume $x < y \wedge y < x$. Using the transitivity of $<$ (axiom V-14) plus modus ponens, this yields $x < x$, which produces a contradiction in conjunction with the axiom $x \nless x$ (V-13).

Example 9: Indirect proof is the most efficient way to prove that the sum of a rational number and an irrational number must be irrational. (Recall that a rational number is one that can be written as a fraction of integers.) Here is a proof:

Assume the claim is false. Then we have $a + b = c$, for some numbers a , b , and c , with a and c rational and b irrational. Simply subtract a from both sides to obtain $b = c - a$. Since the difference of two fractions can always be written as a single fraction, this makes b rational, a contradiction.

We have been emphasizing that almost all the rules of inference in this section are derivable from propositional consequence, without going into much detail. Here is a proof of a similar fact.

Theorem 4.1: The inference rule of indirect proof is derivable from the rules in our axiom system, namely propositional consequence and conditional proof. That is, anything that can be proved using indirect proof (and perhaps the other two rules as well) can be proved with just those other two rules.

Proof: Suppose we have proved statement P by indirect proof. That means we have a proof that begins with the step "Assume $\sim P$ " and, after a certain number of correct steps, reaches a contradiction. Then simply take this proof and add two more steps:

Therefore, $\sim P \rightarrow$ some contradiction (conditional proof).

Therefore, P (from the previous step, by propositional logic). ■

Again, it doesn't matter what form of contradiction is obtained in Theorem 4.1; the statements $\sim P \rightarrow$ some contradiction and P must be propositionally equivalent. One way of interpreting this theorem is that indirect proof is a special type or offshoot of conditional proof.

Rule of Inference: Proof by Cases

If you have a step of the form $Q \vee R$ and the two implications $Q \rightarrow P$ and $R \rightarrow P$, you may conclude the statement P .

In practice, the usual format of a proof by cases is as follows: first you establish a disjunction that you think divides the problem up into cases that are easier to handle than the whole problem at once. Often this disjunction is very simple, perhaps a tautology or other axiom, and is not explicitly stated. Then you must show that each disjunct, or case, implies the statement that is to be proved; this is normally done by conditional proof. To keep the reader clear about what's going on, you might say "Case 1: Assume Q " and derive P , and then say "Case 2: Assume R " and derive P from that assumption. This rule of inference can be diagrammed as follows:

$Q \vee R$	(Proved somehow)
Case 1: Assume Q	
[Some correct intermediate steps]	
P	(End of Case 1)
Case 2: Assume R	
[Some correct intermediate steps]	
P	(End of Case 2)

$\therefore P$

Specifically, this is the diagram for proof by cases with *two* cases. It's fine to have more than two cases. For example, if you want to use a disjunction $Q \vee R \vee S$ to prove P by cases, you have to show that each of the three cases (Q , R , and S) implies P .

Derivation of proof by cases: Proof by cases is a special case of the rule PC, because P is a propositional consequence of the three statements $Q \vee R$, $Q \rightarrow P$, and $R \rightarrow P$. To see this without constructing a truth table, notice that the two implications $Q \rightarrow P$ and $R \rightarrow P$ are together equivalent to the single implication $(Q \vee R) \rightarrow P$, by tautology 26. Using the last implication, we can assert P by modus ponens. A similar derivation can be used to justify proofs by three or more cases (see Exercise 6).

A proof by cases using tautology 26 and propositional consequence was done in Proof Preview 3 in Section 2.3 (see Exercise 7). Proofs by cases in which the cases involve the sign of a real number are extremely common. Here is another example.

Example 10: Theorem A-11 of Appendix 2 shows a very common type of proof by cases. The goal is to prove the inequality $x^2 \geq 0$, for an arbitrary real number x . It is not easy to do this all at once. But if we look separately at the three cases $x = 0$, $x > 0$, and $x < 0$, it is pretty easy to show the desired conclusion holds in each case. (The disjunction of the three cases, which is a necessary part of the proof, is based on axiom V-15.)

Example 11: Let's consider a real-life example. Imagine that your girlfriend has told you that she doesn't want to see you tonight because she needs to stay home and study all evening. You want to believe but you're suspicious, so you're tempted to phone. Suppose you reason: "Well, either she's home or she's not. If she is, I'll be better off if I call because I'll be reassured. If she's not, I'll also be better off if I call because at least I'll know the score. So I should call." You are using proof by cases. The disjunction being used to define the two cases is an instance of the law of the excluded middle, $Q \vee \sim Q$, with Q representing "She's home."

Rule of Inference: Biconditional Rule

If you have implications $P \rightarrow Q$ and $Q \rightarrow P$, you may conclude the biconditional $P \leftrightarrow Q$.

The biconditional rule can be diagrammed

$$\begin{array}{l} P \rightarrow Q \\ Q \rightarrow P \\ \hline \therefore P \leftrightarrow Q \end{array}$$

In practice, this is by far the most common way to prove a biconditional, as we did in Proof Preview 1 in Section 2.2. A proof by this rule has two separate parts, or directions, which should *not* be called "cases" since that term is best reserved for proofs by cases. Each direction is an implication, which of course can be proved by conditional proof.

Some mathematicians use arrows to indicate the directions in this type of proof. So a proof of some statement $P \leftrightarrow Q$ could begin " \rightarrow : Assume P " and derive Q . It would then say " \leftarrow : Assume Q " and derive P .

Example 12: Theorem A-13 of Appendix 2 illustrates the typical use of this rule. We want to prove a biconditional: a number is positive iff its reciprocal is positive. So for the forward direction, we must prove $x^{-1} > 0$ from the assumption $x > 0$; for the reverse direction, we must prove $x > 0$ from the assumption $x^{-1} > 0$.

In simple proofs of this type, it may be obvious that the two directions of the proof are exactly the reverse of each other, step by step. In this situation, it's considered fine to show just one direction and to mention that the reverse direction can be proved by exactly the reverse sequence of steps. What this means is that you've established a sequence of iffs between the two statements whose biconditional you are trying to prove. We use this shortcut frequently, notably in Section 5.2.

Before we state our next rule of inference, we need some explanation and a bit of notation. Suppose that in a proof we have a step of the form $P \leftrightarrow Q$. This says that P and Q are equivalent, which ought to mean that P and Q are *interchangeable*. So suppose we also have some long statement that contains P as a substatement, such as one of the form $(R \rightarrow (\sim S \wedge P)) \vee (P \rightarrow T)$. Should we be allowed to conclude the same statement with one or both of the occurrences of P replaced by Q ? Intuitively, we should be. Can this conclusion be justified by propositional consequence? Yes, it certainly can. But note that to do so would require a 32-line truth table! The next rule of inference lets you draw such conclusions without having to construct a truth table.

Notation: The notation $S[P]$ denotes a statement S that contains some statement P as a substatement (which could be the whole statement S).

The notation $S[P/Q]$ denotes a statement that results from $S[P]$ by replacing *one or more* of the occurrences of the statement P within $S[P]$ by the statement Q .

Rule of Inference: Substitution

From statements $P \leftrightarrow Q$ and $S[P]$, you may conclude $S[P/Q]$ as long as no free variable of P or Q becomes quantified in $S[P]$ or $S[P/Q]$.

The diagram for substitution is

$$\begin{array}{c} P \leftrightarrow Q \\ S[P] \\ \hline \therefore S[P/Q] \end{array}$$

Remarks: (1) The restriction on quantifiers in the use of substitution is meant to guarantee that $S[P]$ is built up from P using connectives only, not quantifiers (or at least not quantifiers that matter). In practice, it is not necessary to worry about this restriction very often. For a specific example, see Exercise 18 of Section 4.3.

(2) There is some deliberate ambiguity in the notation introduced here. Even when S , P , and Q are known, there may be more than one possibility for what statement $S[P/Q]$ represents, for if P occurs within S more than once, then there's a choice as to which occurrence(s) of P within S are to be replaced by Q . The example in the paragraph before the definition of substitution illustrates this.

(3) Why is the notation $S[P]$, rather than $S(P)$, being used here? The reason is that P is *not* a mathematical variable appearing in the statement S . It's a substatement of S , which is totally different. To emphasize this, it makes sense to introduce a different-looking notation.

(4) Do not confuse substitution with the familiar principle of substitution of equals, which says that if two numbers or other mathematical quantities are equal, then they are interchangeable. This principle appears in our axiom system as axiom III-4. In flavor, it is certainly very similar to substitution. The main difference between them is what they talk about: Substitution is about the interchangeability of *statements*, whereas Axiom III-4 is about the interchangeability of *numbers* or other mathematical *objects*.

As we show in Section 4.4, the principle of substitution of equals is what allows us to do the same thing to both sides of an equation. Similarly, with the rule of inference substitution, it becomes permissible in most situations to do the same thing to both sides of an equivalence.

Example 13: Here is a simple example of substitution from real life. Suppose you say, "If Harry shows up at my party, I'll call the police." Then your friend says, "But Harry and your boss do everything together; if one shows up, so will the other." Then you say, "Well, I guess that means that if my boss shows up, I'll call the police." You have just used substitution, because the second part of your friend's statement means, "Harry will show up if and only if your boss does."

Example 14: Suppose we know some biconditional $P \leftrightarrow Q$. Then, if we also know P , we can use substitution to conclude Q . On the other hand, if we know $\sim P$, we can conclude $\sim Q$. If we know a conjunction $P \wedge R$, we can conclude $Q \wedge R$. If we know $P \vee R$, $P \rightarrow R$, $R \rightarrow P$, or $P \leftrightarrow R$, we can conclude, respectively, $Q \vee R$, $Q \rightarrow R$, $R \rightarrow Q$, or $Q \leftrightarrow R$. Of course, these conclusions also follow by propositional logic (see Exercise 1).

We have now discussed all the important propositional rules of inference. Here are a few more rules, all of which are pretty obvious and follow easily from the ones we have listed so far. We won't give derivations or examples of these because of their simplicity.

Rule of Inference: Conjunction

If you have, as separate steps, any two statements P and Q , you may conclude the single statement $P \wedge Q$.

This rule of inference follows trivially from propositional consequence. See Exercise 9(a) for a related and not-so-easy problem.


Rule of Inference: Modus Tollens

If you have a step of the form $P \rightarrow Q$ and also have $\sim Q$, you may assert $\sim P$.

“Modus tollens” is Latin for “method of denying.” It is based on tautology 10, but it can also be thought of as an offshoot of modus ponens, based on the equivalence of $P \rightarrow Q$ and its contrapositive $\sim Q \rightarrow \sim P$. Similarly, the next rule is a straightforward contrapositive offshoot of conditional proof.

Rule of Inference: Contrapositive Conditional Proof

If you can produce a proof of $\sim P$ from the assumption $\sim Q$, you may conclude the single statement $P \rightarrow Q$.

 We have now discussed two ways to prove an implication $P \rightarrow Q$, but there are at least three common ways:

- (1) Direct conditional proof: Assume P , and derive Q .
- (2) Contrapositive conditional proof: Assume $\sim Q$, and derive $\sim P$.
- (3) Indirect proof: Assume $\sim (P \rightarrow Q)$, or equivalently assume $P \wedge \sim Q$, and derive a contradiction.

Table 4.1 Summary of propositional proof methods

Statement	Way(s) to Prove	Way(s) to Use
A negation $\sim P$	Indirectly: Assume P , and derive a contradiction.	Move negation sign inward.
A conjunction $P \wedge Q$	Conjunction rule: Prove P and also prove Q .	Assert P and/or Q , separately.
A disjunction $P \vee Q$	(1) Prove P or prove Q . (2) Indirectly, by De Morgan's laws: Assume $\sim P \wedge \sim Q$, and derive a contradiction.	Proof by cases.
A conditional $P \rightarrow Q$	Three methods, listed on Page 88.	(1) Modus ponens. (2) Modus tollens.
A biconditional $P \leftrightarrow Q$	Biconditional rule: Prove $P \rightarrow Q$ and prove $Q \rightarrow P$.	Substitution.

In a sense, the last method is the most powerful, because it lets you begin with two assumptions instead of one.

We conclude this section with a chart summarizing the most important propositional proof methods (see Table 4.1). This chart shows the most natural ways to *prove* and to *use* each type of statement (with type based on the connective). Table 4.1 is meant to help you when doing proofs, and you are urged to study it carefully. At the same time, don't fall into the trap of thinking that any single chart can teach you everything there is to know about proofs in mathematics or even everything about the proof methods based on propositional logic. There may be only a finite number of rules of inference in common use, but there are literally an infinite number of different ways to apply them.

Exercises 4.2

(1) Show that each of the conclusions in the first paragraph of Example 14 could have been made using propositional consequence instead of substitution.

(2) Suppose that we have proved steps of the form $P \leftrightarrow Q$ and $(P \wedge R) \vee (\sim P \wedge S)$ in a proof. State the three different conclusions that may be made from these two steps using substitution.

(3) Redo Exercise 5 of Section 2.3 as formal proofs from hypotheses. You may use all tautologies as axioms, and you may use all the rules of inference discussed in this section, *except* propositional consequence.

(4) For each of the following, state what seems to be the logical conclusion and also state which rule of inference (other than propositional consequence) could be used to reach that conclusion:

(a) To make you happy today, I'd have to be in two places at once, which is impossible. Therefore, ...

(b) If it's Saturday, I don't go to school. If I don't go to school, I'm very sad. It's Saturday. Therefore, ...

(c) If the Mets win, I'll come out ahead in my bets. If the Mets don't win, I'll also come out ahead in my bets. Therefore, ...

(d) If a function is continuous, it's integrable. This function is not integrable. Therefore, ...

(5) Turn each of the arguments in Exercise 4 into formal proofs. To do this, you should introduce propositional variables for the atomic substatements and rewrite each given statement symbolically, as in Section 2.1. Then write formal proofs to prove each conclusion from the givens.

(6) Precisely state the inference rule of proof by cases with *three* cases, and derive this rule from propositional consequence.

(7) Redo Proof Preview 3 (Section 2.3), explicitly using the method of proof by cases.

(8) Derive the rule modus ponens from propositional consequence.

*(9) In the text, we have shown that all tautologies and the modus ponens rule of inference are derivable from the propositional consequence rule. This exercise establishes the converse result that propositional consequence is derivable from all tautologies and modus ponens. Consider an axiom system that has modus ponens as a rule of inference instead of propositional consequence and has all tautologies as axioms.

(a) Show that the conjunction rule of inference is derivable from this axiom system. (You must *formally* prove $P \wedge Q$ from the two assumptions P and Q . This can be done in very few steps, but it's tricky to find the proof.)

(b) Using the result of part (a), whether or not you were able to show it, show that propositional consequence is derivable from this axiom system.

Exercises 10 through 12 ask you to prove various results. Since we have not yet discussed methods of proof involving quantifiers, do not attempt to make these proofs rigorous or formal. For the most part, you may assume familiar results from high school algebra about integers, equations and inequalities. However, where *real* numbers are involved, you should not assume anything beyond the axioms in group V of our axiom system, unless instructed to do so.

(10) Prove: If n is any integer, then $n^2 - 3n$ must be even.

(11) Prove: For any real numbers x , y , and z , if $x < y$ and $y \leq z$, then $x < z$.

(12) Prove: For any *positive* real number x , $x + 1/x \geq 2$. It might be tempting to use calculus to prove this, but you may not do so; use information from high school algebra only. **Hint:** You might try a *reverse proof*, as explained in the previous section. That is, start with the inequality you're trying to prove and try to transform it into something that you know to be true. But then you have to make sure you can turn this into a *forward* proof. You may assume, without proof, that the square of any real number is at least 0. By the way, is there any point in your proof where you need to know that x is positive?

The rest of the exercises in this section are of a type that occurs frequently throughout the rest of this book. These problems ask you to "critique the following proof," and then give a supposed theorem and proof. To do such an exercise, you should carefully read and consider the given proof and come to one of four conclusions:

- (i) The theorem and its proof are correct (and the proof has no major omissions).
- (ii) The theorem is correct; the proof has no mistakes but does omit one or more substantial step(s). In this situation, you should supply the missing step(s) in the proof.
- (iii) The theorem is true, but the proof is substantially flawed. In this situation, you should point out the error(s) in the given proof *and* provide a correct proof of the theorem.
- (iv) The stated theorem is false. In this situation, you should point out the error(s) in the given proof *and*, if appropriate, provide a concrete example to show that the claimed theorem is in fact false. (Such examples are called **counterexamples**; see Section 4.3.)

There may be problems of this type that you believe to be borderline between two of the four choices, for example, between (i) and (ii). If that occurs, feel free to say so.

Critique the proofs in Exercises 13 through 17.

(13) **Theorem:** If x is any real number, then $x^2 \geq x$.

Proof: Assume x is any real number. We proceed by cases:

Case 1: Assume $x \leq 0$. Then $x^2 \geq 0$, while $x \leq 0$. Therefore, $x^2 \geq x$.

Case 2: Assume $x \geq 1$. Then multiply both sides of this inequality by x to obtain $x^2 \geq x$. Since both possible cases lead to the desired conclusion, we have proved it.

(14) **Theorem:** For any real numbers x and y , $x < y$ implies $x^2 < y^2$.

Proof: We use conditional proof. So assume that $x < y$. Then simply square both sides of the inequality. This gives us the desired conclusion $x^2 < y^2$.

(15) **Theorem:** For any real number x , if $x^2 = 0$ then $x = 0$.

Proof: We use indirect proof. Assume the entire implication is false. By tautology 19, that means $x^2 = 0$ and $x \neq 0$. Then we can multiply both sides of the equation $x^2 = 0$ by $1/x$. The left side becomes $x^2(1/x)$, which equals x ; and the right side becomes $0(1/x)$, which equals 0. Thus $x = 0$. This contradicts our assumption that $x \neq 0$, and so we are done.

(16) **Theorem:** For any integer n , if n is even, so is n^2 .

Proof: We use indirect proof. Assume n is *not* even; from this we must prove that n^2 is not even either. To say that n is not even means that it is odd. So we have $n = 2m + 1$, for some integer m . Therefore, $n^2 = (2m + 1)^2 = 4m^2 + 4m + 1 = 2(2m^2 + 2m) + 1$, which by definition is an odd number.

*(17) **Theorem:** For any positive integer n , if $2^n - 1$ is prime, then so is n .

Proof: We use contrapositive conditional proof. That is, we begin by assuming that n is *not* prime.


Case 1: Assume $n = 1$. Then $2^n - 1 = 2^1 - 1 = 1$, which is not prime.

Case 2: Assume $n > 1$. Then n must be a composite number, so $n = ab$, where a and b are both integers greater than 1. Let $m = 2^a$. So $2^n - 1 = 2^{ab} - 1 = (2^a)^b - 1 = m^b - 1$. From algebra, $m^b - 1$ is divisible by $m - 1$. (Specifically, a simple computation shows that $m^b - 1 = (m - 1)(1 + m + m^2 + m^3 + \dots + m^{b-1})$.) Now, since $a > 1$, $2^a > 2$, and thus $m - 1 > 1$. Also, since $m > 1$ and $b > 1$, $m - 1$ is certainly less than $m^b - 1$. Therefore, $2^n - 1$ (which equals $m^b - 1$) has a factor strictly between itself and 1, and so it is not prime. This completes the proof.

4.3 The Use of Quantifiers in Proofs

Recall that we have emphasized propositional consequence as the most important rule of inference for propositional logic; in fact, we said it could be viewed as the *only* propositional rule of inference. Why don't we do the same thing for quantifier logic? That is, why not have a rule of inference that says you may assert any statement that is a *logical* consequence of previous steps in the proof? This question was answered in Section 3.3. There is no simple, mechanical procedure (like truth tables) for testing whether a given statement is a logical consequence of certain other statements. Therefore, the abstract concept of logical consequence cannot be used to define axioms or rules of inference.

Before proceeding to specific axioms and rules of inference, we need to state an important convention that is standard throughout all of mathematics.

 **Convention:** Whenever an *axiom* or *theorem* contains free mathematical variables, the statement is understood to begin with universal quantifiers for those variables. (This convention applies to axioms and theorems *only*. It does not apply to definitions or to steps within a proof.)

Axioms: De Morgan's Laws for Quantifiers

$$\sim \forall x P(x) \leftrightarrow \exists x \sim P(x)$$

$$\sim \exists x P(x) \leftrightarrow \forall x \sim P(x)$$

These important equivalences appeared earlier as Theorem 3.2, but that was in an informal setting. Now, we more rigorously classify them as axioms. As we will soon see (Theorem 4.2 and Exercise 15), either of them can be proved from the other, so only one of them *needs* to be an axiom.

Examples showing why these quantifier laws are useful and how to use them were also given in Chapter 3, so we won't give more examples here. Remember that it's useful to be able to move negation signs through quantifiers, especially from outside to inside, but such a move must be accompanied by changing every quantifier through which the negation sign is moved.


De Morgan's laws for quantifiers can be useful for indirect proofs of quantified statements. For instance, if you want to prove a statement of the form $\exists x P(x)$, you can assume $\forall x \sim P(x)$ and derive a contradiction. However, indirect methods are considered a bit less attractive than the more direct methods (UG and EG) to be discussed shortly, especially in the case of statements of the form $\exists x P(x)$.

Our next axiom is an obvious enough principle. If we know a statement is true for all members of a certain domain (real numbers, integers, functions, sets, and so on), then it's true for any particular one. This is a direct consequence of what the universal quantifier means.

Axiom: Universal Specification or US

$$\forall x P(x) \rightarrow P(t)$$

where the letter t here denotes any term or expression (*not* necessarily a single letter like a variable or constant) that represents an object in the domain of the variable x .

 Mathematicians rarely use the terminology “universal specification” and the similar names for the next three proof methods (as well as some of the terminology in Section 4.2). In fact, many fine mathematicians probably don't even know some of these terms. However, *all* mathematicians understand these proof methods and how to use them extremely well. The reason we introduce this terminology is to help you keep these proof methods straight while you are first learning about them. But it is vital that you concentrate on the *content* of these methods, not the words we use to classify them!

Remarks: (1) US allows plenty of freedom in the choice of the term t . It could be a variable, either the same one as in the quantifier or a different one. It could be a constant. Or it could be a more complicated expression.

(2) Technically, there is another restriction in the use of US: no free variable of t may become bound when $P(t)$ is formed. This situation occurs so rarely that we didn't mention it in the definition of US above (see Example 4).

(3) You may wonder why some principles are stated as axioms and others as rules of inference. In some cases the distinction is less important than in others. In particular, any axiom that is an implication can also be viewed (in combination with modus ponens) as a rule of inference. For example, US is primarily used as a rule of inference: if you have a step of the form $\forall x P(x)$, you may then assert $P(t)$, where t is any Usually, if a principle can be considered either an axiom or a rule of inference, we set it up as an axiom.

Example 1: You have been using US ever since you first studied algebra, even if you didn't have a name for it. For instance, consider a typical algebra formula such as $(x + y)^2 = x^2 + 2xy + y^2$. By the convention stated at the beginning of this section, the variables x and y in this formula are assumed to be universally quantified. So when you learned this formula in high school, you learned that it was true for all numbers, and that therefore you could substitute *any* expression for x and/or y . So you knew that you could write

$$(3a + 2)^2 = 9a^2 + 12a + 4$$

$$(x^2 + y^3)^2 = x^4 + 2x^2y^3 + y^6$$

$$(\sin x + \cos x)^2 = \sin^2 x + 2 \sin x \cos x + \cos^2 x, \text{ and so on.}$$

Every time you make this type of substitution for a variable, you are using US (plus modus ponens).

Example 2: Just about every proof in Appendix 2 uses US. A typical but clever use occurs in the proof of Theorem A-1. In the uniqueness part of that proof, we get to assume $\forall x (x + y = x)$ and $\forall x (x + z = x)$. To use these assumptions to prove that $y = z$, the trick is to specify the first quantified x to be z , and the second one to be y .

Example 3: Here is a simple instance of the necessity to have t be of the same sort as x in the use of US. Suppose we know the formula $\forall x (x^2 \geq 0)$, where x is a real variable. Then we cannot conclude that $i^2 \geq 0$, where i is the imaginary unit, since i is not a real number. In fact, $i^2 = -1$, so the conclusion would be false.

Example 4: Here is an example of why the other restriction on t (mentioned in Remark 2) is necessary. Suppose we have proved the true statement (about real numbers) that $\forall x \exists y (y > x)$. Without the restriction we could substitute the term " $y + 1$ " for x and derive the false conclusion $\exists y (y > y + 1)$.

Our next rule of inference is by far the most common way to prove a "For all ..." statement. You may well be familiar with it and know how to use it, even if you don't

know its name. Here's what it says: Suppose you want to prove a statement of the form $\forall x P(x)$. Then it is sufficient to prove the simpler statement $P(x)$, where x is just a variable representing an arbitrary member of the domain in question. It is important that *no assumptions* are made about x except that it is a member of that domain. If you can produce such a proof of $P(x)$, you may conclude $\forall x P(x)$.

Rule of Inference: Universal Generalization or UG

If you can produce a proof of $P(x)$, where x is a free variable representing an arbitrary member of a certain domain, you may then conclude $\forall x P(x)$.

Universal generalization can be diagrammed as follows:

[Proof with no assumptions about x]	
$P(x)$	
$\therefore \forall x P(x)$	

Remarks: (1) Although it is not required, it is helpful to let the reader know that you are intending to use UG by saying something like "Let x be any ..." or "Let x be an arbitrary ..." or "Let a ... x be given" at the beginning of the proof.

(2) People sometimes think that the logical way to prove that something is true for all members of some domain is to prove it for each member individually. That works fine if the domain happens to be finite and small but it's impossible if the domain is infinite. Unfortunately, most mathematical variables have infinite domains, such as the set of all real numbers.

(3) UG is used in the vast majority of theorems in mathematics, although it is almost never mentioned. By the convention stated at the beginning of this section, free variables in theorems are always understood to be universally quantified. In the proofs of such theorems, neither those quantifiers nor the use of UG are likely to be explicitly mentioned. The practical meaning of UG is that if you want to prove a statement that begins with one or more universal quantifiers, you can essentially ignore those quantifiers. This is a handy thing to know!

Example 5: Recall Proof Preview 1 in Section 2.2. There is nothing in the proof about quantifiers or UG. But the variable n is free, so the theorem must be about *all*

integers. The proof is carried out for an arbitrary integer n , and by UG, this establishes the result for all integers. Similarly, in Proof Preview 2 in Section 2.3, the set variables A , B , and C are free. So this theorem is technically about *all* sets A , B and C , and the proof may be viewed as including an implicit use of UG to quantify these variables. To be fully rigorous, Proof Preview 2 should also have dealt with the quantification of the variable x (see Exercise 1).

Example 6: Suppose we want to prove the important theorem that every nonnegative real number has a square root. We would probably begin the proof with the words “Let x be an arbitrary nonnegative real number.” This tells the reader that we intend to prove that this one unspecified x has a square root, which by UG establishes that every such number has one.

It may seem to you that the use of UG in Example 6 is pretty similar to conditional proof and that what’s really being proved is the implication “If x is a nonnegative real number, then x has a square root.” Conditional proof and UG are in fact related, and many mathematicians mix them and blur the distinction between them. They might start the above proof with the words “Assume $x \geq 0$,” as if it were a conditional proof.

Here is the source of the relationship between conditional proof and UG: If a mathematical variable x has some set D as its domain, then the statement $\forall x P(x)$ really means $\forall x \in D P(x)$. But in Section 3.3 we noted that $\forall x \in D P(x)$ is an abbreviation for $\forall x (x \in D \rightarrow P(x))$. In other words, just about every statement that begins with a universal quantifier also contains an implication; so when UG is used, conditional proof is usually used with it. The only exceptions would be when the variable x can denote any object whatsoever, so that it is not restricted to any domain.

In spite of this closeness between conditional proof and UG, they are different and should not be confused. Conditional proof is a rule of inference that involves implications, not quantifiers; UG is a rule of inference that is about universal quantifiers, even though it indirectly involves implications too. Another difference is that when you set up a conditional proof, the assumption made can be any type of statement, even a false one. But when you set up a UG proof, the only thing you get to assume is that some variable is a member of some domain. I usually say “Let x be a ...,” rather than “Assume x is a ...,” when I start a UG proof, because I’m not really assuming anything; I’m just specifying how I’m going to use a certain letter. But as I said above, most mathematicians ignore this distinction, and doing so normally creates no problems.

This discussion is related to the comments made in Section 2.2 about the word “whenever.” This word is very close to the word “if” in mathematics, but there is a difference, namely that the word “whenever” actually combines an implication with a universal quantifier. For instance, the words “Whenever a function is continuous, it’s integrable” should be interpreted as meaning “every continuous function is integrable,” which can be written “ $\forall f (f \text{ is a continuous function} \rightarrow f \text{ is integrable})$.” To be sure, a *theorem* that says “If a function is continuous, it’s integrable” should be interpreted the same way, because of the convention stated earlier in this section. But the word “if” conveys a universal quantifier only some of the time, whereas the word “whenever” *always* does.

Example 7: Here is a simple example of incorrect and correct uses of UG. Suppose that we start a proof by assuming that $x > 7$, and from this we prove that $x > 0$. Can we now apply UG to get the step $\forall x (x > 0)$, followed by conditional proof to reach the conclusion $x > 7 \rightarrow \forall x (x > 0)$? Definitely not! The rule UG may not be applied to the step $x > 0$ because there is an assumption about x in effect at that point in the proof. Furthermore, the conclusion obtained can't be right, since it would be nonsense to claim that information about one number implies an incorrect general statement about all numbers. Instead, after the step $x > 0$, we can assert $x > 7 \rightarrow x > 0$ by conditional proof and then apply UG to reach the correct conclusion $\forall x (x > 7 \rightarrow x > 0)$.

Let us now turn to proof methods involving the existential quantifier. One very important one is based on the following idea. Suppose we know that some object exists. In other words suppose that, in a proof, we have a step of the form $\exists x P(x)$. If we don't know a specific value of x that makes $P(x)$ true, it's convenient (and harmless) to introduce a temporary name for some unknown object satisfying $P(x)$.

Rule of Inference: Existential Specification or ES

If you have a step of the form $\exists x P(x)$, you may assert $P(c)$, where c is a new, temporary constant symbol.

Remarks: (1) It is important to understand the conditions required for the correct use of this rule. The name c introduced for the unknown object represents *one particular* object. Therefore, it must be a constant, not a variable; it cannot be quantified. Since we don't know anything else about this object except that it satisfies $P(x)$, it must be a *new* symbol; that is, it may not appear earlier in the proof. Finally, this rule is meant to be a temporary convenience; the new symbol should not appear in the final conclusion of the proof.

To view it another way, a constant introduced by ES may be viewed as a temporary *definition*. Section 3.4 pointed out that mathematicians normally make a *permanent* definition only when they know that some object exists *uniquely*. The rule ES provides us with a more limited course of action we can take when the uniqueness condition is lacking.

(2) Mathematicians often seem to violate the restrictions on ES. They go from $\exists x P(x)$ to $P(x)$, using the same letter (apparently a variable and not a new symbol) that appeared in the quantifier when they apply ES. However, if you carefully examine these proofs, you see that after they apply ES in this way, they treat x as though it was a constant, not a variable. This avoids any danger of faulty logic, but it takes some

experience to do it this way and keep things straight. Until you are very familiar with doing proofs, I suggest that you always use a new letter whenever you apply ES.

(3) By now you have perhaps figured out the terminology being used for these quantifier proof methods. Specification (also called **instantiation**) means *using* a known quantified statement to assert a statement with the quantifier *removed*. Generalization means *proving* a quantified statement from a known *unquantified* statement.

Example 8: Refer to Proof Preview 1 in Section 2.2. The rigorous definition of “ n is even” is $\exists m (n = 2m)$, where m also stands for an integer. So a more rigorous version of that proof would have used ES to go from this quantified statement to the unquantified statement $n = 2m$ or, more correctly, $n = 2c$ (see Exercise 2).

Many important theorems are so-called **existence theorems**, which means that they say that something exists without telling you how to find it. The rule ES is the main tool for making use of an existence theorem.

Example 9: An existence theorem that is extremely important in calculus is the **mean value theorem for derivatives**. It says that if a function f is continuous on the closed interval $[a, b]$ and differentiable on the open interval (a, b) , then there is a number x strictly between a and b such that $f'(x) = [f(b) - f(a)]/(b - a)$. Let’s see how this might be applied in practice. First of all, the unquantified variables a , b , and f are understood to be universally quantified. Therefore, by applying US, we can give them any particular values we want. For instance, we could let $a = 0$, $b = \pi$, and $f(x) = x^2 + \cos x$. (Note how a function is specified by giving a defining equation for it, rather than a numerical value.) This function is differentiable (and therefore continuous) on the whole real line. So by modus ponens, we get $\exists x (0 < x < \pi \text{ and } f'(x) = \pi - 2/\pi)$. By taking the derivative of $f(x)$, we obtain $\exists x (0 < x < \pi \text{ and } 2x - \sin x = \pi - 2/\pi)$.

If you try to solve this equation for x , you quickly run out of things to do; it’s essentially impossible to solve. (One could use Newton’s method or some other approximation technique to compute a solution to as many decimal places as desired, but that is not our purpose here.) So the only way to eliminate the existential quantifier is to use ES; we can say “Let c be a number such that $0 < c < \pi$ and $2c - \sin c = \pi - 2/\pi$.” (The use of ES is usually accompanied by the word “let” in this manner.) Then, even though we don’t know the exact value of c , we have a convenient temporary symbol that denotes a solution to this equation between 0 and π .

Example 10: Here is an incorrect use of ES. Suppose we have a step of the form $\exists x P(x)$. Consider this proof:

- | | |
|----------------------|------------------|
| (1) $\exists x P(x)$ | [Proved somehow] |
| (2) $P(x)$ | ES on step 1 |
| (3) $\forall x P(x)$ | UG on step 2 |

Of course, the mistake here is that the letter x in step 2 gets treated as a variable in step 3, but after the use of ES it has to be treated as a constant. Note that if the above proof were correct, it would (by conditional proof) yield an absurd conclusion: that if there's one object with a certain property, then all objects have that property.

Example 11: Here is another, more subtle, error to watch out for when using ES. Suppose that we have a step of the form $\exists x (P(x) \wedge Q(x))$. Then we can certainly apply ES to get $P(c) \wedge Q(c)$. But suppose instead that we have the step $\exists x P(x) \wedge \exists x Q(x)$. Can we still apply ES to obtain $P(c) \wedge Q(c)$? No, this is not allowed! The rule ES can only be applied to one quantifier at a time. To eliminate the quantifiers from this latter statement, we must break it up as follows:

- | | | |
|-----|--|------------------------------|
| (1) | $\exists x P(x) \wedge \exists x Q(x)$ | [Proved somehow] |
| (2) | $\exists x P(x)$ | From step 1 |
| (3) | $P(c)$ | ES on step 2 |
| (4) | $\exists x Q(x)$ | From step 1 |
| (5) | $Q(b)$ | ES on step 4 |
| (6) | $P(c) \wedge Q(b)$ | Conjunction on steps 3 and 5 |

Note that we can't write $Q(c)$ at step 5 because c would not be a new constant symbol. In fact, there is no way to prove $P(c) \wedge Q(c)$ from step 1, because step 1 does not say that there is a single object that satisfies *both* P and Q simultaneously (even though the same variable x is used in both conjuncts of step 1). To make this more concrete, let $P(x)$ be " $x > 0$ " and let $Q(x)$ be " $x < 0$." Then step 1 of the above argument is true in the real number system, but $c > 0 \wedge c < 0$ cannot be true.

We have one more important quantifier proof method to discuss.

Axiom: Existential Generalization or EG

$$P(t) \rightarrow \exists x P(x)$$

where t is a term with the same restrictions as in the rule US.

It is no accident that US and EG both mention a term t with exactly the same restrictions. In fact, either of them can be derived from the other (see Theorem 4.3 and Exercise 16).

Since EG is an implication, it can be combined with modus ponens to form a new rule of inference. In this form, it is *by far* the most natural and common way to prove an existential statement. A more down-to-earth name for this proof method would be

proof by example, since what it says is that if you want to prove that something exists, it suffices to find one actual example of whatever it is. For instance, if you want to show there's a real number with a certain property, the cleanest way by far is to find a specific number with that property.

However, recall Theorem 3.1, which said that existential quantifiers normally describe functions of the universally quantified variables to the outside of them. For this reason, the example you find in a proof by example won't necessarily be a constant; it might have to depend on some variables. That is why EG refers to a term or expression t , rather than a constant. In a sense, EG is just a restatement of Theorem 3.1.

Example 12: Proof Preview 4 (in Section 3.3) provides a typical example of the use of EG. We want to prove $\forall x, y \exists z (z > x \wedge z > y)$, with all three variables being real variables. We could begin the proof with the words "Let x and y be given," reflecting the fact that we can ignore the universal quantifiers, by UG. Then we need to prove the statement $\exists z (z > x \wedge z > y)$. To prove this by EG, we must find a term or expression that works when substituted for the variable z , and this term may involve x and y . As we showed in Proof Preview 4, the term $z = |x| + |y| + 1$ works. In standard mathematical writing, the second sentence of this proof would probably be simply "Let $z = |x| + |y| + 1$." (In practice mathematicians don't use the letter t to represent a term in such proofs, any more than they use the letter P to represent a statement.)

Another way to prove this statement is by cases (see Exercise 4).

Example 13: Let's prove, assuming basic results from calculus, that given any number, there's a real-valued function (other than the zero function) whose derivative is that (constant) number times itself. The statement to be proved has the form $\forall k \exists f (f \neq 0 \wedge f' = kf)$. Note that, in terms of the logical structure of this statement, both k and f are mathematical variables, even though k has been called a constant (because it's not a variable of the function f), and f represents a function, not a number. So let k be any real number. By EG, we just need to find a nonzero function $f(x)$, which can depend on k , with the desired property. Perhaps you've already figured out that we should let $f(x) = e^{kx}$. It can then be easily verified by differentiation that $f'(x) = kf(x)$.

Most of the time, when one uses EG or Theorem 3.1, the term found for the existentially quantified variable must involve all the universally quantified variables to the left of it. So, in Example 12, it's not possible to find an expression for z that does not involve both x and y . Similarly, in Example 13, the expression we find for f must involve k . But it's permissible to omit a variable, as the next example shows.

Example 14: Suppose we want to prove $\forall x \exists y (x + y = x)$, where x and y are real variables. The simplest expression that works for y in the equation $x + y = x$ is 0. This certainly satisfies the requirement of Theorem 3.1. So, assuming that we know that $x + 0 = x$, we can consider this result proved, by EG and UG. The fact that we could find an expression for y that does not involve x tells us that we could just as easily prove the stronger (that is, better) result $\exists y \forall x (x + y = x)$.

Counterexamples

The words (as opposed to the methods) “existential generalization” and “proof by example” are rarely used by mathematicians. In contrast, mathematicians frequently talk about **counterexamples**, primarily as a method of *disproving* statements. This method is just a special case of EG, but it is used so often that it deserves separate discussion.

One standard type of mathematics problem asks the reader to prove or disprove some statement. This often involves more work than a problem that just asks the reader to prove a statement: first you have to determine (or at least guess) whether the statement is true or false; then you must prove the statement or its negation.

Now, imagine that you are asked to prove or disprove a statement of the form $\forall x P(x)$. If you think the statement is true, you probably try to prove it by UG. But if you think it's false, how do you disprove it? Well, *disproving* $\forall x P(x)$ means *proving* $\sim \forall x P(x)$, which is equivalent to $\exists x \sim P(x)$. And by EG, we know we can prove this if we can find a term t such that $\sim P(t)$ holds. That is, we want to find a specific example of an object for which P is false. Such an example is called a **counterexample** to the statement $\forall x P(x)$.

When a statement of the form $\forall x P(x)$ is involved, the words “prove or find a counterexample” are more commonly used than “prove or disprove.”

Example 15: Suppose we are asked to prove or disprove that $n^2 - n + 41$ is prime for every nonnegative integer n (recall Exercise 5 of Section 1.3). If this were true, it might be very difficult to prove. But it's not true, and all it takes to show this is a single counterexample. Interestingly, $n^2 - n + 41$ is prime for all integers from 0 to 40, but 41 is obviously (in retrospect, anyway) a counterexample.

Example 16: Chapter 1 discussed Goldbach's conjecture and de Polignac's conjecture. Both of these are almost certainly true, but neither has been proved. What would it take to disprove these conjectures? A counterexample to Goldbach's conjecture would be a positive even number (greater than 2) that is not the sum of any two prime numbers. Since there are only a *finite* number of ways to express a given integer as a sum of two positive integers, a counterexample to Goldbach's conjecture, if it exists, could be identified by simple arithmetic (but lots of it! A powerful computer would probably be required).

Similarly, a counterexample to de Polignac's conjecture would be an even number that is not the *difference* of any two prime numbers. But there are an infinite number of ways to express a given integer as a difference of two integers, and there are also an infinite number of primes. So there would be no way to verify that a given number was a counterexample merely by arithmetic computation. For instance, suppose that you believed that 6 was not the difference between any two prime numbers. How would you attempt to verify this, even with a powerful computer?

We conclude the main part of this section with two tables that you should look through now and refer to as needed in the future. Table 4.2 contains a list of some laws of logic, analogous to the list of tautologies in Appendix 3. All these are provable from

Table 4.2 Some useful laws of logic

In the following statements, it is always assumed that the proposition R does not contain the variable x . The propositions P and Q may be assumed to contain, as free variables, the variables in the quantifiers that precede them. The restrictions on the term t are as described previously in the definition of universal specification.

- (1) $\forall x P \rightarrow \exists x P$
- (2) $\forall x \forall y P \leftrightarrow \forall y \forall x P$
- (3) $\exists x \exists y P \leftrightarrow \exists y \exists x P$
- (4) $\exists x \forall y P \rightarrow \forall y \exists x P$
- (5) $\forall x R \leftrightarrow R$
- (6) $\exists x R \leftrightarrow R$
- (7) $\forall x P(x) \rightarrow P(t)$ (Universal specification)
- (8) $P(t) \rightarrow \exists x P(x)$ (Existential generalization)

De Morgan's laws for quantifiers

- (9) $\sim \forall x P \leftrightarrow \exists x \sim P$
- (10) $\sim \exists x P \leftrightarrow \forall x \sim P$

Replacement of one quantifier by the other

- (11) $\forall x P \leftrightarrow \sim \exists x \sim P$
- (12) $\exists x P \leftrightarrow \sim \forall x \sim P$

Distributing quantifiers over connectives

- (13) $\forall x (P \wedge Q) \leftrightarrow (\forall x P \wedge \forall x Q)$
- (14) $\exists x (P \vee Q) \leftrightarrow (\exists x P \vee \exists x Q)$
- (15) $(\forall x P \vee \forall x Q) \rightarrow \forall x (P \vee Q)$
- (16) $\exists x (P \wedge Q) \rightarrow (\exists x P \wedge \exists x Q)$
- (17) $\forall x (P \vee R) \leftrightarrow (\forall x P \vee R)$
- (18) $\exists x (P \wedge R) \leftrightarrow (\exists x P \wedge R)$
- (19) $\exists x (P \rightarrow Q) \leftrightarrow (\forall x P \rightarrow \exists x Q)$
- (20) $\forall x (R \rightarrow Q) \leftrightarrow (R \rightarrow \forall x Q)$
- (21) $\forall x (P \rightarrow R) \leftrightarrow (\exists x P \rightarrow R)$

Table 4.3 Summary of quantifier proof methods

Statement	Ways to Prove	Ways to Use
A universally quantified statement $\forall x P(x)$	UG: Prove $P(x)$, for an arbitrary x . Indirectly: Assume $\exists x \sim P(x)$ and derive a contradiction.	US: Assert $P(t)$, for any appropriate term t .
An existentially quantified statement $\exists x P(x)$	EG: Prove $P(t)$, for some appropriate term t . Indirectly: Assume $\forall x \sim P(x)$, and derive a contradiction.	ES: Assert $P(c)$ for a new constant c .

our axiom system, and the exercises ask you to prove a few of them. Better yet, look through the list yourself, pick a couple of statements in the list that look the *least* obvious to you, and see if you can prove them.

Even though most of the laws of logic in Table 4.2 are not included as axioms in our system, you should feel free to consider them as known results or theorems, unless, of course, you're asked to do a proof using only what's in the axiom system. Pay particular attention to the laws in Table 4.2 that involve the propositional variable R . It is specified that, in these laws, R does not contain x as a free variable, and this restriction allows steps that otherwise would not be correct.

Example 17: An important feature of the real number system is that between any two distinct real numbers there's another one. Suppose we want to prove this. The statement can be symbolized, a bit loosely, as $\forall x, y (x < y \rightarrow \exists z (x < z < y))$.

It is not difficult to prove this statement as it stands, using methods of proof from this section and the previous one. But a slightly different approach is to note that the variable z does not occur in the inequality $x < y$. This means we can use laws 5 and 19 of Table 4.2 to rewrite $(x < y \rightarrow \exists z (x < z < y))$ in the equivalent form $\exists z (x < y \rightarrow x < z < y)$. So the statement we want to prove becomes $\forall x, y \exists z (x < y \rightarrow x < z < y)$. The point is that this revised statement fits the conditions needed to apply Theorem 3.1, which provides a straightforward way of completing the proof (see Exercise 10).

Table 4.3 is a quantifier version of Table 4.1 in the previous section. Since there are only two quantifiers as compared to five connectives, Table 4.3 is shorter and simpler than Table 4.1. Probably the most important thing to get from Table 4.3 is what the roles of the four principles US, UG, ES, and EG are.

Some Theorems Involving Quantifiers (Optional Material)

We now prove some simple theorems or laws of logic, all of which appear in Table 4.2. The first two provide the derivations of the two asterisked quantifier axioms, and the

next two prove extremely simple laws of logic. Theorem 4.6 is probably the most useful of these results, since it clarifies some of the subtleties involved when quantifiers are combined with connectives. The fact that the content of these theorems is pure logic gives their proofs a rather artificial, nonmathematical flavor.

Theorem 4.2: Quantifier axiom II-3, $\sim \exists x P(x) \leftrightarrow \forall x \sim P(x)$, is derivable from the nonasterisked portion of our axiom system.

Proof: For the forward direction, assume $\sim \exists x P(x)$. Using tautology 23 (and substitution), we can replace $P(x)$ by $\sim \sim P(x)$. This gives us the step $\sim \exists x \sim \sim P(x)$. But this step contains a substatement of the form $\exists x \sim \dots$. Therefore, we can apply axiom II-2, plus substitution, to yield the step $\sim \sim \forall x \sim P(x)$. Applying tautology 23 again lets us delete the double not in front of this statement and gives us $\forall x \sim P(x)$. This completes the conditional proof of the implication $\sim \exists x P(x) \rightarrow \forall x \sim P(x)$.

For the reverse direction, assume $\forall x \sim P(x)$. Again applying tautology 23 plus substitution, we get $\sim \sim \forall x \sim P(x)$. Then by applying axiom II-2 plus substitution to the substatement $\sim \forall x \dots$, we get the step $\sim \exists x \sim \sim P(x)$. With tautology 23 plus substitution used one final time, we obtain $\sim \exists x P(x)$. Therefore, by conditional proof, the reverse implication is also established. By the biconditional rule, we conclude $\sim \exists x P(x) \leftrightarrow \forall x \sim P(x)$. (See Exercise 15 for the converse of this theorem.) ■

Theorem 4.3: Existential Generalization, $P(t) \rightarrow \exists x P(x)$, is derivable from the nonasterisked portion of our axiom system.

Proof: We prove the contrapositive of the desired statement. So assume $\sim \exists x P(x)$. By Theorem 4.2, this gives us $\forall x \sim P(x)$. By US, this implies $\sim P(t)$, which is what we want. (See Exercise 16 for the converse result.) ■

Theorem 4.4: For any statement $P(x)$,

- (a) $\forall x P(x) \leftrightarrow \sim \exists x \sim P(x)$
- (b) $\exists x P(x) \leftrightarrow \sim \forall x \sim P(x)$

Proof: These statements are just the result of negating both sides of De Morgan's Laws for quantifiers (see Exercise 11). ■

Theorem 4.5: (a) For any statement $P(x)$, $\forall x P(x) \rightarrow \exists x P(x)$.

(b) For any statement $P(x, y)$, $\exists x \forall y P(x, y) \rightarrow \forall y \exists x P(x, y)$.

Proof: (a) Assume $\forall x P(x)$. By US, $P(x)$. Then by EG, $\exists x P(x)$. By conditional proof, we are done.

(b) Assume $\exists x \forall y P(x, y)$. Then by ES, let c be an object satisfying $\forall y P(c, y)$. To prove the right side of the implication, let y be given. By US applied to $\forall y P(c, y)$, we have $P(c, y)$. Then we can use EG to get $\exists x P(x, y)$. Finally, since y was arbitrary, we can apply UG to get $\forall y \exists x P(x, y)$. ■

You might get more out of the proof of Theorem 4.5(b) if you try to use similar logic to prove its converse. Recall from Section 3.3 that the converse of this statement is not true in general, and so of course it can't be proved.

Example 18: In the real number system, $\exists x \forall y (x + y = y)$ is true, since the constant value 0 works for x . So, by Theorem 4.5(b), $\forall y \exists x (x + y = y)$ must also be true.

In the real number system, the statement $\forall x \exists y (x + y = 0)$ is also true, since for any x we can let $y = -x$. But since this value of y depends on x and there is no constant value of y that works for all x , the statement $\exists y \forall x (x + y = 0)$ is false.

Theorem 4.6: For any statements $P(x)$ and $Q(x)$:

$$(a) \forall x P(x) \wedge \forall x Q(x) \leftrightarrow \forall x [P(x) \wedge Q(x)]$$

$$(b) \forall x P(x) \vee \forall x Q(x) \rightarrow \forall x [P(x) \vee Q(x)]$$

$$(c) \exists x P(x) \vee \exists x Q(x) \leftrightarrow \exists x [P(x) \vee Q(x)]$$

$$(d) \exists x [P(x) \wedge Q(x)] \rightarrow \exists x P(x) \wedge \exists x Q(x)$$

Proof: (a) See Exercise 12.

(b) Assume the left side. Since this is a disjunction, we can do a proof by cases.

Case 1: Assume $\forall x P(x)$. Let x be given. By US, we have $P(x)$. By propositional logic, we then get $P(x) \wedge Q(x)$. Finally, UG yields $\forall x [P(x) \wedge Q(x)]$. Case 2: Assume $\forall x Q(x)$. The argument for this case is almost identical to the other case and so we omit it. (It is common practice to say something like this instead of practically repeating a proof.)

(c) See Exercise 13.

(d) Assume the left side. Using ES, we can write $P(c) \wedge Q(c)$. By propositional logic, we then get $P(c)$. Then EG yields $\exists x P(x)$. An almost identical argument proves $\exists x Q(x)$. Finally, the conjunction rule yields the desired statement. ■

Note that Theorem 4.6(b) and (d) are just conditionals, not biconditionals. Exercise 14 asks you to show that the converses of these conditionals are not valid.

Exercises 4.3

(1) The rigorous definition of $A \subseteq B$ is $\forall x (x \in A \rightarrow x \in B)$. Using this definition, write a more correct version of the proof in Proof Preview 2 (Section 2.3), dealing properly with the quantification of x as well as that of A , B , and C .

(2) Write a more rigorous version of the proof in Proof Preview 1 (Section 2.2). Use the definition of “ n is even” given in Example 8 of this section, and write a similar definition of “ n is odd.” Because of the existential quantifiers involved, you need to use ES and EG in your proof.

(3) Carefully prove the argument given in Example 1 of Section 3.2.

(4) Prove the result discussed in Example 12, using cases that are based on the relationship between x and y .

(5) In this exercise, n and k denote integers.

(a) Write a symbolic statement that captures the meaning of “ n is divisible by k ,” or, equivalently, “ n is a multiple of k .” Your solution should include a quantifier and should mention multiplication rather than division or fractions.

(b) Prove that if n is a multiple of k , then so is n^2 . Make sure to treat quantifiers rigorously.

(c) Prove that if n is one greater than a multiple of k , then so is n^2 . (Same warning as in part (b).)

(d) Prove or find a counterexample to the converse of part (b).

(6) Redo the argument of Example 4 of the previous section, carefully including all the implicit quantifiers and the reasoning needed to deal with them.

(7) Give a more rigorous, axiomatic proof of Theorem 3.4 than was given on page 68.

*(8) Make up at least three symbolic statements, *not* already in Table 4.2 and *not* tautologies, that you believe are laws of logic. Then prove them from our axiom system.

(9) Prove the following laws from Table 4.2: 2, 19, and 20. Do not make these informal proofs; rather, make them axiomatic and quite formal. You may use any theorems in your proofs, but you may not use any of the laws in Table 4.2.

(10) (a) Prove the result discussed in Example 17 by the first method outlined in that example (without directly referring to any of the laws in Table 4.2).

(b) Prove this result by the second method outlined in Example 17 (using laws 5 and 19 of Table 4.2, and Theorem 3.1). Don't try to make this proof very rigorous.

Exercises 11 through 16 pertain to the optional material at the end of this section.

(11) Prove Theorem 4.4.

(12) Prove Theorem 4.6(a).

(13) Prove Theorem 4.6(c).

(14) Give examples to show that the converses of Theorem 4.6(b) and (d) are *not* laws of logic. An example in this situation should consist of a specific domain for the variable x and specific statements for P and Q .

(15) Prove the converse of Theorem 4.2. In other words, derive axiom II-2 from the rest of the axiom system, including axiom II-3.

(16) Prove the converse of Theorem 4.3. In other words, derive axiom II-1 from the rest of the axiom system, including axiom II-4.

Critique the following proofs. (If necessary, review the instructions for such problems in Exercises 4.2.)

(17) **Theorem:** In the real number system, $\exists x \forall y (y + x = 3)$.

Proof: Let y be any real number. We know from basic algebra that $y + (3 - y) = 3$, and by UG we can conclude that $\forall y (y + (3 - y) = 3)$. Now, the expression $3 - y$ certainly denotes a real number, so we can apply EG to obtain $\exists x \forall y (y + x = 3)$. (Technically, we are applying EG with $P(x)$ being $\forall y (y + x = 3)$ and the term t being $3 - y$.)

(18) The purpose of this problem is to illustrate the quantifier-related restriction in the definition of the substitution rule of inference.

Theorem: If x is a positive real number, then all real numbers are positive.

Proof: Assume $x > 0$. We also know that $1 > 0$, and therefore $x > 0$ iff $1 > 0$. By substitution, we can do the same thing to both sides of this equivalence; in particular, we can say $\forall x (x > 0)$ iff $\forall x (1 > 0)$. But the inequality $1 > 0$ is true for every value of x , so $\forall x (1 > 0)$ is true. Therefore, we can conclude $\forall x (x > 0)$.

4.4 The Use of Equations in Proofs

The last logic-based part of our axiom system consists of the equality axioms. These were discussed briefly in Section 3.4. Now we examine them, and their use in proofs, in detail. Axioms III-1 and III-2 (the reflexive and symmetric properties of equality) are so simple that they are not often used overtly in proofs, and there isn't a whole lot to say about them. Here is a small point involving symmetry. Remember that every equation in mathematics is a two-way street (the same phrase that we applied to biconditionals) and that even if a particular equation is usually applied in one direction, it must be allowable to apply it in the other direction as well.

Example 1: An equation that is frequently used in both directions is the distributive law for real numbers: $x(y + z) = xy + xz$. When you see an algebraic law like this, it's natural to think that the equation is usually used to change expressions that look like the left side into expressions that look like the right side. In this direction, the distributive law provides the main rule for multiplying out algebraic expressions. But the distributive law is equally important when used from right to left, for in that direction it provides the main rule for *factoring* algebraic expressions.

The transitive property of equality, axiom III-3, is more substantial than the previous two axioms. It is this axiom that allows you to write an extended equation and then conclude that the first expression equals the last expression. This technique is used in proofs as well as in problem solving, as the following example illustrates.

Example 2: Suppose you want to factor the expression $x^3 - 6a^2 - 2ax^2 + 3ax$, and you decide to try factoring by grouping. Then you might write

$$\begin{aligned} x^3 - 6a^2 - 2ax^2 + 3ax &= x^3 - 2ax^2 + 3ax - 6a^2 \\ &= x^2(x - 2a) + 3a(x - 2a) \\ &= (x^2 + 3a)(x - 2a) \end{aligned}$$

Of course, when you write this, you mean to say that the final, factored expression is equal to the original expression, and almost anyone reading it would interpret it in that way. But how can this conclusion be proved rigorously? Let's take a look.

The extended equation above has the structure $A = B = C = D$, where A , B , C , and D are particular expressions. To turn this into a formal proof, we would first need to prove the three separate equations $A = B$, $B = C$, and $C = D$. For now, we won't concern ourselves with how to prove these separate equations (see Exercise 3). From the two separate equations $A = B$ and $B = C$, we can use transitivity (plus propositional logic) to get $A = C$. From $A = C$ and $C = D$, we can then prove $A = D$ in a similar manner.

By the way, this example illustrates another point. Many people think of doing proofs in mathematics as a very different (and much harder) activity from ordinary problem solving. But it's really a false distinction. Every time you solve a problem in mathematics, you must have some justification or rationale for your steps; this means that you must have some sort of proof for your solution. Realizing this may help you to see that writing proofs is not such a strange activity.

To illustrate this, suppose you were given the task to find the value of such-and-such. You would probably view this as a problem, as opposed to a proof, and would use whatever means you could think of to simplify or evaluate the given expression. But suppose instead that you were asked to prove that such-and-such equals 2. You would probably think of this as a proof, and thus harder than an ordinary problem. The irony of this is that the second version should be easier than the first, because it tells you what the answer is.

Example 3: The transitive property of equality is used in just about every proof in Appendix 2. A typical example occurs in Theorem A-5. Step 2 of the formal proof is the equation $x(x + 0) = x \cdot x$, and step 3 says that $x(x + 0) = x \cdot x + x \cdot 0$. By transitivity of equality (plus axiom III-2 and a bit of propositional logic), this enables us to assert that $x \cdot x + x \cdot 0 = x \cdot x$.

The last equality axiom, substitution of equals, has already been discussed a bit. Very simply, if two things are equal, then they are *interchangeable*. Remember that this axiom and the propositional rule of inference substitution (described in Section 4.2) are based on a similar idea, but they are grammatically very different.

Example 4: Suppose we start with the known fact that $\sin x \leq 1$, for any number x . Then, by using US, we can derive $\sin(x + y) \leq 1$. Now, we also have the trigonometric formula

$$\sin(x + y) = \sin x \cos y + \cos x \sin y$$

Then we can use axiom III-4 to replace the expression $\sin(x + y)$ with the right side of this equation, to obtain

$$\sin x \cos y + \cos x \sin y \leq 1$$

Since x and y were arbitrary in this derivation, this inequality can be concluded for all real numbers x and y .

The following theorem often gives us a more usable form of axiom III-4.

Theorem 4.7: Let $t(x)$ be an expression or term containing the free variable x , and let $t(y)$ denote the same term with *some or all* of the occurrences of x replaced with the variable y . Then $x = y$ implies $t(x) = t(y)$.

Proof: See Exercise 4. ■

A term or expression in mathematics may be undefined, in the sense that it can't denote an actual value or object. The most common examples of this are fractions with denominator 0. The implication in Theorem 4.7 makes no sense if the terms $t(x)$ and $t(y)$ are undefined (unless we want to say that two undefined things equal each other, which is a consistent viewpoint but best avoided since it can create confusion). Therefore, the usual convention is that *Theorem 4.7 applies only when the expressions on the right side of the implication are defined*. The significance of this restriction becomes apparent in Example 11 at the end of this section.

Doing the Same Thing to Both Sides of an Equation

As it's stated, it's hard to see just how powerful Theorem 4.7 is. But in fact it is the basis of the fundamental rule that you can do just about anything to both sides of an equation, provided that you do precisely the same thing to both sides. To see how Theorem 4.7 says this, remember that the variables x and y in it are universally quantified. Therefore, using US, they can be replaced by *any expressions whatsoever*. So the practical significance of Theorem 4.7 is that whenever we know any equation, we can conclude any other equation in which the two sides of the original equation appear in the same way within the two sides of the new equation. The next few examples illustrate how this works.

Example 5: Let $t(x)$ be the expression " $x + z$," so that $t(y)$ is the expression " $y + z$." Then the theorem for this t reads: $x = y \rightarrow x + z = y + z$. Since x , y , and z are all understood to be universally quantified, this implication holds with any three expressions in their place. In other words, an immediate corollary of Theorem 4.7 is that *you can add any expression you want to both sides of an equation*. If the original equation is true, the new one must be also. Similar reasoning shows that you can subtract, multiply, or divide both sides of an equation by the same thing.

Example 6: Let $t(x)$ be the expression x^2 . Then Theorem 4.7, with this particular t , becomes $x = y \rightarrow x^2 = y^2$. In other words, you can square both sides of an equation.

Example 7: Now let's use the variables A , B , and C to stand for sets. We can let $t(A)$ be the expression $A \cup C$, and $t(B)$ be the expression $B \cup C$. Theorem 4.7 then yields

the implication $A = B \rightarrow A \cup C = B \cup C$. In words, you can form the union of both sides of an equation between sets, with the same set.

Example 8: There are places in Appendix 2 where axiom III-4 is stated as a reason for a step, but it might be more to the point to use Theorem 4.7 as the justification. For instance, step 2 of the formal proof of Theorem A-5 is obtained by multiplying both sides of step 1 by x .

Reversibility

Our discussion so far has indicated that you can do anything you want to both sides of an equation, but there are some subtleties involved with this rule. You may remember these subtleties from precalculus algebra. They have to do with reversibility of steps used in solving or simplifying an equation.

Not all steps allowed by Theorem 4.7 are reversible; sometimes that matters, and sometimes it doesn't. Let's clarify this with some examples.

Example 9: Suppose we want to solve the equation $x + 2 = 8 - 2x$. Our solution might look like this:

$x + 2 = 8 - 2x$	
$2x + x + 2 = 8$	Adding $2x$ to both sides
$3x + 2 = 8$	Combining terms
$3x = 8 - 2 = 6$	Subtracting 2 from both sides
$x = 6/3 = 2$	Dividing both sides by 3

So we would say that $x = 2$ is the solution of the equation. But if we view the above solution as a sort of proof, then what exactly have we proved? Have we proved that if a number x satisfies the given equation, then $x = 2$? Or have we proved the converse, that if $x = 2$, then x satisfies the given equation? Or have we proved both directions, that is, a biconditional: x satisfies the equation if and only if $x = 2$?

Whenever you solve an equation, your solution should establish a biconditional if possible. In other words, to solve an equation (in one variable) means to find a set of numbers (the **solution set**) such that any number in that set satisfies the equation, and no other numbers do. In the above example, we're not just saying that 2 works in the equation; we're also saying that no others work.

Thus, when solving or simplifying an equation, in a sense more care is required than when doing an ordinary proof: you have to make sure your steps are *reversible*. In the above example, the steps are definitely reversible: adding something to both sides can be reversed by subtracting that thing from both sides, dividing both sides by 3 can be reversed by tripling both sides, and so on. So the solution shown establishes a biconditional, as it should.

What sorts of steps used to solve equations would not be reversible? Two standard ones are squaring both sides and multiplying both sides by an expression that could equal zero, as the next two examples illustrate.

Example 10: Suppose we want to solve the equation $\sqrt{x+3} + x = 3$. The obvious steps to solve this equation are as follows:

$\sqrt{x+3} + x = 3$	
$\sqrt{x+3} = 3 - x$	Subtracting x from both sides
$x + 3 = 9 - 6x + x^2$	Squaring both sides
$0 = x^2 - 7x + 6$	Subtracting $x + 3$ from both sides
$0 = (x - 1)(x - 6)$	Factoring
$x = 1 \text{ or } x = 6$	See Exercise 12

But on checking these numbers in the original equation, we find that $x = 1$ works, but $x = 6$ doesn't. (The value $x = 6$ works if we set $\sqrt{9} = -3$, but that's not how the symbol $\sqrt{}$ is normally used.) You might remember that 6 is called an **extraneous solution** to this equation; but why does a "wrong solution" crop up here?

The reason that our solution method to this equation can lead to extraneous solutions is that squaring both sides of an equation is not reversible. It might seem that the reverse of squaring both sides would be taking the square root of both sides; but it's not, because if you start with a negative number, square it, and then take the square root of that number, you don't get the original number back. To put it another way, if the equation $A^2 = B^2$ is true, we can't conclude that $A = B$. We can only conclude that $|A| = |B|$, or $A = \pm B$.

Because one step is not reversible, the above solution to this equation only shows a forward implication, not a biconditional. So if a number satisfies the equation, it must be either 1 or 6. But we can't conclude from the steps shown that 1 and 6 do satisfy the equation. Note that this is not such a terrible situation: it just means we need to check for extraneous solutions. On the other hand, a solution to an equation that represented a reverse implication but not a forward one would be useless, since it would mean that you might not have found all the correct solutions.

For example, if you try to solve the equation $x^2 = 4$ by taking the square root of both sides to yield $x = 2$, you've made a mistake. The correct solution is $x = \pm 2$. In general, while squaring both sides of an equation is OK (provided you remember to check for extraneous solutions), taking the square root of both sides of an equation is a step that can lead to trouble. By the way, none of this complication would occur with cubing or taking the cube root of both sides of an equation, since these operations are truly the reverse (or *inverse*) of each other.

Example 11: Suppose you were asked to solve the equation $(\sin x)/x = 1$. You would probably multiply both sides by x and obtain $\sin x = x$. It can be shown that the only number satisfying this latter equation is 0. But you can't have a denominator of 0, so 0 is an extraneous solution, and the given equation has no solution.

Why does an extraneous solution occur here? The reason is that, in our solution, we multiplied both sides by an expression that turns out to be zero. This step is not reversible, since division by zero is impossible. (Remember the remark after Theorem 4.7, which said that what you do to both sides of an equation must keep both sides

defined.) Therefore, a solution that includes multiplying both sides of an equation by an expression that could be zero is only a forward implication, and extraneous solutions can appear.

Like squaring both sides of an equation, multiplying both sides by an expression that could be zero is fine, provided you check for extraneous solutions at the end. But dividing both sides by such an expression is dangerous and should be avoided if possible. For example, it's wrong to try to solve the equation $x^2 = 5x$ by dividing both sides by x . The correct method is to put everything on one side and then factor.

We conclude this section by justifying the fact that axioms III-2 and III-3 are marked with asterisks.

Theorem 4.8: Axioms III-2 and III-3 are superfluous; that is, they can be proved from the rest of the axiom system.

Proof: First let's prove symmetry: assume $x = y$. Applying axiom III-4 with $S(x)$ being the statement $x = x$, we get $x = x \leftrightarrow y = x$. But we know $x = x$ (from axiom III-1 and US), and so $y = x$ follows by propositional logic. Thus we have proved that $x = y$ implies $y = x$. ■

Exercises 4.4

(1) Complete the proof of Theorem 4.8, by proving axiom III-3, transitivity, from the rest of our axiom system.

(2) Prove: if $x = y$ and $u = v$, then $x + u = y + v$.

(3) Prove the three separate equations involved in Example 2. You may use any of the results of Appendix 2, as well as what's in the axiom system.

(4) Prove Theorem 4.7.

The remaining exercises of this section do not pertain specifically to the equality axioms. Rather, they involve the material in Appendix 2 and are intended to give you practice with all the proof methods that have been discussed in this chapter.

(5) Prove Theorem A-2.

(6) Prove Theorem A-4.

(7) Redo the formal proof given for Theorem A-8 in good, nonformal style.

(8) Redo the formal proof given for the forward direction of Theorem A-13 in good, nonformal style.

(9) Identify the omissions in the proof of Theorem A-7, and complete the proof by filling in the gaps.

In Exercises 10 through 17, prove the statement from the field axioms. (When such a statement is made, it is understood that you may also use all rules of inference and all axioms that are based on logic and equality.) You may also use any results in Appendix 2, unless stated otherwise. Also remember that all free variables in these statements are understood to be universally quantified. Consult your instructor if you have questions about style, format, or how much detail to show.

(10) The number 0 has no multiplicative inverse. (First write the statement in symbols.)

$$(11) (-1)x = -x$$

$$(12) xy = 0 \text{ iff } (x = 0 \text{ or } y = 0)$$

$$(13) x \neq 0 \text{ iff } x^{-1} \neq 0$$

$$(14) \text{ If } x \neq 0 \text{ and } y \neq 0, \text{ then } (xy)^{-1} = y^{-1}x^{-1}$$

$$(15) \text{ If } x \neq 0 \text{ and } y \neq 0, \text{ then } u/x + v/y = (uy + vx)/xy$$

$$(16) \forall x, y \exists z (x + z = y)$$

$$(17) (x + y)^2 = x^2 + 2xy + y^2$$

(18) Replace field axiom V-12 with the statement that $0 = 1$, and prove from this alternate set of axioms that there is only one number.

(19) Prove Theorem A-10(b).

(20) Prove Theorem A-10(c).

(21) Prove the last case ($x < 0$) of Theorem A-11.

(22) Prove Corollary A-12.

(23) Prove the right-to-left implication of Theorem A-13. Your proof may be formal, but it doesn't have to be.

(24) (a) Fill in the details of the proof of Theorem A-15(a). In particular, show that the cases used in that proof really do cover all possible cases.

(b) Prove Theorem A-15(b).

(c) Prove Theorem A-15(c).

*(25) Fill in the missing details in the proof of Theorem A-16.

The instructions for Exercises 26 through 31 are the same as those for Exercises 10 through 17, except that now you may use all the ordered field axioms.

- (26) If $x \leq y$ and $y \leq x$, then $x = y$. This is called the **antisymmetry property** of \leq .
- (27) (a) $x < y$ iff $-x > -y$
 (b) $x > 0$ iff $-x < 0$
- (28) If $x < y$ and $z < 0$, then $xz > yz$. Note that this is the usual rule about multiplying both sides of an inequality by a negative number.
- (29) $1 + 1 \neq 0$.
- (30) If $0 < x < y$, then $x^2 < y^2$. ($0 < x < y$ is an abbreviation for $0 < x$ and $x < y$.)
- *(31) If $x < y$, then $x^3 < y^3$.
- (32) Rewrite axioms V-13 through V-17 and the definitions that follow them, so that the only inequality symbol mentioned in the axioms is \geq , and the other three symbols are defined in terms of this one. Of course, make sure all your axioms and definitions are correct.
- (33) Critique the following proof. If necessary, review the instructions for such problems in Exercises 4.2.

Theorem: $1 + 1 \neq 0$.

Proof: Assume, on the contrary that $1 + 1 = 0$, that is, $2 = 0$. By Theorem 4.7, we can multiply both sides of this equation by $1/2$ to obtain $2(1/2) = 0(1/2)$. The left side equals 1, by axiom V-11; and the right side equals 0, by Theorem A-5. Therefore, $1 = 0$. But this is a contradiction (in conjunction with axiom V-12), so we are done, by indirect proof.

4.5 Mathematical Induction

This section is devoted to a single method of proof, known as the principle of mathematical induction (PMI), or simply induction. Induction is undoubtedly one of the most important proof techniques in mathematics. Mathematical induction is quite different from the axioms and rules of inference described in the previous three sections. It is not based on logic. Technically, it is an axiom for just one particular number system, the natural numbers. This would presumably make it less useful than very general methods like conditional proof, indirect proof, and so on, but natural numbers occur so universally in mathematics that proofs by induction are a vital part of every branch of the subject.

The term **natural numbers** refers to the numbers we all study first in grade school: 1, 2, 3, and so on. These, the positive integers, are the simplest type of number. (Some

books, particularly older ones, include 0 as a natural number and use the term **counting numbers** to denote the positive integers.)

Notation: The letter \mathbb{N} denotes the set of all natural numbers. For now, we use the letters m, n, k , and j as natural number variables, that is, variables whose domain is \mathbb{N} . (Later, we may use one or more of these letters to stand for *any* integer, not necessarily a positive one.)

Neither the notation just introduced nor the paragraph preceding it constitutes a mathematical definition of the set of natural numbers. Therefore, this notation and discussion may not be used to *prove* things about \mathbb{N} . The only basis for proving things about the natural numbers is the axioms pertaining to them. For example, to prove even the “obvious” fact that every natural number is positive, induction is required. You can’t prove it “by definition” because no definition has been given.

Axioms for the Natural Numbers

Our axioms for \mathbb{N} comprise group VI of Appendix 1. There are only three axioms in this group, and the first two are extremely simple. Axiom VI-1 just says that the number 1 is a natural number. Axiom VI-2 says that if you add 1 to any natural number, the result is still a natural number. These two axioms, taken together, can be thought of as describing how \mathbb{N} is generated.

The principle of mathematical induction can be stated in two equivalent forms, the set form and the statement form. In the axiom system, we have listed both. Since we have not studied sets yet, this section concentrates on the statement form.

Definition: The **principle of mathematical induction** (statement form) consists of all statements of the form

$$[P(1) \wedge \forall n (P(n) \rightarrow P(n+1))] \rightarrow \forall n P(n)$$

where $P(n)$ is any statement containing a free natural number variable n .

The Meaning of Mathematical Induction

Some people learn only the *method* of induction proofs without ever learning the reasoning behind induction. This approach can be successful with straightforward induction proofs, but it falls apart when the problems get more involved. So let’s spend some effort now to analyze the content of this principle.

To visualize what induction says, imagine that the natural numbers 1, 2, 3, ... are arranged vertically, in a sort of infinite ladder (see Figure 4.1). Let $P(n)$ be the statement that it’s possible to reach the n th step of the ladder. Then what does induction claim with this $P(n)$? Well, $P(1)$ says it’s possible to reach the first step, and $\forall n (P(n) \rightarrow P(n+1))$ says that, for any n , if you can reach the n th step, you can also

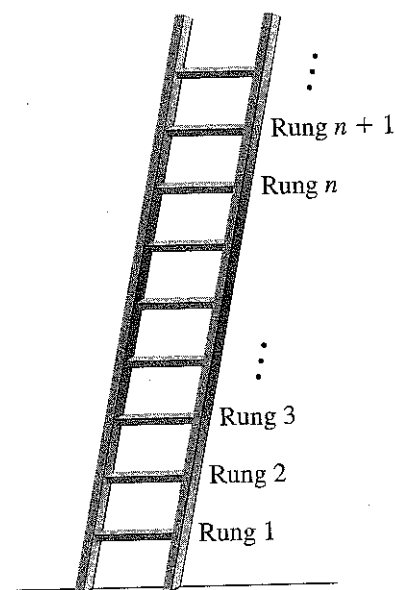


Figure 4.1 Ladder image for mathematical induction

reach the $(n + 1)$ -th step. In other words, it says that it is always possible to go one step higher than you already are. So the entire implication can be paraphrased as follows: if you can get to the first step of the ladder and you can always take one more step, then you can go as high as you want.

PMI can also be illustrated nicely with a horizontal image instead of a vertical one. Imagine an infinite row of standing dominos, as in Figure 4.2. Let $P(n)$ be the statement that the n th domino falls. Suppose we know that the first domino will be knocked over and that each domino is close enough to the next one so that, when it falls, the next one will also fall. Then induction, applied to this situation, states that every domino will fall—a fact that even most young children are aware of.

Now let's see what PMI says for an arbitrary statement $P(n)$. The hypothesis of induction says two things: first, it says that P is true for the particular number 1. Second, it says that whenever P is true for a number n , it must also be true for $n + 1$. Induction says that if these two things both hold, then P must be true for every natural number n . To see that this is valid, assume the hypothesis is true, and let's start listing numbers for which P must be true. It must be true for 1, because that's specifically stated. But then, since it's true for 1, it must be true for 2. Then, since it's true for 2, it must be true for 3, and so on. Continuing in this way, we see that P must be true for every value of n .

Remember that mathematical induction is an axiom based on the particular way

that the natural numbers are structured. It holds only because \mathbb{N} consists of a single infinite sequence of numbers. For example, there's no such thing as a direct induction proof on the set of all real numbers.

It is instructive to think about the relationship between induction and the other natural number axioms. It is similar to the relationship between modus ponens and conditional proof: induction is essentially the *converse* of axioms VI-1 and VI-2. Axioms VI-1 and VI-2 say that if you start at 1 and count by ones, you stay within \mathbb{N} . Induction says that if you start at 1 and count by ones, you eventually reach *every* member of \mathbb{N} .

This distinction can be made even clearer by referring to sets. Let A denote the set of all numbers that can be reached by starting at 1 and counting by ones. Then axioms VI-1 and VI-2 together say that A is a subset of \mathbb{N} , whereas PMI says that \mathbb{N} is a subset of A . The combined meaning of all the natural number axioms is that $\mathbb{N} = A$. This makes sense, because the way we defined the set A is the most sensible way to describe the natural numbers rigorously.

While we're at it, let's clarify another potential source of confusion. In Chapter 1, we talked about inductive reasoning (or the inductive method), which is the main tool for acquiring knowledge in science. What is the relationship between inductive reasoning and mathematical induction? The answer is simple: *There is no connection between inductive reasoning and mathematical induction*. It's best to view the similarity in wording as an historical accident and leave it at that. When mathematicians use the term "induction," they almost always mean mathematical induction, not inductive reasoning.

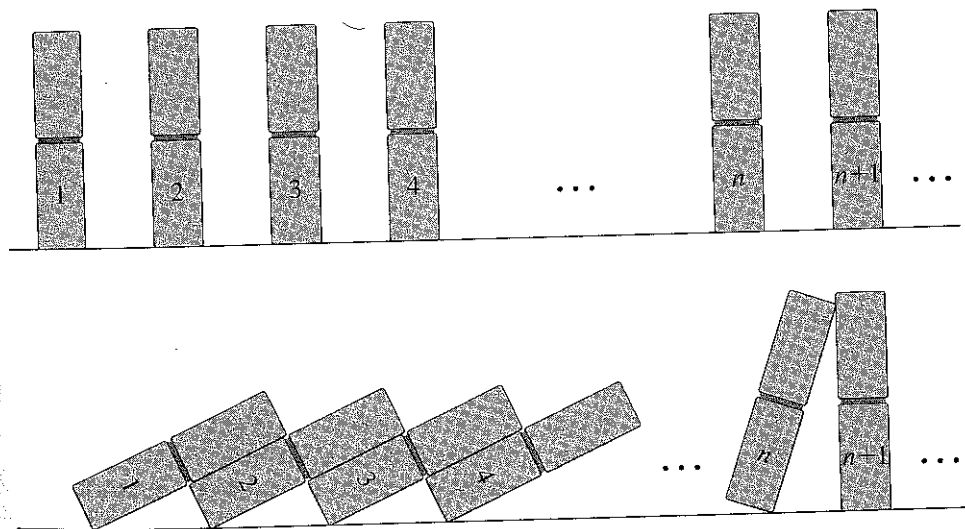


Figure 4.2 Domino image for mathematical induction

The Structure of Proofs by Mathematical Induction

In a proof by induction, the goal is to prove a statement of the form $\forall n P(n)$. By PMI and modus ponens, it suffices to prove $P(1)$ and $\forall n (P(n) \rightarrow P(n+1))$. So the first part of an induction proof is usually to prove $P(1)$. This is often a very short, obvious step. The second and major part is to prove $\forall n (P(n) \rightarrow P(n+1))$; this is called the **induction step** of the proof. As usual, to prove this quantified statement, it suffices to prove the unquantified implication $P(n) \rightarrow P(n+1)$. And this will usually be proved by conditional proof, which means the induction step starts with assuming $P(n)$. (This assumption of $P(n)$ is sometimes called the **induction hypothesis** of the proof). It is then required to prove $P(n+1)$ from this assumption. If these two parts can be done, then the desired statement can be asserted.

People sometimes are surprised by the fact that, in an induction proof, you get to assume $P(n)$, which is very similar to what you are trying to prove. But note that the statement to be proved is $\forall n P(n)$, whereas the induction hypothesis is just $P(n)$. It's vital that you *don't* put the quantifier $\forall n$ in your assumption for the induction step. Still, what you get to assume is quite close to what you are trying to prove; there is no other type of proof in mathematics where you can assume something so close to what you are trying to prove. But, as has already been pointed out, mathematical induction is a very special axiom, based on the particular arrangement of \mathbb{N} . If you understand the logic behind PMI, then there should be nothing surprising about how induction proofs are structured.

Now let's look at some examples of theorems proved by mathematical induction.

Theorem 4.9: Every natural number is a real number; that is, \mathbb{N} is a subset of \mathbb{R} .

Proof: We want to prove the statement $\forall n (n \in \mathbb{R})$. [Remember that n automatically denotes a natural number.] We prove this by induction, with $P(n)$ being " $n \in \mathbb{R}$."

We first must prove $1 \in \mathbb{R}$. This is implied by axiom V-9.

To prove the induction step, assume $n \in \mathbb{R}$. We already know that $1 \in \mathbb{R}$, and so by axiom V-1 (plus US), $n+1 \in \mathbb{R}$. ■

Theorem 4.10: The set \mathbb{N} is closed under addition; that is, $\forall m, n (m+n \in \mathbb{N})$.

[We've said that induction is used to prove statements of the form $\forall n P(n)$. But now we are asked to prove a statement that begins with two universally quantified natural number variables. How should we approach that? Do we have to do a separate induction proof for each of the variables m and n ? There are proofs in which it is necessary to do a double induction; but this is a complicated technique that is not required very often. Even with more than one variable present, it is permissible to use induction on just one of them. This simple approach works quite often, and we use it here.]

Proof: Let m be any natural number. We do induction on n only, with $P(n)$ being the statement " $m+n \in \mathbb{N}$." (So m is viewed as fixed.)

$P(1)$ is just a special case of axiom VI-2, so it holds.

To prove the induction step, assume $m + n \in \mathbb{N}$. We want to show that $m + (n + 1)$ is also a natural number. But note that m , n , and 1 are all real numbers, by Theorem 4.9. Therefore $m + (n + 1) = (m + n) + 1$ by axiom V-3 (associativity). It follows, by axiom VI-2 (and axiom III-4) that $m + (n + 1)$ is a natural number.

So we have $\forall n (m + n \in \mathbb{N})$, by induction. Since m was arbitrary, we can use UG to conclude $\forall m, n (m + n \in \mathbb{N})$, as desired. ■

Several of the exercises in this section are related to Theorem 4.10. Exercise 3 asks you to prove that \mathbb{N} is closed under multiplication, a proof very similar to Theorem 4.10's. And Exercise 22 asks you to critique an alternative proof of Theorem 4.10.

In our proof of Theorem 4.10, we used induction on n for any one given value of m ; so m was unquantified in the statement $P(n)$. It is possible to take a different approach, in which $P(n)$ is $\forall m (m + n \in \mathbb{N})$. Exercise 2 asks you to do this. Usually, it doesn't matter whether $P(n)$ contains universal quantifiers of variables besides n , but occasionally an induction proof becomes much easier if an extra universal quantifier is included in $P(n)$.

Mathematical induction is the main tool used to prove formulas for sums and products of sequences of numbers. We turn now to several examples of this. A rigorous treatment of this material requires the use of functions. For now we instead use the familiar ellipsis notation (three dots), which most mathematicians consider rigorous enough.

Theorem 4.11: $1 + 2 + 3 + \dots + n = n(n + 1)/2$ (that is, $\sum_{i=1}^n i = n(n + 1)/2$)

Proof: Since the variable n is unquantified in the statement of this theorem, we're supposed to prove it for all n , and we do so by induction.

For $n = 1$, the left side equals 1. [You have to be sensible about how you read something like $1 + 2 + 3 + \dots + 1$. The numbers to be added only go up to n , so the 2 and the 3 would not be included here.] And the right side equals $1(1 + 1)/2$, which also equals 1, so the equation holds.

For the induction step, assume

$$1 + 2 + 3 + \dots + n = n(n + 1)/2$$

Now add $n + 1$ to both sides, as Theorem 4.7 allows:

$$(1 + 2 + 3 + \dots + n) + n + 1 = n(n + 1)/2 + n + 1$$

The left side of this is just $1 + 2 + 3 + \dots + (n + 1)$, while the right side is easily simplified (by a couple of high school algebra steps) to $(n + 1)[(n + 1) + 1]/2$. ■

To prove the next theorem, we need to assume some basic properties of exponents, which are proved in Chapter 7.

Theorem 4.12: $1 + 2 + 4 + \dots + 2^{n-1} = 2^n - 1$

Proof: We proceed by induction on n . For $n = 1$, the statement says $1 = 1$, so it's true. For the induction step, assume $1 + 2 + 4 + \dots + 2^{n-1} = 2^n - 1$. Now add 2^n to both sides, and obtain

$$\begin{aligned} 1 + 2 + 4 + \dots + 2^n &= 2^n - 1 + 2^n \\ &= 2(2^n) - 1 \\ &= 2^{n+1} - 1 \end{aligned}$$

as desired. ■

Theorems 4.11 and 4.12 are particular cases of the formulas for the sum of **arithmetic** and **geometric** series. Here are the general formulas, with the proofs left as exercises.

Theorem 4.13 (arithmetic series formula): For any real numbers a and d and any natural number n ,

$$a + (a + d) + (a + 2d) + \dots + (a + (n - 1)d) = n[2a + (n - 1)d]/2$$

(Note that Theorem 4.11 is this formula with $a = d = 1$.)

Proof: See Exercise 4. ■

Theorem 4.14 (geometric series formula): For any real numbers a and r (provided $r \neq 1$) and any natural number n ,

$$a + ar + ar^2 + \dots + ar^{n-1} = a(1 - r^n)/(1 - r)$$

(Note that Theorem 4.12 is this formula with $a = 1$ and $r = 2$.)

Proof: See Exercise 8. ■

In Theorems 4.13 and 4.14, d stands for “difference” and r stands for “ratio.” An arithmetic sequence is one in which the difference between successive terms is constant, whereas a geometric sequence has a constant ratio between successive terms. Note that in both theorems, the formula given is for the sum of n terms of a sequence. (The word “series” always means a sum of terms.)

We turn now to a special case of the **division algorithm**, one of the most important basic results in number theory. We prove the general form of the division algorithm in Section 8.2. (We have occasionally referred to odd and even *integers*. When natural numbers are being discussed, it is best to define n to be even iff it is of the form $2m$ and odd iff it is of the form $2m - 1$, where m must also be a natural number in both cases. Do you see why this is preferable to saying an odd number is of the form $2m + 1$?)

Theorem 4.15: Every natural number is either even or odd.

Proof: By induction on n : for $n = 1$, note that $1 = 2(1) - 1$, so 1 is odd, by EG. Of course, this makes it even or odd. Now assume that n is even or odd. We use proof by cases. Case 1: Assume n is even. That means $n = 2m$, for some $m \in \mathbb{N}$. Then $n + 1 = 2m + 1 = 2(m + 1) - 1$, so $n + 1$ is odd. Case 2: Assume n is odd. That means $n = 2m - 1$, for some $m \in \mathbb{N}$. Then $n + 1 = 2m$, so $n + 1$ is even. ■

Compare this proof to Proof Preview 1 (Section 2.2). Exercise 11 asks you to prove that the “or” in this theorem is actually exclusive.

Theorem 4.16 proves a few important facts about natural numbers. It is tempting to just assume these “obvious” facts without proof, but that would be sloppy mathematical practice.

Theorem 4.16: (a) Every natural number is equal to or greater than 1.

(b) If $n > 1$, then $n - 1$ is in \mathbb{N} (and so $n \geq 2$).

(c) There is no natural number (strictly) between n and $n + 1$.

Proof: (a) and (b): These straightforward inductions are left for Exercise 12.

(c) We proceed by induction on n . For $n = 1$, we must prove that there is no natural number between 1 and 2. But part (b) says that any natural number greater than 1 is at least 2, so we are done.

For the induction step, assume there is no natural number between n and $n + 1$; we must show there is none between $n + 1$ and $n + 2$. We proceed by cases on an arbitrary natural number m . If $m = 1$, then m is not between $n + 1$ and $n + 2$, because part (a) guarantees that $n + 1$ is at least 2. On the other hand, if $m > 1$ and we assume that $n + 1 < m < n + 2$, then $m - 1$ would be strictly between n and $n + 1$, and $m - 1$ is a natural number by part (b). This would violate the induction hypothesis, and so it is impossible. ■

It takes a while to learn just how powerful and versatile mathematical induction is and how often it can be used. As a rule of thumb, when you are asked to prove a statement for all values of a variable and that variable can *only* be a natural number, you should consider using induction. Note that Theorems 4.11 through 4.15 all fall into this category, in that n has to be a positive integer for these statements to make sense. For example, an expression like $1 + 2 + 3 + \dots + 8.32$ is meaningless.

But there are also many theorems in mathematics for which it's much harder to see that induction must be used. There are also several variants of mathematical induction, and it is not always obvious that one of these is required. Numerous such situations are encountered later in this book, and it seems appropriate to list some of them here, for future reference:

(1) The well-ordering property of \mathbb{N} (Theorem 5.6) and results whose proofs use this property instead of ordinary induction (for example, Theorems 8.14 and 10.10).

(2) Theorems that seem to be about objects other than natural numbers, such as sets or polynomials, but that can somehow be classified in terms of a natural number

variable, and proved by induction (for example, Theorems 5.8 and 6.1). This important method is discussed further before Theorem 5.8.

(3) Induction proofs that begin at 0 or some other integer, instead of at 1. The induction proofs for the theorems mentioned in item 2 begin at 0. See also Exercises 12 and 13 of Section 5.3.

(4) Definitions by induction (Section 7.4).

(5) Double induction (Exercise 20 of Section 8.2).

(6) Complete induction (Lemma 8.20 and Theorem 8.21).

Mathematical Discovery Revisited

In Section 1.2 we discussed the important ideas of discovery and conjecture in mathematics. It was mentioned that in some situations the discovery process and the proof process are very closely linked, and in other situations they are totally separate. Induction is an excellent example of the latter situation. Because of the way induction proofs must be structured, it is impossible to begin an induction proof without already knowing what it is that you are trying to prove. Therefore, in most situations induction is not helpful as a discovery tool.

With this in mind, it is reasonable to ask what sorts of methods are used to discover new information about natural numbers. This is too complex a question to tackle in depth, but it is worth considering briefly. Note that Theorems 4.11 through 4.14 state formulas for the sums of various series. Mathematicians have many techniques, some of which are viewed as tricks (in a positive sense), for discovering such formulas.

Probably the most famous such trick relates to Theorems 4.11 and 4.13. Supposedly, when Carl Friedrich Gauss was about nine years old, his teacher was annoyed with the class and ordered everyone to add up all the numbers from 1 to 100. While the rest of the students toiled away, Gauss found the answer in a few seconds, without knowing Theorem 4.11. How? Quite simply (but ingeniously), he regrouped the numbers in the series, writing them in pairs as follows:

$$(1 + 100) + (2 + 99) + (3 + 98) + \dots + (49 + 52) + (50 + 51)$$

It then becomes a simple matter to compute the sum of these hundred numbers, and the same trick can be used to prove Theorem 4.13, the formula for the sum of an arbitrary arithmetic series (see Exercise 6). Other exercises in this section guide you through the discovery process for various sum formulas, including Theorem 4.14.

Does it seem to you that Gauss's computation should be considered a proof of Theorem 4.11? Most mathematicians would agree that it is a somewhat informal but essentially correct proof. However, they still usually prefer to write induction proofs for these sorts of formulas, even when the formula has already been derived by manipulating the terms of the series.

Carl Friedrich Gauss (1777–1855) was one of the greatest mathematicians of all time and also one of the most brilliant child prodigies in the history of the subject. One story has it that he found a mistake in his father's ledger book when he was three. Gauss's genius was recognized early; this enabled him to accelerate his education and be sponsored by the Duke of Brunswick.

Gauss's first important mathematical result, at the age of nineteen, was a proof that a regular 17-sided polygon can be constructed with straightedge and compass. This problem had evaded solution for over two thousand years. Five years later he completed his doctoral dissertation, the "Fundamental Theorem of Algebra." During his career he contributed to many branches of mathematics, notably differential geometry, number theory, and probability theory. His name is attached to many important concepts, such as gaussian curvature and the gaussian distribution.

Gauss chose to publish only the most significant of his results and only when they were quite complete, rigorous, and polished. The motto on his seal was "Pauca sed matura" ("Few but ripe"). But he left behind an enormous amount of valuable mathematical writing, in twelve volumes of diaries.

Gauss did important work in fields other than mathematics, notably astronomy and physics. He was one of the founders of the modern theory of electromagnetism, and the standard (metric) unit of magnetic strength bears his name. Many people would name Gauss as the last person to reach the very highest ranks of research in both mathematics and physics.

Exercises 4.5

(1) Evaluate by the formulas given in Theorems 4.13 and 4.14:

- (a) $1 + 3 + 5 + 7 + \dots + 399$
- (b) $2 + 5 + 8 + 11 + \dots + 200$
- (c) $1 + 2 + 4 + 8 + \dots + 1024$
- (d) $27 + 9 + 3 + 1 + \dots + 1/81$

(2) Write an induction proof of Theorem 4.10 in which $P(n)$ is $\forall m (m + n \in \mathbb{N})$ instead of just $m + n \in \mathbb{N}$.

(3) Prove that \mathbb{N} is closed under multiplication, that is, $\forall m, n (mn \in \mathbb{N})$. You may assume Theorem 4.10.

- (4) (a) Prove Theorem 4.13 by induction.
(b) The content of Theorem 4.13 can be expressed simply in words, as "The sum of an arithmetic series equals ...," where the "..." is an expression built up from the first term of the series, the last term, and the number of terms. Complete this statement.
- (5) Restate Theorems 4.12 through 4.14 in sigma notation.
- (6) (a) Complete Gauss's derivation of the formula in Theorem 4.11, as discussed at the end of the section.
(b) Generalize part (a) to derive Theorem 4.13. Make sure to include the possibility that the series has an odd number of terms.
- (7) (a) Noting that the first positive odd number is 1, and odd numbers differ by 2, find a formula for the n th positive odd number.
(b) Use induction to prove this formula.
(c) Derive a formula for the sum of the first n positive odd numbers.
- (8) Prove Theorem 4.14 by induction.
- (9) Here is the classic trick that can be used to derive Theorem 4.14 without induction:
(a) Start by writing $S = a + ar + ar^2 + \dots + ar^{n-1}$. We want to find a concise expression for S . What quantity could be multiplied by both sides of this equation so that all terms on the right side of the new equation, except one, would be the same as in the original equation?
(b) Now carry out the step determined in part (a), subtract the new equation from the original one, and solve for S .
- (10) In the geometric series formula of Theorem 4.14, if $|r| < 1$ and n gets large, what happens to the term r^n ? From this, make a conjecture about the formula for the sum of an *infinite* geometric series with $|r| < 1$.
- (11) Prove that the "or" in Theorem 4.15 is exclusive; that is, no natural number is both odd and even. (Don't just assume that 1 is not even or that $1/2$ is not in \mathbb{N} ; prove these things)
- (12) Prove parts (a) and (b) of Theorem 4.16.
- (13) Consider the series $1/2 + 1/6 + 1/12 + \dots + 1/n(n+1)$.
(a) Directly evaluate the sum of this series for several small values of n , and use these results to form a conjecture for the sum of this series in general.
(b) Using the fact that $1/n(n+1) = 1/n - 1/(n+1)$, find a trick for deriving the sum of this series without induction.
(c) Use induction to prove the formula for the sum of this series.

(14) Prove the surprising result (predicted in Exercise 6 of Section 1.2) that the sum of the cubes of the first n natural numbers equals the square of the sum of these numbers; that is,

$$1^3 + 2^3 + \dots + n^3 = (1 + 2 + \dots + n)^2 \quad \left(\text{that is, } \sum_{i=1}^n i^3 = \left(\sum_{i=1}^n i \right)^2 \right)$$

You may use Theorem 4.11.

(15) (a) Consider Theorem 4.11. If the right side of this formula is expanded, what is its leading (highest power) term?

(b) Repeat part (a) for the formula in Exercise 14, after using Theorem 4.11 to write the right side of that formula in closed form.

(c) On the basis of parts (a) and (b), complete this conjecture: For any natural numbers k and n , the sum $1^k + 2^k + \dots + n^k$ equals a polynomial in n whose leading term is ____.

*(16) The goal of this problem is to derive the formula for $1^2 + 2^2 + 3^2 + \dots + n^2$.

(a) Applying Exercise 15(c) to this series, write a polynomial with unknown coefficients for its sum. How many unknown coefficients must be included?

(b) Now substitute four small values of n into the result of part (a), to get a system of linear equations whose variables are the unknown coefficients. Surprisingly, it's OK and in fact advisable to use 0 as one of your values for n .

(c) Solve this system of equations to determine the exact polynomial that represents the sum of this series.

(d) Show that this polynomial equals $n(n+1)(2n+1)/6$.

(e) Now use induction to prove that $1^2 + 2^2 + 3^2 + \dots + n^2 = n(n+1)(2n+1)/6$.

(17) Prove that $n < 2^n$, for any natural number n . You may assume basic facts about the algebra of exponents, as was done in the proof of Theorem 4.12.

(18) Prove that $1 + 1/4 + 1/9 + \dots + 1/n^2 < 2 - 1/n$.

The next two problems provide examples of the **fundamental counting principle**, which is discussed further in Sections 6.1 and 7.6

(19) (a) At meetings of the Oxnard Pataphysics Club, every person present is required to say hello to every other person there, exactly once. Use trial and error and/or common sense to arrive at a conjecture about how many hellos are spoken at a meeting with n people present.

(b) Use induction to prove this conjecture.

(20) The English alphabet has 26 letters. Prove by induction that, for any n , there are 26^n n -letter words, where a "word" just means any sequence of letters.

(21) Prove the following calculus formulas, where n is any natural number. Use induction and the indicated formulas for each one:

$$(a) \quad \frac{d}{dx} (x^n) = nx^{n-1} \quad \text{Use the product rule and the derivative of } x.$$

$$(b) \quad \frac{d^n}{dx^n} (e^{kx}) = k^n e^{kx} \quad \text{Use the chain rule, the derivative of } e^x, \text{ and the}$$

derivative formula for a constant multiple of a function.

$$(c) \quad \frac{d^n}{dx^n} (xe^x) = (x+n)e^{kx} \quad \text{Use the product rule and the derivative of } e^x.$$

(22) Critique the following proof of Theorem 4.10. (If necessary, review the instructions for this type of problem in Exercises 4.2.)

Proof: First we must prove the theorem for 1. To do this, note that $1 + 1 \in \mathbb{N}$, because $1 \in \mathbb{N}$ (by axiom VI-1), and thus, so is $1 + 1$ (by axiom VI-2). Now assume the theorem is true for n ; that is, $n + n \in \mathbb{N}$. Then we must prove that $(n + 1) + (n + 1) \in \mathbb{N}$. But $(n + 1) + (n + 1) = [(n + n) + 1] + 1$, by the commutative and associative laws of addition. So the desired result follows by two applications of axiom VI-2.

(23) Critique the following proof.

Theorem: For every nonnegative integer n , $\sin n + \cos n = 2^n$.

Proof: For $n = 0$, the statement $\sin 0 + \cos 0 = 2^0$, which is true. For the induction step, assume $\sin n + \cos n = 2^n$. By substituting (that is, specifying) the expression $n + 1$ for the variable n , we get $\sin(n + 1) + \cos(n + 1) = 2^{n+1}$, as desired.

4.6 Hints for Finding Proofs

With the exception of Section 4.1, which dealt with different styles of proofs, the purpose of this chapter has been to explain the axioms and rules of inference that are commonly used in mathematical proofs. But we haven't said much about how to *find* proofs of statements. This section discusses how mathematicians go about proving things. It is as if we have just explained the rules of some game, like chess; this section starts to explain how to play the game competently.

Section 4.1 urged you to try to write proofs that are outlines or summaries of formal proofs. But most mathematicians write clear, logical proofs without ever consciously considering formal proofs. How is that done?

Here is another rule of thumb, based on a somewhat different (but not conflicting) perspective from that presented in Section 4.1: *A good proof of a statement should be a clear explanation of why the statement must follow from what you already know.* In other words, if you have a clear understanding of why a statement must be true, then you should be able to convert that understanding into a good proof of that statement. But to make that conversion requires careful analysis of your own understanding, and the ability to explain the sources of that understanding. Understanding why something is true entails more than merely seeing *that* it's true.

Is your understanding based in part on logic and common sense? Then the corresponding part of your written proof will use logical axioms and/or rules of inference. Is your understanding based in part on things you know about the subject matter of the statement? Then the corresponding part of your written proof will use proper axioms and/or previous theorems. Is your understanding based in part on knowing what certain words or symbols mean? Then your written proof will probably need to use the definitions of those words or symbols. Is your understanding of why the statement is true based on some reason why it *couldn't* be false? Then you will want to use indirect proof to prove the statement. And so on.

It is not always easy to analyze your own understanding in this way. But although mathematicians sometimes come up with proofs of difficult theorems without being consciously aware of how they did it, most would agree that proofs are based on understanding, and with enough analysis you can usually turn your understanding into a proof.

Gaining Insight into a Proof

We've been discussing how to convert your understanding of why a statement is true into a proof of that statement. But what if you are trying to prove something and you don't see why it's true? In fact, you may barely understand what the statement is *saying*, let alone that it's a true statement, let alone *why* it's true. Believe it or not, this happens frequently to all mathematicians, even the best. How are you supposed to prove something when you don't see what's going on?

Writing proofs when you have insight into the problem can already be hard; it is not something anyone can learn completely from one book or one course or in one year. Learning how to find proofs when you don't have that insight or understanding is that much harder, and it would be absurd to pretend that it's possible to learn this skill quickly. Nonetheless, it's possible to give some hints or guidelines for tackling proofs. Here are some of the more useful ones:

(1) *Analysis of the Structure of the Statement:* What is the logical structure (in terms of connectives and quantifiers) of the statement you are trying to prove? Answering this is not something that usually provides much mathematical understanding, but it may help you choose the right proof technique. Is the statement an implication? Then you almost certainly want to try conditional proof. Is it a negation? Then indirect proof might be a good try. Does the statement begin with one or more quantifiers? Then you probably need universal and/or existential generalization, remembering that existentially quantified variables are to be chosen as functions of the outer universal quantifiers. Tables 4.1 and 4.3 are meant to help you with this process.

(2) *Forward Reasoning:* This term refers to attempting to write a forward proof, perhaps more or less by trial and error. First you must start somewhere. How to start? If you are trying to prove an implication, conditional proof provides you with an assumption to start with. If not, you can try indirect proof, which also provides you with a starting assumption. If that seems inappropriate, you need to start with an axiom or

theorem. Which one? Obviously, you need to find an axiom or theorem that somehow is relevant to what you're trying to prove, but it can take some luck to find the most appropriate one.

Once you have one or more steps to start with, you have to go forward. How? One procedure is to look for an axiom or theorem that says that what you already have implies some other statement. Then you can use modus ponens to get a new step. Another procedure is to look for a rule of inference that would use one or more of the steps you have so far to get a new step. Of course, it is important not to be too random in generating new steps; rather, you need to constantly remember what you are trying to prove and try to keep getting closer to it.

(3) *Reverse Reasoning*: We discussed reverse proofs in Section 4.1, and this is another term for the same technique. If you can't see how to start your proof forward, start at the end and work in reverse. But, as was emphasized in our earlier discussion, you have to be careful when you do this. It's useless to find a new statement that is *implied by* the statement you want to prove; you need to find a new statement that *implies* the statement you are trying to prove. In other words, when trying a proof in reverse, you're always looking for statements that are *sufficient* for something you've already stated. Except for this one important difference, reverse reasoning is similar to forward reasoning.

Sometimes you can attempt a proof by a combination of forward and reverse reasoning. If you are fortunate, your two partial proofs will meet in the middle, and then you have a complete proof.

By the way, do not confuse the idea of reverse proofs with the method of indirect proof. Indirect proof is a valid rule of inference in which you assume the negation of what you're trying to prove. In a reverse proof, you start with the statement you're trying to prove, but you certainly are *not* assuming this statement!

(4) *Definition Unraveling*: This is a fairly simple process that can sometimes make a difficult-looking proof very easy. When the statement you are trying to prove involves defined words or symbols, it's always legitimate to replace them with whatever they are defined to mean. This usually makes the statement longer but may make it easier to understand, since you have replaced some words or symbols with simpler ones. Sometimes, you may be able to repeat this process two or more times, until the original statement has been unraveled into one involving only very basic symbols. At that point, you may see that the statement is one that is easy to prove; it may even be a tautology. Note that unraveling the statement you want to *prove* is a type of reverse reasoning. Unraveling statements that you are *using* in a proof—assumptions, axioms, previous theorems, and so on—is a type of forward reasoning. Both processes are quite legal.

As we see in Chapter 5, this procedure is very useful in basic set theory, a subject in which definitions tend to be used more than axioms. Often, in order to prove a biconditional involving sets, all you need to do is unravel all the defined symbols on both sides of the biconditional, and it then turns out that the two sides are logically equivalent.

(5) *Trying Special Cases:* This is an extremely important technique used by *all* successful mathematicians when they get stuck on a proof or problem of any sort. Most students do not realize how important it is; the sooner you learn to appreciate it and use it fully, the better off you will be. The idea is that most statements that are to be proved begin with at least one universal quantifier (even if it's not explicitly stated). So if you don't see how to prove the statement for an arbitrary value of whatever variable, first try to do it for one or more particular values. Although this can't constitute a complete proof, it's amazing how often doing some simple special cases provides enough insight to enable you to do the problem in full generality.

Example 1: Suppose you are asked to prove a statement of the form "Given any point in the domain of a real-valued function of several variables," This statement contains at least three variables (f for a function, n for the number of variables of f , and c for a point in the domain of f), which are understood to be universally quantified. If you don't see how to prove the statement, it may be because it's hard to visualize things in higher dimensions. So perhaps you should first try to prove it when $n = 1$, the simplest possible special case. If that works, you might then try $n = 2$, and if that works also, perhaps you can see how to do it for arbitrary n . What if you can't do it for $n = 1$? Don't panic; there are other variables to specify. Within the special case $n = 1$, you might try the proof for a simple particular function like $f(x) = x^2$, $f(x) = x$, or even $f(x) = c$. You might even want to choose a specific number for c .

Example 2: Suppose you are asked to prove some geometric statement involving an arbitrary triangle. Geometry problems can be very difficult. Instead of beating your brains out over the full problem, why not try it for a few very special triangles, like equilateral ones and isosceles right triangles. If you succeed with those, you might then try it for all right triangles and/or all isosceles triangles, or you might tackle the general problem.

Typically, when you succeed with one special case, you then try a more general or a harder special case (or the whole problem). But when you can't do one, you then try a more specific or a simpler special case.

No serious mathematician ever gives up on a problem or proof without trying at least one or two special cases, where applicable. The limitation "where applicable" is necessary because some statements have no variables from which to form special cases; but these statements are a minority. A mathematics instructor can do his students a service by refusing to help them with proofs until they have at least looked at a couple of special cases. But why lay this responsibility on your teacher? Establish this practice on your own!

(6) *Trying a Simpler Problem:* This is related to the trying special cases, but can be quite different. Suppose you are asked to prove something about any 3 by 3 matrix. If you first try it for some particular 3 by 3 matrix, that's a special case. But what if you first try it for some particular 2 by 2 matrix or perhaps all 2 by 2 matrices? This could not be called a special case of the problem, because 2 is not in any sense a special case

of 3. (That is, 3 is not a variable.) It's just a different problem. What you are asked to prove may not be true for or even apply to 2 by 2 matrices. But if it does, you might find it's much easier to see what's going on with 2 by 2 matrices, and solve the problem for those. Then the insight gained from that may help you do the 3 by 3 version. This technique is not quite as powerful as the use of special cases, but it's still useful.

Suggestions for Further Reading: The references listed at the end of Chapter 2 all include some coverage of the major ideas of this chapter: formal systems, proofs, rules of inference, and so on. So do Stoll (1979) and Wilder (1965). For an in-depth treatment of the method of mathematical induction, see Sominskii (1961). For elaboration of the discussion in Section 4.6, you are encouraged to read the lucid observations on mathematical insight in the classic works by Polya (1945, 1954, and 1965).