Handbook of Mathematical Proof

Edward D. Kim
University of Wisconsin-La Crosse

Edward D. Kim
ekim@uwlax.edu
Department of Mathematics and Statistics
University of Wisconsin-La Crosse
1725 State Street
La Crosse, WI 54601
https://edward-kim-math.github.io/

Printed in the United States of America

Kim, Edward D.
Handbook of Mathematical Proof

Contents

Preface to the student

This is a different kind of class and math book. You may be used to mathematics where you compute something, whether you solved an equation, differentiated a function, simplified an expression, determined a limit, or evaluated an integral. You may have worked on applying theorems (such as the Squeeze Theorem for limits) and have built strategies for a certain problem type (sorting out when to integrate using substitution versus when to integrate by parts).

The majority of your mathematical experience so far may have been computational in nature. However, when you rely on theorems from calculus, how do you know that what you rely upon is solid? If this text is in front of you, it is because you are now at a place in your mathematical career where computation can be put aside for a moment so that you can learn how to read and write mathematical proofs.

Saying the word "proof" may sound scary to you. In fact, due to previous experiences, you may have some extremely negative feelings associated to proofs. In the past, you may have dealt with ε-δ proofs in a first semester calculus class. You may have preferred computing derivatives of functions over applying the Squeeze Theorem, the Intermediate Value Theorem, or the Mean Value Theorem. Applying convergence tests for infinite series may have seemed like such a strange experience in calculus. The portions of a class where you were expected to "do proofs" may have been "coached" in the following sense: you "knew" when an exam question needed the Intermediate Value Theorem and you could even apply the theorem for full credit, but you never quite felt sure about what you were doing.

If that resonates with you, then this handbook should be really refreshing. As a student who has taken at least a semester of calculus, you're now at a stage where a complete foundation in mathematical proof can and should be discussed. There is a complete framework that needs to be learned. If you're reading this, someone gave you this text with the confidence that *you* have the background and technical skills to learn this complete framework. (If you're discovering this text on your own, successful completion of one semester of calculus is about the right level of experience for reading this text.)

This text is, therefore, an invitation to you! No longer do proofs need to be some big, scary monster. You (yes, *you*) can understand everything there is to know about proofs. By going through this handbook, you will learn all that is necessary to prove and use mathematical statements. This will take some work, but you should read through this handbook *thoroughly* to get there. It is tempting to keep a highlighter close at hand to highlight what you consider "important." However it is all important, so please read every sentence. It would be tempting to read only the contents of the framed boxes. However, the boxes are provided to use their color functionality (red for warnings, blue for definitions, and so on) and so that using the numbering, it is easy to refer back to previous ideas: when a previous warning or definition is referenced, be sure to go back and read the referenced thing, and create the connections for yourself that need to be made.

This handbook is really written for you, an individual who has had past success in a computation-based mathematics course. Mathematics has a level of consistency. For example, recall the struggle involved with using $a(b+c) = ab + ac$ to rewrite $x^2(ac + b^3)$. This level of consistency is gained by taking classes such as algebra and calculus. Quite surprisingly, learning mathematical proof does not require you to recall all sorts of facts from algebra or calculus. Instead, you will need to apply a certain level of consistency. In the past, this consistency of methods was used to compute a derivative or apply an algebra identity. Now, this consistency of thought is used to combine old truths into new truths in what is called a proof.

Practice speaking and writing like a mathematician. You must pay attention to the terminology used. This will be hard at first. Pay attention to mathematical vocabulary. What are the nouns? What are the verbs? What are the adjectives? The presentation of your mathematical work in a past calculus class may have involved no sentences at all! Proofs are written in complete sentences. Definitions are written in complete sentences. In fact, see the webpage `http://thatsmathematics.com/mathgen/` which generates a random (satirical) math paper.

The webpage mentioned may seem silly, but if you show this to any mathematician, they would assure you that the text generated, while actually being gibberish, truly sounds like mathematics. This is more than a silly experiment: a lot is to be learned from this. In algebra and calculus, your instructors likely emphasized the difference between an expression and an equation, since these are different (but related) creatures. For example, $3+2$ is an expression, while $3+2=5$ is an equation. You are aware that equations can have variables, such as $3+x=5$. An equation with a variable may have solutions. For example, you know that $3+x=5$ has the solution $x=2$. The purpose of bringing up such an elementary equation is not about the procedures involved in solving the equation: you can solve such an easy equation without coffee. The purpose of this is to highlight the phrase $\boxed{\text{is a solution of the equation}}$ in the sentence "$x=2$ is a solution of the equation $3+x=5$."

You may have been successful in previous math classes while allowing the grammar and usage slip by. Beyond calculus, this method won't lead to success anymore. This handbook will often highlight certain specimens of writing in a box, as was used above in $\boxed{\text{is a solution of the equation}}$. Pay careful attention to how mathematics is written and spoken. Many of these issues will be highlighted in Habit boxes and Language Discussion boxes throughout the text.

Please read this handbook (and any future proof-based math book) very carefully. Every word counts. Every symbol counts. Writing in complete sentences is the community-wide standard for writing definitions and proofs in mathematics. It will seem intimidating at first, but the only way to get better at writing mathematics in the manner your instructor expects is to practice thoroughly reading mathematics. (If it helps you, read each sentence aloud.) In second-semester calculus, there is a test about infinite series which mentions both $\sum |a_n|$ and $\sum a_n$. More specifically, knowing information about $\sum |a_n|$ allows you to conclude something about $\sum a_n$. Students in calculus may not have paid close attention to the presence or absence of the absolute values, but would have been more successful if they had. When reading, please try to pay *that* close attention to notation. Though the example is very late in the book, see Example 846 which discusses the situation of k vectors in n-dimensional space. The point here is to say that a k is used and an n is used, so they may be the same value, or they may be different. In other settings, the same variable will appear more than once. If that occurs, the variable must take on the *same* value each time it occurs.

The book has a number of unique, fun features. The methods of proof are presented visually using flowcharts. Minesweeper is used as a way of understanding proof by cases and proof by contradiction. Students in previous semesters have really enjoyed exercises where a proof to a theorem such as Theorem 283 must be developed. As silly as this theorem sounds, these are the best ways to simulate what is covered in math classes which this handbook leads toward.

Preface to the instructor

While the preferred method of using this handbook is to read from cover to cover, this guide was written recognizing that some students will need to use this as a reference (in a later class), there are lots of references (by number) back to previous material. This may seem a little overdone to the reader going through this handbook methodically. (Then again, if a reader rolls their eyes while thinking, "You really didn't need to recall that: I knew what to do!" then they are getting affirmation of their understanding.)

Some students will have passed a course on how to do proofs and find themselves in a topology, real analysis, or abstract algebra class and feel like they still don't understand proofs yet. These students will find the constant use of references by number helpful. Other students need a brief primer on proof because they're finding themselves in a linear algebra class where some small proofs are required, but they have no training. These students will also find the heavy use of referencing useful.

For day-to-day proofs, mathematicians do not think about induction from the formal statement of the Principle of Mathematical Induction. So, this formal statement is left out. A section of this handbook covers how mathematicians think about the Pigeonhole Principle in actual proofs. While a certain level of formality is needed, there is a strong tendency to explain mathematics in the way mathematicians *truly* think about mathematics: stuffiness has been removed where possible.

Designing any guide or text on mathematical proof leads to a discussion of sets first or propositions first. Sets can be introduced first from the computational viewpoint of algebra on sets, but would need to be revisited after propositions are discussed in order to understand a set like $\{x \in \mathbb{R} : x < 30 \text{ or } x \text{ divides } 25\}$, due to the use of the word "or," among other issues. It would seem natural, then, to discuss propositions first, but then the set memberships in a statement such as $\forall x \in \mathbb{R}\, \exists y \in \mathbb{Z}\, [z > x]$ often get postponed, so students become first exposed to quantified statements such as $\forall x\, \exists y\, [z > x]$ which do not look like actual statements encountered in mathematics texts. We have adopted a bit of a *both* first approach: Chapter 1 introduces just the bare necessities of both sets and propositions so that a statement like $\forall x \in \mathbb{R}\, \exists y \in \mathbb{Z}\, [z > x]$ never has to be introduced.

The goal of Chapter 3 is to quickly get the reader to actual proofs. Section 3.1 is organized in a non-traditional manner to achieve this goal. Students will find it confusing to discuss proving existentially-quantified statements and using existentially-quantified statements, and proving universally-quantified statements and using universally-quantified statements. Because the definitions of even and odd use existential quantifiers, the rules of inference for universally-quantified statements was postponed to a later subsection of Section 3.1.

Because proof is so different from students' computational courses, there are frequent warnings, language discussions, and discussions regarding habits. To prepare students for their subsequent proof-based mathematics course, the proofs in Sections 4.8 and 5.1 are particularly challenging. Some of the presentation in Section 5.1 is perhaps a bit unusual: the more informal idea that countability should be thought of as enumerability is de-emphasized so that fairly challenging yet formal proofs can be practiced.

Acknowledgments

I am grateful for the support from my colleagues Robert Allen, Tushar Das, Whitney George, and Nathan Warnberg in the development of this handbook. A huge thank you is due to Kaisa Crawford-Taylor for extensive comments, corrections, and suggestions in an early version of this manuscript. I also wish to thank Kean Fallon, Mitch Haeuser, Sam Haeuser, Madisyn Janusiak, and Avery McLain for conversations in the brainstorming phase of this manuscript. Some of the work presented here grew out of conversations with fellow Project NExT Gold'14 dots. Comments from MTH 309 in Spring 2019 and MTH 225 in Fall 2019 have been helpful in improving the handbook. I thank Jenn Kosiak and Megan Litster for conversations on active learning that have influenced this handbook. I am also grateful for Deb Hoskins for our conversations on teaching and learning which have influenced this handbook as well. The danger of writing an acknowledgments section is the error of omission. If I have missed you, I apologize, and ask that you kindly ask for acknowledgment in the next iteration of the handbook.

Chapter 1

Introduction

Before getting too far into the details, a word of encouragement: you are best suited to learn this material. The contents you'll see are really different compared to a math class had primarily focused on computation. However, when you've been in computation-foused math classes, you've been succesful with pattern recognition and seeing how things fit together. For this material, plesae bring those same habits to the table: pattern recognition and seeing how things fit together. The material is going to feel very different at first, and it will require your patience, but we will build everything needed from the ground up. Finally, because the material is going to look so different (we will see lots of sentences, and include some goofy, humorous sentences in the process, because, well, why not?), please be advised that you shouldn't "wait to take notes until the real math shows up." Please note that, in a broader sense, all of what we introduce here is math. Mathematics is more than just computations with numbers (and if we can just admit now that the variables in algebra are numbers, then tha's what algebra was too). Mathematics is a style of argumentation with water-tight thought arguments.

1.1 Definitions

Definitions are the most fundamental sentences for mathematicians. Without definitions, we could not prove anything. Definitions are probably the most neglected aspect by students new to mathematical proof. You must commit to understanding each definition thoroughly. You must have more than "just the gist" of definitions. We are about to discuss how definitions appear in sentences, but don't let that give the false impression that definitions in math are ambiguous. Instead, definitions in math are very technical and precise, in the same way that sine of $\frac{\pi}{6}$ is exactly $\frac{1}{2}$, not 0.49.

Definitions in mathematics are always written in complete sentences. In contrast, definitions in an English language dictionary are not written in complete sentences. Sometimes, the term being defined is more than one word, so look to see if there's a larger phrase being defined. (Often, it will be necessary to incorporate the new word as part of a more complete phrase.) In mathematics, the word or phrase being defined *must* be written in the definition. In contrast, definitions in an English language dictionary do not write the word being defined.

Habit 1

When reciting a definition, be sure to write the word (or phrase) being defined.

Notice the two differences to dictionary definitions: complete sentences and mentioning the word (or phrase) when stating the definition. Mathematicians follow these conventions because stating mathematical definitions in this way is how mathematicians communicate the part of speech of the new word, in addition to how the new word is used in context.

Habit 2

Every time you encounter a new definition, determine if the thing being defined is a noun, an adjective, a verb, etc.

One thing that catches students off guard in a class that introduces a proof is the expectation that teachers have that students will need to recite definitions in written form. That is, is is very typical for quizzes and tests in a class based on this book to have students write a definition for terminology introduced in class in a closed-book closed-note setting such as a quiz. It is besd to avoid "memorizing" a definition word-for-word and symbol-for-symbol. Instead, focus on inventorying all of the essential points about a definition when first seeing one, and then be sure that all of the essential details get included when reciting a definition. When you write a definition, your instructor expects you to write in complete sentences and include the word (or phrase) in your sentence(s) because your written definition must clearly convey whether the term being defined is a noun, adjective, verb, etc. As a brief review, nouns are persons, places, or things (such as your math instructor, Central Park, or your favorite board game), adjectives are descriptive words (such as fast, green, or smooth), and verbs are actions (such as runs, sings, or talks). This process takes getting used to, so keep working on it!

If the term being defined is a noun, your writing should make clear whether the noun is a vector, or a number, or a matrix, or a set, or an equation, etc.

If the term being defined is an adjective, what is the noun being modified? For example, the definition of continuous from calculus applies to functions. Keeping track of what adjective applies to which noun is a *critical* aspect of mathematical language and totally unlike how adjectives work in English. While you'd probably never use the adjective $\boxed{\text{sleepy}}$ in front of the noun $\boxed{\text{car}}$, it is technically okay. As far as English grammar is concerned, the phrase $\boxed{\text{sleepy car}}$ is allowed. You would most likely never use the adjective 'sad' to describe the noun 'pickle,' but this is technically okay. The mathematical adjective $\boxed{\text{continuous}}$ can *only* be used for a $\boxed{\text{function}}$. So, you can speak of a $\boxed{\text{continuous function}}$, but the phrase $\boxed{\text{continuous equation}}$ doesn't make sense. As another example that you're familiar with, saying that "two lines are parallel" has meaning, but saying that "two points are parallel" does not make sense.

Warning 3

Always be mindful of what noun an adjective may modify. Never use an adjective to modify an inappropriate noun.

Warning 4: Do not ignore language issues and grammar

Be sure to write definitions in complete sentences. Think of whether a word being defined is a noun, adjective, verb, etc. and clearly convey this when you write definitions. Adjectives should only be applied to *appropriate* nouns. Do not ignore this advice regarding language. You might have been quite successful in a previous math class without paying attention to language, but it is important to now start paying attention to definitions in a very deep way.

You *can* succeed at using mathematical language appropriately – many before you have been successful at this – but only if you try! Be willing to work on mathematical language. The author of this handbook understands that this is probably new to you, so there are boxes about language discussion: read and re-read these.

After reading each definition, pause and think of examples and non-examples. Do not just rely on the examples and non-examples from a book. Try to make your own examples and non-examples. Come up with good examples and non-examples. What makes a good example and good non-example? For an adjective (example: continuous function), think of a function which *is* continuous and a function which is *not* continuous. For a noun, ensure your non-example just barely breaks the defining property. For example, when making a function that is not continuous, try making a function that is not continuous just at one input. In fact, it is good to try to think of several non-examples for each definition. Often, to have an example of something of a definition, several things must all be true. So, in each non-example, try to have all but one of the things be true. (And in the next non-example, have a different thing be untrue.)

If your previous math classes (algebra, calculus, etc.) have been more computational in nature, then when you think of the word “example,” you may be thinking of a very different type of thing than when your instructor uses the word “example” in a math class where computation is not the main focus. In calculus, an example is a worked out calculation of a derivative or an integral. In classes based on this handbook, an example is a description of an instance (or occurrence) of an object satisfying all the parts of a definition.

Do not memorize a definition “word for word.” When given a definition, understand what it is saying. Then, work on reconstructing a sentence that captures all of that meaning. (While you should not memorize anything “word for word” do note that the order of words matters. For example, “AJ found a gold person’s necklace” versus “AJ found a person’s gold necklace” communicate very different things.)

In the subsequent sections, we will introduce definitions, and illustrate some of the key pieces of advice, habits, and warnings.

1.2 Mathematical language

It is tempting for new mathematicians to write phrases like “The sphere, S, is centered at the origin.” Leave out the commas. It is correct to write “The sphere S is centered at the origin” instead. Why is this? In English, consider the following sentence: “My dog Thor likes to play frisbee.” There is no need to surround Thor (the name of the dog) by commas. Likewise, there is no need to surround S (the name/label of the sphere) by commas. Putting “Thor” immediately after “dog” makes it clear that Thor is a dog. Similarly, putting S immediately after “sphere” makes it clear that S is a sphere. There is something very important to learn from this:

Language Discussion 5: Notation placement

Notation for an object is placed immediately after the noun which describes the object, with no surrounding commas.

Placement of notation within a complete mathematical sentence is challenging at first. However, clear understanding of written mathematics (whether you are the one reading or you are the one writing) requires this skill. It will be helpful to work on this early in the process while there are still not yet too many moving parts.

1.3 Main objects

Every math class has to start somewhere: a course in Euclid’s geometry begins with Euclid’s five axioms. For this text, we introduce two types of objects. Everything else that we will study is built on top of these two objects.

1.3.1 Propositions

The first of two primitive objects we introduce is the proposition:

Definition 6: Proposition

A **proposition** is a declarative sentence that is true or false, but not both.

Language Discussion 7

Recall from Habit 2 we should ask: Is a proposition a noun, an adjective, or a verb? The definition says that a proposition is a sentence, so a proposition is a noun.

If we only consider propositions, we are talking about sentences which states something in a factual manner, whether or not the actual sentence is true or false.

Example 8. *The sentence "What time is it right now?" is not a proposition because this is a question.*

Example 9. *The sentence "George Washington was the first president of the United States." is a proposition.*

Example 10. *The sentence "George Washington was the first president of Mali." is also a proposition, even though this proposition is false.*

Example 11. *The sentence "$2 + 3 = 5$" is a proposition. This proposition is true.*

Example 12. *The sentence "The function $f(x) = \frac{1}{x}$ is continuous on $(-\infty, \infty)$" is a proposition. This proposition is false, since f is not continuous at $x = 0$.*

Pause and think of your own examples and non-examples. To make sure your examples cover all the possibilities, think of propositions which are true, and think of propositions which are false.

We will thoroughly cover propositions in Chapter 2 after a brief introduction to sets. A thorough coverage of sets is in Chapter 4.

Exercise 13. *Consider each sentence. Is each sentence a proposition? Why or why not?*

- *January 1, 1990 was a Tuesday.*
- *Fridays are better than Mondays.*
- *Would you like soup or a sandwich?*
- *Arizona is a continent.*

Exercise 14. *Consider each sentence. Is each sentence a proposition? Why or why not?*

- *Elephants are mammals.*
- *Wednesdays are part of the weekend.*
- *Red is an ugly color.*
- *Where did the day go?*

1.3.2 Sets

The second of two primitive objects we introduce is the set:

Definition 15: Set

A **set** is a collection of objects. The objects in the set are called members or elements.

Language Discussion 16

Recall from Habit 2 we should ask: Is a set a noun, an adjective, or a verb? The definition says that a set is a "collection," so a set is a noun.

Each object in a set is called a member or an element. If A is a set, we will write $\boxed{m \in A}$ to mean that m is a member of A and write $\boxed{m \notin A}$ to mean that m is not a member of A. Since the words "member" and "element" are synonymous, $m \in A$ means that m is an element of A and $m \notin A$ means that m is not an element of A.

Warning 17

A set is not a proposition. Therefore, a set is neither true nor false.

Though a set is not a proposition, the statement $\boxed{m \in A}$ is either true or false, but not both. So $\boxed{m \in A}$ is a proposition, while A is not. Thus $m \in A$ is either true or false, but not both (depending on what m and A are), while A, being a set, can neither be true nor false. If m is not an element of A, we may write $m \notin A$.

Habit 18: Is the noun a proposition or a set?

When encountering a new definition, if the new term being defined is a noun, determine whether what is being defined is a proposition or a set.

A set is not a proposition, and a proposition is not a set.

One way to define a specific set is in words. The typical language for this will be "Let [a new letter] be the set of all [...]"

Example 19. *Let F be the set of all fruits. In this case, the elements of the set F are all fruits. So, apple, banana, orange, dragon fruit, and pear are all elements of the set F.*

Thus, we can write apple $\in F$, banana $\in F$, orange $\in F$, dragon fruit $\in F$, and pear $\in F$. However, car, truck, bus, and bicycle are not elements of the set F. Thus, we can write car $\notin F$, truck $\notin F$, bus $\notin F$, and bicycle $\notin F$.

Before moving on to another example, us talk about words that we can say for $x \in F$, where here x represents an unknown thing. We could say "x is in the set F" and while this is accurate, this doesn't bring up what we know about F. We could say "x is in the set of all fruits" and this is also technically accurate, since F is the set of all fruits, but what this version lacks is plain language. It is much clearer if we write "x is a fruit." Is this the right grammar? Let's go back and look: we wrote apple $\in F$ in our example, and by taking $x \in F$ to be "x is a fruit" in words means that apple $\in F$ translates to "apple is a fruit." That works!

Example 20. *Let U be the set of all car brands. Then for example Ford $\in U$ and Nintendo $\notin U$. Translating $z \in U$ in plain English, we'd write "z is a car brand." For example, Ford $\in U$ translates to "Ford is a car brand."*

Example 21. *Let W be the set of all countries. Then Canada $\in W$ and California $\notin W$. To say $m \in W$ in plain English, we'd write "m is a country."*

Exercise 22. *Define a set called S. Write the set you define in a complete sentence following the format of: Let S be the set of all...*

Then, write in plain English what $n \in S$ means, and write in plain English what $m \notin S$ means.

Besides defining sets using words, there are several ways to define sets using notation. One of the ways of writing notation to define sets is a comma-separated list of elements surrounded by curly braces.

Example 23. *Let us consider a set where the only elements are apple, banana, orange, dragon fruit, and pear. These five fruits can be the five elements of a set, and that set can be described in standard mathematical notation by writing $\{$apple, banana, orange, dragon fruit, pear$\}$.*

The difference between this example (a set with just fix specific fruits) and our earlier example labeled with F is that F has *all* the fruits, while the set we defined here has just *some* of the fruits.

Example 24. *Let $J = \{Chevy, Toyota, Subaru\}$. Then Chevy $\in J$. In fact, from our set U earlier which was the set of all car brands, Chevy $\in U$ as well. While Fiat $\in U$, we have Fiat $\notin J$, since J only has three car brands.*

Example 25. *For instance, $\{2, 4, 6, 8\}$ is a set. For convenience, we may name this set by writing $S = \{2, 4, 6, 8\}$. However, $\{2, 4, 6, 8\}$ is a set whether it is named or not.*

Example 26. *Let $T = \{1, 4, ☺\}$. Then $4 \in T$ and $☺ \in T$, while $2 \notin T$.*

A set is not the same as an element. What's the difference between a set and an element? As an analogy, think about the difference between a city and a resident. A city is comprised of many residents. Similarly, a set is comprised of elements. The difference between a set and an element is like a club and its members. If AJ Smith belonged to the Chess Club, you wouldn't say that AJ *is* the Chess Club: you'd say that AJ is a *member of* the Chess Club.

On one extreme, take a city like Seattle, which has over one million residents. If AJ is a resident of Seattle, you wouldn't say that AJ *is* Seattle: instead, you'd say that AJ *belongs to* Seattle. For the other extreme, consider the set $C = \{r\}$. Then C is a set, while r is an element of the set C. In this example, there is only one element that belongs to C. Yet, we cannot write $r = C$ when we need to write $r \in C$ instead.

Some sets which occur are so common that we will introduce them here and use them throughout the handbook. The set of integers is denoted by $\mathbb{Z}$. An integer is any real number whose expansion past the decimal point is all zeroes.

Example 27. *Thus,* $3 \in \mathbb{Z}$ *and* $-3 \in \mathbb{Z}$ *and* $0 \in \mathbb{Z}$*, however,* $\frac{3}{2} \notin \mathbb{Z}$*.*

The set of all rational numbers (or the set of rationals) is denoted $\mathbb{Q}$, where the notation is inspired by the word quotient. A rational number is any number which is the quotient of an integer by a non-zero integer.

Example 28. *Thus* $\frac{3}{2} \in \mathbb{Q}$*. We also have* $\frac{12}{8} \in \mathbb{Q}$*, and the peculiarity that this is really the same rational number (in numerical value) is a concern that is not really addressed here: that matter is sorted out in abstract algebra.*

The set of all real numbers (or the set of reals) is denoted $\mathbb{R}$.

Example 29. *So,* $3 \in \mathbb{R}$ *and* $\pi \in \mathbb{R}$ *and* $\sqrt{\pi} \in \mathbb{R}$*, but* $\sqrt{-1} \notin \mathbb{R}$*.*

The set of all complex numbers (or the set of complexes) is denoted $\mathbb{C}$. A complex number is the sum of a real number with a real number times the imaginary unit i, where $i = \sqrt{-1}$.

Example 30. *Thus* $3 + 4i \in \mathbb{C}$ *and* $e^\pi - 5i \in \mathbb{C}$*.*

In this handbook, we will avoid saying natural number. While all mathematicians are in agreement that every positive integer is a natural number, there is disagreement over whether 0 is a natural number. Thus, for some mathematicians, $0 \in \mathbb{N}$ while for others, $0 \notin \mathbb{N}$. When picking up a textbook for the first time, you should read to figure out whether the author includes 0 as a natural number or not. You might find an author write $\mathbb{N} = \{0, 1, 2, 3, \dots\}$ early in a text (perhaps on a reference page of symbols/notation) where another author might write $\mathbb{N} = \{1, 2, 3, \dots\}$. For the purposes of reading a particular text, go with how the author defines $\mathbb{N}$. Neither is "right" or "wrong" in a general or moral sense: however the author defines $\mathbb{N}$ is right for that text. This follows the general principle of needing definitions, as discussed in Section 1.1.

To say a bit more about this, those who say that the set of natural numbers is $\mathbb{N} = \{1, 2, 3, \dots\}$ are likely to call $\{0, 1, 2, 3, \dots\}$ the set of whole numbers, while those who say that the set of natural numbers is $\mathbb{N} = \{0, 1, 2, 3, \dots\}$ are likely to call $\{1, 2, 3, \dots\}$ the set of positive integers. Various authors will define sets such as $\mathbb{Z}_+$ and $\mathbb{Z}_0$ to include or possibly exclude the number 0. Again, refer to how the author defines these sets in a particular book.

For better or worse, we will rectify the inconsistency in this text by providing unambiguous notation which is clear from context, even if it is a bit stuffier to write this way: using $\mathbb{Z} = \{\dots, -2, -1, 0, 1, 2, \dots\}$ as the inspiration of our notation, we will write $\mathbb{Z}_{>0}$ to mean the set $\{1, 2, 3, \dots\}$, and we will write $\mathbb{Z}_{\geq 0}$ to mean the set $\{0, 1, 2, 3, \dots\}$. While writing $\mathbb{N}$ is quicker than writing either $\mathbb{Z}_{>0}$ or $\mathbb{Z}_{\geq 0}$, our notation has the benefit of being clear and unambiguous.

Some of the early discussions in this handbook are easier if we use $\mathbb{H}$ to denote the set of all humans. Since I, the author of this text, am a human, I am a member of the set $\mathbb{H}$. Since you, dear reader, are also a human, you are also a member of the set $\mathbb{H}$. An important lesson here is that not all sets have the property that the elements are numbers. The set $\mathbb{H}$ is a set of people. In light of this, it is good practice to follow this habit:

Habit 31: What types of objects are elements of a set?

Every time you encounter a set, ask yourself what kind of object are elements of the set.

The habit just mentioned is a clue into the following important fact: the elements of a set are not necessarily numbers. For example, $\mathbb{H}$ is a set, even though there are no numbers in this set.

Whenever you have determined you have a set, determine what kind of objects belong to your set. As evidenced by $\mathbb{H}$, not all sets in this class will be sets of numbers. In this handbook, you will see a set of numbers, a set of humans, a set of functions, and so on. You will even see a set of sets! For example, $\{\mathbb{Z}, \mathbb{Q}, \mathbb{R}\}$ is a set of sets, while $\mathbb{Z}$ is a set of numbers. If you have a set S and you know that each element of S is a function, then you should use the phrase "S is a set of functions."

On the other hand, if you know that S is a set, refer to the set S as S alone. In this case, avoid saying "set of S." (The phrase "set of S" sounds like a set of sets.)

Let's recall interval notation. Recall that writing $[3, 7)$ means we want all real numbers between 3 and 7, though we are including 3 but excluding 7. Well, $[3, 7)$ is a set. More specifically, in keeping true to Habit 31, notice that $[3, 7)$ is a set of numbers, or more specifically, a set of real numbers. Note that $\pi \in [3, 7)$.

Habit 32: Handwriting and typing for sets

For the sets $\mathbb{Z}, \mathbb{Q}, \mathbb{R}$, and $\mathbb{C}$ respectively, do not write Z, Q, R, and C respectively as substitutions, whether in handwritten work or in typewritten work. This may seem like an odd request, but it's a good habit to form now, since there are proofs where both R and $\mathbb{R}$ appear. When typing in LaTeX, type `\mathbb{Z}` in math mode after including the `amsfonts` package to get $\mathbb{Z}$.

1.4 Connection to the past

This chapter introduced you to the two main objects of mathematical proof: propositions and sets. The next chapter will introduce special kinds of propositions. (For example, the conjunction is a proposition.)

While conjunctions are propositions, sets are not propositions, and propositions are not sets. Sets and propositions are different. Similarly, in the past, you had to consider equations and functions, which are *different*.

Example 33. *For example, let f be the sine function from trigonometry. (So $f(\frac{\pi}{4}) = \frac{\sqrt{2}}{2}$.) Consider an equation in the variable x, for example, $3 + x^2 = 28$. Then the function f we have described is not an equation, and the equation $3 + x^2 = 28$ is not a function.*

Definition 34

An equation is **consistent** if it has a solution.

Here's afriendly reminder to apply Habit 2. That is, because we just came across a new definition, let's get in the habit of determining the part of speech of the new word: the word consistent defined here is an adjective. Challenge yourself to pay attention to the part of speech in each new definition.

Example 35. *The equation $3 + x^2 = 28$ is consistent because $x = -5$ is a solution.*

Example 36. *The equation $x + 5 = x$ is inconsistent because it has no solution.*

What about the function f, the sine function? Is f consistent? Recall Warning 3, that we must be mindful of noun an adjective will modify. The definition of consistent shows that this adjective is *only* to be used on equations. However, f is *not* an equation, so it makes no sense to even *ask* if f is consistent or not!

Each adjective defined in mathematics only applies to a specific type of noun, which will be mentioned in the definition. This is very different from how adjectives are used in English: while it is natural to say "windy road," it isn't grammatically incorrect (or devoid of meaning) to say "windy speech." While you can say "that speech was windy" and have some meaning (albeit awkward), you cannot say "the function f is consistent." Similarly, yellow is a color, and speech is a noun, but we wouldn't be making any sense by saying "yellow speech." Spoken everyday language still allows for this kind of thing (and is much more flexible about which adjectives can pair up with which nouns), but mathematical terminology is extremely strict about adjectives: an adjective should only be attached to the noun it was first introduced with.

Challenge yourself to pay attention to which noun an adjective will modify. This is an extremely important issue that students often overlook. While this may be a new mental task for you, it isn't *entirely* new either: in calculus, you learned about taking the derivative. But whether you realized it then or not, you took derivatives of *functions*. You did not, for example, take derivatives of geometric objects (even though these are mathematical). Would you be puzzled if somebody asked you, "How do you take the derivative of a tetrahedron?" It really doesn't make sense whenever we write a sentence where we have applied an adjective on an inappropriate noun.

1.5 Fun logic puzzles that don't seem relevant, but they are!

Exercise 37. *After rubbing a magic lantern, a genie appears. The genie will grant you one million dollars if you can identify the fake coin. There are* 9 *coins, all except one are the same weight, the fake one is heavier than the rest. You must determine which is fake using an old fashioned balance. You may use the balance three times. (The scales are of the old balance variety. That is, a small dish hangs from each end of a rod that is balanced in the middle. The device enables you to conclude either that the contents of the dishes weigh the same or that the dish that falls lower has heavier contents than the other.) Explain how this can be done.*

Exercise 38. *The next day, another genie appears. Again, there are* 9 *coins:* 8 *genuine coins and a fake coin which is heavier. The genie will give you one billion dollars if you can identify the fake coin, but you can only use the balance twice.*

Exercise 39. *You walk home from a friend's house and run into another genie. "Oh no, not you again!" you exclaim. The genie lays out* 12 *coins and a balance while saying, "I'll give you one trillion dollars if you can identify which one of these twelve coins is fake, but you can only use the balance three times." You think for a minute, without even using scratch paper to say, "Oh hey, that's easy now! Give me that balance and I'll identify which of these coins is heaviest in no time!" The genie grabs the balance from you and says, "Not so fast! I didn't say whether the fake coin is lighter than the rest or heavier than the rest. Here, have some scratch paper." The genie gives an evil laugh, only to say, "Oh, for the trillion dollars, you have to identify* which *coin is fake, and whether the fake coin is heavier than a genuine coin or lighter than a genuine coin." Explain how using the balance at most three times, you can identify the fake, and whether it is heavier or lighter than the typical coin. Good luck!*

Exercise 40. *Inside of a dark closet are five hats: three maroon and two gray. Knowing this, three logicians go into the closet, and each selects a hat in the dark and places it unseen upon their own head. Once outside the closet, nobody can see their own hat. The first logician looks at the other two, thinks, and says, "I cannot tell what color my hat is." The second logician hears this, looks at the other two, and says, "I cannot tell what color my hat is either." The third logician is blind. The blind logician says, "Well, I know what color my hat is." What color is the third logician's hat?*

Exercise 41. *You are the front desk manager at The Count's Hotel at Transylvania Beach. The hotel has an infinite number of rooms in the following sense: each hotel room has a plaque with a positive integer on it, with no duplication, and for each positive integer, there is a hotel room with that number. (So, there is a room* 1*, and there is a room* 2*, and there is a room* 23487965987*, but there is no room* $\sqrt{7.14}$ *or room* π*. Room* 1 *is the lowest-numbered room. If you think of your social security number, that is one of the rooms of this hotel, as is the square of your SSN.) Using the PA system, you can use the microphone at the front desk to speak to the occupant in each room. Oh! Each room is occupied, so you have no vacancy. A weary traveler (named Jonathan Harker?) shows up to check in to the hotel. Do you turn the traveler away due to no vacancy? Or, can you accommodate Jonathan?*

Exercise 42. *You are the front desk manager at The Count's Hotel at Transylvania Beach. The hotel has an infinite number of rooms in the following sense: each hotel room has a plaque with a positive integer on it, with no duplication, and for each positive integer, there is a hotel room with that number. Using the PA* *system, you can use the microphone at the front desk to speak to the occupant in each room. Oh! Each room is occupied, so you have no vacancy. Ten weary travelers (ten of them!) show up to check in to the hotel, and each want their own hotel room. Do you turn them away? What can you do?*

Exercise 43. *You are the front desk manager at The Count's Hotel at Transylvania Beach. The hotel has an infinite number of rooms in the following sense: each hotel room has a plaque with a positive integer on it, with no duplication, and for each positive integer, there is a hotel room with that number. Using the PA system, you can use the microphone at the front desk to speak to the occupant in each room. Oh! Each room is occupied, so you have no vacancy.*

Suddenly, a bus from Van Helsing's Charter Vans, Inc. with an infinite number of people pulls up. The number of people in the bus is infinite in the following sense: each person on the bus has an index card with a positive integer written on it (with no duplication), and for each positive integer, there is a person who is assigned that number.

How can you accommodate all infinite people already in the hotel and all infinite people on the bus? (Note, you can't just tell all the people in the hotel to move "an infinite number of spots". Your instructions should give the occupant in hotel room 54601 *a specific hotel room to use, and should also give the person number* 608 *on the bus a specific hotel room to use!*

Exercise 44. *You are the front desk manager at The Count's Hotel at Transylvania Beach. The hotel has an infinite number of rooms in the following sense: each hotel room has a plaque with a positive integer on it, with no duplication, and for each positive integer, there is a hotel room with that number. Using the PA system, you can use the microphone at the front desk to speak to the occupant in each room. Oh! Each room is occupied, so you have no vacancy.*

Suddenly, an INFINITE number of buses from Van Helsing's Charter Vans, Inc. (each bus corresponding to a positive integer) pull up, each bus with an infinite number of people. So there is a person number 223897698 *on bus number* 98716234.

How can you accommodate all the people who want hotel rooms? Note, you cannot just try to empty bus 1 *first. Why not? The first person in bus* 2 *would be waiting for eternity!*

Exercise 45. *There is a town consisting of all people who need to shave. The barber in the town is the person who (by definition) only shaves the people who don't shave themselves. Who shaves the barber?*

Exercise 46. *A three-card trick: For this trick you need only three cards: an ace (which we will treat as a "one"), a two, and a three. Line them up in increasing order, left to right, facing up. Turn your back so that you can't see the cards. Ask a friend to choose one of the cards and remember which number it is (If your friend has a faulty memory, you can ask the friend to write the number on a piece of paper and hide the paper from you.). Tell the friend to pick up the chosen card and turn it over (in place). Then tell the friend to switch the places of the other two cards and turn them over, too. Now pick up the cards so that the rightmost card is on top, the middle card is in the middle, and the leftmost card is on the bottom, keeping them face down. Without looking at the cards, move the top card to the bottom of the stack. Repeat, until you have done this exactly ten times. Keep the stack face down and lay the cards out, face up, putting the top card in the middle, the second card on the right, and the bottom card on the left. Magically, exactly one of the following will be true: the three will be on the left, the two will be in the middle, or the one will be on the right. Whichever one it is, that is the card your friend picked!*

Prove that the trick always works. (Hint: Let c be the card chosen by the friend. Then $c = 1$ or $c = 2$ or $c = 3$.) If you need help visualizing this, write the numbers 1, 2, *and* 3 *on some index cards or even little pieces of paper.*

Exercise 47. *You work at the deli counter of a grocery store. Four people (Person 1, Person 2, Person 3, Person 4) line up to get sandwiches. In what order would you serve the people to ensure that everyone gets a sandwich? (That is, who will get the first sandwich? Who will get the third sandwich?)*

Exercise 48. *You work at the deli counter of a grocery store. In one line, there are an infinite number of people (Person 1, Person 2, Person 3, Person 4, etc.) In the VIP line, one person is waiting. In what order would you serve the people to ensure that everyone gets a sandwich? (That is, who will get the first sandwich? Who will get the third sandwich?)*

Exercise 49. *You work at the deli counter of a grocery store. In one line, there are an infinite number of people (A1, A2, A3, A4, etc.) In a second line, there are an infinite number of people (B1, B2, B3, B4, etc.) In what order would you serve the people to ensure that everyone gets a sandwich? (That is, who will get the first sandwich? Who will get the third sandwich?)*

Exercise 50. *You work at the deli counter of a grocery store. In one line, there are an infinite number of people (A1, A2, A3, A4, etc.) In a second line, there are an infinite number of people (B1, B2, B3, B4, etc.) In a third line, there are an infinite number of people (C1, C2, C3, C4, etc.) In what order would you serve the people to ensure that everyone gets a sandwich? (That is, who will get the first sandwich? Who will get the third sandwich?)*

Exercise 51. *For the table below, construct a formula for y that, given a value of x, gives the value of y. The formula should express y in terms of x.*

x	1	2	3	4	5	6
y	1	4	6	8	10	12

Exercise 52. *For the table below, construct a formula for y that, given a value of x, gives the value of y. The formula should express y in terms of x.*

x	2	4	6	8	10	12
y	1	2	3	4	5	6

Exercise 53. *For the table below, construct a formula for y that, given a value of x, gives the value of y. The formula should express y in terms of x.*

x	3	6	9	12	15	18
y	1	2	3	4	5	6

Exercise 54. *For the table below, construct a formula for y that, given a value of x, gives the value of y. The formula should express y in terms of x.*

x	4	7	10	13	16	19
y	1	2	3	4	5	6

Exercise 55. *For the table below, construct a formula for y that, given a value of x, gives the value of y. The formula should express y in terms of x.*

x	2	5	8	11	14	17
y	1	2	3	4	5	6

Exercise 56. *For the table below, construct a formula for y that, given a value of x, gives the value of y. The formula should express y in terms of x. The forumla will need to be piece-wise and non-recursive (that is, non self-referential)*

x	1	2	3	4	5	6	...
y	1	1	2	2	3	3	...

Chapter 2

Propositions

This chapter is about understanding what propositions mean, and especially understanding the meaning of new propositions that are built out of old propositions. In the next chapter, we learn how to prove each kind of proposition we introduce here. Prior to proving propositions, we must understand the types of propositions which we will need to prove by clarifying the typical grammar used in propositions, and how the propositions should be read. The process described in this chapter is very formulaic: you are used to applying formulas where the variables are numbers (or functions). Just as algebra involves a precise manipulation of numbers (and variables which can have numeric values), logic is a precise manipulation of propositions (and variables which can have truth values). Throughout the first several sections of this chapter, the "variables" which you will substitute are propositions.

You have seen examples of these prpositions before. For example, from calculus consider:

Theorem 57 (Mean Value Theorem)**.** *If f is a function that is continuous on $[a, b]$ and f is differentiable on (a, b), then there exists $c \in (a, b)$ such that*

$$f'(c) = \frac{f(b) - f(a)}{b - a}.$$

An example you saw even earlier than this (though not typically stated this formally) is from algebra:

Theorem 58 (Distributive Law)**.** *For all $a, b, c \in \mathbb{R}$, the equation $a(b + c) = ab + ac$ holds.*

You have likely seen the Distributive Law without all of the extra "for all" language, yet this is the type of text we will focus on in this chapter.

Before moving forward, we pause to discuss the difference between a definition and a theorem. In mathematics, a definition introduces a reader to a new mathematical concept, while a theorem is a factual statement which can be proved. Compare the definition we gave of a proposition with the words of the Mean Value Theorem. The definition of a proposition introduces a new idea (what a proposition is), while the text of the Mean Value Theorem states a guarantee about the slope of a tangent line at $x = c$.

2.1 Logical operations

Just like addition is an operation which allows us to take two integers such as 2 and 5 and make meaning of $2 + 5$, which is defined to be another integer, we have logical operations which allow us to take propositions and combine them in some meaningful way to obtain a new proposition. Recall from Definition 6 that a proposition is a declarative sentence that is true or false, but cannot be both. Therefore, if p is a proposition and p is not true, then p must be false.

2.1.1 Negation

Definition 59: Negation

Let p be a proposition. The **negation** of p, denoted $\neg p$, is the proposition that is false when p is true, and is true when p is false. In other words, $\neg p$ has the opposite truth value of p. The negation of p is read aloud "not p."

Notice the word proposition appears twice in the definition of negation. The first use of proposition is important because the second sentence says that we take the negation of p, where p is a proposition. Recall from Warning 17 that sets are not propositions, so it makes no sense to talk about the negation of a set. The second use of the word proposition is useful in handling Habit 2 which asks is the negation a noun, an adjective, or verb? Because the negation of a proposition is a proposition, the negation is a noun. When you are asked to recite the definition of negation, you do not need to use the exact words above in the order given, but you will be expected to use the word proposition twice.

While a good definition is the collection of complete sentences above, the information about the negation is nicely summarized in a truth table:

p	$\neg p$
T	F
F	T

Example 60. *Let us define a to be the proposition "February has 31 days," define b to be "Canada is in North America," define c to be "Canada is in Asia," and let d be "Tasmania is an island." Since a is false, $\neg a$ is true. Since b is true, $\neg b$ is false. We will use the propositions a, b, c, and d as examples to illustrate examples throughout this section.*

2.1.2 Conjunction

Definition 61: Conjunction

Let p and q both be propositions. The **conjunction** of p and q, denoted $p \wedge q$, is the proposition that is true when both p and q are true, and is false otherwise. The conjunction of p and q is read aloud "p and q."

Habit 2 urges us to ask: is the conjunction a noun, an adjective, or verb? Due to the second use of "proposition" in the definition, conjunction is a proposition, which is a noun. To make a conjunction, we need two propositions.

There are two uses of the word $\boxed{\text{and}}$ here. The first sentence "Let p and q both be propositions" could have been written as two separate sentences: "Let p be a proposition. Let q also be a proposition." The phrase "p and q" at the end of the definition uses the word $\boxed{\text{and}}$ in its *mathematical* sense, while the first use of the word $\boxed{\text{and}}$ in the sentence "Let p and q both be propositions" uses $\boxed{\text{and}}$ in the colloquial sense. When $\boxed{\text{and}}$ is placed between two propositions, it will generally be the mathematical conjunction. Context will give away whether "p and q" is to be thought of as "one new thing," the conjunction of two propositions, or whether "p and q" is an attempt to refer to two separate things.

The conjunction is summarized in a truth table. We start with p and with q being the basic propositions (to the left of the double vertical bar) and there are now four possible situations, shown in four rows:

p	q	$p \wedge q$
T	T	T
T	F	F
F	T	F
F	F	F

Example 62. *With a, b, c, and d as defined in Example 60, $b \wedge d$ is true while $b \wedge a$ is false. The proposition $a \wedge d$ is false. Finally, $c \wedge a$ is false as well. (The four examples here illustrated all four situations in the truth table.)*

Example 63. *Consider the proposition "Pizza is food," which is true. Consider a second proposition "Cats are animals," which is true. The conjunction of these two propositions is the proposition "Pizza is food and cats are animals," which is true.*

Example 64. *Consider the true proposition "Pizza is food" and the false proposition "Cats are plants." The conjunction of these two propositions is the proposition "Pizza is food and cats are plants," which is false.*

Example 65. *The conjunction of the false proposition "Pizza is a vehicle" and the true proposition "Cats are animals" is the proposition "Pizza is a vehicle and cats are animals," which is false.*

Example 66. *The conjunction of the proposition "Pizza is a vehicle" and the proposition "Cats are plants" is the proposition "Pizza is a vehicle and cats are plants," which is false.*

We have several other logical operations to introduce, but before doing so, the definition of conjunction that we introduced may be one of the first kinds of propositions that pose a unique kind of challenge when you are asked to recite a definition on a quiz. So, let us give several nearly correct definitions of conjunction, and then comment on how to fix each almost-correct definition:

- Given propositions p and q, we define $p \wedge q$ to be the proposition that is true when p and q are both true, and false otherwise. This is said "p and q."
  - **Comment:** In this text, the new terminology (conjunction) is not introduced. The new word should be inserted at minimum one time somewhere amongst the sentences. One option would be to change the last sentence to say: This is called the conjunction, and is said "p and q."
- Let p and q both be propositions. The conjunction is written $p \wedge q$ and is true when p and q are both true, and false otherwise. This is read "p and q."
  - **Comment:** This text never mentions that a conjunction is a proposition. The second sentence could be modified by writing: The conjunction is written $p \wedge q$ and is the proposition that is true when p and q are both true, and false otherwise.
- The conjunction of p and q is the proposition that is true when p and q are both true, and false in all remaining situations, and is said "p and q."
  - **Comment:** This text does not mention that p is a proposition and that q is a proposition. One way to fix this would be by keeping the senteces that we already have, but inserting one of the following texts in front of all the sentences already written:
    * Let p be a proposition. Let q be a proposition.
    * Let p and q both be propositions.
    * Suppose we have a proposition p, and we have a proposition q.

2.1.3 Disjunction (and symmetric difference)

Definition 67: Disjunction

Let p and q both be propositions. The **disjunction** of p and q, denoted $p \vee q$, is the proposition that is false when both p and q are false, and is true otherwise. The disjunction of p and q is read "p or q."

Following Habit 2, the disjunction is a noun (due to the second use of the word proposition). The word $\boxed{\text{and}}$ in this definition is in the non-technical everyday English sense of the word "and." The word $\boxed{\text{or}}$ illustrates the precise mathematical use of the word. We summarize the disjunction in a truth table:

p	q	$p \vee q$
T	T	T
T	F	T
F	T	T
F	F	F

Example 68. *With a, b, c, and d as defined in Example 60, $b \vee c$ is true, while $a \vee c$ is false.*

For $p \vee q$ to be true, we only need a minimum of one of p or q to be true: if both p is true and q is true, then $p \vee q$ is true. Sometimes, the word $\boxed{\text{or}}$ is used in English in a more exclusive way. For instance, when at a car dearleship, a customer that says "I'll take the red car or the yellow car" is probably not intending to buy both. This other kind of $\boxed{\text{or}}$ is the symmetric difference:

Definition 69: Symmetric difference

Let p and q both be propositions. The **symmetric difference**, denoted $p \oplus q$, is the proposition that is true when exactly one of p and q is true, and is false otherwise.

Following Habit 2, the symmetric difference is a noun, summarized in the following truth table:

p	q	$p \oplus q$
T	T	F
T	F	T
F	T	T
F	F	F

The symmetric difference is not extremely common: we mention it only to contrast this with the disjunction.

Example 70. *With a, b, c, and d as defined in Example 60, $b \vee d$ is true while $b \oplus d$ is false.*

Example 71. *Below is a complete truth table for $p \oplus (q \vee r)$.*

p	q	r	$q \vee r$	$p \oplus (q \vee r)$
T	T	T	T	F
T	T	F	T	F
T	F	T	T	F
T	F	F	F	T
F	T	T	T	T
F	T	F	T	T
F	F	T	T	T
F	F	F	F	F

We have presented the basic propositions p and q and r before the double vertical line. Since there are 3 basic propositions, there are 2^3 situations/rows in the truth table. If you examine the columns for q and r only, you'll see two copies (the first four rows versus the last four rows) of a truth table for two basic propositions.

It is helpful to be systematic in this way. Notice that we have the four situations for q and r when p is true first, followed by the four possible combinations for q and r when p is false next. An alternate way to think about this is that the r column alternates every one row, the q column alternates in truth value every two rows, while the p column alternates every four rows.

How many situations would there be with four basic propositions? What would be an order which ensures that the rows are presented systematically?

2.1.4 Implication

Definition 72: Implication

Let p and q both be propositions. The **implication** from p to q, denoted $p \to q$, is the proposition that is false when p is true and q is false, and is true otherwise. The implication from p to q is read "if p, then q."

The proposition p which appears in the implication is called the **hypothesis** or the **premise**, while the proposition q is called the **conclusion**.

Since an implication is a proposition, an implication is a noun. This is our first definition where accompanying words are defined as well, so here is a new habit to follow:

Habit 73

Every time you encounter a new definition, notice if additional words/phrases accompany the main definition, and identify their part of speech as well.

Besides the main word **implication**, the hypothesis/premise of an implication is a noun (since it is a proposition), and the conclusion of an implication is also a noun. We summarize the implication in a truth table:

p	q	$p \to q$
T	T	T
T	F	F
F	T	T
F	F	T

Remark 74. *The implication $p \to q$ should be read "if p then q" but is sometimes informally read "p* **implies** *q." On rare occasion, you might see $p \to q$ read "p is* **sufficient** *for q" or read "q is* **necessary** *for p."*

Example 75. *The implication "If Canada is in Asia, then February has 31 days" is true (based on the last row of the truth table), while "If Canada is in North America, then February has 31 days" is false.*

Example 76. *Let's examine the implication "If cats are plants, then dogs fly." This is an implication: it is text in the form "if p then q". Note that the p text is "cats are plants" which is false, and that the q text is "dogs fly" which is also false. So what about the truth value of the proposition "if p then q" which is denoted $p \to q$? We go to the truth table that was just presented, and look for the row in which there is an F in the p column and an F in the q column. The row that has this happens to the the final row. In that final row, we look up the truth value of $p \to q$ in the final column, and see a T there in the final row. Thus, by examining that T, the proposition "If cats are plants, then dogs fly" is true.*

This example tends to really bother people. I suspect this is because people want to think "How can two falses make a true?" But in the end, we're just mechanically following what the truth table tells us about $p \to q$, and the truth table is meant to summarize the text of the definition, so ultimately, we're basing this off of the definition of implication.

Again, it may seem bothersome, but we have to stick with the definition of implication as it is universally accepted. If you want to change where the Ts and Fs are located for the final column, you can do this, but you'd have to name your logical operation something *other* than implication and speak it using something *other* than "if ... then ..." (To try to convince you further that what we have is okay, consider an operation from arithmetic and algebra: a negative times a negative is a positive. That's just how the multiplication operation works.)

Example 77. *Let's determine the truth value of "If* 125 *people were on the Eiffel Tower at midnight on New Years Day in 1950, then kanagroos are marsupials." The p text is:* 125 *people were on the Eiffel Tower at midnight on we Years Day in 1950. It is not likely we can know for certain whether this was true or false.*

However, look at the q text: kangaroos are marsupials. We know that q is true. Now, going back to the truth table, in both rows where q is true, we see that $p \to q$ is true. Therefore, we have successfully determined that "If 125 people were on the Eiffel Tower at midnight on New Years Day in 1950, then kanagroos are marsupials" is true, in spite of the fact that we don't know the truth value of "125 people were on the Eiffel Tower at midnight on New Years Day in 1950."

If we swap the hypothesis and conclusion of the previous implication, we get the implication "If February has 31 days, then Canada is in North America," which is true. This illustrates the implication is "one-sided," also evidenced by the column for $p \to q$ in the truth table having three Ts and only one F. So $p \to q$ and $q \to p$ are different propositions. Starting from the implication $p \to q$, the new implication $q \to p$ obtained by swapping the premise and the conclusion has a name:

Definition 78: Converse

The **converse** of the implication $p \to q$ is the implication $q \to p$.

We can only talk about the converse of an implication. It makes no sense to talk about the converse of $p \wedge q$.

Example 79. *The converse of $a \to b$ is $b \to a$, and the converse of $b \to a$ is $a \to b$.*

Let us put the truth values for $p \to q$ in one column, and the truth values for $q \to p$ in another column:

p	q	$p \to q$	$q \to p$
T	T	T	T
T	F	F	T
F	T	T	F
F	F	T	T

The last two columns of the truth table are not identical so:

Warning 80: General warning about converses

The converse of an implication is not the same as the original implication.

Throughout mathematics courses, after theorems which are implications are presented, a warning that the converse is not necessarily true is often presented. (The reason that these types of warnings are valid to bring up in the first place is due to Warning 80.)

Example 81. *From calculus, if f is differentiable at c, then f is continuous at c. However, the converse is not necessarily true: a function f can be continuous at c without being differentiable at c.*

The fact that the implication "if f is differentiable, then f is continous" is true is presented in calculus. What it means is that if we are in the situation that f is differentiable, then we know for sure that f is continuous. Here's another example of this, based on everyday life:

Example 82. *The implication "If it is snowing, then it is cold" is a true implication. That is to say, when it's snowing, it has to be cold. We just can't have snow in a non-cold situation.*

However, the converse is the implication "If it is cold, then it is snowing." That implication is not true. It can be cold, but that doesn't force *snow.*

This example is really convincing that an implication $p \to q$ and its converse $q \to p$ are generally different: in the example we gave, one implication was true, while the converse was false.

Starting with the implication $p \to q$, there is another implication which needs to be defined:

Definition 83: Contrapositive

The **contrapositive** of the implication $p \to q$ is the implication $\neg q \to \neg p$.

The contrapositive of an implication is an implication. As with the converse, it makes no sense to talk about the contrapositive of $r \wedge s$, since $r \wedge s$ is not an implication. In a combined truth table for $p \to q$ and its contrapositive $\neg q \to \neg p$,

p	q	$p \to q$	$\neg q \to \neg p$
T	T	T	T
T	F	F	F
F	T	T	T
F	F	T	T

the last two columns are identical, so $p \to q$ is saying the same thing as $\neg q \to \neg p$, though the words used may be different. Just as $x(y+z)$ always has the same numerical value as $xy+xz$ in algebra, the truth values of $p \to q$ and of $\neg q \to \neg p$ will always be the same, an idea we will explore more thoroughly in Section 2.3.

2.1.5 Biconditional

There is one final logical operation we discuss:

Definition 84: Biconditional

Given proposition p and proposition q, the **biconditional** of p and q is the proposition that is true when p and q have the same truth value, and is false otherwise. The biconditional is denoted $p \leftrightarrow q$ and is spoken "p if and only if q."

The biconditional is a noun. If formality is not required, the phrase "if and only if" can be abbreviated by writing "iff" instead. (You may see $p \leftrightarrow q$ read aloud as "q is a **characterization** of p" or more rarely, "p is **necessary and sufficient** for q") A truth table for the biconditional summarizes the definition:

p	q	$p \leftrightarrow q$
T	T	T
T	F	F
F	T	F
F	F	T

To connect ideas, note that the negation of $p \leftrightarrow q$ would have the same truth value as $p \oplus q$ in all situations. The idea that the proposition $\neg(p \leftrightarrow q)$ and the proposition $p \oplus q$ always have the same truth value is another example of propositional equivalence, which previews Section 2.3.

2.1.6 Additional exercises

Exercise 85. *Construct a complete truth table for* $(p \wedge q) \vee r$ *[key]*

Exercise 86. *Construct a complete truth table for* $a \oplus (b \to c)$ *[key]*

Exercise 87. *Construct a complete truth table for* $(p \to \neg q) \to r$ *[key]*

Exercise 88. *Construct a complete truth table for* $(r \leftrightarrow s) \to (t \vee u)$ *[key]*

Exercise 89. *Build a complete truth table for* $(p \vee \neg q) \wedge (r \vee s)$

Exercise 90. *Build a complete truth table for* $(p \vee q) \wedge (q \oplus s)$

Exercise 91. *If p is the proposition "JJ ate peas", q is the proposition "JJ ate pasta", and r is the proposition "JJ may have dessert", express each proposition below as an English sentence:*

- $\neg p$
- $(p \vee q) \to r$
- $\neg r \to \neg p$

- $\neg p \wedge (q \vee r)$
- $p \leftrightarrow q$

Exercise 92. *Let p be the proposition "It is below freezing". Let q be the proposition "It is snowing". Write the propositions below symbolically using p and q and logical operations.*

- *It is below freezing and snowing.*
- *It is below freezing but not snowing.*
- *It is not below freezing and it is not snowing.*
- *It is either snowing or below freezing (or both).*
- *If it is below freezing, it is also snowing.*
- *It is either below freezing or it is snowing, but it is not snowing if it is below freezing.*

Exercise 93. *Determine whether the propositions below are true or false. Briefly explain why.*

- *If* $2 + 2 = 4$*, then* $1 + 1 = 2$.
- *If* $1 + 1 = 2$*, then* $2 + 3 = 4$.
- *If it is winter, then it is not spring, summer, or fall.*
- *If* $1 + 1 = 3$*, then pigs can fly.*
- *If* $0 > 1$*, then* $2 > 1$.

Exercise 94. *Let p be the proposition "If George Washington owned exactly 145 books, then rabbits are animals." Answer each question below:*

- *Is p true or false?*
- *Identify two propositions q and r in words such that p is $q \rightarrow r$.*
- *In symbols, what is the converse of p? In English, what is the converse of p?*
- *In symbols, what is the contrapositive of p? In English, what is the contrapositive of p?*

Exercise 95. *Determine whether these biconditionals are true or false. Briefly explain why.*

- $2 + 2 = 4$ *if and only if* $1 + 1 = 2$.
- $1 + 1 = 2$ *if and only if* $2 + 3 = 4$.
- *It is winter if and only if it is not spring, summer, or fall.*
- $1 + 1 = 3$ *if and only if pigs can fly.*
- $0 > 1$ *if and only if* $2 > 1$.

2.2 Mathematical language: definitions

Because we now have enough topics to discuss, this is a good time to pause to discuss mathematical language before moving on. Recall our definition of conjunction from Definition 61. Instead of trying to memorize the definition of conjunction word-for-word, ensure that a definition that you write conveys all the required meaning. For examples, we show other perfectly good ways to write the definition of conjunction, with the first sample being the original definition we gave:

1. Let p and q both be propositions. The **conjunction** of p and q, denoted $p \wedge q$, is the proposition that is true when both p and q are true, and is false otherwise. The conjunction of p and q is read aloud "p and q."

2. Let p be a proposition. Let q be a proposition. The **conjunction** of p and q, denoted $p \wedge q$, is the proposition that is true when both p and q are true, and is false otherwise. The conjunction of p and q is read aloud "p and q."

3. The **conjunction** of the proposition p and the proposition q, denoted $p \wedge q$, is the proposition that is true when both p and q are true, and is false otherwise. The conjunction of p and q is read "p and q."

4. The **conjunction** $p \wedge q$ of the proposition p and the proposition q is the proposition that is true when both p and q are true, and is false otherwise. The conjunction of p and q is read aloud "p and q."

5. Given a proposition p and a proposition q, the **conjunction** $p \wedge q$ of p and q, read "p and q" is the proposition that is true when both p and q are true, and is false otherwise.

6. Given propositions p and q, the **conjunction** of p and q, read "p and q" is the proposition that is true when both p and q are true, and is false otherwise, and is denoted $p \wedge q$.

These are six good examples of how to write the definition of conjunction.

Warning 96

Do not memorize the text of a definition word-for-word. (Why does this matter? In practice, some students who have focused on memorizing each definition word-for-word and symbol-for-symbol down to the exactness of which letter(s) appeared when a definition was first stated, tended to de-emphasize focus on the meaning that's behind the words, and that led to more struggle in the end)

Instead of memorizing, understand what key things you are supposed to learn from the definition (and thus, what key things are expected every time you write a definition). We'll use the conjunction as an example. All six samples above make clear that the conjunction requires two "things" – one called p, and one called q. However, p is not just any thing. All samples make clear that p is a proposition. Similarly, all samples make clear that q is a proposition.

All six samples make clear that the conjunction is a proposition. An incomplete definition of conjunction would leave this out and make the reader "guess" that the conjunction is a proposition based on the discussion of "true" and "false" which appears in the text.

Of course, since the thing being defined is a proposition, the text in each of the six samples must make clear when $p \wedge q$ is true and when $p \wedge q$ is false.

Finally, the full phrase "conjunction of p and q" appears, whether it is exactly this text unbroken or that word order is part of some nonconsecutive text. What of this broken text? Sample 1 has a "set up sentence" while sample 2 has two set up sentences. Sample 3 does not start with set up, so that is included "along the way" by saying "of the proposition p and the proposition q". It is a matter of personal taste, but one way or another, before p really gets used, it should be explained that p is a proposition.

Similar discussion can be held for all definitions introduced so far. You should try this yourself.

With every new definition always figure out the part of speech (see Habit 2) and for adjectives determine what type of noun is modified (Warning 3). In addition, for a noun, determine what kind of object is being defined. The vast majority of nouns defined so far have been propositions. This advice may not seem important now, but later, your nouns might be things that are not propositions (some examples of this we'll see later on include sets, relations, functions, and ordered pairs). There will be so many moving parts later, it is important to practice these habits now.

Appealing to previous mathematical experience, it is vital to mathematical proof in the next chapter that you pay close attention to grammar. For instance, a function can be continuous, but an equation cannot. So, we can't say "The equation $3x^2 + 5 = 8$ is continuous." You have equations with variables (such as $3x^2 + 5 = 8$) and equations without variables (such as $3 + 6 = 9$). An equation with variables might have solutions, but we wouldn't speak of an equation *without* variables as having a solution: pay attention, even to the language surrounding this algebra. We would say, "$x = -1$ is a solution to the equation $3x^2 + 5 = 8$" because substituting -1 for x would make $3x^2 + 5$ have the same value as 8.

Samples 4 and 5 of the definitions for conjunction from earlier in this section place the notation $p \wedge q$ immediately after the word conjunction, following Language Discussion 5.

Instead of running away from the grammar and language, run towards it! By the end of the next chapter, you will see that a large component of the thought process for proofs stimulates the same part of the brain you used when you evaluated integrals or took a derivative. However, the new component is language: it is *impossible* to be successful with proofs if you ignore the grammar and language.

One lesson to take away from this is that there is more than one correct way to write a definition for **conjunction**. (Likewise, the same principles apply for every definition.) It is tempting to simply memorize a definition word-for-word, but this is dangerous.

After the six examples of definitions for conjunction, we discussed the important features that a good definition of conjunction would have. While you can view this section as a sort of "checklist" of things to cover in your own definition of conjunction, that view is too limiting. Instead, look at the discussion of this section as an example of the type of discussion you should create in your *own* head for all the *other* definitions you have learned and will learn in this handbook.

In other words, prior to the discussion in this section, it would have been very natural for you to not know exactly what to look for when reading a definition of conjunction. The discussion helps with hints on what kinds of things mathematicians are looking for when reading definitions. Try to do this with every definition. There are at least three benefits to doing this: first, you'll have a more thorough understanding of each definition you're presented; second, you'll likely have a more complete definition when called upon to recite a definition; and third, a well-constructed checklist like this is probably very close to your instructor's rubric for grading the recitation of a definition. In other words, these are probably the things that your instructor is looking for when grading your definition work. So, knowing what to look for as you read each definition should translate into more points!

Exercise 97. *What are the key things you must address when writing out a definition of disjunction?*

2.3 Propositional equivalence

In Section 2.1, logical operations such as $\wedge$ were used to take **simple propositions** such as p and such as q to make **compound propositions** such as $p \wedge q$. (By *compound* proposition, we mean that instead of a proposition being represented as a single letter such as p, we have a larger proposition constructed from prior propositions and logical operations.) Some propositions are always true:

Definition 98: Tautology

A **tautology** is a compound proposition that is always true, regardless of the truth values of the simple propositions which appear.

Definition 99: Contradiction

A **contradiction** is a compound proposition that is always false, regardless of the truth values of the simple propositions which appear.

Example 100. *The truth table below shows that the proposition $[a \to b] \leftrightarrow [(\neg b) \leftrightarrow (\neg a)]$ is a tautology.*

a	b	$a \to b$	$\neg b$	$\neg a$	$\neg b \to \neg a$	$[a \to b] \leftrightarrow [(\neg b) \leftrightarrow (\neg a)]$
T	T	T	F	F	T	T
T	F	F	T	F	F	T
F	T	T	F	T	T	T
F	F	T	T	T	T	T

Example 101. *By negating the previous example, $\neg[[a \to b] \leftrightarrow [(\neg b) \leftrightarrow (\neg a)]]$ is a contradiction.*

Definition 102: Logical equivalence

The propositions p and q are **logically equivalent** if the proposition $p \leftrightarrow q$ is a tautology. We write $p \equiv q$ to denote that p and q are logically equivalent.

In algebra, the fact that the expression $x(y+z)$ always has the same numerical value as $xy+xz$ no matter what the numerical values of the simple expressions x, y, and z are is denoted $x(y+z) = xy+xz$ using an equal sign, and we say that the expressions $x(y+z)$ and $xy+xz$ are **equal**. In complete analogy to this:

Example 103: Every implication is logically equivalent to its contrapositive

The proposition $a \to b$ always has the same truth value as $(\neg b) \to (\neg a)$ no matter what the truth values of the simple propositions a and b are. Since the truth table below shows that $[a \to b] \leftrightarrow [(\neg b) \to (\neg a)]$ is a tautology, we say that the proposition $a \to b$ is logically equivalent to the proposition $(\neg b) \to (\neg a)$, which we denote by $[a \to b] \equiv [(\neg b) \leftrightarrow (\neg a)]$.

a	b	$a \to b$	$\neg b$	$\neg a$	$\neg b \to \neg a$	$[a \to b] \leftrightarrow [(\neg b) \leftrightarrow (\neg a)]$
T	T	T	F	F	T	T
T	F	F	T	F	F	T
F	T	T	F	T	T	T
F	F	T	T	T	T	T

Warning 104: Be careful where $\equiv$ appears

In the truth table above, we have a column heading of $[a \to b] \leftrightarrow [(\neg b) \leftrightarrow (\neg a)]$. This is fine, as the symbols used are the logical operators introduced in the previous section. It is a bit of abuse of notation if we had a column heading which said $[a \to b] \equiv [(\neg b) \leftrightarrow (\neg a)]$. Instead, we write $[a \to b] \equiv [(\neg b) \leftrightarrow (\neg a)]$ to say that $[a \to b]$ is logically equivalent to $[(\neg b) \leftrightarrow (\neg a)]$, which, if we peel apart the definition of logical equivalence, is that $[a \to b] \equiv [(\neg b) \leftrightarrow (\neg a)]$ is a tautology, which, if we peel apart the definition of tautology says that $[a \to b] \equiv [(\neg b) \leftrightarrow (\neg a)]$ is always true, which is evidenced by having the column for $[a \to b] \equiv [(\neg b) \leftrightarrow (\neg a)]$ have only "T"s in it.
Having a $\equiv$ symbol as part of a column heading is missing the point of what $\equiv$ is for. Let's write $\equiv$ between one propsition and another to indicate that two propositions are equal, which we can verify using a truth table, but in truth table columns, let's only stick to writing compound propositions using the logical operations such as $\neg, \wedge, \vee, \to$, and $\leftrightarrow$.

Practice showing logical equivalences using the examples below as exercises by writing out appropriate truth tables, as shown in Example 103.

Exercise 105. *Show the* **identity laws** *by writing out appropriate truth tables:*

- $p \wedge T \equiv p$
- $p \vee F \equiv p$

Exercise 106. *Show the* **domination laws** *by writing out appropriate truth tables:*

- $p \vee T \equiv T$
- $p \wedge F \equiv F$

Exercise 107. *Show the* **repetition removal laws** *by writing out appropriate truth tables:*

- $p \vee p \equiv p$
- $p \wedge p \equiv p$

Exercise 108. *Show the* **double negation law** $\neg(\neg p) \equiv p$ *by writing out an appropriate truth table.*

Exercise 109. *Show the* **commutative laws** *by writing out appropriate truth tables:*

- $p \vee q \equiv q \vee p$
- $p \wedge q \equiv q \wedge p$

Exercise 110. *Show the* **associative laws** *by writing out appropriate truth tables:*

- $(p \vee q) \vee r \equiv p \vee (q \vee r)$
- $(p \wedge q) \wedge r \equiv p \wedge (q \wedge r)$

Exercise 111. *Show the* **distributive laws** *by writing out appropriate truth tables:*

- $p \vee (q \wedge r) \equiv (p \vee q) \wedge (p \vee r)$
- $p \wedge (q \vee r) \equiv (p \wedge q) \vee (p \wedge r)$

Exercise 112. *Show* **De Morgan's laws** *by writing out appropriate truth tables:*

- $\neg(p \wedge q) \equiv (\neg p) \vee (\neg q)$
- $\neg(p \vee q) \equiv (\neg p) \wedge (\neg q)$

Exercise 113. *Show* **constant laws** *by writing out appropriate truth tables:*

- $p \vee \neg p \equiv T$
- $p \wedge \neg p \equiv F$

Exercise 114. *Show the* **implication conversion law** $p \to q \equiv \neg p \vee q$ *by writing out an appropriate truth table.*

Method 115: Two ways to show logical equivalence

There are two ways to show that one proposition is equivalent to another. Suppose you are tasked with showing that the proposition $p \to (q \to r)$ is logically equivalent to the proposition $(p \wedge q) \to r$. You can:

1. Write a truth table, showing that the column for $[p \to (q \to r)] \leftrightarrow [(p \wedge q) \to r]$ only has Ts, which means that $[p \to (q \to r)] \leftrightarrow [(p \wedge q) \to r]$ is a tautology and, therefore, $[p \to (q \to r)] \equiv [(p \wedge q) \to r]$. This way was done in Example 103 to show that two propositions are logically equivalent.

2. Start with the proposition $p \to (q \to r)$ and use previous logical equivalences (in a manner of substitution) to obtain the proposition $(p \wedge q) \to r$ in the same format as trigonometric identities. So $[p \to (q \to r)] \equiv [(p \wedge q) \to r]$ is shown by using previous logical equivalences such as $\neg(p \wedge q) \equiv (\neg p) \vee (\neg q)$ in the same way that a trig identity like $\tan\theta + \cot\theta = \sec\theta \csc\theta$ is verified using standard (previous) trig identities such as $\sin^2\theta + \cos^2\theta = 1$.

As an example of what me mean by the format of verifying a trigonometric identity, let's examine the following example where we verify $\sec\theta - \cos\theta = \sin\theta\tan\theta$.

$$\begin{aligned}
\sec\theta - \cos\theta &= \frac{1}{\cos\theta} - \cos\theta \\
&= \frac{1}{\cos\theta} - \frac{\cos\theta}{1} \\
&= \frac{1}{\cos\theta} - \frac{\cos^2\theta}{\cos\theta} \\
&= \frac{1-\cos^2\theta}{\cos\theta} \\
&= \frac{\sin^2\theta}{\cos\theta} \\
&= \sin\theta\frac{\sin\theta}{\cos\theta} \\
&= \sin\theta\tan\theta
\end{aligned}$$

Our process took the left side of $\sec\theta - \cos\theta$ and used the standard list of trigonometric identities to arrive at the right side. Our first step took $\sec\theta-\cos\theta$ and rewrote it as $\frac{1}{\cos\theta}-\cos\theta$ using the the standard identity $\sec\theta = \frac{1}{\cos\theta}$. The second and third steps were algebra. The next step took $1-\cos^2\theta$ and rewrote this as $\sin^2\theta$ using the standard trig identity $\sin^2\theta+\cos^2\theta = 1$. The last step applied the standard trig identity $\tan\theta = \frac{\sin\theta}{\cos\theta}$. In summary, by applying standard trig identities (and some algebra identities) we verified the new identity $\sec\theta - \cos\theta = \sin\theta\tan\theta$, though this identity is not "standard" in the sense that it's not on everyone's list to memorize, but it is nonetheless true. We apply that same kind of formatting and thought process (use what we know is true to get something new to be true) for logical equivalences.

Example 116. *Here, we show how that second method works. Using the logical equivalences introduced earlier, let's verify that $p \to (q \to r) \equiv (p \wedge q) \to r$.*

$$\begin{aligned}
p \to (q \to r) &\equiv p \to ((\neg q) \vee r) \\
&\equiv \neg p \vee ((\neg q) \vee r) \\
&\equiv ((\neg p) \vee (\neg q)) \vee r \\
&\equiv (\neg(p \wedge q)) \vee r \\
&\equiv (p \wedge q) \to r.
\end{aligned}$$

One of the main purposes of the idea of logical equivalences is to know when two propositions always have the same truth value, as one proposition could then justifiably be replaced with the other. This is the same as saying the algebra student should know that $5a + 15b$ can be replaced with $5(a + 3b)$.

A common mistake of algebra students is to replace $(a+b)^2$ with a^2+b^2. We know $(a+b)^2 = a^2+b^2$ is not true in general. That is, the expressions $(a+b)^2$ and a^2+b^2 are not numerically equivalent. The algebra student should know that these two expressions are different. In the same way, the logic student should know of logical equivalence to know of the idea that two propositions may *not* be logically equivalent:

Warning 117: Conjunction is not the same as implication

Let p be a proposition and let q be a proposition. The conjunction $p \wedge q$ and the implication $p \to q$ are not logically equivalent, as seen in their truth tables. Therefore, we shouldn't mentally equate $p \to q$ with $p \wedge q$, since they are different propositions.

p	q	$p \to q$	$p \wedge q$
T	T	T	T
T	F	F	F
F	T	T	F
F	F	T	F

Exercise 118. *Show that $\neg(p \vee (\neg p \wedge q))$ and $\neg p \wedge \neg q$ are logically equivalent without using a truth table.* [key]

Exercise 119. *Show that $(p \wedge q) \rightarrow (p \vee q)$ is a tautology without using a truth table.* [key]

Exercise 120. *Show that $\neg(p \rightarrow (p \vee q))$ is a contradiction without using a truth table.* [key]

Exercise 121. *Show that the implication $[(p \rightarrow q) \wedge (q \rightarrow r)] \rightarrow (p \rightarrow r)$ is a tautology using truth tables.*

Exercise 122. *Show that the proposition $\neg((p \wedge q) \rightarrow q)$ is a contradiction WITHOUT using truth tables.*

Exercise 123. *Show that the proposition $\neg(p \leftrightarrow q)$ and the proposition $p \leftrightarrow \neg q$ are logically equivalent.*

2.4 Predicates and quantifiers

2.4.1 Predicates

Definition 124: Predicate

A **predicate** (with variable x) is a declarative sentence $P(x)$, which satisfies the property that upon substituting x with an appropriate entity, the sentence is a proposition.

In other words, $P(x)$ is a predicate if, after substituting an appropriate value of x, the resulting text is a declarative sentence that is either true or false, but not both. (The result of true or false may depend on the choice of value to substitute for x.)

Definition 125: Universe of discourse

A **universe of discourse** for a predicate $P(x)$ is a set U such that all elements of U are appropriate entities to substitute for x in $P(x)$.

Example 126. *Let $P(x)$ be $\boxed{x \text{ was born in the year 1918}}$. Then $P(x)$ is a predicate. If m denotes Nelson Mandela, then $P(m)$ is true. If n denotes Chuck Norris, then $P(n)$ is false.*

Recall that we defined $\mathbb{H}$ to be the set of all humans in Section 1.3.2. In fact, $\mathbb{H}$ could be used as a universe of discourse for $P(x)$. If T is the set of all turtles that ever roamed this planet, then T could also be used as a universe of discourse for $P(x)$.

Just as a function such as $f(x) = \sqrt{x-10}$ has a domain, if we think of a predicate as a type of function, its universe of discourse is meant to serve the role of the domain of such a function. The outputs of this function would be either true or false, depending on what input is used.

Remark 127

When an algebra student struggles with finding $f(x+h)$ if $f(x) = \sqrt{x-10}$ due to the notation clash, we encourage the student to write $f(y) = \sqrt{y-10}$ instead to find $f(x+h)$. Just as $f(x) = \sqrt{x-10}$ has a "placeholder variable" x and the same function could be rewritten $f(y) = \sqrt{y-10}$, the x appearing in our example predicate is a placeholder.
So, think $P(y)$ defined by $\boxed{y \text{ was born in the year 1918}}$ as being the same predicate as defined in our earlier example.

The purpose of a universe of discourse is to set our minds in the right direction regarding a specific predicate. Our earlier substitutions included people, examples of nouns. Thus, it would be grammatically correct to substitute any noun for x in the sentence "x was born in the year 1918." While it might be amusing to consider the sentence "Happiness was born in the year 1918" when playing a game of Mad Libs, we no longer have a proposition (a sentence which is either true or false, but not both). Specifying a universe of discourse is intended to avoid these awkward situations where the resulting sentence, though correct in English grammar, has no readily apparent truth value.

If $Q(x)$ is a predicate, then $Q(x)$ is neither true nor false. If x is substituted with an element from the universe of discourse, this results in a proposition, which is either true or false.

Example 128. *Suppose $Q(x)$ is the predicate $\boxed{x^2 < 82}$ with $\mathbb{R}$ as the universe of discourse. Then, it makes no sense to say that $Q(x)$ is true or false. However, $Q(9)$ is true, while $Q(10)$ is false. To have a truth value, you must plug in something for x. We just saw that $Q(x)$ didn't have a truth value, while $Q(9)$ did. By analogy, using $f(x) = \sqrt{x - 10}$ from earlier, $f(x)$ does not have a numerical value, but $f(26)$ does. In fact, $f(26) = 4$.*

The process of taking a predicate and obtaining a proposition involves taking a **free** variable and **binding** it to have a **bound** variable. There are two ways to **bind** a variable. The first way is the one that we have seen. Using the previous example, $Q(x)$ is a predicate, but if the variable x is bound to have value 9, then $Q(9)$ is a proposition, specifically the true proposition $9^2 < 82$. Continuing this example, we could instead take the predicate $Q(x)$ and bind the variable x to the value 10, and get the proposition $Q(10)$, which is $10^2 < 82$, which is a false proposition. A second way to bind a variable is using a quantifier. In a moment, we will learn about both quantifiers.

Warning 129

Due to the grammatical role of binding a variable, a variable may not be bounded more than once. In other words, binding can only be applied to free variables. The process of binding turns a free variable into a bound variable.

2.4.2 Universal quantifier

Starting with a predicate $P(x)$, substituting x with an element from the universe of discourse is one way to get a proposition. Another way to get a proposition is to do the thought experiment of substituting *every* element from the universe of discourse for x, which uses a **quantifier**:

Definition 130: Universal quantification

Given a predicate $P(x)$, the **universal quantification** of $P(x)$ over the set U is the proposition that is true if $P(x)$ is true for all $x \in U$ and is false otherwise. The universal quantification is written $\forall x \in U\, [P(x)]$ and is read, "For all x in U, $P(x)$."

Occasionally, the universal quantification is read, "$P(x)$ for all x in U." In Section 2.6, we will explain why we use this language sparingly.

Example 131. *Recall $\mathbb{H}$ denotes the set of all people. Define $P(x)$ to be the predicate "x was kung fu fighting." Then $\forall x \in \mathbb{H}\, [P(x)]$ is the proposition "For all x which is a human, x was kung fu fighting." This can be said more succinctly: Everybody was kung fu fighting. It is probably not the case that* everybody *was kung fu fighting, so $\forall x \in \mathbb{H}\, [P(x)]$ is false.*

Stating "For all x which is a human, for all x which is a human, x was kung fu fighting" would be grammatically incorrect (see Warning 129) due to binding x more than once.

Our first several examples are focused on making sure that the meaning of these statements is clear. In order to do that, we have to focus on taking symbolic language and translating this into plain but accurate English. At first, this will feel challenging and confusing to do, but it is very important to focus on what things mean instead of hoping to slip by just by doing symbolic manipulation.

Example 132. *Suppose that $P(x)$ is the predicate "x has a logo" and that C is the set of all car brands. We can take $\forall x \in C[P(x)]$ and say "For all $x \in C$, x has a logo" and while this is technically accurate, it's much more helpful to understand if we can use plain language. Here are other examples that are accurate, but not plain:*

- *For all x in the set of car brands, x has a logo*
- *For all x in C, where C is the set of all car brands, x has a logo.*

- *For all x in C, x has a logo, where C is the set of all car brands.*
- *For all things in the set of car brands, that thing has a logo*

A truly accurate and plain way to say this is to say "Every car brand has a logo" or to say "Each car brand has a logo." If there's even one car brand out there without a logo, then this statement would be false. But if every car brand has a logo, then this statement would be true.

Example 133. *Let $T(x)$ be the predicate "x has an orange cowboy hat" and let N be the set of all New York residents (or if we wish, we could say that N is the set of all New Yorkers). When we discuss $\forall x \in N[T(x)]$ we tend to say and write "For all $x \in N$, x has an orange cowboy hat." If we were trying to say plainly what this means, it's not enough to say "For all x in the set of New Yorkers, x has an orange cowboy hat." Instead, we should say "Every New Yorker has a cowboy hat" or say "Each New Yorker has a cowboy hat." The proposition is false: to be false, we only need information that at least one New Yorker doesn't have a cowboy hat, and that seems like easy enough information for us to ascertain.*

Example 134. *Let Z be the set of all Canadians. Let $J(x)$ be the predicate "x likes to drink boba tea." Then when writing $\forall x \in Z[J(x)]$, we tend to say "For all $x \in Z$, x likes to drink boba tea" so that our speaking mimicks our writing. However, to plainly (but still accurately) describe what $\forall x \in Z[J(x)]$ means, we could say "Every Canadian likes to drink boba tea." I don't know about you, but I'd believe that "Every Canadian likes to drink boba tea" is false.*

Example 135. *Suppose that there was a small town (called Uruapan) which had the following twelve residents:*

Name	Address	Phone	Blood type	Birthday
Al	3064 Ross Street	910-390-9402	O-	May 12, 1960
Bo	2901 Florence Street	870-432-0142	AB-	Mar. 4, 1935
Cloyne	3290 Hickory Ridge Drive	501-216-6782	AB+	Jul. 16, 1978
Destiny	3195 Pine Street	605-643-1492	A+	Jan. 30, 1933
Echo	3064 Ross Street	334-200-6032	B-	Dec. 25, 1961
Finley	543 Station Street	301-386-8822	O+	Jan. 12, 1998
Gal	998 Armory Road	443-447-1792	AB-	Apr. 3, 1970
Hollis	3229 Pine Street	469-576-3672	A-	Sep. 29, 1992
Irving	3064 Ross Street	254-563-7422	A-	Apr. 21, 1952
Jackson	543 Station Street	580-643-7212	B+	Jun. 1, 1966
Kylar	4248 Rosebud Avenue	580-461-6572	O-	Jul. 24, 1997
Lee	773 Hart Ridge Road	217-202-2432	A+	May 13, 1948

That is, Uruapan (for which we will use U as notation) consists of these twelve people, and only these twelve people, and nobody other than these twelve people.

The proposition "For all $p \in U$, the last digit in p's phone number is a 2" is true.

Remark 136. *To preview something in the next chapter, let's point out that in order to* prove *the proposition "For all $p \in U$, the last digit in p's phone number is a 2", well, I would ask you (the person reading) to select to represent p any object from U that you like. That is, you could have p be* whichever *resident of Uruapan that you'd like. You don't even have to tell me who. You might have selected Finley. You might have selected Hollis. No matter who you picked, though, if I could write down that person's phone number accurately, and it ended in a 2, then I will have proved my point. The key thing though is that I'd have to perform this feat* no matter *who you selected. Or, if we had the time, you could keep selected elements of U (and there's only 12 anyway), and after exhausting all possibilities, if I demonstrate each resident has a 2 as th elast digit of their phone number, then I've proved my point.*

Example 137. *Let $M(x)$ be the predicate "$x^2 > 100$" and let $S = \{11, 12, 13\}$. Then $\forall x \in S\,[M(x)]$ is true. However, with $T = \{10, 20, 30, 40\}$, the proposition $\forall x \in T\,[M(x)]$ is false. Since S only had three elements, we could check whether this "for all" statement was true by plugging in each element of S for x and verifying whether $x^2 > 100$ was true or not. Similarly, T only had four elements, so we could plug in* *each of the four elements of T into x. Plugging in $x = 10$ made $x^2 > 100$ false, and therefore $\forall x \in T\,[M(x)]$ is false.*

Remark 138

Revisiting the last example, specifying the set used (for the universe of discourse) made a difference. A "for all" statement (a universally-quantified statement) is very "strong" in the sense that $M(x)$ must be true for *each* element of the set specified.

In fact, $\forall x \in S\,[M(x)]$ is logically equivalent to $M(11) \wedge M(12) \wedge M(13)$. Likewise, $\forall x \in T\,[M(x)]$ is logically equivalent to $M(10) \wedge M(20) \wedge M(30) \wedge M(40)$.

A universally-quantified statement is very "strong" in that many individual propositions have to be true. In fact, if the set L consists of 27 elements, then $\forall x \in L\,[P(x)]$ can be rewritten using the word "and" a total of 26 times. This idea extends for a set such as $\mathbb{Z}_{>0}$ with an unlimited number of elements. So, if $P(x)$ is some proposition, we can think of $\forall x \in \mathbb{Z}_{>0}\,[P(x)]$ as the same as $P(1) \wedge P(2) \wedge P(3) \wedge P(4) \wedge \cdots$, but we won't have time to write out the whole proposition using conjunctions.

Again, consider $\forall x \in S\,[M(x)]$ but for a new purpose. This says we should write $M(x)$ replacing x with 11, then with 12, and then with 13. So we get $M(11) \wedge M(12) \wedge M(13)$. What would $\forall y \in S\,[M(y)]$ say? We should write $M(y)$ replacing y with 11, then with 12, then with 13. So we get $M(11) \wedge M(12) \wedge M(13)$, the same as before.

Remark 139: The quantified variable is a placeholder variable

Recall Remark 127 where the predicate $P(x)$ variable x should be thought of the same as $P(y)$ with variable y. In the same way, if $P(x)$ is a predicate, then $\forall x \in S\,[P(x)]$ is the same proposition as $\forall y \in S\,[P(y)]$. The quantified variable is a placeholder variable.

If an algebra student struggles with applying $a(b + c) = ab + ac$ on the expression $5(a + 2)$ due to the notation clash, we would encourage the student to apply $x(y + z) = xy + xz$ on the expression $5(a + 2)$. Similarly, if x is already used in a proof and you encounter $\forall x \in S\,[P(x)]$, this remark tells us it's okay to replace this with $\forall y \in S\,[P(y)]$. This isn't strictly required, but removing the notation clash may be helpful at first.

Example 140. *Let $C(x)$ be the predicate "x was born in New York." Then $\forall x \in \mathbb{H}\,[C(x)]$ is the same proposition as $\forall s \in \mathbb{H}\,[C(s)]$. In plain English, both propositions say "Every person was born in New York." Clearly, $\forall t \in \mathbb{H}\,[C(t)]$ is false.*

Like the last example shows, practice writing the same universally-quantified statement using *different* variables.

Example 141. *Given S is the set of all tuba players and $P(x)$ is the predicate "x likes pumpkin pie" then $\forall x \in S[P(x)]$ is "For all $x \in S$, x likes pumpkin pie" or in plain English "Every tuba player likes pumpkin pie." Switching variables, $\forall y \in S[P(y)]$ is "For all $y \in S$, y likes pumpkin pie" which in plain English is still "Every tuba player likes pumpkin pie."*

2.4.3 Existential quantifier

We started with a predicate $P(x)$ and examined the universal quantification. There is another **quantifier** which yields a proposition:

Definition 142: Existential quantification

Given a predicate $P(x)$, the **existential quantification** of $P(x)$ over the set U is the proposition that is true if $P(x)$ is for at least one $x \in U$ and is false otherwise. The existential quantification is written $\exists x \in U\,[P(x)]$ and is read, "There exists x in U such that $P(x)$."

Occasionally, the existential quantification is read, "$P(x)$ for some x in U." In Section 2.6, we will explain why we use this language sparingly. The phrase "There exists" is used at the beginning of the sentence, while the phrase "for some" is used at the end, due to English grammar.

Language Discussion 143: What does the phrase "such that" mean?

The phrase "such that" is used to indicate that the text before and the text after are connected in some way. While mathematicians don't usually read $\exists x \in U\,[P(x)]$ as, "There exists x in U, and $P(x)$ is true," this would technically be a possible way to comprehend this proposition. The phrase "α such that β" is used to say α is true, β is true, and in some way, α and β are connected. In this case, the connection between the sentence "There exists $x \in U$" and the sentence "$P(x)$ is true" is that both mention x. The universal quantification does not use the phrase "such that."

Example 144. *Let Y be the set of all U.S. states. Let $P(s)$ be the predicate "s begins with the letter Z." Then the proposition $\exists s \in Y\,[P(s)]$ in plain English says "There exists a U.S. state which begins with the letter Z" which is clearly false. Using the word "which" was natural for the English language, but a slightly more mathematical verbiage of this would have been "There exists a U.S. state y such that y begins with the letter Z." The use of the phrase "such that" is a connecting phrase. In a sense, it would have been the same to say "There exists a U.S. state y" and "y begins with the letter Z" but because the two sentences are linked (by both having y) the phrase "such that" is preferred over "and" here.*

Example 145. *Let Y be the set of all U.S. states. Let $Q(t)$ be the predicate "The state vegetable of t is the watermelon." The proposition $\exists t \in Y\,[Q(t)]$ is true because the state vegetable of Oklahoma is the watermelon.*

Before digging in too far, we want to make sure that we practice what the existential quantifier looks like when put into plain English.

Example 146. *Suppose that $P(x)$ is the predicate "x has a logo" and that C is the set of all car brands. We can take $\exists x \in C[P(x)]$ and say "There exists $x \in C$ such that x has a logo" and while this is technically accurate, it's much more helpful to understand if we can use plain language. Here are other examples that are accurate, but not plain:*

- *There exists an x in the set of car brands such that x has a logo*
- *There exists x in C, where C is the set of all car brands, such that x has a logo.*
- *There exists x in C such that x has a logo, where C is the set of all car brands.*

A truly accurate and plain way to say this is to say "There exists a car brand has a logo" or to say "At least one car brand has a logo." To be true, we just need at least one car brand to have a logo. To be false would mean that there isn't even one car brand out there with a logo.

Example 147. *Let $T(x)$ be the predicate "x has an orange cowboy hat" and let N be the set of all New York residents (or if we wish, we could say that N is the set of all New Yorkers). When we discuss $\exists x \in N[T(x)]$ we tend to say and write "There exists $x \in N$ such that x has an orange cowboy hat." If we were trying to say plainly what this means, it's not enough to say "There exists x in the set of New Yorkers such that x has an orange cowboy hat." Instead, we should say "There exists a New Yorker that has a cowboy hat" or say "There is a New Yorker that has a cowboy hat" or "Some New Yorker has a cowboy hat." To be true, among all the New Yorkers, we would just need one person to have an orange cowboy hat. To be false, nobody from New York could own an orange cowboy hat.*

Example 148. *Let Z be the set of all Canadians. Let $J(x)$ be the predicate "x likes to drink boba tea." Then when writing $\exists x \in Z[J(x)]$, we tend to say "There exists $x \in Z$ such that x likes to drink boba tea" so that our speaking mimicks our writing. However, to plainly (but still accurately) describe what $\exists x \in Z[J(x)]$ means, we could say "There exists a Canadian who likes to drink boba tea." I don't know about you, but I'd believe that "There exists a Canadian likes to drink boba tea" is true: after interviewing every Canadian person and asking each to answer "Do you like boba tea?" it seems likely that a response of "Yes" will happen at least once.*

An existentially-quantified statement is "weak" in the sense that only one of many individual propositions has to be true:

Example 149. *Let $M(x)$ be the predicate "$x^2 > 100$" and let $T = \{10, 20, 30, 40\}$. The proposition $\exists x \in T\,[M(x)]$ is true. Since T only has four elements, we could plug in each of the four elements of T into x. Plugging in $x = 20$ made $x^2 > 100$ true, and therefore $\exists x \in T\,[M(x)]$ is true.*

In fact, with the notation set up from the last example, $\exists x \in T\,[M(x)]$ is logically equivalent to $M(10) \vee M(20) \vee M(30) \vee M(40)$.

Remark 150: The quantified variable is a placeholder variable

Just as Remark 139 says for universally-quantified statements, the quantified variable in an existentially-quantified statement is a placeholder variable. Thus, if $P(x)$ is a predicate and D is a set, then $\exists x \in D\,[P(x)]$ is the same proposition as $\exists t \in D\,[P(t)]$. In the next chapter on proofs, there will be times when writing $\exists t \in D\,[P(t)]$ to replace $\exists x \in D\,[P(x)]$ helps us avoid notation clashes.

As with universally-quantified statements, practice writing existentially-quantified statements using new placeholder variables.

Comparing the quantifiers

Warning 151: Do not use the word "for" alone

Suppose $P(x)$ is a predicate. In their alternate wordings, "$P(x)$ for all $x \in M$" is the universal quantification while "$P(x)$ for some $x \in M$" is the existential quantification. These are *different* propositions. Therefore, it is ambiguous to say "$P(x)$ for $x \in M$." Never use the word "for" alone right before "$x \in M$" or right before "$y \in N$" or similar. Always use "for all" or "for some" to clearly communicate whether you are talking about the universal quantification or the existential quantification.

Language Discussion 152: Grammar of quantified statements

A universally-quantified statement and an existentially-quantified statement have different grammar. If $P(x)$ is the predicate "Google knows x's phone number" then $\forall x \in \mathbb{H}\,[P(x)]$ is the proposition "For all $x \in \mathbb{H}$, Google knows x's phone number" while $\exists x \in \mathbb{H}\,[P(x)]$ is the proposition "There exists an $x \in \mathbb{H}$ such that Google knows x's phone number." The first sentence does not use the phrase "such that" and requires a comma (because of the dependent clause: the phrase "for all $x \in \mathbb{H}$" is not a sentence on its own). The second sentence uses the phrase "such that" and should not have a comma (because the phrase "There exists an $x \in \mathbb{H}$" is a sentence).

Example 153. *This example refers back to the table in Example 135.s The proposition "There exists $m \in U$ such that m has a phone number that starts with a 7" is false. The proposition "There exists $m \in U$ such that m has a B- blood type" is true. (Namely, Echo has this blood type.) The proposition "There exists $n \in U$ such that n lives on Fillmore Street" is false.*

The proposition "There exists $b \in U$ such that the birth year of b is less than or equal to 1930" is false. Another way to express this is to say that "For all $b \in U$, the birth year of b is greater than 1930" is true. Notice that one of these sentences has the phrase "less than or equal to 1930" while the other sentence has the phrase "greater than 1930." Negating quantified statements is the subject of the next section.

Warning 154: Do not ignore quantifiers

When learning definitions, reciting definitions, and eventually when proving statements, it is dangerous to ignore quantifiers. Let's consider the following example, where $S = \{6, 7, 8, 9\}$. The statement $\forall x \in S[x^2 < 40]$ is false, while the statement $\exists x \in S[x^2 < 40]$ is true. So, one should not write 'for all' where 'there exists' is appropriate and vice versa. Do not switch which quantifier appears, but in addition, do not just "ignore" quantifiers altogether. The pair of examples here that $\forall x \in S[x^2 < 40]$ has a different truth value from $\exists x \in S[x^2 < 40]$ gives all the evidence that writing $\boxed{x^2 < 40}$ alone is insufficient as an attempt to mean $\exists x \in S[x^2 < 40]$ if that is what is meant, or to mean $\forall x \in S[x^2 < 40]$ if that is what is meant. So we can't ignore or leave out quantifiers: they are not optional, and should not be treated as optional.

2.4.4 Revisiting what quantifiers mean

For this subsection, let us fix the set S to be $S = \{1, 2, 3, 4\}$. Suppose that $P(k)$ is a predicate and that S is its universe of discourse, so that $P(k)$ isn't true or false, but $P(1)$ is either true or false, $P(2)$ is either true or false, and so on.

In the same way that

$$\sum_{k=1}^{4} f(k) = f(1) + f(2) + f(3) + f(4)$$

we have

$$\forall k \in S[P(k)] \equiv P(1) \wedge P(2) \wedge P(3) \wedge P(4),$$

so just as sigma notation is a compact way of writing a sum of many numbers, "for all" can be thought of as a compact way of writing a conjunction of many propositions. Likewise, we can think of "there exists" in the following manner:

$$\exists k \in S[P(k)] \equiv P(1) \vee P(2) \vee P(3) \vee P(4).$$

When writing

$$\sum_{k=1}^{4}$$

before $f(k)$, the notation is asking us to have the value of k take on each number between 1 and 4, including the endpoints. In other words, we are to take the values of $f(k)$ for each $k \in \{1, 2, 3, 4\}$ and add the results together. Similarly, in the case of

$$\forall k \in S[P(k)]$$

we are to take the [logical] values of $P(k)$ for each $k \in \{1, 2, 3, 4\}$ and then "and" the results together.

2.4.5 Negating quantified statements

Section 2.3 introduced logical equivalence, which allows us to convert (if done carefully) an implication into a disjunction. In addition, De Morgan's laws give us a way to simplify the writing of the negation of either a conjunction or a disjunction. How can we negate a quantified statement?

Consider, for example, the proposition $\forall x \in A\,[P(x)]$. It would be tempting to say that the negation of this proposition is $\forall x \in A\,[\neg P(x)]$, but that would be incorrect.

Example 155. *Recall the notation from Example 131, where $\forall x \in \mathbb{H}\,[P(x)]$ said succinctly is "Everybody was kung fu fighting." I once saw a bumper sticker that said, "Surely* somebody *was not kung fu fighting." In fact, $\neg\forall x \in \mathbb{H}\,[P(x)]$ is true, and as we see from the bumper sticker text, $\exists x \in \mathbb{H}\,[\neg P(x)]$ captures the same meaning, but in different words.*

Method 156: Negating a universally-quantified statement

The negation of $\forall x \in A\,[P(x)]$ is $\exists x \in A\,[\neg P(x)]$. In symbols,

$$\neg\forall x \in A\,[P(x)] \equiv \exists x \in A\,[\neg P(x)].$$

One can consider $\neg(a \wedge b \wedge c \wedge d)$ being logically equivalent to $(\neg a) \vee (\neg b) \vee (\neg c) \vee (\neg d)$ as an extended version of De Morgan's law. While we provided the kung fu example to convince readers that Method 156 is telling the truth, here is a more complete way to reason through this. Suppose $A = \{1, 2, 3, 4\}$. Then $\forall x \in A\,[P(x)]$ is logically equivalent to $P(1) \wedge P(2) \wedge P(3) \wedge P(4)$. Thus, using the extended version of De Morgan's law just mentioned, the negation of the previous proposition should be $[\neg P(1)] \vee [\neg P(2)] \vee [\neg P(3)] \vee [\neg P(4)]$ which is logically equivalent to $\exists x \in A\,[\neg P(x)]$.

Note $\forall x \in \mathbb{H}\,[\neg P(x)]$ says "Everybody was not kung fu fighting." This is not the correct negation of $\forall x \in \mathbb{H}\,[P(x)]$. Similarly, we don't say that the reason "Everybody has green eyes" is false is because "Everybody does not have green eyes." Instead, we say that the reason "Everybody has green eyes" is false is because "There is somebody that does not have green eyes."

Warning 157

The negation of $\forall x \in A\,[P(x)]$ is not $\forall x \in A\,[\neg P(x)]$.

To ensure all of the discussion above is clear, let's put into plain English the propositions that have come up:

- $\forall x \in \mathbb{H}\,[P(x)]$ is: Everybody was kung fu fighting.
- $\forall x \in \mathbb{H}\,[\neg P(x)]$ is: Everybody was not kung fu fighting.
- $\exists x \in \mathbb{H}\,[P(x)]$ is: Somebody was kung fu fighting.
- $\exists x \in \mathbb{H}\,[\neg P(x)]$ is: Somebody was not kung fu fighting.

Now, if we negate each of these four propositions (in the same order as presented), we have:

- $\neg\forall x \in \mathbb{H}\,[P(x)]$ is: It isn't true that everybody was kung fu fighting. (If we interviewed everybody and asked "Were you kung fu fighting?" at least one person would answer no.)
- $\neg\forall x \in \mathbb{H}\,[\neg P(x)]$ is: It isn't true that everybody was not kung fu fighting. (If we interviewed everybody and asked "Were you kung fu fighting?" at least one person would answer yes.)
- $\neg\exists x \in \mathbb{H}\,[P(x)]$ is: Not even one person was kung fu fighting. (If we interviewed everybody and asked "Were you kung fu fighting?" at everyone wwould answer no.)
- $\neg\exists x \in \mathbb{H}\,[\neg P(x)]$ is: Not even one person was not king fu fighting. (If we interviewed everybody and asked "Were you kung fu fighting?" every person doesn't answer no, meaning every person answers yes.)

Let us consider another incorrect negation: $\forall x \notin \mathbb{H}\,[P(x)]$ would in plain English become "Every non-human was kung fu fighting." This is not the correct negation of $\forall x \in \mathbb{H}\,[P(x)]$. (Why, in discussing whether every human was kung fu fighting or not, would we start up a discussion of whether chameleons were kung fu fighting?)

Warning 158: Do not negate the set membership

The negation of $\forall x \in A\,[P(x)]$ is not $\forall x \notin A\,[P(x)]$.

Now that we have discussed how to negate universally-quantified statements, how would we negate an existentially-quantified statement such as $\exists x \in A\,[P(x)]$?

Example 159. *Let A be the set of all people born before the year 1800. So Martha Washington is in the set A, while Neil Armstrong is not in the set A. Let $P(x)$ be the predicate "x has been to the moon." In plain English, the proposition $\exists x \in A\,[P(x)]$ says "There exists a person born before 1800 who has been to the moon," which is of course false. Another way to capture the same meaning is by saying "Every person born before 1800 has not been to the moon."*

From this example, we gather the pattern:

Method 160: Negating an existentially-quantified statement

The negation of $\exists x \in A\,[P(x)]$ is $\forall x \in A\,[\neg P(x)]$. In symbols,

$$\neg \exists x \in A\,[P(x)] \equiv \forall x \in A\,[\neg P(x)].$$

Besides looking at our moon example, you should convince yourself that this is the proper negation via an extended De Morgan's law which says that $\neg(a \vee b \vee c \vee d)$ is logically equivalent to $(\neg a) \wedge (\neg b) \wedge (\neg c) \wedge (\neg d)$. We'll skip these details.

It is tempting to say that the negation of $\exists x \in A\,[P(x)]$ is $\exists x \in A\,[\neg P(x)]$, but this is incorrect. Note the last sentence says "There is a person born before 1800 who has not been to the moon." While this sentence is true, it does not truly express why the sentence "There is a person born before 1800 who has been to the moon" is false! Therefore:

Warning 161

The negation of $\exists x \in A\,[P(x)]$ is not $\exists x \in A\,[\neg P(x)]$.

Let us consider another incorrect negation: $\exists x \notin A\,[P(x)]$ would in plain English become "There exists an x not born before 1800 who has been to the moon." This is not the correct negation of $\exists x \in A\,[P(x)]$. As earlier, we do not negate the set membership:

Warning 162: Do not negate the set membership

The negation of $\exists x \in A\,[P(x)]$ is not $\exists x \notin A\,[P(x)]$.

2.5 Mathematical language

Be patient with the process, but keep challenging yourself to work on mathematical language. If you are at this point in the handbook, you are so close to having all the knowledge/skills to fully connect with (for example) the proof of the Intermediate Value Theorem, a statement which you learned in calculus and had to just accept with blind faith.

The main struggle for students at this point tends to be the language surrounding quantified statements. Work on understanding these phrases *completely*. Let S be the set of all people who attend your school. Then consider these two propositions:

- For all $a \in S$, a prefers mechanical pencils over wood pencils.
- There exists $a \in S$ such that a prefers mechanical pencils over wood pencils.

Think about the grammar: why is there a comma in the first proposition, but not the second? By replacing the text before the comma in the first proposition yet try to retain the meaning, we might say, "No matter which $a \in S$ that you think of, a prefers mechanical pencils over wood pencils." That comma is a natural pause even in speaking that sentence aloud.

The second proposition (the existentially-quantified statement) might be restated, "There's someone (named a) in S who prefers mechanical pencils over wood pencils." Is there a place one would pause in speaking that sentence? Probably not. The second sentence has the phrase "such that." While the second sentence could *functionally* be thought of as "There exists $a \in S$ and a prefers mechanical pencils over wood pencils," the use of "such that" stresses that there is a link between the part which says "There exists $a \in S$" and the part which says "a prefers mechanical pencils over wood pencils." What is the common link? It is a.

Even further, why the phrase "such that"? Well, if S has one or more elements, then it doesn't really say much to say "There exists $a \in S$," now, does it? So we could think of the fact that the same a appears in both places and say that a (much more drawn-out) version of the sentence is "There exists an $a \in S$, but not only is there an a in S, but in addition, a prefers mechanical pencils over wood pencils."

Take time to thoroughly think about the grammar of these sentences (down to the appearance/absence of a comma, and the appearance/absence of the phrase "such that", where notation appears relative to words, etc.). This will be crucial in the next section on nested quantifiers, and a clear understanding of what these sentences are saying will be required in order to *prove* these propositions in Chapter 3.

2.6 Nested quantifiers

While it is possible to give a full definition of a **predicate** of two variables, as a variant to Definition 124, it will be natural to just give some examples instead:

Example 163. *We define $P(x, y)$ to be the predicate "x has given y a high five." Then $P(x, y)$ is a two-variable predicate, and a reasonable universe of discourse to use for both the variable x and the variable y is $\mathbb{H}$, the set of all humans.*

If a is Katy Perry and b is Vincent Van Gogh, then $P(a, b)$ is a false proposition. If u is you and v is your math instructor, you can determine for yourself if $P(u, v)$ is true or false. If w is your best friend, then I'm bet that $P(u, w)$ is true.

Example 164. *Let $Q(s, t)$ be the two-variable predicate "s graduated with a t degree." Based on the text for $Q(s, t)$, it does not seem so natural to use the same set as the universe of discourse for the variable s and for the variable t. Let E be the set of all people who have graduated from your university, and let M be the set of all the majors offered at your university. Then it would make sense to use E as the universe of discourse for the variable s, and to use M as the universe of discourse for the variable t.*

From the previous example, notice that a two-variable predicate does not have to have the same set used as universe of discourse for each variable.

Example 165. *Let $R(m, n)$ be "m has sent n an email." If v is your math instructor and u is you, then I'd guess that $P(v, u)$ is likely true. If o and p are two randomly chosen people in $\mathbb{H}$, it is quite possible that $R(o, p)$ is true while $R(p, o)$ is false. (That is, o has sent p an email, but p has not sent o an email.)*

As with predicates with one variable, one way to create a proposition is to substitute a single element from the universe of discourse and the other way is to quantify a variable. When we look at predicates with two variables, we can quantify each variable. (In fact, each variable may only be quantified once.)

Example 166. *Using the set up of Example 163, the proposition $\forall x \in \mathbb{H}\, [\forall y \in \mathbb{H}\, [P(x, y)]]$ says "For all humans x, for all humans y, x has given y a high five" or in more plain language, "Everyone has given everyone a high five." While it's amusingly silly, for this proposition to be true, each person must have given themselves a high five (a "self-five"?), but much more than that would need to be true. (This proposition is definitely false.) It is possible to convey the same meaning while completely leaving out the square brackets, so we can write $\forall x \in \mathbb{H}\, \forall y \in \mathbb{H}\, P(x, y)$ instead.*

The proposition $\exists x \in \mathbb{H}\, [\exists y \in \mathbb{H}\, [P(x, y)]]$ says "There exists a person x such that there exists a person y such that x has given y a high five." In plain language, "Someone has given someone a high five." If ever a high five occurred in human history, then this proposition would be true, so this proposition is true. It may be written without the square brackets, as we do in the next example:

The proposition $\exists x \in \mathbb{H}\, \forall y \in \mathbb{H}\, P(x, y)$ says "There exists a person x such that for all people y, x has given y a high five." More plainly, this says, "Someone has given everyone a high five." For this to be true, there must be an individual (which we are calling x) who has performed a high five with every *single person on the planet. Certainly, $\exists x \in \mathbb{H}\, \forall y \in \mathbb{H}\, P(x, y)$ is false.*

The proposition $\forall x \in \mathbb{H}\, \exists y \in \mathbb{H}\, P(x, y)$ says "For all people x, there is a person y such that x has given y a high five." More plainly, this says, "Everyone has given someone a high five." For this to be true, each person in the world only needs to have given one high five in their life (more than one is fine too). While you might say it is plausible for this to be true, there are places around the word for which high fiving is not part of the culture, so it is likely the case that $\forall x \in \mathbb{H}\, \exists y \in \mathbb{H}\, P(x, y)$ is false.

In the last example, we discussed that $\exists x \in \mathbb{H}\, \forall y \in \mathbb{H}\, P(x, y)$ is false and we also guessed that $\forall x \in \mathbb{H}\, \exists y \in \mathbb{H}\, P(x, y)$ is false. These two propositions were false for *different* reasons precisely because they really are *different* statements. The general warning to are to learn from this is:

Warning 167: We cannot generally switch the order of quantifiers

Suppose $P(x, y)$ is a predicate of two variables x and y. In general, the proposition $\forall x \in A\, \exists y \in B\, P(x, y)$ and the proposition $\exists y \in B\, \forall x \in A\, P(x, y)$ are in general different:

Both may be true (but for different reasons), both may be false (but for different reasons), or it may even be the case that one proposition is true while the other proposition is false.

In general, the order of the quantifiers may not be swapped, as this creates a change in meaning. More specifically, the order of a universal quantifier and an existential quantifier may not be swapped, because there will be a change in meaning.

Consider the case of a single quantifier. If $P(x)$ is "x has used a 3D printer" and C the set of all students at your school, then $\forall x \in C\, P(x)$ could be spoken aloud either "For all x at your school, x has used a 3D printer" or "x has used a 3D printer for all x at your school." We mentioned that we'll generally avoid this second form. The reason is due to the previous warning:

Phrasing like "There exists $a \in C$ such that $P(a, b)$ for all $b \in D$" makes it truly ambiguous whether this is $\exists a \in C\, \forall b \in D\, P(a, b)$ or if this is $\forall b \in D\, \exists a \in C\, P(a, b)$. Recall from the last two paragraphs of Example 2.6 that these two propositions are different. Therefore:

Language Discussion 168

When writing a statement with multiple quantifiers in words, the quantified text (whether that text is "For all .... in ...," or that text is "There exists ... in ... such that") will *always* be placed in the same order as the symbols appear. Thus, the text appears before the predicate text when there are multiple quantifiers, never after.

Thus, $\exists a \in C\, \forall b \in D\, P(a, b)$ will have words that stick to the symbol order and become "There exists $a \in C$ such that for all $b \in D$, $P(a, b)$" while the proposition $\forall b \in D\, \exists a \in C\, P(a, b)$ will become "For all $b \in D$, there exists an $a \in C$ such that $P(a, b)$."

How does $\forall b \in D\, \exists a \in C\, P(a, b)$ become true? Imagine a video camera recording a conversation between you (the speaker) and another person (the skeptic). A video plays and to interpret the part which says "For all $b \in D$" the skeptic chooses anything in D that they would like, but so that you can refer to it, both people agree to call it b. (The skeptic may keep the identity of b hidden.) For the proposition $\forall b \in D\, \exists a \in C\, P(a, b)$ to be true, among other things, with the b in D chosen, the statement $\exists a \in C\, P(a, b)$, and to interpret the $\exists a \in C$ part, the speaker must now choose an $a \in C$. However, the a that you choose must be chosen in such a way that $P(a, b)$ is true (so the choice must be made carefully).

Remark 169. *At any point, the skeptic may "rewind the tape." (In other words, if we think of the conversation like a video playing, the skeptic can bring the video back to an earlier point.) So the skeptic can go back and choose a new b in D, but then the skeptic must allow the speaker to choose a new $a \in C$ in reaction to the newly-chosen $b \in D$. If each individual can make their choices in a way that $P(a, b)$ ends up being true (that is, the speaker can "react" to the skeptic's choice of $b \in D$ by selecting an appropriate $a \in C$), then the proposition $\forall b \in D\, \exists a \in C\, P(a, b)$ is true. ao*

The discussion in the last two paragraphs is a bit abstract at first, but the idea of playing and rewinding a video is meant to give a strong intuition for what the logic is when there are multiple (that is, nested) quantifiers.

Example 170. *Let D be the set of all people with a cell phone. Let C be the set of all positive integers. Let $P(a,b)$ be the predicate "a can be contacted by calling the number b."*

Consider the proposition $\forall b \in D\, \exists a \in C\, P(a,b)$. A video starts. The skeptic chooses a person with a cell phone, and does not have to tell the speaker which person was chosen. (However, the speaker may need to refer to this individual, so both people agree to call the selected person b.) Then, the speaker assigns to a an appropriate value by decreeing "Let a be the cell phone number for b."

Now, it could be the case that there is another phone number (a work phone?) by which b can be reached. However, the speaker's decree is an appropriate sentence which causes $P(a,b)$ to be true.

Whether before the decree or after the decree, imagine rewinding the tape. That is to say, imagine the skeptic saying, "No, I change my mind! I want to select a different individual in D instead!" Well, that's fine, but after the skeptic has made their choice, the speaker would then again decree, "Let a be the cell phone number for b." and thus $P(a,b)$ would be true.

Again, the example may seem as weird as the discussion before it, but there is a strong foreshadowing of what happens in the next chapter.

Here is an actual example of a definition (from calculus!) involving multiple quantifiers:

Definition 171. *We say the function f **has a limit** L at the x-value a if for all $\varepsilon \in \mathbb{R}_{>0}$, there exists $\delta \in \mathbb{R}_{>0}$ such that if $0 < |x-a| < \delta$, then $|f(x) - L| < \varepsilon$.*

Each time a positive real number is chosen to be ε, a new δ may be chosen. (That is, δ is a function of ε.)

How should we negate a statement with multiple quantifiers? One at a time. Recall that the negation of a universally-quantified statement is an existentially-quantified statement, and similarly, the negation of an existentially-quantified statement is a universally-quantified statement.

Example 172. *To negate $\forall b \in D\, \exists a \in C\, P(a,b)$, we place a negation symbol in front and simplify:*

$$\begin{aligned}\neg\forall b \in D\, \exists a \in C\, P(a,b) &\equiv \exists b \in D\, \neg\exists a \in C\, P(a,b)\\ &\equiv \exists b \in D\, \forall a \in C\, \neg P(a,b).\end{aligned}$$

Example 173. *To negate $\forall a \in X\, \forall b \in Y\, Q(a,b)$, we place a negation symbol in front and simplify:*

$$\begin{aligned}\neg\forall a \in X\, \forall b \in Y\, Q(a,b) &\equiv \exists a \in X\, \neg\forall b \in Y\, Q(a,b)\\ &\equiv \exists a \in X\, \exists b \in Y\, \neg Q(a,b).\end{aligned}$$

We refer the reader to John Quintanilla's article "Name That Tune: Teaching Predicate Logic with Popular Culture" in *MAA Focus* August/September 2016 for a fun way to practice nested quantifiers.

Warning 174: A variable may not be re-quantified

Writing $\forall x \in A\,\exists x \in B\,P(x)$ makes no sense. A variable may not be quantified more than once. While there is a technical approach (discussing a thing called **binding**), we choose to take a more casual approach.

The purpose of introducing a variable is so that we can clearly communicate quantification. Phrasing that we tend to use in everyday life (example: "Everybody has a crush on someone.") tends to have a fairly clear sense of meaning. This is the case, even though, in the previous example, we haven't made any variables. If $Q(c,d)$ is "c has a crush on d" then it is fairly clear that the example in notation is $\forall a \in \mathbb{H}\,\exists b \in \mathbb{H}\,Q(a,b)$ instead of being $\exists b \in \mathbb{H}\,\forall a \in \mathbb{H}\,Q(a,b)$.

In a more long-winded approach, the beginning of $\forall a \in \mathbb{H}\,\exists b \in \mathbb{H}\,Q(a,b)$ can be said, "For all humans, and so we are clear which human we are addressing, let refer to that human as a no matter which human you (or anybody) should pick...'

The more long-winded phrasing makes clear why $\forall a \in \mathbb{H}\,\exists a \in \mathbb{H}\,Q(a,a)$ could not have any meaning. In the second quantification of a, we have already placed an identity on a earlier in the first quantification. A variable may not be quantified twice, as there is a complete lack of meaning in doing so. (In fact, each time we take a statement with a quantified variable in it, our plain English sentence to capture the meaning no longer even has that variable appear anymore, so this really goes to show why we can't quantify a variable more than once.)

Exercise 175. *Negate $(p \wedge q) \vee r$ and simplify so that all negation symbols immediately precede the basic propositions (symbols such as p, q, r, etc.).* [key]

Exercise 176. *Negate $(p \to q) \wedge \neg r$ and simplify so that all negation symbols immediately precede the basic propositions (symbols such as p, q, r, etc.).* [key]

Exercise 177. *Negate $(p \to q) \vee (r \to s)$ and simplify so that all negation symbols immediately precede the basic propositions (symbols such as p, q, r, etc.).* [key]

Exercise 178. *Negate the proposition $\forall a \exists b \forall c \exists d[P(a,b) \wedge Q(b,c,d) \wedge R(a,d)]$. Rewrite/simplify so that negations do not appear to the left of a quantifier.*

Exercise 179. *Negate and simplify $\exists a \forall b \forall c \exists d[P(a,b) \to Q(c,d)]$ so that all negation symbols immediately precede predicates.*

Exercise 180. *Define your own predicate $P(x,y)$ of two variables. You may not use one from class, the book, or the Internet. Provide a reasonable universe of discourse for the two variables. You should pick $P(x,y)$ in such a way that $\forall x \exists y[P(x,y)]$ is true and $\exists y \forall x[P(x,y)]$ is false. Briefly explain why of the two propositions, one is true and one is false. This shows that the $\exists y$ and the $\forall x$ cannot be swapped.*

Exercise 181. *Let $P(x,y)$ be the predicate "x sent y a thank you card in the mail." For both variables x and y, we use "all people" as the universe of discourse. Write $\forall x \exists y[P(x,y)]$ as an English sentence. Be careful to include any commas and phrases "such that" where they apply, and do not include them where they are inappropriate.*

Exercise 182. *Let $P(x,y)$ be the predicate "x sent y a thank you card in the mail." For both variables x and y, we use "all people" as the universe of discourse. Write $\forall y \exists x[P(x,y)]$ as an English sentence. Be careful to include any commas and phrases "such that" where they apply, and do not include them where they are inappropriate.*

Exercise 183. *Translate each of the following propositions into standard mathematical English, where the universe of discourse for each variable consists of all real numbers. Argue whether the statement would be true or false with any of the quantifiers swapped.*

- $\exists x \forall y[xy = y]$
- $\forall x \forall y[((x \geq 0) \wedge (y < 0)) \to (x - y > 0)]$

- $\forall x \forall y \exists z [x = y + z]$

NOTE: For the third proposition, note that there are TWO swaps to discuss: discuss swapping $\forall x$ with $\forall y$, and discuss swapping $\forall y$ with $\exists z$.

2.7 Examples of propositions

In this section, we introduce some propositions which are true, which (due to your previous mathematical experiences) we are certain you would agree to their veracity. Many of these facts (commutative laws for numbers, associative laws for numbers, etc.) will be referenced and used in proofs in the next chapter. To start, here are some facts about integers:

- Closure of addition: For all $a \in \mathbb{Z}$, for all $b \in \mathbb{Z}$, one has $a + b \in \mathbb{Z}$.
- Closure of subtraction: For all $a \in \mathbb{Z}$, for all $b \in \mathbb{Z}$, one has $a - b \in \mathbb{Z}$.
- Closure of multiplication: For all $a \in \mathbb{Z}$, for all $b \in \mathbb{Z}$, one has $ab \in \mathbb{Z}$.
- Addition is commutative: For all $a \in \mathbb{Z}$, for all $b \in \mathbb{Z}$, one has $a + b = b + a$.
- Multiplication is commutative: For all $a \in \mathbb{Z}$, for all $b \in \mathbb{Z}$, one has $ab = ba$.
- Addition is associative: For all $a \in \mathbb{Z}$, for all $b \in \mathbb{Z}$, for all $c \in \mathbb{Z}$, one has $(a + b) + c = a + (b + c)$.
- Multiplication is associative: For all $a \in \mathbb{Z}$, for all $b \in \mathbb{Z}$, for all $c \in \mathbb{Z}$, one has $(ab)c = a(bc)$.
- Distributive law: For all $a \in \mathbb{Z}$, for all $b \in \mathbb{Z}$, for all $c \in \mathbb{Z}$, one has $a(b + c) = ab + ac$.

While there is closure for addition, subtraction, and multiplication, the quotient of two integers may not be an integer. Similar facts hold for rational numbers:

- Closure of addition: For all $a \in \mathbb{Q}$, for all $b \in \mathbb{Q}$, one has $a + b \in \mathbb{Q}$.
- Closure of subtraction: For all $a \in \mathbb{Q}$, for all $b \in \mathbb{Q}$, one has $a - b \in \mathbb{Q}$.
- Closure of multiplication: For all $a \in \mathbb{Q}$, for all $b \in \mathbb{Q}$, one has $ab \in \mathbb{Q}$.
- Addition is commutative: For all $a \in \mathbb{Q}$, for all $b \in \mathbb{Q}$, one has $a + b = b + a$.
- Multiplication is commutative: For all $a \in \mathbb{Q}$, for all $b \in \mathbb{Q}$, one has $ab = ba$.
- Addition is associative: For all $a \in \mathbb{Q}$, for all $b \in \mathbb{Q}$, for all $c \in \mathbb{Q}$, one has $(a + b) + c = a + (b + c)$.
- Multiplication is associative: For all $a \in \mathbb{Q}$, for all $b \in \mathbb{Q}$, for all $c \in \mathbb{Q}$, one has $(ab)c = a(bc)$.
- Distributive law: For all $a \in \mathbb{Q}$, for all $b \in \mathbb{Q}$, for all $c \in \mathbb{Q}$, one has $a(b + c) = ab + ac$.

Similar facts hold for reals:

- Closure of addition: For all $a \in \mathbb{R}$, for all $b \in \mathbb{R}$, one has $a + b \in \mathbb{R}$.
- Closure of subtraction: For all $a \in \mathbb{R}$, for all $b \in \mathbb{R}$, one has $a - b \in \mathbb{R}$.
- Closure of multiplication: For all $a \in \mathbb{R}$, for all $b \in \mathbb{R}$, one has $ab \in \mathbb{R}$.
- Addition is commutative: For all $a \in \mathbb{R}$, for all $b \in \mathbb{R}$, one has $a + b = b + a$.
- Multiplication is commutative: For all $a \in \mathbb{R}$, for all $b \in \mathbb{R}$, one has $ab = ba$.
- Addition is associative: For all $a \in \mathbb{R}$, for all $b \in \mathbb{R}$, for all $c \in \mathbb{R}$, one has $(a + b) + c = a + (b + c)$.
- Multiplication is associative: For all $a \in \mathbb{R}$, for all $b \in \mathbb{R}$, for all $c \in \mathbb{R}$, one has $(ab)c = a(bc)$.

- Distributive law: For all $a \in \mathbb{R}$, for all $b \in \mathbb{R}$, for all $c \in \mathbb{R}$, one has $a(b + c) = ab + ac$.

We will use these facts above (and similar facts such as $a + 0 = a$ for all $a \in \mathbb{R}$) often in Chapter 3, where we learn how to use facts above alongside new definition in order to learn how to write our first proofs. (By "similar facts" we mean that $a^b a^c = a^{b+c}$ for all $a \in \mathbb{R}_{\neq 0}$ and for all $b, c \in \mathbb{R}$ and other similar facts will also be taken to be true.) There are some additional facts for real numbers which are helpful:

- For all $a, b \in \mathbb{R}$, we have $a \leq b$ if and only if $a < b$ or $a = b$.
- For all $a, b, c \in \mathbb{R}$, if $a \leq b$ and $b \leq c$, then $a \leq c$.
- Fix a positive number k. For all $x \in \mathbb{R}$, we have $|x| < k$ if and only if $x > -k$ and $x < k$.
- Fix a positive number k. For all $x \in \mathbb{R}$, we have $|x| \leq k$ if and only if $x \geq -k$ and $x \leq k$.
- For all $a, b \in \mathbb{R}$ and for all $c > 0$, we have $a + c \leq b$ if and only if $a < b$.

In addition to the facts above, we will also use these facts:

- If an integer s is not odd, then s is even.
- If an integer s is not even, then s is odd.

Chapter 3

Methods of proof

The previous chapters laid a foundation. Understanding how mathematical proofs work requires those previous chapters, but we haven't proved anything yet. This chapter introduces the main methods of proof. Work on thoroughly understanding each method introduced in this chapter.

Just as definitions are written in complete sentences, proofs are written in complete sentences. A proof starts with certain propositions (the **hypotheses** or the **premises**) assumed to be true, and ends with a proposition (the **conclusion**). A proof leads the reader from the hypotheses to the conclusion using water-tight arguments. The arguments used along the way are applications of **syllogisms**, which are also referred to as the **rules of inference**.

All of the rules of inference which we introduce in the next sections are informed by the definitions from Chapter 2. Consider the proposition $p \wedge q$. In some proofs, the proposition $p \wedge q$, having already been established to be true, will *lead* to another truth. This is *using* $p \wedge q$. In other occasions, truths will be combined together to *lead* to establishing $p \wedge q$. This is *proving* $p \wedge q$.

In due time, we will go through some complete proofs. Every proof will consist of using some propositions and proving other propositions. To reiterate, the truth of some propositions will lead to establishing other propositions as being true. So, some propositions are *used* to *prove* other propositions.

Warning 184: Using versus proving a proposition

Using a proposition and proving a proposition are very different tasks. Do not confuse these two tasks.

We will preview the details in a second, but it would be good to first pause and give the big picture. We will be given a *statement* to prove. (That statement is actually a proposition.) Then, we will write sentences that constitute a *proof* of that proposition. The previous chapter (about propositions) was important to make sure that we understood the *statement*, as most statements are compound propositions, built using language from logical operations and quantifiers. Thus, we first built up the foundation to know what is *meant* by the statements we will soon see. What about the proof? The proof is text that is written that follow the logic (the rules given by the rules of inference) which forms a convincing argument that the statement is true. The text of the proof is often several sentences, but let's start with an everyday analogy where the "proof" is only one sentence:

Again, what we are about to state is meant to be a helpful analogy. If someone makes the *statement* that "Someone went to the kitchen and ate a cookie." This person might *prove* this statement by showing the top of the cookie jar is no longer level, and there are clearly crumbs on the counter top. The proof is evidence Of the claimed statement. The statement says "Here's something that's true" while the proof is text that convinces the reader that the statement is true. However, the difference with this everyday version is that the "proof" provided isn't really a proof: it's evidence of plausibility. That is, someone might have just taken a cookie clumsily without having eaten it. The analogy breaks down, because what we will do is write proofs based on the rules of inference, which leave absoluetly no room for an alternative explanation.

So, what does a proof like? Even that, let's discuss by analogy first. The general structure of proofs that we will write will be to start with allowable assumptions, and then reach a final conclusion. From the assumption(s) that we are allowed, we will *use* these to *prove* an additional proposition. The new propositon that was just proved is a new fact that we know to be true, and the fact that this new proposition is true can then be *used* to *prove* a further proposition. While the analogy isn't perfect, it is helpful to imagine the following: Say that we start a journey by airplane at New York City and our final destination is Singapore. The itenerary we have might be a first flight from New York City to Los Angeles, and second flight from Los Angeles to Tokyo, and a third flight from Tokyo to Singapore. Now, is New York City a departure or arrival airport? It's a departure airport. Is Singaproe a departure or arrival ariport? It's our arrival airport; in fact it's our final destination. What about Los Angeles? Is it a departure or arrival airport? Well, both, but at different points in the story: in the early part of the story, when we just get on our first flight, Los Angeles is an arrival airport. A little bit later in the story, when we hop on our second flight, Los Angeles is (to us in that moment) a departure airport. Think of NYC where we started like the assumption(s) that we are allowed to make in our proof. Think of Singapore like the final conclusion of the proof. We will need to "stop" at some places along the way: think of the stops in Los Angeles and Tokyo like the *intermediate conclusions* in a proof: these are statements that are proved at first because they are helpful steps in our proof, but once proved, we are looking for a way to *use* these newly-true-to-us statements to prove something further.

The anology is helpful, but now let's actually explain what a proof looks like. Again, recall that there's a statement (the claim that a certain fact is true) and the proof of the statement (the argument that the claimed fact actually *is* true). Each kind of compound proposition to prove has a format for how to prove that proposition. (We will carefully study these.) While each of these will look different in their implementation, for a big picture, we will be allowed to assume one or more propositions. (By "assume", we mean that we will get to accept certain propositions as being true for free.) Then, based on these assumptions, we will apply any/all applicable rules of inference to end up learning a new proposition is true, and the new proposition is called an intermediate proposition. (We will say that we *used* the assumptions to *prove* this intermediate conclusion.) Then, we take the new intermediate conclusion (which we now know to be true) alongside our assumptions (which we were allowed to pretend for free are true) and then work to discover another fact (also called an intermediate conclusion) obtained through the rules of inference. This new intermediate conclusion is also called proved. We keep continuing in this way: any facts that we know to be true (either because we just proved them, proved them earlier in the text of the proof, or got to assume were true at the beginning of the proof) are *used* to *prove* additional intermediate conclusions. At some point, we end up proving a proposition that we needed to be true (again, according to the process of rules of inference which we will describe soon) and then our proof is complete.

One major aspect of confusion that we just want to bring up right away is that in the process of writing a proof, there are a bunch of individual steps along the way where one or more propositions previously in the proof are *used*, and they are used to *prove* another proposition. That is, a proof (which might be several paragraphs of text) has within it a bunch of little steps of proving: yes, a proof has "mini-proofs" inside it. Imagine the individual steps of going from proposition to proposition like the individual flights: using one proposition to prove another proposition should be thought of like the flight that departs from New York and arrives at Los Angeles. Then, the next part of the proof that uses the new proposition to get an even newer proposition can be thought of as the second flight which departs Los Angeles and arrives at Tokyo. Then using the newest proposition to get an even newer proposition can be thought of like that third (and in our analogy, final) flight from Tokyo to Singapore. To complete this analogy (where "using" a proposition is like departing from a certain airport, and "proving" a proposition is like arriving at a certain airport), the proof (that is, the overall multi-paragraph text) should be thought of as the entire flight itinerary: the proof is like whole journey (plane 1 from New York to LA, plane 2 from LA to Tokyo, and plane 3 from Tokyo to Singapore).

The big questions at this point might be: what are we allowed to assume? What has to be the final fact that we get? How do we do the steps in between? That's what the point of this chapter is: to carefully introduce this for each kind of logical operation and quantifier we introduced in the previous chapter.

Do not attempt to merely memorize proofs word-for-word. There are a lot of often-used steps (introduced below), so don't memorize those steps! If there's anything to memorize, it's the key *idea* behind large proofs.

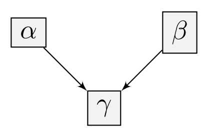
The last section of this chapter (Section 3.10) highlights some moments from previous math classes where the ideas in this Chapter may have been used in bite-sized chunks. You may consider reading the last section first as a way to appreciate how past math experiences (even if they were awful at the time) serve as a warm up for a formal study that we are about to undertake. Alternatively, you may wish to keep that section for the end, where references to methods throughout this chapter help paint a complete picture of your past exposure to proof.

3.1 Basic methods of proof

This section introduces the rules of inference. The presentation does not parallel the order from Chapter 2, and while that may seem peculiar, several ideas have to come together and our first complete proofs appear in Section 3.1.6. (There are a lot of things that need to get put together, but we intentionally do not cover the logical operations in the same order as the previous chapter: our goal is to have students see a complete proof as soon as possible, because it will be helpful to see the big picture.)

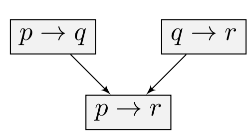
3.1.1 Introduction to rules of inference

A rule of inference will be presented visually like this:



Each rule of inference will tell us how knowing certain propositions to be true leads us to learn other propositions are true. The diagram above is generic: in place of α and β and γ will be actual propositions, instead of these Greek symbols. How we are to read this is that in any situation where we know that α is true and we also know that β is true, then we get to conclude that γ is true. We can't just draw any of these diagrams to our heart's content. Instead, for the diagram to legitimately be a rule of inference, we need to argue why we sohuld accept the rule of inference being presented. We do this by examining truth tables (and recall that truth tables themselves are actually just visual summaries of the definitions of each of the logical operations). Once we have provided an argument of why we should accept a rule of inference, it becomes a useful tool for us: it allows us to say that in a situation where we know that α is true and also know that β is true, we end up learning that γ has to be true (and so γ is called a conclusion, or an intermediate conclusion). Now, why might α be true? It might be that we were allowed to assume α is true for free, or it might be that we just earlier in our proof got α as an intermediate conclusion. Similarly, why might we have β being true? It might be that β was an allowable assumption, or it might be that β was an intermediate conclusion that we obtained earlier in the proof. Then, the proposition in the spot where γ is (note that both upper boxes point down to the box containing γ) ends up being our newest intermediate conclusion. We will say that we *use* α and *use* β, and by using both, we ended up *proving* γ.

The discussion above was generic. Let's now present a specific rule of inference:



How do we read this rule of inference? Wheneer we are in a situation where we know that $\boxed{p \to q}$ is true and we also know that $\boxed{q \to r}$ is true, then we can conclude that $\boxed{p \to r}$ is true.

Why should we accept this rule of inference? Consider the following truth table which presents the three propositions $\boxed{p \to q}$ and $\boxed{q \to r}$ and $\boxed{p \to r}$.

p	q	r	$p \to q$	$q \to r$	$p \to r$
T	T	T	T	T	T
T	T	F	T	F	F
T	F	T	F	T	T
T	F	F	F	T	F
F	T	T	T	T	T
F	T	F	T	F	T
F	F	T	T	T	T
F	F	F	T	T	T

Although there are typically eight situations, let us remove any situations where $p \to q$ is false. We end up removing rows 3 and 4, resulting in

p	q	r	$p \to q$	$q \to r$	$p \to r$
T	T	T	T	T	T
T	T	F	T	F	F
F	T	T	T	T	T
F	T	F	T	F	T
F	F	T	T	T	T
F	F	F	T	T	T

We also only want to consider situations where $q \to r$ is true, so from the six rows above, let us remove the two rows where $q \to r$ is false. We are left with

p	q	r	$p \to q$	$q \to r$	$p \to r$
T	T	T	T	T	T
F	T	T	T	T	T
F	F	T	T	T	T
F	F	F	T	T	T

From our original 8 rows, by removing any rows where $p \to q$ was false and also removing any rows where $q \to r$ is false, we are left with 4 rows. In all four situations which remain, $p \to r$ is true. While we accept this rule of inference, we do not usually apply this rule of inference. (The typical rules of inference in a mathematician's toolbox start in Section 3.1.2.)

Before jumping into that section, though, we present another (rare) rule of inference and argue why it should be trusted.

$$\begin{array}{ccc} \boxed{p \lor q} & & \boxed{r \lor \neg p} \\ & \searrow \quad \swarrow & \\ & \boxed{q \lor r} & \end{array}$$

How do we read this? Say that we happen to be in a situation where we know $\boxed{p \lor q}$ is true and we also know $\boxed{r \lor \neg p}$ is true. In that situation, we can conclude that $\boxed{q \lor r}$ is true. Why do we trust this rule of inference? Why do we accept it? The truth table below presents $\boxed{p \lor q}$ and $\boxed{r \lor \neg p}$ and $\boxed{q \lor r}$.

p	q	r	$p \lor q$	$r \lor \neg p$	$q \lor r$
T	T	T	T	T	T
T	T	F	T	F	T
T	F	T	T	T	T
T	F	F	T	F	F
F	T	T	T	T	T
F	T	F	T	T	T
F	F	T	F	T	T
F	F	F	F	T	F

Since we only want to consider situations where $p \lor q$ is true, let's remove rows where $p \lor q$ is false. We get:

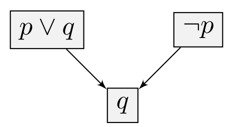
p	q	r	$p \vee q$	$r \vee \neg p$	$q \vee r$
T	T	T	T	T	T
T	T	F	T	F	T
T	F	T	T	T	T
T	F	F	T	F	F
F	T	T	T	T	T
F	T	F	T	T	T

From these remaining 6 rows, we only want to look at situations where $r \vee \neg p$ is true, so let's remove the rows where $r \vee \neg p$ is false:

p	q	r	$p \vee q$	$r \vee \neg p$	$q \vee r$
T	T	T	T	T	T
T	F	T	T	T	T
F	T	T	T	T	T
F	T	F	T	T	T

In all the rows that remain, the proposition $q \vee r$ is true. Thus, this rule of inference is valid. If $p \vee q$ is true and $r \vee \neg p$ is true, then it is automatically always the case that $q \vee r$ is true as well.

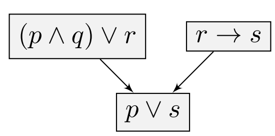
Exercise 185. *Consider the following rule of inference:*



Answer the following questions:

1. *What is this rule of inference stating?*
2. *Using a truth table (and removing rows from it), show that this rule of inference can be trusted as valid.*

Exercise 186. *Consider the following rule of inference:*



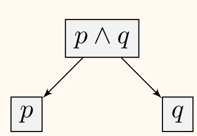
Answer the following questions:

1. *What is this rule of inference stating?*
2. *Using a truth table (and removing rows from it), show that this rule of inference can be trusted as valid.*

3.1.2 Rules of inference for conjunctions

We use the proposition $p \wedge q$ according to the following flowchart:

Method 187: Using a conjunction



Visually, we have written $p \wedge q$ at the top, and have an arrow down to p as well as an arrow down to q. In a situation where we know that $p \wedge q$ is true (whether we were told to assume it in the beginning, or through the course of writing a proof, we discovered this fact to be true), how can we use $p \wedge q$? We can conclude that p is true. We can also conclude that q is true.

Why should we accept this rule of inference? This rule of inference is based on the definition of conjunction. Take the truth table of conjunction, and note that there are three rows where $p \wedge q$ is false. If we remove these three rows, we are left with

p	q	$p \wedge q$
T	T	T

and notice that p is true, and also that q is true. Think of the flowchart above as a small puzzle piece used in a larger puzzle. (In this analogy, think of the entire proof as being the entire puzzle.)

Example 188. *Suppose we knew "c is a snark and d eats vegetables." Then we can conclude "c is a snark." We can also conclude "d eats vegetables." (To be clear what we mean by* conclude*, we mean that we have a proposition that we now know for certain is true for sure, and before this moment, we didn't know whether the propositien "c is a snark" was true or false.)*

Example 189. *Suppose we knew "a is even and b is even." Then we can conclude that "a is even." We can also conclude that "b is even."*

Example 190. *Suppose we knew "a is odd and b is odd." Then we can conclude that "a is odd." We can also conclude that "b is odd."*

The puzzle piece we saw has $p \wedge q$ at the top, and we saw what facts can be obtained from having $p \wedge q$ be true. Another "puzzle piece" is the flowchart for *proving* the proposition $p \wedge q$. Notice that in this flowchart, $p \wedge q$ is at the bottom:

Method 191: Proving a conjunction

$$\begin{tikzcd}
\boxed{p} \arrow[dr] & & \boxed{q} \arrow[dl] \\
& \boxed{p \wedge q} &
\end{tikzcd}$$

How should we read this? As soon as p is established to be true and q is also established to be true, we can then conclude that $p \wedge q$ is true. Why should we accept this rule of inference? Start with the truth table for conjunction and remove any rows where there is an F in the column for p, since we are in the situation where p is true. Similarly, because q is true, remove any rows where there is an F for the q column. In all remaining rows, the column for $p \wedge q$ always has a T.

Example 192. *Suppose we knew "s has a dog." Suppose we also knew "t has a cat." Then we can conclude "s has a dog and t has a cat."*

3.1.3 Rules of inference for implications

There are two main ways to use the implication $p \to q$. The first way of using $p \to q$ is called **modus ponens**:

Method 193: Using an implication (modus ponens)

$$\begin{tikzcd}
\boxed{p \to q} \arrow[dr] & & \boxed{p} \arrow[dl] \\
& \boxed{q} &
\end{tikzcd}$$

Modus ponens is saying that, in the process of writing a proof, when you know that $p \to q$ is true and you know that p is true, then you can conclude that q is true. Why should we accept modus ponens as one of our rules of inference? Start with the truth table for implication:

p	q	$p \to q$
T	T	T
T	F	F
F	T	T
F	F	T

Since we are in the situation where $p \to q$ is true, let us remove the one situation where $p \to q$ is false. So we have

p	q	$p \to q$
T	T	T
F	T	T
F	F	T

We are also in the situation where p is already true, so of the three remaining situations, let us remove the rows where p is false. We are left with

p	q	$p \to q$
T	T	T

and q must therefore be true.

Example 194. *Suppose we were allowed to assume "If x is a borogove, then x is a tove" and we were also instructed to assume that "x is a borogove" is true. Then, we can conclude "x is a tove" via modus ponens.*

Example 195. *Suppose we had just proved the implication "If s^2 is even, then s is even" and we were instructed to assume that "s^2 is even" is true. Then, we can conclude "s is even" via modus ponens.*

Another way to use the implication $p \to q$ is **modus tollens**:

Method 196: Using an implication (modus tollens)

$$\begin{tikzcd} \boxed{p \to q} \arrow[dr] & & \boxed{\neg q} \arrow[dl] \\ & \boxed{\neg p} & \end{tikzcd}$$

You should convince yourself that it is reasonable to accept modus tollens as a rule of inference: start with the truth table for $p \to q$. Remove any rows where $p \to q$ is false. Remove any rows where $\neg q$ is false. In all remaining rows, note that $\neg p$ is true.

Example 197. *Suppose we were allowed to assume "If x is a borogove, then x is a tove" and we were also instructed to assume that "x is not a tove" is true. Then, we can conclude "x is not a borogove" via modus tollens.*

Example 198. *Suppose we had just proved the implication "If s^2 is even, then s is even" and we were instructed to assume that "s^2 is not even" is true. Then, we can conclude "s is not even" via modus tollens.*

Warning 199

An implication can never be used on its own. To use $p \to q$, we need at least one other proposition, and the typical proposition we need is p.

The warning above just stated that in addition to $p \to q$, we would need p. It is tempting to fall into the trap of saying that $p \to q$ together with q will allow us to conclude something. However:

Warning 200

It is not possible to conclude anything from the propositions $p \to q$ and q, though someone new to proof is often tempted to conclude p. In other words, following the flowchart

$$\begin{tikzcd} \boxed{p \to q} \arrow[dr] & & \boxed{q} \arrow[dl] \\ & \boxed{p} & \end{tikzcd}$$

is not valid reasoning.

Example 201. *Suppose we were allowed to assume "If x is a borogove, then x is a tove" and we were also instructed to assume that "x is a tove" is true. That we cannot conclude anything, because neither modus ponens nor modus tollens will apply. In particular, we do not obtain "x is a borogove" as a conclusion.*

It is important to note that $p \to q$ cannot be used on its own, although $p \wedge q$ can be used on its own.

There are several ways to prove the implication $p \to q$. We introduce the **direct proof** of $p \to q$ here, and save the other methods for Sections 3.4 and 3.5:

Method 202: Proving an implication (direct proof)

$$\begin{tikzcd} \boxed{\text{Assume } p} \arrow[d] \\ \cdots \arrow[d] \\ \boxed{\text{Prove } q} \end{tikzcd}$$

We will use the direct proof so often, let us highlight this strategy:

Method 203: Direct proof of $p \to q$

To prove the implication $p \to q$, assume that p is true. Then use rules of inference to (eventually) obtain the fact that q is true.

To show it visually in the format of a cartoon, text that we work on might have this shape:

- Theorem: if p, then q.
- Proof of theorem: Assume p. (Additional text here based on rules of inference.) Final conclusion q.

Please note that in the proof of the theorem, there is no mention of the phrase "if p, then q" itself. Instead, the format of this text tells us where to look: what we assume is the text that is after "if" and before "then", while the text that we need as our final conclusion is the proposition appearing after the word "then." As a bit of an analogy, in algebra, we teach students that in $y = mx + b$, the value of m is the slope. So, when a student new to this is given $y = 3x + 4$ and says that $3x$ is the slope, we have to gently correct for this, and say that m represents the slope, and (in terms of position), m is the content after the equal sign but *before* the x, meaning the slope doesn't include x itself. (I have used language that's helpful like saying that x tells us where to look, but that we need the content *before* that x.) This is helpful here, so much so, that let's box things: when tasked with proving "if $\boxed{p}$, then $\boxed{q}$", we get to assume $\boxed{p}$ to be true for free (for free meaning we don't have to do any work: we just get to assume p is true), then work to conclude $\boxed{q}$. As in the analogy with identifying the slope in $y = 3x + 4$, be very literal about what's being assumed and what's being proved. Part of the confusion is that in the process of proving the overall proposition "if p then q" the proof text we write needs to end up proving q (though proving q ends up happening because we get to start the proof text by positioning ourselves in the hypothetical world where p must be true.)

Example 204. *Consider the implication "If Jo buys a car, then Jo lives in an apartment." To prove this proposition via the direct proof method, the author of the proof would start by writing "Assume that Jo buys a car." The goal is to then prove that "Jo lives in an apartment."*

The word "assume" and the word "suppose" used to start a sentence in a proof have the same role. Here is an example.

Example 205. *Consider the implication "If a is odd and b is odd, then $a + b$ is even." To prove this implication via the direct proof method, the first sentence of the proof should be "Suppose a is odd and b is odd." The goal in the remainder of the proof is to obtain the proposition "$a + b$ is even." To get to that point, there will be many other sentences which advance the logic forward, so the writer of the proof might indicate this eventual goal by writing "We want to show that $a + b$ is even."*

3.1.4 Our first complete proofs

We can now apply what we've learned to write our first complete proofs.

Example 206. *Use the following hypotheses*

- *H1: If Dave is not a physics major, then Steve is a chemistry major.*
- *H2: If Alice is a math major, then Carol is a math major.*
- *H3: If Steve is a chemistry major, then Bob is a computer science major.*
- *H4: If Dave is a physics major, then Carol is not a math major.*

to prove: if Alice is a math major, then Bob is a computer science major.

Because this process is new, we'll take it slowly. The four hypotheses are propositions that we are allowed to assume to be true for free. We will visually represent these as boxes at the top of our diagram:

$$\boxed{\text{H1}} \qquad \boxed{\text{H2}} \qquad \boxed{\text{H3}} \qquad \boxed{\text{H4}}$$

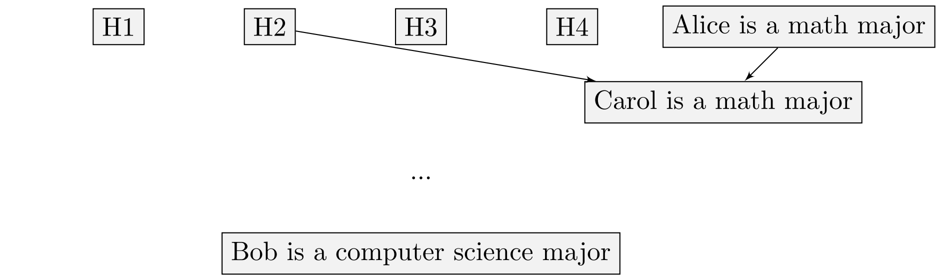
We wrote "H1" inside the box instead of the acutual full text "If Dave is not a physics major, then Steve is a chemistry major" to keep the diagram from being too cluttered. Now, we were asked to prove "If Alice is a math major, then Bob is a computer science major." Because we are being asked to prove an implication, we will assume the hypothesis of the implication (the part after "if") and work to prove the conclusion of the implication (the part after "then"). So, we will have a box that says "Alice is a math major" at the top of our diagram, and place our goal of "Bob is a computer science major" in a box at the bottom of our diagram:

$$\boxed{\text{H1}} \qquad \boxed{\text{H2}} \qquad \boxed{\text{H3}} \qquad \boxed{\text{H4}} \qquad \boxed{\text{Alice is a math major}}$$

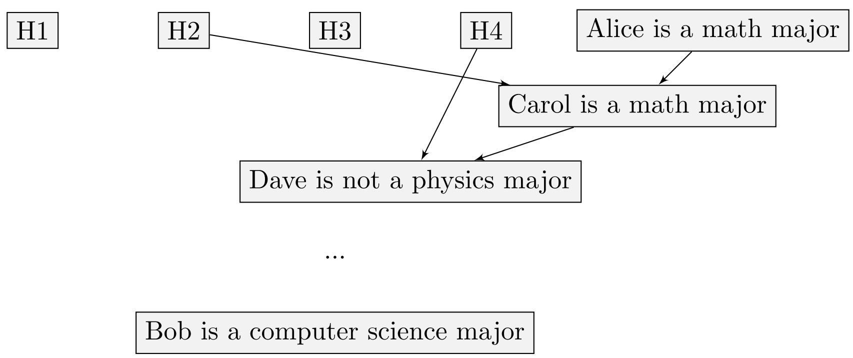
$$\ldots$$

$$\boxed{\text{Bob is a computer science major}}$$

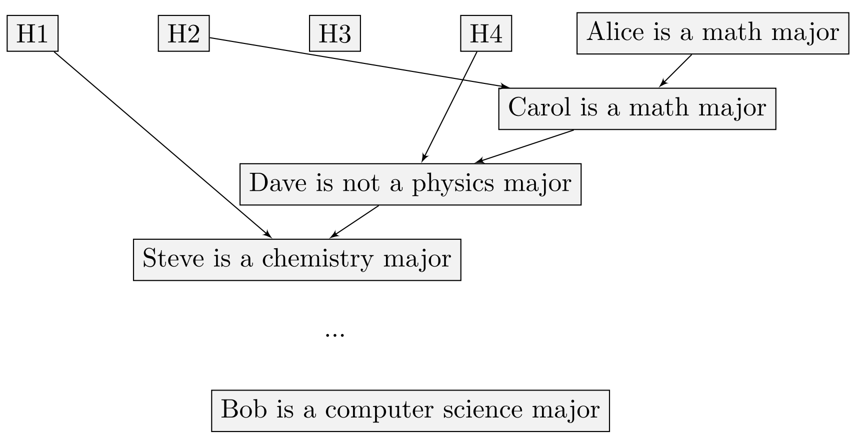
Now, we need to take the five boxes at the top and see how to connect them to the goal at the bottom, by applying rules of inference. (This is what needs to be filled in, represented by the three dots in the diagram above.) Starting with the box that says "Alice is a math major," we can use H2 (If Alice is a math major, then Carol is a math major") via modus ponens to conclude "Carol is a math major." We add this to our diagram:



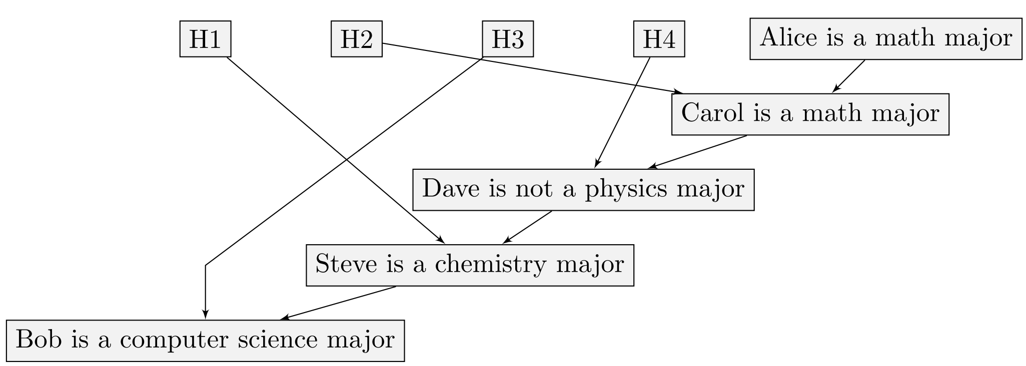
The fact that "Carol is a math major" is a new fact for us. Now that we know this to be true, we can use H4 (If Dave is a physics major, then Carol is not a math major") via modus tollens to conclude "Dave is not a physics major." We add this to our diagram:



New we can use H1 (If Dave is not a physics major, then Steve is a chemistry major") via modus ponens to conclude "Steve is a chemistry major." We add this to our diagram:



Finally, we can use H3 (If Steve is a chemistry major, then Bob is a computer science major") via modus ponens to conclude "Bob is a computer science major," which is our goal. We add this to our diagram:



The diagram is connected, so now we can write the proof in complete sentences. (In what we write below, we will place the propositions that appeared in the flowchart in quotation marks, just to be able to see clearly the propositions being discussed, but these should actually be removed.)

Proof. Assume the four numbered hypotheses. Assume that "Alice is a math major." We want to prove that "Bob is a computer science major." Because "If Alice is a math major, then Carol is a math major" and "Alice is a math major" we conclude "Coral is a math major." Because "If Dave is a physics major, then Carol is not a math major" and "Carol is a math major," we conclude "Dave is not a physics major." Because "If Dave is not a physics major, then Steve is a chemistry major" and "Dave is not a physics major," we conclude "Steve is a chemistry major." Because "If Steve is a chemistry major, then Bob is a computer science major" and "Steve is a chemistry major," we conclude "Bob is a computer science major." ☐

Notice that the third sentence of the proof said that *We want to prove that "Bob is a computer science major."* At this moment in the proof, we haven't actually proved that Bob is a computer science major yet, but this is what *will* be proved by the the end of the proof. In each of the places in the proof where we wrote a full implication, we could instead have written H1, H2, H3, or H4. Let's provide a proof that does that (and then removes any/all quotation marks).

Proof. Assume the four numbered hypotheses. Assume that Alice is a math major. We want to prove that Bob is a computer science major. Because H2 and Alice is a math major, we conclude Carol is a math major. Because H4 and Carol is a math major, we conclude Dave is not a physics major. Because H1 and Dave is not a physics major, we conclude Steve is a chemistry major. Because H3 and Steve is a chemistry major, we conclude Bob is a computer science major. ☐

Noitec that we have a sentence like "Because H4 and Carol is a math major, we conclude Dave is not a physics major." which states both of the propositions which were used (namely H4 as well as "Carol is a math major") and while there are times that writing a proof with this amount of thoroughness makes sense, because we had just obtained the fact that "Carol is a math major" one sentence before, we could (for stylistic reasons) write the proof slightly shorter. We will apply this shortening in the same way to all sentences where this makes sense to do:

Proof. Assume the four numbered hypotheses. Assume that Alice is a math major. We want to prove that Bob is a computer science major. By H2, we conclude Carol is a math major. By H4, we conclude Dave is not a physics major. By H1, we conclude Steve is a chemistry major. By H3, we conclude Bob is a computer science major. ☐

We switched from the word "Because" to the word "By" in the last proof, which is a stylistic choice. Both words are acceptable. We could actually insert the information when modus ponens applied versus when modus tollens applied, but this is not required. Just to see what it looks like to do this, here's a slightly longer proof that includes this information:

Proof. Assume the four numbered hypotheses. Assume that Alice is a math major. We want to prove that Bob is a computer science major. By H2, we conclude Carol is a math major applying modus ponens. By H4, we conclude Dave is not a physics major applying modus tollens. By H1, we conclude Steve is a chemistry major applying modus ponens. By H3, we conclude Bob is a computer science major applying modus ponens. ☐

This proof is a slightly longer text. If we wanted to shorten the shortest that we had by silghtly further, we could even remove "we conclude" from each sentence:

Proof. Assume the four numbered hypotheses. Assume that Alice is a math major. We want to prove that Bob is a computer science major. By H2, Carol is a math major. By H4, Dave is not a physics major. By H1, Steve is a chemistry major. By H3, Bob is a computer science major. ☐

This is a very concise proof. Each sentence is short and to the point. However, it is still clear what is happening in each step.

Remark 207

There are many ways to write a proof. The proofs above illustrate some of the choices that can be made. As you write proofs, you will develop your own style. The important thing is that the logic is clear and easy to follow.

Let us comment on one way that this proof shouldn't be written: when we know $p \to q$ and we also know p, then applying modus ponens means that we need to conclude q, instead of concluding $p \to q$. Notice the subtle difference: the implication $p \to q$ is not the new fact that we get. The new fact that we get is q. So after the first three sentences "Assume the four numbered hypotheses. Assume that Alice is a math major. We want to prove that Bob is a computer science major." sometimes when people are new to this process of writing a proof, the following sentence gets included: "If Alice is a math major, then Carol is a math major." Writing this sentence never actually states that Carol is a math major. By writing a sentence in the form "If p then q" all that's doing is providing a complete sentence version of $p \to q$. The conclusion that we need to reach is q, and that conclusion itself (that is, the text after the word "then" in the implication) needs to be stated as a standalone conclusion.

Exercise 208. *Using the hypotheses*

- *If it does not rain or if it is not foggy, then the sailing race will be held and the lifesaving demonstration will go on.*
- *If the sailing race is held, then the trophy will be awarded.*
- *The trophy was not awarded.*

prove (in full sentences) the statement "It rained." *[key]*

3.1.5 Rules of inference for existentially-quantified statements

To use the existentially-quantified statement $\exists x \in U\,[P(x)]$, we introduce a new rule of inference:

Method 209: Using an existentially-quantified statement

$$\begin{tikzcd}[column sep=small] & \boxed{\exists x \in U\,[P(x)]} \arrow[d] & \\ & \boxed{\text{Let } b \in U \text{ satisfying } P(b).} \arrow[dl] \arrow[dr] & \\ \boxed{b \in U} & & \boxed{P(b)} \end{tikzcd}$$

If we are to start with $\exists x \in U\,[P(x)]$ being true, since there exists something in U, and if that something were called x, then $P(x)$ is true, we should make such an object. In the flowchart above, we intentionally use a new variable b. It may be the case that you have already used b in your proof. Then, use another letter. The word "Let" should always be followed by a new variable that has not yet been used. For example, the sentence $\boxed{\text{Let } b \in U \text{ satisfying } P(b).}$ written in a proof creates b, and b happens to be an element of U (Generically, b is an object. Now, if U were the set of all car brands, then b would be a car brand. If U were instead the set of all countries, then b would be a country.) and, in addition, $P(b)$ is true. We can make such an object (called b) because it was previously established that there exists $x \in U$ such that $P(x)$.

Example 210. *Say we know* $\boxed{\exists m \in Y\,[P(m)]}$ *is true. Then, as long as b has not been written in our proof yet, we could write* $\boxed{\text{Let } b \in Y \text{ satisfying } P(b)}$*, then from this conclude that* $\boxed{b \in Y}$ *and* $\boxed{P(b)}$ *are true.*

Why does this work? Suppose that Y was the set of all cats, and $P(m)$ was the property "m has green eyes." Then, the statement $\exists m \in Y[P(m)]$ means that there exists a cat with green eyes, which we can also state as: there is a cat with green eyes. Because we are starting from the fact that there exists a cat with green eyes, writing "Let $b \in Y$ satisfying b has green eyes" is giving us access to a cat (that we may not know the name of, so we just call it b) and this cat has green eyes. Though we might not know the specific identity of the cat with green eyes, we know that a cat with green eyes must exist, and now that we have labeled this cat b, we can continue in our proof using this information: any time after the sentence "Let $b \in Y$ satisfying $P(b)$" is written, we can use the fact that b is a cat (because $b \in Y$) and can also use the fact that b has green eyes (because $P(b)$ is true).

Example 211. *Say we know $\boxed{\text{there exists } h \in L \text{ such that } P(h)}$ is true. Then, as long as b has not been written in our proof yet, we could write $\boxed{\text{Let } b \in L \text{ satisfying } P(b)}$, then from this conclude that $\boxed{b \in L}$ and $\boxed{P(b)}$ are true.*

Why does this work? Suppose L is the set of all lakes and $P(h)$ is the property "h has a depth greater than 100 feet." Then, the statement "there exists $h \in L$ such that $P(h)$" means that there exists a lake with a depth greater than 100 feet. Because we are starting from the fact that there exists a lake with a depth greater than 100 feet, writing "Let $b \in L$ satisfying $P(b)$" is giving us access to talking about a specific lake (that we may not know the name of, so we just call it b) that has a depth greater than 100 feet.

Example 212. *Say we know $\boxed{\text{there exists } m \in S \text{ such that } P(m)}$ is true. Then, as long as y has not been written in our proof yet, we could write $\boxed{\text{Let } y \in S \text{ satisfying } P(y)}$, then from this conclude that $\boxed{y \in S}$ and $\boxed{P(y)}$ are true.*

Why does this work? Suppose S is the set of all smartphones and $P(m)$ is the property "m has a battery life longer than 24 hours." Then, the statement "there exists $m \in S$ such that $P(m)$" means that there exists a smartphone with a battery life longer than 24 hours. Because we are starting from the fact that there exists a smartphone with a battery life longer than 24 hours, writing "Let $y \in S$ satisfying $P(y)$" is giving us access to a specific smartphone (that we may not know the identity of, so we just call it y) that has a battery life longer than 24 hours.

Example 213. *Say we knew that $\boxed{\text{there exists } c \in D \text{ such that } c \text{ plays soccer}}$ is true. Then, as long as b has not been written in our proof yet, we could use this by writing $\boxed{\text{Let } b \in D \text{ satisfying } b \text{ plays soccer}}$, then conclude from this that $\boxed{b \in D}$ and $\boxed{b \text{ plays soccer}}$ are true.*

Why does this work? Suppose D is the set of all dogs. Then the statement "there exists $c \in D$ such that c plays soccer" means that there exists a dog that plays soccer. Because we are starting from the fact that there exists a dog that plays soccer, writing "Let $b \in D$ satisfying b plays soccer" is giving us access to a specific dog (that we may not know the name of, so we just call it b). Then, we know that b is a dog (because $b \in D$) and we also know that b plays soccer, in spite of potentially not knowing the specific dog that b refers to.

Example 214. *Say we knew that $\boxed{\text{there exists } c \in D \text{ such that } c \text{ plays soccer}}$ is true. Then, if b has been used in our proof, but k has not, we could use this by writing $\boxed{\text{Let } k \in D \text{ satisfying } k \text{ plays soccer}}$, then conclude from this that $\boxed{k \in D}$ and $\boxed{k \text{ plays soccer}}$ are true.*

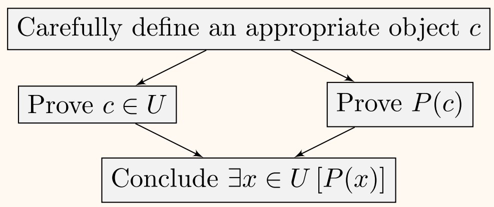
Example 215. *Suppose we know the proposition "There exists an $x \in \mathbb{H}$ such that x has dual citizenship" is true. To use this proposition, we would include in our proof the sentence "Let b be in $\mathbb{H}$ satisfying the property that b has dual citizenship." Then, we could draw two additional conclusions. We can conclude $b \in \mathbb{H}$. We can also conclude b has dual citizenship.*

Why does this work? Because we are starting from the fact that there exists an $x \in \mathbb{H}$ such that x has dual citizenship, which in words is "Someone has dual citizenship" writing "Let b be in $\mathbb{H}$ satisfying the property that b has dual citizenship" is giving us access to a specific b (that we may not know the identity of, so we just call it b): that is, b is a specific person who has dual citizenship.

Example 216. *Suppose we were to assume "There exists $k \in \mathbb{Z}$ such that $a = 2k + 1$." Then, we could write "Let $x \in \mathbb{Z}$ satisfying $a = 2x + 1$." From here, we could conclude $x \in \mathbb{Z}$. We could also conclude $a = 2x + 1$.*

To prove the existentially-quantified statement $\exists x \in U\,[P(x)]$, follow the flowchart:

Method 217: Proving an existentially-quantified statement



It is extremely intentional that c is used in three locations. First, c should be defined. Then, using how c was defined, prove that $c \in U$. Third, using the c that was defined, prove $P(c)$. The emphasis is that it must be the same c. In other words, both $c \in U$ and $P(c)$ must simultaneously be true. This is the only way we could rightly conclude that there exists an $x \in U$ such that $P(x)$ is true. Now, if c had already appeared in the proof, then a different letter would be written instead of c. That is, replace all of the cs in the flowchart above with a letter that has not yet been used in the proof.

Remark 218: Proving something exists

This is essentially a restating of the flowchart above. To prove that something exists and satisfies a certain property, make an object (say we call it c) and prove that all the necessary properties are true for that object: in this case, prove that $c \in U$ and prove that $P(c)$ is true. In short, to convince someone that a certain thing exists, make the thing! To prove $\exists m \in U$ such that $P(m)$ is true, you should find an object c in U be carefully defining c, then show $P(c)$.

In order to prove something exists (and has a certain behavior), you should find it (and show that it has that behavior).

Example 219. *Say we need to prove $\boxed{\exists m \in Y[P(m)]}$ is true. Then, as long as b has not been written in our proof yet, we could define b. Then, with this b, prove that $\boxed{b \in Y}$ and $\boxed{P(b)}$ are true.*

Why does is work? Suppose that Y is the set of all cats, and $P(m)$ is the property "m has green eyes." Then, the statement $\exists m \in Y[P(m)]$ means that there exists a cat with green eyes. To prove to someone that there exists a cat with green eyes, we should find such a cat. By defining b to be such a cat, we can then prove that b is a cat (because $b \in Y$) and also prove that b has green eyes. Once we have defined b, and proved b is a cat, and proved b has green eyes, then the reader should be convinced that there exists a cat with green eyes.

Example 220. *Say we need to prove $\boxed{\text{there exists } h \in L \text{ such that } P(h)}$ is true. Then, as long as b has not been written in our proof yet, define b. Then, based on what you have said b is, prove that $\boxed{b \in L}$ and $\boxed{P(b)}$ are true.*

Why does this work? Suppose L is the set of all lakes and $P(h)$ is the property "h has a depth greater than 100 feet." Then, the statement "there exists $h \in L$ such that $P(h)$" means that there exists a lake with a depth greater than 100 feet. To prove to someone that there exists a lake with a depth greater than 100 feet, we should find such a lake. By defining b to be such a lake, we can then prove that b is a lake (because $b \in L$) and also prove that b has a depth greater than 100 feet. Once we have defined b, and proved b is a lake, and proved b has a depth greater than 100 feet, then the reader is convinced that there exists a lake with a depth greater than 100 feet.

Example 221. *Say we need to prove $\boxed{\text{there exists } m \in S \text{ such that } P(m)}$ is true. Then, as long as y has* *not been written in our proof yet, define y. Then, based on how you defined y, prove* $\boxed{y \in S}$ *and also prove* $\boxed{P(y)}$.

Why does this work? Suppose S is the set of all smartphones and $P(m)$ is the property "m has a battery life longer than 24 hours." Then, the statement "there exists $m \in S$ such that $P(m)$" means that there exists a smartphone with a battery life longer than 24 hours. To prove to someone that there exists a smartphone with a battery life longer than 24 hours, we should find such a smartphone. By defining y to be such a smartphone, we can then prove that y is a smartphone (because $y \in S$) and also prove that y has a battery life longer than 24 hours. Once we have defined y, and proved y is a smartphone, and proved y has a battery life longer than 24 hours, then the reader is convinced that there exists a smartphone with a battery life longer than 24 hours.

Example 222. *Say we need to prove* $\boxed{\textit{there exists } c \in D \textit{ such that } c \textit{ plays soccer}}$ *is true. Then, as long as b has not been written in our proof yet, define something to be b. If b has been defined appropriately, then it will be possible to prove* $\boxed{b \in D}$ *and* $\boxed{b \textit{ plays soccer}}$ *are true.*

Why does this work? Suppose D is the set of all dogs. Then the statement "there exists $c \in D$ such that c plays soccer" means that there exists a dog that plays soccer. To prove to someone that there exists a dog that plays soccer, we should find such a dog. By defining b to be such a dog, we can then prove that b is a dog (because $b \in D$) and also prove that b plays soccer. Once we have defined b, and proved b is a dog, and proved b plays soccer, then the reader is convinced that there exists a dog that plays soccer.

Example 223. *Say we need to prove* $\boxed{\textit{there exists } c \in D \textit{ such that } c \textit{ plays soccer}}$ *is true. Then, if b has been used in our proof, but k has not, carefully define k, and based on how you defined k, prove* $\boxed{k \in D}$ *and* $\boxed{k \textit{ plays soccer}}$ *are true.*

Example 224. *Suppose we have the task of needing to prove $\exists k \in \mathbb{Z}\,[a+b=2k]$. If so, we should define z. Then, based on how we have defined z, we would need to prove $z \in \mathbb{Z}$. We also would need to prove $a+b=2z$. If we are successful proving both of these things, we would then be able to conclude $\exists k \in \mathbb{Z}\,[a+b=2k]$.*

As another example, let us consider the proof of this short theorem.

Theorem 225. *There exists a real number b such that $b^2 - 10 = 3b + 30$.*

Proof. Let $b = 8$. Then $b \in \mathbb{R}$ because every integer is a real number. Because $b^2 - 10 = (8)^2 - 10 = 64 - 10 = 54$ and $3b + 30 = 3(8) + 30 = 54$, we have $b^2 - 10 = 3b + 30$. □

Of course, this is not the only possible proof of this theorem. Here is another proof:

Proof. Let $a = -5$. Then $a \in \mathbb{R}$ because every integer is a real number. Because $a^2 - 10 = (-5)^2 - 10 = 25 - 10 = 15$ and $3a + 30 = 3(-5) + 30 = 15$, we have $a^2 - 10 = 3a + 30$. □

3.1.6 Rules of inference for definitions

Suppose we had a predicate $P(z)$, and using this predicate, suppose that a definition had previously been established that said $\boxed{\text{Definition: We say } z \text{ is a } \textbf{borogove} \text{ if } P(z).}$

We introduce flowcharts for proving g is a borogove and – just to use a different letter – using the fact that m is a borogove.

Method 226: Proving g is a borogove

$$\boxed{P(g)}$$
$$\downarrow$$
$$\boxed{g \text{ is a borogove.}}$$

Method 227: Using m is a borogove

$$\boxed{m \text{ is a borogove.}} \to \boxed{P(m)}$$

Language Discussion 228: Every definition behaves like an "if and only if" statement

By staring at the rules of inference for using and proving a definition, you may notice that we are using the text $\boxed{z \text{ is a } \textbf{borogove} \text{ if } P(z)}$ as if it says $\boxed{z \text{ is a } \textbf{borogove} \text{ if and only if } P(z)}$. While there are some authors that insist on writing the phrase "if and only if" in each definition for this very reason, the majority of authors only write the word "if." Nevertheless, when reading definitions from these authors, one use of the word "if" in the definition truly operates as an "if and only if."

This handbook on proof follows the convention that the majority of mathematical authors use: one instances of the word 'if' in each definition operates as an 'if and only if,' but the word 'if' is used nonetheless.

While you are certainly familiar with integers that are even and integers that are odd, to manipulate odd and even numbers in proofs, we will need definitions:

Definition 229: Even

The integer n is **even** if there exists an integer k such that $n = 2k$.

Definition 230: Odd

The integer n is **odd** if there exists an integer k such that $n = 2k + 1$.

By referencing the definition of even, here are flowcharts (based on the template flowcharts in Methods 226 and 227) for using a is even and for proving c is even:

$$\boxed{a \text{ is even}} \to \boxed{\exists k \in \mathbb{Z}\,[a = 2k]} \qquad \boxed{\exists y \in \mathbb{Z}\,[c = 2y]} \to \boxed{c \text{ is even}}$$

Notice the intentional writing of different letters than the usual ones appearing in the definition of even.

Exercise 231. *Write for yourself flowcharts for using r is odd and for proving s is odd.*

Example 232. *Here's a flowchart for proving $r + s$ is even:*

$$\boxed{\exists m \in \mathbb{Z}\,[r + s = 2m]} \to \boxed{r + s \text{ is even}}$$

This flowchart is based on Method 226 and the definition of even. How is that so? In the flowchart above, replace each c with $r + s$, and replace each y with m.

Exercise 233. *In Exercise 97, you were asked to consider what are things that must be addressed in writing a definition of disjunction. Similarly, what are things that you must address when writing a definition for odd?*

Exercise 234. *In Exercise 97, you were asked to consider what are things that must be addressed in writing a definition of disjunction. Similarly, what are things that you must address when writing a definition for even?*

In this class, your instructor will likely ask you to recite definitions such as the definitions of odd and even. Like we did for the definition of conjunction, let's give some sample texts that are *almost* acceptable definitions for even and for odd, then explain what's missing in each (and provide a fixed version).

- **Incomplete:** Let m be an integer. Then m is odd if there exists k such that $m = 2k + 1$.
  - This text does not specify that k is an integer. A definition must specify the universe of discourse for each variable that appears in the definition. As written, k could be a real number, a rational number, a complex number, or something else like an elephant. Here is a fixed version: Let m be an integer. Then m is odd if there exists an integer k such that $m = 2k + 1$.
- **Incomplete:** Let n be an integer. There exists an integer k such that $n = 2k + 1$.
  - This text does not introduce the word "odd" itself. A definition must introduce a new word or phrase. Why is this important? By reading the text given, the reader does not know if the word "odd" is going to be the word describing n or the word describing k. So we should clarify *which* thing is odd. Here is a fixed version: Let n be an integer. We call n odd if there exists an integer k such that $n = 2k + 1$.
- **Incomplete:** Let n be an integer. Even is when there exists an integer a such that $n = 2a$.
  - This text introduces the word "even" but by reading the text given, the reader does not know if the word "even" is going to be the word describing n or the word describing a. So we should clarify *which* thing is even. Here is a fixed version: Let n be an integer. Then n is even if there exists an integer a such that $n = 2a$.
- **Incomplete:** We will call z even if there exists b such that $z = 2b$.
  - This text introduces the word "even" and clarifies that z is the thing being described as even. However, the text does not specify that z is an integer, nor does it specify that b is an integer. If we take the text written at its word, then $z = 2\pi$ and $b = \pi$ would satisfy the definition, and we should conclude $2\pi \approx 6.28$ is even, but we don't want that to happen. We should specify that z should be an integer (so that the reader can't attempt to apply the definition of even to anything that isn't an integer in the first place such as 6.28), and we should also always specify a set for a quantified variable such as b. Here is a fixed version: We will call the integer z even if there exists $b \in \mathbb{Z}$ such that $2b = z$.

Finally, here's another incorrect definition: "An integer n is even if n can be divided by 2 with no remainder." This isn't really text that we can say is "incomplete" and just add text to it to fix it. This text is problematic because the phrase "can be divided by 2 with no remainder" is not precise enough to be useful in proofs. The text after the first "if" in a definition must be written in the language of logical operations and quantifiers, because these are the types of text that we have tools for using and proving. We have no tools for using or proving the phrase "can be divided by 2 with no remainder." It turns out that all mathematical definitions can be written using this language:

- ... and ...
- ... or ...
- not ...
- if ... then ...
- ... if and only if ...
- there exists ... such that ...
- for all ... , ...

Because the types of text above are the ones that we have developed tools for using and proving, this is the type of text we need to have appearing in definitions.

Here are definitions of even that are complete:

- An integer a is even if there exists an integer k such that $a = 2k$.
- The integer b is even if there exists an integer m such that $b = 2m$.
- We say that the integer n is even if there exists $t \in \mathbb{Z}$ such that $2t = n$

The first two examples are similar: writing "an" or "the" are both acceptable here.

Example 235. *Consider the following definition: A real number m is* **duporous** *if there exists an integer c such that $m = 2^c$.*

If we were given the fact that $\boxed{r \text{ is duporous}}$ *is true, then we can conclude that there exists an integer d such that $r = 2^d$. This follows Method 227.*

On the other hand, suppose we needed to prove that y is duporous. (For the sake of simplicity, suppose we already knew that y is real.) Then, we should first prove the statement "there exists an integer k such that $y = 2^k$." After proving this, we could conclude that y is duporous. This follows Method 226.

Notice in this example that the text of the definition of dupourous only has the word "if" instead of the phrase "if and only if." However, we operate pretending the sentence were "A real number m is **duporous** if and only if there exists an integer c such that $m = 2^c$." See Language Discussion 228.

Warning 236

This is an important discussion about the definition of even and the definition of odd. Please do not ignore these two definitions.Notice that both definitions include the phrase "there exists." We will need to deal with the phrase "there exists" in the manner introduced in Section 3.1.5. It will be impossible to do proofs if you think of "n is even" as being "n can be divided by 2 with no remainder" or similar notions such as "2 goes evenly into n." These phrases are unhelpful for proof (because we have no rules of inference that deals with the phrase "goes evenly into."

We are now ready to prove our first theorem:

Theorem 237. *If a is odd and b is odd, then $a + b$ is even.*

Your intuition might tell you this statement is true. After all, $3 + 5 = 8$ and $11 + 9 = 20$. However, thinking of a million examples where this seems to be true does not make a proof! Examples are helpful as intuition, but they are not substitutes for the proof. (It is nice that we have intuition, but the big idea of this chapter is that there is a "one size fits all" process of proving statements based on the logical operations and quantifiers that appear in the relevant definitions.)

The first several times you work on a proof, it will feel like a lot of work. It will be tempting to think, "Why should I follow such dry methods of proof when I can use my intuition on this statement?" In future math classes, you will be expected to prove statements for which you will have no intuition at all. However, the methods of proof will apply. Even if the theorem above seems silly, this is a chance to practice the methods of proof.

Notice that Theorem 237 is an implication. To prove an implication, we follow Method 203, so we will assume that $\boxed{a \text{ is odd and } b \text{ is odd}}$ and work to prove $\boxed{a + b \text{ is even}}$. (See Example 205.) If we were to organize this in a flowchart, the flowchart of logic so far would look like this:

$$\begin{array}{c}\boxed{a \text{ is odd and } b \text{ is odd}}\\ \downarrow \\ \cdots \\ \downarrow \\ \boxed{a+b \text{ is even}}\end{array}$$

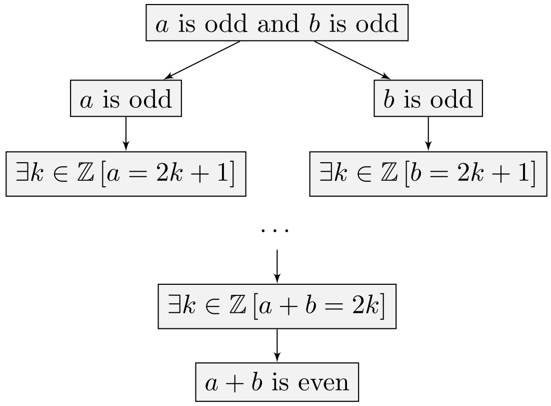
We have more that we need to fill in for the flowchart above, but we already get to see what box is at the top and what box is at the bottom. Moreover, the placement of the boxes already tells us how some of the sentences in the eventual proof will look, as well as when they will appear. For example, since "a is odd and b is odd" is at the top of the flowchart, our first sentence will be "Assume a is odd and b is odd" or will be "Suppose a is odd and b is odd" and both of these are functionally equivalent. Similarly, since "$a+b$ is even" is at the bottom of the flowchart, our last sentence will be "Therefore, $a+b$ is even" or some version of this such as "We finally conclude that $a+b$ is even." Because we need the final fact that's true to be "$a+b$ is even," after our starting sentence (the sentence that begins with "Assume" or "Suppose"), we can optionally include as our second sentence something like "We will show that $a+b$ is even" or "Our goal is to prove that $a+b$ is even" or similar, since a sentence like this helps us see where we need to go. We need to use the rules of inference to connect these two propositions. (This is what the dots represent – details which we need to fill in. Note that this is how nearly all proofs go in their flowchart form: we have one or more boxes and the top, we have a box at the bottom, and we have to use the box or boxes at the top and connect to the box at the bottom based on the rules of inference.) The statement $\boxed{a \text{ is odd and } b \text{ is odd}}$ is a conjunction, and since we are assuming this to be true, let's use this statement, following Method 187. (In fact, we have already done this work in Example 190.) Now our flowchart is

$$
\begin{tikzcd}
& \boxed{a \text{ is odd and } b \text{ is odd}} \arrow[dl] \arrow[dr] & \\
\boxed{a \text{ is odd}} & & \boxed{b \text{ is odd}} \\
& \cdots \arrow[d] & \\
& \boxed{a+b \text{ is even}} &
\end{tikzcd}
$$

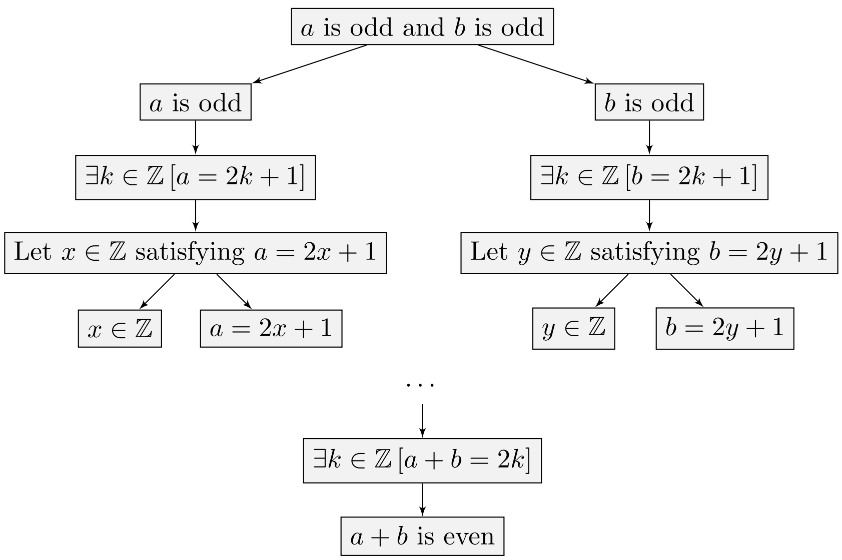
We should use the fact that a is odd, a flowchart which we already wrote out and can borrow. We similarly use the fact that b is odd, also modeled after Method 227. With these additions, our flowchart is now:

$$
\begin{tikzcd}
& \boxed{a \text{ is odd and } b \text{ is odd}} \arrow[dl] \arrow[dr] & \\
\boxed{a \text{ is odd}} \arrow[d] & & \boxed{b \text{ is odd}} \arrow[d] \\
\boxed{\exists k \in \mathbb{Z}\,[a = 2k+1]} & & \boxed{\exists k \in \mathbb{Z}\,[b = 2k+1]} \\
& \cdots \arrow[d] & \\
& \boxed{a+b \text{ is even}} &
\end{tikzcd}
$$

It is helpful to work on the lower portion of the flowchart. What is it going to take for $a+b$ to be even? (In other words, what will need to be the box immediately before?) Here, Method 226 applies. To give more detail, examine Example 232, which gives a flowchart for proving $r+s$ is even: by replacing each $r+s$ with $a+b$ and replacing each m with k, we see that to prove $a+b$ is even, we need to prove $\boxed{\exists k \in \mathbb{Z}\,[a+b = 2k]}$. Our updated flowchart is:



We have two existence statements which we have not yet used. We update the flowchart, informed by Method 209, keeping in mind that each "Let" sentence should use a *new* variable. (See our work in Example 216.)



We have already done a lot of work, and this may seem overwhelming at first, but this process will get easier. Up to this point, everything (other than choosing to use the letters x and y) done was very methodical and did not require creativity. (What we mean by "did not require creativity" is that there weren't any choices to be made. It might not feel that way yet, but this is just an attempt to point out that there wasn't really anything that the proof writer had to decide. If we were to consider examples in previous math classes, to solve $3x^2 = 4x - 5$ using the quadratic formula, there is very little creativity involved, as you'd just need to decide whether you want all the terms on the left or all the terms on the right before applying the Quadratic Formula. By contrast, given a typical integral, you have a lot of creativity: you'd have to choose between substitution, integration by parts, and so on.) We now enter the creative phase of this proof. We still need to prove $\boxed{\exists k \in \mathbb{Z}\,[a + b = 2k]}$ and to prove this statement, Method 217 says we need to define an appropriate object. Let's say that the object we define will be called z, though we could have used any letter other than x, y, a, or b. (See Example 224.) Not *only* do we have to define z, but z needs to be defined in such a way that two things are true. The first of two things we will need is $z \in \mathbb{Z}$. The second of two things we will need is $a + b = 2z$. Note that we are allowed to use the facts that $a = 2x + 1$ and $b = 2y + 1$. In fact, we should define z in *in terms of* these already-defined variables x, y, a, and b. Because we need z to be an integer, it is probably not good for us to define z to be something like $z = \frac{a+y}{3}$ or $z = \sqrt[7]{a}$, because defining z in either of those two ways probably does not make an integer. Instead, we will probably need to define z as something like $z = a - b + y$. Now, it won't actually be this, but $a - b + y$ is certainly an integer because a, b, and y are all integers. How do we actually determine how we should define z? There are several explanations that feel very different, but are actually the same. We will provide several explanations: try reading them all to see which one makes the most sense to you, but then go back to see how the other explanations are really saying the same thing:

- If we consider $a+b$, we can rewrite this as $a+b=(2x+1)+(2y+1)$ and using further properties of algebra, this can be rewritten as $2x+2y+1+1=2x+2y+2=2(x+y+1)$. In short, using the facts $a=2x+1$ and $b=2y+1$, we see that $a+b=2(x+y+1)$, yet recall we need to define an integer z such that $a+b=2z$. It seems we should define z to be $x+y+1$. In fact, note that because x and y are integers, z is an integer (since the sum of integers is an integer, the closure of addition stated in Section 2.7).

- The equation that we need to be true is $a+b=2z$, which will end up being true by picking z appropriately. In the equation $a+b=2z$, let's substitute $2x+1$ for a and $2y+1$ for b. This gives us

  $$(2x+1)+(2y+1)=2z.$$

  We can rewrite the left side

  $$2x+2y+2=2z.$$

  We can factor a 2 out of the left side to get

  $$2(x+y+1)=2z.$$

  Dividing both sides by 2 gives

  $$x+y+1=z.$$

  This suggests defining z to be $x+y+1$. Because x and y are integers, defining z to be $x+y+1$ will mean that z is an integer.

- We can take the view of "several equations". We have equations $a=2x+1$ and $b=2y+1$ that we know to be true, and the equation $a+b=2z$ that we want to be true. Since we need to define z, we can think of it like we need to solve for z. From $a+b=2z$, by dividing by 2 on both sides, we have

  $$z=\frac{a+b}{2}.$$

  The problem is that we don't actually want to define z as this fraction, because it's hard to argue that a fraction is an integer. So, in the equation we just wrote, let's substitute $2x+1$ for a and $2y+1$ for b. This gives us

  $$z=\frac{(2x+1)+(2y+1)}{2}.$$

  We can simplify the right side:

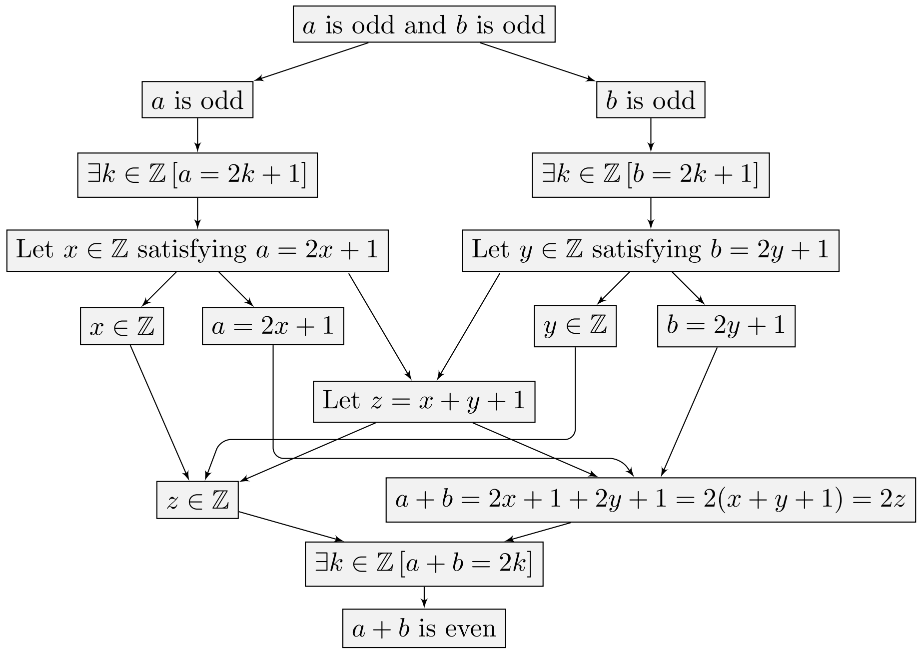
  $$z=\frac{2x+2y+2}{2}.$$

  $$z=\frac{2(x+y+1)}{2}.$$

  $$z=x+y+1.$$

  This suggests defining z to be $x+y+1$. Because x and y are integers, defining z to be $x+y+1$ will mean that z is an integer. (That is, when we define $z=x+y+1$, there is no fraction written: we have said that z should be the sum of the integer x, the integer y, and the integer 1.)

Defining $z=x+y+1$ will complete the flowchart:



The box "Let $x \in \mathbb{Z}$ satisfying $a = 2x+1$" defined x and the box "Let $y \in \mathbb{Z}$ satisfying $b = 2y+1$" defined y. Those definitions of x and y are all that are needed to be able to say "Let $z = x + y + 1$" which ends up defining z. (This is why these are the only two arrows pointing toward the box for "Let $z = x + y + 1$." While the *inspiration* for *how* to define z came from $a+b = 2(x+y+1)$, the only things needed to define z were x and y.) The box defining z as the upper part of the flowchart in Method 217. The fact that z is an integer relied on the definition of z (which was the text "Let $z = x + y + 1$") as well as the fact that x and y are integers (which explains the three arrows pointing inwards), since adding integers will always produce an integer. Within the flowchart, we included the mechanics of proving that $a + b = 2z$. The first step of that was a substitution of a and b. The second step was algebra. The final step used the definition of z. The flowchart is now complete.

Warning 184 tells us not to confuse using a proposition with proving a proposition. Use the direction of the arrows (and the "level" at which we have written the boxes) of the flowchart to visually see which proposition(s) were used to prove which proposition(s). For example, the proposition "a is odd" is used to prove the proposition "$\exists k \in \mathbb{Z}[a = 2k + 1]$," while the proposition "$\exists k \in \mathbb{Z}[a + b = 2k]$" is used to prove the proposition "$a + b$ is even."

Think of the flowchart as being a big puzzle, while all of the methods introduced in this chapter as being little puzzle pieces. The design of a proof is to put together the puzzle pieces in ways that they fit. Though we have now visually written out the logical flow of our proof, a flowchart itself is not the proof. Proofs are written in complete sentences. The flowchart helps us organize what order the sentences should appear in our proof: we cannot discuss what's in a certain box until we've first discussed the boxes pointing in. Here is a proof of Theorem 237.

Proof. Suppose a is odd and b is odd. We want to show that $a+b$ is even. From the hypothesis, we conclude a is odd. We also conclude b is odd. Since a is odd, there exists $k \in \mathbb{Z}$ such that $a = 2k + 1$. So, let $x \in \mathbb{Z}$ satisfy $a = 2x + 1$. Then $x \in \mathbb{Z}$ and $a = 2x + 1$. Similarly, since b is odd, there exists $k \in \mathbb{Z}$ such that $b = 2k + 1$. Let $y \in \mathbb{Z}$ satisfy $b = 2y + 1$. Then $y \in \mathbb{Z}$ and $b = 2y + 1$.

Let $z = x + y + 1$. So $z \in \mathbb{Z}$, since $x \in \mathbb{Z}$ and $y \in \mathbb{Z}$. By substitution and algebra,

$$\begin{aligned} a + b &= (2x + 1) + (2y + 1) \\ &= 2x + 2y + 1 + 1 \\ &= 2x + 2y + 2 \\ &= 2(x + y + 1) \\ &= 2z. \end{aligned}$$

Since we have defined $z \in \mathbb{Z}$ and proved that $a+b = 2z$, it is true that there exists $k \in \mathbb{Z}$ such that $a+b = 2k$. Therefore, $a + b$ is even. □

If we wanted to include more detail, when saying that $z \in \mathbb{Z}$, we could have said: So $z \in \mathbb{Z}$, since the sum of integers is an integer, and since x, y, and 1 are all integers. While mentioning that 1 is an integer is true, and similarly saying that the sum of integers is an integer is true, we hinted at these things with our shorter sentence.

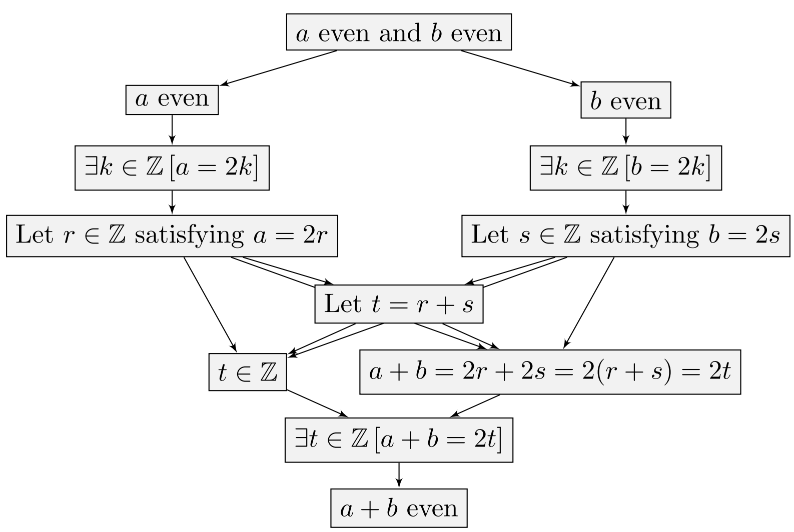
Our first example took a while to build, and it is important for you to work on building some of the top of the flowchart, then some of the bottom of the flowchart, then going back to the top, and so on. Very often, it is helpful to bounce back and forth like this, so that you can see where you are heading. In fact, in our proof, we wrote the sentence "We want to show that $a + b$ is even." as a reminder to ourselves and to our reader, but the proof would be complete even without this sentence. (When you are new to proofs, it is helpful to include these sentences because it serves as a good self-reminder of your goals: that is, what sentence do you need to logically arrive at so that we can consider the proof complete.)

It is important to emphasize Warning 184. Proving something and using something are very different: in the proof above, we used a is odd, used b is odd, and proved $a + b$ is even. After applying the definition that a is odd, we *used* the statement $\exists k \in \mathbb{Z}\,[a = 2k + 1]$. This is unlike later in the proof where we *proved* the statement $\exists k \in \mathbb{Z}\,[a + b = 2k]$ by defining an appropriate integer called z. In every proof, some statements get used and some statements get proved. Some statements which are proved along the way (the intermediate conclusions) are then *used* to prove something later on in the proof. (Think of the airline travel analogy we gave. A first flight might go from New York City to Los Anegeles, meaning that Los Angeles is an arrival location. But the second flight might be from Los Angeles to Tokyo, meaning that Los Angeles is now thought of as a departure location. The same location can be an arrival location at one time and a departure location at another time. Similarly, the same statement can be something that is proved at one time and something that is used at another time. Once we have proved a proposition to be true, we know it to be true, and so we can then use that proposition later on. This is analogous to the idea that we have to first arrive at an airport to then depart from it.)

We will introduce a more compact form of using an existentially-quantified statement through an example. We will present three complete flowcharts and their corresponding proofs, each a little shorter than the previous. Our case study is the following theorem.

Theorem 238. *If a is even and b is even, then $a + b$ is even.*

In the style of our previous example, a complete flowchart would be

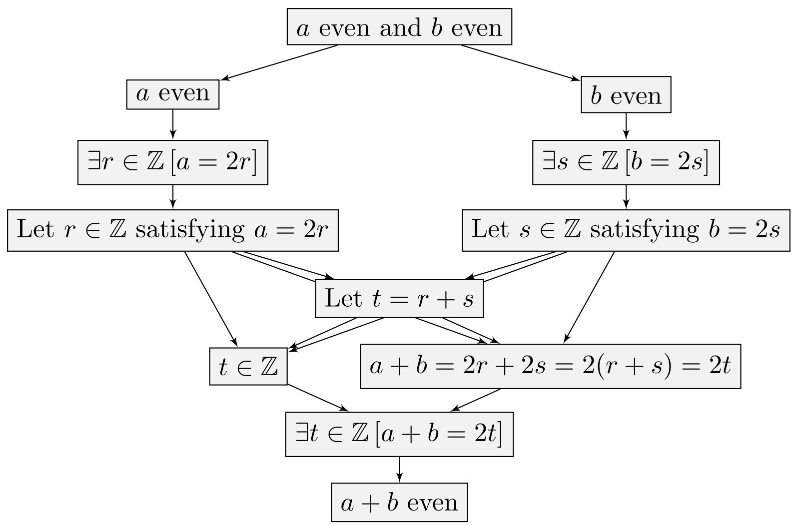


The corresponding proof follows.

Proof. Let a be even and b be even. We want to prove $a + b$ is even. Since a is even, there exists an integer k such that $a = 2k$. So, let r be an integer satisfying $a = 2r$. Since b is even, there exists an integer k such that $b = 2k$. So let s be an integer satisfying $b = 2s$.

Let $t = r + s$. Since r and s are integers, t is an integer. By substitution, $a + b = 2r + 2s = 2(r + s) = 2t$. So there exists an integer t such that $a + b = 2t$. Therefore, $a + b$ is even. $\square$

Recall from Section 2.4 that the quantified variable is a placeholder variable, so writing $\exists k \in \mathbb{Z}[c = 2k]$ is the same as writing $\exists u \in \mathbb{Z}[c = 2u]$. Let's insist on writing a new variable every time we have an existential quantifier. Then the flowchart is:

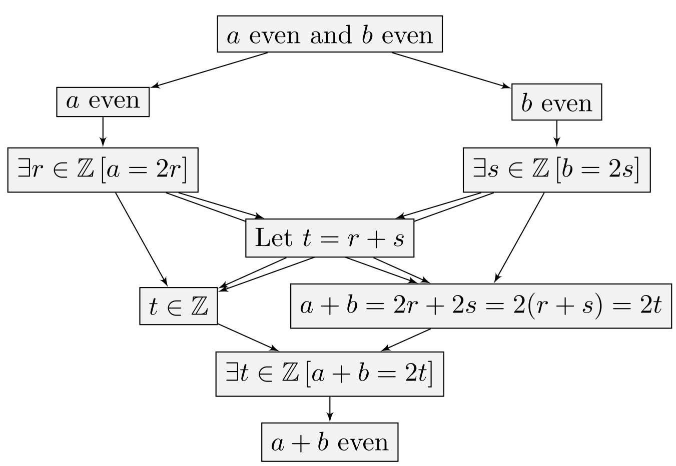


The corresponding proof follows:

Proof. Let a be even and b be even. We want to prove $a + b$ is even. Since a is even, there exists an integer r such that $a = 2r$. So, let r be an integer satisfying $a = 2r$. Since b is even, there exists an integer s such that $b = 2s$. So let s be an integer satisfying $b = 2s$.

Let $t = r + s$. Since r and s are integers, t is an integer. By substitution, $a + b = 2r + 2s = 2(r + s) = 2t$. So there exists an integer t such that $a + b = 2t$. Therefore, $a + b$ is even. $\square$

Then, it seems a bit redundant to have "there exists..." as well as the "let ... satisfying ..." ... so let's just keep one, but not the other. (We'll keep the "there exists" sentence, but now we must write with a new variable each time.)



Proof. Let a be even and b be even. We want to prove $a + b$ is even. Since a is even, there exists an integer r such that $a = 2r$. Since b is even, there exists an integer s such that $b = 2s$.

Let $t = r + s$. Since r and s are integers, t is an integer. By substitution, $a + b = 2r + 2s = 2(r + s) = 2t$. Because there exists an integer t such that $a + b = 2t$, we conclude $a + b$ is even. $\square$

Before moving on, I encourage you to re-read the introductory paragraphs that begin this chapter, especially the paragrph that made an analogy between a multi-segmint flight itinerary and the process of proof. We need to be clear about the difference between proving and using, so we will point this out in the proof that we just wrote. We started with the assumptions that a is even and b is even. Note that a being even and b being even are propositions that are never proved. These are just assumptions, meaning these are accepted as being true for free. For the rest of the proof text:

- We use the fact that a is even to prove that there exists an integer r such that $a = 2r$.
- We use the fact that b is even to prove that there exists an integer s such that $b = 2s$.
- We defined t. Nothing is proved in doing this.
- We used the fact that r is integer and used that s is an integer and used the definition of $t = r + s$ to prove that t is an integer.
- We used $a = 2r$ and $b = 2s$ and $t = r + s$ to prove $a + b = 2t$, and this proving this was several algebra steps in a row.
- We used the fact that there exists an integer t such that $a + b = 2t$ to prove that $a + b$ iseven.

So one big proof is actually a bunch of mini-proofs strung together in a meaningful way. The list above is discussing each of the mini-proofs (though the exception that was pointed out is that defining t didn't prove anything). In each of the remaining steps of the proof, one or more propsitions that is alerady true (whether by assumption, or previously proved) is used to prove the next proposition. The newly-proved propsition is new information, and the things that were *used* to get there are the *reasons* why the new proposition is true. In terms of the flowchart diagrams, when one or more boxed propositions point down to a proposition, the boxes above (where the arrows come *from*) are the propositions thate are *used*, and the box below (where the arrows point *to*) is the proposition that is *proved*. Keep in mind, though, that many boxes in the middle (not at the top, not at the bottom) serve both the role of being *proved* first, only to be *used* later. The timeline matters. Back to our flight analogy of an itnerary from New York City to Los Angeles to Tokyo to Singapore, Los Angeles is an arrival (analogy to a proposition that is proved), but then later Los Angeles is our departure point (analogy to a proposition being used). We keep working in this way until we get to the *final* conclusion.

For additional practice, prove these two theorems:

Theorem 239. *If m is even and n is odd, then $m + n$ is odd.*

Theorem 240. *If j is even and k is even, then $j - k$ is even.*

The following definition is a generalization of even:

Definition 241: Divides

We say the integer a **divides** the integer b if there exists an integer k such that $b = ak$. We write $a \mid b$ to denote a divides b.

This notion generalizes evenness since b is even if and only if 2 divides b.

Language Discussion 242

This is a good time to remind all readers to follow the main habits for a definition. Following Habit 2, is "divides" a noun, verb, or adjective? The word "divides" is a verb. Moreover, it is a transitive verb (meaning this is a verb which performs an "action" onto an object).

Warning 243: Divides versus divided by

We have defined the transitive verb "divides" above, but do not confuse this with the phrase "divided by." While there is a strong similarity in spelling, treat these as *completely* different. The statement $\boxed{\text{3 divides 12}}$ is a proposition, and a true one at that. In contrast, the phrase $\boxed{\text{3 divided by 12}}$ is not a proposition because this is neither true nor false. In fact, 3 divided by 12 is just the number $\frac{1}{4}$, once reduced.
In English, both "contepmt" and "contemplate" start similarly, agreeing in their first seven letters, but their meanings are totally unrelated. In the same way, try to keep "divides" and "divided by" separate.

Exercise 244. *In Exercise 97, you were asked to consider what are things that must be addressed in writing a definition of disjunction. Similarly, what are things that you must address when writing a definition for divides?*

Theorem 245. *If m divides c and m divides d, then m divides $c - d$.*

Proof. Suppose m divides c and m divides d. We want to prove m divides $c - d$. Since m divides both c and d, there exist integers a and b such that $c = ma$ and $d = mb$. Let $u = a - b$. Then $u \in \mathbb{Z}$ and $c - d = ma - mb = m(a - b) = mu$, so m divides $c - d$. □

Warning 246: Do not divide when working just with integers

It is tempting to want to write things like $r = \frac{t}{u}$ in the course of a proof involving only integers (as is the case when dealing with the definitions of even, odd, and divides). By applying division, however, you end up making it less clear which quantities are integers. If you are tempted to write $r = \frac{t}{u}$, then write $ru = t$ instead: every equation involving division/fractions can always be rewritten in terms of multiplication instead.
When we proved Theorem 237, we showed three separate ways of thinking that led to $z = x + y + 1$. The third way ended up introducing fractions. This was good in our scratch work, but the final proof never introduced fractions. If you find a situation where you feel fractions are unavoidable in your final work, keep in mind that you can always take an equation that has fractions and multiply both sides by the denominator. For example, if you wrote $\frac{a}{b} + c = \frac{d}{e}$ in your scratch work, you can multiply both sides by be, so you could write $ae + bce = bd$ in the proof you write.

Remark 247: Existence does not always use the set of integers

Because our first full examples of proof deal with the existence of quantified variable in $\mathbb{Z}$, it sometimes becomes a habit to always assume that existence deals with integers. However, if S is a set (even if S is not a set of integers, and even if S is not even a set of a numbers), it is possible to prove statements that begin "there exists $m \in S$ such that ..." and it would be incorrect (in general) to say "there exists an integer m such that ...". This kind of language could be used if S was known to be a set of integers (or the set of all integers).

Exercise 248. *Prove: If x is even and y is even, then xy is even.* *[key]*

Exercise 249. *Prove: If k is even and n is odd, then kn is even.* *[key]*

Exercise 250. *Prove: If k is odd and p is odd, then kp is odd.* *[key]*

Exercise 251. *Prove: if a divides b and b divides c, then a divides c.*

3.1.7 Rules of inference for universally-quantified statements

Let $P(x)$ be the predicate "x has slept." Recall that we use $\mathbb{H}$ denote the set of all [living] humans. Let us assume that $\forall x \in \mathbb{H}\,[P(x)]$ is true. (Actually, it probably doesn't need to be *assumed*: it is true that every human has slept.) Let y denote your math instructor. Then $y \in \mathbb{H}$. (In other words, your math instructor is a human.) Given the fact $\forall x \in \mathbb{H}\,[P(x)]$ and given the fact $y \in \mathbb{H}$, what can we conclude? We can conclude that your math instructor has slept. In other words, we conclude $P(y)$.

Every time you have a universally-quantified statement which is true, as soon as you know you have an element y in the universe of discourse which was mentioned, then you get a fact about the individual y. In a flowchart:

Method 252: Using a universally-quantified statement

To use $\forall x \in U\,[P(x)]$, follow:

$$\begin{tikzcd} \boxed{\forall x \in U\,[P(x)]} \arrow[dr] & & \boxed{y \in U \text{ was already established}} \arrow[dl] \\ & \boxed{P(y)} & \end{tikzcd}$$

Warning 253

To use a universally-quantified statement, we emphasize that you must already have an element y which was already known to be in U. This is why we stated the phrase "was already established" in the flowchart above. You cannot just say "Let $y \in U$."

Recall that $p \to q$ cannot be used on its own. We saw that one way to use $p \to q$ is to also know p is true. There is an analogy to this here: we cannot use $\forall x \in U[P(x)]$ on its own. In order to use $\forall x \in U[P(x)]$, we need to have something already established in the set U, and that thing was labeled y in Method 252.

The flowchart presented in Method 252 allows us to start with a general fact ("$P(x)$ holds for all $x \in U$") and combined with knowing of an object y in the universe of discourse (example: y is known to belong to U) allows us to conclude that $P(y)$ holds. This takes us from a general fact to a specific fact about a specific individual named y. It is general to say, "Everyone has slept" and it is specific to say, "Your math teacher has slept." There will be times in our proof we will need to be that specific, and this flowchart helps us do this.

Example 254. *Suppose we know* $\boxed{\forall v \in D[P(v)]}$ *is true. Suppose we also know* $\boxed{s \in D}$ *is true. Then we can conclude* $\boxed{P(s)}$ *is true. Notice that our conclusion does not have the "for all" symbol.*

Why does this work? Say that D is the set of all dogs, and $P(v)$ is the predicate "v is cute." Then $\forall v \in D[P(v)]$ in plain English is: every dog is cute. Let's say that we have "Every dog is cute" is true. Saying that we also know $s \in D$ is true is means to say that we know "s is a dog" to be true. The logical way to combine "Every dog is cute" and "s is a dog" is to conclude "s is cute" which is exactly what was stated earlier to say $P(s)$.

Example 255. *Suppose we know* $\boxed{\forall h \in L[P(h)]}$ *is true. Suppose we also know* $\boxed{z \in E}$ *is true. Then we cannot conclude anything based on these two facts, because we need something in L to use $\forall h \in L[P(h)]$.*

What happened? Say that L is the set of all turtles, and $P(h)$ is the predicate "h eats chicken." Then $\forall h \in L[P(h)]$ in plain English is "Every turtle eats chicken" or "All turtles eat chicken." Knowing $z \in E$ doesn't help, because E is some other set that might not be related to L.

Example 256. *Suppose we know* $\boxed{\text{for all } b \in C\text{, the manager gives } b \text{ a raise}}$ *is true. Suppose we also know* $\boxed{m \in C}$ *is true. Then we can conclude* $\boxed{\text{the manager gives } m \text{ a raise}}$ *is true. Notice that the words "for all" are not in the conclusion.*

Why does this work? Let's say that C was the set of all Dallas-based employees. Then "for all $b \in C$, the manager gives b a raise" translates to: every Dallas-based employee is given a raise by the manager. Knowing $m \in C$ means knowing that m is a Dallas-based employee. The conclusion we should reach is that the manager gives m a raise.

Example 257. *Suppose we know* $\boxed{\textit{for all } b \in C\textit{, the manager gives } b \textit{ a raise}}$ *is true. Suppose we also know* $\boxed{b \in S}$ *is true. Then we cannot conclude anything based on these two facts, because we need something in* C*.*

Example 258. *Suppose we know* $\boxed{\textit{for all } x \in S\textit{, a delivery for } x \textit{ is available}}$ *is true. Suppose we also know* $\boxed{y \in S}$ *is true. Then we can conclude* $\boxed{\textit{a delivery for } y \textit{ is available}}$ *is true. Notice that the words "for all" are not in the conclusion.*

Why does this work? Suppose that S *is the set of all geckos. Then "for all* $x \in S$*, a delivery for* x *is available" translates into "Every gecko has a delivery available" and* $y \in S$ *translates into "*y *is a gecko" so we can conclude: a delivery is available for* y*.*

Example 259. *Suppose we know* $\boxed{\textit{for all } j \in J\textit{, if } \pi > 2\textit{, then Alex buys a bike from } j}$*. Even though we know* $\pi > 2$ *is true, we can't really use* $\boxed{\textit{for all } j \in J\textit{, if } \pi > 2\textit{, then Alex buys a bike from } j}$ *yet, because it's stuck "behind" a 'for all' statement. (We write this word "behind" informally: as an analogy, in the expression* $a(b+c)^2$*, we cannot distribute* a *because of the exponent* 2*. For another analogy, while we would call* $j+k$ *a sum, we wouldn't call* $(j+k)(m+n)$ *a sum, because this is a product, but then "behind" each of the factors of the product is a sum.) Say that we also knew* $c \in J$*. Then we could first conclude the implication* $\boxed{\textit{if } \pi > 2\textit{, then Alex buys a bike from } c}$ *and by applying modus ponens, we further conclude* $\boxed{\textit{Alex buys a bike from } c}$ *is true.*

Example 260. *Suppose we know* $\boxed{\textit{for all } m \in A\textit{, there exists } r \in B \textit{ such that } m \textit{ gives } r \textit{ money}}$ *is true. Suppose we also know* $\boxed{z \in A}$ *is true. Then we can conclude* $\boxed{\textit{there exists } r \in B \textit{ such that } z \textit{ gives } r \textit{ money}}$ *is true. From this conclusion, Method 209 applies and further conclusions can be made. (The big idea here is to take things one step at a time. Because we were able to conclude a statement that begins "there exists", we know now this existentially-quantified statement is true, and then follow the rule of inference for using an existentially-quantified statement as we has previously learned.)*

Example 261. *If we know* $\forall x \in G, xy = yx$ *and we also know* $z \in G$*, then we can conclude* $zy = yz$*. In more detail, the predicate* $P(x)$ *is* $xy = yx$*. Since* z *is already known to be in* G*, we can conclude* $P(z)$*, which is* $zy = yz$*.*

Example 262. *If we know* $\forall x \in G, xy = yx$ *and we know* $y \in G$*, we can conclude* $yy = yy$*. Note we replaced all the* x*'s with* y*'s.*

Example 263. *What can we conclude if we know* $\forall x \in G, xy = yx$ *when we also have* $x \in G$*? Here, the* x*'s are different in meaning. Since the* x*'s in* $\forall x \in G, xy = yx$ *are placeholder variables, it may help to replace the* x*'s with* u*'s, so that we can say we know* $\forall u \in G, uy = yu$*. Then with* $\forall u \in G, uy = yu$*, combined with the fact* $x \in G$*, we can conclude* $xy = yx$*.*

Example 264. *What can we conclude if we know* $\forall x \in G, xy = yx$ *and we know* $x \in K$*? Nothing. The universally-quantified statement says that something is true for each member of the set* G*. However, the* x *we know about is in* K*. Do not let the fact that both of these statements use* x *cause confusion. In fact, the first statement can be restated as* $\forall u \in G, uy = yu$ *by replacing all the placeholder variables.*

Example 265. *For a more stark example, if we know* $\forall a \in B, a^4 = 1$ *and we know that* $c \in D$*, then we can conclude nothing. Note that* $\forall a \in B, a^4 = 1$ *is the same statement as* $\forall c \in B, c^4 = 1$*. For the same reason that the* a*'s in* $\forall a \in B, a^4 = 1$ *are unrelated to the* $c \in D$*, the* c*'s in* $\forall c \in B, c^4 = 1$ *are unrelated to* $c \in D$*.*

In writing proofs, some people like to use the phrase "in particular" when using a universally-quantified statement.

Example 266. *Let's look back to the motivating example for using a universally-quantified statement. That is, let's start from knowing to be true: for all* $x \in \mathbb{H}$*,* x *has slept. From here, we could note that Desmond Tutu is a human. (That is, Desmond Tutu is an element of* $\mathbb{H}$*.)*

Then in our proof, we could write "In particular, Desmond Tutu has slept" in place of just writing "Desmond Tutu has slept." What role do the words "in particular" have here? Think of this as trying to *point out that we are looking at a particular human (or a specific human). The initial fact we start with was universal – it was about* all *humans, stating that* all *humans have slept. The concluding sentence was to state a fact about* one particular *human: we ended up pointing out that* one specific *human (named Desmond Tutu) has slept.*

Example 267. *To use the statement $\forall r \in \mathbb{R}\, \exists s \in \mathbb{Z}\, s > r$ we will need to have a real number. For example, let's have in mind the real number π. In the form of a flowchart, here's what we can conclude:*

$$\begin{tikzcd}
\boxed{\forall r \in \mathbb{R}\, \exists s \in \mathbb{Z}\, s > r} \arrow[dr] & & \boxed{\pi \in \mathbb{R}} \arrow[dl] \\
& \boxed{\exists s \in \mathbb{Z}\, s > \pi} &
\end{tikzcd}$$

So note that $\forall r \in \mathbb{R}\, \exists s \in \mathbb{Z}\, s > r$ is a general statement. No matter what real number r is, the statement $\exists s \in \mathbb{Z}\, s > r$ is true. But that statement is so general. For a specific statement, we should think of a particular real number, and we get the statement $\exists s \in \mathbb{Z}\, s > \pi$. We have already used the word particular (which has practically synonymous with "specific" as used here.)

How would the wording in a proof look? Suppose we had the sentence "For all reals r, there exists an integer s such that $s > r$." Then, the next sentence of our proof might be written, "In particular, there exists an integer s such that $s > \pi$."

We have discussed how to use a universally-quantified statement. How would we prove a universally-quantified statement?

Let V be the set of all mountains in the United States with an elevation above 15000 feet. Let $Q(x)$ be the predicate "x has a tree on it." How would we prove $\forall x \in V\, [Q(x)]$? In other words, how would we convince someone that every mountain with an elevation above 15000 feet in the United States has a tree on it? Fortunately, there are not so many mountains with an elevation above 15000 feet, so we could start with the Denali (the tallest) and keep going through each possible choice for x one by one until we have seen all seven possible values for x. Each time, we can name the mountain and verify for the reader of our proof that x has a tree on it.

Let U be the set of all squirrels in Russia. Let $P(x)$ the predicate "x has brown eyes." How would we prove a statement like $\forall x \in U\, [P(x)]$? Maybe "prove" is too strong of a word to use here, but how would we convince someone that $\forall x \in U\, [P(x)]$ is true? Could we approach this in the same way as the tall mountains in the United States? Probably not. There are probably too many squirrels in Russia. Instead, imagine that we ask the reader to select any Russian squirrel they like. (In order for us to refer to the squirrel they picked, we'll call that squirrel r, and we'll state this in words by writing, "Let $r \in U$ be arbitrary.") Note that r is a single squirrel. Then, suppose we write a proof which convinced the reader that $P(r)$ is true, in other words, that r has brown eyes.

If we succeed in doing this, the reader should be convinced that $\forall x \in U\, [P(x)]$ is true. Now either the reader of our proof or an independent observer might say, "But wait! The reader only picked one squirrel!" The point is that the reader can go back and "rewind the video" in the sense of Remark 169 and select a new squirrel: the reader goes back to the point in the proof which said "Let $r \in U$ be arbitrary" and selects a different squirrel to be r. Then, if the argument still applies and the reader sees that $P(r)$ is true, the reader can rewind again and select a new choice for r. But if the reader does the thought experiment to see that no matter which $r \in U$ is selected, $P(r)$ is true, then the reader is now convinced that for *all* x in U, we have $P(x)$. This discussion leads to the following general method for proving a universally-quantified statement:

Method 268: Proving a universally-quantified statement

To prove $\forall x \in U\,[P(x)]$, follow:

$$\boxed{\text{Let } r \in U \text{ be arbitrary.}} \rightarrow \cdots \rightarrow \boxed{\text{Prove } P(r)}$$

Method 269: Proving a universally-quantified statement: a summary

Method 203 said to prove the implication $p \to q$, you first assume p is true, then use rules of inference to prove q is true. Something is assumed (namely p).

In the same way, when proving $\forall x \in U\,[P(x)]$ we start by assuming we have a "random" element from the set U. This is accomplished by writing "Let $r \in U$ be arbitrary" as long as r is a variable that has not yet been used in our proof. Then, using rules of inference, we must prove $P(r)$.

Language Discussion 270: What does the word "arbitrary mean?

The purpose of the word "arbitrary" is really meant to be a reminder that the person *reading* the proof (not the person writing the proof) may really choose whichever element from U they like to be r. Instead of writing "Let $r \in U$ be arbitrary," this may be shortened to just writing "Let $r \in U$." There is no change in meaning. The point is that the sentence "Let $r \in U$" – should it end there with a period – already allows the reader to pick whichever element from U that they want to be called r, even without the subsequent words "be arbitrary." However, the word "arbitrary" is meant to serve as *emphasis* that the reader may pick.

Because the reader gets to pick *whichever* element they want from U, and for the practical reason that a proof will need to be written, the reader's choice is called r, and then the proof writer then proves that $P(r)$ holds, this ends up proving $\forall x \in U\,[P(x)]$. The idea that the reader may make *any* choice of $r \in U$ is why the for *all* statement is true. Let's emphasize that the reader can pick any element from U that they want, and we don't actually end up knowing *which* element from U gets chosen. However, we need to talk about the choice they made, so we call their selection r.

As an analogy, there are many magic tricks where a magician asks a member of the audience to pick any card from a deck of cards. While some tricks start with sleight of hand, there are some tricks where the magician genuinely does not know which card the audience member picked. In that case, you'll hear the magician say things like "Now, put your card on the top of this half of the deck." In a sentence like that, the phrase "your card" is the way the magician refers to the card that they don't know (but that the audience member knows). This is similar to what happens in our proof: we ask the reader to pick from U (think of U as "the deck"), and the reader picks an element (think of this as "the card") which we refer to as r.

Warning 271: Proving versus using a universally-quantified statement

To prove $\forall x \in U\,[P(x)]$, the first sentence to write would be "Let $r \in U$ be arbitrary" as long as r is so far unused in the present proof. Instead the first sentence could be "Let $s \in U$ be arbitrary" but then the proof is done once $P(s)$ is proved.

This is unlike *using* the proposition $\forall x \in U\,[P(x)]$. To *use* this proposition, you must already have an element in U that you know of (which we have called y in our earlier flowchart and discussion). Using $\forall x \in U\,[P(x)]$ should *never* have a sentence of the form "Let $y \in U$ be arbitrary."

Back in Remark 136, we previed the idea of proving "For all $p \in U$, the last digit in p's phone number is a 2." At that time, we didn't have the complete set up to talk about proving anything (after all, that's what we're doing right now!). In fact, at that time, the discussion was really to just try to make as clear as possible what "for all" really means. But since "the last digit in p's phone number is a 2" needed to be true for *every* person p in U, it makes sense that the proof strategy would start with the sentence "Let $p \in U$" which allows the reader to pick any person from U that they want.

Recall that in Method 202, the most typical way to prove the implication $p \to q$, is to start by assuming p is true, then prove that q is true. We assume something (namely p), which gives us a fact to work with, and then we prove something (namely q). Proving a universally-quantified statement is similar in that we start by assuming we have an arbitrary element from the set U (which we have named here r), which gives us something to work with, and then we prove something, namely $P(r)$.

Example 272. *Say we need to prove $\forall v \in D[P(v)]$. As long as s has not been written before in the proof, write* $\boxed{\text{Let } s \in D}$*. Then work to prove* $\boxed{P(s)}$.

Why does this work? Imagine that D is the set of all dolphins, and $P(v)$ is the predicate "v can jump over a boat." Then $\forall v \in D[P(v)]$ in plain English is: every dolphin can jump over a boat. To prove this statement, we start by writing "Let $s \in D$" which allows the reader to pick any dolphin they like to be called s. Then we work to prove that $P(s)$ is true, which in plain English is: the dolphin s can jump over a boat. If we succeed in proving this, then the reader can rewind the tape and pick a different dolphin to be called s, and if our argument still applies, then the reader can be convinced that every dolphin can jump over a boat. The reader got to pick which dolphin they'd like, so they should be convinced that all dolphins can jump over a boat.

This process is useful language for writing proofs. We need to address *every* dolphin, but it gets difficult to keep writing sentence after sentence that keeps mentioning all dolphins in the world. So, the sentence "Let $s \in D$" instructs the reader to pick any one dolphin, and then from there and further, we don't need to refer to *all* dolphinsa anymore: we can just refer to s. This is a slick way of getting to talk about all dolphins (plural) by just talking about one dolphin (singular), namely s.

Example 273. *Say we need to prove $\forall v \in D[P(v)]$. As long as v has not appeared in the proof, write* $\boxed{\text{Let } v \in D}$*. Then work to prove* $\boxed{P(v)}$.

Why does this work? Imagine that D is the set of all dolphins, and $P(v)$ is the predicate "v can jump over a boat." Then $\forall v \in D[P(v)]$ in plain English is: every dolphin can jump over a boat. To prove this statement, we start by writing "Let $v \in D$" which allows the reader to pick any dolphin they like to be called v. Then we work to prove that $P(v)$ is true, which in plain English is: the dolphin v can jump over a boat. Since we allowed the reader to pick any dolphin that they want to be called v, and then we proved that v can jump over a boat, the reader is convinced that every dolphin can jump over a boat.

Example 274. *Say we need to prove* $\boxed{\text{for all } b \in C\text{, the manager gives } b \text{ a raise}}$ *is true. As long as b has not been written in your proof, write* $\boxed{\text{Let } b \in C}$*. Then work to prove* $\boxed{\text{the manager gives } b \text{ a raise}}$.

Why does this work? Let's say that C is the set of all Canadians. Then "for all $b \in C$, the manager gives b a raise" translates to: the manager gives a raise to every Canadian. When we write "Let $b \in C$," we are allowing the reader to pick any Canadian that they want to be called b. Then we work to prove that the manager gives b a raise. Since the reader could have picked any Canadian to be called b, and we proved that the manager gives b a raise, the reader is convinced that the manager gives a raise to every Canadian.

Example 275. *Say we need to prove* $\boxed{\text{for all } b \in C\text{, the manager gives } b \text{ a raise}}$ *is true. Suppose that the letter b is already a letter that appears in the proof. As long as m has not been written in your proof, write* $\boxed{\text{Let } m \in C}$*. Then work to prove* $\boxed{\text{the manager gives } m \text{ a raise}}$.

Example 276. *Let's say the task is that we need to prove* $\boxed{\text{for all } x \in S\text{, a delivery for } x \text{ is available}}$ *is true. Let's say that the letter r has not been written yet in your proof. Then write* $\boxed{\text{Let } r \in S \text{ be arbitrary}}$ *which allows the reader to pick any element from the set S that they want, and you will refer to their choice by the variable r. Then you need to prove* $\boxed{\text{a delivery for } r \text{ is available}}$ *is true.*

Why does this work? Suppose that S is the set of all geckos. Then "for all $x \in S$, a delivery for x is available" translates into "Every gecko has a delivery available" and when we write "Let $r \in S$ be arbitrary," we are allowing the reader to pick any gecko that they want to be called r. Then we work to prove: a delivery is available for r. Since the reader could have picked any gecko to be called r, and we proved that a delivery is available for r, the reader is convinced that every gecko has a delivery available.

Example 277. *Say we need to prove* $\boxed{\text{for all } a \in A\text{, if } a \text{ likes to sew, then } a \text{ rides motorcycles}}$. *As long as b hasn't been written in the proof yet, we would write* $\boxed{\text{Let } b \in A \text{ be arbitrary.}}$ *then we would need to prove* $\boxed{\text{if } b \text{ likes to sew, then } b \text{ rides motorcycles}}$, *so we would assume* $\boxed{b \text{ likes to sew}}$ *and work to prove* $\boxed{b \text{ rides motorcycles}}$.

The point of this example is to say that we take care of one thing at a time. Since the statement to prove starts with "for all" we do the usual thing by writing "Let $b \in A$ be arbitrary." Then, since the statement to prove is an implication, we follow Method 202 by assuming the hypothesis of the implication ("b likes to sew") and working to prove the conclusion of the implication ("b rides motorcycles").

3.1.8 Multiple quantifiers

When we see nested quantifiers, start slow. Address one quantifier at a time, in the order in which they appear. All of what we've discussed applies, and we just need to process one quantifier at a time. Let's look at some examples.

Example 278. *Say we need to prove* $\boxed{\forall x \in U \exists y \in V[Q(x,y)]}$. *It may help to write extra brackets for this, for now, taking it one step at a time. To prove* $\boxed{\forall x \in U[\exists y \in V[Q(x,y)]]}$ *we should first write* $\boxed{\text{Let } r \in U \text{ be arbitrary}}$ *as long as r hasn't been written in our proof. Then we would need to prove* $\boxed{\exists y \in V[Q(r,y)]}$. *Connecting this back to Method 268, the $P(x)$ that appears there is* $\boxed{\exists y \in V[Q(r,y)]}$ *here.*

It is important again to take care of things one step at a time. First, we address the "for all" quantifier by writing "Let $r \in U$ be arbitrary." Then, we would need to prove $\boxed{\exists y \in V[Q(r,y)]}$, *and this means that we'd have to find a specific thing (which we could name y), prove that thing is in V, then prove that $Q(r,y)$ is true for that specific thing.*

Example 279. *Say we need to prove* $\boxed{\text{for all } m \in A\text{, there exists } r \in B \text{ such that } m \text{ gives } r \text{ money}}$. *This is really in the same general structure/shape as the previous example, but in words. Then we would start by writing* $\boxed{\text{Let } m \in A \text{ be arbitrary}}$ *as long as m hasn't been written. (Otherwise, pick a variable that hasn't been mentioned in the proof yet.) Then we would work to prove* $\boxed{\text{there exists } r \in B \text{ such that } m \text{ gives } r \text{ money}}$.

To do the next step, we'd define something (which could be called r as long as r hasn't been written yet in the proof), then we'd prove $\boxed{r \in B}$ *and then we'd prove* $\boxed{m \text{ gives } r \text{ money}}$.

Example 280. *Say we need to prove* $\boxed{\text{for all } h \in S\text{, if } h \text{ has a cat, then } h \text{ is photographer}}$. *As long as h is hasn't appeared so far in our proof, then our proof of this statement would start* $\boxed{\text{Let } h \in S \text{ be arbitrary.}}$. *This allows the reader of our proof to pick their favorite element of S and call it h. Then we would need to prove* $\boxed{\text{if } h \text{ has a cat, then } h \text{ is photographer}}$ *which we could signal in the text of our proof by writing* $\boxed{\text{We will prove if } h \text{ has a cat, then } h \text{ is photographer.}}$ *with emphasis here on the words "we will prove" because we haven't proven this implication at this point yet. Then to prove this implication, following Method 202, we would write* $\boxed{\text{Suppose } h \text{ has a cat.}}$ *and then we would work to prove* $\boxed{h \text{ is a photographer}}$.

Example 281. *Suppose that we were given the task to prove "For all $m \in U$, there exists an $n \in L$ such that $n \geq m$." This task will be a bit impossible here only because we were not told what set U was nor what set L was. Nevertheless, the format of proving this statement would be as follows. We would start by writing "Let $m \in U$ be arbitrary" or can just write "Let $m \in U$" for short. Then, one can include the* *optional sentence "We will show that there exists an $n \in L$ such that $n \geq m$."*

At the end of the proof, there will probably a sentence like "Therefore, there exists an $n \in L$ such that $n \geq m$," but this very likely not be the sentence immediately after the "We will show..." sentence.

The examples above show that in a proof, it is necessary to address the quantifiers in the order in which they appear. In Section 2.6 we learned it is not okay to swap quantifiers (in general).

For a complete example, let us examine the following theorem:

Theorem 282. *For all $a \in \mathbb{Z}$, if b divides c, then b divides ac.*

Proof. Let $a \in \mathbb{Z}$ be arbitrary. We will prove that if b divides c, then b divides ac.

To prove that b divides c implies that b divides ac, let us suppose that b divides c. We will prove that b divides ac. Since b divides c, both b and c are integers and there exists an integer d such that $bd = c$. Let $z = ad$. Then z is an integer, and $bz = b(ad) = bad = abd = a(bd) = ac$. Since z is an integer and $bz = ac$, we have proved that b divides ac.

The previous paragraph proved that if b divides c, then b divides ac. Since the selection of $a \in \mathbb{Z}$ was arbitrary, we have proved that for all $a \in \mathbb{Z}$, if b divides c, then b divides ac. ☐

Theorem 283. *If*

- *For all $a \in B$, if a does not play chess, then a does not bike to school.*
- *For all $j \in T$, if j likes ketchup, then j bikes to school.*
- *Every element of S is an element of T.*
- *Every element of T is an element of B.*

then: For all $r \in S$, if r likes ketchup, then r plays chess.

Proof. Let $r \in S$ be arbitrary. We want to prove if r likes ketchup, then r plays chess. So suppose r likes ketchup. We will prove r plays chess. Recall $r \in S$, so by the third hypothesis, we can say $r \in T$ by modus ponens. Since $r \in T$, by the second hypothesis, if r likes ketchup, then r bikes to school. Since r likes ketchup, we conclude r bikes to school. Recall $r \in T$, so by the fourth hypothesis, $r \in B$. Since $r \in B$, and the first hypothesis is a "for all" statement we now get to use, we know if r does not play chess, then r does not bike to school. But we saw r bikes to school, so by modus tollens, r plays chess. ☐

Theorem 284. *Assuming the hypotheses*

- *H1: For all $r \in M$, if r is vegetarian, then r makes ceramic vases.*
- *H2: For all $w \in X$, if w makes ceramic vases, then there exists an $f \in Q$ such that w mails a gift to f.*
- *H3: For all $z \in X$, z is vegetarian.*
- *H4: Every element of Q is an element of Y.*
- *H5: Every element of X is an element of M.*

then, for all $b \in X$, there exists $c \in Y$ such that b mails a gift to c.

Proof. Let $b \in X$ be arbitrary. We want to prove there exists $c \in Y$ such that b mails a gift to c, for then we would be done. Since $b \in X$, by H3, we get b is vegetarian. Since $b \in X$, by H5 we have $b \in M$. Since $b \in M$, applying H1, we learn that b is vegetarian implies b makes ceramic vases. Since we already saw b is vegetarian, by modus ponens, b makes ceramic vases. Since $b \in X$, by H2, we get the following implication: if b makes ceramic vases, then there exists an $f \in Q$ such that b mails a gift to f. Using the implication we just got (via modus ponens) together with the fact that b makes ceramic vases, we conclude there exists an $f \in Q$ such that b mails a gift to f. Let $c = f$. Since $c \in Q$, we conclude $c \in Y$ by H4. Since $c \in Y$ and since b mails a gift to c, we have shown that there exists a $c \in Y$ such that b mails a gift to c. ☐

Exercise 285. *Prove: For all integers a and b, if there is an integer c such that $10c = a - b$, then there is an integer d such that $5d = b - a$.* [key]

Exercise 286. *Use the following hypotheses:*

- *H1: For all $r \in M$, if r likes Disneyland, then r is an element in the set S.*
- *H2: For all $c \in P$, if c walks to school, then c is a rock climber.*
- *H3: Every element in the set M is an element in the set P.*
- *H4: For all $b \in M$, if b does not like Disneyland, then b is not an element in the set P.*
- *H5: For all $s \in S$, the person s walks to school.*

to prove the proposition: For all $m \in M$, the person m is a rock climber.

Exercise 287. *Prove the following statement: If x is even and y is odd, then $x - y$ is odd.*

Exercise 288. *Prove: For all integers s, if s is even, then s^3 is even.*

Exercise 289. *Prove: For all integers s, if s is odd, then s^2 is odd.*

Exercise 290. *What is the contrapositive of $\boxed{\textit{If } s \textit{ is odd, then } s^2 \textit{ is odd}}$? You may assume the fact that an integer that is not odd is even, as found on page 38.*

Exercise 291. *Prove: If a and b are integers satisfying $2b^2 = a^2$, then a is even and b is even. Hint: use Exercise 290.*

Exercise 292. *Assume the two hypotheses*

- $\forall y \in B, \exists x \in A$ *such that* $P(x, y)$
- $\forall z \in C, \exists y \in B$ *such that* $Q(y, z)$

to prove: for all $c \in C$, there exists an $a \in A$ such that $\exists b \in B$ such that $P(a, b) \wedge Q(b, c)$.

Exercise 293. *Assume the two hypotheses*

- $\forall y_1 \in B\, \forall y_2 \in B\, \forall x_1 \in A\, \forall x_2 \in A$, *if* $y_1 = y_2$ *and* $P(x_1, y_1)$ *and* $P(x_2, y_2)$, *then* $x_1 = x_2$.
- $\forall z_1 \in C\, \forall z_2 \in C\, \forall y_1 \in B\, \forall y_2 \in B$, *if* $z_1 = z_2$ *and* $Q(y_1, z_1)$ *and* $Q(y_2, z_2)$, *then* $y_1 = y_2$.

to prove: for all $c_1 \in C$, for all $c_2 \in C$, for all $b_1 \in B$, for all $b_2 \in B$, for all $a_1 \in A$, for all $a_2 \in A$, if $c_1 = c_2$ and $P(a_1, b_1)$ and $P(a_2, b_2)$ and $Q(b_1, c_1)$ and $Q(b_2, c_2)$, then $a_1 = a_2$.

3.2 Intermission: comments on proofs

By rereading the three proofs of Theorem 238 and the proof of Theorem 245, notice each proof is written in complete sentences. A proof is a water-tight argument that starts from propositions which are assumptions and, following the rules of inference, convinces a reader of new propositions, eventually leading to a final proposition.

In a previous math class such as algebra or calculus, you were probably not expected to write in complete sentences. However, this old style of writing mathematics, while it can convey all of the key logic involved in a computation, is too primitive as a method of writing to convince a reader of the type of facts (theorems) that are truly proved in upper-level mathematics. We need complete sentences.

This type of writing, because it is new, will be challenging at first. You are encouraged to revisit this section and the previous sections in this chapter often, so that everything becomes more familiar. Chapter 3 is the core material of this handbook, and is the prerequisite to further chapters. This section states many pieces of advice, tips, and expectations:

Habit 294

Write your proofs in complete sentences. (Some of the things we mention later will help with this process.)

Habit 295

Every sentence in a proof ends in a period.

In fact, each sentence of your proof should end with a period, even if the last portion of the sentence is notation, and even if the last portion is center-lined as larger block of notation. So, instead of writing $\boxed{\text{Let } a = x + y}$ as a sentence, write $\boxed{\text{Let } a = x + y.}$ with the period at the end.

Habit 296

If you are concerned about writing in complete sentences, remember that there are only several types of sentences that you will write.

For example, if p, q, r, and s are all propositions, your sentences will likely be of one of the following formats:

- p.
- Since p, we have q.
- q since p.
- Therefore, q because p.
- Because p, q.
- So, q.
- Thus, q.
- Because r and s, we conclude q.

It is understandable that those who feel comfortable with mathematical calculations as done in calculus might be squeamish at first regrading writing proofs. You might want to say, "but I'm not a sentences-making person!" Take heart in knowing that the sentences written for a proof are little more than propositions surrounding a few connecting words such as thus, because, since, therefore, we see that, hence, etc. For example, the proof of Theorem 238 had the sentence "Since a is even, there exists an integer k such that $a = 2k$." which is in the grammatical format of "Since p, we have q." from our list above.

Habit 297

Remember that most of the hard work of a proof is in building the flowchart which is *very* much using the same type of thinking you did when doing computational mathematics.

The consistency of thought you have trained yourself to do in deciding between the Chain Rule versus the Quotient Rule has made you ready to build flowcharts of proof. Just as you can take the distributive law $a(b + c) = ab + ac$ to turn $5(a + b)$ into $5a + 5b$, you can take the flowchart for modus ponens and see that if you already have $q \to s$ is true and you already have q is true, you can conclude s is true. Think of the "variables" you see in the flowcharts as propositions to substitute in, just the same as you need to substitute to use $a(b + c) = ab + ac$.

Habit 298

The order of sentences matters.

Though you may have built the flowchart for your proof by discovering things in a different order, when turning your flowchart into the sentences of a proof, think about what relies on what. Consider a box in the flowchart, and note that you can only convince the reader that the proposition in that box is true only after having introduced the contents of the boxes pointing into this box: the arrows of your diagram can be thought of like the prerequisites for various classes in a major.

Habit 299

Keep sentences short.

You do not need to worry about overly flowery language. Try to communicate one idea (or maybe two ideas) in a sentence. Often, long sentences are trying to communicate too many things, and often, the things that are attempting to be discussed wouldn't even belong in consecutive sentences! Break up long sentences into smaller sentences. (Once you have it so that each sentence is communicating one idea, it is easier to also think about what order the sentences should appear.)

Habit 300

Don't be afraid of having lots of scratch work, though you'll probably never present your scratch work.

Clarify your task, then build a flowchart organizing the argument of the proof. A flowchart of proof logic involves *lots* of scratch work! Do not be shy about having scratch work! You may have become so accustomed to the calculations done in calculus that you forgot a time when you used a lot of scratch work. (Remember the first time you computed slopes of secant lines, there was a lot of work to write out!) A cleanly-written proof is nice, but often so nice that it hides the fact that there were probably a lot of dead ends, restarts, and so on.

Habit 301

Build the pieces of the flowchart which you can. Then work on connecting pieces.

Identifying what you need to prove (and looking up its flowchart) and identifying what you have not yet used (and looking up its flowchart). In our first example of a proof, we started with the first box and the box. Then we worked inward by creating the second box and the second-to-last box. Getting the main skeleton of the proof is half the battle!

Habit 302

Always keep track of whether you need to use a proposition or need to prove a proposition.

For each type of proposition (conjunction, implication, existentially-quantified statement, universally-quantified statement, and so on), the flowchart templates for proving and using are different. If you need to use a statement but are attempting to prove that statement, things will not work out. Similarly, if you need to prove a statement but are attempting to use that statement, things will not work out.

Habit 303

If you're stuck, build intuition from a concrete example.

As an example, let's say we were stuck at proving that the product of perfect squares is a perfect square. (We will need one new definition: an integer r is a **perfect square** if there exists an integer x such that $x^2 = r$.) We are looking to prove for all $r, s \in \mathbb{Z}$, if r is a perfect square and s is a perfect square, then rs is a perfect square.

Of course the proof of this universally-quantified statement is going to start with "Let $r \in \mathbb{Z}$ be arbitrary and let $s \in \mathbb{Z}$ be arbitrary" followed by a sentence such as "Suppose r and s are perfect squares." Because r is a perfect square, there exists an integer x such that $r = x^2$. Similarly, there exists an integer y such that $s = y^2$. If we feel stuck in proving that rs is a perfect square, we should build intuition from an example.

As our example, suppose $r = 9$ and $s = 25$. Then, the reason r is a perfect square is that there exists an integer (namely 3) such that $r = 3^2$. Similarly, $s = 5^2$. Connecting our example to the notation already introduced in our proof, $x = 3$ and $y = 5$. We need to prove rs is a perfect square, and in our current example, that is the number 225. When trying to prove rs is a perfect square, need to prove that there exists $z \in \mathbb{Z}$ such that $rs = z^2$. Connecting to our example, we need to find an integer z such that $225 = z^2$. While z needs to be 15 in our specific example, how does this provide a hint for how we should define z in our specific proof, only having access to r, s, x, and y? Notice that $15 = 3 \cdot 5$, so this suggests we should define z to be xy. After checking, it seems this will work, so here is a complete proof:

Let $r \in \mathbb{Z}$ be arbitrary and let $s \in \mathbb{Z}$ be arbitrary. Suppose r and s are perfect squares. Because r is a perfect square, there exists an integer x such that $r = x^2$. Similarly, there exists an integer y such that $s = y^2$. Let $z = xy$. Since the product of integers is an integer, $z \in \mathbb{Z}$. Then $rs = x^2 \cdot y^2 = (xy)^2 = z^2$. Since $z \in \mathbb{Z}$ and $rs = z^2$, we have proved that rs is a perfect square.

Note that the example is not a proof. (The proof was the eight sentences of the previous paragraph.) However, the example helped us in defining z to be xy, the key place where one usually gets stuck. Think of what is going on as a conversation. As the reader reads $\boxed{\text{Let } r \in \mathbb{Z} \text{ be arbitrary}}$, they might pick $r = 31$. They might similarly pick $s = 16$. Then, when the next sentence says "Suppose r and s are perfect squares" they need to go back, and choose differently. They might choose $r = 9$ and $s = 16$. From their choice of $r = 9$, they would determine that x could either be 3 or -3, and in this manner, they could use their own example to help guide their *reading* of the proof. However, the proof stands apart from any example. In fact, from the fact that they can pick r and s to be arbitrary perfect squares, they might "rewind" and choose $r = 100$ and $s = 121$. And once they read through the proof using $r = 100$ and $s = 121$, if they are still are still unconvinced, they could choose $r = 144$ and $s = 49$ and read through the proof again.

Habit 304

The language used to state propositions differs from the language in proofs.

For example, to *prove* a "for all" statement, your proof is going to contain text similar to "Let $x \in A$ be arbitrary." However, keep in mind that the phrase "For all $x \in A, \ldots$" is language used in *stating* a proposition that you will prove, while "Let ... be arbitrary" is very different language, and is used in *proving* that statement.

As another example, suppose the statement to prove is an implication. Then, you might be *proving* the statement "If p then q" but your proof of this will probably not focus on using the words "if" and "then." Instead, your proof will say something like "Suppose p is true." The word "suppose" is *not* going to appear when you state the proposition.

Habit 305

Remember that using a proposition and proving a proposition are very different.

$$\begin{tikzcd} \boxed{a} \arrow[d] \\ \boxed{b} \end{tikzcd}$$

In the small flowchart above the proposition a is being used, and the proposition b is being proved. In fact, a is used to prove b. On a very practical level, the methods used for proving and using an implication are very different, as seen from the very different looking flowcharts in Section 3.1.3. Similarly, proving versus using a conjunction have very different flowcharts. Proving verses using an existentially-quantified statement have very different flowcharts. For each type of proposition, this occurs. You will waste time if you try to prove something that you should use. Similarly, you will waste time trying to use something that you should prove.

Habit 306: Do not use a characterization as a definition

Do not use an "alternate definition" as a definition.

Recall the definition of odd: an integer m is odd if there exists an integer s such that $m = 2s+1$. While it turns out to be true that an integer m is odd if and only if there exists an integer k such that $m = 2k-1$, do not use this as the definition. The fact that m is odd if and only if there is an integer k such that $m = 2k-1$ is known as a **characterization** of being odd. A characterization should not replace the definition. (It is this characterization which we have called an "alternate definition" above.)

Think of each definition given as the *original* source material. With each newly-introduced term, we have to start somewhere, and that is the definition. When you are asked to recite a definition, do not provide a characterization instead.

Habit 307

Do not bother writing "by definition" all the time.

Perhaps this is a matter of taste, but if you were required to write "by definition" each time you needed a definition in your proofs, you would likely need to write "by definition" at least once (if not two or three times) per sentence. This would be too much clutter to include this all the time. There may be some key points of your proof where you would like to use this, but even in those cases, you might want to say something like "by definition of φ" or state "by definition of the normal subgroup."

Part of minimizing the expectation of writing "by definition" all the time points out a rather important aspect of proofs. You will need to reference definitions a lot! (This is why you hear your math instructor say "Know your definitions" so often: it is an essential part of proof.)

Habit 308

Be patient with yourself.

The list above has many suggestions on things to work on, and it may take a while to master them all. Though the explanation of each point may be longer, the expectation or the advice consists of one or two short sentences. Connect the advice/expectations to the examples of proofs already given, and challenge yourself to work on one of these things at a time. You'll likely want to revisit this list and keep working on these things.

3.2.1 The word "let"

We have used the word "let" in several different settings. We will discuss the various ways we have used the word "let" and provide a unified understanding of why the word is being used. There are three settings in our proofs where the word "let" has been used so far:

1. Using an existentially-quantified statement (Method 209)

   Initially, when using an existentially-quantified statement, the flowchart suggested writing a sentence like "Let $b \in U$ satisfying $P(b)$." However, by the time we got to shorter proofs of Theorem 238, we saw ways to avoid writing in such a lengthy manner. This was included for completeness, but it means that we are really only down to the next two uses of the word "let."

2. Proving an existentially-quantified statement (Method 217)

   To prove that there exists an $x \in U$ such that $P(x)$, the flowchart says to define an object (which we called c). This is often done by writing something like "Let $c = 2r^3$" or a similar statement. In the proof of Theorem 238, we wrote "Let $t = r + s$." At the time that we wrote this, note that r and s were already defined.

   Later, once you are comfortable with proving existentially-quantified statements, you may get to a point where you define an object as needed, but without naming it using a variable such as c. In this sense, it would be possible to say that you can avoid writing things like "Let $t = r + s$" but for now, it would be good practice to write things like "Let $t = r + s$."

3. Proving a universally-quantified statement (Method 268)

   To prove $\forall x \in U\,[P(x)]$, you must have a sentence such as "Let $r \in U$ be arbitrary" and then you would work towards proving $P(r)$. Writing "Let $r \in U$ be arbitrary" allows the reader to select any element of U they want to, and their selection is referred to as r, so that you have a name by which to call the reader's selection from U.

   Note that there is no use of the word "let" when *using* a universally-quantified statement.

These are the three situations in which you would use "let" in a proof, though you can avoid the first of these three. Thus, two situations remain. What does "let" mean? We use the word "let" in a proof to refer to some object later. When writing "Let $t = r + s$" there was already an r and an s which were introduced in the proof, and their sum was useful later on in the proof, so it helped to call $r+s$ something, and we chose to call it t. Instead of writing "Let $t = r + s$" we could have written "We define t to be $r + s$." Then, t is mentioned later on in the proof. Finally, if you write the sentence "Let $r \in U$ be arbitrary" then the reader may fix r to be any element from the set U. Just as in the previous example where t would be mentioned later on in the proof, here, r will be mentioned later in the proof.

In both of these situations, the word "let" allows the proof writer to define things which will be useful later. There was a time in the past when you used "let" although it might have been implicit. When evaluating

$$\int (1 - \sin^2 x)^3 \cos x \, dx$$

you had used substitution. While you may have just written $\boxed{u = \sin x}$ in the past, this is really short for "Let $u = \sin x$" after which you ended up writing u again when you wrote

$$\int (1 - u)^3 \, du.$$

The word let should always be followed by variable that has not been used yet in your proof:

- Ensure that the variable you are defining appears *immediately* after the word "let."

  Write "Let $t = r + s$" instead of writing "Let $r + s = t$." We are introducing t, and defining t to be $r + s$.

- Ensuring that the variable appears immediately after the word "let" also helps avoid a peculiar issue that might arise otherwise:

  Suppose, we are in the middle of a proof, where a, b, c, and d are all real numbers which have already been defined. Then if a proof has the sentence "Let $3a + b^5 + \ln(3 + c^2) - \sin(m) = m^5 + d$," the reader would (probably) interpret this to mean that the writing is trying to define m. However, it is not readily apparent that there even is *any* possible choice for m where $3a + b^5 + \ln(3 + c^2) - \sin(m) = m^5 + d$ is true. In other words, what if the equation $3a + b^5 + \ln(3 + c^2) - \sin(m) = m^5 + d$ has no solution for m? Writing like this becomes a difficult way to define the variable m, and insisting on the new variable appearing immediately after the word "let" will avoid this situation.

  As a slightly different version of this, sometimes people will write "Let m be an integer such that $3a + b^5 + \ln(3 + c^2) - \sin(m) = m^5 + d$." While this sentence follows the expectations written for the word "let" it has the same danger which was just mentioned. In fact, it is generally good to avoid writing a sentence in the form "Let $m \in U$ be such that $P(m)$" because this generally leads the reader to ask, "Wait, is there even any $m \in U$ where $P(m)$ is true in the first place?" The *only* time to write this type of sentence is immediately after the reader has been convinced that $\exists z \in U\,[P(z)]$.

- Be sure that "let" is followed by an *unused* variable.

  Otherwise, you would be redefining the variable. If your proof at some point said "Let $m = 3k + 4$" then the same proof should not later have the sentence "Let $m = 3k + 5$." The reader would think, "Wait! Earlier you told me that m would be defined to be $3k+4$. Now you're saying that m should be defined to be $3k+5$? Which is it?"

This would be redefining m. If you do need to define something to be $3k+5$, use a new variable.

Warning 309

While it is okay to write $\boxed{\text{Let } x = 7^2 + 3}$ it is not okay to write $\boxed{\text{Let } x \in \mathbb{Z} \text{ such that } 3x + 5 = 29 - x^2}$ (unless it comes immediately after an appropriate "there exists" statement) because there might not exist an integer x which satisfies $3x + 5 = 29 - x^2$.

Warning 310

As another example of this, it is incorrect to write "Let x be a real such that $x^2 + 49 = 3 - 2x$." Note that even among real numbers x, there is no solution to the equation $x^2 + 49 = 3 - 2x$. Thus, writing a sentence in this format (Let x ... such that ...) is dangerous because you as the writer may feel you have established something and want the reader to move on to the next thing, but nothing was established because there might not exist such an x. This is why the format of using 'let' is to have 'let' followed by a new variable (say x) followed by an equal sign, followed by an expression that is already defined (and therefore doesn't mention x). For example, writing "Let $x = u^2 + 3$" is okay, provided u already is established.

Contrast this with Method 209: if you already know $\exists x \in U$ such that $P(x)$, then it is okay to write $\boxed{\text{Let } x \in U \text{ such that } P(x)}$.

3.2.2 Other concerns

The higher-priority comments have been addressed earlier in this section, but there are other things to consider (some stylistic, some not) when writing mathematics.

- In formal mathematical writing, a sentence should not begin with notation.

  Instead of writing $\boxed{f \text{ is continuous on the interval } [3,9].}$ in formal writing, one should write $\boxed{\text{The function } f \text{ is continuous on the interval } [3,9].}$. The additional words "the function" help to clarify what type of object we have.

  There are situations when a less-than-formal style of writing is acceptable, but you will need to discuss this with your instructor. The typical situations for this are when time is of the essence (instructor's lecture, or students taking a quiz/test). In my own proof-based classes, writing $\boxed{f \text{ is continuous on the interval } [3,9].}$ would be okay on a quiz but not in submitted homework.

- Proofs will involve scratch work.

  Be sure not to turn in scratch work as the proof. Your own process of discovering will involve writing propositions (in a flowchart) in a different order than you might present them in a proof. Your proof should demonstrate careful flow of logic. Related to this:

- Each sentence should be readily apparent to the reader.

  When proving the implication following Method 202, we would first suppose that p is true. Then we eventually need to convince the reader that q is true, but this might not occur right away. To indicate that this is where we are headed, we might include "We will show that q is true." A sentence of this kind is not strictly *required*, but it helps the reader see where you are going in writing your proof. (In addition, while you are early in your proof-writing career, sentences like this help you on track of what you need to prove.)

As an example, see the proof of Theorem 237. After starting by supposing that a is odd and that b is odd, the next sentence cannot be "The integer $a+b$ is even" because the reader would not be convinced at this point that $a+b$ is even. We wrote "We want to prove $a+b$ is even" which is another way of saying "We will show that $a+b$ is even." This tells the reader that they should expect a sentence similar to "The integer $a+b$ is even" much later. (The words "prove" and "show" are synonymous in the context of writing proofs.)

The same can be said for Method 268. You will start by writing a sentence such as "Let $r \in U$ be arbitrary." Then, it will not usually be immediately clear to the reader that $P(r)$ is true. Convincing the reader of this might take several sentences. So instead of writing "Therefore, $P(r)$" (with the reader saying, "I'm unconvinced that $P(r)$ is true") you could write "We will prove that $P(r)$ holds" or something similar.

For example, see the proof of Theorem 282. The second sentence of the proof was "We will prove that if b divides c, then b divides ac." which is the specific version of the generic sentence "We will prove that $P(r)$ holds." The second-to-last sentence of the proof stated, "The previous paragraph proved that if b divides c, then b divides ac." It is at that moment in the proof that the reader is *finally* convinced that the implication is true. Between the second sentence and the second-to-last sentence is the proof of the implication.

When writing formal mathematics, it makes sense to write out complete phrases (such as "we want to show" or "we want to prove"). When writing less formally, one might use the abbreviations "WTS" and "WTP" for "want to show" and "want to prove," respectively.

- Recall (from Language Discussion 5) that notation for an object follows immediately after the noun which describes the object.

  For example, write "The maximum M of the set S is" and continue the sentence. Do not use a comma before and after the M by writing "The maximum, M, of the set..."

- As you describe the process of logic to yourself and to others, be clear about the difference between an assumption and a conclusion.

  In fact, think about using the phrases "assumption" and "conclusion" and "intermediate conclusion." The box at the very top of your flowchart is (likely) going to be your assumption. The box at the bottom will be your conclusion. All the boxes in the middle are intermediate conclusions. So, it would be incorrect to take the flowchart for Theorem 238 and describe your work to a fellow student or to the instructor by saying "I assumed that a is even and that b is even, and then I assumed that a is even, and then assumed that there exists an integer k such that $a = 2k$." In this sentence, the first use of "assumed" was correct, but the other pieces are intermediate conclusions: one could say, "I assumed that a is even and that b is even, and from there concluded that a is even, and from this, obtained another intermediate conclusion that there exists an integer k such that $a = 2k$."

- Avoid notation both right before and right after a comma if the comma indicates a break between clauses

  Because we often write "For all $a, b \in \mathbb{R}$" as short hand to mean "For all $a \in \mathbb{R}$ and for all $b \in \mathbb{R}$" having notation right before and after a comma may be a little ambiguous. Consider the sentence "Since the integer a divides b, ac divides br." Reading the b before the comma and the ac after the comma may be confusing. Is the author suggesting that a divides b and that a also divides ac? Certainly if the sentence said "The integer a divides b, ac" then this would have to be the interpretation. However, based on the sentence structure, it appears that reader meant something else.

  Instead of writing "Since the integer a divides b, ac divides br" it is preferable include some text right after the comma: the sentence "Since the integer a divides b, the integer ac divides br" removes the ambiguity of notation immediately before and after the comma. Moreover, adding the phrase "the integer" before ac helps clarify that ac is an integer, following Language Discussion 5. While the sentence is slightly longer, not only is the sentence clearer, there is also a benefit to the proof writer in being forced to think about "What kind of thing is ac anyway?"

- Writing the word “assume” and the word “suppose” in a proof are essentially synonymous.

Warning 311

Note that:

- proving $\forall x \in U[P(x)]$
- using $\forall x \in U[P(x)]$
- proving $\exists x \in U[P(x)]$
- using $\exists x \in U[P(x)]$

are four very different-looking tasks. For example, if you need to prove $\exists x \in U[P(x)]$, don’t start by writing $\boxed{\text{Let } x \in U}$ as that is what you’d do to start proving $\forall x \in U[P(x)]$.

3.3 Intermission: comments on definitions

You will need to reference definitions to be successful at proofs. Imagine trying to prove Theorem 237 without having Definition 230. It would be impossible. (People have tried to ignore the definitions while writing a proof of Theorem 238. They end up writing things like “Suppose a is odd” but then have no where to really go. While people then end up trying to write things like “Thus, a has a remainder of 1 when divided by 2,” none of our methods of proof which we have described are helpful. By applying the definitions using the methods we have given, one is led to the useful intermediate conclusion of $\exists k \in \mathbb{Z}\,[a = 2k + 1]$. Then, Method 209 can apply. The usefulness of this approach is that essentially every area in proof-based mathematics will require being familiar with Method 209, so even in what appears to be a very “familiar” statement in Theorem 237 is a good place to practice referencing definitions and using the methods that we have presented in this chapter.

The new definitions introduced in this chapter (even, odd, divides) and the definitions you will see in the future are written from the perspective that the definitions in the previous chapter (conjunction, logically equivalent, implication, etc.) are the language foundation. Every time you encounter a new definition, you should:

1. Keep in mind that the text following the first “if” is the defining property. Convert this proposition into symbols. Write its negation in symbols.
2. Identify the new word.
3. Identify its part of speech. (noun, adjective, verb)
   - For an adjective what is the noun which is being modified? (example: an injective function is a special kind of function).
4. What type of object is being defined? (a proposition, a set, a function, a relation?)
5. Are there other words being defined in the process? (Example: hypothesis and conclusion are words that are defined when defining implication)
6. Pay attention to the notation and grammar/usage in the text of the definition.
7. Make your own examples and non-examples.
   - For an adjective (for example, an injective function) think of a function which *is* injective and a function which is *not* injective.
   - For a noun, ensure your non-example just barely breaks the defining property. (example: insert an ordered triple into a binary relation to make a set which is not a binary relation)

8. Read book's examples and non-examples.

9. Pay attention to (and mimic) the notation and grammar/usage in the text of the examples.

If your work in reading definitions has been thorough, then you will have more success with writing definitions. When writing definitions, do the following:

1. Try writing something from memory.

2. Write out wordy (something you might say out loud, yet imprecise), mid-length and precise (as found in books/class), and compact (beyond the defining word, only using logical symbols). The wordy version may be imprecise, but serves as the "memory hooks',' needed for the mid-length definition. Here is an example from abstract algebra:

   - Wordy: A group is **abelian** if any two elements in the group commute.
   - Mid-length: A group G is **abelian** if $ab = ba$ for all $a, b \in G$.
   - Compact: A group G is **abelian** if $\forall a, b \in G$, $ab = ba$.

   The compact statement (example: the proposition $\forall a, b \in G$, $ab = ba$) or any logically equivalent statement is the has all the technical pieces of the definition, but we tend to expand this out a little (in the mid-length version) and present this as a working definition when communicating with fellow mathematicians. The mid-length statement should capture all the precision of the compact statement, but be more readable (can mix symbols and English). All of the definitions that we have provided (and will continue to provide in future sections) in boxes in this handbook are of the mid-length style, the typical style found in mathematics textbooks. The wordy statement is *never* acceptable in written work, but is often the language used verbally, and should capture the essence (you should be able to use the wordy to help you remember the compact).

   As a habit, with every definition encountered in Chapter 4 and beyond, practice reading the formally-presented definition (likely of mid-length style) and write something wordy (to help serve as a memory hook), but also write something compact (because it will fit the methods of proof in this chapter better).

3. Check that you have "set up" information, either as its own earlier sentence or incorporated into the defining sentence. (Example: Let S and T be sets.)

4. Does your definition make clear the part of speech? (noun, adjective, verb)

5. Does your definition make clear the type of object being defined? (a set, a function, a relation, a group)

6. Did you introduce necessary notation before defining?

7. Did you use all notation introduced? (Otherwise, can you leave out that notation?)

8. Does the language you use suggest the accepted notation and grammar/usage?

Habit 312

Every time you encounter a new definition, determine if the word or phrase being defined is part of a more complete phrase.

Do not be distracted by the *names* of definitions, which sometimes lead students to create inaccurate definitions.

Recall Definition 229 which states that an integer n is even if there exists an integer k such that $n = 2k$. The text $\boxed{\text{there exists an integer } k \text{ such that } n = 2k}$ defines what it means that $\boxed{\text{integer } n \text{ is even}}$. The main thing to rely on is $\boxed{\text{there exists an integer } k \text{ such that } n = 2k}$. "Even" is just a name.

Later, in Definition 607, you will learn the definition of the word **onto**. Based on the word alone, it is tempting to say that **onto** is a preposition (as in, "they went up onto the ridge"). However, we will see that onto is an adjective. Treating **onto** like a preposition would be dangerous, and is using our everyday English understanding of the word, which is a distraction.

Similarly, from second-semester calculus, there is no immediate reason why the definitions of "convergent" and "absolutely convergent" are connected at all. Sure, their *naming* is similar, but if you look at the definitions of the two in any calculus book, you'll see that there's almost nothing similar about them in the defining text. The fact that the latter implies the former is settled due to the proof of a theorem.

As another example, from third-semester calculus, if D is a subset of the Cartesian plane, we discuss the definition of the **boundary** of D, and whether D is **bounded** or not. While both words start with the same five letters, the two concepts are completely unrelated. Any attempt to connect the two (simply because the words sound similar) creates misconceptions in third-semester calculus! (In fact, **boundary** is a noun, while **bounded** is an adjective.)

3.4 Indirect proof

Method 313: Indirect proof of $p \to q$

To prove $p \to q$, one can prove its contrapositive $\neg q \to \neg p$, since the contrapositive is logically equivalent:

$$\boxed{\text{Assume } \neg q} \downarrow \cdots \downarrow \boxed{\text{Prove } \neg p}$$

Theorem 314. *For all $s \in \mathbb{Z}$, if s^2 is even, then s is even.*

To start a proof, we would write "Let $s \in \mathbb{Z}$ be arbitrary." Then we would need to prove if s^2 is even, then s is even. Our usual way of proving this would be to suppose that s^2 is even. So, there would exist an integer k such that $s^2 = 2k$. But then how would we ever express s? It seems awkward to square root both sides and have $s = \pm\sqrt{2k}$. Where would we go from here? It appears that a direct proof of the implication is awkward. Let us try an indirect proof (with the selection of $s \in \mathbb{Z}$ still arbitrary). Before starting the proof, recall from Section 2.7 that an integer that is not odd is even, and that an integer that is not even is odd.

Proof. Let $s \in \mathbb{Z}$ be arbitrary. We wish to show that if s^2 is even, then s is even. To prove this, suppose s is not even. Then, due to Section 2.7, s is odd. We wish to prove that s^2 is odd. Since s is odd, there exists an integer a such that $s = 2a+1$. Let $b = 2a^2 + 2a$. Then b is an integer, and $s^2 = (2a+1)^2 = 4a^2 + 4a + 1 = 2(2a^2 + 2a) + 1 = 2b + 1$. Since $s^2 = 2b + 1$ and b is an integer, s^2 is odd. $\square$

If you need to prove an implication, always try a direct proof. However, if you get stuck in a direct proof, consider proving the contrapositive of the implication instead (which is called the **indirect proof**).

Exercise 315. *Prove: for all $k \in \mathbb{Z}$, if k^2 is odd, then k is odd.*

Warning 316

Notice that to prove $p \to q$ using the indirect method, we end up assuming $\neg q$. We *never* assume q. (That is, never assume what you need to prove.)

3.5 Proof by contradiction

There are situations where the usual proof techniques do not lead to the required propositions. A new method – proof by contradiction – offers another opportunity to prove certain statements by allowing the proof writer to add an additional assumption.

Suppose that we need to prove q. If we feel that we are stuck having used all the propositions we were allowed to assume, we might add a new assumption, namely $\neg q$.

Method 317: Prove q by contradiction

$$\boxed{\text{Assume } \neg q.} \downarrow \cdots \downarrow \boxed{\text{Obtain contradiction}}$$

By adding $\neg q$, if we can obtain a contradiction, then it must have been erroneous to assume $\neg q$, so we could conclude that q is true. In other words, due to having a contradiction, we would then conclude that our original assumption of $\neg q$ must have been false, which means that q is true.

Theorem 318. *The real number $\sqrt{2}$ is irrational.*

Recall that a real number r is rational if there exists an integer b and a non-zero integer c such that $r = \frac{b}{c}$. In the proof we look at below, we will even look at the stronger situation where $\frac{b}{c}$ is reduced as much as possible. So, for instance, if r is $\frac{4}{12}$, we would need to reduce the fraction to $\frac{1}{3}$. Note that, once $r = \frac{4}{12}$ is reduced, then $b = 1$ and $c = 3$ do not share any common prime factors.

Proof. In order to obtain a contradiction, suppose that $\sqrt{2}$ is rational. Then, there exists an integer b and a non-zero integer c such that $\sqrt{2} = \frac{b}{c}$, and we consider the situation where $\frac{b}{c}$ is reduced, so that b and c do not share any common factors greater than 1.

Since $\sqrt{2} = \frac{b}{c}$, we have $c\sqrt{2} = b$ by multiplication, and $2c^2 = b^2$ by squaring. But since $b^2 = 2(c^2)$ proves that b^2 is even, by Theorem 314, b is even. Since b is even, there exists an integer k such that $b = 2k$. Then $2c^2 = (2k)^2$, so $2c^2 = 4k^2$, and by division, $c^2 = 2(k^2)$. Since c^2 is even, c is even by Theorem 314.

But if b and c are both even, then they share common factors greater than 1, contradicting the earlier statement that b and c do not share any common factors greater than 1. Our original assumption that $\sqrt{2}$ is rational must have been incorrect: therefore, $\sqrt{2}$ is irrational. □

Warning 319

To prove q by contradiction, note that we do not assume q is true. Rather, we assume that $\neg q$ is true then work towards obtaining a contradiction.

Now, we examine a special case of proof by contradiction. The method we are about to describe is just a special case of the general proof by contradiction that we have already described, but it occurs so frequently that we write extensively about this special case.

Suppose that we need to prove the implication $p \to q$. Following the method of direct proof, we would assume p. Then, we would need to prove q. However, if we feel that we are getting nowhere in proving q, we can prove q by contradiction. Following the method already mentioned, we would assume $\neg q$, then work to obtain a contradiction. In other words, to prove $p \to q$ by contradiction, we assume two propositions: first, we assume p, and then we assume $\neg q$. Then, once we obtain a contradiction, we would then conclude that, our assumption of $\neg q$ must have been false.

Method 320: Proving $p \to q$ by contradiction

$$\boxed{\text{Assume } p. \text{ Assume } \neg q.} \to \cdots \to \boxed{\text{Obtain contradiction}}$$

Warning 321

Notice that to prove $p \to q$ by contradiction, we end up assuming $\neg q$. We *never* assume q. (That is, never assume what you need to prove.) This mirrors the warnings mentioned as Warning 316 and Warning 319. In this warning as in the previous two warnings, we never assume the proposition q.

As an example, let us visit the game of Minesweeper. A typical cell (those found in the middle of the Minesweeper board) is surrounded by 8 other cells. (The cells on the edge are surrounded by only five other cells, while the four corner cells are surrounded by only three other cells.) The "raised" cells are unknowns. The more "indented" cells are known. (Indented cells without a number are zeros: none of the 8 surrounding cells are mines.) The cells which have a red triangular flag are marked by the game player (usually by right-clicking) to indicate that these are mines.

To carefully understand, let us consider the example in Figure 3.1. The cell labeled 2 to the right of Cell Y should have two mines. The cell to the upper right (with a 2), the cell to the right (with a 1), the cell to the lower right (blank, so a zero), and the cell below (the 1 to the right of Z) have all been cleared. The cell above the 2 is already marked as a mine. While that spot is "technically not known", the person playing felt confident enough to mark that spot a mine. In addition, Cell X, Cell Y, and Cell Z are unknowns. For an example of the language of proof that can be used, let us consider the same example. The cell labeled 2 to the right of Cell Y should have two mines. Since there is already one marked mine, among the three cells Cell X, Cell Y, and Cell Z, exactly one of these is a mine.

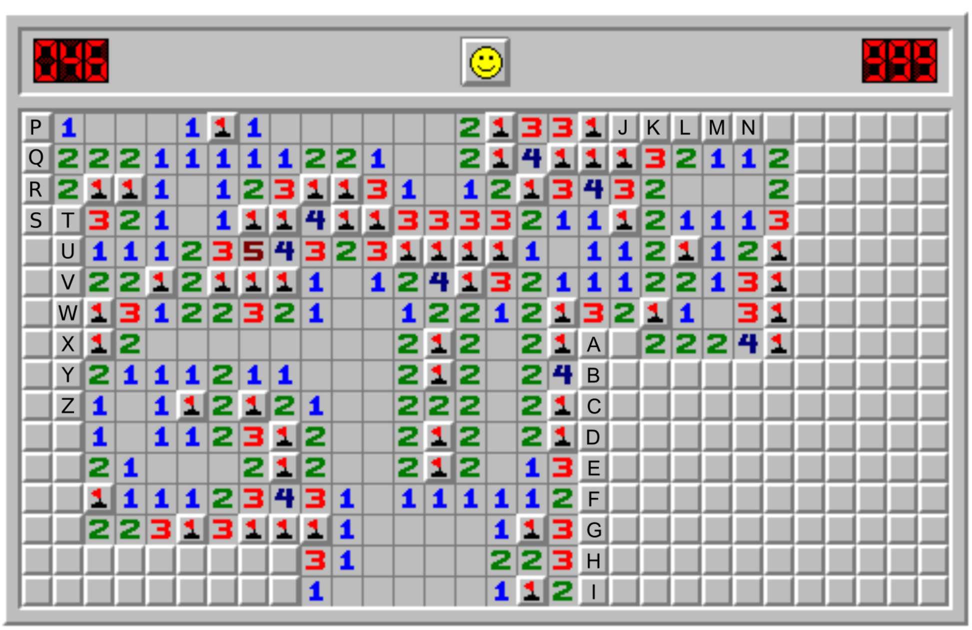


Figure 3.1: A player who is "stuck" in Minesweeper

The game of Minesweeper gives us a fun way to examine proving $p \to q$ by contradiction. How would one prove the statement "If the Minesweeper configuration is as given in Figure 3.1, then Cell I is safe" by contradiction? Following Method 320, we should first assume we have the mine configuration given in Figure 3.1. Then, to prove that Cell I is safe by contradiction, we assume that Cell I is not safe – in other words, we assume that Cell I is a mine.

Theorem 322. *If the Minesweeper configuration is as given in Figure 3.1, then Cell I is safe.*

Proof. Suppose we have the given mine configuration from Figure 3.1. To obtain a contradiction, suppose that Cell I is a mine. Then the 3 to the left of H has all of its mines so, Cell H and Cell G are both safe. Then the 3 to the left of G has one discovered mine, and one possible mine at F, leaving only at most two neighboring mines to a cell labeled with 3, a contradiction. ☐

In the previous proof, the contradiction obtained is that both "F has only two neighboring mines" and "F has exactly three neighboring mines" cannot both be true.

Theorem 323. *If the Minesweeper configuration is as given in Figure 3.1, then Cell V is safe.*

Proof. Suppose we have the given mine configuration from Figure 3.1. In order to obtain a contradiction, assume that Cell V is a mine. Then because the cell labeled 1 to the right of U already has its 1 mine at V, the cells T and U must be safe. Then, the cell labeled 3 to the right of Cell T only has the two-previously discovered mines, and since the T and U are not mines, the 3 is not satisfied, a contradiction. ☐

Exercise 324. *Prove by contradiction: If the Minesweeper configuration is as given in Figure 3.2, then Cell B is safe.*

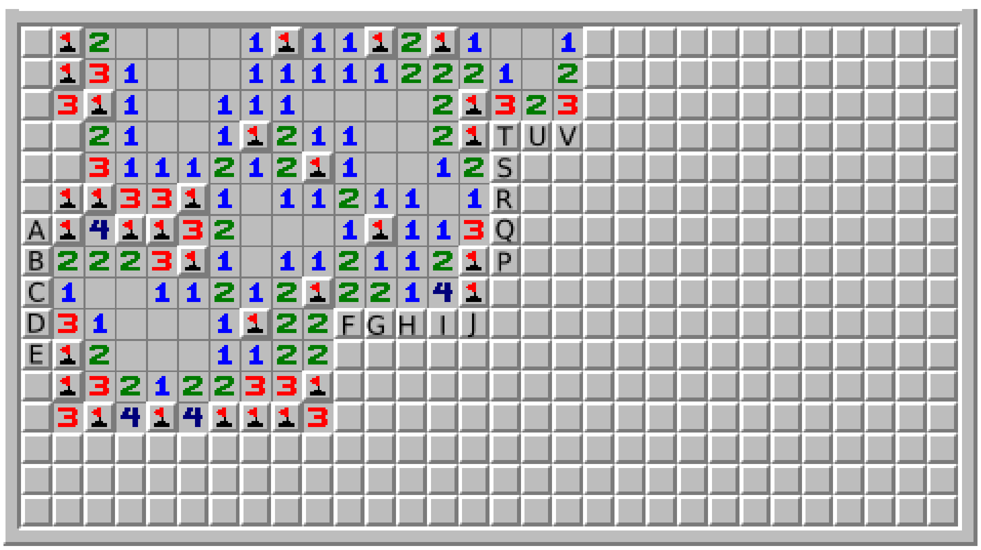


Figure 3.2: A player who is "stuck" in Minesweeper

When proving $p \to q$ by contradiction, you first assume p, and then because you want to prove q by contradiction, you assume $\neg q$. There are three ways that proving $p \to q$ may end:

- You end up with $\neg p$ which means you end up with the contradiction $p \wedge \neg p$. Typically, though not always, a proof of this kind can be restructured to an indirect proof of $p \to q$ which avoids a proof by contradiction.
- You end up with q, which means you end up with the contradiction $q \wedge \neg q$. These typically feel bizarre, because you assumed $\neg q$ and you somehow used p and $\neg q$ to obtain q.
- You end up with $r \wedge \neg r$, for some proposition r which is different from p and different from q.

3.6 Proof by cases

In the last section, we saw the method of proof by contradiction as one way forward when it seems like we are stuck in a proof. Another method that is useful in such "sticky" situations is a proof by cases.

Suppose that one wishes to prove $p \to q$. The natural method is to assume p first, then attempt to prove q. However, if you get stuck proving q, you might try to identify k propositions, which we will name $h_1, h_2, \ldots, h_k$, where $k \geq 2$. Typically, k should be small: perhaps no more than 5, ideally. The goal is to have propositions $h_1, h_2, \ldots, h_k$ where the following can be proved:

- $p \to (h_1 \vee h_2 \vee \cdots \vee h_k)$
- $h_1 \to q$
- $h_2 \to q$
- $h_3 \to q$
- and so on until
- $h_k \to q$

If all the proofs above are successful, then one can validly conclude that q is true.

Why does this work? The proof of $p \to (h_1 \vee h_2 \vee \cdots \vee h_k)$ is the identification of the cases. For simplicity, let us consider a situation where $k = 3$. That is, let us examine $p \to (h_1 \vee h_2 \vee h_3)$. This is saying that from p, one of three things must occur: either h_1 must be true, or h_2 must be true, or h_3 must be true. At least one of these three propositions must be true. In other words, we want to ensure that it is inevitable that at least one of h_1 or h_2 or h_3 are true. Then, no matter which of these three things are true, we will reach our destination. If h_1 is true, then we get to apply the proof that we present of $h_1 \to q$. If instead h_2 is true, then we apply the proof of $h_2 \to q$. If h_3 is true, then we apply the proof of $h_3 \to q$. Here we saw an example of $k = 3$, which is a proof in three cases.

For a proof in four cases ($k = 4$), we must identify four propositions h_1, h_2, h_3, h_4. The proof's author should first argue that $p \to (h_1 \vee h_2 \vee \vee h_3 \vee h_4)$ is true. Then there are four additional proofs to do, namely $h_1 \to q$ and $h_2 \to q$ and $h_3 \to q$ and $h_4 \to q$.

Let us consider the following example:

Theorem 325. *If the Minesweeper configuration is as given in Figure 3.1, then Cell N is safe.*

Proof. Assume we have the given mine configuration from Figure 3.1. Because of the 3 below K, of the three cells J, K, and L, exactly two must be mines. That is, either J and K are mines with L safe, or J and L are mines with K safe, or K and L are mines with J safe. We break into three cases:

- If Cell J and Cell K are mines and Cell L is safe, then Cell M must be a mine so that the 2 below Cell L has its second mine. Then the 1 below Cell M has all its mines discovered, so Cell M is safe and Cell N is safe.
- If Cell J and Cell L are mines and Cell K is safe, then because of the 2 under Cell L, Cell M must be a mine. However, this case is impossible because the 1 below M is surrounded by mines at both Cell L and Cell M.
- If Cell K and Cell L are mines and Cell J is safe, then the 1 under Cell M has its mine located at Cell L, so Cell M and Cell N are both safe.

No matter which of the three cases occur, we can see that Cell N is safe. □

In this example, $k = 3$. Note that "we have the given mine configuration from Figure 3.1" is p, following the notation from earlier. Let us identify the three situations:

- h_1 is "J and K are mines with L safe"
- h_2 is "J and L are mines with K safe"

- h_3 is "K and L are mines with J safe"

and our proof of $p \to (h_1 \vee h_2 \vee h_3)$ is embedded in the short sentence, "Because of the 3 below K, of the three cells J, K, and L, exactly two must be mines." In the proof, after the sentence "We break into three cases" we have presented proofs of $h_1 \to q$ and $h_2 \to q$ and $h_3 \to q$. No matter what, we reach the conclusion of q, where q is namely "Cell N is safe." (One word about the second case, where K was assumed safe: it turned out that this was impossible, so while it seemed that we had three cases in the beginning, in retrospect, we only had two cases.)

The idea of proof by cases is to branch out from the situation of p and look at one of a number $(h_1, \ldots, h_k)$ situations, at least one of which *must* hold. Then, assuming each of these situations, we either determine that the situation was actually impossible (as we had in our second case above) or that we reach the conclusion that q is true (as we did in cases one and three above).

Let us revisit Theorem 323, which we earlier proved by contradiction, and now provide a proof of the same theorem by cases. In this proof by cases, the proof of the first case is an embedded proof by cases.

Theorem 326. *If the Minesweeper configuration is as given in Figure 3.1, then Cell V is safe.*

Proof. Suppose we have the given mine configuration from Figure 3.1. Because of the 1 in row 1 column 2, either Cell P is a mine or Cell Q is a mine. We proceed by cases.

- In the first case, P is a mine. So Q is safe. Since P is a mine, the 2 in $(2, 2)$ has all necessary mines, so R is safe. Then exactly one of S or T is a mine, due to the 2 in $(3, 2)$. We proceed in two subcases:

  – If S is a mine and T is safe, then U must be a mine for the third mine of the 3 located at $(4, 3)$. Since the 1 to the right of U has all its mines, V must be safe.

  – If T is a mine and S is safe, then U is safe because the 3 located at $(4, 3)$ has all its mines. In addition, the 1 to the right of U already has its mine at T, so V is safe.

  In both subcases V is safe.

- In the second case, Q is a mine, so P is safe. Since Q is a mine, the 2 in $(3, 2)$ has all necessary mines, so R, S, and T are all safe. Since T is safe, due to the 3, U is a mine. Since the 1 to the right of U has all its mines, V is safe.

In both cases, V is safe. □

Let us analyze the structure of this proof and see how it fits our framework. Following our notation from earlier, we have an initial hypothesis p that "we have the given mine configuration from Figure 3.1" and then identified two propositions, where h_1 was "Cell P is a mine" and h_2 was "Cell Q is a mine." Our proof of $p \to (h_1 \vee h_2)$ is given in the sentence "Because of the 1 in row 1 column 2, either Cell P is a mine or Cell Q is a mine."

After the sentence "We proceed by cases" in the proof above, the proof presented a proof of $h_1 \to q$ and a proof of $h_2 \to q$, where q was "Cell V is safe." We postpone the discussion of the proof of $h_1 \to q$ momentarily. The proof of $h_2 \to q$ was the text that started with "In the second case, Q is a mine, ..." and ended with "Since the 1 to the right of U has all its mines, V is safe."

What about the proof of $h_1 \to q$? We proved this by cases. Here, the initial hypothesis was "P is a mine", though we may of course use our earlier initial hypothesis p as well, which is how we immediately obtained the next sentence "So Q is safe," and further along obtained the fact that R is safe as well. Within the proof, we identified two possible situations: let us call i_1 the situation of "S is a mine an T is safe" and call i_2 the situation "T is a mine and S is safe." The proof that the initial hypothesis implies $(i_1 \vee i_2)$ is proved in the sentence "Then exactly one of S or T is a mine, due to the 2 in $(3, 2)$." The inner-indented items provide the proof of i_1 implies that V is safe and the proof of i_2 implies that V is safe.

Warning 327: The error of missing cases

When proving by cases, it is important to state the case hypotheses $h_1, \ldots, h_k$. It is important to explain why at least one case hypothesis holds. (That is, be sure to include an argument – often very short – that $p \to (h_1 \vee \cdots \vee h_k)$ is true.) Otherwise, you run into the situation of having missing cases.

When initially starting, you might identify propositions h_1, h_2, and h_3. However, if you have trouble proving $p \to (h_1 \vee h_2 \vee h_3)$, it might be because $p \to (h_1 \vee h_2 \vee h_3)$ is not true. Perhaps there is a "missing case." Sometimes, by keeping h_1,h_2, and h_3 exactly as you thought of them, but also identifying a new proposition h_4, it is possible to prove $p \to (h_1 \vee h_2 \vee h_3 \vee h_4)$. When it is time to write your formal proof, imagine the immediately before writing "We proceed in cases" you would write the sentence "Either h_1 or h_2 or h_3 or h_4." You want the reader of your proof to be convinced that "Either h_1 or h_2 or h_3 or h_4" is true.

Warning 328: The error of not reaching the required conclusion

If you have k cases, beyond the proof of $p \to (h_1 \vee \cdots \vee h_k)$, be sure to provide the k proofs of the form $h_j \to q$. In particular, be sure to reach the conclusion q in each case.

Exercise 329. *Prove by cases: If the Minesweeper configuration is as given in Figure 3.2, then Cell V is a mine. (While it is possible to provide a proof by contradiction, it would be good to practice proof by cases. As a hint, start where Cell P, Cell Q, and Cell R are, and create the three cases. Note that one of the cases will be impossible, but there will still be the remaining cases to consider.)*

Exercise 330. *Prove: If the Minesweeper configuration is as given in Figure 3.2, then Cell F is a mine.*

We started this chapter by discussing how one uses a conjunction, but did not discuss using a disjunction. Often, using a disjunction fits the pattern of a proof by cases. Some details are provided in the next section.

3.7 Rules of inference for disjunctions

To use $p \vee q$, if we knew $p \vee q$ and we knew $\neg p$, then we can conclude q. More generally, if we know $p_1 \vee p_2 \vee p_3$ and we knew $\neg p_1$ and we know $\neg p_2$, then we can conclude p_3.

Often, we use a disjunction in a way that essentially follows a proof by cases, because we are often in the situation of needing to prove the implication $(p_1 \vee p_2 \vee \cdots \vee p_k) \to q$. Of course, we'd immediately assume $(p_1 \vee \cdots \vee p_k)$. Then if we prove $p_1 \to q$ and prove $p_2 \to q$ and so on, all the way up to proving $p_k \to q$, then we can conclude q by cases, completing our proof of $(p_1 \vee p_2 \vee \cdots \vee p_k) \to q$.

Suppose we had to prove the disjunction $p \vee q$. Of course, if we knew p was true, we can conclude $p \vee q$.

$$\begin{tikzcd} \boxed{p} \arrow[d] \\ \boxed{p \vee q} \end{tikzcd}$$

Likewise, if we knew q was true, we can conclude $p \vee q$.

Typically, a bit more may be needed to prove $p \vee q$. To prove $p \vee q$, one can add in the assumption of $\neg p$ and then attempt to prove q. Alternately, one can add in the assumption of $\neg q$ and then attempt to prove p. If one approaches a proof of $p \vee q$ by adding in the assumption of $\neg p$, this can be thought of as a proof by cases: either p or $\neg p$. If p holds, then $p \vee q$ is automatic. If $\neg p$ holds, then the provided proof of $(\neg p) \to q$ is the proof of the second case.

3.8 Rules of inference for biconditionals

Suppose you need to prove the biconditional $p \leftrightarrow q$. Since $p \leftrightarrow q$ is logically equivalent to $(p \to q) \wedge (q \to p)$, one can leave proving $p \leftrightarrow q$ to proving $p \to q$ and proving $q \to p$. The proof of $p \to q$ is often informally called the **forward direction** with the proof of $q \to p$ called the **reverse direction**. Of course, either or both implications can be proved by considering the contrapositive instead. That is, here are four possible methods for proving $p \leftrightarrow q$.

- Directly prove $p \to q$. That is, assume p, then prove q. Then, directly prove $q \to p$. That is, assume q, then prove p.

- Directly prove $p \to q$. That is, assume p, then prove q. Then, indirectly prove $q \to p$. That is, assume $\neg p$, then prove $\neg q$.

- Indirectly prove $p \to q$. That is, assume $\neg q$, then prove $\neg p$. Then, directly prove $q \to p$. That is, assume q, then prove p.

- Indirectly prove $p \to q$. That is, assume $\neg q$, then prove $\neg p$. Then, indirectly prove $q \to p$. That is, assume $\neg p$, then prove $\neg q$.

To use $p \leftrightarrow q$, we again appeal to its logical equivalence to $(p \to q) \land (q \to p)$. Thus,

- Knowing $p \leftrightarrow q$ and knowing p means we can conclude q.

- Knowing $p \leftrightarrow q$ and knowing q means we can conclude p.

- Knowing $p \leftrightarrow q$ and knowing $\neg p$ means we can conclude $\neg q$.

- Knowing $p \leftrightarrow q$ and knowing $\neg q$ means we can conclude $\neg p$.

There is a generalization of $p \leftrightarrow q$. One might read a theorem whose statement says "**the following are equivalent** (TFAE)

- a

- b

- c

- d"

where a, b, c, and d are all propositions. This is taken to mean that $a \leftrightarrow b$, and $a \leftrightarrow c$, and $a \leftrightarrow d$, and $b \leftrightarrow c$, and $b \leftrightarrow d$, and $c \leftrightarrow d$.

There are many ways to prove such a statement. For example, one can prove $a \to b$ and prove $b \to c$ and prove $c \to d$ and prove $d \to a$. Alternately, one can prove $a \to b$ and prove $b \to c$ and prove $c \to a$ and prove $a \to d$ and prove $d \to a$. Here are diagrams that show both of these major proof outlines:

$$\begin{tikzcd} a \arrow[d] & d \arrow[l] \\ b \arrow[r] & c \arrow[u] \end{tikzcd} \qquad \begin{tikzcd} a \arrow[d] \arrow[r, bend left] & d \arrow[l] \\ b \arrow[r] & c \arrow[ul] \end{tikzcd}$$

These are schematic diagrams of proof outlines. As long as one can go from any proposition to any other proposition along a directed arrow, then a proof of a "the following are equivalent" statement is complete.

Suppose you knew that a, b, c and d were all equivalent. (An example is the Invertible Matrix Theorem from linear algebra.) Then, if you also knew b, you could conclude c. If you know d, you could conclude a. If you knew $\neg c$, you could conclude $\neg d$, and so on.

3.8.1 A characterization is not a substitute for a definition

Recall Definition 230, that an integer b is odd if there exists an integer s such that $b = 2k + 1$. As a nice review problem, prove the following theorem:

Theorem 331. *Let b be an integer. Then b is odd if and only if there exists an integer u such that $b = 2u-1$.*

Theorem 331 is known as a **characterization** of being odd. However, when asked to state the definition of odd, it would be incorrect to write "an integer b is odd if there exists an integer u such that $b = 2u - 1$" as the definition.

Why not? Well, this is really a form of circular thinking. The definition (the original definition) needs to be referenced in proving Theorem 331 in the first place. Even after Theorem 331 is proved, the text of the definition should remain as it was, and should not be changed just because we have now proved Theorem 331. This principle applies in general:

Warning 332

After a characterization is known (through proof, or accepted as true for free), do not replace the definition with text from the characterization theorem.

3.9 Uniqueness

Suppose U is a set and $P(x)$ is a predicate. There are times when we will need to prove "There exists a unique $b \in U$ such that $P(b)$." (Notice the inclusion of the new word "unique" here.) How does one prove **uniqueness**? Or more generally, how does one prove existence and uniqueness? The sentence in quotes is sometimes denoted by including an exclamation point as in: $\exists! b \in U\, [P(b)]$.

Example 333. *We revisit Example 135. There exists a unique resident x of Uruapan whose blood type is O+. (Namely, Finley is the only resident of Uruapan with blood type O+.)*

How one should prove this is informed by how we can rewrite $\exists! b \in U\, [P(b)]$ in terms of earlier symbols. In fact, $\exists! b \in U\, [P(b)]$ is logically equivalent to $[\exists b \in U\, [P(b)]] \wedge [\forall a \in U \forall c \in U[(P(a) \wedge P(c)) \rightarrow (a = c)]]$. The portion after the $\wedge$ symbol is the unique part. Thus,

Method 334

To prove that x satisfying $P(x)$ is unique, suppose a satisfies $P(a)$, and suppose c satisfies $P(c)$. Then prove $a = c$.

Example 335. *Suppose we need to prove "A mayor of Proofville is unique." A proof could start by writing the sentence "Let a be a mayor of Proofville." The following sentence could be "Let c be a mayor of Proofville." Then, the proof author should work towards proving $a = c$.*

Even in this example, it is ideal to just write "Let a be a mayor of Proofville Let c be a mayor of Proofville." in two consecutive sentences instead of trying to combine these into one sentence. The point is that after these two sentences are written, the proof writer is not saying that a and c are different, and the proof writer is also not saying that a and c are the same. Without assuming (even implicitly in how it is written) that $a = c$ or that $a \neq c$, it is much easier to then lead the reader to the eventual conclusion that $a = c$. (Imagine that it is more problematic to go towards the conclusion of $a = c$ if one has the erroneous sentence "Let a and c be different mayors of Proofville" at the beginning of the proof.)

Theorem 336. *There exists a unique real number b such that $2b + 7 = 1$.*

Proof. To prove existence, let $b = -3$. Then $2b+7 = 2(-3)+7 = -6+7 = 1$. To prove uniqueness, suppose that a satisfies $2a + 7 = 1$, and suppose that c satisfies $2c + 7 = 1$. (At this point in the proof, we are not saying that $a = c$. We are also not saying that $a \neq c$ for that matter.) Then, by transitivity of equality, $2a + 7 = 2c + 7$. By subtraction on both sides, $2a = 2c$. By division, $a = c$, so uniqueness is proved. □

By contrast, it would be impossible to prove that there exists a unique real number b such that $b^2-3b=40$. Existence is true (see Theorem 225), but uniqueness is not true. (This is because $a=-5$ is a solution to $x^2-3x=40$ and $c=8$ is also a solution to $x^2-3x=40$.)

Remark 337. *In English, "a" and "and" are indefinite articles, while "the" is the only definite article. An indefinite article is used when there is possibly uncertainly regarding how many objects there are, or if there is known to be only one, emphasizing that there's only one is unimportant. (Example: Here is an apple.) The word "the" is used the communicate that there's only one of a certain object. (Example: This is the [only] apple in the store.)*

Now that we have discussed uniqueness, challenge yourself to pay attention to how mathematicians speak and write when using articles. While the use of definite versus indefinite articles is often a matter of taste for most definitions, you will begin to notice that the word "the" only makes certain appearances (especially in certain theorems and definitions) only after a uniqueness statement is proved.

3.10 Connection to the past: examples from previous classes

Your previous math classes have offered small examples of many of the methods presented in the previous sections of this chapter:

- In algebra, you used the implication "If $a=b$ then $a+c=b+c$." Actually, a more complete version of this has quantifiers: you used "For all reals a, b, and c, if $a=b$, then $a+c=b+c$." This procedure was called "adding c to both sides" in your algebra class. In fact, when you were asked to solve the equation $x-5=4$, the idea that x was real was implied. To use the quantified statement (following Method 252, we'd consider $x-5$ to be the real a, have $b=4$, then use $c=5$. Then, since we have $a=b$ which is specifically $x-5=4$ for us, we could apply modus ponens (following Method 193) to get $a+c=b+c$ which in this example is $(x-5)+5=4+5$, and we are now well on our way to solving for x.

- Actually, a more complete statement is "For all reals a, b, and c, we have $a=b$ if and only if $a+c=b+c$." The biconditional shows that adding c to both sides is reversible. In contrast, "For all reals a and b, if $a=b$, then $a^2=b^2$" is true, but trying to change the implication to a biconditional will not work. (It is possible to have $a^2=b^2$ with $a\neq b$.) This is why, when solving equations such as $\sqrt{x-3}=10$ which requires squaring both sides, students are admonished to check their solutions, because not every "algebra move" they apply is reversible.

- Also from algebra, the statement "For all reals a and b, if $ab=0$, then $a=0$ or $b=0$" known as the Zero Product Property is used to solve an equation such as $x^2-x-6=0$, by first writing $(x-3)(x+2)=0$ and applying the statement using $a=x-3$ and $b=x+2$.

- As a general comment, nearly all properties from algebra such as $a(b+c)=ab+ac$ which are used frequently by an algebra student actually come with universal quantifiers.

- In precalculus, verifying a trigonometric identity often involves using other trigonometric identities. Some trig identities get *used* in the process of *proving* one new trig identity.

- In calculus, given a function f, students may be asked to prove that the function f is continuous at the x-value 6. This is typically done by verifying the three-part definition of continuity, an example of following Method 226.

- In calculus, students may be asked to show that a given function f has a limit of L as x approaches a given number c using the precise ε-δ definition of limit, which states the limit of the function f as x approaches c is L if for all $\varepsilon>0$, there exists $\delta>0$ such that if $0<|x-c|<\delta$, then $|f(x)-L|<\varepsilon$. In the process of verifying that the definition holds following Method 226, one ends up proving a universally-quantified statement, proving a (smaller) existentially-quantified statement, and (eventually) proving an implication.

- From calculus, the Squeeze Theorem states that if $\lim_{x\to c} f(x) = L$ and $\lim_{x\to c} h(x) = L$ and $f(x) \leq g(x) \leq h(x)$ for all x, then $\lim_{x\to c} g(x) = L$. To use the Squeeze Theorem, after verifying that the three requirements hold, the student applies modus ponens with the Squeeze Theorem to conclude a limit value for the function g.

- From calculus, the Intermediate Value Theorem states that if f is a function that is continuous on the interval $[a, b]$, then for all y between $f(a)$ and $f(b)$, there exists c in the interval $[a, b]$ such that $f(c) = y$. Students in calculus use this theorem by first verifying the given function f is continuous on $[a, b]$. By modus ponens, a universally-quantified statement is true. After selecting (writer's choice) any y-value between $f(a)$ and $f(b)$, this universally-quantified statement is used following Method 252.

- From calculus, Rolle's Theorem states that if f is continuous on the interval $[a, b]$ and differentiable on the interval (a, b) and $f(a) = f(b)$, then there exists c in the interval (a, b) such that $f'(c) = 0$. To use Rolle's Theorem, students in calculus verify that a given function f is continuous on $[a, b]$ and is differentiable on (a, b) and that $f(a)$ equals $f(b)$. Then modus ponens allows the student to conclude that there exists $c \in (a, b)$ such that $f'(c) = 0$. Using the Mean Value Theorem is similar.

- Typically in a second-semester calculus course, students are expected to apply convergence tests for infinite series. An example is the p-series Test, which states that (a) if $p > 1$, then the p-series $\sum_{n=1}^{\infty} \frac{1}{n^p}$ converges and (b) if $p \leq 1$ then the p-series $\sum_{n=1}^{\infty} \frac{1}{n^p}$ diverges. Like any series convergence test, students are expected to verify the requirements of the test. (These happen to be the hypotheses of an implication statement.) Then, via modus ponens, the student may conclude that a series diverges or converges. For example, when considering $\sum_{n=1}^{\infty} \frac{1}{n^5}$, since $p > 1$, modus ponens allows us to conclude (by the p-series Test) that $\sum_{n=1}^{\infty} \frac{1}{n^5}$ converges.

3.11 Some end-of-chapter notes about proofs

Warning 338

Don't just add in extra assumptions just because it's convenient. If you add in assumptions that are not allowed by the methods of proof, then you have proved something with an extra assumption. Assumptions can only be made when one of our methods of proof allow.

For example, when proving a statement that starts $\boxed{\text{For all } x \in A, \text{ there exists } y \in B}$ it is correct to assume $m \in A$, but it is tempting for people to then want to assume an arbitrary element $b \in B$, but this is not how a "there exists" statement is proved.

Always, always, always keep track of what you need to use and what you need to prove. Revisit the content of this chapter (the most important chapter!) frequently, and convince yourself all the methods of using and proving here align with what each proposition means, and if you are unconvinced, consult with your instructor as soon as possible.

There may be times when you have a fact you need to use, but you don't have the other accompanying fact to be able to use your proposition. (For example, you may know $p \to q$, but you don't know p yet.) Then, don't use that as the reason to try to prove $p \to q$. If you know you need to use $p \to q$, perhaps you need to work on some other aspect of the proof, and then you'll get p to pop up.

When proving a "there exists" statement, use the set that's involved to give you hints. Recalling Method 217, to prove $\exists x \in U\,[P(x)]$ you must start by defining a specific thing to take the place of x, which was named c in the Method. Now, coming up with what you should define c to be is challenging (and frustrating) at first. However, note that you later need to take what you defined c to be and then prove that c is in U. Use the fact that U is a certain kind of set to help give you a hint for how to define c. That is, if U is a set of rational numbers bigger than 100, then you know that you must define c to be a rational number bigger than 100. (Look at what has already been established in the problem; do you have a rational number bigger than 100 already established? If not, do you have a way to get a rational number bigger than 100? For example, if you had a rational number d that is bigger than 50, then multplying this number by

2 would give you a rational number bigger than 100. In this case, defining $c = 2d$ would allow you to get $c \in U$. This might not be the right thing to define c to be, as it depends on the problem, but you can use what the set U to help give you hints.) As another example, if U is a set of matrices, then you know that you must define c to be a matrix.

Where do we go from here? In the next chapter, you will learn new definitions. Every section will build upon the last section, but each new definition will go back to using an implication, proving an implication, using a "for all" statement, proving a "for all" statement, and so on.

As we transition from the elementary proofs of this chapter to the more involved proofs of the next chapter, it will become especially important to take careful note of what you need to use and what you need to prove. (Anything that you are allowed to assume are things that you will eventually need to use.) It is highly recommended that you take a minute at the beginning of every proof to write down exactly what is being asked of you in the proof you are trying to construct. There is no benefit to rushing through a proof. Take your item, move methodically, and be patient with yourself.

Exercise 339. *Student X has written* $\boxed{\text{Let } y \in Y \text{ such that } y^2 - y = 0 \text{ and } y = |y|.}$ *in their proof. Thoroughly explain why this type of sentence is problematic (unless it follows one specific format of sentence, which itself must be true.)*

Chapter 4

Sets

Sets were first introduced in Section 1.3.2. Recall that a set is a collection of objects, and that the objects that belong to a set are called its members or elements. We discussed some common sets (such as $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$, and $\mathbb{C}$) and have thus far only really used sets in stating quantified statements, proving quantified statements, and using quantified statements. Recall that we write $x \in A$ to mean that x is an element (or member) of the set A, and we write $x \notin A$ to indicate that x is not an element of the set A.

This chapter will introduce the most common notations used for sets, then discuss how to use and prove membership in a set, and consider other common mathematical objects made from sets and used throughout proof-based mathematics. We remind the reader that there are two primitive objects: sets and propositions. Sets are not propositions, and propositions are not sets. Informally, propositions are sentences which are true or false (but not both), while sets are "bags" which collectively hold some objects (called elements). Finally, recall Habit 18: when encountering a new noun, determine whether what is being defined is a proposition or a set. Every time you encounter a set, follow Habit 31 and determine if you have a set of numbers, or a set of people, or a set of ordered pairs, etc.

If p is a proposition, then p will be true or p will be false. If S is a set, it makes no sense to say S is true, and it makes no sense to say S is false. So, if A is a set, we cannot say something like "Therefore, A." This suggests that A has been proved to be true.

4.1 Set notations

There are three extremely common notations used for sets which involve curly braces. Upon seeing a set written using curly braces, it will be important to determine which of the three notations is being used.

4.1.1 Comma-separated format

The first of these three notations is the **comma-separated format**, a listing of elements of the set surrounded by a set of curly braces.

Example 340. *The set $\{2, 3, 5, 7\}$ is written in the comma-separated format. This set has exactly four elements: 2 is an element of the set, 3 is an element of the set, 5 is an element of the set, and 7 is an element of the set. No other object is an element of this set. If we assign a letter to this set by writing* $\boxed{\text{Let } P = \{2,3,5,7\}}$*, then $7 \in P$ while $8 \notin P$.*

Remark 341. *Note that $\{2, 3, 5, 7\}$ is a set whether it was named P or not. Writing P just becomes shorthand for writing $\{2, 3, 5, 7\}$. Thus, we could have written $7 \in \{2, 3, 5, 7\}$. As a general principle, objects are examples of the definitions introduced whether they are named using a new letter or not. The naming does not matter: what matters is if the object satisfies all the defining characteristics given in a definition. In addition, we typically use a capital letter when naming sets, but this is not required: if we write $d = \{2, 3, 5, 7\}$, then d is still a set.*

Example 342. *The set $L = \{1, 2, 3, \ldots, 100\}$ is written in the comma-separated format. The pattern dots are used to tell the reader to follow the pattern (as writing out the set in full would not be an effective use of paper). Based on the pattern, we surmise that $79 \in L$.*

Example 343. *The set $T = \{10, 20, 30, \ldots\}$ is written in the comma-separated format. This time, by having pattern dots (yet no number at the end like 100 in the previous example), we are to follow the writer's intent that this listing elements in the set T goes on forever. Assuming there is no intent to deceive us, $270 \in T$.*

Remark 344. *Repetition does not matter in the comma-separated format. So $\{1, 2, 3, 4, 5, 6\}$ is considered the same set as $\{1, 2, 3, 3, 4, 4, 5, 6\}$. We do not think of 3 belonging to a set "twice": 3 either belongs to a set, or does not belong to a set.*

4.1.2 Set builder with criterion format

While the comma-separated format is an extremely concrete way to describe a set in writing, it is inadequate to describe the majority of sets mathematicians need.

The second of the three notations for sets using curly braces is called the **set builder with criterion format** in this handbook. Suppose that a set has already been defined, and that $P(z)$ is a predicate whose universe of discourse is T. Then, a new set can be defined in the set builder with criterion format by writing $S = \{z \in T : P(z)\}$? This says "Let me introduce to you a set called S." The part before the colon is read *differently* than the part after:

- In this example, the part before the colon says $z \in T$. This says that *each* element of S comes from another set called T. The use of the variable z is a "placeholder variable". It is notation to help refer to a "typical element" of the set S in the same way that (x, y) represents a "typical point" on the line defined by $y = 3x + 5$. Since z is a placeholder variable, you can also write S as $\{m \in T : P(m)\}$ with no change in meaning.

- The part *after* the colon is read differently. The condition there must be satisfied for z to be an element of S. Put all together, S is the set consisting of elements from T, and to refer to a typical element, let's call it z. To be in S, not only should z be in T, but in addition, $P(z)$ must also be true. We think of $P(z)$ as a condition or a criterion which must be satisfied for z to belong to S.

Example 345. *Let $A = \{z \in \mathbb{R} : z - 3 \geq 0 \text{ and } z^2 < 19\}$. The portion before the colon says $\boxed{z \in \mathbb{R}}$ so in considering what elements belong to A, we will only consider elements in $\mathbb{R}$. Said differently, only elements of $\mathbb{R}$ belong to A. However, not all elements of $\mathbb{R}$ belong to A. Which elements $z \in \mathbb{R}$ are kept for A? Only those elements z which satisfy the predicate appearing after the colon, namely $\boxed{z - 3 \geq 0 \text{ and } z^2 < 19}$. For instance the real number 4 satisfies the property that $4 - 3 \geq 0$ and $4^2 < 19$. Therefore, $4 \in A$.*

Since $\pi \in \mathbb{R}$, and in addition, $z - 3 \geq 0$ and $z^2 < 19$ is true when substituting $z = \pi$, we also get that $\pi \in \mathbb{R}$. It turns out that the set A is really the interval $[3, \sqrt{19})$.

Example 346. *Though the set A in the previous example was defined by writing $A = \{z \in \mathbb{R} : z - 3 \geq 0 \text{ and } z^2 < 19\}$, we would describe exactly the same set by writing $A = \{m \in \mathbb{R} : m - 3 \geq 0 \text{ and } m^2 < 19\}$. The variable z really is a "placeholder" in the same way that in algebra, we read $f(x) = 3 + \sin(x)$ and $f(t) = 3 + \sin(t)$ in the same way.*

It may seem strange to bring up that you have the same set when replacing all of the zs with ms, but this is particularly useful if z has already been used in your proof somewhere. Suppose you were told $A = \{z \in \mathbb{R} : z - 3 \geq 0 \text{ and } z^2 < 19\}$. While it is not required to rewrite A, if z is already in use, you might find it useful to rewrite the definition of the set A by using another letter (such as m).

Example 347. *Let $B = \{c \in \mathbb{Z} : c > 100 \text{ or } 20 \text{ divides } c\}$. Due to the $c \in \mathbb{Z}$, the only things that belong to B are going to be integers. (However, not all integers belong to B.) Which integers are elements of B?*

Consider $c = 117$. Since $c \in \mathbb{Z}$ and since c satisfies the condition $\boxed{c > 100 \text{ or } 20 \text{ divides } c}$, we see that $177 \in B$.

Now, consider $c = 60$. Since $c \in \mathbb{Z}$ and since c satisfies the condition $\boxed{c > 100 \text{ or } 20 \text{ divides } c}$, we see that $60 \in B$.

Finally, consider $c = 25$. While $c \in \mathbb{Z}$, because c does not satisfy the condition $\boxed{c > 100 \text{ or } 20 \text{ divides } c}$, *we see that $25 \notin B$.*

Remark 348

Instead of using a colon, some authors will write sets in essentially this format replacing the colon with a vertical bar. Thus, $S = \{z \in T : P(z)\}$ is the same as $S = \{z \in T \mid P(z)\}$.

Example 349. *Some authors would write the previous example as $B = \{c \in \mathbb{Z} \mid c > 100 \text{ or } 20 \text{ divides } c\}$.*

Whether a the symbol used is a vertical bar or a colon, read this aloud as "such that."

Example 350. *The previous example could be read aloud, "We defined B to be the set consisting of all c in $\mathbb{Z}$ such that c is greater than 100 or 20 divides c."*

Let us consider a complete example:

Example 351. *Let $D = \{r \in \mathbb{R} \mid \sin(r) \geq 0\}$. Since $\frac{\pi}{2} \in \mathbb{R}$ and $\sin(\frac{\pi}{2}) \geq 0$, we see that $\frac{\pi}{2} \in D$. While $\frac{7\pi}{6} \in \mathbb{R}$, because $\sin(\frac{7\pi}{6}) < 0$, we see that $\frac{7\pi}{6} \notin D$. Finally, $\sqrt{-1} \notin D$, because the condition before the colon already fails. In short D, consists of those real numbers whose sine value is non-negative, and D consists* **only** *of these numbers.*

The same set could also have been written $D = \{r \in \mathbb{R} : \sin(r) \geq 0\}$ using a colon instead of a vertical bar. Either way, we could read aloud, "Let D be the set of all r in $\mathbb{R}$ such that sine of r is greater than or equal to 0."

Warning 352

The portion before the colon (or vertical bar) and the text appear after are standardized in this order. It is incorrect to swap the order of the two texts.

Using the previous example to illustrate, it is incorrect to write $D = \{\sin(r) \geq 0 \mid r \in \mathbb{R}\}$. It is similarly incorrect to write $D = \{\sin(r) \geq 0 : r \in \mathbb{R}\}$.

Definition 353 (Interval)**.** *Suppose that a and b are real numbers satisfying $a \leq b$. Then we define the* **intervals***:*

- $(a, b) = \{x \in \mathbb{R} : a < x \text{ and } x < b\}$
- $(a, b] = \{x \in \mathbb{R} : a < x \text{ and } x \leq b\}$
- $[a, b) = \{x \in \mathbb{R} : a \leq x \text{ and } x < b\}$
- $[a, b] = \{x \in \mathbb{R} : a \leq x \text{ and } x \leq b\}$.

One can similarly define (a, ∞) and $[a, \infty)$ and $(-\infty, b)$ and $(-\infty, b]$ and $(-\infty, \infty)$, the last of these which is the same as $\mathbb{R}$ itself.

4.1.3 Build running through set format

The third of the three notations for sets using curly braces is called the **build running through set format** in this handbook. Suppose that a set U has already been defined. Then a new set can be defined in build running through set format by writing $S = \{f(c) : c \in U\}$. As with the previous format, the part before the colon is read *differently* than the part after:

- This time, we first look to the text after the colon. In our example of the format, we have $\boxed{c \in U}$. Read this as an instruction to yourself, "As you run through each element of U (and to name a typical element in U, let's call it c)."
- Then look at the text before the colon, which in this case is $\boxed{f(c)}$. Read this as an instruction as well: "Compute what $f(c)$ is and throw that into the set S."

Example 354. *Let $J = \{5c + 2 : c \in \mathbb{Z}\}$. We should first look to the portion after the colon, which says $c \in \mathbb{Z}$. Then, every time you think of a number $c \in \mathbb{Z}$, compute $5c+2$ and the result of that computation is an element of J.*

For instance, due to the fact that $3 \in \mathbb{Z}$, we learn that $17 \in J$. Due to the fact that $10 \in \mathbb{Z}$, we learn that $52 \in J$. We call this the "build running through set format" because as you run through eat element in the set $\mathbb{Z}$, take that element c from $\mathbb{Z}$ and what you get for $5c + 2$ is an element of J.

Example 355. *The use of c is a placeholder variable. So, the previous set could also have been described by writing $J = \{5d + 2 : d \in \mathbb{Z}\}$.*

Example 356. *Let $K = \{m^2 : m \in \mathbb{R}\}$. As you pick any element in $\mathbb{R}$ and call it m, take the value of m^2 and make it a member of K. In other words, K consists of the squares of every real number.*

Due to the fact that $m = 5$ is a real number, we have learned that $25 \in K$. Due to the fact that $m = -5$ is a real number, we relearn that $25 \in K$. Due to the fact that $\sqrt{3}$ is a real number, we learn that $3 \in K$. Since $-10 \in \mathbb{R}$, we learn $100 \in K$.

It appears to be the case that K consists of all real numbers greater than or equal to zero.

Example 357. *Let $L = \{\sqrt{p} : p \in \mathbb{Z}\}$. Since $5 \in \mathbb{Z}$, we get $\sqrt{5} \in L$.*

To ensure that we have a complete understanding of this notation, let us pretend to run the following experiment. In a large auditorium, have peach person write an integer on an index card right. One person might write $\boxed{p = 365}$ while another person might write $\boxed{p = 12345}$. Then, take each index card, and when picking up the card that says $p = 365$, you discover that $\sqrt{365}$ is in L. When picking up the card that says $p = 12345$, you learn that $\sqrt{12345}$ is in L.

Through this thought experiment, we are discovering what numbers can conceivably belong to L. Of course, a room only holds a finite number of people (whereas there are an infinite number of integers), but through this experiment, you'll never conclude that $\sqrt{\pi}$ belong to L. In fact, $\sqrt{\pi} \notin L$.

If we use $p = -1$, then $p \in \mathbb{Z}$, so $\sqrt{-1} \in L$. In other words, $i \in L$.

This experiment may help with the naming of this format of set notation: as you run through each element p in the set $\mathbb{Z}$, we discover that $\sqrt{p}$ is an element of L.

Example 358. *Let $M = \{\cos(n\pi) : n \in \mathbb{Z}\}$. By picking $n = 0$, we learn that $1 \in M$. By picking $n = 1$, we learn that $-1 \in M$. By picking other integers to be n, we will discover no other elements belonging to M, other than the two we already discovered.*

Remark 359

Similar to Remark 348, some authors write a vertical bar instead of a colon. Thus, $S = \{f(c) : c \in U\}$ should be read the same as $S = \{f(c) \mid c \in U\}$

Example 360. *As a concrete example, the previous set can be written $M = \{\cos(n\pi) \mid n \in \mathbb{Z}\}$.*

The set in this example could be read aloud, "M is the set of all values of the form $\cos(n\pi)$ as n runs through all elements of $\mathbb{Z}$."

Warning 361

Similar to Warning 361, the portion before the colon (or vertical bar) and the text appear after are standardized in this order. It is incorrect to swap the order of the two texts, since the part before the colon (or vertical bar) is handled so differently from the portion after.

Using the previous example to illustrate, it is incorrect to write $M = \{n \in \mathbb{Z} \mid \cos(n\pi)\}$. It is similarly incorrect to write $M = \{n \in \mathbb{Z} : \cos(n\pi)\}$.

4.1.4 Important notes about the three set notations

When reading sets others have written, be aware that some authors use colons and some will use vertical bars. In fact, some authors use the colon notation in one portion of a paper or book and vertical bars in another portion. (The author of this handbook is guilty of this inconsistency.) We point out this matter to say that colons and vertical bars (at least in the context of set notation) are used interchangeably. If you are familiar with probability, we should remark that the vertical bar which may be used in set notation is *unrelated* to the notation $P(A|B)$ used for conditional probability.

Of the three formats, it is fairly easy to distinguish the first format from the last two – just look for the commas. The last two formats appear to be very similar, since both have either a colon or a vertical bar right in the middle, with text both before and after.

Method 362: How can I tell apart the two similar-looking set formats

Our first example in set builder with criterion format was $A = \{z \in \mathbb{R} : z - 3 \geq 0 \text{ and } z^2 < 19\}$. Our first example of build running through set format was $J = \{5c + 2 : c \in \mathbb{Z}\}$.

In the former, the portion before the colon is in the form "element in set" where this was the form of the portion after the colon in the latter.

While this is useful as evidence, the key to distinguishing these two formats is seen elsewhere: notice that the portion before the colon (or vertical bar) in defining the set J was $5c+2$. Notice that this is *not* a proposition/predicate: instead, this is an expression (in our example, a number). If the portion before the colon (or vertical bar) is an expression (whether that is a number, a matrix, a vector, etc.), this is the key sign that the set is written in build running through set format. Otherwise, the set is likely written in set builder with criterion format.

For proofs, it is more useful to have a set written in set builder with criterion format:

Method 363: Converting a set in build running through set format

Any set that is written in the build running through set format can be converted into the set builder with criterion format. Let us consider the set from Example 354, which was

$$J = \{5c + 2 : c \in \mathbb{Z}\}.$$

Choose a new variable to use – one that has not been mentioned within the definition of the set. For example, let us choose to use w. Then the variable in the portion after the colon (or vertical bar) must be quantified existentially. The portion before the colon is equated with the new variable. This one example is a perfect model to follow for all the rest:

$$J = \{w : \text{There exists } c \in \mathbb{Z} \text{ such that } w = 5c + 2\}.$$

Due to context, we can see that each time c is an integer, $w = 5c + 2$ is going to be an integer, so we don't change anything by writing

$$J = \{w \in \mathbb{Z} : \text{There exists } c \in \mathbb{Z} \text{ such that } w = 5c + 2\}.$$

Example 364. *From Example 356 recall we defined $K = \{m^2 : m \in \mathbb{R}\}$. Using the placeholder variable b, we can rewrite $K = \{b : \exists m \in \mathbb{R} \text{ such that } b = m^2\}$. Using the placeholder variable c, we can rewrite $K = \{c \mid \exists m \in \mathbb{R} \text{ such that } c = m^2\}$.*

Example 365. *Recall the set $L = \{\sqrt{p} : p \in \mathbb{Z}\}$ from Example . Using the placeholder variable u, we can rewrite $L = \{u : \exists p \in \mathbb{Z} \text{ such that } u = \sqrt{p}\}$. Using the placeholder variable m, we can rewrite $L = \{m \mid \exists p \in \mathbb{Z} \text{ such that } m = \sqrt{p}\}$.*

Example 366. *Recall the set $M = \{\cos(n\pi) \mid n \in \mathbb{Z}\}$ from Example . We can write $M = \{a : \exists n \in \mathbb{Z} \text{ such that } a = \cos(n\pi)\}$. We can write $M = \{b \mid \exists n \in \mathbb{Z} \text{ such that } b = \cos(n\pi)\}$. We can write $M = \{b \mid \exists k \in \mathbb{Z} \text{ such that } b = \cos(k\pi)\}$.*

If Method 363 can convert any set from build running through set format into set builder with criterion format, why even have the build running through set format? Some sets are only naturally described using the set builder with criterion format. This might further cause you to ask why we should have the build running through set format. It is a natural way to describe some sets, and it is easier to get a sense for what belongs to the set. For example, by writing

$$J = \{5c + 2 : c \in \mathbb{Z}\}.$$

if we were "running through" $\mathbb{Z}$ and chose $c = 100$, then we quickly see that $502 \in J$. Choose a different integer for c and you'll quickly see another element of the set J. While this format for a set is good for an intuitive sense of what belongs to the set, it's generally not a good format for proof.

For proof, we will see in the next section that the set builder with criterion format is preferable to the build running through set format. Fortunately, Method 363 makes it routine to convert when needed.

It is essential to understand how to read sets presented to you in all three formats. You will be expected to write sets, and based on the things which are elements of your set, you will need to pick the most appropriate format for writing your set, and follow the conventions of this section in writing your set using one of the standard three notations. Math books frequently employ all three formats. If a set is being described to you in one of these standard three formats and you haven't bothered to take the time to understand how to read each of the three formats, any proof involving sets (which is nearly all of them!) will be downright impossible. Finally, it is essential to understand how to read these standard set notations in order to prove that something (say x) belongs to a set, or to use the fact that x is an element of a set. This is the focus of the next section.

4.2 Rules of inference for set membership

Suppose S is a set, and that S is defined using one of the three set notations from the previous section. The three set notations lead to rules of inference for using $x \in S$ and for proving $x \in S$. The rules of inference are inextricably linked to *how* each of the set notations must be read and understood.

4.2.1 Comma-separated format

Due to how we must read a set written in comma-separated format such as $\{1, 3, 5\}$, we obtain methods for using an element belongs to such a set and proving an element belongs to such a set.

Method 367: Using $x \in S$ if S is in comma-separated format

Suppose $S = \{a, b, c, d, \ldots\}$. To use $x \in S$,

$$\begin{tikzcd} \boxed{x \in S} \arrow[d] \\ \boxed{(x = a) \vee (x = b) \vee (x = c) \vee (x = d) \vee \ldots} \end{tikzcd}$$

Example 368. *Let $S = \{2, 4, 6, 7, 13\}$. Suppose we were given the information that $x \in S$. Then we can conclude that $x = 2$ or $x = 4$ or $x = 6$ or $x = 7$ or $x = 13$.*

Example 369. *Let $T = \{2, 4, 6, 8, 10, 12, \ldots\}$. If we were told that $y \in T$, then we can conclude that $y = 2$ or $y = 4$ or $y = 6$ or so on. In other words, we could conclude that y is a positive even integer.*

Method 370: Proving $x \in S$ if S is in comma-separated format

Suppose $S = \{a, b, c, d, \ldots\}$. To prove $x \in S$,

$$\begin{tikzcd} \boxed{x = a} \arrow[d] & \boxed{x = b} \arrow[d] & \boxed{x = c} \arrow[d] & \boxed{x = d} \arrow[d] & \\ \boxed{x \in S} & \boxed{x \in S} & \boxed{x \in S} & \boxed{x \in S} & \ldots \end{tikzcd}$$

Example 371. *Let $S = \{2, 4, 6, 7, 13\}$. Suppose we were given the information that $a = 4$. Then we can conclude that $a \in S$. If we were given the information that $b = 7$, we can conclude $b \in S$.*

Example 372. *Let $T = \{2, 4, 6, 8, 10, 12, \dots\}$. If we were told that $y = 38$, then we can conclude that $y \in T$.*

4.2.2 Set builder with criterion format

Due to how we must read a set written in set builder with criterion such as $\{x \in \mathbb{Z} : \exists y \in \mathbb{Z} \text{ such that } x = y^2\}$, we obtain methods for using an element belongs to such a set and proving an element belongs to such a set.

Method 373: Using $x \in S$ if S is in set builder with criterion format

Suppose $S = \{z \in T : P(z)\}$, where $P(z)$ is a predicate. To use $x \in S$,

$$\begin{tikzcd} & \boxed{x \in S} \arrow[dl] \arrow[dr] & \\ \boxed{x \in T} & & \boxed{P(x)} \end{tikzcd}$$

Example 374. *Let $S = \{y \in B : P(y)\}$. Say we were told that $\boxed{x \in S}$ is true. Then we can conclude that $\boxed{x \in B}$ and $\boxed{P(x)}$ are true. If we were given information that $\boxed{y \in S}$ is true, then we conclude $\boxed{y \in B}$ and $\boxed{P(y)}$.*

Example 375. *Let $U = \{m \in C : P(m)\}$. Say we were told that $\boxed{a \in U}$ is true. Then we can conclude that $\boxed{a \in C}$ and $\boxed{P(a)}$ are true.*

Example 376. *Let $Y = \{m \in C : m \text{ bakes cookies}\}$. Say we were told that $\boxed{t \in Y}$ is true. Then we can conclude that $\boxed{t \in C}$ and $\boxed{t \text{ bakes cookies}}$ are true.*

Example 377. *Let $L = \{u \in B : u \text{ is a doctor}\}$. Say we were told that $\boxed{e \in L}$ is true. Then we can conclude that $\boxed{e \in B}$ and $\boxed{e \text{ is a doctor}}$ are true. On the other hand, if we knew that f is a doctor we could not conclude from this that $f \in B$. Similarly, if we knew $h \in B$, we could not conclude from this that h is a doctor.*

Example 378. *Let $S = \{z \in \mathbb{R} : z - 3 \geq 0 \text{ and } z^2 < 19\}$. Say we were told that $\boxed{x \in S}$ is true. Then we can conclude that $\boxed{x \in \mathbb{R}}$ and $\boxed{x - 3 \geq 0 \text{ and } x^2 < 19}$ are true. If we were given information that $\boxed{y \in S}$ is true, then we conclude $\boxed{y \in \mathbb{R}}$ and $\boxed{y - 3 \geq 0 \text{ and } y^2 < 19}$.*

Example 379. *Let $B = \{c \in \mathbb{Z} : c > 100 \text{ or } 20 \text{ divides } c\}$. Upon given the information that $\boxed{a \in B}$ is true, we conclude $\boxed{a \in \mathbb{Z}}$ is true. We also conclude $\boxed{a > 100 \text{ or } 20 \text{ divides } a}$ is true.*

Example 380. *Let $D = \{r \in \mathbb{R} \mid \sin(r) \geq 0\}$. If we learn $\boxed{s \in D}$ is true, we can conclude $\boxed{s \in \mathbb{R}}$ and $\boxed{\sin(s) \geq 0}$ are true.*

Method 381: Proving $x \in S$ if S is in set builder with criterion format

Suppose $S = \{z \in T : P(z)\}$, where $P(z)$ is a predicate. To prove $x \in S$,

$$\begin{tikzcd} \boxed{x \in T} \arrow[dr] & & \boxed{P(x)} \arrow[dl] \\ & \boxed{x \in S} & \end{tikzcd}$$

Example 382. *Let $S = \{y \in B : P(y)\}$. Say we were told* $\boxed{x \in B}$ and $\boxed{P(x)}$ *are true. Then we could conclude* $\boxed{x \in S}$ *is true. If we were given information that* $\boxed{y \in B}$ and $\boxed{P(y)}$ *are true, we could conclude* $\boxed{y \in S}$ *is true.*

Example 383. *Let $U = \{m \in C : P(m)\}$. Say we were told that* $\boxed{a \in C}$ and $\boxed{P(a)}$ *are true. Then we can conclude that* $\boxed{a \in U}$ *is true.*

Example 384. *Let $Y = \{m \in C : m \text{ bakes cookies}\}$. Say we were told that* $\boxed{t \in C}$ *and* $\boxed{t \text{ bakes cookies}}$ *are true. Then we can conclude that* $\boxed{t \in Y}$ *is true.*

Example 385. *Let $L = \{u \in B : u \text{ is a doctor}\}$. Say we were told that* $\boxed{e \in B}$ and $\boxed{e \text{ is a doctor}}$ *are true. Then we can conclude that* $\boxed{e \in L}$ *is true. On the other hand, if we knew that f is a doctor we could not conclude from this that $f \in B$. Similarly, if we knew $h \in B$, we could not conclude from this that h is a doctor.*

Example 386. *Let $S = \{z \in \mathbb{R} : z - 3 \geq 0 \text{ and } z^2 < 19\}$. Suppose we know* $\boxed{x \in \mathbb{R}}$ *is true and also know* $\boxed{x - 3 \geq 0 \text{ and } x^2 < 19}$ *is true. Then we could conclude* $\boxed{x \in S}$ *is true. If we proved* $\boxed{y \in \mathbb{R}}$ *and also proved that* $\boxed{y - 3 \geq 0 \text{ and } y^2 < 19}$ *is true, then we could use these two new facts to conclude that* $\boxed{y \in S}$ *is true.*

Example 387. *Let $B = \{c \in \mathbb{Z} : c > 100 \text{ or } 20 \text{ divides } c\}$. Upon given the information* $\boxed{a \in \mathbb{Z}}$ *along with the information that* $\boxed{a > 100 \text{ or } 20 \text{ divides } a}$, *we could conclude that* $\boxed{a \in B}$ *is true.*

Example 388. *Let $D = \{r \in \mathbb{R} \mid \sin(r) \geq 0\}$. If we learn or prove that* $\boxed{m \in \mathbb{R}}$ *and also prove that* $\boxed{\sin(m) \geq 0}$, *then we can use this information to conclude that* $\boxed{m \in D}$ *is true.*

4.2.3 Build running through set format

Instead of being burdened with another "using $x \in S$" flowchart and another "proving $x \in S$" flowchart, you may prefer ignoring the new flowcharts introduced here and instead converting the "gather using running set" set notation $S = \{f(c) : c \in R\}$ into the "set-builder with criterion" set notation $S = \{z : \exists c \in R \text{ s.t. } z = f(c)\}$.

Method 389: Using $x \in S$ if S is in build running through set format

If $S = \{f(c) : c \in R\}$, first rewrite S as $S = \{z : \exists c \in R \text{ s.t. } z = f(c)\}$. To use $x \in S$,

$$\begin{tikzcd} \boxed{x \in S} \arrow[d] \\ \boxed{\exists c \in R \text{ such that } x = f(c)} \end{tikzcd}$$

Example 390. *Let $J = \{5c + 2 : c \in \mathbb{Z}\}$. We may directly use the method just mentioned, but it is nice to rewrite the set J by writing $J = \{v :$ there exists $c \in \mathbb{Z}$ such that $v = 5c + 2\}$. Then if we knew (or were told) that $r \in J$, then we could conclude that there exists $c \in \mathbb{Z}$ such that $m = 5c + 2$.*

Example 391. *From Example 356 recall we defined $K = \{m^2 : m \in \mathbb{R}\}$ which we could rewrite $K = \{b : \exists m \in \mathbb{R}$ such that $b = m^2\}$. If we knew that $\boxed{y \in K}$ we could conclude $\boxed{\text{there exists } m \in \mathbb{R} \text{ such that } y = m^2}$ is true.*

Example 392. *Recall the set $L = \{\sqrt{p} : p \in \mathbb{Z}\}$ from Example could be rewritten $L = \{u : \exists p \in \mathbb{Z}$ such that $u = \sqrt{p}\}$. If we knew that $\boxed{c \in L}$ we could conclude $\boxed{\text{there exists } p \in \mathbb{Z} \text{ such that } c = \sqrt{p}}$.*

Example 393. *Recall the set $M = \{\cos(n\pi) \mid n \in \mathbb{Z}\}$ from Example could be rewritten $M = \{a : \exists n \in \mathbb{Z}$ such that $a = \cos(n\pi)\}$. Say we knew $\boxed{h \in M}$ is true. Then we could conclude $\boxed{\text{there exists } n \in \mathbb{Z} \text{ such that } h = \cos(n\pi)}$ is true.*

Method 394: Proving $x \in S$ if S is in build running through set format

If $S = \{f(c) : c \in R\}$, first rewrite S as $S = \{z : \exists c \in R \text{ s.t. } z = f(c)\}$. To prove $x \in S$,

$$\begin{array}{c} \boxed{\exists c \in R \text{ such that } x = f(c)} \\ \downarrow \\ \boxed{x \in S} \end{array}$$

Example 395. *Let $J = \{5c + 2 : c \in \mathbb{Z}\}$. We may directly use the method just mentioned, but it is nice to rewrite the set J by writing $J = \{v :$ there exists $c \in \mathbb{Z}$ such that $v = 5c + 2\}$. Then if we knew (or were told) that $\boxed{\text{there exists } c \in \mathbb{Z} \text{ such that } m = 5c + 2}$, we could then conclude that $\boxed{m \in J}$.*

Example 396. *From Example 356 recall we defined $K = \{m^2 : m \in \mathbb{R}\}$ which we could rewrite $K = \{b : \exists m \in \mathbb{R}$ such that $b = m^2\}$. If we knew that $\boxed{\text{there exists } m \in \mathbb{R} \text{ such that } y = m^2}$ we could conclude $\boxed{y \in K}$.*

Example 397. *Recall the set $L = \{\sqrt{p} : p \in \mathbb{Z}\}$ from Example could be rewritten $L = \{u : \exists p \in \mathbb{Z}$ such that $u = \sqrt{p}\}$. If we knew that $\boxed{\text{there exists } p \in \mathbb{Z} \text{ such that } c = \sqrt{p}}$ we could conclude $\boxed{c \in L}$.*

Example 398. *Recall the set $M = \{\cos(n\pi) \mid n \in \mathbb{Z}\}$ from Example could be rewritten $M = \{a : \exists n \in \mathbb{Z}$ such that $a = \cos(n\pi)\}$. If we knew that $\boxed{\text{there exists } n \in \mathbb{Z} \text{ such that } h = \cos(n\pi)}$ we could conclude $\boxed{h \in M}$.*

Example 399. *Let $R = \{\sin(x) : x \in \mathbb{Q}\}$. We may directly use the method just mentioned, but it is nice to write $R = \{c :$ there exists $x \in \mathbb{Q}$ such that $c = \sin(x)\}$. Then if we knew (or were told) that there exists $x \in \mathbb{Q}$ such that $c = \sin(x)$, we could then conclude that $c \in R$.*

4.2.4 Summary

Later sections of the book will refer to these two methods (which in turn refer back to previous methods in this chapter).

Method 400: Using the fact that an object is an element of a set

Determine which format of set notation is being used. (Method 362 may help.) If the set is written in comma-separated format, refer to Method 367. If the set is written in set builder with criterion format, refer to Method 373. If the set is written in build running through set format, rewrite the set in set builder with criterion format using Method 363, or for a direct method, refer to Method 389.

Method 401: Proving that an object is an element of a set

Determine which format of set notation is being used. (Method 362 may help.) If the set is written in comma-separated format, refer to Method 370. If the set is written in set builder with criterion format, refer to Method 381. If the set is written in build running through set format, rewrite the set in set builder with criterion format using Method 363, or for a direct method, refer to Method 394.

Example 402. *Let* $Z = \{x \in G \mid \forall g \in G, gx = xg\}$*. Notice that this set is in set builder with criterion format.*

Suppose you needed to prove that $g \in Z$*. Unfortunately, the definition of* Z *already has a* g *written, but this is just a notation clash. Think back to the confusion of a struggling algebra student in trying to evaluate* $f(x^2 + 3x)$ *if* f *was defined by* $f(x) = \sin(\sqrt{x} + 3)$*. In this situation, we would encourage the confused student to write* $f(y) = \sin(\sqrt{y} + 3)$*.*

In the same way, noting that the universally-quantified g *in the definition of* Z *is a placeholder variable, let us rewrite* Z *by writing* $Z = \{x \in G \mid \forall y \in G, yx = xy\}$*. Then, it becomes clearer to us what we have to do to prove that* $g \in Z$*.*

Noting that g *would take the place of* x*, we need to a post-substituted version of the statements before and after the vertical bar. Namely, if we prove that* $g \in G$ *and also prove that for all* $y \in G$*, we have* $yg = gy$*, then we can conclude that* $g \in G$*.*

4.2.5 Assorted examples and exercises for set membership rules

Given $S = \{z \in T : P(z)\}$, Method 373 says knowing $x \in S$ leads to knowing both $x \in T$ and $P(x)$, while Method 381 says knowing both $x \in T$ and $P(x)$ lets us conclude $x \in S$. These are the only rules of inference. People who are new to proof may be tempted to make the following errors:

Warning 403

Suppose the set S is defined by $S = \{z \in T : P(z)\}$. Then knowing $\boxed{z \in T}$ is true does not allow you to conclude that $\boxed{P(z)}$ is true. Similarly, knowing $\boxed{y \in T}$ does not mean that we know anything about whether $\boxed{P(y)}$ is true.

Warning 404

Suppose the set S is defined by $S = \{z \in T : P(z)\}$. Then knowing $\boxed{P(z)}$ is true does not allow you to conclude that $\boxed{z \in T}$ is true. Similarly, knowing $\boxed{P(y)}$ does not mean that we know anything about whether $\boxed{y \in T}$ is true.

Warning 405

Method 363 is a procedure that starts with a set in build running through set format and rewrites the set in set builder with criterion format. Do not apply the method to a set that is already in set builder with criterion format.

Example 406. *Consider the set $S = \{x \in \mathbb{Z} : x = x^2\}$. Note that the only integers which satisfy the property that they are equal to their square happen to be $x = 1$ or $x = 0$. More importantly, note that $x \in \mathbb{Z}$ does not imply that $x = x^2$.*

Example 407. *Let's say that someone defined the set S for us by writing $S = \{z \in L : z \text{ has cool sneakers}\}$. Note that S is written in set builder with criterion format.*

- *Say we knew $\boxed{x \in S}$ is true. Following Method 373, we can conclude $\boxed{x \in L}$ and $\boxed{x \text{ has cool sneakers}}$ are both true.*
- *Say we knew $\boxed{c \in S}$ is true. Following Method 373, we can conclude $\boxed{c \in L}$ and $\boxed{c \text{ has cool sneakers}}$ are both true.*
- *Say we knew $\boxed{x \in L}$ and $\boxed{x \text{ has cool sneakers}}$ are both true. Following Method 381, we can conclude that $\boxed{x \in S}$ is true.*
- *Say we knew $\boxed{h \in L}$ and $\boxed{h \text{ has cool sneakers}}$ are both true. Following Method 381, we can conclude that $\boxed{h \in S}$ is true.*
- *If we know $\boxed{m \in L}$ is true, it is an error to conclude $\boxed{m \text{ has cool sneakers}}$ which is addressed in Warning 403.*
- *If we know $\boxed{j \text{ has cool sneakers}}$ is true, it is an error to conclude $\boxed{j \in L}$ which is addressed in Warning 404.*

Example 408. *Let $B = \{x \in A : \text{Sam writes a check to } x\}$, a set written in set builder with criterion format.*

- *Say we knew $\boxed{m \in B}$ is true. By Method 373, we conclude $\boxed{m \in A}$ and $\boxed{\text{Sam writes a check to } m}$ are true.*
- *Say we knew $\boxed{g \in A}$ and $\boxed{\text{Sam writes a check to } g}$ are true. Following Method 381, we conclude $\boxed{g \in B}$ is true.*
- *If we know $\boxed{r \in A}$ is true, it is an error to conclude $\boxed{\text{Sam writes a check to } r}$ which is addressed in Warning 403.*
- *If we know $\boxed{\text{Sam writes a check to } k}$ is true, it is an error to conclude $\boxed{k \in A}$ which is addressed in Warning 404.*

Warning 409

Say you are given the set $S = \{z \in L : z \text{ has cool sneakers}\}$. Say that you know $u \in S$. Then it is not truly advancing your proof forward to write

$$S = \{u \in L : u \text{ has cool sneakers}\}.$$

This is simply rewriting the definition of S using a new placeholder variable, and runs close to the danger of accidentally "redefining" S (note: nothing should ever be redefined). Instead of doing this kind of symbol manipulation that does not advance your proof forward, you should look at

$$S = \{z \in L : z \text{ has cool sneakers}\}$$

and based on $u \in S$, you should conclude two things:

- $u \in L$.
- u has cool sneakers.

Example 410. *You are given the set $U = \{y \in \mathbb{Z} : y^2 < 15\}$. If you are told that $x \in U$, then you can conclude $x \in \mathbb{Z}$ and can also conclude $x^2 < 15$. Warning 409 tells us it's pointless (and borderline dangerous) to write $U = \{x \in \mathbb{Z} : x^2 < 15\}$. If we know that $a \in \mathbb{Z}$, we cannot conclude $a^2 < 15$. If we know $b^2 < 15$, we cannot conclude $b \in \mathbb{Z}$. If we know $c \in \mathbb{Z}$ and we know $c^2 < 15$, then we can conclude $c \in U$.*

Example 411. *You are given the set $U = \{y^2 + 15 : y \in \mathbb{Z}\}$. This is in build running through set format, so we rewrite the set as $U = \{h : \exists y \in \mathbb{Z}$ such that $h = y^2 + 15\}$. If we are told that $s \in U$, it is pointless (and dangerous) to write $U = \{s : \exists y \in \mathbb{Z}$ such that $s = y^2 + 15\}$. Instead, we simply conclude that there exists $y \in \mathbb{Z}$ such that $s = y^2 + 15$. If we know that there exists $t \in \mathbb{Z}$ such that $c = t^2 + 15$, this is the same as knowing that there exists $y \in \mathbb{Z}$ such that $c = y^2 + 15$, so we can conclude that $c \in U$.*

Example 412. *Say that $L = \{a^2 + b^2 : a \in \mathbb{Z}, b \in \mathbb{Z}\}$. Then L is built running through two sets. If we know that $h \in L$, we can conclude that there exists $a \in \mathbb{Z}$ and there exists $b \in \mathbb{Z}$ such that $h = a^2 + b^2$.*

Exercise 413. *If the set Q is defined to be $Q = \{x \in X : d(x, a) < r\}$ what can you conclude if you know $\boxed{w \in Q}$?*

Solution to exercise. Since $w \in Q$, we can conclude $w \in X$ and we can also conclude $d(w, a) < r$. ☐

Exercise 414. *If the set Q is defined to be $Q = \{x \in X : d(x, a) < r\}$ what do you have to do to prove $\boxed{c \in Q}$?*

Solution to exercise. To prove $c \in Q$, we must prove $c \in X$ and we must also prove $d(c, a) < r$. ☐

Exercise 415. *If the set E is defined to be $E = \{g \cdot x \mid g \in G\}$ what can you conclude if you know $\boxed{j \in E}$?*

Solution to exercise. First, we rewrite the set E in set builder with criterion format. So,

$$E = \{s \mid \exists g \in G \text{ such that } s = g \cdot x\}.$$

Note that s is placeholder variable, and all of the s's above could have been t's (but not g's or x's).

Since $j \in E$, we get to conclude $\exists g \in G$ such that $j = g \cdot x$. ☐

Exercise 416. *If the set E is defined to be $E = \{g \cdot x \mid g \in G\}$ what do you have to do to prove $\boxed{k \in E}$?*

Solution to exercise. First, we rewrite the set E as

$$E = \{t \mid \exists g \in G \text{ such that } t = g \cdot x\}.$$

Note that t is just a placeholder variable, and all of the t's above could have been u's (but not g's or x's).

To prove $k \in E$, we have to first prove $\exists g \in G$ such that $k = g \cdot x$. ☐

Exercises from actual mathematical literature

Consider the following sets:

- $A = \{a \in G \mid ax = xa \text{ for all } x \in G\} = \{a \in G \mid \text{for all } x \in G,\ ax = xa\}$ from pg. 66 of Joseph Gallian. *Abstract Algebra.* (8th ed.)
- $B = \{z \in (\mathbb{C}^*)^n : f(z) = 0 \text{ for all } f \in I\}$ from pg. 17 of Diana Maclagan, Bernd Sturmfels. *Introduction to Tropical Geometry.*
- $C = \{f(x) : x \in [x_{i-1}, x_i]\}$ from pg. 276 of Matthew A. Pons. *Real Analysis for the Undergraduate.*
- $D = \{x \mid x \in A \text{ or } x \in B\}$ from pg. 5 of James R. Munkres. *Topology.*
- $E = \{g \cdot x : g \in G\}$ from pg. 116 of Joseph Rotman. *Galois Theory.*
- $F = \{T(v) : v \in V\}$ from pg. 43 of Sheldon Axler. *Linear Algebra Done Right.* (2nd ed.)
- $G = \{4, 3, 2\}$ from pg. 217 of Branko Grünbaum. *Configurations of Points and Lines.*

- $H = \{x \in P : cx = c_0\}$ from pg. 130 of Günter M. Ziegler. *Lectures on Polytopes.*
- $I = \{y_F + \lambda(x - y_F) : x \in G, \lambda \geq 0\}$ from pg. 32 of Rekha R. Thomas. *Lectures in Geometric Combinatorics.*
- $J = \{p, \rho(p), \rho^2(p), \ldots, \rho^{n-1}(p)\}$ from pg. 39 of Ronald Solomon. *Abstract Algebra.*
- $K = \{B \in W^P \mid B \neq A, \mu^{-1}[A] \text{ and } \mu^{-1}[B] \text{ are adjacent}\}$ from pg. 168 of Alexandre V. Borovik, I. M. Gel'fand, and Neil White. *Coxeter Matroids.*
- $L = \{(0, y) \mid y \in Y\}$ from pg. 61 of Gerald Teschl. *Topics in Real and Functional Analysis.*
- $M = \{r \in \mathbb{R}^n : x + \lambda r \in P \text{ for all } x \in P \text{ and } \lambda \in \mathbb{R}_{\geq 0}\}$ from pg. 97 of Michele Conforti, Gérard Cornuéjols, Giacomo Zambelli. *Integer Programming.*
- $N = \{\mu_1 b_1 + \mu_2 b_2 : (\mu_1, \mu_2) \in \mathbb{Z}^2_+\}$ from pg. 110 of Jesús A. De Loera, Raymond Hemmecke, Matthias Köppe. *Algebraic and Geometric Ideas in the Theory of Discrete Optimization.*
- $O = \{(a_1, \ldots, a_n) \in k^n \mid f(a_1, \ldots, a_n) = 0 \text{ for all } f \in S\}$ from pg. 2 of William W. Adams, Philippe Loustaunau. *An Introduction to Gröbner Bases.*
- $P = \{f \in V^* : |f(x)| \leq 1 \text{ for all } x \in U\}$ from pg. 116 of Alexander Barvinok. *A Course in Convexity.*
- $Q = \{x \in X : d(x, a) < r\}$ from pg. 16 of Joseph Muscat. *Functional Analysis.*
- $R = \{x \mid 0 \leq T(x) \leq s\}$ from pg. 77 of Manfred Einsiedler, Thomas Ward. *Ergodic Theory with a view towards Number Theory.*
- $S = \{r \in \mathbb{Q} : \text{for some } a \in A \text{ and } c \in C,\ r = a + c\}$ from pg. 14 of Charles Chapman Pugh. *Real Mathematical Analysis.*
- $T = \{x : \bar{A}x = 0\}$ from Éva Tardos, A Strongly Polynomial Algorithm to Solve Combinatorial Linear Programs, *Operations Research* 34(2):250–256.
- $U = \{i_0 + i_1\omega + i_2\omega^2 + i_3\omega^3 : i_o, i_1, i_2, i_3 \in \mathbb{Z}, |i_j| \leq m\}$ from pg. 51 of Jiři Matoušek. *Lectures on Discrete Geometry.*

For each set defined above,

- Determine if the notation is comma-separated format, set builder notation with criterion, or set builder format with a running set used to "gather" elements.
- In the case of the running set, rewrite as set builder notation with criterion.
- Write out a flowchart of *using* x in the set. Write a flowchart for *proving* x in the set. (Because of the converting of "running set" notation, you should *never* use the last row of the reference.) What about using g in the set? Proving g in the set? (Try *any* letter – not just x or g.)

4.3 Properties of sets

Definition 417: Subset

A set S is a **subset** of the set T if every element of S is an element of T. In other words, S is a **subset** of T if the implication "if $x \in S$, then $x \in T$" is true. We write $S \subseteq T$ to denote S is a subset of T.

Example 418. *Let $S = \{3, 4\}$ and $T = \{3, 4, 5, 6\}$. Then S is a subset of T since every element of S is an element of T. Note that T is not a subset of S.*

Example 419. *Let $C = \{3, 4, 5\}$ and $D = \{3, 4, 5\}$. Then $C \subseteq D$ since every element of C is an element of D. Also, $D \subseteq C$.*

Example 420. *Recall $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$, and $\mathbb{C}$ are the sets of integers, rationals, reals, and complexes, respectively. Then $\mathbb{Z} \subseteq \mathbb{Q}$ and $\mathbb{Q} \subseteq \mathbb{R}$ and $\mathbb{R} \subseteq \mathbb{C}$.*

Exercise 421. *In Exercise 97, you were asked to consider what are things that must be addressed in writing a definition of disjunction. Similarly, what are things that you must address when writing a definition for subset? What things* don't *matter?*

Method 422: How to prove a set A is subset of the set B

Let A and B be sets. Informed by Definition 417, to prove that A is a subset of B, we should show that every element of the set A is an element of the set B. To do this, write $\boxed{\text{Let } x \in A \text{ be arbitrary.}}$ Then prove that $x \in B$. Said differently, since $A \subseteq B$ is logically equivalent to $\boxed{s \in A \text{ implies } s \in B}$ you should assume $s \in A$ and then prove $s \in B$.

Once you say "Let $x \in A$ be arbitrary" you will need to use the fact that x is an element of A, following Method 400. You will likely need to use the facts obtained through that method to prove that $x \in B$, following Method 401.

If x has already been used in your proof, it is good practice to select a different variable. Perhaps you can let $\vartheta \in A$, and then prove that $\vartheta \in B$.

As an example, let us prove the following:

Theorem 423. *Let A be the set of all multiples of 10. Let B be the set of all multiples of 5. Then $A \subseteq B$.*

Proof. Let A be the set of all multiples of 10. Let B be the set of all multiples of 5. Let $c \in A$. Then c is a multiple of 10. In other words, 10 divides c. Thus, there exists an integer k such that $10k = c$. Rewriting the left side, we have $5 \cdot 2 \cdot k = c$. Let $r = 2k$. Then $5r = 5(2k) = (5 \cdot 2)k = 10k = c$. Since $5r = c$ and r is an integer, 5 divides c, which proves that c is a multiple of 5. Therefore $c \in B$.

We proved that $c \in A$ implies that $c \in B$, so A is a subset of B. □

Other facts follow from the subset inclusions mentioned in Example 420, such as $\mathbb{Z} \subseteq \mathbb{R}$. In fact, you should prove:

Theorem 424. *Let A, B and C be sets. If $A \subseteq B$ and $B \subseteq C$, then $A \subseteq C$.*

Exercise 425. *Prove: for all sets S, one has $S \subseteq S$.*

Method 426: How to use a set A is subset of the set B

Let A and B be sets. Informed by Definition 417, to prove that A is a subset of B, since $A \subseteq B$ is logically equivalent to $x \in A$ implies $x \in B$, as soon as you know $m \in A$, then you can conclude $m \in B$. If you do not know of a specific element of A, then you cannot use $A \subseteq B$.

Warning 427

It is an error to take the fact $A \subseteq B$ and then write $\boxed{x \in A \text{ and } x \in B}$. This is turning a conjunction into an implication.

Warning 428: Membership versus subset

Recall that $x \in B$ means that x is a member of the set B, and that $x \subseteq B$ means that x is a subset of the set B. While this may look strange, it is only because we generally stick to the convention that sets are named with capital letters, but there are times that this convention must be suspended. Aside from how these two look, $x \in B$ and $x \subseteq B$ really say different things, and one should not be written for the other. For instance, if $B = \{7, 9\}$, then we can write $7 \in B$. We cannot write $7 \subseteq B$, because 7 is not a set.

Warning 429: Do not "assume A"

When tasked with proving that the set A is a subset of the set B, it is tempting to say that you need to "assume A" and then "prove B." However, this does not make sense. Saying "assume A" makes it sound like A is either true or false. However, we should recall Warning 17: a set is not a proposition, so a set is neither true nor false. The same comment applies to saying "prove B." A set cannot be proved: a set is a collection of objects, and is neither true nor false. If A and B are both sets, then the same error of thought occurs in saying "if A is true, then B is true."

Warning 430

When proving that A is a subset of B, the second sentence in Definition 417 shows we are proving an implication. Therefore, near the end of the proof, it is incorrect to say "Since x is in A and x is in B, we conclude A is a subset of B."

Recall from Warning 117 that $p \wedge q$ does not have the same meaning as $p \to q$. The correct way to conclude would be to write "Since we proved that $x \in A$ implies that $x \in B$, we conclude A is a subset of B."

In our concluding sentence in the proof of Theorem 423, we said $c \in A$ implies $c \in B$. We did *not* say, for example, $c \in A$ and $c \in B$, since an implication is not the same as a conjunction. Similarly, we didn't say that A is true, because a set is neither true nor false.

Definition 431: Set equality

Two sets A and B are equal if each set is a subset of the other. In this case, we write $A = B$. In other words, if A and B are sets, we say that A is equal to B and write $A = B$ if $A \subseteq B$ and $B \subseteq A$.

Method 432: How to prove that two sets are equal

To prove that the set A is equal to the set B, informed by Definition 431, you must prove that A is a subset of B and that B is a subset of A. In light of Method 422, you should take an arbitrary element $c \in A$ and prove $c \in B$. Then take an arbitrary element $d \in B$ and then prove $d \in A$.

As an example, let us look at proving the following theorem:

Theorem 433. *Let $X = \{c \in Y : c \text{ plays tennis}\}$ and let $Y = \{r \in L \mid r \text{ plays volleyball}\}$. If*

- *for all $z \in Y$, if z does not play tennis, then z is not an interior designer,*
- *for all $p \in L$, if p plays volleyball, then p is an interior designer; and*

then $X = Y$.

We will annotate this proof with many references. (The references should not really be included in the proof, but are there for your convenience.) While the statement to prove is $X = Y$, the set up of the theorem comes with two definitions (the definition of the set X and the definition of the set Y) alongside two hypotheses (provided in the bullet list).

Proof. To prove that $X = Y$, following Method 432 we need to prove $X \subseteq Y$ and also prove $Y \subseteq X$.

First, we will prove that $X \subseteq Y$. To prove this, following Method 422, let $m \in X$ be arbitrary. Following Method 400, we are directed to Method 373, based on format/notation of the definition of X. Thus, $m \in Y$ and m plays tennis. Specifically, following Method 187, $m \in Y$. Since we have proved that $m \in X$ implies $m \in Y$, this concludes the proof that $X \subseteq Y$.

We now will prove that $Y \subseteq X$. To prove this, following Method 422, let $u \in Y$ be arbitrary. Following Method 400, we are directed to Method 373, since Y is also written in set builder with criterion format. From this, we conclude $u \in L$. We also conclude u plays volleyball. Since $u \in L$, this fact together with the second hypothesis (following Method 252), we conclude that the implication "if u plays volleyball, then u is an interior designer" is true. From this implication and the earlier fact that u plays volleyball, by Method 193, we conclude that u is an interior designer.

Since $u \in Y$, combined with the first hypothesis, following Method 252, we get the implication "if u does not play tennis, then u is not an interior designer." But earlier, we already learned that u is an interior designer, so combined with the new implication, following Method 196, we get that u plays tennis. Since $u \in Y$ and u plays tennis, following Method 381, we conclude that $u \in X$. We have now proved that $u \in Y$ implies $u \in X$, so our proof that $Y \subseteq X$ is complete.

To conclude, since we have proved $X \subseteq Y$ and also proved $Y \subseteq X$, we conclude that $X = Y$. □

Definition 434

A set S is a **proper subset** of the set T if S is a subset of T and $S \neq T$. We write $S \subsetneq T$ to denote that S is a proper subset of T.

Example 435. *Let $A = \{2, 3, 4\}$ and $B = \{2, 3, 4, 5, 6\}$ and $C = \{2, 3, 4\}$. Then A is a proper subset of B. While A is not a proper subset of C, though A is a subset of C.*

Example 436. *The set $\mathbb{Z}$ is a proper subset of $\mathbb{Q}$, since $\mathbb{Z} \subseteq \mathbb{Q}$ and $\mathbb{Z} \neq \mathbb{Q}$. One piece of evidence for why $\mathbb{Z} \neq \mathbb{Q}$ is that $\frac{2}{3} \in \mathbb{Q}$ but $\frac{2}{3} \notin \mathbb{Z}$.*

Example 437. *Let $D = \{5, 6\}$. Then D has four subsets, which are $\{\}$ and $\{5\}$ and $\{6\}$ and $\{5, 6\}$. Of these four, the first three are proper subsets of D while the fourth is not a proper subset of D.*

Definition 438: Empty set

A set consisting of no elements is called an **empty set**, and is denoted $\emptyset$.

It may have seemed strange to say "an empty set" instead of "the empty set." Let us prove that an empty set is unique, following Method 334.

Theorem 439. *An empty set is unique.*

Proof. Let A be an empty set. Let B be an empty set. To prove uniqueness, based on the discussion in Section 3.9, we must prove that $A = B$. To prove $A = B$, following Method 432, we need to prove $A \subseteq B$ and also prove $B \subseteq A$.

To prove $A \subseteq B$, following Method 422, we prove that every element of A is an element of B. However since A is an empty set (and has no elements according to Definition 438), there is nothing to prove to show that every element of A is an element of B. This is already true because there are no elements in A. Thus $A \subseteq B$. The argument to prove $B \subseteq A$ is similar. □

Since an empty set is unique, we can now say "the empty set." The empty set is a subset of every set. In other words, for every set A, we have $\emptyset \subseteq A$. (You should prove: for every set A, it is true that $\emptyset \subseteq A$ is true.)

Definition 440: Singleton

A set consisting of exactly one element is called a **singleton**.

Example 441. *The set $\{7\}$ is a singleton.*

Example 442. *The set $\{7,9\}$ is not a singleton, since this set has more than one element.*

It is important to note that $\{7\}$ is different from 7. For example, if $B = \{7,9\}$, it makes sense to write $\{7\} \subseteq B$. However, writing $7 \subseteq B$ doesn't make any sense, because 7 is not a set. This is a good chance to revisit Warning 428.

A finite set generalizes the notion of singleton:

Definition 443: Finite set, infinite set, cardinality

Let A be a set. If there exists a non-negative integer $n \in \mathbb{Z}_{\geq 0}$ such that A has exactly n elements, then we say that A is a **finite set** and the **cardinality** of A is n, denoted $|A| = n$. If there is no such integer n, then A is an **infinite set**.

Example 444. *The set $A = \{2,4,6,8,10\}$ is finite and $|A| = 5$.*

Example 445. *The set $\{7,8,9\}$ is finite and has cardinality 3.*

Example 446. *There is no non-negative integer n such that $\mathbb{Q}$ has exactly n elements. So $\mathbb{Q}$ is an infinite set. Similarly, the set $\mathbb{Z}$ is infinite.*

Example 447. *The sets $\mathbb{R}$ and $\mathbb{C}$ are also infinite.*

Definition 448: Power set

Given a set B, the **power set** of B is the set of all subsets of B, and is denoted $P(B)$.

Note that $A \in P(B)$ if and only if $A \subseteq B$.

Example 449. *Let B be the set $\{7,8\}$. Then $P(B) = \{\{\},\{7\},\{8\},\{7,8\}\}$, which we could also have written $\{\emptyset,\{7\},\{8\},B\}$. Every element of $P(B)$ is a subset of B. Any set which is a subset of B is an element of $P(B)$.*

Example 450. *Let B be the set $\{7,8,9\}$. Then $P(B) = \{\{\},\{7\},\{8\},\{9\},\{7,8\},\{7,9\},\{8,9\},\{7,8,9\}\}$. Each of the eight elements of $P(B)$ is a set.*

Remark 451. *Let B be a set. We have introduced $P(B)$ as the notation for the power set of B. Other texts use $\mathbb{P}(B)$, though for those who have studied probability, do not confuse this for the probability of an event B, which is usually clear by context.*

Remark 452. *Another common notation is 2^B. Note that there will never be confusion whether 2^B denotes the power set of B or is the process of exponentiation from arithmetic, as long as you first ask yourself, "Is B a set or a number?"*

This notation is used because when B is a finite set, $|P(B)| = 2^{|B|}$. In Example 450, we had $|B| = 3$ and $|P(B)| = 2^3$.

Remark 453

The objects of a set may be sets themselves! In Example 450, we saw the set $\{\{\},\{7\},\{8\},\{9\},\{7,8\},\{7,9\},\{8,9\},\{7,8,9\}\}$. In this set, one of the members is the set $\{7,9\}$.

Exercise 454. *Let $X = \{a \in Y : a \text{ teaches geometry}\}$ and let $Y = \{b \in Z \mid b \text{ plays World of Warcraft}\}$. With the assumptions*

- *For all $a \in Z$, if a plays World of Warcraft, then a has a Pinterest account.*
- *For all $z \in Y$, if z does not teach geometry, then z does not have a Pinterest account.*

Prove $X = Y$.

Proof. In the first paragraph, we prove $X \subseteq Y$. Let $c \in X$. We will prove that $c \in Y$. Since $c \in X$, we conclude $c \in Y$ and that c teaches geometry. In particular, $c \in Y$. So $X \subseteq Y$.

In the second paragraph, we prove $Y \subseteq X$. So, let $r \in Y$. We will prove $r \in X$. Since $r \in Y$, we obtain the facts $r \in Z$ and r plays World of Warcraft. Since $r \in Z$, the first hypothesis allows us to conclude that if r plays World of Warcraft, then r has a Pinterest account. Since r plays World of Warcraft, this and the previous implication give us r has a Pinterest account, by modus ponens. Since $r \in Y$, the second hypothesis gives us the fact that if r does not teach geometry, then r does not have a Pinterest account. Since r has a Pinterest account, by modus tollens, r teaches geometry. Since $r \in Y$ and r teaches geometry, we conclude $r \in X$. □

Exercise 455. *Let $C = \{x \in E : x \text{ goes fishing}\}$ and let $B = \{y \in F \mid y \text{ does not like beets}\}$. Using the hypotheses:*

- *$A = P(\emptyset)$, in other words: A is the power set of $\emptyset$.*
- *If E is not an element of the power set of F, then $A \neq \{\emptyset\}$.*
- *$C \subseteq K$*
- *For all $k \in K$, if k likes beets, then k does not go fishing.*

Prove: $C \subseteq B$.

Proof. To prove that $C \subseteq B$, let $a \in C$. We will prove that $a \in B$. Since $a \in C$, we conclude $a \in E$ and a goes fishing.

Consider the first premise. Now, because A is the power set of $\emptyset$, we know that A is non-empty. More specifically, $A = \{0\}$. By applying modus tollens to the second premise, E is an element in the power set of F. In other words, $E \subseteq F$. Since $a \in E$ and $E \subseteq F$, we conclude $a \in F$.

The third premise says $C \subseteq K$ and we already know $a \in C$, so $a \in K$. Since we have an element in K, namely a, we can apply the fourth premise to learn that if a likes beets, then a does not go fishing. Since a goes fishing, by modus tollens, a does not like beets.

Since $a \in F$ and a does not like beets, $a \in B$. Therefore $C \subseteq B$. □

Exercise 456. *Let $X = \{c \in G : c \text{ still has MySpace}\}$ and let $Z = \{a \in Q \mid \forall b \in B, \text{ if } b \text{ likes to skateboard, then } a \text{ follows } b \text{ on Instagram}\}$. Assuming:*

- *$\forall s \in G$, $\exists t \in Z$ such that s follows t on Twitter*
- *$X \subseteq B$*
- *If Q is not an element of the power set of F, then $Z = \emptyset$.*

Prove: $\forall m \in X$, if m likes to skateboard, then $\exists y \in F$ such that y follows m on Instagram. [key]

Proof. Let $m \in X$ be arbitrary. We need to prove if m likes to skateboard, then $\exists y \in F$ such that y follows m on Instagram. Suppose m likes to skateboard. We will prove that there exists a $y \in F$ such that y follows m on Instagram.

Now $m \in X$, so $m \in G$ and m still has MySpace. Since $m \in G$, combining this with the first assumption gives $\exists t \in Z$ such that m follows t on Twitter. Since such a t in Z exists, let t be defined so that $t \in Z$ and m follows t on Twitter.

Because $t \in Z$, we can see that $Z \neq \emptyset$. Then applying modus tollens to the third assumption, we conclude that Q is an element of the power set of F. So $Q \subseteq F$.

As another consequence of $t \in Z$, by definition of Z, we have $t \in Q$ and also now know that for all $b \in B$, if b likes to skateboard, then t follows b on Instagram.

Since $m \in X$ and the second premise is $X \subseteq B$, we now know $m \in B$. Since we know "for all $b \in B$, if b likes to skateboard, then t follows b on Instagram" in particular since $m \in B$, we can conclude that if m likes to skateboard, then t follows m on Instagram.

Since we earlier showed that $t \in Q$ and also showed that $Q \subseteq F$, we have $t \in F$. Since $t \in F$ and t follows m on Instagram, we have proved that there exists a $y \in F$ such that y follows m on Instagram. □

Exercise 457. *Let $A = \{x \in C : x$ writes books and x is a barista$\}$, let $B = \{y \in D \mid y$ sings or y has a goldfish$\}$, and let $C = \{y \in E : y$ likes Culvers$\}$. Using the hypotheses*

- *For all $z \in E$, if z writes books, then z has a goldfish*
- *C is an element of the power set of E*
- *$\forall j \in D$ if j sings then j runs marathons*
- *For all $g \in D$, if g has a goldfish, then g runs marathons.*
- *$\forall b \in B$, if b is not a barista, then b does not run marathons*
- *$\forall x \in B$, if x runs marathons, then x writes books*
- *$\forall z \in C$, if z likes Culvers, then z lives in a pineapple*
- *$C \subseteq D$.*
- *$B \subseteq C$.*

prove $A = B$. *[key]*

Proof. To prove that the sets A and B are equal, we need to prove $A \subseteq B$ and $B \subseteq A$.

To prove that $A \subseteq B$, let $x \in A$ be arbitrary. We will prove that $x \in B$. Since $x \in A$, we conclude $x \in C$ and x writes books and x is a barista. Just to note this now, to prove that $x \in B$, we will need to prove that $x \in D$ and that x sings or x has a goldfish. The second premise is that $C \in P(E)$. In other words, $C \subseteq E$. Since $x \in C$ and $C \subseteq E$, we get $x \in E$. The first hypothesis applies to any element in E, in particular, to $x \in E$, so we conclude that if x writes books, then x has a goldfish. Since we earlier discovered that x writes books, we use modus ponens to conclude that x has a goldfish.

Now $C \subseteq D$ and since $x \in C$, we conclude $x \in D$. Since we saw that x has a goldfish, it is certainly true that x sings or x has a goldfish, because a disjunction is true when at least one of the propositions is true. Since $x \in D$ and since x sings or x has a goldfish, $x \in B$. This completes the proof of $A \subseteq B$.

To prove that $B \subseteq A$, let $x \in B$ be arbitrary. (Note, we reuse x, but have to ignore *everything* from earlier.) We will prove that $x \in A$. In other words, we need to prove that $x \in C$ and that x writes books and x is a barista. Since $x \in B$, by how B is defined, $x \in D$. In addition, x sings or x has a goldfish. (Note, we cannot assume that both x sings and x has a goldfish. We can assume at least one of these two things is true.)

Since at least one of the propositions "x sings" and "x has a goldfish" is true, we prove that x runs marathons by cases. Either x sings, or x has a goldfish:

- Case 1: suppose x sings. Since $x \in D$, the third premise tells us that if x sings, then x runs marathons. By modus ponens, x runs marathons.
- Case 2: suppose x has a goldfish. Since $x \in D$, the fourth proposition tells us that if x has a goldfish, then x runs marathons. By modus ponens, x runs marathons.

In either case (in either situation) x runs marathons. (In other words, no matter what, we know that x runs marathons.) Since $x \in B$, the 6th hypothesis tells us if x runs marathons, then x writes books. So x writes books. Since $x \in B$, the 5th hypothesis tells us if x is not a barista, then x does not run marathons. By modus tollens, x is a barista. So x writes books and x is a barista. Since $x \in B$ and $B \subseteq C$, we conclude $x \in C$. Since $x \in C$ and x writes books and x is a barista, $x \in A$. □

Exercise 458. *Let B be the set $B = \{x \in D : x$ has a puppy$\}$, let F be the set $F = \{z \in E \mid z$ commutes by bike$\}$ and $G = \{z \in B : z$ commutes by train$\}$. Using the hypotheses:*

- *$C \subseteq E$.*
- *for all $c \in C$ and for all $d \in D$, if c supervises d, then c commutes by bike and d commutes by train.*

- $A \subseteq C$.

Prove: for all $x \in A$, for all $y \in B$, if x supervises y, then $x \in F$ and $y \in G$.

Exercise 459. *Use the definitions*

- $X = \{u \in L : u \text{ is a coin collector}\}$
- $Y = \{m \in M \mid m \text{ likes soda if and only if } m \text{ plays video games}\}$

and the hypotheses

- *if x plays video games, then x researches chemistry.*
- $L \subseteq Y$
- *for all $c \in M$, if c researches chemistry, then for all $d \in M$, d gives c pizza.*

to prove: for all $y \in Y$, for all $x \in X$, if x likes soda, then y gives x pizza.

Exercise 460. *Use the definitions*

- $A = \{a \in M : a \text{ likes to study math}\}$
- $B = \{z \in D : \text{ if } z \text{ has a pet hippo, then } z \text{ does not like to study math}\}$
- $C = \{c \in D \mid c \text{ writes poetry}\}$
- $D = \{s \in S : s \text{ likes radishes if and only if } s \text{ writes poetry}\}$

and the hypotheses

- *For all $d \in D$, if d does not have a pet hippo, then d likes radishes.*
- $B \subseteq A$

to prove: $B \subseteq C$.

Exercise 461. *Use the definitions*

- $A = \{x \in U : x \text{ works weekends}\}$
- $C = \{y \in B \mid \text{if } y \text{ works weekends, then } y \text{ manages employees}\}$
- $D = \{m \in U : m \text{ likes pretzels and } m \text{ pickles veggies}\}$
- $K = \{k \in S : k \text{ runs a restaurant}\}$
- $N = \{n \in C : n \text{ likes pretzels if and only if } n \text{ manages employees}\}$
- $U = \{z \in L \mid z \text{ plays poker}\}$

and the hypotheses

- *D is an element of the power set of K*
- $U \subseteq N$

to prove: for all $x \in A$, if x pickles veggies, then x runs a restaurant.

4.4 Set operations

Just as Section 2.1 introduced logical operations (such as negation, conjunction, disjunction, implication, biconditional) which allowed us to meaningfully combine propositions to get new propositions, set operations (such as union and intersection) allow us to meaningfully combine sets to define new sets:

Definition 462: Union

Let A and B be sets. The **union** of A and B, denoted $A \cup B$, is the set that satisfies the property that $x \in A \cup B$ if and only if $x \in A$ or $x \in B$. In symbols,

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

The definition clearly states that $A \cup B$ is a set.

Example 463. *Let $A = \{1, 2, 3, 4\}$ and $B = \{3, 4, 5, 6\}$. Then $A \cup B = \{1, 2, 3, 4, 5, 6\}$. Duplicates may be written optionally, as sets don't keep track of "how many times" an element appears. (Sets are only concerned with* if *an element appears.) See Remark 344.*

Definition 464: Intersection

Let A and B be sets. The **intersection** of A and B, denoted $A \cap B$, is the set that satisfies the property that $x \in A \cap B$ if and only if $x \in A$ and $x \in B$. In symbols,

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

The definition states that $A \cap B$ is a set.

Example 465. *Let $A = \{1, 2, 3, 4\}$ and $B = \{3, 4, 5, 6\}$. Then $A \cap B = \{3, 4\}$.*

Example 466. *Let $C = \{7, 8, 9\}$ and $D = \{9, 10, 11, 12\}$. It is tempting to write $C \cap D = 9$, but this is incorrect. We noted that the intersection is a set. Instead, one must write $C \cap D = \{9\}$.*

Definition 467: Disjoint

Two sets A and B are **disjoint** if $A \cap B = \emptyset$.

Sometimes, this is phrased as A is **disjoint** from B. Two sets are disjoint if and only if they have no elements in common.

Example 468. *The sets $\{1, 2, 3, 4\}$ and $\{7, 8\}$ are disjoint.*

Example 469. *The sets $\{1, 2, 3, 4\}$ and $\{2, 8\}$ are not disjoint.*

Warning 470

The definition of disjoint references the definition of intersection, but the two should not be confused with each other. Given sets $C = \{3, 8\}$ and $D = \{7, 13, 18\}$ and $E = \{8, 18\}$ it is common for students to say things like $C \cap D$ doesn't exist, but this is not true. The intersection of C and D does exist, and is the empty set. We would say C and D are disjoint because $C \cap D = \emptyset$, but we don't say $C \cap D$ doesn't exist. By analogy, we know that $-5 + 5 = 0$, but simply because the result of $-5 + 5$ is the number 0, we don't say $-5 + 5$ doesn't exist.

To restate, both $C \cap E$ and $C \cap D$ exist. In the first case, $C \cap E$ is the set $\{8\}$, while in the second case, $C \cap D$ is the set $\{\}$.

The intersection of two sets is always defined, and is always a set (even if the resulting set from intersection is the empty set). Nothing about the definition of intersection suggests that the intersection of a set A and a set B is only "sometimes" defined, so don't add in this extra language. (In addition, avoid language like "A intersects B" to mean "A and B are not disjoint." Instead, just say/write "A and B are not disjoint."

Our two primitive objects in this handbook have been propositions and sets. (All objects introduced in this handbook thus far have been examples of one or the other.) While it is possible to define an ordered pair through the language of sets, this may add unnecessary confusion to our discussion, so we'll consider this a new primitive object.

Definition 471: Ordered pair

An **ordered pair** is an ordered listing of two elements: (a, b).

The order in an ordered pair matters. Thus, in general, $(a, b) \neq (b, a)$. In fact, the only time $(a, b) = (b, a)$ is when $a = b$.

Example 472. *The ordered pair $(3, 5)$ is not equal to the ordered pair $(5, 3)$.*

Definition 473: Cartesian product

Let A and B be sets. The **Cartesian product** of A and B, denoted $A \times B$, is the set of all possible ordered pairs where the first coordinate of the ordered pair is an element of A and the second coordinate of the ordered pair is an element of B. In symbols,

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

Following Habit 31, the set $A \times B$ is a set of ordered pairs. Due to the Cartesian product, it will be a common occurrence to see sets where the elements are ordered pairs. If you have a set, determine why types of elements are in the set. Do you have a set of numbers, functions, dogs, and now, perhaps ordered pairs? Moreover, think of what type of object you have as the coordinate of each ordered pair.

Example 474. *Let $C = \{1, 2, 4\}$ and $D = \{5, 9\}$. Then $C \times D = \{(1, 5), (1, 9), (2, 5), (2, 9), (4, 5), (4, 9)\}$. Then $C \times D$ is a set of ordered pairs. In more detail, $C \times D$ is a set of ordered pairs of integers.*

Example 475. *Let $M = \{1, 2\}$ and $N = \{2, 3\}$. Then $M \times N = \{(1, 2), (1, 3), (2, 2), (2, 3)\}$. Note that $M \times N$ is a set. In fact, the ordered pair $(2, 3)$ is an element of the set $M \times N$.*

Example 476. *Let $M = \{steak, carnitas, chicken, barbacoa, sofritas\}$ and $T = \{burrito, bowl, taco, salad\}$. If we define $C = M \times T$, then C has a total of $5 \cdot 4 = 20$ ordered pairs, representing the different configurations of menu options available at Chipotle. For instance, $(carnitas, taco) \in C$. Note that C is a set of ordered pairs of kitchen ingredients.*

Example 477. *Let $P = \{A, 2, 3, 4, 5, 6, 7, 8, 9, 10, J, Q, K\}$ and let $S = \{\diamondsuit, \clubsuit, \heartsuit, \spadesuit\}$. Then $P \times S$ would have 52 ordered pairs, each ordered pair representing a card in a standard deck of cards with no jokers. For* *instance,* $(Q, \heartsuit) \in P \times S$.

Warning 478

Suppose that M and N are sets, and suppose you know that $z \in M \times N$. It is an error to then conclude $z \in M$ and $z \in N$. (In fact, note that you would conclude $z \in M$ and $z \in N$ if you had known that $z \in M \cap N$, but the point is to not treat $M \times N$ the same way as $M \cap N$, because $M \times N$ and $M \cap N$ are different.)

In fact, examine Example 475. Using the sets M and N presented there, say that someone told us $z \in M \times N$. If we erroneously concluded $z \in M$ and $z \in N$, then I suppose the only possibility was that $z = 2$. However, the elements of $M \times N$ are not integers such as 2. The elements of $M \times N$ are ordered pairs. Recall, for example, that $(2, 3) \in M \times N$.

Instead of taking $z \in M \times N$ and erroneously concluding $z \in M$ and $z \in N$, apply the rules of inference for set membership.

Warning 479

The previous warning applies even if the same set is used twice in the Cartesian product: suppose that M is a set, and suppose you know that $z \in M \times M$. It is an error to then conclude $z \in M$. It is also an error to say that an element in $M \times M$ must be of the form (a, a). If M has more than one element, then there are elements of the form $(a, b) \in M \times M$, where $a \neq b$.

Warning 480: Avoid writing too much

Some definitions rely on other definitions. When stating a definition, write by using previous definitions, but do not include the text of an earlier definition. This would become burdensome.

Example 481. *When stating the definition of the Cartesian product, you will need to mention "ordered pair" but you should not* define *ordered pair when defining Cartesian product.*

Definition 482: Ordered triple

An **ordered triple** is an ordered listing of three elements or objects: (a, b, c).

Example 483. *The ordered triple* $(4, 5, 5)$ *is not equal to the ordered triple* $(5, 5, 5)$.

Definition 484: Triple Cartesian product

Let A, B, and C be sets. The **Cartesian product** of A, B, and C, denoted $A \times B \times C$, is the set of all possible ordered triples where the first coordinate of the ordered triple is an element of A, the second coordinate of the ordered triple is an element of B, and the third coordinate of the ordered triple is an element from C. In symbols,

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

One can similarly define quadruple Cartesian products, and so on. A fast way to generalize this is:

Definition 485: Ordered n-tuple

An **n-tuple** is an ordered listing of n objects: $(a_1, a_2, \ldots, a_n)$.

Definition 486: n-fold Cartesian product

Let A_1, A_2, ..., and A_n be sets. The **Cartesian product** of A_1, A_2, ..., and A_n, denoted $A_1 \times A_2 \times \cdots \times A_n$, is the set of all possible ordered n-tuples where the ith coordinate of the ordered triple is an element of A_i. In symbols,

$$A_1 \times A_2 \times \cdots \times A_n = \{(a_1, a_2, \ldots, a_n) : a_1 \in A_1 \text{ and } a_2 \in A_2 \text{ and } \ldots \text{ and } a_n \in A_n\}.$$

For clarity, the definition above discusses n sets: the first set is A_1, the second set is A_2, the third set is A_3, and so on. The final set discussed (the nth set) is called A_n.

We often want to take the Cartesian product of a set with itself. (That is, we seek to look at a special case of the previous example where $A_1 = A_2 = \cdots = A_n$. In other words, all n sets are the same set.) Some notation will be convenient for this.

Definition 487: Iterated Cartesian product

Let A be a set. Define $A^2 = A \times A$, define $A^3 = A \times A \times A$, and so on. More generally, if n is a positive integer, then A^n is defined to be $A \times A \times \cdots \times A$, the set of all possible n-tuples, where each coordinate of the n-tuple is any element from the set A.

Example 488. *Here is an important example in linear algebra: $\mathbb{R}^2$. The set $\mathbb{R}^2$ is defined to be $\mathbb{R} \times \mathbb{R}$. If we go back to Definition 473, this is the set of all ordered pairs where the first coordinate is an element of $\mathbb{R}$ and the second coordinate is an element from $\mathbb{R}$. For instance, $(\frac{\pi}{5}, -\sqrt{7}) \in \mathbb{R}^2$. In linear algebra, an element in $\mathbb{R}^2$ such as $(\frac{\pi}{5}, -\sqrt{7})$ is called a* **vector**.

Similarly, $\mathbb{R}^3$ consists of all ordered triples of real numbers. For example, $(\sqrt{6}, 0, -7.89) \in \mathbb{R}^3$. More generally, $\mathbb{R}^n$ is the set of all n-tuples of real numbers.

Definition 489: Set difference

Let A and B be sets. The **set difference** of A and B, denoted $A \setminus B$, is the set that satisfies the property that $x \in A \setminus B$ if and only if $x \in A$ and $x \notin B$. In symbols,

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

In other words, $A \setminus B$ has all elements of A which are not in B. In addition to the notation $A \setminus B$, the notation $A - B$ is used by many authors. There should be no confusion in writing $A - B$ being confused with the subtraction of numbers, since a reader should have already established that A and B were sets (and not numbers) through reading with good habits.

Definition 490: Universal set

Let $A_1, A_2, \ldots, A_n$ be sets. Then a set U is a **universal set** for $A_1, A_2, \ldots, A_n$ if $A_1 \cup A_2 \cup \cdots \cup A_n \subseteq U$.

Definition 491: Complement

Given a set $A \subseteq U$, the **complement** of A with respect to U is the set $U \setminus A$. The complement is denoted $\overline{A}$.

Some authors denote the complement of A by A^c.

Example 492. *Let A_k be the set of integers which are multiples of k. Then $\mathbb{Z}$ is a universal set for A_2 and A_3. However, $\{3, 6, 7\}$ is not a universal set for A_2 since A_2 is not a subset of $\{3, 6, 7\}$.*

Example 493. *Let $A = \{2, 4, 5, 6\}$ and $U = \{1, 2, 3, 4, 5, 6, 7, 8, 9\}$. Then U is a universal set for A since $A \subseteq U$. The complement of A with respect to U is $\overline{A} = A^c = \{1, 3, 7, 8, 9\}$.*

Given sets $A_1, A_2, \ldots, A_n$, we may write

$$\bigcup_{i=1}^{n} A_i$$

as notation to mean $A_1 \cup A_2 \cup \cdots \cup A_n$, reminiscent of summation notation

$$\sum_{i=1}^{n} r_i.$$

Similarly,

$$\bigcup_{i=1}^{n} A_i$$

is notation for $A_1 \cap A_2 \cap \cdots \cap A_n$. For example, we could have written Definition 490 by writing "Let $A_1, A_2, \ldots, A_n$ be sets. Then a set U is a **universal set** for $A_1, A_2, \ldots, A_n$ if $\bigcup_{i=1}^{n} A_i \subseteq U$."

Exercise 494. *Let $C = \{x \in \mathbb{R} \mid x \geq 0\}$. Let $D = \{x^2 : x \in \mathbb{R}\}$. Prove $C = D$.*

Proof. To prove two sets are equal, we have to do two subset proofs.

To prove $C \subseteq D$, let $c \in C$. Then $x \geq 0$. Let $y = \sqrt{x}$. Then $y \in \mathbb{R}$ and because we have shown that there exists a $y \in \mathbb{R}$ such that $y^2 = x$, we have proved that $c \in D$. So $C \subseteq D$.

To prove that $D \subseteq C$, let $d \in D$. Then, there exists a $z \in \mathbb{R}$ such that $d = z^2$. Because the square of a real number is non-negative, z^2 is non-negative. Since $d = z^2$, then d is non-negative as well. Since $d \geq 0$, this proves that $d \in C$. So $D \subseteq C$. $\square$

Exercise 495. *Let $C = \{x \in E : x$ has a dog$\}$ and $D = \{x \in J \mid x$ has a cat$\}$. With the assumptions*

- $E \subseteq J$
- $E \cup Y \subseteq Z$
- $\forall z \in Z$, *if z does not have a cat, then z does not have a dog*

Prove $C \subseteq D$.

Proof. We need to prove $C \subseteq D$. Let $m \in C$. We will prove that $m \in D$. Since $m \in C$, we conclude $m \in E$ and m has a dog. Since $m \in E$ and $E \subseteq J$, we learn that $m \in J$.

Now $m \in E$, so $m \in E$ or $m \in Y$. Thus $m \in E \cup Y$. Since $m \in E \cup Y$ and $E \cup Y \subseteq Z$, we conclude $m \in Z$. Because m is in Z, the third hypothesis tells us that if m does not have a cat, then m does not have a dog. This together with our earlier fact that m has a dog allows us to conclude, by modus tollens, that m has a cat.

Since $m \in J$ and m has a cat, this proves that $m \in D$. Since we took $m \in C$ arbitrarily and proved that $m \in D$, we have shown that C is a subset of D. $\square$

Exercise 496. *Let $R(x, y)$ be the two-variable predicate "If y is an author and $x \in Z$, then x is a painter and $Y \subseteq B$". Let F be the set $F = \{a \in B \mid a$ is an author$\}$. Using the hypotheses:*

- $Z \subseteq Y$.
- $\forall m \in Y\, \exists n \in F$ *s.t.* $R(m, n)$.
- $B \subseteq M$

Prove: for all $a \in Z$, there exists $c \in B$ such that $(a, c) \in M \times M$ and a is a painter.

Exercise 497. *Using the set definitions:*

- $D = \{y \in L : y$ *likes to wave to strangers and* y *is a trendsetter*$\}$.
- $E = \{m \in M \mid m$ *likes to wave to strangers and* m *raises cats*$\}$.

- $K = \{p \in A : p$ *does not raise cats if and only if* p *is not a trendsetter*$\}$.

and the hypotheses:

- *for all* $m \in L$, *if* m *does not give money to charity, then* m *does not like cheese.*
- *if* $E \subseteq D$ *and* $G \neq \emptyset$, *then* D *is an element of the power set of* E.
- *if there exists* $g \in G$ *such that* g *gives money to charity, then* $g \in K$.
- $L \subseteq G$.
- *for all* $p \in M$, *if* p *is a trendsetter, then* p *raises cats.*
- $M \subseteq K$.
- *The set* M *is an element of the power set of* L.

Prove: for all $d \in D$, *if* d *likes cheese, then* d *gives money to charity and* $D = E$.

Exercise 498. *Let us define:*

- $A = \{x \in E : x$ *is an optimist or* x *is famous*$\}$
- $B = \{y \in F \mid y$ *plays solitaire or* y *bakes cookies*$\}$
- $C = \{x \in G \mid x$ *plays solitaire*$\}$
- $D = \{y \in G : y$ *has good manners*$\}$
- $P(y)$ *is the predicate "*y *is not an optimist and* y *is not famous"*

Use the hypotheses

- $G = E \cup F$
- $\forall g \in G$, *if* g *does not have good manners, then* g *does not bake cookies or* $P(g)$.

to prove the statement $A \cap B$ *is an element of the power set of* $C \cup D$. *[Hint: do a proof by cases. Note that when doing a proof by cases, before breaking into your cases, you should (1) identify your cases [the case conditions], and (2) explain why at least one of the case conditions must hold!]*

Exercise 499. *Prove: if* $A \subseteq X$ *and* $B \subseteq Y$, *then* $A \times B \subseteq X \times Y$.

4.5 Algebra of sets

Use Method 432 for the exercises below. Using some of the propositional equivalences in Section 2.3 may be very helpful.

Exercise 500. *Prove: for all sets* A *and* B, *the* **commutative laws** *hold:*

- $A \cup B = B \cup A$
- $A \cap B = B \cap A$

Exercise 501. *Prove: for all sets* A, B *and* C, *the* **associative laws** *hold:*

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

Exercise 502. *Prove: for all sets* A, B *and* C, *the* **distributive laws** *hold:*

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

Exercise 503. *Let* U *be a universal set for* A *and* B. *Prove:* **De Morgan's laws** *hold:*

- $\overline{A \cup B} = \overline{A} \cap \overline{B}$
- $\overline{A \cap B} = \overline{A} \cup \overline{B}$

4.6 Relations

Definition 504: Binary relation

Given a set A and a set B, a **binary relation** from A to B is a subset of $A \times B$.

Example 505. *Let* $A = \{1, 2, 3, 4\}$ *and* $B = \{4, 5, 6, 7\}$. *Then* $R = \{(1, 5), (1, 6), (2, 4), (3, 6), (3, 7), (4, 4)\}$ *is a binary relation from* A *to* B, *because* $R \subseteq A \times B$.

Example 506. *The set* $S = \{(a, b) \in \mathbb{R} \times \mathbb{R} : a - b \text{ is positive}\}$ *is a binary relation from* $\mathbb{R}$ *to* $\mathbb{R}$. *In this example,* $(\pi, 2) \in S$ *while* $(3, \sqrt{26}) \notin S$.

Mathematics uses several notations for the same notions, for example, $\ln(z)$ is alternate notation for $\log_e(z)$. If R is a binary relation, we write aRb for $(a, b) \in R$. Similarly, $a\not{R}b$ is notation for $(a, b) \notin R$. In our second example, we can write $\pi S2$. Writing the symbol representing the relation between is inspired by the standard notation $\pi > 2$. In fact, note that cSd is true if and only if $c > d$ is true.

Definition 507: Binary relation

Given a set A, a **binary relation** on A is a binary relation from A to A.

This definition of binary relation relies on our previous definition of binary relation in Definition 504. There should be no ambiguity because the new definition uses the phrase “binary relation on” a set, while the earlier definition of a binary relation uses the phrase “binary relation from ... to ...”

Example 508. *Let* R *be the binary relation on* $\mathbb{R}$ *defined by* aRb *if and only if* $a - b$ *is negative. Notice* $3R5$ *and* $9\not{R}5$. *To say this differently,* $(3, 5) \in R$ *and* $(9, 5) \notin R$. *There is a difference between* R *versus* aRb: *if* $a = 3$ *and* $b = 5$, *then note that* aRb *is a proposition, while* R *by itself is the binary relation.*

Example 509. *Let* $\spadesuit$ *be the relation on* $\mathbb{Q}$ *defined by* $\spadesuit = \{(c, d) \in \mathbb{Q} \times \mathbb{Q} : c - d \geq 0\}$. *Then* $4 \spadesuit \frac{11}{9}$ *is true, while* $1 \spadesuit \frac{11}{9}$ *is not true. Also,* $5 \spadesuit 5$ *because* $5 - 5 \geq 0$.

Example 510. *Let* $\sim$ *be a binary relation on* $\mathbb{Z}$ *defined by* $a \sim b$ *if and only if* a *divides* b. *Then* $3 \sim 12$ *but* $5 \not\sim 12$. *Note that the proposition* $13 \sim 13$ *is also true.*

Definition 511: n-ary relation

Given sets $A_1, A_2, \ldots, A_n$, an ***n*-ary relation** is a subset of $A_1 \times A_2 \times \cdots \times A_n$.

An n-ary relation is a generalization of a binary relation. (The specific case of binary relations is when $n = 2$.) The most common relations are binary relations. We follow the convention of most texts that a **relation** is assumed to be a binary relation.

Habit 512

The main nouns we have now are propositions, sets, and relations. When new terminology is introduced in a definition, if the term being defined is a noun, ask yourself whether the new term is a proposition, a set, or a relation. More specifically, since a relation is a set, if you have determined that something being defined is a set, sort out whether you have a set that is a relation, or a set that is not a relation.

Due to the existence of binary relations, it will be a common occurrence to see sets where the elements are ordered pairs. If you have a set, determine why types of elements are in the set. Do you have a set of numbers, functions, dogs, and now, perhaps ordered pairs? In addition, if you have a set of ordered pairs, ask yourself what kind of object you have as each coordinate of an ordered pair.

Definition 513: Reflexive

A binary relation R on the set A is **reflexive** if for all $a \in A$, one has aRa.

Remark 514. *The word "reflexive" is an adjective that applies to binary relations.*

Example 515. *The relation described in Example 509 is reflexive. The divisibility relation in Example 510 is also reflexive.*

Example 516. *The relations in Examples 506 and 508 are not reflexive.*

Example 517. *Let $C = \{1, 2, 3\}$. Then $R = \{(1, 1), (3, 3)\}$ is not a reflexive relation C because the ordered pair $(2, 2)$ is not an element of R, and thus the property* $\boxed{\forall c \in C, \text{ the ordered pair } (c, c) \text{ is in } R}$ *fails to hold. Both $S = \{(1, 1), (2, 2), (3, 3)\}$ and $T = \{(1, 1), (2, 2), (3, 3), (3, 1)\}$ are reflexive relations on C. The fact that $(3, 1) \in T$ does not matter.*

Example 518. *Let A be any set. Let R be a binary relation on A. Then R is a* reflexive *binary relation on A if and only if the set $\{(m, m) : m \in A\}$ is a subset of A.*

Method 519: Proving a binary relation is reflexive

Since Definition 513 starts with a universal quantifier, a proof that the relation R on the set A is reflexive should start (informed by Method 268) by writing "Let $x \in A$ be arbitrary." Then, the proof writer should use how the relation R was defined to prove that xRx is true.

As an example, let us consider the following theorem:

Theorem 520. *Let $\sim$ be the relation on $\mathbb{Z}$ defined by $a \sim b$ if and only if a divides b. Then $\sim$ is reflexive.*

Proof. Let $u \in \mathbb{Z}$ be arbitrary. We need to prove $u \sim u$. Since $u = 1u$, this proves that u divides u. Thus, $u \sim u$. Therefore, $\sim$ is reflexive. □

Warning 521

Notice that the previous proof ends by saying that "$\sim$ is reflexive," which is grammatically correct, instead of saying "$u \sim u$ is reflexive" would is grammatically incorrect.

Definition 522: Symmetric

A binary relation R on the set A is **symmetric** if for all $a \in A$ and for all $b \in A$, if aRb, then bRa.

Example 523. *Recall that $\mathbb{H}$ is the set of all people. Then $R = \{(x, y) \in \mathbb{H} \times \mathbb{H} : x \text{ is a blood relative of } y\}$ is a symmetric binary relation on $\mathbb{H}$.*

Example 524. *Let $C = \{1, 2, 3, 4\}$. Then $\{(1, 3), (3, 3)\}$ and $\{(1, 1), (2, 2), (3, 3)\}$ are both symmetric relations on C, but $\{(1, 3), (3, 1), (2, 3)\}$ is not a symmetric relation on C.*

Method 525: Proving a binary relation is symmetric

Since Definition 522 starts with two universal quantifiers, a proof that the relation R on the set A is symmetric should start (informed by Method 268) by writing "Let x and y in A be arbitrary." Then, the proof writer should use how the relation R was defined to prove the implication "if xRy then yRx."

As an example, let us consider the proof of the following theorem:

Theorem 526. *Let $\equiv$ be the relation on $\mathbb{Q}$ defined by $u \equiv w$ if and only if $uw > 0$. Then $\equiv$ is a symmetric relation.*

Proof. Let a and b in $\mathbb{Q}$ be arbitrarily chosen. We wish to prove if $a \equiv b$, then $b \equiv a$. To prove this, suppose that $a \equiv b$, which means that $ab > 0$. Since multiplication of rational numbers is commutative, from $ab > 0$ we get $ba > 0$. Thus $b \equiv a$. Therefore, $\equiv$ is symmetric. □

Warning 527

Note that $\equiv$ is the name of relation, while $a \equiv b$ is a predicate (with variables a and b). Similarly, $\equiv$ is a relation, while $3 \equiv 5$ is a proposition! The word symmetric is an adjective which only applies to relations. So, due to Warning 3, do not say that $a \sim b$ is symmetric. Instead, we finished our proof by saying $\equiv$ is symmetric (without mentioning a or b).

Definition 528: Transitive

A binary relation R on the set A is **transitive** if for all $a \in A$, for all $b \in A$, and for all $c \in A$, if aRb and bRc, then aRc.

Example 529. *The relations described in Examples 506, 508, 509, and 510 are all transitive.*

Example 530. *The relation* $\{(1, 2), (2, 5), (3, 7)\}$ *on the set* $\mathbb{Z}$ *is not transitive the ordered pair* $(1, 5)$ *is missing. On the other hand, the relation* $\{(1, 2), (3, 5), (3, 7)\}$ *on the set* $\mathbb{Z}$ *is transitive.*

Method 531: Proving a binary relation is transitive

Since Definition 528 starts with three universal quantifiers, a proof that the relation R on the set A is symmetric should start (informed by Method 268) by writing "Let x, y, and z in A be arbitrary." Then, the proof writer should use how the relation R was defined to prove the implication "if xRy and yRz, then xRz."

As an example, let us consider the following theorem:

Theorem 532. *Let* $\sim$ *be the relation on* $\mathbb{Z}$ *defined by* $a \sim b$ *if and only if* a *divides* b*. Then* $\sim$ *is transitive.*

Proof. The proof is left as an exercise for the reader. □

Warning 533

One should feel free to use the phrase "$\sim$ is transitive" if it has been proved, but saying "$a \sim b$ is transitive" is grammatically incorrect.

4.7 Equivalence relations

Definition 534: Equivalence relation

A binary relation R on the set A is an **equivalence relation** if R is reflexive, symmetric, and transitive.

Theorem 535. *The relation* $R = \{(u, w) \in \mathbb{Z} \times \mathbb{Z} : 5 \text{ divides } u - w\}$ *on* $\mathbb{Z}$ *is an equivalence relation.*

Proof. To prove that R is an equivalence relation on $\mathbb{Z}$, we need to prove that R is reflexive, symmetric, and transitive.

To prove that R is reflexive, suppose $u \in \mathbb{Z}$ is arbitrary. Then define $d = 0$. Because $d \in \mathbb{Z}$ and $5d = u-u$, we have proved 5 divides $u - u$, so uRu, which proves R is reflexive.

To prove that R is symmetric, let $u, w \in \mathbb{Z}$ be arbitrary. Suppose uRw. Then 5 divides $u - w$, so there exists an integer $z \in \mathbb{Z}$ such that $5z = u - w$. Let $y = -z$. Then $5y = 5(-z) = -5z = -(u - w) = w - u$.

Since $5y = w - u$ and y is an integer, 5 divides $w - u$, which proves wRu. Since we supposed uRw and concluded wRu, for arbitrary $u, w \in \mathbb{Z}$, we have proved that the relation R is symmetric.

To prove R is transitive, let $a, b, c \in \mathbb{Z}$ arbitrary. Suppose aRb and bRc. We must prove aRc. Since aRb and bRc, we know 5 divides both $a - b$ and $b - c$. So there exist integers r and s such that $5r = a - b$ and $5s = b - c$. Let $t = r + s$. Since the sum of integers is an integer, t is an integer. By substitution, $5t = 5(r + s) = 5r + 5s = (a - b) + (b - c) = a - c$. Since t is an integer satisfying $5t = a - c$, we have proved 5 divides $a - c$, so aRc. □

Definition 536: Equivalence class

Suppose that R is an equivalence relation on the set A. Let $a \in A$. Then the **equivalence class** of a under the relation R is

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

and a is called a **representative** of the equivalence class $[a]_R$.

If it is clear which equivalence relation we are talking about from context (because only one was mentioned), then $[a]$ is often used as notation in place of $[a]_R$.

Remark 537. *In order to have an equivalence class, you need two things: first, you need an equivalence relation on a set. (For sake of clarifying the second thing you need, let us call the set A, so that we have an equivalence relation on A.) What is the second thing you need? Note that the second sentence in Definition 536 said "Let $a \in A$" so we need an element of A. The second sentence could have instead said "Fix $a \in A$."*

Example 538. *Consider the equivalence relation R on $\mathbb{Z}$ from Theorem 535. Then what is $[2]_R$? Well, $2 \in [2]_R$ because $(2, 2) \in R$. Also, $7 \in [2]_R$ because $(2, 7) \in R$. Also, $12 \in [2]_R$ because $(2, 12) \in R$. In fact, $[2]_R$ is the set*

$$[2]_R = \{\ldots, -28, -23, -18, -13, -8, -3, 2, 7, 12, 17, 22, 27, 32, 37, \ldots\}$$

Remark 539. *Let us see Remark 537 in action in Example 538. The first of two things we needed was an equivalence relation, which is R. Remark 537 mentioned that we also needed an element from A, and in this example, A is $\mathbb{Z}$. In Example 538, that element was $2 \in \mathbb{Z}$. Those were the two things needed (namely, R and 2) in order to talk about $[2]_R$.*

Exercise 540. *Using the equivalence relation from Theorem 535, describe the set $[1]_R$ in roster format. Describe $[3]_R$ in roster format. Describe $[13]_R$ in roster format. What do you notice about $[3]_R$ and about $[13]_R$? Describe $[0]_R$ in roster format.*

Exercise 541. *Let M be the set of all multiples of 5. Using the equivalence relation from Theorem 535, prove $[0] = M$. Be sure to follow Method 432.*

Theorem 542. *Suppose R is an equivalence relation on the set S. For all $a \in S$ and for all $b \in S$, if $b \in [a]$, then $[a] = [b]$.*

Proof. Suppose R is an equivalence relation on the set S. Let $a \in S$ be arbitrary and let $b \in S$ be arbitrary. We want to prove if $b \in [a]$, then $[a] = [b]$.

Suppose $b \in [a]$. We want to prove $[a] = [b]$. Since $b \in [a]$, we conclude $(a, b) \in R$. Since we need to prove $[a] = [b]$, we follow Method 432.

To prove $[a] \subseteq [b]$, let $m \in [a]$ be arbitrary. We will show that $m \in [b]$. Since $m \in [a]$, we have $(a, m) \in R$. From $(a, b) \in R$, we may conclude $(b, a) \in R$ since R is symmetric. Since $(b, a) \in R$ and $(a, m) \in R$, by transitivity of R, we get $(b, m) \in R$. Thus $m \in [b]$, which concludes the proof of $[a] \subseteq [b]$.

To prove $[b] \subseteq [a]$, let $k \in [b]$ arbitrary. We will show $k \in [a]$. Since $k \in [b]$, we obtain $(b, k) \in R$. Since $(a, b) \in R$ and $(b, k) \in R$, we obtain $(a, k) \in R$ since R is transitive. Thus, $k \in [a]$. Therefore $[b] \subseteq [a]$.

Since $[a]$ and $[b]$ have been proven to be subsets of each other, the sets $[a]$ and $[b]$ are equal. □

Definition 543: Partition

Let S be a set. If $U_1, \ldots, U_k$ are subsets of S, we say that $U_1, \ldots, U_k$ form a **partition** of S if

1. For all $i \in \{1, \ldots, k\}$, the set U_i is non-empty.
2. $\displaystyle\bigcup_{i=1}^{k} U_i = S$.
3. The sets $U_1, \ldots, U_k$ are pair-wise disjoint: that is, if $i \neq j$, then $U_i \cap U_j = \emptyset$.

Example 544. *Consider the equivalence relation from Theorem 535. Let $U_1 = [0]_R$ and $U_2 = [1]_R$ and $U_3 = [2]_R$ and $U_4 = [3]_R$ and $U_5 = [4]_R$. Then $U_1, \ldots, U_5$ form a partition of $\mathbb{Z}$.*

More generally, the set of equivalence classes of an equivalence relation R on A forms a partition of A.

Definition 545: Disjoint Union

Given sets A, B, and C, we say that C is the **disjoint union** of A and B if A and B are disjoint and $C = A \cup B$.

The second and third conditions in the definition of a partition are reminiscent of a disjoint union (of more than two sets).

Example 546. *Continuing with Example 544, the sets $[0]_R$ and $[1]_R$ and $[2]_R$ and $[3]_R$ and $[4]_R$ form a disjoint union of $\mathbb{Z}$.*

Exercise 547. *Let $I = \mathbb{Z}_{>0} \times \mathbb{Z}_{>0}$. Define $\sim$ by saying that $(r,s) \sim (t,u)$ if $r + u = s + t$. Prove that $\sim$ is an equivalence relation on I.*

Exercise 548. *Let $P = \mathbb{Z}_{>0} \times \mathbb{Z}_{>0}$. Define the relation $\sim$ on P by $(a,b) \sim (c,d)$ if and only if $ad = bc$. Prove that $\sim$ is an equivalence relation on P. Hint: before trying to prove anything, what kind of things belong to P? If one is to think of $\sim$ as a set, what kind of ordered pairs belong to $\sim$?*

Exercise 549. *Let $R = \{(x,y) \in \mathbb{R} \times \mathbb{R} : x - y \in \mathbb{Z}\}$. Prove that R is an equivalence relation on $\mathbb{R}$. You do NOT need to prove that R is a relation first.*

Exercise 550. *Let $R = \{(u,w) \in \mathbb{Z} \times \mathbb{Z} : 321 \text{ divides } u - w\}$. Prove that R is an equivalence relation on $\mathbb{Z}$. You do NOT need to prove that R is a binary relation.*

4.8 Functions

Prior to a course in mathematical proof, you had been introduced to the idea of a function. The definitions you had seen in the past (often something like "a function is a rule where...") were adequate at the time, but the proof-based mathematics you will see in the future will require this precise definition instead:

Definition 551: Function, domain, codomain

A **function** f from a set A to a set B is a relation from A to B satisfying (1) for all $a \in A$, there is a $b \in B$ such that $(a,b) \in f$, and (2) for all $a \in A$ and all $b, c \in B$, if $(a,b) \in f$ and $(a,c) \in f$, then $b = c$.

The set A is the **domain** of f and the set B is the **codomain** of f.

Instead of the word **function** other texts may use the word **map**, **mapping**, or **transformation**. For consistency, we will only use the word **function**.

Because a function is a relation, and a relation is a set, a function is ultimately a set (which is a noun).

Habit 552

This habit is meant to update (and enhance) what was said in Habit 512. The main nouns we have now are propositions, sets, relations, and functions. When reading any definition, if the thing being defined is a noun, ask yourself whether the new term is a proposition, a set, a relation, or a function. Since a relation is a set, you should sort out whether the new term is a set that is a relation, or is a set that is not a relation.

If you read the definition of something and have determined it is a relation, since a function is a relation, you should sort out whether the new term is a relation that is a function, or is a relation that is not a function.

Habit 553

If you have a set, determine what types of objects are elements of the set. Do you have a set of numbers, functions, dogs, ordered pairs, or people?

Example 554. *The relation* $\{(1,2),(3,4),(5,4),(7,4),(9,10)\}$ *is a function from domain* $\{1,3,5,7,9\}$ *to codomain* $\{2,4,10\}$.

Example 555. *The relation* $\{(1,2),(3,4),(5,4),(7,4),(9,10)\}$ *is a function from domain* $\{1,3,5,7,9\}$ *to codomain* $\{2,4,6,8,10\}$. *While the set of ordered pairs is the same as the previous example, notice that the codomain has more elements. For every ordered pair, the second coordinate is an element of the codomain. However, not every element of the codomain appears as a second coordinate of an ordered pair.*

Example 556. *Let* $A=\{1,2,3,4,5\}$. *Let* $B=\{4,5,6,7,8\}$. *Let* $f=\{(1,8),(2,7),(3,6),(4,5),(5,4)\}$. *Then* f *is a function from* A *to* B. *Our previous examples did not name the function. Here, we have used* f *as the name/label of the function.*

Example 557. *Let* $A=\{1,2,3,4,5\}$. *Let* $B=\{4,5,6,7,8\}$. *Then* $\{(1,8),(2,7),(3,6),(4,5)\}$ *is* **not** *a function from* A *to* B *because there is no ordered pair with first coordinate* 5*, breaking condition (1) in the definition of a function.*

Example 558. *Let* $A=\{1,2,3,4,5\}$. *Let* $B=\{4,5,6,7,8\}$. *Then* $\{(1,8),(2,7),(3,6),(4,5),(5,4),(5,8)\}$ *is* **not** *a function from* A *to* B *because the ordered pairs* $(5,4)$ *and* $(5,8)$ *are both elements of the relation, breaking condition (2) in the definition of a function.*

Example 559. *Let* $A=\{1,2,3,4,5\}$. *Let* $B=\{4,5,6,7,8\}$. *Then* $n=\{(1,8),(2,7),(3,6),(4,5),(5,9)\}$ *is* **not** *a function from* A *to* B *because* $9\notin B$*, and thus* n *is not a relation from* A *to* B *(in that* n *is not a subset of* $A\times B$*.)*

Example 560. *Let* $D=\{1,2,3\}$. *Let* $C=\{\clubsuit,\spadesuit\}$. *Then* $\{(1,\clubsuit),(2,\spadesuit),(3,\spadesuit)\}$ *is a function from* D *to* C. *The set* D *is the domain of this function, and the set* C *is the codomain of this function.*

We often write $f:A\to B$ as notation to mean that f is a function from A to B. (As a stylistic note, writing $\boxed{f \text{ is a function from } A\to B}$ where the arrow replaces the word "to" should be avoided: if the word "from" is written out in words, the word "to" should be written out in words for balance.)

If f is a function, we often write $f(a)=b$ as notation to mean $(a,b)\in f$. Writing $(a,b)\in f$ provides the reminder that a function is a special kind of relation while writing $f(a)=b$ provides a connection to the past with familiar notation used for functions in calculus and algebra. If it is clear which function is being discussed (that is, an exercise or proof does not deal with two functions), then $a\mapsto b$ is also used as notation to mean $(a,b)\in f$. Notice that the symbol between a and b is a short vertical stick attached to an arrow.

Example 561. *We revisit the function described in Example 554, with the only difference being that we will name the function by the letter* h. *Let* h *be the function from* $\{1,3,5,7,9\}$ *to* $\{2,4,10\}$ *by the rule* $h(1)=2$ *and* $h(3)=4$ *and* $h(5)=4$ *and* $h(7)=4$ *and* $h(9)=10$.

Example 562. *We revisit the function described in Example 560 using new notation. Consider the function from* $D=\{1,2,3\}$ *to* $C=\{\clubsuit,\spadesuit\}$ *where* $1\mapsto\clubsuit$ *and* $2\mapsto\spadesuit$ *and* $3\mapsto\spadesuit$.

Method 563: Defining a function

To fully define a function, you must describe (1) the function's domain, (2) the function's codomain, and (3) the rule for the function.

Warning 564

Do not think of $f(x) = x^2 + \sin(x)$ as the only thing being a function. Notice that Example 554 is a function, although there is no formula such as $x^2+\sin(x)$ written. Moreover, writing $f(x) = x^2+\sin(x)$ actually breaks the principle of Method 563 by leaving the domain and codomain implicit. While this type of writing was fine in precalculus, it does not properly define a function in a proof-based math course.

Warning 565

Do not think that having a formula such as $f(x) = x^2 + \sin(x)$ is necessary to have a function be a function. For example, examine Example 560. Do not think that a function has to be a continuous function (in the sense of first semester calculus). Any relation that satisfies Definition 551 is a function.

Warning 566

When defining a function $f : A \to B$, your rule must ensure that $f(a) \in B$ for each $a \in A$.

Example 567. *If someone writes "Let us define $f : [0, 1] \to [4, 6]$ by the rule $f(x) = 3x + 4$" though a domain, codomain and rule have been provided, this does not properly define a function due to Warning 566. Notice that $f(1) = 7$, but $7 \notin [4, 6]$.*

To contrast, if someone writes "Let us define $g : [0, 1] \to [4, 8]$ by the rule $g(x) = 3x + 4$" they have properly defined a function. Warning 566 has been properly heeded, because for every $a \in [0, 1]$, notice that $f(a) \in [4, 8]$.

Example 568. *If someone writes "Let us define $h : \mathbb{R} \to [0, 1]$ by the rule $h(x) = \sin(x)$" this does not properly define a function due to Warning 566. Note that $h(\frac{3\pi}{2}) \notin [0, 1]$.*

By contrast, it is okay to define a function by writing: let $f : \mathbb{R} \to \mathbb{R}$ by the rule $f(x) = \sin(x)$.

Example 569. *Let $f : \mathbb{R} \to \mathbb{R}$ by the rule $f(x) = 1 + \cos(x)$ and let $g : \mathbb{R} \to [0, 2]$ by the rule $g(x) = 1 + \cos(x)$. Both f and g are well-defined functions. Notice that every element in the codomain of g is an output (for some x), while not every element in the codomain of f is an output. This example will be used later to discuss the difference between codomain and range. (We have not yet discussed range.)*

Example 570. *Let $A = \{1, 2, 3\}$ and $B = \{4, 5, 6\}$ and $C = \{4, 5, 6, 7\}$ and $D = \{4, 5\}$.*

Writing "Let us define $h : A \to B$ by the rule $h = \{(1, 4), (2, 5), (3, 7)\}$" does not properly define a function due to Warning 566. Note that $h(3) \notin B$.

Let $f : A \to B$ by the rule $f(1) = 4$ and $f(2) = 4$ and $f(3) = 5$. Let $g : A \to D$ by the rule $g(1) = 4$ and $g(2) = 4$ and $g(3) = 5$. Both f and g are well-defined functions. Notice that every element in the codomain of g is an output, while not every element in the codomain of f is an output. This example will be used later to discuss the difference between codomain and range. (We have not yet discussed range.)

Remark 571. *We could have also defined the same function f by writing: Let $f : A \to B$ defined by the rule $f = \{(1, 4), (2, 4), (3, 5)\}$. Writing $f(1) = 4$ and $f(2) = 4$ and $f(3) = 5$ is more reminiscent of how function notation was used in the past, while $f = \{(1, 4), (2, 4), (3, 5)\}$ creates the connection that a function is a special kind of relation. In fact, the view of $f = \{(1, 4), (2, 4), (3, 5)\}$ is that the function (written as a set) is what used to be called the* graph *of the function.*

Warning 572: f versus $f(x)$

In calculus and earlier, you may have been used to using the notations f and $f(x)$ interchangeably. In calculus and prior, it is possible to get away with calling both f and $f(x)$ a function. In proof-based mathematics, the distinction between f and $f(x)$ must be made. Here, f is the name of a function, while $f(x)$ is an element of the codomain (if x is in the domain).

We now define what it means for two functions to be equal:

Definition 573: Function equality

Two functions f and g are **equal** if

1. f and g have the same domain,
2. f and g have the same codomain, and
3. for all x in their common domain, $f(x) = g(x)$.

Notice in this definition how Warning 572 is applicable. The first and second items of the definition refer to the functions f and g, while the third item in the definition refers to outputs $f(x)$ and $g(x)$. In other words, function equality is defined in terms of (due to the third item) function value equality.

Warning 574: Function equality versus set equality

Suppose you are asked to prove that $\clubsuit = \spadesuit$. Now, it becomes all the more important to follow Habit 552. If $\clubsuit$ and $\spadesuit$ are both sets, then follow Method 432 to prove $\clubsuit = \spadesuit$. If $\clubsuit$ and $\spadesuit$ are both functions, then prove the three items in Definition 573.

Definition 575: Composition

Given functions $f : A \to B$ and $g : B \to C$, the **composition** (or **composite function**) is the function $(g \circ f) : A \to C$ given by the rule $(g \circ f)(x) = g(f(x))$.

Remark 576

Sometimes, a function is not fully defined, but enough information is given. Even if the rule for a function is not explicitly mentioned, it may be invoked.

The definition of composition defines a function, and notice how Method 563 was followed. For the function $g \circ f$, a domain was specified (namely A), a codomain was specified (namely C), and a rule was specified. In fact, the rule $(g \circ f)(x) = g(f(x))$ was given for $g \circ f$, even though the rules for the functions f and g were not explicitly given. This is an example of Remark 576: the functions f and g were not fully defined, but we still invoked the rules of f and g when we wrote $g(f(x))$.

Example 577. *Let $A = \{1, 2, 3, 4\}$ and $B = \{4, 5, 6\}$ and $C = \{7, 8, 9\}$. If $f = \{(1, 4), (2, 4), (3, 6), (4, 5)\}$ and $g = \{(4, 9), (5, 8), (6, 7)\}$, then $g \circ f = \{(1, 9), (2, 9), (3, 7), (4, 8)\}$.*

Warning 578

Be careful with notation surrounding compositions. It is tempting to write $\boxed{g\, of\, f\, x}$ and similar things, but this mixes notation and words in an awkward manner. Instead, write $g(f(x))$ or write $(g \circ f)(x)$. In addition, writing $\boxed{g(f)}$ and other similar things makes no grammatical sense: see Warning 572.

Theorem 579. *Function composition is associative.*

Proof. Let $f : A \to B$ and $g : B \to C$ and $h : C \to D$ be any functions (with the specified domains and codomains). We need to prove that $(h \circ g) \circ f = h \circ (g \circ f)$. Making note of Warning 574, we are directed to Definition 573.

For convenience, we will let $j = h \circ g$ and $k = g \circ f$. By carefully following the notation in Definition 575, j is a function from B to D, and k is a function from A to C. Then $(h \circ g) \circ f = j \circ f$, a function from A to D, and $h \circ (g \circ f) = h \circ k$, a function from A to D. Since both the function $(h \circ g) \circ f$ on the left on the function $h \circ (g \circ f)$ on the right are from A to D, to prove the functions are equal, we are only left with checking that they have the same rule.

So, we wish to prove that for all $x \in A$, one has $((h \circ g) \circ f)(x) = (h \circ (g \circ f))(x)$. Let $x \in A$ be arbitrary. We want to prove $((h \circ g) \circ f)(x) = (h \circ (g \circ f))(x)$. In other words, we want to prove $(j \circ f)(x) = (h \circ k)(x)$.

First, consider the expression $(j \circ f)(x)$, and for convenience, let $y = f(x)$. Note that $(j \circ f)(x) = j(f(x)) = j(y) = (h \circ g)(y) = h(g(y)) = h(g(f(x))$.

Second, consider the expression $(h \circ k)(x)$, but first note that $k(x) = (g \circ f)(x) = g(f(x))$. Then $(h \circ k)(x) = h(k(x)) = h(g(f(x))$.

Since $(h \circ g) \circ f$ and $h \circ (g \circ f)$ both have domain A and both have codomain D and have the same rule, namely that for all $a \in A$, we have $((h \circ g) \circ f)(a) = (h \circ (g \circ f))(a)$, we conclude that $(h \circ g) \circ f = h \circ (g \circ f)$. ☐

Examine the proof of this Theorem 579 carefully to see if Warning 572 was properly followed. Is there any place where an f should be replaced with an $f(x)$ or vice versa? Is there any place where $(h \circ g) \circ f$ should be replaced with $((h \circ g) \circ f)(x)$ or vice versa?

Definition 580: Preimage of an element

Let $f : A \to B$. Let $b \in B$. Then the **preimage of b** is

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

Definition 581: Preimage of a set

Let $f : A \to B$. Let $Z \subseteq B$. Then the **preimage of Z** is

$$f^{-1}(Z) = \{a \in A : \text{there exists } b \in Z \text{ such that } f(a) = b\}.$$

Warning 582

Do not confuse the preimage of an element with the preimage of a set (see Definitions 580 and 581). When encountering $f^{-1}(☺)$, first figure out if this is preimage of an element or preimage of a set. From the standpoint of notation, these look identical, but the notation is never ambiguous! Say $f : A \to B$. If $☺ \in B$, then use Definition 580. If $☺ \subseteq B$, then use Definition 581.

There are authors that use different notations for the preimage of an element versus the preimage of a set, but the majority of authors use what appears to be the same notation. If you first ask yourself whether the thing in the parentheses after the f^{-1} is an *element* of the codomain or a *subset* of the codomain, then there is no ambiguity!

Example 583. *Let $f : \mathbb{R} \to \mathbb{R}$ be defined by $f(x) = 1 + x$. If $b = 10$, since $b \in B$, we use Definition 580, so $f^{-1}(b) = \{9\}$. If $b = \{10\}$, since $b \subseteq B$, we use Definition 581, so $f^{-1}(b) = \{9\}$. If $b = [10, 26]$, since $b \subseteq B$, we use Definition 581, so $f^{-1}(b) = [9, 25]$.*

Recall (see Remark 341) that the name of a set does not have to use a capital letter. For instance, we could have presented the last example by writing if $\spadesuit = 10$, then $f^{-1}(\spadesuit) = \{9\}$ using Definition 580, and instead if $\clubsuit = [10, 26]$, then $f^{-1}(\clubsuit) = [9, 25]$ using Definition 581.

Example 584. *Let $f : \mathbb{R} \to \mathbb{R}$ be defined by $f(x) = 1 + x^2$. If $b = 10$, since $b \in B$, we use Definition 580, so $f^{-1}(b) = \{-3, 3\}$. If $b = \{10\}$, since $b \subseteq B$, we use Definition 581, so $f^{-1}(b) = \{-3, 3\}$. If $b = [10, 26]$, since $b \subseteq B$, we use Definition 581, so $f^{-1}(b) = [-5, -3] \cup [3, 5]$.*

Example 585. *If we name the function from Example 554 by f, then $f^{-1}(4) = \{3, 5, 7\}$ and $f^{-1}(10) = \{9\}$ using Definition 580. Note that $f^{-1}(10) \neq 9$. Using Definition 581, $f^{-1}(\{2, 10\}) = \{1, 9\}$.*

Example 586. *If we name the function from Example 555 by g, then $g^{-1}(4) = \{3, 5, 7\}$ and $g^{-1}(6) = \emptyset$ using Definition 580. Using Definition 581, $g^{-1}(\{4\}) = \{3, 5, 7\}$ and $g^{-1}(\{2, 4, 6\}) = \{1, 3, 5, 7\}$.*

Example 587. *If we name the function from Example 560 by h, then $h^{-1}(\spadesuit) = \{2, 3\}$ using Definition 580.*

Example 588. *With the setup of Example 570, $f^{-1}(\{4\}) = \{1, 2\}$ and $f^{-1}(\{5\}) = \{3\}$ and $f^{-1}(\{4, 5\}) = \{1, 2, 3\}$ using Definition 581. By contrast, $f^{-1}(4) = \{1, 2\}$ and $f^{-1}(5) = \{3\}$ using Definition 580.*

Definition 589: Image of a set

Let $f : A \to B$. Let $Z \subseteq A$. Then the **image of Z** is

$$f(Z) = \{f(a) : a \in Z\}.$$

As defined, the set $f(Z)$ is written in build running through set notation, so we may convert this and write instead

$$f(Z) = \{b : \text{there exists } a \in Z \text{ such that } f(a) = b\}.$$

Due to the fact that the definition of a function tells us that $f(a)$ is always in the codomain, we could even write

$$f(Z) = \{b \in B : \text{there exists } a \in Z \text{ such that } f(a) = b\}.$$

In terms of notation, this is very close to the definition of the preimage of a set in Definition 581. Instead of starting with a subset of Z the codomain and obtaining a subset of the domain, in the definition of image, we start with a subset Z of the domain and obtain a subset of the codomain.

Warning 590

Do not confuse the output of an element with the image of a set. Say $f : A \to B$ and you encounter $f(☺)$. If $☺ \in A$, then $f(☺)$ is an element of B, following the definition of function. If $☺ \subseteq A$, then use Definition 589, and $f(☺)$ will be a subset of the codomain.

Example 591. *Let $f : \mathbb{R} \to \mathbb{R}$ be defined by $f(x) = 1 + x$. If $a = 10$, since $a \in A$, we use Definition 551, so $f(a) = 11$. If $a = \{10\}$, since $a \subseteq A$, we use Definition 589, so $f(a) = \{11\}$.*

Example 592. *If we name the function from Example 554 by f, then $f(1) = 2$ using Definition 551, while $f(\{1\}) = \{2\}$ and $f(\{1, 3\}) = \{2, 4\}$ using Definition 589.*

Example 593. *If we name the function from Example 555 by g, then $g(9) = 10$ using Definition 551, while $g(\{9\}) = \{10\}$ and $g(\{7, 9\}) = \{4, 10\}$ using Definition 589.*

Example 594. *If we name the function from Example 560 by h, then $h(1) = \clubsuit$ using Definition 551, while $h(\{1\}) = \{\clubsuit\}$ and $h(\{1, 3\}) = \{\clubsuit, \spadesuit\}$ using Definition 589.*

Example 595. *With the setup of Example 570, then $f(1) = 4$ using Definition 551, while $f(\{1, 2\}) = \{4\}$ and $f(\{2, 3\}) = \{4, 5\}$ and $f(\{1, 2, 3\}) = \{4, 5\}$ using Definition 589.*

Definition 596: Range

Let $f : A \to B$. Then the **range of f** is

$$\{f(a) : a \in A\}.$$

As defined, the range is written in build running through set notation, so we may convert this and write instead

$$\{b : \text{there exists } a \in A \text{ such that } f(a) = b\}.$$

Due to the fact that the definition of a function tells us that $f(a)$ is always in the codomain, we could even write

$$\{b \in B : \text{there exists } a \in A \text{ such that } f(a) = b\}.$$

Example 597. *The range of $f : \mathbb{R} \to \mathbb{R}$ defined by $f(x) = 1 + x$ is $\mathbb{R}$. Note that the codomain of f is also $\mathbb{R}$.*

Example 598. *The range $f : \mathbb{R} \to \mathbb{R}$ defined by $f(x) = 1 + x^2$ is $[1, \infty)$. Note that the codomain of f is $\mathbb{R}$.*

Example 599. *The range $f : \mathbb{R} \to [1, \infty)$ defined by $f(x) = 1 + x^2$ is $[1, \infty)$. Note that the codomain of f is also $[1, \infty)$.*

Example 600. *The range of the function from Example 554 is $\{2, 4, 10\}$. Notice that the codomain is also $\{2, 4, 10\}$.*

Example 601. *The range of the function from Example 555 is $\{2, 4, 10\}$, but the codomain is $\{2, 4, 6, 8, 10\}$.*

Example 602. *The range and the codomain of the function from Example 560 are both $\{\clubsuit, \spadesuit\}$.*

Example 603. *With the setup of Example 569, the range of f is $[0, 2]$ while the codomain of f is $\mathbb{R}$. The range of g is $[0, 2]$ and the codomain of g is also $[0, 2]$.*

Example 604. *With the setup of Example 570, the range of f is $\{4, 5\}$ while the codomain of f is $\{4, 5, 6\}$. The range of g is $\{4, 5\}$ and the codomain of g is also $\{4, 5\}$.*

Example 605. *Let $f : \mathbb{Q} \to \mathbb{R}$ defined by $f(x) = \sin(x)$. Then the codomain of f is $\mathbb{R}$, while the range of f is the set R defined in Example 399. Note that R is a proper subset of $\mathbb{R}$.*

Definition 606: Surjective

A function $f : A \to B$ is **surjective** if for all $y \in B$, there exists an $x \in A$ such that $f(x) = y$.

Following Habit 2, surjective is an adjective. Since surjective is an adjective which applies to functions, following Warning 3, we should not apply this adjective to anything which is *not* a function.

If f is a surjective function, we can also call f a **surjection**. Other texts will say that f is **onto** instead of saying f is surjective. The phrases "f is onto" and "f is surjective" mean the same thing.

Definition 607: Onto

A function $f : A \to B$ is **onto** if for all $y \in B$, there exists an $x \in A$ such that $f(x) = y$.

Example 608. *The range of $f : \mathbb{R} \to \mathbb{R}$ defined by $f(x) = 1 + x$ is $\mathbb{R}$. Note that the codomain of f is also $\mathbb{R}$. Thus, f is surjective.*

Example 609. *The range $f : \mathbb{R} \to \mathbb{R}$ defined by $f(x) = 1 + x^2$ is $[1, \infty)$. Note that the codomain of f is $\mathbb{R}$. Note that f is not surjective since there does not exist an $x \in \mathbb{R}$ such that $f(x) = \frac{1}{2}$, yet $\frac{1}{2}$ is in the codomain.*

Example 610. *The range $f : \mathbb{R} \to [1, \infty)$ defined by $f(x) = 1 + x^2$ is $[1, \infty)$. Note that the codomain of f is also $[1, \infty)$. Thus, f is onto.*

Example 611. *The function from Example 554 is onto.*

Example 612. *The function from Example 555 not onto, since there is no input which has 6 as an output.*

Example 613. *The function from Example 560 is surjective, since the range and codomain are equal.*

Example 614. *Both in Example 569 and in Example 570, f is not surjective and g is surjective.*

How would Warning 480 apply here? When stating the definition of onto, you should mention the word onto, but you should not (at the same time) define what a function is. Write to an audience who already knows what the definition of function is.

Method 615: Proving a function is surjective

To prove that the function $f : A \to B$ is surjective, since the definition of surjectivity (Definition 606) is a universally-quantified statement, informed by Method 268, start by writing "Let $y \in B$ be arbitrary." Then, we need to prove that there exists an $x \in A$ such that $f(x) = y$. To prove this existentially-quantified statement, following Method 217 we need to define something which we'll name x (and how x is defined likely depends on y), then prove $x \in A$ and also prove $f(x) = y$. Proving $f(x) = y$ will require the use of the definition of f.

Method 616: Using a function is surjective

To use the fact that the function $f : A \to B$ is surjective, since the definition of surjectivity (Definition 606) is a universally-quantified statement, informed by Method 252, we must already have an element in B, or else we cannot use the fact that f is surjective. Suppose we are in the middle of a proof and have $m \in B$ already established. Then, we would be able to conclude that there exists an $n \in A$ such that $f(n) = m$.

Theorem 617. *If $f : A \to B$ and $g : B \to C$ are surjective, then $(g \circ f)$ is surjective.*

Proof, with annotations and comments. Suppose that $f : A \to B$ and $g : B \to C$ are surjective. From Definition 575, $g \circ f$ is a function from A to C. We want to prove $g \circ f$ is surjective, so following Method 615, we let $z \in C$ be arbitrary. (Note that C is the codomain of $g \circ f$.) While we want to use the fact that f is surjective, Method 616 warns us that we need to have an element of B already established, and we don't. So, at the moment, we cannot use the fact that f is surjective.

Since we have an element of C, we can use the fact that g is surjective, following Method 616. So, there exists a $y \in B$ such that $g(y) = z$. Now, that we have $y \in B$, we can follow Method 616 and use the fact that f is surjective, so there exists an $x \in A$ such that $f(x) = y$.

Then $(g \circ f)(x) = g(f(x)) = g(y) = z$, so $g \circ f$ is surjective. □

Proof, without annotations or comments. Suppose that $f : A \to B$ and $g : B \to C$ are surjective. Let $z \in C$. Since g is surjective, there exists a $y \in B$ such that $g(y) = z$. Since f is surjective, there exists an $x \in A$ such that $f(x) = y$. Then $(g \circ f)(x) = g(f(x)) = g(y) = z$, so $g \circ f$ is surjective. □

Notice that the proof follows Method 616 twice (in the only order that works) and follows Method 615 once.

Definition 618: Injective

A function $f : A \to B$ is **injective** if for all $w, x \in A$, if $f(w) = f(x)$, then $w = x$.

Following Habit 2, injective is an adjective. What kind of noun does injective modify? Based on the definition, injective is an adjective which applies to functions. As an example of Warning3, it is forbidden to use the adjective injective on anything which is *not* a function.

To describe the same notion, in the form of a noun, if f is an injective function, we can refer to f as an **injection**. Other texts will say that f is **one-to-one** instead of saying f is injective. The phrases "f is one-to-one" and "f is injective" say exactly the same thing.

Definition 619: One-to-one

A function $f : A \to B$ is **one-to-one** if for all $w, x \in A$, if $f(w) = f(x)$, then $w = x$.

Habit 620

It is tempting to think of the definition injective/one-to-one as 14 or so separate words, phrases, or bits of notation. Thinking of f, then A, then arrow, then B, then "injective" then, "if", then "for all", and so on is not sustainable. Instead, consider the advice of Section 3.3. Think of something "wordy" to serve as your memory hook for the definition. As an example, a function is injective if the same outputs lead to the same inputs.

Remark 621. *Due to the logical equivalence of an implication and its contrapositive, other textbooks will state the definition of injective as follows: A function $f : A \to B$ is injective if for all $w, x \in A$, if $w \neq x$, then $f(w) \neq f(x)$.*

Example 622. *The function $f : \mathbb{R} \to \mathbb{R}$ defined by $f(x) = 1 + x$ is injective.*

Example 623. *The function $f : \mathbb{R} \to \mathbb{R}$ defined by $f(x) = 1 + x^2$ is not injective since $f(3) = f(-3)$ yet $3 \neq -3$.*

Example 624. *The function from Example 554 is not one-to-one since $f(3) = f(5)$.*

Remark 625. *When you need to unpack what it means if a function f is one-to-one, it is not helpful to think that f is injective. Think about the condition "for all $w, x \in A$, if $f(w) = f(x)$, then $w = x$" instead.*

How would Warning 480 apply here? When stating the definition of injective, you should mention the word injective, but you should not (at the same time) define what a function is. Write to an audience who already knows what the definition of function is.

Method 626: Proving a function is injective

To prove that the function $f : A \to B$ is injective, since the definition of injectivity (see Definition 618) is a universally-quantified statement, informed by Method 268, start by writing "Let $w \in A$ and $x \in A$ be arbitrary." This is assuming that the variables w and x are unused in your proof so far. Then, since we need to prove $\boxed{\text{If } f(w) = f(x)\text{, then } w = x}$, following Method 202, we should assume $f(w) = f(x)$. Then, from this fact, we should prove that $w = x$, which will require the definition of f.
Once we have written $\boxed{\text{Let } w \in A \text{ and } x \in A \text{ be arbitrary.}}$ we can prove $\boxed{\text{If } f(w) = f(x)\text{, then } w = x}$ by proving its contrapositive instead, following Method 313 by assuming $w \neq x$ and then proving $f(w) \neq f(x)$ using the definition of f.

Method 627: Using a function is injective

To use the fact that the function $f : A \to B$ is injective, since the definition of injectivity (see Definition 618) is a twice universally-quantified statement, informed by Method 252, we must have two elements of A, which we will call here c and d. If we have that (in other words, we already have a $c \in A$ and a $d \in A$), then we need to use the implication "if $f(c) = f(d)$, then $c = d$." Most of the time, this will be achieved by following Method 193, where we must have the situation that $f(c)$ equals $f(d)$ and we then get to conclude that $c = d$. However, sometimes, we might use the implication by following Method 196, where we must have the situation that $c \neq d$, in which case we get to conclude that $f(c)$ and $f(d)$ are different elements of the codomain of f.

Theorem 628. *If $f : A \to B$ and $g : B \to C$ are injective, then $(g \circ f)$ is injective.*

Proof. Suppose $f : A \to B$ and $g : B \to C$ are injective. To prove $(g \circ f) : A \to C$ is injective, let $w, x \in A$ be arbitrary. Suppose $(g \circ f)(w) = (g \circ f)(x)$. We want to prove that $w = x$. We can rewrite our earlier equation as $g(f(w)) = g(f(x))$. Now, since $f(w)$ and $f(x)$ are in B and g is injective, $f(w) = f(x)$. Then, since f is injective, $w = x$. $\square$

Exercise 629. *Define four functions from $\mathbb{R}$ to $\mathbb{R}$. Have examples of functions which are injective, surjective, both, or neither. (Hint: you do not have to stick with continuous functions.) Having these examples will be* *useful in building intuition regarding other statements about functions.*

Definition 630

A function f is **bijective** if f is injective and f is surjective.

We use the phrases "f is bijective" and "f is a bijective function" and "f is a **bijection**" interchangeably. Using the other typical terminology:

Definition 631

A function f is **one-to-one correspondence** if f is one-to-one and f is onto.

Example 632. *The function $f : \mathbb{R} \to \mathbb{R}$ defined by $f(x) = 3x + 1$ is bijective.*

Example 633. *The function $f : \mathbb{R} \to \mathbb{R}$ defined by $f(x) = 3x^2+1$ is not bijective, since f is neither injective nor surjective.*

Example 634. *The function $f : \mathbb{R} \to [1, \infty)$ defined by $f(x) = 3x^2 + 1$ is not injective, but is surjective. Thus f is not bijective. Said differently, f is not a one-to-one correspondence.*

Example 635. *The function $f : \{1, 2, 3\} \to \{4, 5, 6\}$ defined by $f(1) = 5$ and $f(2) = 4$ and $f(3) = 6$ is bijective. Said differently, f is a one-to-one correspondence.*

Definition 636. *Absolute valueabsolute value The* **absolute value function** *is the function $|\cdot|: \mathbb{R} \to \mathbb{R}$ with the rule*

$$|x| = \begin{cases} x & \textit{if } x \geq 0 \\ -x & \textit{if } x < 0. \end{cases}$$

In the stated definition above, we have written absolute value using its typical notation, representing the input by a single dot. To describe the function using more standard function notation, we could write: The **absolute value function** is the function $a : \mathbb{R} \to \mathbb{R}$ with the rule

$$a(x) = \begin{cases} x & \text{if } x \geq 0 \\ -x & \text{if } x < 0. \end{cases}$$

Example 637.

Method 638: Proving a function is bijective

To prove that a function f is bijective, prove f is injective following Method 626 and prove f is surjective following Method 615. The proofs of injectivity and surjectivity can be done in either order, but it is helpful to be clear. (Perhaps you have a paragraph that starts with the phrase "To prove f is injective" and another paragraph that starts "For surjectivity" or similar phrasing.)

Method 639: Using a function is bijective

If it has been established that a function f is bijective, you will either need to use the fact that f is injective following Method 627 or use the fact that f is surjective following Method 616. It is highly likely that you will need to use both facts one or more times.

Theorem 640. *If $f : A \to B$ and $g : B \to C$ are bijective, then $g \circ f$ is bijective.*

Exercise 641. *Prove the theorem above.*

Definition 642

Let $f : A \to B$ be a bijective function. The **inverse function** of f is the function $f^{-1} : B \to A$ with the rule $f^{-1} = \{(b, a) : (a, b) \in f\}$.

Note that the domain and codomain have swapped when going from f to its inverse f^{-1}.

Example 643. *The function f defined in Example 635 is bijective. Its inverse is $f^{-1} : \{4,5,6\} \to \{1,2,3\}$ defined by $f^{-1}(5) = 1$ and $f^{-1}(4) = 2$ and $f^{-1}(6) = 3$.*

Example 644. *Let $f : \mathbb{R} \to (0,\infty)$ be defined by $f(x) = e^x$. In other words, $f = \{(x, e^x) : x \in \mathbb{R}\}$. Then f is bijective, so f^{-1} exists, and $f^{-1} : (0,\infty) \to \mathbb{R}$ has rule $f^{-1} = \{(e^x, x) : x \in \mathbb{R}\}$. Another way to state the rule for f^{-1} is to write $f^{-1}(z) = \ln(z)$.*

Warning 645

Not every function has an inverse. Only bijective functions have inverses. If you have a function f, you cannot immediately speak of "the inverse of f." You can only talk about the inverse of f if you have proved that f is bijective, or you were told to assume that f is bijective.

Example 646. *The function from Example 555 not bijective, so does not have an inverse. The function from Example 635 is bijective, thus has an inverse.*

Warning 647: Multiple meanings for f^{-1}

Let f be a function from A to B. If you encounter the notation f^{-1}, you should not automatically think of "the inverse function." In fact, if we examine Definition 642 closely, f has an inverse only if f is bijective. If f is not bijective, then f^{-1} is not referring to the inverse function (because there *is* no inverse function in this case)!

What then? Keep in mind that we had two definitions of preimage. (See Definition 580 and 581.) If $y \in B$ and f is not bijective (or not *known* to be bijective), then only Definition 580 applies, and $f^{-1}(y)$ is a subset of the domain A. If f is known to be bijective, then $f^{-1}(y)$ may be read using Definition 580 to obtain a subset of A, or $f^{-1}(y)$ may be read using Definition 642, obtaining an element of A.

Example 648. *Consider the function $f : \{1,2,3\} \to \{4,5,6\}$ by the rule $f(1) = 4$ and $f(2) = 4$ and $f(3) = 5$. Since f is not bijective, the only possible interpretation of $f^{-1}(4)$ is to look at Definition 580 and get $f^{-1}(4) = \{1,2\}$. Since f is not bijective, the only possible interpretation of $f^{-1}(5)$ is to look at Definition 580 and get $f^{-1}(5) = \{3\}$. Note, $f^{-1}(5) \neq 3$. Since f is not bijective, we may not apply Definition 642.*

Example 649. *The function $f : \{1,2,3\} \to \{4,5,6\}$ defined in Example 635 is bijective. When applying Definition 580, $f^{-1}(4) = \{2\}$, a subset of $\{1,2,3\}$. When applying Definition 642, $f^{-1}(4) = 2$, an element of $\{1,2,3\}$.*

Exercise 650. *Let $A = \{1,2,3,4,5,6,7,8,9,10\}$. Let F be the set of all functions from A to $\mathbb{R}$. Define the binary relation $\sim$ on F by the following rule:*

$$\text{For } f \text{ and } g \text{ in } F, \text{ we say } f \sim g \text{ if } 0 \leq \sum_{x=1}^{10} f(x)g(x).$$

Prove that $\sim$ is reflexive.

Before showing the proof, it is important to stop and understand each thing being defined before moving on to the next thing. What is F? Note F is a set. A set of what? (See Habit 553.) The elements of F are functions. (Said differently, F is a set of functions.) Can you come up with examples of elements in F? For example, let $f : A \to \mathbb{R}$ be defined by

$$f = \{(1,3),(2,4),(3,2098),(4,-3847),(5,-2),(6,-1),(7,\pi^2),(8,103),(9,-123456789),(10,\tfrac{2}{6})\}.$$

Then $f \in F$. As another example, let $g : A \to \mathbb{R}$ be defined by

$$g = \{(1,0),(2,0),(3,0),(4,0),(5,0),(6,0),(7,0),(8,0),(9,0),(10,0)\}.$$

Then $g \in F$.

Proof. Let $f \in F$ be arbitrary. We want to show $f \sim f$. In other words, we need to show that

$$\sum_{i=1}^{10} f(i)f(i) \geq 0.$$

Note that the sum is really

$$f(1)f(1) + f(2)f(2) + \cdots + f(10)f(10)$$

which can be rewritten

$$[f(1)]^2 + [f(2)]^2 + \cdots [f(10)]^2.$$

Since we are squaring real numbers, each term above is greater than or equal to zero. Therefore, their sum is also non-negative. Thus,

$$\sum_{i=1}^{10} f(i)f(i) \geq 0.$$

which proves $f \sim f$. □

Exercise 651. *Let $f : X \to M$. Prove that the range of f is a subset of M.*

Exercise 652. *Suppose $f : \mathbb{R} \to \mathbb{R}$ satisfies $f(x+y) = f(x) + f(y)$ for all real numbers x and y. Prove that $f(0) = 0$.*

Exercise 653. *Suppose $f : \mathbb{R} \to \mathbb{R}$ satisfies $f(x+y) = f(x) + f(y)$ for all real numbers x and y. Prove for all $x \in \mathbb{R}$, the equation $f(-x) = -f(x)$ holds.*

Exercise 654. *Let $f : A \to B$. Prove that $A = f^{-1}(B)$.*

Exercise 655. *Let $f : X \to Y$ and $g : Y \to Z$. Prove: if $g \circ f$ is one-to-one and f is onto, then g is one-to-one.*

Exercise 656. *Let $f : X \to Y$ and $g : Y \to Z$. Prove: if $g \circ f$ is onto and g is one-to-one, then f is onto.*

Exercise 657. *Let $f : A \to B$. Prove that f is surjective if and only if for all $Y \subseteq B$, the equation $Y = f(f^{-1}(Y))$ holds.*

Exercise 658. *Let $f : A \to B$ and let $X \subseteq A$. Prove $X \subseteq f^{-1}(f(X))$.*

Exercise 659. *Let $f : A \to B$ and let $W, X \subseteq A$. Prove $f(W \cap X) \subseteq f(W) \cap f(X)$.*

Exercise 660. *Let $f : A \to B$ and let $W, X \subseteq A$. Prove $f(W \cup X) = f(W) \cup f(X)$.*

Exercise 661. *Let $f : A \to B$ and let $Y, Z \subseteq B$. Prove $f^{-1}(Y \cap Z) = f^{-1}(Y) \cap f^{-1}(Z)$.*

Exercise 662. *Let $f : A \to B$ and let $Y, Z \subseteq B$. Prove $f^{-1}(Y \cup Z) = f^{-1}(Y) \cup f^{-1}(Z)$.*

Exercise 663. *Let $f : A \to A$ and $g : A \to A$ both be bijective functions. Prove $(f \circ g)^{-1} = g^{-1} \circ f^{-1}$.*

Exercise 664. *Let $f : X \to Y$ be a function with $S \subseteq X$. Prove that $S \subseteq f^{-1}(f(S))$.*

Exercise 665. *Let $f : X \to Y$ be a function with $T \subseteq Y$. Prove that $f(f^{-1}(T)) \subseteq T$.*

Exercise 666. *Let $f : X \to Y$. Prove that f is one-to-one if and only if for all $y \in f(X)$, there exists a unique $x \in X$ such that $f(x) = y$.*

Exercise 667. *Given any function $h : \mathbb{R} \to \mathbb{R}$, let us define $U(h) = \{x \in \mathbb{R} : h(x) = 1\}$ and $S(h) = \{x \in \mathbb{R} : (h(x))^2 = 1\}$. Prove: for any function h from $\mathbb{R}$ to $\mathbb{R}$, the set $U(h)$ is a subset of $S(h)$.*

Exercise 668. *If A is a set of functions from $\mathbb{R}$ to $\mathbb{R}$, let us define $Z(A) = \{x \in \mathbb{R} : h(x) = 0$ for all $h \in A\}$. Let U be a set of functions from $\mathbb{R}$ to $\mathbb{R}$. Let V be a set of functions from $\mathbb{R}$ to $\mathbb{R}$ as well. Prove: if $U \subseteq V$, then $Z(V) \subseteq Z(U)$.*

Note that the exercise above ends $Z(V) \subseteq Z(U)$, which is correct: do not accidentally read this as $Z(U) \subseteq Z(V)$.

Exercise 669. *Let F be the set of all functions from $A = \{1, 2, \ldots, 225\}$ to $\mathbb{R}$. If $r, s \in F$, then we define $r * s$ to be*

$$r * s = \sum_{i=1}^{225} r(i)s(i).$$

Using C and D defined by

$$C = \{f \in F : f(a) \geq 0 \text{ for all } a \in A\} \quad \text{and} \quad D = \{f \in F \mid f * g \geq 0 \text{ for all } g \in C\},$$

prove that $C = D$.

Exercise 670. *Let $f : X \to Y$ be a function. Let A, B, and C be subsets of X. Prove: if $A \subseteq B$, then $f(A) \subseteq f(B)$. Then prove $f(B \cap C) \subseteq f(C)$.*

Exercise 671. *Prove: If $f : X \to Y$ is injective and C and D are any subsets of X, then $f(C) \cap f(D) = f(C \cap D)$.*

Exercise 672. *Let $f : \mathbb{Z} \to \mathbb{R}$ be given by the rule $f(x) = x^3 + \sqrt{1 + x^2}$. Explain the difference between f and $f(x)$. As a hint, think about what kind of thing f is, and what kind of thing $f(10)$ is.*

Chapter 5

Additional proof topics

5.1 Cardinality

Definition 673: Equicardinality

Two sets A and B are **equicardinal** if there exists a bijection $f : A \to B$.

Informally, we might say that A and B **have the same cardinality**. Informally, think of A and B as "having the same number of elements."

Method 674

If you are asked to prove that two sets have the same cardinality, informed by Definition 673, you must prove that there exists a bijection from A to B. To prove that a bijection $f : A \to B$ exists, informed by Remark 218, we should define a function $f : A \to B$ and prove that f is bijective.
Thus, we should define a function (according to Method 563) by specifying its domain, codomain, and rule. Then, use the rule to prove that f is bijective, following Method 638.

Theorem 675. *Let $E = \{2s : s \in \mathbb{Z}\}$. Then $\mathbb{Z}$ and E are equicardinal.*

Proof. Let $f : \mathbb{Z} \to E$ be the function defined by $f(z) = 2z$. Note that for each $z \in \mathbb{Z}$, we have $f(z) \in E$.

To prove that f is surjective, let $m \in E$. Then there exists an $s \in \mathbb{Z}$ such that $m = 2s$. Since $f(s) = 2s = m$, we have found an element in $\mathbb{Z}$ (namely s) such that $f(s) = m$. Thus f is surjective.

To prove that f is injective, let $a, b \in \mathbb{Z}$ be arbitrary. Suppose $f(a) = f(b)$. Then $2a = 2b$. By division, $a = b$. So f is injective.

Since f is a bijection from A to B, the sets $\mathbb{Z}$ and E are equicardinal. □

It is peculiar that the set $\mathbb{Z}$ is equicardinal to its proper subset E. It seems strange that $E \subsetneq \mathbb{Z}$, yet $\mathbb{Z}$ and E have "the same number of elements."

Definition 676: Countable

A set A is **countable** if A is finite or if $\mathbb{Z}_{>0}$ and A are equicardinal.

Definition 677: Uncountable

A set A is **uncountable** if A is not countable.

Due to De Morgan's Law, A is uncountable if and only if A is not finite, and $\mathbb{Z}_{>0}$ and A are not equicardinal.

Definition 678: Countably infinite

A set A is **countably infinite** if A is countable and A is not finite.

Theorem 679. *Let A and B be disjoint sets. Suppose A is finite. Suppose B is countably infinite. Then $A \cup B$ is countable.*

Imagine that you showed up to work at your job. (You help serve customer inquiries at the deli counter of your local grocery store.) Suppose that two lines had formed. The short line had seven people. The long line had people as far as the eye can see. In what order could you help people to ensure that everybody was served? Start with the line with seven people. Help all of them first. Then, you can help the remaining people in the very long line in the order in which they are standing. Think of the short line representing the set A in the statement of the theorem, and B representing the very long line. In the proof below, we will quickly introduce (and use) a bijection f from $\mathbb{Z}_{>0}$ to B, which represents the order in which the people are standing. In fact, if B were a set of people, think of $f(1)$ as the first person in line, $f(2)$ as the second person in line, and so on.

Proof. Let A and B be disjoint sets. Suppose A is finite. Since there exists a non-negative integer n such that A has exactly n elements, let $a_1, \ldots, a_n$ denote the n distinct elements of A. Thus, $A = \{a_1, \ldots, a_n\}$. Suppose B is countably infinite. Thus, there is a bijection $f : \mathbb{Z}_{>0} \to B$. To prove $A \cup B$ is countable, we define the function $g : \mathbb{Z}_{>0} \to A \cup B$ by the rule

$$g(x) = \begin{cases} a_x & \text{if } x \leq n \\ f(x-n) & \text{if } x > n. \end{cases}$$

To prove that g is surjective, let $y \in A \cup B$ be arbitrary. Then $y \in A$ or $y \in B$, which leads to two cases:

- Suppose $y \in A$. Then $y = a_i$ for some $i \in \{1, \ldots, n\}$. Let $x = i$. Then $g(x) = a_x = a_i = y$.
- Suppose $y \in B$. Since $f : \mathbb{Z}_{>0} \to B$ is surjective, there exists $m \in \mathbb{Z}_{>0}$ such that $f(m) = y$. Let $x = m + n$. Then $x > n$, and $g(x) = g(m+n) = f((m+n)-n) = f(m) = y$.

In both cases, $g(x) = y$, so g is surjective.

To prove that g is injective, let $u, w \in \mathbb{Z}_{>0}$ be arbitrary. Suppose $g(u) = g(w)$. We will prove that $u = w$. Either $u \leq n$ or $u > n$. Similarly, $w \leq n$ or $w > n$. Then, one of the following must occur

- $u \leq n$ and $w \leq n$
- $u \leq n$ and $w > n$
- $u > n$ and $w \leq n$
- $u > n$ and $w > n$

so we prove that $u = w$ in four cases:

- Suppose $u \leq n$ and $w \leq n$. From $g(u) = g(w)$ with u and w both less than or equal to n, we get $a_u = a_w$, and since $a_1, \ldots, a_n$ denoted the n distinct elements of A, we have $u = w$.
- Suppose $u \leq n$ and $w > n$. Then $g(u) \in A$ and $g(w) \in B$. But then $g(u) = g(w)$ is impossible, since A and B are disjoint.
- Suppose $u > n$ and $w \leq n$. Then $g(u) \in B$ and $g(w) \in A$. But then $g(u) = g(w)$ is impossible, since A and B are disjoint.
- Suppose $u > n$ and $w > n$. From $g(u) = g(w)$, we get $f(u-n) = f(w-n)$. Since f is injective, $u - n = w - n$. By addition, $u = w$.

Of the four cases, two cases turned out to be impossible. In the remaining two cases, we proved $u = w$. Thus g is injective.

Since g is both surjective and injective, g is a bijection. Since g is a bijection from $\mathbb{Z}_{>0}$ to $A \cup B$, the set $A \cup B$ is countable. $\square$

Some things about this proof may look bizarre, but let's use the deli counter to help. If it helps, replace each n in the proof with a 7. (However, it makes sense for the actual proof to not mention the number seven.) Think of $g(x)$ as representing the order in which you help each customer: $g(1)$ is the first customer you help. Following our example of $n = 7$, note $g(1) = a_1$, the first of seven people in the short line. In the same way, $g(2)$ is the second customer you help, and $g(7)$ is the seventh customer (the last person in the short line) whom you help. Who should be $g(8)$? This should be the first person in the second line. So, we want $g(8)$ to be assigned $f(1)$. We want $g(9)$ to be $f(2)$, and so on. If $x > 7$, we want $g(x) = f(x-7)$. Now, all we have to do is replace all the appearances of 7 with n: this helped us successfully define a function g. Since g was defined carefully (that is, we found a rule that works), it was possible to prove that g is bijective.

Theorem 680. *Let A and B be disjoint countably infinite sets. Then $A \cup B$ is countable.*

To prove this theorem, let's return to the deli counter story. Suppose the next day you go to work, there are two lines as long as the eye can see. How could you decide on an order in which to help people to ensure that everyone gets helped? It would not make sense to help everyone in one line first: then the people in the next line would never get helped! Helping all of the people in the second line first would have the same type of issue. If I were the tenth person in the first line and I saw an employee only helping the other line, I'd panic.

What is a reasonable compromise – one that would ensure everybody gets served at the deli counter? The employee could alternate: serve one person in one line, then serve a person in the other line. If I'm tenth in one line, I probably won't be the tenth person served. (In fact, I'd estimate being the 20th person served, or close to it: perhaps the 19th person or the 21st person.) However, I would realize that, if the deli counter employee was methodical about alternating lines, I'd get served! In the proof below, think of the function h which we define as helping set an order for serving customers. The fact that the h from $\mathbb{Z}_{>0}$ to $A \cup B$ which we define will be surjective translates to the idea that every customer in each line will get served eventually. (The analogy to injectivity is weirder: it says that no customer will be served multiple times, but it seemed implicitly built into our story that a customer who is served will immediately leave the store and not return.)

Proof. Let A and B be disjoint countably infinite sets. Since A is countably infinite, there is a bijection $f : \mathbb{Z}_{>0} \to A$. Similarly, there is a bijection $g : \mathbb{Z}_{>0} \to B$. Let us define $h : \mathbb{Z}_{>0} \to A \cup B$ by the rule

$$h(x) = \begin{cases} f(\frac{x+1}{2}) & \text{if } x \text{ is odd} \\ g(\frac{x}{2}) & \text{if } x \text{ is even.} \end{cases}$$

To prove h is surjective, let $z \in A \cup B$. Then $z \in A$ or $z \in B$. We have two cases:

- Suppose $z \in A$. Since f is surjective, there exists $m \in \mathbb{Z}_{>0}$ such that $f(m) = z$. Let $j = 2m - 1$. Then $j > 0$ and since $j = 2(m-1)+1$ is odd, $h(j) = f(\frac{j+1}{2}) = f(\frac{2m-1+1}{2}) = f(m) = z$.

- Suppose $z \in B$. Since g is surjective, there exists $m \in \mathbb{Z}_{>0}$ such that $g(m) = z$. Let $j = 2m$. Since j is even, $h(j) = g(\frac{j}{2}) = g(\frac{2m}{2}) = g(m) = z$.

In both cases there exists $j \in \mathbb{Z}_{>0}$ such that $h(j) = z$, so h is surjective.

We now prove h is injective. Let $r, s \in \mathbb{Z}_{>0}$ be arbitrary and suppose $h(r) = h(s)$. Either r is even or odd, and s is either even or odd. If we match the possibilities, we will have four cases:

- Suppose r and s are both even. Then $h(r) = g(\frac{r}{2})$ and $h(s) = g(\frac{s}{2})$. From our earlier supposition, $g(\frac{r}{2}) = g(\frac{s}{2})$. Since g is injective, $\frac{r}{2} = \frac{s}{2}$. By multiplication, $r = s$.

- Suppose r is even and s is odd. Then $h(r) \in B$ while $h(s) \in A$, and $h(r) = h(s)$ contradicts the fact that A and B are disjoint, so this case is impossible.

- Suppose r is odd and s is even. Then $h(r) \in A$ while $h(s) \in B$, and $h(r) = h(s)$ contradicts the fact that A and B are disjoint, so this case is impossible.

- Suppose r and s are both odd. Then $f(\frac{r+1}{2}) = f(\frac{s+1}{2})$. Since f is injective, $\frac{r+1}{2} = \frac{s+1}{2}$, and with some algebra, $r = s$.

In the cases which are possible, we proved $r = s$, so h is injective.

Since h is a bijection from $\mathbb{Z}_{>0}$ to $A \cup B$, its codomain $A \cup B$ is countably infinite. $\square$

Exercise 681. *Use the ideas of the proof of Theorem 680 to prove that $\mathbb{Z}$ is countably infinite. However, see if you can do so without having a function f, a function g, and a function h. Since you know the elements of $\mathbb{Z}$, see if you can define function h without making any reference to functions called f or g.*

Exercise 682. *Prove: Let A and B and C be disjoint sets. Suppose A and B are finite. Suppose C is countably infinite. Then $A \cup B \cup C$ is countable.*

Exercise 683. *Let A, B and C be pair-wise disjoint sets. (That is, $A \cap B = \emptyset$ and $A \cap C = \emptyset$ and $B \cap C = \emptyset$.) Suppose A, B, and C are countably infinite. Prove $A \cup B \cup C$ is countable.*

If you'd like a challenging problem, consider the exercise below:

Exercise 684. *Prove: Let A and B and C be disjoint sets. Suppose A is finite. Suppose B and C are countably infinite. Then $A \cup B \cup C$ is countable.*

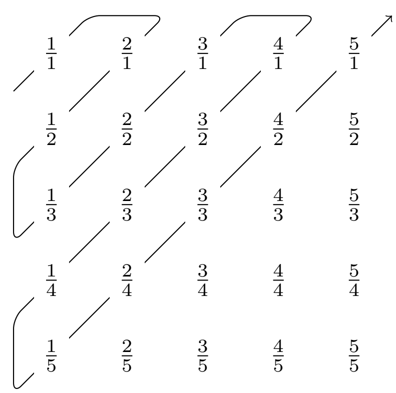
If you'd like a warm up to the previous question, try the two exercises below first:

Exercise 685. *Prove: Let A and B and C be disjoint sets. Suppose $|A| = 7$. Suppose B and C are countably infinite. Then $A \cup B \cup C$ is countable.*

Exercise 686. *Prove: Let A and B and C be disjoint sets. Suppose $|A| = 8$. Suppose B and C are countably infinite. Then $A \cup B \cup C$ is countable.*

Theorem 687. *The set $\mathbb{Q}_{>0}$ is countable.*

Proof. Consider the following picture:



Note that every positive rational number appears in this diagram. If p and q are positive integers, then $\frac{p}{q}$ appears in the qth row and pth column. Note also the shape of the path (following the direction of the arrow). We define a function h from $\mathbb{Z}_{>0}$ to $\mathbb{Q}_{>0}$ by assigning a value of h each time we encounter a rational number whose value we have not encountered before.

Thus, $h(1) = \frac{1}{1}$, and $h(2) = \frac{2}{1}$, and $h(3) = \frac{1}{2}$, and $h(4) = \frac{1}{3}$, but $h(5) \neq \frac{2}{2}$, because we have already encountered the value $\frac{2}{2}$ as $h(1)$. So, we skip over $\frac{2}{2}$ and instead say $h(5) = \frac{3}{1}$. Continuing, $h(6) = \frac{4}{1}$ and $h(7) = \frac{3}{2}$ and $h(8) = \frac{2}{3}$ and $h(9) = \frac{1}{4}$ and $h(10) = \frac{1}{5}$. Note that $h(11) \neq \frac{2}{4}$ because we have already encountered the value $\frac{2}{4}$ when we said that $h(3) = \frac{1}{2}$. So, instead, $h(11) = \frac{5}{1}$, since we also need to skip over the values $\frac{3}{3} = 1$ and $\frac{4}{2} = 2$. Continue defining the outputs of h in this way.

We see that h is bijective by design. The function h is surjective (every positive rational number appears in this diagram) and h is injective (we skip over any rational number whose value we have already encountered). $\square$

Theorem 688. *The set $\mathbb{Q}$ is countable.*

The proof of the previous theorem is left as an exercise. Prove this theorem using the previous theorem (that is, assume you have a bijection h from $\mathbb{Z}_{>0}$ to $\mathbb{Q}_{>0}$) and use the ideas which made Exercise 681 successful.

Exercise 689. *Prove the theorem above, that $\mathbb{Q}$ is countable.*

Theorem 690. *The set $\mathbb{R}$ is uncountable.*

Proof. The proof we present is known as Cantor's Diagonalization Argument. To obtain a contradiction, suppose that $\mathbb{R}$ is countable. Then, $[0,1]$, which is a proper subset of $\mathbb{R}$ would also be countable. Since $[0,1]$ is not finite, then $[0,1]$ is countably infinite, thus there exists a bijection $f : \mathbb{Z}_{>0} \to [0,1]$.

Now, each element in $[0,1]$ has a decimal expansion of the form $0.d_1d_2d_3d_4d_5\dots$. For example, $\frac{1}{4} = 0.25000\dots$. We will write out the decimal expansion of $f(1)$ as $0.d_{11}d_{12}d_{13}d_{14}d_{15}$..., and $f(2)$ as $0.d_{21}d_{22}d_{23}d_{24}d_{25}$..., and so on. More generally, $f(j) = 0.d_{j1}d_{j2}d_{j3}d_{j4}d_{j5}\dots$, and thus $d_{jk} \in \{0,1,2,\dots,0\}$ is the kth digit in the decimal expansion of $h(j)$.

We will construct a real number ℓ whose decimal expansion we denote $\ell = 0.\ell_1\ell_2\ell_3\ell_4\ell_5\dots$ in the following manner:

$$\ell_j = \begin{cases} 8 & \text{if } d_{jj} = 5 \\ 5 & \text{if } d_{jj} \neq 5. \end{cases}$$

That is, if the first digit of $f(1)$ is a 5 then the first digit of ℓ is 8, and if the first digit of $f(1)$ is not a 5 then the first digit of ℓ is 5. Likewise, if the second digit of $f(2)$ is a 5 then the second digit of ℓ is 8, and if the second digit of $f(2)$ is not a 5 then the second digit of ℓ is 5. In this manner, ℓ differs from $f(j)$ in the jth digit, so for all $j \in \mathbb{Z}_{>0}$, we have $\ell \neq f(j)$. But $\ell \in [0,1]$, so this proves that f is not surjective, a contradiction to f being bijective. ☐

The work of equicardinality involves the following: given two sets A and B, define a function from A to B that is bijective. An important intuition-building exercise is to consider when $A = \mathbb{R}$ and $B = \mathbb{R}$:

Exercise 691. *In this exercise, there is not* one *correct answer:*

1. *Define a function from $\mathbb{R}$ to $\mathbb{R}$ which is injective but not surjective.*

2. *Define a function from $\mathbb{R}$ to $\mathbb{R}$ which is surjective but not injective.*

3. *Define a function from $\mathbb{R}$ to $\mathbb{R}$ which is injective and surjective.*

4. *Define a function from $\mathbb{R}$ to $\mathbb{R}$ which is neither injective nor surjective.*

Theorem 692. *The subsets $A = [3,7]$ and $B = [5,20]$ of $\mathbb{R}$ are equicardinal.*

First, note that 3 is the smallest input and 5 is the smallest output. Second, note 7 is the largest input, while 20 is the largest output. One can write $y - 5 = \frac{15}{4}(x-3)$ as an equation for the line going through the points $(3,5)$ and $(7,20)$. This was how we determined the rule for the function h in the proof to be $h(x) = \frac{15}{4}(x-3)+5$. Now, we prove the theorem:

Proof. Let $h : [3,7] \to [5,20]$ be the function defined by the rule $h(x) = \frac{15}{4}(x-3)+5$.

First we ensure that $[5,20]$ as stated for the domain is accurate. In other words, we would have a "problem" if the value of $\frac{15}{4}(x-3)+5$ is *not* between 5 and 20 for any x-value between 3 and 7.

- Note that $h(3) = \frac{15}{4}(3-3)+5 = 0+5 = 5$.

- Note that $h(7) = \frac{15}{4}(7-3)+5 = 15+5 = 20$.

- Note $h'(x) = \frac{15}{4}$. Since $h'(x) > 0$ for all $x \in [3,7]$, the function h is increasing on this interval, and with $h(3) = 5$ and $h(7) = 20$, this shows that $h(x)$ is in $[5,20]$ for all $x \in [3,7]$.

Therefore, the stated codomain of $[5,20]$ is acceptable, following the requirement mentioned in Warning 566.

To prove h is injective, let $a, b \in [3,7]$ both be arbitrary. Suppose $h(a) = h(b)$. Then $\frac{15}{4}(a-3)+5 = \frac{15}{4}(b-3)+5$. So $\frac{15}{4}(a-3) = \frac{15}{4}(b-3)$. Multiplying both sides by $\frac{4}{15}$, we learn $a-3 = b-3$. Therefore $a = b$, and h is injective.

To prove h is surjective, let $z \in [5,20]$ be arbitrary. We want to prove that there exists $m \in [3,7]$ such that $h(m) = z$.

Scratch work. Leave everything in this box out of the proof. Want $h(m) = z$. In other words, want $\frac{15}{4}(m-3)+5 = z$. Solving for m, we get $m = 3 + \frac{4}{15}(z-5)$.

Let $m = 3 + \frac{4}{15}(z-5)$. We will prove $m \in [3,7]$ and we will prove $h(m) = z$.

- To prove $m \in [3,7]$, note $5 \le z \le 20$. By subtracting, $0 \le z-5 \le 15$. By multiplication (by a positive), $0 \le \frac{4}{15}(z-5) \le 4$. By addition, $3 \le 3+\frac{4}{15}(z-5) \le 7$. We replace the quantity in the middle with the definition of m, so $3 \le m \le 7$, which proves $m \in [3,7]$.

- To prove $h(m) = z$, note

$$\begin{aligned} h(m) &= \tfrac{15}{4}(m-3)+5 \\ &= \tfrac{15}{4}([3+\tfrac{4}{15}(z-5)]-3)+5 \\ &= \tfrac{15}{4}\cdot\tfrac{4}{15}(z-5)+5 \\ &= (z-5)+5 \\ &= z, \end{aligned}$$

  and since $m \in [3,7]$ and $h(m) = z$, we have proved that there exists $m \in A$ such that $h(m) = z$. Since the choice of $z \in B$ was arbitrary, h is surjective.

Since h is bijective, $[3,7]$ and $[5,20]$ are equicardinal. □

Given $A = [3,7]$ and $B = [5,20]$, what was the intuition behind using the rule $h(x) = \frac{15}{4}(x-3)+5$? Both A and B are subsets of $\mathbb{R}$, but since A is the domain, imagine thickening the portion of the x-axis where $3 \le x \le 7$. Similarly, thicken the portion of the y-axis where $5 \le y \le 20$. The line through the points $(3,5)$ and $(7,20)$ would be the graph of a bijective function from $\mathbb{R}$ to $\mathbb{R}$. An equation for that line is $y-5 = \frac{15}{4}(x-3)$ in point-slope form. The line *segment* with endpoints $(3,5)$ and $(7,20)$ would be the graph of a bijective function from A to B.

Exercise 693. *Let A, B, C, and D be sets. Suppose A and C have the same cardinality. Suppose B and D have the same cardinality. Prove $A \times B$ has the same cardinality as $C \times D$.*

Exercise 694. *Let A and B be sets. Prove that if A and B have the same cardinality, then $P(A)$ and $P(B)$ have the same cardinality.*

Exercise 695. *Let $f : X \to Y$ be an injective function. Prove that X and $f(X)$ have the same cardinality.*

Exercise 696. *Prove that $(3,7)$ and $(5,12)$ have the same cardinality.*

Exercise 697. *Prove that the set $[2,7] = \{a \in \mathbb{R} : 2 \le a \le 7\}$ and the set $[3,14] = \{p \in \mathbb{R} : 3 \le p \le 14\}$ are equicardinal.*

Exercise 698. *Prove the relation "A has the same cardinality as B" is an equivalence relation on sets.*

Exercise 699. *Prove: if B is countable and $A \subseteq B$, then A is countable.*

Exercise 700. *Prove: if A and B are countable, then $A \cap B$ is countable. (You cannot assume that A and B are disjoint.) For good practice, prove this statement without using the previous exercise.*

5.2 Induction

Let $P(n)$ be a predicate. The traditional way (Method 268) to prove $\boxed{\forall k \in \mathbb{Z}_{>0},\, P(k)}$ is to select $k \in \mathbb{Z}_{>0}$ arbitrarily and work to prove $P(k)$. Sometimes, one can get stuck trying this traditional proof method. The method of induction provides an alternate means for proving $P(k)$ is true for all $k \in \mathbb{Z}_{>0}$.

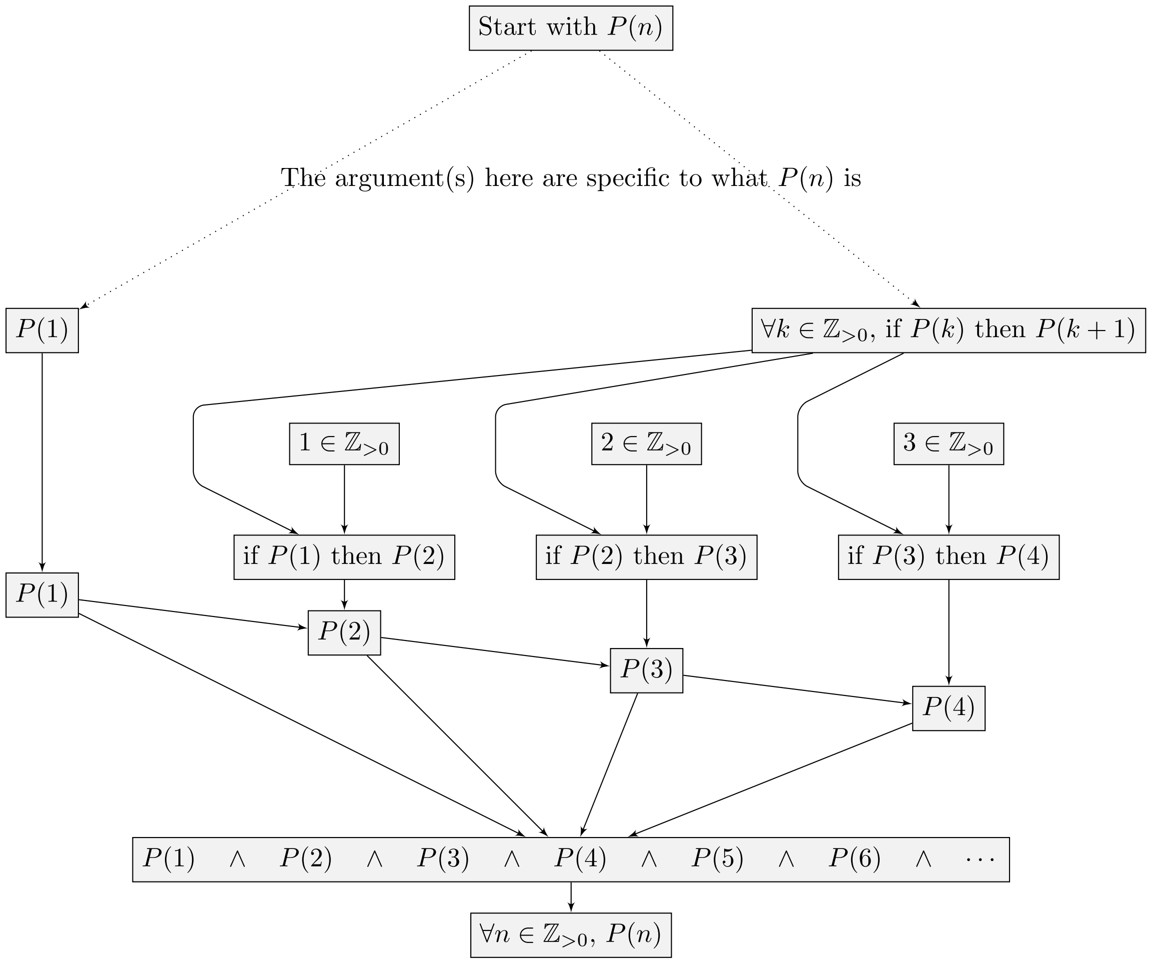


Figure 5.1: Justification of proof by induction

Method 701: Proof by induction

To prove $\forall n \in \mathbb{Z}_{>0}$, $P(n)$ by induction, one must prove:

- **base case**: prove $P(1)$
- **inductive step**: prove $\forall k \in \mathbb{Z}_{>0}$, if $P(k)$ then $P(k+1)$

A proof by induction consists of the proof writer filling in the details for dotted arrows in Figure 5.1, taking the defined $P(n)$ and establishing $P(1)$ is true (**base case**), and $\boxed{\forall k \in \mathbb{Z}_{>0}, \text{ if } P(k) \text{ then } P(k+1)}$ is true (**inductive step**). If those are established, the generic part of the argument (called the **Principle of Mathematical Induction**) is the bottom half of the flowchart below, and is not written out in standard proofs, because the argument is always the same.

Let us consider an example:

Theorem 702. *The identity* $\dfrac{n(n+1)}{2} = \displaystyle\sum_{j=1}^{n} j$ *holds for all* $n \in \mathbb{Z}_{>0}$.

Proof. We provide more detail than is typical, as this is our first induction proof. First, note that $P(n)$ is the predicate $\dfrac{n(n+1)}{2} = \displaystyle\sum_{j=1}^{n} j$.

For the base case, we need to prove $P(1)$. In other words, we need to prove $\dfrac{1(1+1)}{2} = \displaystyle\sum_{j=1}^{1} j$. Note that $\sum_{j=1} 1j = 1 = \frac{2}{2} = \frac{1(1+1)}{2}$, so the base case is proved.

The inductive step is to prove $\forall k \in \mathbb{Z}_{>0}$, if $P(k)$ then $P(k+1)$. As we traditionally would do following Method 268, we let $k \in \mathbb{Z}_{>0}$ be arbitrary. We want to prove if $P(k)$ then $P(k+1)$. So, assume $\dfrac{k(k+1)}{2} = \displaystyle\sum_{j=1}^{k} j$. We want to prove $\dfrac{(k+1)((k+1)+1)}{2} = \displaystyle\sum_{j=1}^{k+1} j$.

How does the statement that we want to prove, namely $\dfrac{(k+1)((k+1)+1)}{2} = \displaystyle\sum_{j=1}^{k+1} j$, relate to the statement we assumed was true, namely $\dfrac{k(k+1)}{2} = \displaystyle\sum_{j=1}^{k} j$? Notice that $\displaystyle\sum_{j=1}^{k+1} j = (k+1) + \sum_{j=1}^{k} j$, and note that $\sum_{j=1}^{k} j$ can be replaced with $\frac{k(k+1)}{2}$. So,

$$\begin{aligned}
\sum_{j=1}^{k+1} &= \left[\sum_{j=1}^{k} j\right] + (k+1) \\
&= \frac{k(k+1)}{2} + (k+1) \\
&= \frac{k^2+k}{2} + \frac{2k+2}{2} \\
&= \frac{k^2+3k+2}{2} \\
&= \frac{(k+1)(k+2)}{2} \\
&= \frac{(k+1)((k+1)+1)}{2},
\end{aligned}$$

and thus the inductive step is proved.

Using mathematical induction, for all $n \in \mathbb{Z}_{>0}$, we have $\sum_{j=1}^{n} j = \frac{n(n+1)}{2}$. □

For a shortened version of the proof:

Proof. We prove by induction. For the base case, $\sum_{j=1} 1j = 1 = \frac{2}{2} = \frac{1(1+1)}{2}$. For the inductive step, $k \in \mathbb{Z}_{>0}$ be arbitrary, and assume $\dfrac{k(k+1)}{2} = \displaystyle\sum_{j=1}^{k} j$. So,

$$\sum_{j=1}^{k+1} j = \left[\sum_{j=1}^{k} j\right] + (k+1) = \frac{k(k+1)}{2} + (k+1) = \frac{(k+1)(k+2)}{2}.$$

□

When we talk about **mathematical induction** or **induction**, we mean this proof method. The word "induction" is used in everyday language to mean "making a conclusion that (vaguely) seems likely to be true" based on repeated observation. We do not mean this at all.

In the proof above, $P(k)$ is called the **inductive hypothesis**. In our specific example, the inductive hypothesis was the assumed statement $\frac{k(k+1)}{2} = \sum_{j=1}^{k} j$.

Theorem 703. *For all positive integers n, the integer 3 divides $n^3 + 2n + 9$.*

Before proving Theorem 703, some comments are in order. Note that $P(n)$ is the predicate "3 divides $n^3 + 2n + 9$."

Warning 704

If you use $P(n)$ for a predicate, but $P(n)$ includes a function p, do not say that $p(n)$ is true!

Example 705. *We will not say that $P(n)$ is n^3+2n+9. If we'd like to, we can define a function $p : \mathbb{Z}_{>0} \to \mathbb{Z}$ by saying that $p(n) = n^3+2n+9$, but if we do so, let us not confuse $P(n)$ with $p(n)$. Here, $p(n)$ is a function and $P(n)$, following the notation from earlier, is a predicate. In fact, $P(n)$ is the statement that 3 divides the number $p(n)$. While we could speak of $P(5)$ being true, we could not speak of $p(5)$ being true, since $p(5)$ is just a number. Similarly, we cannot talk about $p(n)$ being true or false.*

We now prove Theorem 703.

Proof. For our base case, we need to prove 3 divides $1^3 + 2(1) + 9$. Let $c = 4 \in \mathbb{Z}$. Then $3c = 3 \cdot 4 = 12 = 1 + 2 + 9 = 1^3 + 2(1) + 9$. Since $c \in \mathbb{Z}$ and $3c = 1^3 + 2(1) + 9$, we have proved 3 divides $1^3 + 2(1) + 9$.

For the inductive step, let $k \in \mathbb{Z}_{>0}$ be arbitrary. Suppose 3 divides the integer $k^3 + 2k + 9$. We want to show 3 divides $(k + 1)^3 + 2(k + 1) + 9$. Since 3 divides $k^3 + 2k + 9$, there exists an integer w such that $3w = k^3 + 2k + 9$. Let $r = w + k^2 + k + 1$. Then,

$$\begin{aligned}
(k+1)^3 + 2(k+1) + 9 &= k^3 + 3k^2 + 3k + 1 + 2k + 2 + 9 \\
&= k^3 + 2k + 9 + 3k^2 + 3k + 3 \\
&= 3w + 3k^2 + 3k + 3 \\
&= 3(w + k^2 + k + 1) \\
&= 3r.
\end{aligned}$$

Since $r \in \mathbb{Z}$ and $3r = (k + 1)^3 + 2(k + 1) + 9$, this proves that 3 divides $(k + 1)^3 + 2(k + 1) + 9$, completing the inductive step. □

Exercise 706. *Prove: for all $n \in \mathbb{Z}_{>0}$, the equality $\displaystyle\sum_{i=1}^{n} i^2 = \frac{n(n+1)(2n+1)}{6}$ holds.*

Exercise 707. *Prove: $n! < n^n$ for all integers $n > 1$.*

Exercise 708. *Prove $1 + 3 + 5 + \cdots (2n - 1) = n^2$ for all $n \in \mathbb{Z}_{>0}$.* [key]

Exercise 709. *Prove for all $n > 0$ which are integer, one has*

$$\sum_{j=1}^{n} \frac{1}{j(j+1)} = \frac{n}{n+1}$$

[key]

Exercise 710. *Suppose $f : \mathbb{R} \to \mathbb{R}$ satisfies $f(x + y) = f(x) + f(y)$ for all real numbers x and y. Fix a real number s. Prove for all $n \in \mathbb{Z}_{>0}$, the equation $f(ns) = n\,f(s)$ holds.* [key]

Exercise 711. *Prove 5 divides $6^n + 4$ for all $n \in \mathbb{Z}_{>0}$.* [key]

Exercise 712. *Prove 4 divides $n(n+2)$ if n is any even positive integer. (Hint: create for yourself a modified type of mathematical induction)* [key]

Exercise 713. *Prove 8 divides $3^{2k} - 1$ for all $k \in \mathbb{Z}_{>0}$.* [key]

Exercise 714. *Prove for every positive integer n, the integer 9 divides $n^3 + (n + 1)^3 + (n + 2)^3$.*

5.2.1 Strong induction

Let $P(n)$ be a predicate. Perhaps the traditional way (Method 268) to prove $\forall k \in \mathbb{Z}_{>0}$, $P(k)$ fails to work, and perhaps the method of mathematical induction just introduced in the previous section also fails to work. Proof by strong induction provides yet another means for proving $P(k)$ is true for all $k \in \mathbb{Z}_{>0}$.

Method 715: Proof by strong induction

To prove $\forall n \in \mathbb{Z}_{>0}$, $P(n)$ using strong induction, one must prove:

- **base case**: prove $P(1)$
- **inductive step**: prove for all $k \in \mathbb{Z}_{>0}$, if $P(1) \wedge P(2) \wedge \cdots \wedge P(k)$, then $P(k+1)$.

Here, the **inductive hypothesis** is $P(1) \wedge P(2) \wedge \cdots \wedge P(k)$, where as earlier, the inductive hypothesis (in standard induction) was just $P(k)$.

Before looking at an example, we will need a definition.

Definition 716

Let f_n denote the **Fibonacci sequence**, defined by

$$\begin{aligned} f_1 &= 1 \\ f_2 &= 1 \\ f_n &= f_{n-1} + f_{n-2}, \quad \text{if } n > 2. \end{aligned}$$

Theorem 717. *If f_n denotes the nth Fibonacci number, then for all $n \in \mathbb{Z}_{>0}$, the equation $f_{n+6} = 4f_{n+3} + f_n$ is true.*

Proof. Let $P(n)$ be the predicate $f_{n+6} = 4f_{n+3} + f_n$. We will proceed with a proof by strong induction.

For the base case, we want to show that $P(1)$ is true. Since $f_7 = 13$ and $4f_4 + f_1 = 4(3) + 1 = 13$, we see that $P(1)$ holds.

For the inductive step, let $k \in \mathbb{Z}_{>0}$ be arbitrary. Suppose that $P(1)$, $\ldots$, and $P(k)$ are all true (the inductive hypothesis). We want to show that $P(k+1)$ is true. In other words, we want to show that $f_{k+7} = 4f_{k+4} + f_{k+1}$ holds. We have

$$\begin{aligned} f_{k+7} &= f_{k+6} + f_{k+5} \\ &= (4f_{k+3} + f_k) + (4f_{k+2} + f_{k-1}) \\ &= 4(f_{k+3} + f_{k+2}) + (f_k)f_{k-1}) \\ &= 4f_{k+4} + f_{k+1}, \end{aligned}$$

where the first and last equalities come from the recursive definition of the Fibonacci sequence, and the second equation are from $P(k-1)$ and $P(k)$, which came from the inductive hypotheses. $\square$

Exercise 718. *Let $f_0 = 1, f_1 = 1$, and if $n > 1$, then $f_n = f_{n-1} + f_{n-2}$. Let $\phi = \frac{1}{2}(1 + \sqrt{5})$. Prove that $f_n \leq \phi^n$ for all $n \in \mathbb{Z}_{\geq 0}$.* [key]

Exercise 719. *Let $g : \mathbb{Z}_{>0} \to \mathbb{Z}$ be the function defined by the rule*

$$g(n) = \begin{cases} 1 & \text{if } n = 1, \\ 5 & \text{if } n = 2, \\ 5g(n-1) - 6g(n-2) & \text{if } n > 2. \end{cases}$$

Prove: $g(n) = 3^n - 2^n$ for all $n \in \mathbb{Z}_{>0}$.

Chapter 6

Counting

This chapter investigates the principles of mathematical counting. The chapter largely does not rely on much of the content of the previous chapters, though there is some language of sets and functions which will appear. The word "counting" makes it sound like we will walk through "one, two, three, four." Don't worry: our counting problems will be more interesting than that.

6.1 The Product and Sum Rules

The Product Rule does not always apply, but when it does, it takes a more complex task (which we call a procedure) and counts how many ways there are to do the procedure in terms of simpler tasks:

Method 720: Product Rule

Suppose that a procedure can be broken down into a sequence of two independent tasks. If there are n_1 ways to do the first task and n_2 ways to do the second task, then there are $n_1 n_2$ ways to do the procedure.

To justify this rule which claims that there are $n_1 n_2$ ways to do the procedure, consider the following set up. Let F be a set consisting of n_1 elements: the elements of F are precisely the n_1 ways to do the first task. Let S be a set consisting of n_2 elements: the elements of S are precisely the n_2 ways of doing the second task. The procedure is done by selecting one of the n_1 options for the first task, followed by selecting one of the n_2 options for the second task. Then $|F \times S| = |F| \times |S|$, and each element of $F \times S$ is an ordered pair, which represents a different way to do the procedure.

When does the Product Rule apply? The Product Rule applies when there are two *independent* tasks, and *both* tasks must be performed.

Example 721. *In a certain psychology experiment, to protect the identities of the subjects, each participant is to be identified only using a pairing of a letter and a number. How many possible participant identifiers are there?*

The procedure of determining a participant identifier requires a first task of choosing one of $n_1 = 26$ *letters, and a second task of choosing one of* $n_2 = 10$ *digits (from* 0 *through* 9*, inclusive). The procedure is incomplete just doing one task or the other: both tasks must be completed. Thus, the total number of participant identifiers is* 260*.*

Example 722. *In Spain, a popular way to do lunch is called the menú del dia. A typical restaurant's menú del dia will consist of the diner selecting exactly one choice from the "first plates" list, exactly one item from the "second plates" list, and one item from the "desserts" list, with the price fixed to be constant (independent of the choices made). Suppose one restaurant offers the following choices: First plates are sopa castellana, sopa de pescado, or djudias blancas. Second plates are cordero asado, pescadilla a la romana, chuleta de aguja, or trucha ala navarra. Dessert choices are crema catalana or helado.*

How many different meals can be made? Since there are 3 *choices for first plate,* 4 *choices for second plate, and* 2 *choices for desert, there are* $3 \cdot 4 \cdot 2 = 24$ *possible meals. (That is,* 24 *visitors to this restaurant can have unique dining experiences. We have avoided using the word "combination" which has a specific meaning in a future section.)*

In cases where the Product Rule does not apply, the Sum Rule might apply. Like the Product Rule, the Sum Rule takes a more complex task (which we call a procedure) and counts how many ways to do this procedure in terms of simpler tasks.

Method 723: Sum Rule

Suppose a procedure is done by picking exactly one of two tasks to do. If there are n_1 ways to do the first task and n_2 ways to do the second task, then there are $n_1 + n_2$ ways to do the procedure.

To justify this rule which claims that there are $n_1 + n_2$ ways to do the procedure, consider the following set up. Let F be a set consisting of n_1 elements: the elements of F are precisely the n_1 ways to do the first task. Let S be a set consisting of n_2 elements: the elements of S are precisely the n_2 ways of doing the second task. If there are truly n_1 ways to do the first task and n_2 ways to do the second task, then $F \cup S$ represents the set of ways to do exactly one task, assuming F and S are disjoint. That is, if $F \cap S = \emptyset$, then $|F \cup S| = |F| + |S|$.

When does the Sum Rule apply? The Sum Rule applies when there are two tasks, but only *one* of the two tasks is performed.

Example 724. *At a restaurant downtown, any choice of dinner place comes with either soup or salad (but not both). There are* 3 *soup selections, and* 2 *salad choices. How many possible accompaniments (without the up-charge of both) are there to a dinner? Is the correct answer* 5 *or* 6*?*

There are three possible ways to complete the first task (choose one of the three soups). There are two possible ways to finish the second task (choose one of the two salads). For this particular thought experiment, the procedure is complete by doing one *of the two tasks. Therefore, there are* 5 *ways to complete the procedure.*

Example 725. *At another restaurant downtown, any choice of dinner place comes with both soup and salad. There are* 3 *soup selections, and* 2 *salad choices. How many possible accompaniments are there to a dinner? Is the correct answer* 5 *or* 6*?*

There are three possible ways to complete the first task (choose one of the three soups). There are two possible ways to finish the second task (choose one of the two salads). For this particular thought experiment, the procedure is complete by doing both *of the tasks. Therefore, there are* 6 *ways to complete the procedure.*

Warning 726: Do not confuse the Product Rule and the Sum Rule

Do not apply the Sum Rule when only the Product Rule can be used. Do not apply the Product Rule when only the Sum Rule can be used.

Method 727: Product Rule versus Sum Rule

How can you stop confusing the Product Rule and the Sum Rule? Take the [larger] procedure and identify two smaller tasks. Do you have to complete *both* tasks to do the procedure, or are you *only* allowed to do one task to do the procedure? If you must do both tasks, you probably need to use the Product Rule.

Consider this another way: imagine that you completed only *one* of the two tasks. If the procedure would be incomplete in this situation, it is probably because both tasks need to be completed, and the Product Rule probably applies.

As an example of Method 727, let us refer back to Example 721. What would happen if only one procedure were complete? Then, we'd have a choice of a letter (such as picking "W") but if we stopped there, we could not have completed the procedure of creating a participant identifier.

Example 728. *A binary string is a string consisting only of* 0*s and* 1*s. For instance,* 10111 *and* 0101101010001 *are examples of binary strings. The first example is a binary string of length* 5 *while the second example is a binary string of length* 13*. How many binary strings of length* 13 *are there?*

We will answer the more general question of how many binary strings of length n there are. To make our argument, we focus on the number of binary strings of length 3*. Imagine that there are four boxes, and in each box, we must make the choice of either a* 0 *or a* 1*. Imagine repeating this procedure in every way possible. We would discover each of the binary strings of length* 3*.*

There are four tasks. The first task is choosing a 0 *or a* 1 *for the first location. The second task is choosing a* 0 *or a* 1 *for the second location. The third task is choosing a* 0 *or a* 1 *for the third location. All three tasks are similar (choosing a* 0 *or a* 1*), and each task has two ways to complete it.*

Does the Product Rule apply or does the Sum Rule apply? That is, is the final answer $2 \cdot 2 \cdot 2$ or $2+2+2$? Use Method 727 to determine the answer. To get a binary string of length 3*, notice that all three tasks must be completed. Thus, there are 2^3 binary strings of length* 4*. Convince yourself for sure by writing out all* 8 *binary strings of length* 3*. More generally, there are 2^n binary strings of length n.*

Example 729. *How many binary strings are there of length* 5 *or* 6*? Recall that in Example 728, we determined that there are 2^n binary strings of length n.*

In this problem, we should think of the [large] procedure as "making a binary string whose length is either 5 *or* 6*." While there are several ways to think of how to do the procedure, to connect this to the work we have already done, suppose there are two tasks. The first task is to "make a binary string of length* 5*" and the second is to "make a binary string of length* 6*." Then, to complete the procedure, we must complete exactly one of the two tasks, so the Sum Rule applies. Therefore, the total number of binary strings of length* 5 *or* 6 *is exactly $2^5 + 2^6$.*

Presented differently, there are $2 \cdot 2 \cdot 2 \cdot 2 \cdot 2 + 2 \cdot 2 \cdot 2 \cdot 2 \cdot 2 \cdot 2$ strings. Written this way, we can see all the uses of the Product Rule and the one use of the Sum Rule.

Theorem 730: Inclusion-Exclusion Principle

If A is a finite set and B is a finite set, then $|A \cup B| = |A| + |B| - |A \cap B|$.

Warning 731

The Inclusion-Exclusion Principle is a theorem that applies to finite sets only. The theorem does not apply if A is infinite or if B is infinite. If any set you encounter is infinite and the question deals with making equal the sizes of given sets, instead of using the Inclusion-Exclusion Principle (which does not apply), it is likely that the methods in Section 5.1 on cardinality apply.

The purpose of subtracting $|A \cap B|$ is to address the double-counting that occurs when considering $|A| + |B|$.

Example 732. *How many two-digit numbers have exactly one* 7 *as a digit? Let A be the set of all two-digit numbers where the starting digit is a* 7*. Let B be the set of all two-digit numbers where the ending digit is a* 7*. Then $|A| = 10$ and $|B| = 10$. What is $|A \cap B|$? This is the number of two-digit numbers where the starting digit is a* 7 *and the ending digit is a* 7*. There is exactly one number that is described like this, namely* 77*. Thus $|A \cap B| = 1$. Therefore, the number of two-digit numbers which have exactly one* 7 *as a digit is exactly $10 + 10 - 1$.*

Theorem 733: Inclusion-exclusion principle for three sets

If A, B, and C are all finite sets, then $|A \cup B| = |A| + |B| + |C| - |A \cap B| - |A \cap C| - |B \cap C| + |A \cap B \cap C|$.

Exercise 734. *In a standard 8×8 chessboard, a rook may move any number of spaces horizontally or vertically, with the restriction that the rook must stop at the first square to capture a piece of the opposite color. Assuming that a chessboard only contains one black rook and one white rook (and no other pieces),* *how many chessboard configurations are there where neither rook is in danger of being captured in the next move by the other rook?*

Exercise 735. *Now consider using an $n \times n$ chessboard. Assuming that a chessboard only contains one black rook and one white rook (and no other pieces), how many $n \times n$ chessboard configurations are there where neither rook is in danger of being captured in the next move by the other rook?*

Exercise 736. *A banking website requires customers using online banking to choose a password, whose length must be at least 8 characters and at most 20 characters. For simplicity, let's say that passwords are restricted to using the $2 \cdot 26$ upper/lowercase letters and the 10 numerical digits. (So, no special characters like the "at symbol" are allowed, to make sure that things aren't too complicated.) If a valid password must contain both numbers and letters, then how many possible passwords are there?*

Exercise 737. *A hexadecimal digit is a character which is either one of the 10 numeric digits or one of the first six letters (A through F) in the alphabet. For example, 19A0C0BE32D6FF21AB3 is a hexadecimal string. A wifi password must be a string of either 10, 26, or 58 hexadecimal digits. How many different passwords are possible?*

6.2 Permutations, combinations, and binomial coefficients

Suppose that you have all 13 hearts cards from a standard poker deck. In hearts, you have the A, 2, 3, 4, 5, 6, 7, 8, 9, 10, J, Q, and K cards. If you were to shuffle these cards, how many possible finals configurations would there be of these cards? Would it be 13? Or 13^2? Or 2^{13}? It turns out to be none of these.

To start off smaller, say that you had three index cards, labeled 1, 2, and 3. What are the possible results of shuffling? You can have 123, 132, 213, 231, 312, or 321. There are 6 total possibilities. To be a little more precise, we can say the following:

Definition 738

Let X be a finite set of cardinality n. A **permutation** is an n-tuple $(p_1, p_2, \dots, p_n)$ where each p_i is an element of X, and each element of X appears exactly once.

Example 739. *The six permutations of* $X = \{1, 2, 3\}$ *are* $(1, 2, 3)$ *and* $(1, 3, 2)$ *and* $(2, 1, 3)$ *and* $(2, 3, 1)$ *and* $(3, 1, 2)$ *and* $(3, 2, 1)$.

Suppose that you have five cards, labeled 1, 2, 3, 4, and 5. Let's relate how many possible shufflings there are of this 5-card deck to Section 6.1 by looking at a modified version of the Product Rule. Here's an experiment that allows us to consider all shufflings. Place all 5 cards face down. Select a card which will be on top of the deck. (That is, pick the card that's going to be in position 1.) Now, without seeing the card, this models shuffling. It could be the case that we picked up card number 2. Now, let's select one of the remaining 4 cards to take position 2 in the deck. The card we picked up could be card number 5. Then, cards 2 and 5 occupy positions 1 and 2 in the shuffling that we are constructing. Which cards could occupy position 3? It can be any of cards 1, 3, or 4. There are 3 possibilities.

Notice the pattern? When starting with n cards, after selecting the first card, there are $n-1$ possibilities for the second card. Then after selecting a second card, there are $n-2$ possibilities remaining for the third card. Notice that all selections need to be made (so think Product Rule, not Sum Rule), and the number of shufflings is $(n)(n-1)(n-2)\cdots(2)(1)$, which is $n!$.

Theorem 740

The number of permutations of a finite set of cardinality n is exactly $n!$.

We argued the above via the Product Rule and thought about the procedure from a broad view. Alternately (or for more formality), Theorem 740 can be proved by induction.

Exercise 741. *Prove that there are exactly $n!$ permutations of a set of cardinality n by induction.*

Exercise 742. *If there are* 10 *people that run a race, how many possible ways can the "final results" board be written?*

We now turn away from permutations themselves to discuss combinations, and relate the formula for combinations to Theorem 740. Suppose that there are 50 contestants in a game. The game consists of sorting a standard 52 card deck in a specified manner, which is visible to all contestants. The first 3 people to finish are each awarded \$1000. How many possible ways can earnings be distributed?

For this game, it does not matter if you are in first place or second place: either way, you'd win a \$1000 prize. Likewise, it doesn't matter if you're 11th place or 12th place: either way, you would not win a monetary prize. It definitely matters whether you're third place or fourth place. If we don't pay attention to monetary earnings and wanted a complete record of who finished in what order, then there would be 1000! possible outcomes. However, we want to focus our outcomes (for *this* scenario) to who wins money and who does not.

For a smaller version of this problem, suppose 7 contestants were vying for 3 identical prizes, and nobody can win twice. If the 7 people are represented by integers, then we want to consider 1234567 the same situation as 3214567 and as 1327456, even if we didn't want to earlier when discussing permutations. That is, the first three numbers (in any order) are 1, 2, and 3, while the last four numbers (in any order) are 4, 5, 6, and 7. How can we count this? If we start with 7! we have overcounted. By what factor have we overcounted? First, note that the first three numbers can appear in any order (and there are 3! possible orderings of these numbers) so we have overcounted by (at least) a factor of 3!. However, we should also note that the last four numbers can be in any order, and there are 4! such orders. So we have really overcounted by a factor of 4! as well. In fact, we have overcounted by a factor of the product, namely 3!4!.

Returning to our larger example, 1000! overcounts the outcomes we wish to count by a factor of 3!907!. More generally, we have:

Theorem 743

Let X be a finite set of cardinality n. The number of subsets of X whose cardinality is k is exactly

$$\frac{n!}{k!(n-k)!}.$$

We introduce a definition:

Definition 744

Let X be a finite set of cardinality n. A **combination** of X of size k is a k-element subset of X. A combination of size k is also known as a **k-combination**.

Theorem 743 tells us that a set of size n has exactly $\frac{n!}{k!(n-k)!}$ combinations of size k.

Definition 745

Given integers $n \geq k$, the number $\frac{n!}{k!(n-k)!}$ is called a **binomial coefficient**, is denoted by $\binom{n}{k}$, and is spoken, "n choose k."

The number $\binom{n}{k}$ is also denoted ${}_nC_k$ in other texts. The language "n choose k" is used because $\binom{n}{k}$ counts how many ways there are, starting with n times, to choose k of them. (That is, how many ways are there to choose k things when starting with n things? In other language, how many ways are there to make k selections from a list of n items?) The number $\binom{n}{k}$ is surprisingly always an integer (even though $k!(n-k)!$ appears in the denominator of its defining fraction,

We now discuss why $\binom{n}{k}$ is called a binomial coefficient. Consider the binomial $x+y$. If $(x+y)^2$ is expanded, every term has degree 2. If $(x+y)^3$ is expanded, every term has degree 3. If $(x+y)^4$ is expanded, every term has degree 4. If $(x+y)^3$ is expanded, before collecting like terms, how many terms are x^2y? If $(x+y)^4$ is expanded, before collecting like terms, how many terms are x^3y? If $(x+y)^8$ is expanded, before collecting like terms, how many terms are x^3y^5? Of the eight factors $(x+y)$ you have to choose 5 of them

$n=0$:					1				
$n=1$:				1		1			
$n=2$:			1		2		1		
$n=3$:		1		3		3		1	
$n=4$:	1		4		6		4		1

Table 6.1: Pascal's Triangle

to provide y, and the remaining 3 will provide an x. There are $\binom{8}{5}$ ways of doing that. More generally, of the n factors in $(x+y)^n$, to get the number of times the term $x^{n-j}y^j$ appears in the expansion using the distributive law, you have to choose j ys, thus there are $\binom{n}{j}$ ways, and thus that many terms. From this example with $n=8$ and $j=5$, we are seeing evidence of the Binomial Theorem:

Theorem 746: Binomial Theorem

For any reals x and y and for all $n \geq 0$ a non-negative integer

$$(x+y)^n = \sum_{j=0}^{n} \binom{n}{j} x^{n-j} y^j.$$

Exercise 747. *What is the coefficient of $x^{75}y^{25}$ in the expansion of $(x+y)^{100}$?*

Corollary 748. *For all non-negative integers n,*

$$\sum_{k=0}^{n} \binom{n}{k} = 2^n$$

Proof. Let $x=y=1$ in Binomial Theorem. □

Corollary 749. *For all n positive integer,*

$$\sum_{k=0}^{n} \binom{n}{k} (-1)^k = 0$$

Proof. Let $x=1$ and $y=-1$ in Binomial Theorem. □

The binomial coefficient $\binom{n}{k}$ is the kth entry in the nth row of Pascal's Triangle, which is shown in Table 6.1. Note that Pascal's Triangle starts with row $n=0$, and in each row, the starting entry is the 0th entry.

Exercise 750. *Suppose 100 people are trying out to be on the basketball team, but the coach is only going to select 17 people to be on the team. The coach will print a list of the 17 players that made it on the front of her office, with players' names in alphabetical order. How many ways could the team list be printed out?*

Exercise 751. *Suppose 100 people are running a race. The top three runners will be honored on a stage (with first place on the highest pedestal, third place on the lowest pedestal). The fourth through tenth place runners will receive honorable mention. How many ways could the race end?*

Exercise 752. *When writing $(a+b)^2$, there are a total of four terms, if you don't collect like terms. Similarly, if you don't collect like terms in $(a+b+c+d)^{9000}$, how many total terms are there?*

Exercise 753. *How many ways are there to arrange the letters a, b, c, and d such that a is not followed immediately by b? So, bacd is not allowed, but bcad is allowed.* [key]

6.3 Counting via bijections

Consider two sets A and B. When the sets A and B were infinite, the techniques in Section 5.1 were useful in proving that the sets A and B had the same cardinality by defining a bijection from A to B. There is a certain challenge in defining such functions when A and B are infinite. In this section, we consider the same general principle, but to finite sets A and B.

The questions are posed more typically as counting questions. More specifically, in a typical exercise, you will be presented with *two* counting problems, and will need to demonstrate that both counting problems have the same answer. (Sometimes, it is easier to show that the two counting problems have the same answer than it is to find what the actual answer *is*.) To model what is happening, think of the set A as being the set of objects described in the first counting problem, and think of the set B as being the set of objects described in the second counting problem. Even if it is hard to compute what number $|A|$ is or what number $|B|$ is, sometimes it is possible to show that $|A| = |B|$ by construction a bijection from A to B.

Let us dive right into an example:

Example 754. *In this example, we will show that the number of binary strings of length n is equal to the number of subsets of a set of cardinality n by constructing a bijection.*

Fix a positive integer n. Let C be the set of all binary strings of length n. Then, let D be the set of all subsets of $\{1, 2, \ldots, n\}$. In other words, D is the power set of $N = \{1, 2, \ldots, n\}$. We will construct a bijective function from C to D. We define a function $f : C \to D$ by the following rule: given a binary string $\sigma = s_1 s_2 \ldots s_n$ of length n, with each $s_i \in \{0, 1\}$, we assign to this string σ a set $Y = f(\sigma)$, where $i \in Y$ if $s_i = 1$, and $i \notin Y$ if $s_i = 0$.

Based on this definition of f, we prove that f is surjective and injective. To prove that f is surjective, let Y be an arbitrary element of D. That is, Y is a subset of $\{1, 2, \ldots, n\}$. Then, construct the string $\sigma = s_1 s_2 \ldots s_n$ where $s_i = 1$ if $i \in Y$, and $s_i = 0$ otherwise. Then, $f(\sigma) = Y$ follows based on the rule defining f. To prove that f is injective, suppose σ and τ are arbitrary elements of C. That is, suppose that $\sigma = s_1 s_2 \ldots s_n$ and $\tau = t_1 t_2 \ldots t_n$ are two binary strings of length n. Suppose that $f(\sigma) = f(\tau)$. Notice that this is really a statement that two sets are equal. Then, since $f(\sigma) \subseteq f(\tau)$, if $i \in f(\sigma)$, then $i \in f(\tau)$. That is, if $s_i = 1$, then $t_i = 1$. Likewise, if $t_i = 0$, then $s_i = 0$. From this, we can see that the ith character in s must be the same as the ith character of t, so the two binary strings σ and τ must be the same, proving that f is injective. In conclusion, f is a bijection from C to D, so the number of binary strings of length n is equal to the number of subsets of a set of cardinality n.

To focus the example above, the discussion of how large C is or how large D is was left out completely. From Example 728, we had already determined that there are 2^n binary strings of length n. That is, $|C| = 2^n$. By Example 754, $|C| = |D|$. Therefore, $|D| = 2^n$. This puts nice closure to the matter brought up in Remark 452, namely that a set of cardinality n has a power set of cardinality 2^n, and an alternate notation for the power set of N is 2^N, since $|2^N| = 2^{|N|}$.

A second example addresses distributing coins in distinct bins:

Example 755. *Suppose you have 11 identical coins which need to be placed in three bins. The three bins are labeled (and considered different). For instance, the bins are labeled Bin 1, Bin 2, and Bin 3. This question concerns distributing the 11 coins into the three bins. Putting 2 coins in Bin 1, 4 coins in Bin 2, and 5 coins in Bin 3 is considered different than putting 2 coins in Bin 1, 5 coins in Bin 2, and 4 coins in Bin 3. Show that the number of ways of distributing 11 coins in 3 distinct bins is the same as the number of strings of length 13 consisting of 11 stars and 2 vertical lines.*

For this example, we describe the bijection in words, but will not formally define it. (Because we don't formally define the bijection, we will not prove that we have a bijection, but after the description, the reader will likely be convinced that we are describing a bijection.) We consider strings of length 13, with two of the characters being vertical lines and the remaining characters being stars. The 11 stars represent the 11 coins, and the two vertical bars represent "dividing lines." As an example, the string

$$**\,|\,****\,|\,*****$$

represents 2 coins in Bin 1, 4 coins in Bin 2, and 5 coins in Bin 3, whereas

$$**\,|\,*****\,|\,****$$

represents 2 *coins in Bin 1,* 5 *coins in Bin 2, and* 4 *coins in Bin 3. The number of stars before the first vertical bar counts the number of coins in Bin 1, the number of stars between the two vertical bars represents the number of coins in Bin 2, and the number of stars appearing after the second vertical bar represents the number of coins in Bin 3. Thus,* $***||********$ *is the string corresponding to* 3 *coins in Bin 1,* 8 *coins in Bin 3, and no coins in Bin 2. The string* $|*********|**$ *corresponds to no coins in Bin 1,* 9 *coins in Bin 2, and* 2 *coins in Bin 3. The string* $||***********$ *means all* 11 *coins are in Bin 3. How would you represent all coins in Bin 1? How would you represent all coins in Bin 2?*

After some time practicing with this, note that there is one fewer "dividing line" than there are bins. Then, note that every string (with two vertical lines) corresponds to one and only one way of distributing the coins. It follows that the number of ways of distributing 11 *coins in* 3 *distinct bins is the same as the number of strings of length* 13 *consisting of* 11 *stars and* 2 *vertical lines.*

The number of strings of length 13 consisting of 11 stars and 2 vertical lines can be counted by considering the number of ways to choose 2 selections out of 13 items. Consider the 13 locations in a string of length 13 as the 13 items, and choosing 2 of those 13 locations where vertical lines will be placed as the 2 selections. Thus, there are $\binom{13}{2}$ strings of length 13 consisting of 11 stars and 2 vertical lines. The work of the example then allows us to also conclude that there are $\binom{13}{2}$ ways of distributing 11 coins in 3 distinct bins.

Our third example discusses lattice paths, which are paths that can all be drawn on the lined portion of graph paper:

Example 756. *A path from* $(0,0)$ *to* (m,n) *is called a north-and-east path if the path starts at the origin and ends at* (m,n) *traveling in one-unit line segments that go either directly up or directly to the right only. (These paths are not allowed to travel left or travel down.) For example the path from* $(0,0)$ *to* $(0,1)$ *to* $(0,2)$ *to* $(1,2)$ *to* $(1,3)$ *to* $(2,3)$ *to* $(3,3)$ *to* $(4,3)$ *is a north-and-east path. Show that the number of north-and-east paths from* $(0,0)$ *to* (m,n) *is exactly* $\binom{m+n}{n}$.

What is the bijection here? Note that each north-and-east path can be converted to a string consisting Ns and Es only, with n *Ns and* m *Es. For instance, the path already mentioned corresponds to the string* $NNENEEE$. *This is a string of length* $m+n$, *where one only needs to choose which* n *of the* $m+n$ *locations will have the letter N. Thus, there are* $\binom{m+n}{n}$ *such strings, and exactly the same number of paths.*

Exercise 757. *Suppose you have 8 apples and 15 oranges. How many ways can you distribute the fruit to 5 people? (Person 1 ending up with all the fruit is considered different than Person 4 ending up with all the fruit.)*

Exercise 758. *In Figure 6.1, how many paths are there from* $(0,0)$ *to* $(9,5)$*? A path may consist of traveling north one unit at a time or east one unit at a time.*

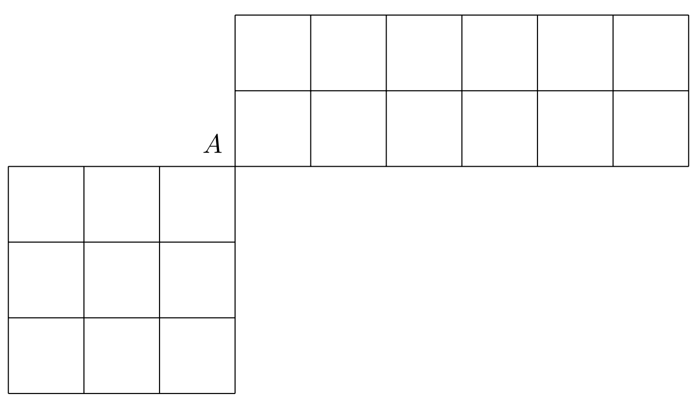


Figure 6.1: All paths must go through the point $(3,3)$

Exercise 759. *In Figure 6.2, how many paths are there from* $(0,0)$ *to* $(9,5)$*? A path may consist of traveling north one unit at a time or east one unit at a time.*

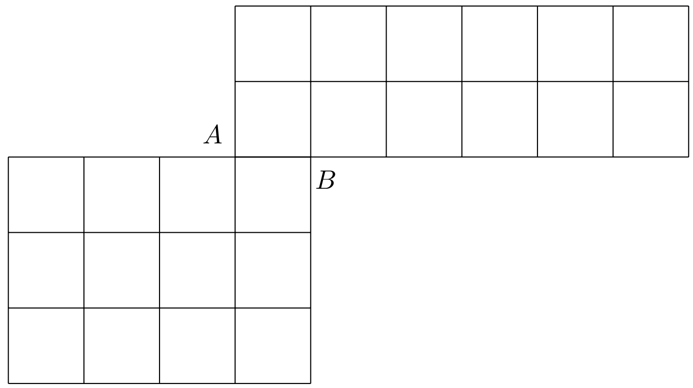


Figure 6.2: All paths must go through the point $(3,3)$ or $(4,3)$ or both

Exercise 760. *Let A be a non-empty finite set. So, there exists a positive integer n such that A has precisely n elements. Let us list the n elements of A as $a_1, \ldots, a_n$, where $a_i \neq a_j$ if and only if $i \neq j$. Thus, $A = \{a_1, \ldots, a_n\}$. Let B be the set of bitstrings of length n. Let $C = \{3,4\}^n$. Let $D = P(A)$. Prove that the sets B, C, and D have the same cardinality.*

Exercise 761. *As the department chair for Physical Education at your local school, you have a total of n basketballs and m large mesh bags to hold basketballs. Each mesh bag is going to a different PE teacher, so putting 8 basketballs in the first bag and 5 basketballs in the second bag is considered different than putting 5 basketballs in the first bag and 8 basketballs in the second bag. However, all the basketballs are exactly of the same brand and quality. If each mesh bag must have at least one basketball (no empty mesh bags allowed), how many ways are there to distribute the n basketballs.*

Exercise 762. *If there are w basic propositions, then how many rows are there in complete truth table?*

6.4 Combinatorial proof

This section introduces combinatorial proof, which is a technique that is sometimes successful to show that a formula involving one or more variables is true. For instance, using the techniques of this section, one can prove that for all positive integers n,

$$2^n + 2^{n+1} = 2^n \cdot 3.$$

Now, of course, the following is a proof, but not a combinatorial proof:

Proof. Let n be an arbitrary positive integer. Then,

$$2^n 3 = 2^n(1+2) = 2^n \cdot 1 + 2^n \cdot 2 = 2^n + 2^n \cdot 2^1 = 2^n + 2^{n+1},$$

as desired. □

The proof above utilizes facts from algebra. After describing what a combinatorial proof is, we will reprove the formula using the new method. While it is tempting to skip this "because we already have a proof," there are many situations where a combinatorial proof is easier, or at least way more natural. (There are other situations where combinatorial proof is natural, and an algebra-based proof is nearly impossible.)

In combinatorial proof, the proof writer is presented with a formula involving one or more variables. For simplicity, let us take a look at a formula involving just one variable, such as $2^n + 2^{n+1} = 2^n 3$ from above.

Method 763: Combinatorial proof

Presented with a formula, to use the method of combinatorial proof, the proof writer should:

1. Describe a relevant counting problem.
2. Use counting techniques from previous sections (in a valid way, of course) to answer the counting problem from Step 1.
3. Use counting techniques to answer the counting problem from Step 1 again, but using different counting techniques.

Steps 2 and 3 above sound rather confusing. While we will clarify using an example in a moment, it will be fruitful to clarify what is meant. Imagine a situation where a counting problem is described (Step 1). Say, for example, that your math instructor describes a counting problem to the class. As students work independently, suppose that two students both go to the instructor to say they have an answer to the question. The first student explains their reasoning to the professor, and the professor agrees that the student properly applied the counting techniques presented from the previous sections. The second student explains their reasoning to the professor, and the professor agrees that the student properly applied the counting techniques presented from the previous sections. Now, the first student and the second student have formulas that look different. How could this be? Is there a student who is incorrect? The reasonable conclusion is that, while the two students' formulas look different, they are actually equal, and can rightly be equated. What happened was that the two students "counted the same thing two different ways," with both ways being correct. The task in combinatorial proof is to take a counting problem and "count the same thing in two different ways."

Let us prove $2^n + 2^{n+1} = 2^n \cdot 3$ using a combinatorial proof. We will follow the three-step method:

1. First, we identify a counting problem. The counting problem needs to be relevant. (This comes with experience.) For the formula we need to prove, we consider the following counting question: how many binary strings are there of length n or length $n + 1$?

2. In Example 729, we worked out that the number of binary strings of length 5 or 6 is $2^5 + 2^6$. A generalized version of that argument proves that the number of binary strings of length n or $n + 1$ is precisely $2^n + 2^{n+1}$. We omit the details here.

3. The final step is to count how many binary strings there are of length n or length $n + 1$ again, but to do this independently. Try to ignore as much as possible the work from Step 2, because we need to do something different (yet valid).

   For the sake of clarity and concreteness, let us consider the number of binary strings of length 5 or 6, noting that our argument will generalize. There are 2 options for the first bit. After selecting that, there are 2 options for the second bit. Then, there are 2 options for the third bit. Then, 2 options for the fourth bit. Then, 2 options for the fifth bit. Finally, we now make a sixth choice, but the options will be different. For our sixth task, the options are 0 or 1 or "nothing." Having our sixth choice be 0 means that we are building a binary string of length 6 where the last bit is 0. Having our sixth choice be 1 means that we are building a binary string of length 6 where the last bit is 1. Having our sixth choice be "nothing" means that we are building a binary string of length 5. By running through all the possibilities where the sixth choice is "nothing" notice that we have discovered all binary strings of length 5. Looking more broadly, we have discovered each binary string whose length is 5 or 6 exactly once. By the Product Rule, there are $2 \cdot 2 \cdot 2 \cdot 2 \cdot 2 \cdot 3 = 2^5 \cdot 3$ such strings. More generally, there are $2^n \cdot 3$ binary strings of length n or $n + 1$.

In summary, here is what we did: In Step 1, we considered the problem of counting binary strings of length n or $n + 1$. In Step 2, we then answered that the number of such strings is $2^n + 2^{n+1}$. In Step 3, we then answered that the number of such strings is $2^n \cdot 3$. Since the number of binary strings of length n or $n + 1$ is $2^n + 2^{n+1}$ and is also $2^n \cdot 3$, it must be the case that $2^n + 2^{n+1} = 2^n \cdot 3$.

As a second example of combinatorial proof, we will prove the following theorem:

Theorem 764 (Pascal's Identity)**.** *Let n and k be positive integers such that $n \geq k$. Then,*

$$\binom{n+1}{k} = \binom{n}{k-1} + \binom{n}{k}.$$

Proof. Let T be a set of cardinality $n+1$. The left side expresses the number of ways of choosing k elements from T.

We now express this number in a different way. Let $a \in T$. (That is, select one element out of the set T to be called a.) Let $S = T \setminus \{a\}$. So S has exactly n elements. That is, $|S| = n$.

To choose k elements from T, we may either include $a \in T \setminus S$ or not:

- if $a \in T$ is one of the k elements, choose $k-1$ elements from S.
- if $A \in T$ is not one of the chosen elements, choose k elements from S.

There are $\binom{n}{k-1}$ ways to do the first task and $\binom{n}{k}$ ways to do the second task. Since we do exactly one of these two tasks, the Sum Rule applies, thus there are $\binom{n}{k-1} + \binom{n}{k}$ ways to complete the task. □

Remark 765: How is this different from counting via bijections?

When comparing this technique to Section 6.3, it sounds like we are discussing the same topic, but we are not. In Section 6.3, there are two different sets A and B (described two different ways), and by constructing a bijective function from A to B, we conclude that $|A| = |B|$. In this section, there is one set A, and $|A|$ is counted two different ways.

Said differently, in Section 6.3, there are *two* sets and there is *one* counting expressions. In this section, there is *one* set and there are *two* counting expressions.

Exercise 766. *Give a combinatorial proof that*

$$\binom{n}{k} = \binom{n}{n-k}.$$

[key]

Exercise 767. *Give a combinatorial proof that*

$$\binom{2n}{n} = \binom{n}{0}^2 + \binom{n}{1}^2 + \cdots + \binom{n}{n}^2.$$

[key]

Exercise 768. *Provide a combinatorial proof of*

$$\binom{n}{r}\binom{r}{k} = \binom{n}{k}\binom{n-k}{r-k}$$

for all non-negative integers $n \geq r \geq k$.

Exercise 769. *Give a combinatorial proof that*

$$\binom{2n}{2} = 2\binom{n}{2} + n^2.$$

6.5 Pigeonhole Principle

The **Pigeonhole Principle** is a useful tool in one's proof toolbox. The name may sound strange, but the theorem is often mentioned by name in proofs involving counting to convince the reader of a situation in which things must be *shared*. More specifically, there are two sets of objects, and as objects from a set A are assigned to objects of another set B, due to "not having enough to go around," some objects from A must *share*:

Theorem 770: Pigeonhole Principle

If $k+1$ or more pigeons fly into k pigeonholes (or cubby holes, or boxes), then there is [at least] one box containing multiple pigeons.

The theorem as stated above is the natural way to think of the Pigeonhole Principle. In its formal version, the Pigeonhole Principle states the following:

Theorem 771 (Pigeonhole Principle, formal)**.** *Let A and B be finite sets satisfying $|A| > |B|$. If $f : A \to B$, then f is not injective.*

How are the informal and formal versions of the Pigeonhole Principle (Theorems 770 and 771) related? The k pigeons from the informal statement are the elements of the set A and the k pigeonholes are the elements of the set B. In the informal version, every pigeon must fly into a pigeonhole. In the formal version, because f is a function from A to B, the definition of function enforces the situation that each element of A is assigned exactly one element of B. If $p \in A$ is a pigeon, then $f(p) \in B$ is the hole that p flew into. If there are more pigeons than there are pigeonholes, and every pigeon must fly into a pigeonhole, then there must be a pigeonhole with more than one pigeon.

Example 772. *In the game of Scrabble, there are 27 types of tiles: one for each letter of the alphabet, and a blank one. If someone selects 28 random Scrabble tiles, by the Pigeonhole Principle, they will have at least one instance of duplicate tiles.*

Example 773. *Suppose that 300 people will participate in a psychology experiment. If subjects are to remain anonymous and use participant identifier consisting of one letter and one digit, by the Pigeonhole Principle, there will be people who need to share participant identifiers, since there are 300 people (pigeons) and from Example 721 there are 260 participant identifiers (pigeonholes).*

Example 774. *If 25 people go for a menú del dia lunch at the restaurant described in Example 722, by the Pigeonhole Principle, there will be people who have identical meals. (The people are the pigeons, and the different meal configurations are the pigeonholes.)*

Example 775. *Many websites allow users to have accounts. To protect security, users are required to sign into their accounts with passwords. If a website were to store people's passwords, this would spell disaster if hackers break in and access the database of passwords, because most people reuse passwords on other sites. Websites need to know if you are really you (did you enter the right password?) but without* storing *your password in their databases. How do they do this?*

The way most websites achieve this is by using a function built into computers called "md5." The function md5 is an example of what programmers call a hash function. The function md5 takes in any string and produces a hexadecimal string of length 32, where a hexadecimal string only uses the symbols 0 through 9 and the letters a through f. For example the md5 of myp@sswoRd is 177f7de747899ada2efba07993e8eb5e while the md5 of myPASSword is 661603f05290ddcaa6697a4b63843ec8. Note that when there are two different strings that have the same md5 output, this is known as a "hash collision" in the computer programmer community.

Since md5 has as domain all finite strings (an infinite set) and codomain of size 16^{32}, by a variant of the Pigeonhole Principle, there must be hash collisions. This means that, if a website is using md5 to store your password, someone can log in using something totally different as your password, but they must guess something which has the same md5 output, which is nearly impossible (as there's only a 1 in 16^{32} chance of this occurring).

We now turn to a more quantified version of the original Pigeonhole Principle:

Theorem 776: Generalized Pigeonhole Principle

If N pigeons fly into k pigeonholes, then there must be one container with at least $\lceil \frac{N}{k} \rceil$ pigeons.

For a real number x, the notation $\lceil x \rceil$ is the **ceiling** which denotes the smallest integer greater than or equal to x. For example, $\lceil 3 \rceil = 3$ and $\lceil \pi \rceil = 4$.

Exercise 777. *At a party, 25 guests mingle and shake hands with some fellow guests. Prove that at least one guest must have shaken hands with an even number of guests.* [key]

Exercise 778. *Show that in any set of six classes, each meeting regularly once a week on a particular day of the week, there must be two that meet on the same day, assuming no classes are held on weekends.* [key]

Exercise 779. *Show that if there are 30 students in a class, then at least two have last names that begin with the same letter.* [key]

Exercise 780. *Let n be a positive integer. Show that in any set of n consecutive integers, there is one which is divisible by n.* [key]

Exercise 781. *A coin is flipped eight times where each flip comes up either heads or tails. How many possible outcomes contain exactly three heads?* [key]

Exercise 782. *How many bit strings of length 10 have at least six 1s?* [key]

Exercise 783. *How many strings of six uppercase letters from our alphabet contain the letter A?* [key]

Exercise 784. *How many bit strings of length 10 contain at most four 0s?* [key]

Exercise 785. *How many ways are there to distribute 100 five-dollar bills amongst 20 friends?* [key]

Exercise 786. *In how many ways can a set of two positive integers less than 100 be chosen?* [key]

Exercise 787. *How many subsets with an odd number of elements does a set with 10 elements have?* [key]

Exercise 788. *Let $f : \mathbb{Z}^2 \to \mathbb{Z}$ be defined by $f(x, y) = x + y$. Prove that f is surjective.* [key]

Exercise 789. *Let A, B, C, D, and E be sets satisfying $A \subseteq B$, and $B \subseteq D \cup E$. Prove: if D and E are elements of the power set of C, then $A \subseteq C$.* [key]

Exercise 790. *How many binary strings of length 7 begin in a 1 or end in a 1 or have exactly four 1s?* [key]

Exercise 791. *Take a standard 52-card deck. A 5-card poker hand is called a flush if all cards are the same suit (for example, all five cards are diamonds). How many different flushes are there?* [key]

Exercise 792. *For this exercise, consider only standard poker decks with 52 cards: there are no jokers. Suppose that you have $52! + 1$ decks of cards, and that you shuffle each deck individually. Prove that at least two of the decks must have exactly the same shuffle result.*

Chapter 7

Loose ends

7.1 Without loss of generality

We discussed proof by cases in Section 3.6. There are examples of proofs done using cases where the cases are nearly identical. (That is, other than swapping each x and y, for example, the proof of each case would be the same.) In those instances, it is typical to only include the proof of once case, and that case usually starts with the language of without loss of generality (WLOG). For example, reading “Either $x \leq y$ or $y \leq x$” sounds like the preamble to a proof by cases using two cases (where the next paragraph will examine the case of $x \leq y$, and the following paragraph will examine the case of $y \leq x$). Instead, one might write “Without loss of generality, we assume $x \leq y$” and then include only the proof of the case when $x \leq y$.

Something like this can really only be done if, up until this point in the proof, x and y have had completely interchangeable roles. That is, you we can't just “get away with” doing one case when there are actually different cases. There really needs to be no *loss* of generality.

7.2 Slicker proofs of existentially-quantified statements

In Section 3.1.6, we proved Theorem 238, which stated “If a is even and b is even, then $a + b$ is even.” Our first proof was long (so that we could illustrate certain points), but we presented shorter and shorter proofs, with the last of these being as follows:

Proof. Let a be even and b be even. We want to prove $a + b$ is even. Since a is even, there exists an integer r such that $a = 2r$. Since b is even, there exists an integer s such that $b = 2s$.

Let $t = r + s$. Since r and s are integers, t is an integer. By substitution, $a + b = 2r + 2s = 2(r + s) = 2t$. Because there exists an integer t such that $a + b = 2t$, we conclude $a + b$ is even. □

The proof above followed Method 217 rather literally in that one *must* define an object to be called t in order to prove a statement of the form “there exists ... in $\mathbb{Z}$ such that ...” but we note that t was simply defined to be $r + s$. We have thusfar taken the approach of “in order to prove that something exists, define it with a new letter” to make the understanding smoother. But essentially the same content appears in this shorter proof which doesn't explicitly define something to be called t.

Proof. Let a be even and b be even. We want to prove $a + b$ is even. Since a is even, there exists an integer r such that $a = 2r$. Since b is even, there exists an integer s such that $b = 2s$.

Since r and s are integers, $r + s$ is an integer. By substitution, $a + b = 2r + 2s = 2(r + s)$. Because there exists an integer t such that $a + b = 2t$, we conclude $a + b$ is even. □

The proof above is slightly shorter, and in fact, the last sentence could even be rewritten from “Because there exists an integer t such that $a + b = 2t$, we conclude $a + b$ is even” to just “We conclude $a + b$ is even.”

Chapter 8

Proof practice

All along, the primary purpose of this handbook has been to build your skills of reading definitions and applying them to create proofs of theorems. These are the main skills applied in courses which have this course as a prerequisite. This chapter provides a preview of the material in those courses taken after this course, as an opportunity to practice these skills that lead toward writing proofs.

8.1 Abstract algebra

Abstract algebra (also called modern algebra) is a systematic study of the behavior of sets equipped with operations, and the functions defined between such sets. The primary subdisciplines in abstract algebra are group theory, ring theory, and field theory. As a preview, in this handbook, we present a short introduction to group theory.

Definition 793. *Let G be a nonempty set. A* **binary operation** *on G is a function from $G \times G$ to G.*

Remark 794. *If the binary operation is $\star$, instead of using typical function notation, typically $a \star b$ is written to mean $\star((a, b))$. Do not confuse $a \star b$, the notation for a binary operation with $a \sim b$, the notation for a binary relation. The set up of any theorem/exercise/etc. will always give enough information away to how you interpret the symbol between the two set elements: note that if $\star$ is a binary relation, then $a \star b$ is a proposition, yet if $\star$ is a binary operation on G, then $a \star b$ is an element of G.*

Definition 795. *Let G be a nonempty set together with a binary operation $\star$ on G. We say G is a* **group under the operation $\star$** *if the following three properties are satisfied:*

- Associativity. *The operation $\star$ is associative, that is $(a \star b) \star c = a \star (b \star c)$ for all $a, b, c \in G$.*
- Identity. *There exists an element e (called the identity) in G such that for all $a \in G$, the equations $a \star e = e \star a = a$ hold.*
- Inverses. *For each element $a \in G$, there is an element $b \in G$ (called an inverse of a) such that $a \star b = b \star a = e$. Note that the e mentioned here is the one which is mentioned in the previous point.*

Warning 796

For the identity and inverses portions of the definition, please pay careful attention to the order of quantifiers! Read carefully. Note that any proving any of these statements and using of these statements *must* stay true to the quantifier order. (You know from Warning 167 that you do not get the same meaning when you switch a "for all" and a "there exists".)

Remark 797. *It turns out that for each element $a \in G$, it is the case that a has a* unique *inverse b. (Try proving this directly from the previous definition.) Because a has a unique inverse, we use a^{-1} to denote* the *inverse of a. This notation of a^{-1} regardless of what the binary operation $\star$ is.*

Definition 798. *Let G be a group under the operation $\star$. We say that G is* **abelian** *if for all $a, b \in G$, the equation $a \star b = b \star a$ holds.*

Definition 799. *Let G be a group under the operation $\star$ and $H \subseteq G$. We say H is a* **subgroup of G** *if H is a group under the operation $\star$.*

Theorem 800 (Subgroup Test)**.** *Let G be a group under the operation $\star$ and let H be a nonempty subset of G. If $a \star b \in H$ for all $a, b \in H$ and $a^{-1} \in H$ for all $a \in H$, then H is a subgroup of G.*

Definition 801. *Let G be a group under the operation $\star$. The* **center** *of group G, denoted by $Z(G)$, is the set of elements in G that commute with every element in G. In other words,*

$$Z(G) = \{g \in G : \forall a \in G,\, a \star g = g \star a\}.$$

Exercise 802. *Prove that the following sets are groups under the indicated operations.*

1. *the set of real numbers under addition*
2. *the set $\mathbb{Q}^2$ using coordinate-wise addition.*
3. *the set $\{1, -1, i, -i\}$ under multiplication*
4. *the set of bijections from a set A to A under composition*

Exercise 803. *Explain why the following sets are not groups under the indicated operations.*

1. *the set of natural numbers under addition*
2. *the set of integers under subtraction*
3. *the set of integers under multiplication*
4. *the set of rationals under multiplication*

Exercise 804. *Prove that the set of non-zero reals under multiplication forms a group.*

Exercise 805. *Let $D = \{d \in \mathbb{R} \mid \text{there is an integer } k \text{ such that } d = 2^k\}$. Prove that D is a group under multiplication.*

Exercise 806. *Prove the identity element of a group is unique.*

Exercise 807. *Prove inverse elements in a group are unique. (In other words, prove that each element in a group has a unique inverse.)*

Exercise 808. *Prove that $Z(G)$ is a subgroup of G.*

8.2 Real analysis

Real analysis (less commonly called advanced calculus) is a proof-based study of the theorems from calculus. When students first take calculus, proofs of some important theorems (such as the Squeeze Theorem, the Intermediate Value Theorem, the Mean Value Theorem, and the Extreme Value Theorem, to name a few) were probably skipped.

Perhaps for something like the Intermediate Value Theorem, a "picture of plausibility" was shown by the instructor. While students may have been expected to write short proofs *using* the Intermediate Value Theorem, almost every calculus instructor skips proving the Intermediate Value Theorem *itself* because a true understanding would require the contents of a book such as this one. In a course on real analysis, equipped with the method of proof, students study topics such as a proof of Intermediate Value Theorem *itself*.

Given a finite set of real numbers such as $\{x_1, x_2, \ldots, x_n\}$, we write $\max\{x_1, x_2, \ldots, x_n\}$ to denote the maximal element of the set and $\min\{x_1, x_2, \ldots, x_n\}$ to denote the minimal element. For example, $\max\{3, 7, 9.7\} = 9.7$ and $\max\{3, 7, 9.7\} = 3$. Note that if $y = \max\{x_1, x_2, \ldots, x_n\}$, then $y \geq x_1$ and $y \geq x_2$ and so on, while there exists an $i \in \{1, \ldots, n\}$ such that $y = x_i$.

Definition 809. *Let $I \subseteq \mathbb{R}$. We say that I is an* **interval** *if for all $a, b \in I$, if $c \in \mathbb{R}$ such that $a < c < b$, then $c \in I$.*

Definition 810. *Let $\varepsilon > 0$ and $c \in \mathbb{R}$. We define the* **ε-neighborhood about c**, *denoted $B_\varepsilon(c)$, to be the set*

$$B_\varepsilon(c) = \{x \in \mathbb{R} : |x - c| < \varepsilon\} = (c - \varepsilon, c + \varepsilon).$$

We call c the **center** *of the neighborhood.*

Definition 811. *Let $\mathcal{O} \subseteq \mathbb{R}$. We say that $\mathcal{O}$ is an* **open set** *if for every $c \in \mathcal{O}$, there exists $\varepsilon > 0$ such that $B_\varepsilon(c) \subseteq \mathcal{O}$.*

Definition 812. *Let $F \subseteq \mathbb{R}$. We say F is a* **closed set** *if $\overline{F}$ is an open set, where $\mathbb{R}$ is the universal set.*

Definition 813. *Let $A \subseteq \mathbb{R}$ be nonempty. We say A is* **bounded above** *if there is an $M \in \mathbb{R}$ such that for all $a \in A$, the inequality $a \leq M$ holds. We say A is* **bounded below** *if there is an $m \in \mathbb{R}$ such that for all $a \in A$, the inequality $m \leq a$ holds.*

Definition 814. *Let $A \subseteq \mathbb{R}$ be nonempty. We say A is* **bounded** *if there exists $M > 0$ such that for all $a \in A$, one has $|a| < M$.*

Definition 815. *Let $f : \mathbb{R} \to \mathbb{R}$. We say that the limit of f as x approaches the real number a* **exists** *if there is a real number L such that for every number $\varepsilon > 0$, there is a number $\delta > 0$ such that if $0 < |x - a| < \delta$, then $|f(x) - L| < \varepsilon$. If this occurs, we also say that the limit of f as x approaches a is L and write*

$$\lim_{x \to a} f(x) = L.$$

Notice that we are really defining the phrase the "limit.... exists" as opposed to defining existence (in the sense of existential quantifier).

Exercise 816. *Let $X \subseteq \mathbb{R}$ be nonempty. Prove that X is bounded if and only if X is bounded below and bounded above.*

Exercise 817. *Suppose $0 < \delta < \varepsilon$. Prove for all $r \in \mathbb{R}$, one has $B_\delta(r) \subseteq B_\varepsilon(r)$.*

Exercise 818. *Prove a singleton set is closed and not open.*

Exercise 819. *Prove or disprove: $\emptyset$ and $\mathbb{R}$ are open sets.*

Exercise 820. *Prove or disprove: $\emptyset$ and $\mathbb{R}$ are closed sets.*

Your proofs/disproofs of the previous two exercises should rely on the <u>definitions</u> and <u>rules of inference</u>. If you simply say that a set cannot be simultaneously open and closed because you are thinking of doors, you are using your intuition: this intuition does not apply when using the words "open" and "closed" on subsets of $\mathbb{R}$.

Exercise 821. *True or False: A set cannot be both open and closed.*

You should base your answer to Exercise 821 on Exercises 819 and 820.

Exercise 822. *Prove the union of a finite collection of open sets is open. (In other words, if n is a finite positive integer, then the union $A_1 \cup A_2 \cup \cdots \cup A_n$ is open, provided that A_1 and A_2 and so on are all open.*

Exercise 823. *Show that the intersection of a countable collection of open sets is not necessarily open. (Give a countable collection of open sets that is open, and give a countable collection of open sets that is not open)*

Exercise 824. *This exercise concerns the definition of limit given in the extremely precise version with three quantifiers. (For most purposes, people typically work with a definition that is slightly less precise with two quantifiers. There is also a version that is more precise: the technically correct version of a definition of a limit should also have the variable x quantified.)*

Let $f : \mathbb{R} \to \mathbb{R}$ be the function defined by the rule $f(x) = 6x+7$. Prove that the limit of f as x approaches 3 exists.

Notice that our extremely-precise definition of limit has three quantifications: the variable L is quantified, the variable ε is then quantified, and then finally the variable δ is quantified. Write the negation of "the limit of f as x approaches a exists". This is good practice to review negation from Section 2.4.

8.3 Linear algebra

Linear algebra is a study of linearity, especially in three or more dimensions. You might be using this handbook as a reference in a linear algebra class, instead of using the handbook section-by-section. In that case, start by reading Chapter 1. It will also help to read Section 4.1 regarding the three formats of set notation. There will be some references to other portions of the handbook (especially the method boxes of the first section of Chapter 3), and refer back as needed.

8.3.1 Systems of linear equations

Definition 825: Linear equation

Fix a positive integer n. Fix real numbers $a_1, a_2, \ldots, a_n$ and a real number b. Then

$$a_1x_1 + a_2x_2 + \cdots + a_nx_n = b$$

is a **linear equation** in the variables $x_1, x_2, \ldots, x_n$.

Following Habit 2, a linear equation is a type of equation, and thus is a noun.

To create a single linear equation, once n is fixed, we need $n+1$ real numbers: the n numbers a_1, a_2, and so on up to a_n are the **coefficients** of the variables $x_1, x_2, \ldots, x_n$, while b is the constant on the right side of the equation. Instead of writing $x_1, x_2, \ldots, x_n$ we will often write $x_1, \ldots, x_n$ which should be understood to mean the same thing: start with x_1, follow the pattern dots, and stop with x_n. It is helpful to the reader to include x_2 in helping to establish the pattern, but there are situations when this can become a lot to write.

Example 826. *Let $n = 3$. Then $7x_1 + 6x_2 + 5x_3 = 4$ is a linear equation in the variables x_1, x_2, and x_3. The coefficients are $a_1 = 7$ and $a_2 = 6$ and $a_3 = 5$. The constant on the right side of the equation is $b = 4$.*

Example 827. *Let $n = 4$. Then $8x_1 - 6x_2 + 7x_4 = 2$ is a linear equation in the variables $x_1, \ldots, x_4$. The coefficient a_2 is negative, while the coefficient a_3 is zero.*

Recall from Definition 34 that an equation is **consistent** if it has a solution.

Example 828. *The linear equation $7x_1 + 6x_2 + 5x_3 = 4$ is consistent because $(x_1, x_2, x_3) = (0, \frac{2}{3}, 0)$ is a solution.*

Example 829. *The linear equation $7x_1 + 6x_2 + 5x_3 = 4$ is consistent because $(x_1, x_2, x_3) = (2, 3, -\frac{19}{5})$ is a solution.*

In the previous two examples, we have used two different solutions to show that the same equation is consistent.

Definition 830: System of linear equations

A **system of linear equations** is a collection of linear equations.

Example 831. *Consider:*

$$3x_1 + 4x_2 - 5x_3 = 21$$

$$2x_1 + 0x_2 + 337x_3 = -\pi$$

Then this is a system of linear equations. There are two equations in three variables.

Example 832. *Consider:*

$$x_1 + x_2 + x_3 = 30$$

$$x_2 + x_3 = 7$$
$$x_3 = 4$$

Then this is a system of linear equations. There are three equations in three variables.

Example 833. *Consider:*

$$x_1 + x_2 = 30$$
$$x_1 - 3x_2 = 48$$
$$x_1 + x_2 = 7$$

Then this is a system of linear equations. There are three equations in two variables.

Definition 834: Consistent system

A system of linear equations in the variables $x_1, \ldots, x_n$ is **consistent** if there is a simultaneous solution: in other words, if there exist real numbers $c_1, \ldots, c_n$ such that each equation is satisfied if $x_1 = c_1$ and $x_2 = c_2$, and so on.

Example 835. *The system in Example 832 is consistent because* $(x_1, x_2, x_3) = (23, 3, 4)$ *is a solution.*

Example 836. *The system in Example 833 is inconsistent because no values of* x_1, x_2, x_3 *will simultaneously satisfy all three equations. In fact, no matter what you pick for* x_1 *and* x_2*, the sum* $x_1 + x_2$ *cannot simultaneously be equal to both* 30 *and* 7*.*

Following Habit 2, the word consistent is an adjective that can apply to a single equation (see Definition 34) or to a system of linear equations (see Definition 834). You can think of replacing the word "consistent" mentally with "has a solution."

Theorem 837. *Consider a fixed system of linear equations. If* $(a_1, a_2, \ldots, a_n)$ *is a solution of the system of linear equations, and* $(b_1, b_2, \ldots, b_n)$ *is a solution of the system of linear equations, then their average* $(\frac{a_1+b_1}{2}, \frac{a_2+b_2}{2}, \ldots, \frac{a_n+b_n}{2})$ *is also a solution of the system of linear equations.*

Proof. Fix a system of linear equations. Let us consider just the first linear equation, which is of the form

$$c_1x_1 + c_2x_2 + \cdots + c_nx_n = d$$

for some scalars $c_1, c_2, \ldots, c_n, d$. Since $(a_1, a_2, \ldots, a_n)$ is a solution of the system,

$$c_1a_1 + c_2a_2 + \cdots + c_na_n = d$$

is true. If we divide both sides by 2 and distribute on the left side, we have

$$c_1\frac{a_1}{2} + c_2\frac{a_2}{2} + \cdots + c_n\frac{a_n}{2} = \frac{d}{2}$$

Similarly, since $(b_1, b_2, \ldots, b_n)$ is a solution of the system,

$$c_1b_1 + c_2b_2 + \cdots + c_nb_n = d$$

and dividing both sides by 2 will give us

$$c_1\frac{b_1}{2} + c_2\frac{b_2}{2} + \cdots + c_n\frac{b_n}{2} = \frac{d}{2}.$$

Adding this equation to a prior equation (and factoring) will give

$$c_1\frac{a_1+b_1}{2} + c_2\frac{a_1+b_2}{2} + \cdots + c_n\frac{a_n+b_n}{2} = d$$

which proves that $(\frac{a_1+b_1}{2}, \frac{a_2+b_2}{2}, \ldots, \frac{a_n+b_n}{2})$ is a solution to the equation

$$c_1x_1 + c_2x_2 + \cdots + c_nx_n = d.$$

While this was an argument for just the first equation in the system of linear equations, the same argument can be copied and used on the second equation in the system, and also used on the third equation in the system, and so on. Thus $(\frac{a_1+b_1}{2}, \frac{a_2+b_2}{2}, \ldots, \frac{a_n+b_n}{2})$ is a simultaneous solution to the system of linear equations. $\square$

8.3.2 Vectors and scalars

For a fixed positive integer n, Example 488 introduced $\mathbb{R}^n$, the set of all vectors in n-dimensional space. For instance, $(4,-5) \in \mathbb{R}^2$ and $(\sqrt{\pi}, 0, -8) \in \mathbb{R}^3$. We could use the build running through set format to write $\mathbb{R}^3$ as

$$\mathbb{R}^3 = \{(a,b,c) : a \in \mathbb{R}, b \in \mathbb{R}, c \in \mathbb{R}\}$$

or we could write

$$\mathbb{R}^3 = \{(x_1,x_2,x_3) : x_1 \in \mathbb{R}, x_2 \in \mathbb{R}, x_3 \in \mathbb{R}\}$$

Following typical convention, we can write $(\sqrt{\pi}, 0, -8) \in \mathbb{R}^3$ or the same element of $\mathbb{R}^3$ can be written vertically, but then this vector with three entries must be written using square brackets:

$$\begin{bmatrix} \sqrt{\pi} \\ 0 \\ -8 \end{bmatrix},$$

but writing

$$\begin{bmatrix} \sqrt{\pi} & 0 & -8 \end{bmatrix}$$

is considered different. In other words,

$$\begin{bmatrix} \sqrt{\pi} \\ 0 \\ -8 \end{bmatrix} = (\sqrt{\pi}, 0, -8),$$

but

$$\begin{bmatrix} \sqrt{\pi} \\ 0 \\ -8 \end{bmatrix} \neq \begin{bmatrix} \sqrt{\pi} & 0 & -8 \end{bmatrix}.$$

Definition 838: Vector

An element in $\mathbb{R}^n$ is called a **vector**.

Definition 839: Scalar

An element in $\mathbb{R}$ is called a **scalar**.

In other words, a scalar is a number. Following Habit 2, a vector is a noun and a scalar is a noun.

Example 840. *Since* $(6,7) \in \mathbb{R}^2$, *we say that* $(6,7)$ *is a vector. Similarly,* $(8,8,9)$ *is a vector in* $\mathbb{R}^3$.

Example 841. *The real number* 6 *is a scalar. The number* $\frac{3\sqrt{e}}{17}$ *is a scalar.*

Notice from our examples that a scalar and vector are different.

Warning 842: Scalar versus vector

A scalar is not the same thing as a vector: a scalar is not a vector, and a vector is not a scalar.

Warning 843: Inappropriate uses of the word consistent

In the previous section, we noted that an equation (and more generally, a system of linear equations) can be consistent. A vector or a scalar cannot be consistent. (In other words, heeding Warning 3, the word "consistent" cannot be applied to a vector or to a scalar.)

While we can say that $(23, 3, 4)$ is a solution to the system described in Example 832, we cannot say that the vector $(23, 3, 4)$ is consistent. We can say that the system described in Example 832 is consistent. Be sure to apply the word "consistent" to the system of linear equations, not to the vector $(23, 3, 4)$. The fact that this vector "works" is connected to all of this, but it is incorrect to speak/write by saying that the vector is consistent.

Typically, vectors are denoted with a bold letter, while scalars are denoted with an non-bold letter.

Example 844. *Consider the vector* $\mathbf{u} = (6, 7)$ *and the vector* $\mathbf{v} = (8, 8, 9) \in \mathbb{R}^3$*.*

Example 845. *Consider the scalar* $c = 6$ *and the scalar* $\lambda = \frac{3\sqrt{e}}{17}$*.*

There are often situations in linear algebra in which we must discuss several vectors in the same problem. For example, if we read "Let $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_s \in \mathbb{R}^m$" then we are asked to consider s vectors. (To clarify, s is the *number* of vectors up for discussion.) Each of those vectors is in $\mathbb{R}^m$. Let us consider other examples:

Example 846. *Suppose we read: Let* $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_k \in \mathbb{R}^n$*. Then there are a total of* k *vectors (the first one is called* $\mathbf{u}_1$*, and the last is called* $\mathbf{u}_k$*), each of which belong to* $\mathbb{R}^n$*.*

Example 847. *Suppose we read: Let* $\mathbf{u}_1, \ldots, \mathbf{u}_n \in \mathbb{R}^m$*. Then there are a total of* n *vectors (the first one is called* $\mathbf{u}_1$*, and the last is called* $\mathbf{u}_n$*), each of which belong to* $\mathbb{R}^m$*.*

Example 848. *Suppose we read: Let* $\mathbf{u}_1, \ldots, \mathbf{u}_n \in \mathbb{R}^n$*. Then there are a total of* n *vectors (the first one is called* $\mathbf{u}_1$*, and the last is called* $\mathbf{u}_n$*), each of which belong to* $\mathbb{R}^n$*. (In this case, the number of vectors and the number of entries in each vector match, as both are* n*. In the previous example,* m *may be equal to* n*, or* m *may not.)*

Example 849. *Suppose we read: Let* $\mathbf{a}, \mathbf{b} \in \mathbb{R}^k$*. Then there are just two vectors, and both are in* $\mathbb{R}^k$*.*

Example 850. *Suppose we read: Let* $\mathbf{a}_1, \ldots, \mathbf{a}_s \in \mathbb{R}^4$*. Then there are* s *vectors, and each vector is in* $\mathbb{R}^4$*. (That is, each vector has four entries.)*

Example 851. *Suppose we read: Let* $\mathbf{a} \in \mathbb{R}^4$*. Then there is one vector, named* $\mathbf{a}$ *and that vector is in* $\mathbb{R}^4$*. It could be that this vector is* $\mathbf{a} = (5, 6, 7, 8)$ *or it could be that this vector is* $\mathbf{a} = (0, -2, 5, -\pi)$*. Until we are told more information,* $\mathbf{a}$ *could be one of the two specific vectors we just mentioned, or* $\mathbf{a}$ *could be many other possible things. However, we know that* $\mathbf{a}$ *could not be* $(7, 8, 9)$*, because* $(7, 8, 9)$ *belongs to* $\mathbb{R}^3$*, not* $\mathbb{R}^4$*.*

Habit 852: Naming a vector's entries

Within a definition or in a proof, there are situations in which it is helpful to name the entries of the vector. (There are situations where this is not needed as well.)

Example 853. *Suppose someone writes: Let* $\mathbf{a} \in \mathbb{R}^4$*. Then, it may be helpful to write "Then* $\mathbf{a} = (a_1, a_2, a_3, a_4)$*." as a way to have* a_1 *and* a_2 *and* a_3 *and* a_4 *as the entries. Note that* a_1 *and* a_2 *and* a_3 *and* a_4 *are each real numbers. Continuing Example 851, it would very well be that* $a_1 = 5$ *and* $a_2 = 6$ *and* $a_3 = 7$ *and* $a_4 = 8$*, but this might not be true. Instead, it might be the case that* $a_1 = 0$ *and* $a_2 = -2$ *and* $a_3 = 5$ *and* $a_4 = -\pi$*, following the second example of what might be possible for* $\mathbf{a}$*.*

Example 854. *Suppose we know that* $\mathbf{u}, \mathbf{v} \in \mathbb{R}^3$*. Then it might be helpful to write "Let* $\mathbf{u} = (u_1, u_2, u_3)$ *and* $\mathbf{v} = (v_1, v_2, v_3)$*." Since* u_1 *is a real number and since* v_1 *is a real number, we could use a fact stated in Section 2.7 to convert* $u_1 + v_1$ *into* $v_1 + u_1$*.*

Example 855. *Suppose we know that* $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$. *Then it might be helpful to write "Let* $\mathbf{u} = (u_1, u_2, \ldots, u_n)$ *and* $\mathbf{v} = (v_1, v_2, \ldots, v_n)$.*" To write slightly less, we might leave out writing* u_2 *and writing* v_2 *and allow the pattern dots to account for them. In other words, it is slightly less to write "Let* $\mathbf{u} = (u_1, \ldots, u_n)$ *and* $\mathbf{v} = (v_1, \ldots, v_n)$.*"*

8.3.3 Vector and scalar arithmetic

Definition 856: Vector equality

Let $\mathbf{u} \in \mathbb{R}^n$ and $\mathbf{v} \in \mathbb{R}^n$. Let $\mathbf{u} = (u_1, u_2, \ldots, u_n)$ and $\mathbf{v} = (v_1, v_2, \ldots, v_n)$. Then we say that the vectors $\mathbf{u}$ and $\mathbf{v}$ are **equal** and write $\mathbf{u} = \mathbf{v}$ if $u_1 = v_1$ and $u_2 = v_2$ and so on, until $u_n = v_n$.

The definition of the equality of vectors given just now applies Habit 852. In fact, the text of the second sentence was taken directly from Example 855.

Definition 857: Vector addition

Let $\mathbf{u} \in \mathbb{R}^n$ and $\mathbf{v} \in \mathbb{R}^n$. Let $\mathbf{u} = (u_1, u_2, \ldots, u_n)$ and $\mathbf{v} = (v_1, v_2, \ldots, v_n)$. Then the **sum** of $\mathbf{u}$ and $\mathbf{v}$ is defined by adding corresponding entries of $\mathbf{u}$ and $\mathbf{v}$. More precisely,

$$\mathbf{u} + \mathbf{v} = \begin{bmatrix} u_1 + v_1 \\ u_2 + v_2 \\ \vdots \\ u_n + v_n \end{bmatrix}.$$

Following Habit 2, the sum of two vectors is a noun. What kind of noun? The sum of two vectors is a vector.

Remark 858. *When adding a vector and a vector, the result is a vector.*

Example 859. *Let* $\mathbf{u} = (4, 5)$ *and* $\mathbf{v} = (2, 7)$. *Then* $\mathbf{u} + \mathbf{v} = (4 + 2, 5 + 7) = (6, 12)$.

Example 860. *We give the same example in the other notation for vectors. Let*

$$\mathbf{u} = \begin{bmatrix} 4 \\ 5 \end{bmatrix} \quad \text{and} \quad \mathbf{v} = \begin{bmatrix} 2 \\ 7 \end{bmatrix}.$$

Then

$$\mathbf{u} + \mathbf{v} = \begin{bmatrix} 4+2 \\ 5+7 \end{bmatrix} = \begin{bmatrix} 6 \\ 12 \end{bmatrix}.$$

Example 861. *If* $\mathbf{a} = (1, 2, 3)$ *and* $\mathbf{b} = (4, 5, -10)$, *then* $\mathbf{a} + \mathbf{b} = (1 + 4, 2 + 5, 3 - 10) = (5, 7, -7)$.

Example 862. *If* $\mathbf{a} = (a_1, a_2, a_3)$ *and* $\mathbf{b} = (b_1, b_2, b_3)$, *then* $\mathbf{a} + \mathbf{b} = (a_1 + b_1, a_2 + b_2, a_3 + b_3)$.

If we compare the last two examples, we were able to simplify $1 + 4$ to become 5, but we were not able to simplify $a_1 + b_1$ any further because we did not know the values of a_1 and b_1.

Warning 863: The sum of vectors is not a scalar

Recall from Remark 858 that if $\mathbf{a}$ is a vector and $\mathbf{b}$ is a vector, then $\mathbf{a} + \mathbf{b}$ is a vector, not a scalar.

Thus, in Example 861, it is correct to take $\mathbf{a} = (1, 2, 3)$ and $\mathbf{b} = (4, 5, -10)$ and write $\mathbf{a} + \mathbf{b} = (1+4, 2+5, 3-10)$, but it would have been *incorrect* to write $\mathbf{a}+\mathbf{b} = 1+4+2+5+3-10$, because $\mathbf{a} + \mathbf{b}$ should be a vector while $1 + 4 + 2 + 5 + 3 - 10$ is a scalar.

Similarly, in Example 862, it is correct to start with $\mathbf{a} = (a_1, a_2, a_3)$ and $\mathbf{b} = (b_1, b_2, b_3)$ and then write $\mathbf{a}+\mathbf{b} = (a_1+b_1, a_2+b_2, a_3+b_3)$, but it would be *incorrect* to write $\mathbf{a}+\mathbf{b} = a_1+b_1+a_2+b_2+a_3+b_3$.

Definition 864: Zero vector

For a fixed positive integer n, the vector where all n entries are zeroes is called the **zero vector** and is denoted $\mathbf{0}$.

Example 865. *If* $\mathbf{u} = (6,5,4)$*, then* $\mathbf{u} + \mathbf{0} = (6,5,4) + (0,0,0) = (6+0, 5+0, 4+0) = (6,5,4)$.

Definition 866: Scalar multiplication

Let $\mathbf{u} \in \mathbb{R}^n$ and $c \in \mathbb{R}$. Let $\mathbf{u} = (u_1, u_2, \ldots, u_n)$. Then the **scalar multiple** of $\mathbf{u}$ by c is defined by multiplying each entry of $\mathbf{u}$ by c. More precisely,

$$c\mathbf{u} = \begin{bmatrix} cu_1 \\ cu_2 \\ \vdots \\ cu_n \end{bmatrix}.$$

Following Habit 2, the scalar multiplication defines a noun. What kind of noun? The result of scalar multiplication is a vector.

Remark 867. *When multiplying a scalar and a vector, the result is a vector.*

Example 868. *Let* $\mathbf{u} = (4,5)$ *and* $c = 3$*. Then* $c\mathbf{u} = 3(4,5) = (3\cdot 4, 3\cdot 5) = (12,15)$.

Example 869. *We give the same example in the other notation for vectors. Let*

$$\mathbf{u} = \begin{bmatrix} 4 \\ 5 \end{bmatrix} \quad \text{and } c = 3.$$

Then

$$c\mathbf{u} = 3\begin{bmatrix} 4 \\ 5 \end{bmatrix} = \begin{bmatrix} 3\cdot 4 \\ 3\cdot 5 \end{bmatrix} = \begin{bmatrix} 12 \\ 15 \end{bmatrix}.$$

Example 870. *If* $\mathbf{a} = (1,2,3)$ *and* $\lambda = 10$*, then* $\lambda\mathbf{a} = (10\cdot 1, 10\cdot 2, 10\cdot 3) = (10,20,30)$.

Example 871. *If* $\mathbf{a} = (a_1, a_2, a_3)$ *and* c *is an unknown real number, then* $c\mathbf{a} = c(a_1,a_2,a_3) = (ca_1, ca_2, ca_3)$.

If we compare the last two examples, we were able to simplify $10\cdot 2$ to become 20, but we were not able to simplify ca_2 any further because we did not know the values of a_2 and c.

Warning 872: The scalar multiplication does not produce a scalar

Recall from Remark 867 that if $\mathbf{v}$ is a vector and c is a vector, then $c\mathbf{v}$ is a vector, not a scalar.

Thus, in Example 870, it is correct to take $\mathbf{a} = (1,2,3)$ and $\lambda = -10$ and write $\lambda\mathbf{a} = (10\cdot 1, 10\cdot 2, 10\cdot 3)$, but it would have been *incorrect* to write $\lambda\mathbf{a} = 10\cdot 1 + 10\cdot 2 + 10\cdot 3$, because $\lambda\mathbf{a}$ should be a vector while $10\cdot 1 + 10\cdot 2 + 10\cdot 3$ is a scalar.

Similarly, in Example 871, it is correct to start with $\mathbf{a} = (a_1,a_2,a_3)$ and $c \in \mathbb{R}$, and then write $c\mathbf{a} = (ca_1, ca_2, ca_3)$, but it would be *incorrect* to write $c\mathbf{a} = ca_1 + ca_2 + ca_3$.

Warnings 863 and 872 are extremely important to keep in mind, both for computations and proofs. To reiterate, multiplying a scalar and a vector results in a vector, not a scalar. By analogy, think about the indefinite integral

$$\int x^8\,dx$$

versus the definite integral

$$\int_3^4 x^8\,dx.$$

While both are, in spoken terms, "integration" of a function, the results are very different. The result of an indefinite integral will be a family of functions (differing from each other by a constant), while the result of a definite integral is a number (representing area). It is proper to write

$$\int x^8\,dx = \frac{1}{9}x^9 + C$$

and

$$\int_3^4 x^8\,dx = \frac{1}{9}(4)^9 - \frac{1}{9}(3)^9.$$

It would be improper to write

$$\int_3^4 x^8\,dx = \frac{1}{9}x^9 + C,$$

and it would be similarly improper to write $c\mathbf{a} = ca_1 + ca_2 + ca_3$.

Definition 873: Short notation for scaling by -1

Given a vector $\mathbf{u}$, define $-\mathbf{u}$ as notation for $(-1)\mathbf{u}$.

Example 874. *If* $\mathbf{u} = (6, 5, 4)$, *then* $-\mathbf{u} = (-1)(6, 5, 4) = (-1 \cdot 6, -1 \cdot 5, -1 \cdot 4) = (-6, -5, -4)$.

Warnings 863 and 872 apply both in computations and in proofs. In linear algebra, some of the first proofs written are about properties of vector addition and scalar multiplication. The statements which are being proved often start with one or more copies of the phrase "for all." In more detail, these statements start with a clause in the form "for all ♣ in ♠," where ♣ is a [new] variable and ♠ is a set. If ♠ is the set $\mathbb{R}$, then ♣ is a scalar. If ♠ is the set $\mathbb{R}^n$, then ♣ is a vector. Thus, if you read a statement that begins "For all $\heartsuit \in \mathbb{R}^n$," then you know that $\heartsuit$ is a vector. Likewise, you might read "For all $\diamondsuit \in \mathbb{R}$" which may be written in slightly more words as "For all $\diamondsuit$ in $\mathbb{R}$" or might have even more words by writing "For all scalars $\diamondsuit$."

If you have to prove a statement that starts "For all scalars k", following Method 268, you should start your proof by writing "Let k be an arbitrary scalar" or "Let $k \in \mathbb{R}$ be arbitrary." Either of those statements invites the reader (as discussed in Language Discussion 270) to pick whatever scalar they want (even if they don't tell you what scalar they picked), but so that you can refer to their choice later, you are using k as notation.

Method 875: Proving a statement that begins "For all k in $\mathbb{R}$"

If you need to prove a statement of the form "For all k in $\mathbb{R}$, ♣" then following Method 268, start by writing "Let k in $\mathbb{R}$ be arbitrary." Then, use previously known (or assumed) statements to prove ♣.

Similarly, if the beginning of the statement you are proving starts "For all $\mathbf{u}$ in $\mathbb{R}^n$" it is because the statement you are proving is supposed to be true no matter what vector from $\mathbb{R}^n$ is chosen. To prove such a statement, following Method 268, you should write "Let $\mathbf{u} \in \mathbb{R}^n$ be arbitrary." As discussed in Language Discussion 270, this is telling the reader "Hey reader of my proof, you can pick anything from $\mathbb{R}^n$ that you want, and you don't even have to tell me what it is. But, so that I can refer to it later in my proof, let's call what you picked $\mathbf{u}$."

Method 876: Proving a statement that begins "For all $\mathbf{u}$ in $\mathbb{R}^n$"

If you need to prove a statement of the form "For all $\mathbf{u}$ in $\mathbb{R}^n$, ♠" then following Method 268, start by writing "Let $\mathbf{u}$ in $\mathbb{R}^n$ be arbitrary." Then, use previously known (or assumed) statements to prove ♠.

Once you have said "Let $\mathbf{u} \in \mathbb{R}^n$ be arbitrary," it will often (but not always) be helpful to refer to the individual entries. So once $\mathbf{u} \in \mathbb{R}^n$ has been established, it is sometimes (but not always) helpful to write "There exist $u_1, \ldots, u_n \in \mathbb{R}$ such that $\mathbf{u} = (u_1, \ldots, u_n)$." This now gives you access to scalars u_1 and u_2 and so on (all the way up to u_n) that can be used in your proof.

Method 877: Access to the entries of a vector

Once it has been established that $\mathbf{u}$ is a vector in $\mathbb{R}^n$, then one can write "There exist $u_1, \ldots, u_n \in \mathbb{R}$ such that $\mathbf{u} = (u_1, \ldots, u_n)$." or it may be shorter to write "Then $\mathbf{u} = (u_1, \ldots, u_n)$." which carries the assumption that $u_1, \ldots, u_n$ are scalars.

Here are the eight main properties of vector addition and scalar multiplication:

Theorem 878. *Fix a positive integer n. For all $\mathbf{u} \in \mathbb{R}^n$, for all $\mathbf{v} \in \mathbb{R}^n$, we have $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$.*

Theorem 879. *Fix a positive integer n. For all $\mathbf{u} \in \mathbb{R}^n$, for all $\mathbf{v} \in \mathbb{R}^n$, and for all $\mathbf{w} \in \mathbb{R}^n$, we have $(\mathbf{u} + \mathbf{v}) + \mathbf{w} = \mathbf{u} + (\mathbf{v} + \mathbf{w})$.*

Theorem 880. *Fix a positive integer n. For all $\mathbf{u} \in \mathbb{R}^n$, we have $\mathbf{u} + \mathbf{0} = \mathbf{u}$ and $\mathbf{0} + \mathbf{u} = \mathbf{u}$.*

Theorem 881. *Fix a positive integer n. For all $\mathbf{u} \in \mathbb{R}^n$, we have $\mathbf{u} + (-\mathbf{u}) = \mathbf{0}$ and $(-\mathbf{u}) + \mathbf{u} = \mathbf{0}$.*

Theorem 882. *Fix a positive integer n. For all $\mathbf{u} \in \mathbb{R}^n$, for all $\mathbf{v} \in \mathbb{R}^n$, and for all $c \in \mathbb{R}$, we have $c(\mathbf{u} + \mathbf{v}) = c\mathbf{u} + c\mathbf{v}$.*

Theorem 883. *Fix a positive integer n. For all $\mathbf{u} \in \mathbb{R}^n$, for all $c \in \mathbb{R}$, and for all $d \in \mathbb{R}$, we have $(c + d)\mathbf{u} = c\mathbf{u} + d\mathbf{u}$.*

Theorem 884. *Fix a positive integer n. For all $\mathbf{u} \in \mathbb{R}^n$, for all $c \in \mathbb{R}$, and for all $d \in \mathbb{R}$, we have $c(d\mathbf{u}) = (cd)\mathbf{u}$.*

Theorem 885. *Fix a positive integer n. For all $\mathbf{u} \in \mathbb{R}^n$, we have $1\mathbf{u} = \mathbf{u}$.*

We will prove Theorems 878 and 882 providing a lot of detail so that you can prove the remaining theorems. Let's start with Theorems 878. If we leave off the text about fixing n, then we are left with:

- For all $\mathbf{u} \in \mathbb{R}^n$, for all $\mathbf{v} \in \mathbb{R}^n$, we have $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$.

Let us compare this with the statement that addition is commutative for reals in Section 2.7:

- For all $a \in \mathbb{R}$, for all $b \in \mathbb{R}$, we have $a + b = b + a$.

Look at how these two statements look similar, but are different. The first of these is about commutativity of addition of *vectors*, while the second statement is about commutativity of addition of *scalars*. We need to use the statement "for all $a \in \mathbb{R}$, for all $b \in \mathbb{R}$, we have $a + b = b + a$" from Section 2.7 to prove the new statement "for all $\mathbf{u} \in \mathbb{R}^n$, for all $\mathbf{v} \in \mathbb{R}^n$, we have $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$." The fact that $a + b = b + a$ if a and b are scalars is to be considered a *known* fact to us, and we should use this fact (alongside any other facts from Section 2.7 that we might need) in order to prove "for all $\mathbf{u} \in \mathbb{R}^n$, for all $\mathbf{v} \in \mathbb{R}^n$, we have $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$."

Since the statement we need to prove starts with a "for all" we should follow Method 876 and write "Let $\mathbf{u} \in \mathbb{R}^n$ be arbitrary" to start our proof.

Proof of Theorem 878, Draft 1. Fix a positive integer n. Let $\mathbf{u} \in \mathbb{R}^n$ be arbitrary. We will prove that for all $\mathbf{v} \in \mathbb{R}^n$, we have $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$. □

The proof is not done (just draft 1), but since the statement that we need to prove now is "for all $\mathbf{v} \in \mathbb{R}^n$, we have $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$" which again starts with "for all" we follow Method 876 and allow the reader to pick $\mathbf{v}$ in $\mathbb{R}^n$ arbitrary.

Proof of Theorem 878, Draft 2. Fix a positive integer n. Let $\mathbf{u} \in \mathbb{R}^n$ be arbitrary. We will prove that for all $\mathbf{v} \in \mathbb{R}^n$, we have $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$. Let $\mathbf{v} \in \mathbb{R}^n$ be arbitrary. We will prove that $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$. □

Now we have to prove $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$, but the statement "For all $a \in \mathbb{R}$, for all $b \in \mathbb{R}$, we have $a + b = b + a$" is about real numbers, not about vectors! Here's where we can apply Method 877. Let's write "Then $\mathbf{u} = (u_1, \ldots, u_n)$." We'll write a similar statement for $\mathbf{v}$.

Proof of Theorem 878, Draft 3. Fix a positive integer n. Let $\mathbf{u} \in \mathbb{R}^n$ be arbitrary. We will prove that for all $\mathbf{v} \in \mathbb{R}^n$, we have $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$. Let $\mathbf{v} \in \mathbb{R}^n$ be arbitrary. We will prove that $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$. Then $\mathbf{u} = (u_1, \ldots, u_n)$. Similarly, $\mathbf{v} = (v_1, \ldots, v_n)$. □

Now, the point is that since u_1 and v_1 are real numbers (not vectors), we can take the statement "For all $a \in \mathbb{R}$, for all $b \in \mathbb{R}$, we have $a + b = b + a$" from Section 2.7 and turn $u_1 + v_1$ into $v_1 + u_1$ or vice versa. (The a would be u_1, and for b we would plug in v_1.) Similarly, we can take $u_2 + v_2$ and replace this with $v_2 + u_2$, and continue in this way, up until the nth time, when we have $u_n + v_n = v_n + u_n$.

Proof of Theorem 878, Draft 4. Fix a positive integer n. Let $\mathbf{u} \in \mathbb{R}^n$ be arbitrary. We will prove that for all $\mathbf{v} \in \mathbb{R}^n$, we have $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$. Let $\mathbf{v} \in \mathbb{R}^n$ be arbitrary. We will prove that $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$. Then $\mathbf{u} = (u_1, \ldots, u_n)$. Similarly, $\mathbf{v} = (v_1, \ldots, v_n)$. Then,

$$\mathbf{u} + \mathbf{v} = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix} + \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} = \begin{bmatrix} u_1 + v_1 \\ u_2 + v_2 \\ \vdots \\ u_n + v_n \end{bmatrix} = \begin{bmatrix} v_1 + u_1 \\ v_2 + u_2 \\ \vdots \\ v_n + u_n \end{bmatrix} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} + \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix} = \mathbf{v} + \mathbf{u},$$

which was what we wanted to prove. □

In the second-to-last line of our proof, there are five equals signs, and let us take a moment to explain each of them. The first one was a substitution: $\mathbf{u}$ was replaced with $(u_1, u_2, \ldots, u_n)$ and $\mathbf{v}$ was replaced with $(v_1, v_2, \ldots, v_n)$, although the vectors were written vertically in the proof (and horizontally in this paragraph). The second equality applies the definition of vector addition given in Definition 857. The third equality was where we had n uses of the commutativity of addition for scalars. This is why the first entry was $u_1 + v_1$ prior to the third equal sign, but is $v_1 + u_1$ after the third equal sign. The fourth equality applies Definition 857, but "backwards" in the sense that one vector turned into the sum of two vectors (while earlier, we had the sum of two vectors turn into one vector). The last equality was substitution, though "backwards" from how we substituted earlier in that $(u_1, u_2, \ldots, u_n)$ was replaced with $\mathbf{u}$ and $(v_1, v_2, \ldots, v_n)$ was replaced with $\mathbf{v}$.

What we gave was a complete proof, but there are ways to shorten it a bit. So, here is fundamentally the same proof, but with slightly fewer words:

Proof of Theorem 878, shorter version. Fix a positive integer n. Let $\mathbf{u} \in \mathbb{R}^n$ and $\mathbf{v} \in \mathbb{R}^n$ be arbitrary. So $\mathbf{u} = (u_1, \ldots, u_n)$ and $\mathbf{v} = (v_1, \ldots, v_n)$. We will prove that $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$. Then,

$$\mathbf{u} + \mathbf{v} = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix} + \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} = \begin{bmatrix} u_1 + v_1 \\ u_2 + v_2 \\ \vdots \\ u_n + v_n \end{bmatrix} = \begin{bmatrix} v_1 + u_1 \\ v_2 + u_2 \\ \vdots \\ v_n + u_n \end{bmatrix} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} + \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix} = \mathbf{v} + \mathbf{u},$$

which was what we wanted to prove. □

In this shorter version, we left out the first "We will prove that" sentence, and then we were able to compress the arbitrary selections into a single sentence. In fact, on some level, it is then natural to bring up $\mathbf{u} = (u_1, \ldots, u_n)$ and $\mathbf{v} = (v_1, \ldots, v_n)$ right away, so we moved this to happen earlier than our declaration of intent to prove that $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$.

For an even shorter proof, it is technically okay to leave out the "We will prove that" sentence. These sentences are optional, but helpful in that they help you see where your destination is. Just to show, we will give an even shorter proof which removes this sentence. While we're at it, we might leave out the sentence about fixing n, leaving this implicit because this was stated already in the two sentences of the statement of Theorem 878.

Proof of Theorem 878, even shorter version. Let $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$ be arbitrary. So $\mathbf{u} = (u_1, \ldots, u_n)$ and $\mathbf{v} = (v_1, \ldots, v_n)$. Then,

$$\mathbf{u} + \mathbf{v} = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix} + \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} = \begin{bmatrix} u_1 + v_1 \\ u_2 + v_2 \\ \vdots \\ u_n + v_n \end{bmatrix} = \begin{bmatrix} v_1 + u_1 \\ v_2 + u_2 \\ \vdots \\ v_n + u_n \end{bmatrix} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} + \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix} = \mathbf{v} + \mathbf{u},$$

as desired. □

See what other differences you spot between the "shorter version" and the "even shorter version."

In general the idea is to follow Method 876 (or Method 875 as appropriate), then take the list of statements in Section 2.7 as assume to be true for the proofs that we will do. In order to apply use the statements from Section 2.7 which are about reals, we may find it necessary to follow Method 877: this will allow us to access entries within a vector (which are scalars).

Depending on spacing, you may wish to write your vectors horizontally instead of vertically:

Proof of Theorem 878, horizontal notation version. Let $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$ be arbitrary. So $\mathbf{u} = (u_1, \ldots, u_n)$ and $\mathbf{v} = (v_1, \ldots, v_n)$. Then,

$$\begin{aligned}\mathbf{u} + \mathbf{v} &= (u_1, u_2, \ldots, u_n) + (v_1, v_2, \ldots, v_n) \\ &= (u_1 + v_1, u_2 + v_2, \ldots, u_n + v_n) \\ &= (v_1 + u_1, v_2 + u_2, \ldots, v_n + u_n) \\ &= (v_1, v_2, \ldots, v_n) + (u_1, u_2, \ldots, u_n) \\ &= \mathbf{v} + \mathbf{u},\end{aligned}$$

as desired. □

From this, though, those who are new to proofs in linear algebra are tempted to write the following as a "proof." What's incorrect about this "proof"?

Warning 886: Find the error in this "proof" of Theorem 878

Not a proof. Let $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$ be arbitrary. So $\mathbf{u} = (u_1, \ldots, u_n)$ and $\mathbf{v} = (v_1, \ldots, v_n)$. Then,

$$\begin{aligned}\mathbf{u} + \mathbf{v} &= (u_1, u_2, \ldots, u_n) + (v_1, v_2, \ldots, v_n) \\ &= u_1 + v_1 + u_2 + v_2 + \cdots + u_n + v_n \\ &= v_1 + u_1 + v_2 + u_2 + \cdots + v_n + u_n \\ &= (v_1, v_2, \ldots, v_n) + (u_1, u_2, \ldots, u_n) \\ &= \mathbf{v} + \mathbf{u},\end{aligned}$$

as desired. □

Discussion of error: While there is more than one error, the first issue is that $(u_1, u_2, \ldots, u_n) + (v_1, v_2, \ldots, v_n)$, which is the sum of two vectors, is turned into $u_1 + v_1 + u_2 + v_2 + \cdots + u_n + v_n$, which is a scalar. See Warning 863, which is illuminating the idea that the writer of this "proof" is not applying Definition 857 properly.

Let us now consider Theorem 882. If we leave off the sentence about fixing n as a positive integer, this theorem stated:

- For all $\mathbf{u} \in \mathbb{R}^n$, for all $\mathbf{v} \in \mathbb{R}^n$, and for all $c \in \mathbb{R}$, we have $c(\mathbf{u} + \mathbf{v}) = c\mathbf{u} + c\mathbf{v}$.

How will the proof start? Method 876 will have us first write something like "Let $\mathbf{u}$ and $\mathbf{v}$ in $\mathbb{R}^n$ be arbitrary." Similarly, Method 875 will have us then write "Let $c \in \mathbb{R}$ be arbitrary." We will then need to prove $c(\mathbf{u} + \mathbf{v}) = c\mathbf{u} + c\mathbf{v}$. While this looks like the distributive law, it is a peculiar mix of vectors and scalars, so the distributive law from Section 2.7 does not immediately apply, though will be relevant:

- For all $a \in \mathbb{R}$, for all $b \in \mathbb{R}$, for all $c \in \mathbb{R}$, one has $a(b + c) = ab + ac$.

The point is that $a(b + c) = ab + ac$ only applies when a, b, and c are real numbers. As with the earlier proof, following Method 877 gives us to access entries of each vector (which are scalars).

Proof of Theorem 882. Let $\mathbf{u}$ and $\mathbf{v}$ in $\mathbb{R}^n$ be arbitrary. So $\mathbf{u} = (u_1, \ldots, u_n)$ and $\mathbf{v} = (v_1, \ldots, v_n)$. Let $c \in \mathbb{R}$ be arbitrary. We will prove $c(\mathbf{u}+\mathbf{v}) = c\mathbf{u} + c\mathbf{v}$. Then,

$$\begin{aligned}
c(\mathbf{u}+\mathbf{v}) &= c\left(\begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix} + \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}\right) \\
&= c\begin{bmatrix} u_1+v_1 \\ u_2+v_2 \\ \vdots \\ u_n+v_n \end{bmatrix} \\
&= \begin{bmatrix} c(u_1+v_1) \\ c(u_2+v_2) \\ \vdots \\ c(u_n+v_n) \end{bmatrix} \\
&= \begin{bmatrix} cu_1+cv_1 \\ cu_2+cv_2 \\ \vdots \\ cu_n+cv_n \end{bmatrix} \\
&= \begin{bmatrix} cu_1 \\ cu_2 \\ \vdots \\ cu_n \end{bmatrix} + \begin{bmatrix} cv_1 \\ cv_2 \\ \vdots \\ cv_n \end{bmatrix} \\
&= c\begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix} + c\begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} \\
&= c\mathbf{u} + c\mathbf{v},
\end{aligned}$$

as desired. □

Now that the proof is done, there is a sequence of seven equalities, which we now describe. The first is substitution. The second applies Definition 857. The third applies Definition 866. In this step, we are careful to write $c(u_1 + v_1)$ so that the crucial work of the next step can be shown. The fourth equality is n total uses of the Distributive Law from Section 2.7. This is what turned $c(u_1 + v_1)$ into $cu_1 + cv_1$ and so on. The fifth equality applies Definition 857, although "backwards" from earlier, and with new vectors. The sixth equality applies Definition 866 "backwards" twice. The second equality is substitution.

While it is visibly helpful to recognize the pattern by including what happens with the second entries of each vector, this can sometimes become a lot to write. Here is the same proof, where the onus is left on the reader a bit more to discover the pattern, as the second entry of each vector is subsumed into the pattern dots.

Proof of Theorem 882, shorter. Let $\mathbf{u}$ and $\mathbf{v}$ in $\mathbb{R}^n$ be arbitrary. So $\mathbf{u} = (u_1, \ldots, u_n)$ and $\mathbf{v} = (v_1, \ldots, v_n)$.

Let $c \in \mathbb{R}$ be arbitrary. We will prove $c(\mathbf{u}+\mathbf{v}) = c\mathbf{u}+c\mathbf{v}$. Then,

$$\begin{aligned}
c(\mathbf{u}+\mathbf{v}) &= c\left(\begin{bmatrix} u_1 \\ \vdots \\ u_n \end{bmatrix} + \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix}\right) \\
&= c\begin{bmatrix} u_1+v_1 \\ \vdots \\ u_n+v_n \end{bmatrix} \\
&= \begin{bmatrix} c(u_1+v_1) \\ \vdots \\ c(u_n+v_n) \end{bmatrix} \\
&= \begin{bmatrix} cu_1+cv_1 \\ \vdots \\ cu_n+cv_n \end{bmatrix} \\
&= \begin{bmatrix} cu_1 \\ \vdots \\ cu_n \end{bmatrix} + \begin{bmatrix} cv_1 \\ \vdots \\ cv_n \end{bmatrix} \\
&= c\begin{bmatrix} u_1 \\ \vdots \\ u_n \end{bmatrix} + c\begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix} \\
&= c\mathbf{u}+c\mathbf{v},
\end{aligned}$$

as desired. □

As a matter of personal taste, the same proof can be written using the horizontal notation for vectors:

Proof of Theorem 882, horizontal vector notation. Let $\mathbf{u}$ and $\mathbf{v}$ in $\mathbb{R}^n$ be arbitrary. So $\mathbf{u} = (u_1,\ldots,u_n)$ and $\mathbf{v} = (v_1,\ldots,v_n)$. Let $c \in \mathbb{R}$ be arbitrary. We will prove $c(\mathbf{u}+\mathbf{v}) = c\mathbf{u}+c\mathbf{v}$. Then,

$$\begin{aligned}
c(\mathbf{u}+\mathbf{v}) &= c\Big((u_1,u_2,\ldots,u_n)+(v_1,v_2,\ldots,v_n)\Big) \\
&= c(u_1+v_1,\ldots,u_n+v_n) \\
&= (c(u_1+v_1),\ldots,c(u_n+v_n)) \\
&= (cu_1+cv_1,\ldots,cu_n+cv_n) \\
&= (cu_1,\ldots,cu_n)+(cv_1,\ldots,cv_n) \\
&= c(u_1,\ldots,u_n)+c(v_1,\ldots,v_n) \\
&= c\mathbf{u}+c\mathbf{v},
\end{aligned}$$

as desired. □

The proof just presented is (other than notation) completely identical to the previous proof. It is a bit harder to see what's going on, as some parentheses are using as grouping symbols, while other parentheses are part of the horizontal vector notation. Based on the proof above using horizontal vector notation, which is correct, it is tempting to write a "proof" which is incorrect. Can you spot what's wrong?

Warning 887: Find the error in this "proof" of Theorem 882

Not a proof. Let $\mathbf{u}$ and $\mathbf{v}$ in $\mathbb{R}^n$ be arbitrary. So $\mathbf{u} = (u_1, \dots, u_n)$ and $\mathbf{v} = (v_1, \dots, v_n)$. Let $c \in \mathbb{R}$ be arbitrary. We will prove $c(\mathbf{u}+\mathbf{v}) = c\mathbf{u} + c\mathbf{v}$. Then,

$$\begin{aligned} c(\mathbf{u}+\mathbf{v}) &= c\Big((u_1, u_2, \dots, u_n) + (v_1, v_2, \dots, v_n)\Big) \\ &= c(u_1+v_1, \dots, u_n+v_n) \\ &= c(u_1+v_1) + \cdots + c(u_n+v_n) \\ &= cu_1 + cv_1 + \cdots + cu_n + cv_n \\ &= (cu_1 + \cdots + cu_n) + (cv_1 + \cdots + cv_n) \\ &= c(u_1, \dots, u_n) + c(v_1, \dots, v_n) \\ &= c\mathbf{u} + c\mathbf{v}, \end{aligned}$$

as desired. □

Discussion of error: While there is more than one error, note that $c(u_1+v_1) + \cdots + c(u_n+v_n)$ and $cu_1 + cv_1 + \cdots + cu_n + cv_n$ and $(cu_1 + \cdots + cu_n) + (cv_1 + \cdots + cv_n)$ are all scalars. Warning 872 reminds us that a scalar times a vector produces a *vector*, not a scalar: the writer of this "proof" is not applying Definition 866 properly.

We have given a detailed treatment of Theorems 878 and 882. Practice yourself by looking at the other six theorems stating properties about vector addition and/or scalar multiplication.

8.3.4 Linear combination and span

Definition 888: Linear combination

Let $\mathbf{v}_1, \dots, \mathbf{v}_s$ be in $\mathbb{R}^n$. The **linear combination** of the vectors $\mathbf{v}_1, \dots, \mathbf{v}_s$ given by the scalars $c_1, \dots, c_s$ (which are called **weights**) is

$$c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_s\mathbf{v}_s.$$

The set up of the definition of linear combination considers s vectors, which we have labeled $\mathbf{v}_1, \dots, \mathbf{v}_s$. Each of these s vectors lives in $\mathbb{R}^n$. In a sense, the scalars/weights $c_1, \dots, c_s$ are also part of the set up.

What kind of thing is obtained in a linear combination? Note the first term: $c_1\mathbf{v}_1$. From Remark 867, $c_1\mathbf{v}_1$ is a vector. Similarly, $c_2\mathbf{v}_2$ is a vector, and so on. Then, the expression defining what a linear combination is is really a sum of s vectors. From Remark 858, this will result in a vector. Thus, a linear combination of vectors is a vector (and thus a noun).

Example 889. *Let us fix $n = 2$. Let $\mathbf{v}_1 = (3,4)$ and $\mathbf{v}_2 = (5,5)$ and $\mathbf{v}_3 = (-1,10)$. If we choose weights $c_1 = 2$ and $c_2 = 0$ and $c_3 = 9$, then*

$$\begin{aligned} c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + c_3\mathbf{v}_3 &= 2(3,4) + 0(5,5) + 9(-1,10) \\ &= (6,8) + (0,0) + (-9,-90) \\ &= (-3,-82). \end{aligned}$$

Thus, the vector $(-3,-82)$ is a linear combination of the vectors $\mathbf{v}_1$, $\mathbf{v}_2$, and $\mathbf{v}_3$.

Language Discussion 890

Notice the language "the vector $(-3, -82)$ is a linear combination of the vectors $\mathbf{v}_1$, $\mathbf{v}_2$, and $\mathbf{v}_3$" with emphasis on the phrase "linear combination of the vectors." It is helpful to mention which vectors are used in the sum, and this appears after the phrase "of the vectors." Notice that the vector $(-3, -82)$ appears before the words "is a linear combination" because it is $(-3, -82)$ that is the linear combination.

Definition 891: Span

Let $\mathbf{v}_1, \ldots, \mathbf{v}_s$ be in $\mathbb{R}^n$. The **span** of the vectors $\mathbf{v}_1, \ldots, \mathbf{v}_s$ is the set of all linear combinations of $\mathbf{v}_1, \ldots, \mathbf{v}_s$. That is, the **span** of the vectors $\mathbf{v}_1, \ldots, \mathbf{v}_s$ is

$$\{c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_s\mathbf{v}_s : c_1, c_2, \ldots, c_s \in \mathbb{R}\}.$$

Notice that the span of the vectors $\mathbf{v}_1, \ldots, \mathbf{v}_s$ is a set, and a set is a noun. This set is written in build running through set format, first mentioned in Section 4.1.3.

Example 892. *Let us fix* $n = 2$. *Let* $\mathbf{v}_1 = (3, 4)$ *and* $\mathbf{v}_2 = (5, 5)$ *and* $\mathbf{v}_3 = (-1, 10)$. *If we choose weights* $c_1 = 2$ *and* $c_2 = 0$ *and* $c_3 = 9$*, then*

$$\begin{aligned} c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + c_3\mathbf{v}_3 &= 2(3, 4) + 0(5, 5) + 9(-1, 10) \\ &= (6, 8) + (0, 0) + (-9, -90) \\ &= (-3, -82). \end{aligned}$$

Thus, the vector $(-3, -82)$ *is in the span of* $\mathbf{v}_1$, $\mathbf{v}_2$, *and* $\mathbf{v}_3$. *However, if we chose different weights* $c_1 = 2$ *and* $c_2 = 2$ *and* $c_3 = 1$*, then*

$$\begin{aligned} c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + c_3\mathbf{v}_3 &= 2(3, 4) + 2(5, 5) + 1(-1, 10) \\ &= (6, 8) + (10, 10) + (-1, -10) \\ &= (15, 8), \end{aligned}$$

so the vector $(15, 8)$ *is also in the span of* $\mathbf{v}_1$, $\mathbf{v}_2$, *and* $\mathbf{v}_3$.

If we fix vectors $\mathbf{v}_1, \ldots, \mathbf{v}_s$ in $\mathbb{R}^n$, then a linear combination of $\mathbf{v}_1, \ldots, \mathbf{v}_s$ is a *single* vector while the span of $\mathbf{v}_1, \ldots, \mathbf{v}_s$ is the set of *all possible* linear combinations of $\mathbf{v}_1, \ldots, \mathbf{v}_s$. A set can have many things. Our last example shows that the span the provided three vectors has the $(-3, -82)$ and $(15, 8)$, and probably many more that were not mentioned – just pick different weights!

Definition 893: Spanned by

Let $\mathbf{v}_1, \ldots, \mathbf{v}_s$ be in $\mathbb{R}^n$. Let $\mathbf{w} \in \mathbb{R}^n$. We say that $\mathbf{w}$ is **spanned by** $\mathbf{v}_1, \ldots, \mathbf{v}_s$ if $\mathbf{w}$ is in the span of $\mathbf{v}_1, \ldots, \mathbf{v}_s$.

Definition 894: Span

Let $\mathbf{v}_1, \ldots, \mathbf{v}_s$ be in $\mathbb{R}^n$. Let H be a set. We say that $\mathbf{v}_1, \ldots, \mathbf{v}_s$ **span** H if for all vectors $\mathbf{w} \in H$, we have that $\mathbf{w}$ is in the span of $\mathbf{v}_1, \ldots, \mathbf{v}_s$.

Both of these definitions use (variants of) the word span. The first of these is using span as a verb in its participle form. The second of these is using span as a transitive verb. The grammar provides the context necessary to distinguish between the three different (but related) definitions.

Example 895. *Let* $\mathbf{v}_1 = (1, 0)$ *and* $\mathbf{v}_2 = (0, 1)$ *and* $\mathbf{v}_3 = (4, 5)$. *Let* $H = \mathbb{R}^2$. *Since every vector in* H *can be written as a linear combination of* $\mathbf{v}_1$, $\mathbf{v}_2$, *and* $\mathbf{v}_3$, *we say that* $\mathbf{v}_1$, $\mathbf{v}_2$, *and* $\mathbf{v}_3$ *span* $H = \mathbb{R}^2$.

8.3.5 Linear independence

Definition 896: Linearly independent and linear dependent

Let $\mathbf{v}_1, \ldots, \mathbf{v}_s$ be in $\mathbb{R}^n$. The vectors $\mathbf{v}_1, \ldots, \mathbf{v}_s$ be in $\mathbb{R}^n$ are **linearly independent** if the only solution to the equation

$$c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_s\mathbf{v}_s = \mathbf{0}$$

is the solution where $c_1 = 0$ and $c_2 = 0$, and so on up to $c_s = 0$. (This solution is called the **trivial solution**.)

The vectors $\mathbf{v}_1, \ldots, \mathbf{v}_s$ be in $\mathbb{R}^n$ are **linearly dependent** if there is a nontrivial solution to the equation

$$c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_s\mathbf{v}_s = \mathbf{0}.$$

Following Habit 2, being linearly independent is an adjective (or a "state of being"). It is an adjective that can only apply to a collection of vectors. Similarly, being linearly dependent is an adjective, which can only apply to a collection of vectors.

Example 897. *It makes no sense to say that the scalars 3, 4, and 5 are linearly independent. This is discussed in Warning 3. The status of being linearly independent should only be applied to a collection of vectors – not to a collection of scalars.*

Example 898. *Similarly, following Warning 3, there is no grammatical meaning to say that an equation is linearly independent. There is no grammatical meaning to saying that a system of linear equations is linearly independent.*

8.3.6 Matrices

Definition 899: Matrix

A **matrix** of size $m \times n$ is a rectangular array of mn real numbers, arranged in the shape of m rows and n columns.

If A is an $m \times n$ matrix, we use $a_{i,j}$ to denote the real number located in the ith row and jth column of A.

Remark 900. *A vector in $\mathbb{R}^n$ is an $n \times 1$ matrix.*

Method 901: Access to the columns of a matrix

Similar to Method 877, once you have established that you have an $m \times n$ matrix called A, there are proofs where it may be helpful to then say "Let $\mathbf{a}_1, \ldots, \mathbf{a}_n$ be the columns of A." (Note, when doing this, each of the vectors $\mathbf{a}_1$, $\mathbf{a}_2$ and so on all belong to $\mathbb{R}^m$.) So each of these vectors (there are n of them) have m entries.

Definition 902: Matrix equality

Let A and B be $m \times n$ matrices. Then A and B are **equal** if their corresponding columns are equal. To clarify, the matrices A and B are equal if $\mathbf{a}_1 = \mathbf{b}_1$ and $\mathbf{a}_2 = \mathbf{b}_2$ and so on, where $\mathbf{a}_j$ is the jth column of A and $\mathbf{b}_j$ is the jth column of B.

It turns out that for matrices to be equal, their corresponding entries need to be equal.

Warning 903

A matrix is not the same as an equation. We cannot speak of a matrix being consistent. We can only speak of an equation (or a system of linear equations) being consistent.

8.3.7 Transformations

A more accurate definition of function is given in Definition 551, which can be referenced for those who have studied binary relations (Section 4.6).

Definition 904: Function (familiar notation), domain, codomain

A **function** f from a set A to a set B is a rule satisfying (1) for all $a \in A$, there is a $b \in B$ such that $f(a) = b$, and (2) for all $a \in A$ and all $b, c \in B$, if $f(a) = b$ and $f(a) = c$, then $b = c$.

The set A is the **domain** of f and the set B is the **codomain** of f.

Instead of the word **function** other texts may use the word **map**, **mapping**, or **transformation**. The word **transformation** is typically used in linear algebra, but the words "function" and "transformation" are synonymous. We often write $f : A \to B$ as notation to mean that f is a function from A to B.

Method 563 and Warning 566 provide some cautions involved in defining a transformation. Warning 572 describes the nuances between f and $f(x)$.

Recall the definition of **range** in Definition 596, which stated: Let $f : A \to B$. Then the **range of f** is

$$\{f(a) : a \in A\}.$$

As defined, the range is written in build running through set notation, so we may convert this and write instead

$$\{b : \text{there exists } a \in A \text{ such that } f(a) = b\}.$$

Due to the fact that the definition of a transformation tells us that $f(a)$ is always in the codomain, we could even write

$$\{b \in B : \text{there exists } a \in A \text{ such that } f(a) = b\}.$$

Example 905. *The range of $T : \mathbb{R} \to \mathbb{R}$ defined by $T(x) = 1 + x$ is $\mathbb{R}$. Note that the codomain of the transformation T is also $\mathbb{R}$.*

Example 906. *The range of $T : \mathbb{R} \to \mathbb{R}$ defined by $T(x) = 1 + x^2$ is $[1, \infty)$. Note that the codomain of the transformation T is $\mathbb{R}$.*

Example 907. *The range of $f : \mathbb{R} \to [1, \infty)$ defined by $f(x) = 1 + x^2$ is $[1, \infty)$. Note that the codomain of the transformation f is also $[1, \infty)$.*

Example 908. *With the setup of Example 569, the range of f is $[0, 2]$ while the codomain of f is $\mathbb{R}$. The range of g is $[0, 2]$ and the codomain of g is also $[0, 2]$.*

The last example dealt with two transformations, named f and g. Many situations deal with only one transformation, such as in Example 905. In that case, instead of naming the transformation T, it is possible to describe the function by placing the symbol $\mapsto$ between input and output. Here is a full example:

Example 909. *Consider the transformation from $\mathbb{R}$ to $\mathbb{R}$ defined by $x \mapsto 1 + x$. The description here defines the same transformation which was described in Example 905, but without naming the transformation.*

The definition of surjective for transformations is copied from Definition 606:

Definition 910: Surjective

A transformation $T : A \to B$ is **surjective** if for all $y \in B$, there exists an $x \in A$ such that $T(x) = y$.

Following Habit 2, surjective is an adjective. Since surjective is an adjective which applies to transformations, following Warning 3, we should not apply this adjective to anything which is *not* a transformation.

Definition 911: Onto

A function $T : A \to B$ is **onto** if for all $y \in B$, there exists an $x \in A$ such that $T(x) = y$.

Example 912. *The range of $f : \mathbb{R} \to \mathbb{R}$ defined by $f(x) = 1 + x$ is $\mathbb{R}$. Note that the codomain of f is also $\mathbb{R}$. Thus, f is surjective.*

Example 913. *The range $f : \mathbb{R} \to \mathbb{R}$ defined by $f(x) = 1 + x^2$ is $[1, \infty)$. Note that the codomain of f is $\mathbb{R}$. Note that f is not surjective since there does not exist an $x \in \mathbb{R}$ such that $f(x) = \frac{1}{2}$, yet $\frac{1}{2}$ is in the codomain.*

Example 914. *The range $f : \mathbb{R} \to [1, \infty)$ defined by $f(x) = 1 + x^2$ is $[1, \infty)$. Note that the codomain of f is also $[1, \infty)$. Thus, f is onto.*

Example 915. *Both in Example 569 and in Example 570, f is not surjective and g is surjective.*

The definition of injective for transformations is copied from Definition 618:

Definition 916: Injective

A transformation $T : A \to B$ is **injective** if for all $w, x \in A$, if $T(w) = T(x)$, then $w = x$.

Following Habit 2, injective is an adjective. What kind of noun does injective modify? Based on the definition, injective is an adjective which applies to transformations. As an example of Warning3, it is forbidden to use the adjective injective on anything which is *not* a transformation.

Definition 917: One-to-one

A function $T : A \to B$ is **one-to-one** if for all $w, x \in A$, if $T(w) = T(x)$, then $w = x$.

Habit 918

It is tempting to think of the definition injective/one-to-one as 14 or so separate words, phrases, or bits of notation. Thinking of T, then A, then arrow, then B, then “injective” then, “if”, then “for all”, and so on is not sustainable. Instead, consider the advice of Section 3.3. Think of something “wordy” to serve as your memory hook for the definition. As an example, a transformation is injective if the same outputs lead to the same inputs.

Example 919. *The transformation $T : \mathbb{R} \to \mathbb{R}$ defined by $T(x) = 1 + x$ is injective.*

Example 920. *The function $T : \mathbb{R} \to \mathbb{R}$ defined by $T(x) = 1 + x^2$ is not injective since $T(3) = T(-3)$ yet $3 \neq -3$.*

8.3.8 Linear transformations

Linear algebra is concerned with a special type of transformation, where the domain is typically $\mathbb{R}^n$ and the codomain is typically $\mathbb{R}^m$, which also obeys some additional behavior:

Definition 921: Linear Transformation

A transformation $T : \mathbb{R}^n \to \mathbb{R}^m$ is **linear** if

- For all $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$, we have $T(\mathbf{u} + \mathbf{v}) = T(\mathbf{u}) + T(\mathbf{v})$.
- For all $c \in \mathbb{R}$, for all $\mathbf{u} \in \mathbb{R}^n$, we have $T(c\mathbf{u}) = cT(\mathbf{u})$.

Example 922. *The transformation T from $\mathbb{R}^3$ to $\mathbb{R}^5$ given by the rule*

$$T\left(\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}\right) = \begin{bmatrix} 8x_1 - 3x_2 + x_3 \\ x_1 + x_2 + x_3 \\ 5x_1 - 302x_3 \\ 4x_1 + x_3 \\ x_2 \end{bmatrix}$$

is linear.

Example 923. *The transformation T from $\mathbb{R}^3$ to $\mathbb{R}^5$ given by the rule*

$$T\left(\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}\right) = \begin{bmatrix} 8x_1 - 3x_2 + x_3 \\ x_1 + x_2 + x_3 \\ 5x_1 - 302x_3 \\ 4(x_1 + x_3)^{3879} \\ x_2 \end{bmatrix}$$

is not linear.

Example 924. *Let A be any $m \times n$ matrix. Then the transformation T from $\mathbb{R}^n$ to $\mathbb{R}^m$ defined by the rule*

$$T(\mathbf{x}) = A\mathbf{x}$$

is linear.

To describe the same example using the $\mapsto$ notation, we can write the following:

Example 925. *Let A be any $m \times n$ matrix. Then the transformation from $\mathbb{R}^n$ to $\mathbb{R}^m$ defined by the rule*

$$\mathbf{x} \mapsto A\mathbf{x}$$

is linear.

How would Warning 480 apply here? When stating the definition of a linear transformation, you should mention the word transformation, but you should not (at the same time) define what a transformation is. Write to an audience who already knows what the definition of transformation is.

8.3.9 Invertibility

Definition 926: Invertible

An $n \times n$ matrix A is **invertible** if there exists an $n \times n$ matrix Z such that $ZA = I$ and $AZ = I$, where I is the $n \times n$ identity matrix.

Theorem 927. *Let A be an $n \times n$ matrix. The following are equivalent:*

1. *The matrix A is invertible.*
2. *The equation $A\mathbf{x} = \mathbf{0}$ has only the trivial solution.*
3. *For all $\mathbf{b} \in \mathbb{R}^n$, the equation $A\mathbf{x} = \mathbf{b}$ has at least one solution.*
4. *The columns of A are linearly independent.*
5. *The columns of A span $\mathbb{R}^n$.*
6. *The linear transformation $\mathbf{x} \mapsto A\mathbf{x}$ is one-to-one.*
7. *The linear transformation $\mathbf{x} \mapsto A\mathbf{x}$ is onto.*

Of course, another way to write the last condition is to write "The linear transformation T defined by $T(\mathbf{x}) = A\mathbf{x}$ is onto." Notice that each of the conditions mentions the matrix A.

Habit 306 mentioned that a characterization (such as Theorem 927 should not replace a definition. If asked to recite the definition of an invertible matrix, the text you write should closely align with Definition 926.

What kind of creatures are mentioned in each condition? Let's inventory:

1. The first condition mentions a matrix A. (The remaining conditions also mention A, but they each mention something else too. For the rest of the list, we will focus on the main creature introduced in the condition.)

2. The second condition mentions an equation $A\mathbf{x} = \mathbf{0}$. In fact, $A\mathbf{x} = \mathbf{0}$ is a matrix equation.

3. The third condition mentions matrix equations of the form $A\mathbf{x} = \mathbf{b}$, one for each and every vector $\mathbf{b} \in \mathbb{R}^n$.

4. The fourth condition mentions the columns of the matrix A. The columns of a matrix form a set of vectors.

5. Similarly, the fifth condition mentions a set of vectors (namely, the columns of A.)

6. The sixth condition mentions a linear transformation.

7. The seventh condition mentions a linear transformation.

Let us stay with this notation, where A is a matrix, T is the linear transformation with rule $\mathbf{x} \mapsto A\mathbf{x}$, and so on. Then, what does Warning 3 say in these specific contexts?

Warning 928

We cannot write $\boxed{A \text{ has a solution}}$ because A is a matrix, while "having a solution" is something that an equation (or a system of equations) can have, not a matrix.

Warning 929

We cannot write $\boxed{A \text{ is linearly independent}}$ because A is a matrix, while being linearly independent is an adjective that applies to a set of vectors, not to a matrix.

Warning 930

We cannot write $\boxed{A \text{ spans } \mathbb{R}^n}$ because A is a matrix, while spanning is something a set of vectors can do, not something a matrix can do.

Warning 931

We cannot write $\boxed{A \text{ is one-to-one}}$ because A is a matrix, while the word $\boxed{\text{one-to-one}}$ only applies to a transformation, not to a matrix.

Warning 932

We cannot write $\boxed{A \text{ is onto}}$ because A is a matrix, while the word $\boxed{\text{onto}}$ only applies to a transformation, not to a matrix.

There are many other versions of violating the idea given in Warning 3. For example:

Warning 933

Even if we had previously established that T was defined to be the transformation $T : \mathbb{R}^n \to \mathbb{R}^n$ given by $T(\mathbf{x}) = A\mathbf{x}$, we cannot write $\boxed{T \text{ has a non-trivial solution}}$ because T is a transformation, not an equation. Having a solution (or having a non-trivial solution) is something that only a linear equation (or a system of linear equations, or a matrix equation) can do. However T is not a linear equation, T is not a system of linear equations, and T is not a matrix equation.

Warning 934

Even if we had previously established that T was defined to be the transformation $T : \mathbb{R}^n \to \mathbb{R}^n$ given by $T(\mathbf{x}) = A\mathbf{x}$, we cannot write $\boxed{T \text{ is linearly independent}}$ because T is a linear transformation, while being linearly independent is an adjective that applies to a set of vectors, not to a linear transformation.

Warning 935

Even if we had previously established that T was defined to be the transformation $T : \mathbb{R}^n \to \mathbb{R}^n$ given by $T(\mathbf{x}) = A\mathbf{x}$, we cannot write $\boxed{T \text{ spans } \mathbb{R}^n}$ because T is a linear transformation, while spanning is something a set of vectors can do, not something a linear transformation can do.

Warning 936

Be careful to distinguish between a [single] vector and a set of vectors. Writing $(4, 5, -7)$ is a single vector in $\mathbb{R}^3$. Writing $\text{Span}(4, 5 - 7)$ denotes a certain *set* of vectors, as defined in Definition 891. The former is just the single vector $(4, 5, -7)$, while the latter is the subset of $\mathbb{R}^3$ consisting of all scalar multiples of $(4, 5, -7)$.

8.3.10 Crash course in linear algebra for proof practice

The definitions and notation in this section are based on *Linear Algebra and its Applications* (5th edition) by Lay, Lay, and McDonald.

Definition 937. *Given positive integers m and n, an $m \times n$* **matrix** *is*[1] *is a rectangular array of [real] numbers with m rows and n columns. If an $m \times n$ matrix is denoted by A, the entry in the ith row and jth column (also called the (i, j)-entry) is denoted $A_{i,j}$. A matrix can alternately be viewed as a function $M : \{1, \ldots, m\} \times \{1, \ldots, n\} \to \mathbb{R}$, with $M(i, j)$ in this notation corresponding to $A_{i,j}$ in the earlier notation.*

Example 938. *Consider*

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix},$$

which is an example of a 2×2 matrix.

Definition 939. *The 2×2 matrix in Example 938 is called the* **identity matrix** *of size 2×2. This matrix is denoted I_2, or just I if the context of matrix size is clear.*

Definition 940. *Two $m \times n$ matrices A and B are* **equal** *if $A_{i,j} = B_{i,j}$ for all $i \in \{1, \ldots, m\}$ and all $j \in \{1, \ldots, n\}$.*

Definition 941. *We define some operations for matrices and real numbers:*

- **matrix addition***: Given $m \times n$ matrices A and B, the (i, j)-entry of the $m \times n$ matrix $A + B$ is $A_{i,j} + B_{i,j}$.*

[1] The plural of matrix is **matrices**, yet "matricey" is not the singular.

- **scalar multiplication**: *Given an $m \times n$ matrix A and a real number r, the (i,j)-entry of the $m \times n$ matrix rA is $r\,A_{i,j}$.*
- **matrix multiplication**: *Given an $m \times n$ matrix A and an $n \times p$ matrix B, the (i,j)-entry of the $m \times p$ matrix AB is $A_{i,1}B_{1,j} + A_{i,2}B_{2,j} + \cdots + A_{i,n}B_{n,j}$.*

Definition 942. *Let V and W be sets for which addition and scalar multiplication are defined. A function $f : V \to W$ is* **linear** *if:*

1. $f(a + b) = f(a) + f(b)$ *for all* $a, b \in V$.
2. $f(cu) = c\,f(u)$ *for all scalars* $c \in \mathbb{R}$ *and for all* $u \in V$.

Theorem 943. *Let m and n be fixed positive integers. Let A be any $m \times n$ matrix. Let X be the vector space of $n \times 1$ matrices*[2]*. Let Y be the vector space of $m \times 1$ matrices*[3]*. Then $f : X \to Y$ defined by the rule $f(x) = Ax$ for each $x \in X$ is linear.*

Definition 944. *An $n \times n$ matrix A is* **invertible** *if there is an $n \times n$ matrix C such that $CA = I$ and $AC = I$, where $I = I_n$ is the $n \times n$ identity matrix. In this case, C is an* **inverse** *of A.*

Definition 945. *The $n \times n$ matrix A is* **similar** *to the $n \times n$ matrix B if there exists an invertible matrix P such that $A = PBP^{-1}$.*

Definition 946. *If A is the 2×2 matrix*

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix},$$

we define the **determinant** *of A to be $\det(A) = ad - bc$.*

Theorem 947. *For all 2×2 matrices A and B, the equation $\det(AB) = \det(A)\det(B)$ holds.*

Theorem 948. *The 2×2 matrix A is invertible if and only if $\det(A) \neq 0$.*

Theorem 949. *The set of 2×2 invertible matrices forms a group using the binary operation of [matrix] multiplication.*

The exercises below make use of the definitions and theorems stated (above) in this section:

Exercise 950. *Let m and n be fixed positive integers. Let A be any $m \times n$ matrix. Let X be the vector space of $n \times 1$ matrices (also called column vectors of dimension n). Let Y be the vector space of $m \times 1$ matrices (also called column vectors of dimension m). Let us denote the function $f : X \to Y$ defined by the rule $f(x) = Ax$ for each $x \in X$. Prove that f is linear.*

Exercise 951. *Prove that a 2×2 matrix is invertible if and only if its determinant is non-zero. Hint: Let A be a 2×2 matrix. Prove if*

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

Exercise 952. *Prove: for all 2×2 matrices A and B, the equation $\det(AB) = \det(A)\det(B)$ holds.*

Exercise 953. *Prove: for all 2×2 matrices A and B, if AB is invertible, then $(AB)^{-1} = B^{-1}A^{-1}$.*

Exercise 954. *Prove the set of $m \times n$ matrices forms a group under matrix addition. (Is this group abelian or not? Give a proof or counterexample.) [Clarification: first fix an m and n. The same argument should work for any m, n.]*

Exercise 955. *Prove the set of 2×2 invertible matrices forms a group under matrix multiplication. (Is this group abelian or not? Give a proof or counterexample.)*

Exercise 956. *Prove that similarity is an equivalence relation on the set of 2×2 matrices. In other words, for two 2×2 matrices A and B, define the [binary] relation $\sim$ by saying that $A \sim B$ if A and B are similar matrices, and prove that $\sim$ is an equivalence relation.*

[2] An $n \times 1$ matrix is also often called a column vector of dimension n.
[3] An $m \times 1$ matrix is also often called a column vector of dimension m.

Exercise 957. *Let T be the set of all 2×2 matrices (with real entries). Let us define the function $\phi : T \to \mathbb{R}$ by the rule $\phi(A) = \sqrt{2}\,\det(A)$. Prove $\phi : T \to \mathbb{R}$ is surjective, but not bijective.*

Index