

The Algebramicon

Patty C. Hill
Jason L. Ermer

The Algebranomicon

Patty C. Hill
Jason L. Ermer

Revised: December 3, 2014

Copyright © 2014 by Patty C. Hill and Jason L. Ermer.

The Algebranomicon is made available under the Creative Commons Attribution-ShareAlike 4.0 International License: <http://creativecommons.org/licenses/by-sa/4.0/>.

www.Algebranomicon.org

Deliciously Evil Publications, a division of YearleighCorp Industries.

Contents

Acknowledgements	iv
Introduction	v
1 Numbers	1
1.1 Story of numbers	1
1.2 Integers	6
1.3 Rational numbers	13
1.4 Order of operations	25
2 Sequences	34
2.1 Sequences and recursion	34
2.2 Geometric sequences	40
2.3 Arithmetic sequences	46
2.4 Other types of sequences	50
3 Graphs and data	57
3.1 Coordinate graphing	57
3.2 Algebraic expressions	61
3.3 Graphing a function	64
3.4 Patterns in data	68
3.5 Interpreting graphs	73
4 Fundamentals of functions	77
4.1 Mathematical relationships	77
4.2 Domain and range	85
4.3 Three families of Algebra 1	89
5 Solving equations	95
5.1 Equivalence	95
5.2 Simplified algebraic expressions	101
5.3 Properties of equality (POEs)	107
5.4 Level 1 and 2 linear equations	111

5.5	Level 3 linear equations: Multiple steps	114
5.6	Level 4 linear equations: Variables on both sides	119
5.7	Level 5 linear equations: Absolute value	124
5.8	Applications of equation solving	128
6	Proportional reasoning	135
6.1	Proportions as equations	135
6.2	Applications of proportional reasoning	141
6.3	Direct variation	146
6.4	Inverse variation	151
7	Linear functions	156
7.1	Rate of change and slope	156
7.2	Point-slope form	165
7.3	Slope-intercept form	170
7.4	Standard form	175
7.5	Parallel and perpendicular lines	181
8	Linear systems	186
8.1	Mathematical modeling	186
8.2	Solving by graphing	192
8.3	Solving by substitution	193
8.4	Solving by elimination	198
8.5	Applications of systems	201
9	Inequalities	208
9.1	Equations versus inequalities	208
9.2	One-variable inequalities	210
9.3	Solving inequalities	213
9.4	Two-variable inequalities	218
9.5	Systems of inequalities	222
9.6	Application: Linear programming	227
10	Exponential functions	232
10.1	Exponential relationships	232
10.2	Exponential growth	239
10.3	Percent change	243
10.4	Personal finance	248
10.5	Exponential decay	254
11	Exponential expressions	261
11.1	Fundamentals of exponents	261
11.2	Simplified algebraic expressions: Exponents edition	263

11.3 Product properties	265
11.4 Quotient properties	269
11.5 Raising sums to powers	275
12 Quadratic equations	279
12.1 Challenges to solving quadratics	279
12.2 Level 1 and 2 quadratics	282
12.3 Level 3 quadratics: Quadrangle method	285
12.4 Level 4 quadratics: Adjust the constant term	289
12.5 Level 5 quadratics: Adjust the linear term	292
12.6 Level 6 quadratics: Adjust the quadratic term	295
13 Radical expressions	299
13.1 Real numbers	299
13.2 Square roots	307
13.3 Simplified radical form	311
13.4 Coordinate geometry	320
14 Quadratic functions	330
14.1 Quadratic Relationships	330
14.2 Vertex form	335
14.3 Standard form of a quadratic	338
14.4 Application: Projectile motion	347
15 Polynomials	356
15.1 Building Polynomials	356
15.2 Adding, subtracting, and multiplying polynomials	361
15.3 Dividing polynomials	366
16 Factoring	369
16.1 Quadratic equations, revisited	369
16.2 Factoring level 1: Greatest common factor	372
16.3 Factoring level 2: Monic quadratics	375
16.4 Factoring levels 3 and 4: Two special cases	377
16.5 Factoring level 5: Factoring by grouping	382
16.6 Solving equations by factoring	386
A Table of prime numbers to 2000	389
Glossary	391
Bibliography	410
Index	412

Acknowledgements

Who needs thank-yous? Math dep't at Kealing... magnet directors (Hovland, Goka, Ramberg)... a general shout out to all of our students over the years? MTCA folks?

Also: should we try to recruit an editor w/ a mathematics background? Altha and Adriana both come to mind...

Thanks to Shawn Tyson for editing help. The serial comma was always a given, naturally, but who knew that “logical punctuation” was a thing?

This book was typeset using \LaTeX , so we want to thank the many program and package authors, designers, and maintainers whose work has made that possible. Thanks also to the online \LaTeX community, in particular all those who have posed and answered questions at tex.stackexchange.com.

The cover image “black scales” was designed by Alex Parker and made available through the website Subtle Patterns by Atle Mo, under the CC BY-SA 3.0 license.

I am always doing that which I cannot do, in order that I may learn how to do it.

Pablo Picasso, Spanish artist

Introduction

The intro is where we describe how our book is different.

He who refuses to do arithmetic is doomed to talk nonsense.

John McCarthy, American computer scientist

Chapter 1

Numbers

The number system that we use every day, both in mathematics class and in our regular lives, developed over many generations. Men and women from all over the world, both famous and anonymous, have helped to make mathematics what it is today.

People argue over whether mathematical ideas are “invented” or “discovered”. For example: **imaginary numbers** (which we will discuss briefly in Algebra 1, and study in more detail in Algebra 2), first appeared on the mathematical scene in the 1500s. Italian mathematician Gerolamo Cardano wrote about these new numbers in his work with certain types of problems that otherwise would have been impossible to solve. Since he was one of the first people to describe this new kind of number, we might say that Cardano *invented* imaginary numbers.

But, we now know that imaginary numbers have practical applications in, for example, electrical engineering. The laws of electromagnetism haven’t changed since the 1500s. (Well, our understanding of the laws has changed, but the physics has not.) So, maybe imaginary numbers have been there all along, lurking within the fabric of the universe. In this case, Cardano *discovered* imaginary numbers.

It’s not clear which is the more accurate description. No matter what side of the debate you find more convincing, there is a certain beautiful interconnectedness to our system of numbers and mathematical laws. To illustrate this, we’d like to tell a story.

1.1 Story of numbers

Once upon a time, there was a simple farmer. Knut Krumbli lived in rural Sweden, raising goats and making goat cheese. He and his family led an uncomplicated life, and they didn’t have much need for mathematics. In fact, they really only needed numbers to count their goats: 1 goat, 2 goats, 3 goats, 4 goats. . .

But one day, after a terrible storm, Knut went to the field to count the goats and discovered, much to his dismay, that there were no goats to count. He hadn’t needed a number to describe this situation before, but

now people were asking him hard questions, like “How many of your goats made it through that crazy storm?” (But, you know, in Swedish.)

Knut and his family couldn’t very well survive without any goats, so he went to his neighbor for help. The neighbor agreed to loan Knut some goats to restart his herd, but, of course, Knut would have to repay his goat-debt later. The village hadn’t needed to do much accounting before the storm, but now they needed a system of numbers that could keep track of debts and credits.

Over time, Knut’s family got back on their feet and thrived. They paid back their debts and eventually grew to raise more goats and to make more goat cheese than they could eat. They began to trade with their neighbors for other goods or services. Of course, everything had relative value: three wheels of goat cheese were worth two bales of hay. So, the village developed a system of numbers for describing exchange rates of this kind.

As the village grew, Knut’s family farm led the development of a booming goat cheese industry. They invested their profits into bank accounts that paid interest. In certain situations, everyone was surprised to discover, interest-bearing accounts led to a new system of numbers that no one had seen before.

Eventually, some of Knut’s ancestors emigrated to America and, years later, a pair of twins – Knut’s great-great-great-grandchildren – would grow up to change the world. But let’s not get too far ahead of ourselves. More about the Krumbli twins later...

1.1.1 Dissecting the story

It’s natural to begin our discussion of numbers exactly where Knut began: counting things. The numbers we use to count are called the **natural numbers** (also known as the counting numbers, for obvious reasons).

Natural number

A member of the list of numbers that starts 1, 2, 3, 4, ... and continues forever.
The set of all natural numbers is denoted using the symbol \mathbb{N} , so we can write

$$\mathbb{N} = \{1, 2, 3, 4, \dots\}.$$

We use the {curly braces} here to indicate that we’re collecting a group of numbers together as a **set**. A set written in this way is written in **set notation**.

The natural numbers have some interesting properties. If we add two natural numbers, their sum will always be a natural number. The same goes for multiplication: the product of two natural numbers is again a natural number.

Mathematically speaking, we call this **closure**. We say that the natural numbers are closed under the operation of addition. Also, the natural numbers are closed under the operation of multiplication.

Notice that we didn't include 0 among the natural numbers. Zero is a bit tricky because it seems like a counting number. For example, "zero" is (probably) the answer to the counting question, "How many goats are there in the room with you right now?" But if there are no goats to count, can we really count them? That's a philosophical question.¹

Practically speaking, we usually exclude 0 from the set of natural numbers. We will always be very clear when we come to situation where we want to consider 0 to be a natural number.

The natural numbers are closed under the operations of addition and multiplication, but they are *not* closed under the operation of subtraction. Sometimes, the difference of two natural numbers is a natural number: for example $8 - 6 = 2$, no problem. But some subtraction sentences don't work: for example $10 - 13 = -3$, and -3 is not a natural number. In cases like this, we don't have a natural number, but an **integer**.

Integer

A natural number, or the opposite of a natural number, or zero. The set of all integers is denoted \mathbb{Z} , so we sometimes write

$$\mathbb{Z} = \{\dots, -3, -2, -1, 0, 1, 2, 3, \dots\}.$$

We could also write $\mathbb{Z} = \{0, \pm 1, \pm 2, \pm 3, \dots\}$.

The symbol \mathbb{Z} comes from *Zahl*, the German word for number.

Note that every natural number is an integer. We say that the set of natural numbers is a **subset** of the set of integers.²

The integers are closed under the operations of addition and multiplication. Plus – bonus! – the integers are closed under the operation of subtraction. Whenever we subtract one integer from another, we always get another integer as the result.

But (as you may have noticed) we have a problem with division. In certain cases, the quotient of two integers is itself an integer: $15 \div -3 = -5$, no problem. But other times we get a quotient that is not an integer: $-3 \div 15 = -0.2$ and -0.2 is not an integer. In other words, the integers are not closed under the operation of division. To achieve closure under division, we must turn to the **rational number**.

¹ Philosophy and mathematics have historically gone hand-in-hand. Many important discoveries (inventions?) in mathematics are attributed to people who are considered both "philosopher" and "mathematician". For example: French mathematician René Descartes, for whom the Cartesian coordinate system is named, is also the philosopher who said "I think therefore I am".

² For those who are into mathematical symbols and notation, the sentence "The natural numbers are a subset of the integers" is denoted $\mathbb{N} \subseteq \mathbb{Z}$.

Rational number

Any number that can be written as the ratio of two integers $\frac{a}{b}$, where b is not zero. This set includes all of your classic fractions, as well as all terminating decimals and all repeating decimals.

The set of all rational numbers is denoted by the symbol \mathbb{Q} , which comes from the word *quotient*.

Fractions have a lousy reputation among math students, but the rational numbers are great because they are closed under all four of the basic operations. When we add, subtract, multiply, or divide any two rational numbers, the result will always be another rational number. What more could we ask for? In some sense, \mathbb{Q} is a complete number system, and in fact the world was content with the rational numbers for a long time.

But, other numbers exist. Yes, the rational numbers include all of the terminating decimals (like 0.5 and 1.678), and all of the repeating decimals (like $0.\overline{3}$ and $-12.34\overline{56}$). But consider the number

$$0.10110111011110111110\dots$$

This number does not terminate, but it does not repeat either.³ So, this number is *not* a rational number. It is an **irrational number**.

Irrational number

A number that cannot be expressed as the ratio of two integers. In decimal form, an irrational number never terminates and never repeats.

You have likely encountered irrational numbers before. A famous example is the number π (pi), which shows up when we study circles. We usually approximate π to be about 3.14, but in fact the decimal representation of π goes on forever without stopping or repeating.⁴

$$\pi \approx 3.141592653589793238462643383279502884197693993751058209749445923078164\dots$$

When we group together all of the rational numbers and all of the irrational numbers, we will have accounted for all possible decimal representations. This combined set of numbers, called the **real numbers**, is going to be of key importance to us in Algebra 1.

³ This number does have a pattern, but that is not the same as “repeating” in the sense of “repeating decimal”. In a repeating decimal the same digit or group of digits recurs over and over again.

⁴ We use the symbol \approx to mean “is approximately equal to”. We write $\pi \approx 3.14$ and not $\pi = 3.14$, because 3.14 is only an approximation of π , not its exact value.

Real number

Any rational or irrational number. The set of all real numbers is denoted \mathbb{R} which, naturally enough, comes from the word *real*.

Like the set \mathbb{Q} , the set \mathbb{R} is closed under the four fundamental operations. \mathbb{Q} and \mathbb{R} are the number systems we will work with in Algebra 1. Other types and sets of numbers exist, like \mathbb{C} , the set of so-called **complex numbers**. Complex numbers will be a topic of study in Algebra 2, but we will spend our time in Algebra 1 with \mathbb{Q} and \mathbb{R} .

1.2 Integers

You have probably been working with positive and negative numbers for a while now, so this section will review the most important terms and algorithms for signed numbers. To get the ball rolling, think about this:

Integer comparisons

// startup exploration

In each of the expressions below, assume that x and y are natural numbers and that $x < y$ (x is less than y). Will the result be greater than 0, less than 0, equal to 0, or is there not enough information to tell? Why?

(a) $x + (-y)$

(b) $x - (-y)$

(c) $x \cdot (-y)$

(d) $x \div (-y)$

1.2.1 Language of signed numbers

As we saw in section 1.1, the integers include the natural numbers, their opposites, and zero. Numbers now include two pieces of information: they have a *size* (or *magnitude*) and a *direction*, either positive or negative.⁵ Sometimes, we care only about the magnitude of a number, in which case we refer to a number's **absolute value**.

Absolute value

The absolute value of a number x is its distance away from zero on the number line. To express this in mathematical symbols, we write $|x|$ to mean “the absolute value of x ”.

For those who like memory aids, it may be helpful to think of the absolute value bars as a little numerical shower stall or car wash: a number goes in and all its negativity gets washed away.

⁵ Later in mathematics and the physical sciences, we'll encounter mathematical objects with both magnitude and direction again. They're called vectors.

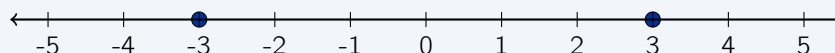
Example 1.1

1. $|4|$ The absolute value of a positive number is just the original number, so $|4| = 4$.
2. $|-8|$ Absolute value ignores the sign and tells us how far a number is from zero. -8 is eight units away from zero, so $|-8| = 8$.
3. $-|-6|$ The absolute value bars only apply to what's inside, but then this negative sign *outside* the absolute value bars will make the final answer negative again! So, $-|-6| = -6$. Note that $-|6|$ would also be -6 .

Every nonzero number has a counterpart that is the same distance away from zero, but on the opposite side of the number line. This is a simple idea, but one that is important enough for us to give it a name.

Opposite

The **opposite** of a number x is the number $-x$. In other words, the opposite of a number is the number with the *same absolute value*, but the *opposite sign*. For example, 3 and -3 are opposites. The number 0 is its own opposite.



The sum of opposites is always 0. For this reason, we sometimes use the term **additive inverse** to describe the opposite of a number. (More on inverses in chapter 5.)

1.2.2 Adding signed numbers

Like matter and antimatter, combining positive and negative numbers leads to annihilation.⁶ For example when we bring together $+8$ and -6 , we can picture 8 units of “matter” and 6 units of “antimatter”. Particles and antiparticles annihilate one another, both disappearing in the process. Since we have more matter than antimatter in this case, all of the antimatter is consumed, leaving behind 2 units of matter.

We might visualize annihilation with a drawing like in fig. 1.1 where black circles are units of matter, red circles are units of antimatter. Or, we could write a number sentence like $8 + -6 = 2$. Of course, no annihilations occur when we scrape together a big pile of matter (or a big pile of antimatter, for that matter). It's only when they mix that anything interesting happens.

⁶ Dramatic, no?

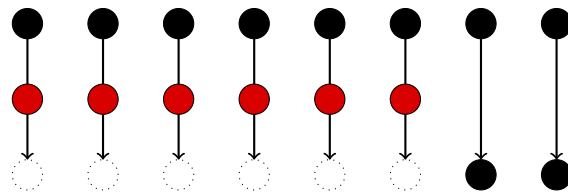


Figure 1.1: Eight units of matter versus six units of antimatter.

Adding signed numbers

When adding two numbers that have the same sign, we add the absolute values of the numbers, and use the sign that they share.

Otherwise, when adding two numbers that have different signs, we find the difference between their absolute values, and use the sign of the number with larger absolute value.

Example 1.2

Compute each of the following.

1. $-8 + -12$ Both numbers are negative. That's a big pile o' antimatter: $-8 + -12 = -20$
2. $-8 + 12$ The numbers have different signs, so prepare for annihilation! $|12|$ is larger than $|-8|$, so we will have matter left over. Therefore, $-8 + 12 = 4$

1.2.3 Subtracting signed numbers

Addition and subtraction are called *opposite* (or *inverse*) *operations* because they “undo” one another. The act of adding 5 of matter can be “undone” by subtracting 5 units of matter. On the other hand, the act of adding 5 units of matter can also be undone by *adding 5 units of antimatter*.

Subtracting signed numbers

Subtracting a number is the same as adding the opposite of the number.

When faced with a subtraction problem, change it to an “addition of the opposite” problem and then follow the rules for adding signed numbers.

The benefit of this approach is that we can avoid having to learn a whole new set of rules for subtracting signed numbers. All we need are the rules for addition, plus one new rule about how to change subtraction problems into addition problems!

Example 1.3

Compute each of the following:

1. $8 - 12$ “Subtracting 12” is the same as “adding negative 12”, so $8 - 12 = 8 + -12$.
Then, we can apply the rules for adding signed numbers: $8 + -12 = -4$
2. $-3 - 14$ This is the same as $-3 + -14$, and we follow the rules for adding numbers with the same sign: $-3 + -14 = -17$
3. $4 - 9$ This is the same as $4 + 9$, which is an easy addition: $4 + 9 = 13$
4. $-6 - 5$ This is the same as $-6 + 5 = -1$

When dealing with addition and subtraction of signed numbers in a problem, a good habit is to simplify the signs by rewriting subtraction as addition-of-the-opposite, before doing any computations.

In the startup exploration for this section, we have two natural numbers x and y where $x < y$. Natural numbers are all positive, so this means that $|x|$ is less than $|y|$.

Part (a) asks us to consider $x + (-y)$. Opposite numbers have the same absolute value, and so $|y|$ is the same as $|-y|$. Therefore when we add, the number with larger absolute value is negative and the sum will take on the negative sign. So, when x and y are natural numbers and $x < y$, we know that $x + (-y)$ is *always negative* so the result is always less than 0.

Part (b) asks us to consider $x - (-y)$. We can change this subtraction expression into addition-of-the-opposite: $x - (-y) = x + y$. Since both x and y are natural numbers, their sum is a natural number. In other words, when x and y are natural numbers, $x - (-y)$ is always positive, in other words always greater than 0.

1.2.4 Multiplying and dividing signed numbers

Like addition and subtraction, the operations of multiplication and division are inverse operations. We'll discuss multiplication below, though the rules about the signs of products also apply to the signs of quotients.

Recall from your elementary school days that one way to think about whole number multiplication is as *repeated addition*. We interpret $a \cdot b$ as “ a groups with b items in each group”. So, $3 \cdot 5$ means “3 groups of 5”, which we can write as an addition sentence: $3 \cdot 5 = 5 + 5 + 5$.

Using this interpretation, we can easily explain the product of a positive number and a negative number. The expression $4 \cdot -8$ means “four groups of negative eight”: $4 \cdot -8 = -8 + -8 + -8 + -8 = -32$. No problem!

But what about $-5 \cdot 6$? What does it mean to have “negative five groups of six”? Or even worse, what about $-3 \cdot -7$, “negative three groups of negative seven”? Rather than try to twist the metaphor to fit these new situations, let's just admit that multiplication cannot *always* be represented by repeated addition.⁷

For the moment, we'll simply state the rules for multiplying (and dividing) signed numbers. We can explain why these rules work using the so-called “field axioms for the real numbers”. More on that in chapter 5.

Multiplying (and dividing) signed numbers

When multiplying signed numbers, we determine the magnitude and sign of the product separately. (The same procedure applies when dividing.)

The absolute value of the product of two numbers is the product of their absolute values: $|a \cdot b| = |a| \cdot |b|$. This gives us the magnitude of the product (or quotient).

To determine the sign, we compare the signs of the numbers we are given. If the two numbers have the same sign, then the answer is positive. If the two numbers have opposite signs, then the answer is negative.

There are several clever ways to remember this rule. Some people remember that every pair of negatives in a product cancel one another. Other people use a triangle with one positive sign and two negative signs drawn on the vertices. We present another way of looking at it in the next section. Choose whichever mnemonic⁸ method is most helpful to you!

Karmic multiplication

Karma, an underlying concept of many Eastern religions, is the belief that a person's actions and intentions shape their future. Performing good deeds will contribute to one's “good karma” and will lead to future happiness. Bad deeds contribute to one's “bad karma” and will lead to future suffering.

So, karma suggests that good things happen to good people, and that bad things happen to bad people. Of course we know that the universe does not always operate in accordance with karma.

⁷ The “multiplication as repeated addition” analogy breaks down when we have a negative number of groups, but also for rational and irrational numbers. What's the repeated addition problem for $\frac{2}{3} \cdot \frac{1}{2}$, or for $\sqrt{2} \cdot \sqrt{3}$?

⁸ mnemonic (*na · MON · ic*, the first “m” is silent): A learning aid that helps to remember or retain information.

Karmic multiplication

When good things happen to good people, that's good!
 When bad things happen to good people, that's bad!
 When good things happen to bad people, that's bad!
 When bad things happen to bad people, that's good!

For example: Mahatma Gandhi used nonviolent means to inspire civil rights movements around the world. Gandhi was a good person. On the other hand, Adolf Hitler was chancellor of Nazi Germany during World War II and orchestrated appalling crimes against humanity. Hitler was a bad person.⁹

In terms of life events, winning the lottery is a good thing. Getting hit by a truck is a bad thing.

If Gandhi had won the lottery, that would have been in accordance with all his good karma. That's good! If Gandhi had been hit by a truck, that would have been in opposition to all of his good karma. That's bad!

If Hitler had won the lottery, that would have been in opposition to his evil karma. That's bad! If Hitler had been hit by a truck, it would have served him right! That's good! Go karma!

Of course, in this metaphor good things and good people represent positive numbers. Bad things and bad people represent negative numbers. When karma is operating as it should, we get a positive result. When the laws of karma are broken, we get a negative result.

Example 1.4

Compute each of the following:

1. $8 \cdot -12$ We multiply the absolute values and, since the two factors have different signs, we know the answer is negative: $8 \cdot -12 = -96$
2. $-72 \div -3$ We divide absolute values and, since the two factors have the same sign (both negative in this case, so that's "Hitler gets his by a truck"), the answer is positive: $-72 \div -3 = 24$

Note: When multiplying (or dividing) more than two numbers, we can approach things in two different ways. We might simplify the product two factors at a time and keep track of the sign as we go. Or, we could treat the signs as a separate problem: first multiply all of the absolute values, then go back and count up the negative signs.

⁹ understatement (*UN · der · state · ment*): The act of representing something in a weak or restrained way, to a lesser degree than is borne out by the facts.

Example 1.5

Multiply: $(2)(-2)(1)(-2)(-2)(1)(1)(-2)(-1)(2)$

Solution: If we count up the negative signs, we find there are five. Pairs of negatives will have a positive product, so we'll have two pairs of negatives plus one left over. Our final answer, then, will be negative. All that remains is to multiply the 2s (of which there are six):

$$(2)(-2)(1)(-2)(-2)(1)(1)(-2)(-1)(2) = -64$$

In the startup exploration for this section, x and y are natural numbers and $x < y$. Since x and y are natural numbers they are both positive, and then $-y$ is negative. Part (c) asks us to consider $x \cdot (-y)$. We have a positive number times a negative number, so the product is *always negative*, or in other words, always less than 0.

Part (d) asks us to consider $x \div (-y)$. Again, we have a positive number and a negative number, so the quotient is *always* less than 0.

1.3 Rational numbers

As with integers, you've probably been working with fractions and decimals for a while now. In this section, we review the key terms and algorithms for working with rational numbers. As we get going, think about this:

Rational comparisons

// startup exploration

In each of the expressions below, a and b are rational numbers where $0 < a < b < 1$. Will the result be greater than 1, less than 1, equal to 1, or is there not enough information to tell? Why?

(a) $a + b$

(b) $a - b$

(c) $a \cdot b$

(d) $a \div b$

1.3.1 Language of rational numbers

Sometimes in life we discover, much to our surprise, that some ridiculous and insignificant thing has been given a name.¹⁰ We find ourselves wondering, "Who decided to give *that* a name? Why bother?" Prepare for one of those moments:

Vinculum

A **vinculum** (plural: *vincula*) is a horizontal bar used in mathematics to show grouping. For example, the fraction bar in the middle of $\frac{5}{2}$ is a vinculum.

Vincula are used in other contexts as well. For example, we use a vinculum to represent a repeating decimal such as $0.\overline{3}$.

With that definition in mind, we can continue with two of the most daunting words in elementary mathematics: **numerator** and **denominator**. You're definitely not alone if you have ever been confused about these.

Numerator and denominator

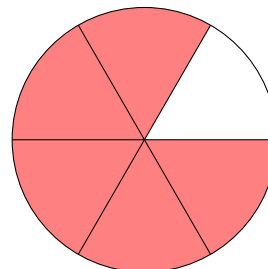
In a fraction, the number above the vinculum is called the *numerator* of the fraction, and the number below the vinculum is called the *denominator* of the fraction.

¹⁰ For instance, did you know that the little plastic sheath at the end of your shoelaces is called an "aglet"? Now you know.

For example in the number $\frac{5}{6}$, the numerator is 5 and the denominator is 6.

If we ask, “*How many sixths?*” – the numerator tells us “*five sixths*”. The word numerator is related to the word “number”, and the numerator counts the pieces.

If we ask, “*Five whats?*” – the denominator tells us “*five sixths*”. The word denominator is related to the word “nominate” (as in “to nominate someone for president”) which means “to name”. The denominator *names* the fraction.



1.3.2 Multiplying rational numbers

Don't worry, your version of the *Algebranonicon* isn't missing any sections. Most textbooks would discuss adding and subtracting rational numbers first, but we're going to start by studying the most helpful of the rational number operations: multiplication.

Suppose that at the cheese market, Knut Krumbli is selling chunks from a 10-pound block of cave-aged goat cheese.¹¹ Half of the original block of cheese is left, and a local weaver asks to buy three-fourths of it. The original block had a value of 800 Swedish kronor. How much should Knut charge the weaver?

Let's draw a picture. In fig. 1.2, the square represents the original block of cheese. We divide the square in half vertically, and the region colored yellow represents how much of the cheese remains. We can then divide the cheese into fourths horizontally and shade in three-fourths (the amount that the weaver wants to buy).

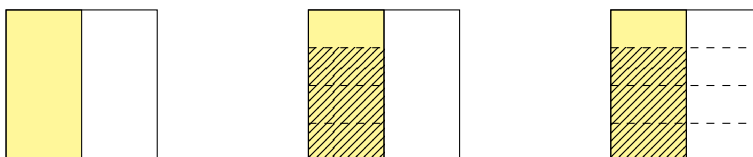


Figure 1.2: Selling three-fourths of one-half of a block of cheese.

What fraction of the whole block does this shaded region represent? If we extend the lines, we can see that we've divided up the whole block of cheese into 8 equally-sized pieces. So the weaver is buying $\frac{3}{8}$ of the whole block and Knut should charge 300 kronor.

In the end, we solved a fraction multiplication problem: “How much is three-fourths of one-half of a whole block of cheese?”

$$\frac{3}{4} \cdot \frac{1}{2} = \frac{3}{8}$$

¹¹ When cheese is aged in the cool, humid air of underground caves, it can develop a denser texture and a more complex flavor, since small salt crystals form throughout its interior.

Multiplying rational numbers

To multiply rational numbers, we multiply numerators to find the numerator of the product. We multiply denominators to find the denominator of the product.

$$\frac{a}{b} \cdot \frac{c}{d} = \frac{a \cdot c}{b \cdot d}$$

In the startup exploration for this section, $0 < a < b < 1$, and part (c) asks us to consider $a \cdot b$. Let's think about this from a block-of-cheese perspective. We start with less than a whole block of cheese (since b is less than 1) and we only want to buy a fraction of what's there (since a is less than 1).

So, we are certainly buying less than a whole block of cheese! Since both a and b are less than 1, we know that their product must be less than one.

Fancy versions of one

Recall from your elementary school days that any number times 1 is itself. This simple fact, along with the fraction multiplication procedure, gives us an extremely powerful tool that we'll use in various different ways throughout algebra. The key idea is that multiplying something by 1 doesn't change its value, even if we use a "fancy version of 1".¹² Consider, for example:

$$\begin{aligned} \frac{5}{8} &= \frac{5}{8} \cdot 1 && \text{multiplication by 1 doesn't change the number} \\ &= \frac{5}{8} \cdot \frac{7}{7} && \text{we can rewrite 1 however we like, here } 1 = \frac{7}{7} \\ &= \frac{35}{56} && \text{multiply fractions} \end{aligned}$$

In the end, we have two equivalent fractions $\frac{5}{8} = \frac{35}{56}$. The representation has changed, but the value is the same! The "fancy one" we chose here is $\frac{7}{7}$, but any other version of 1 would work the same way: $\frac{-140}{-140}$, $\frac{2\pi}{2\pi}$, $\frac{\sqrt{3}}{\sqrt{3}}$... The possibilities are endless.

The first application of the "fancy one" has to do with one of the themes we encounter throughout algebra: the idea of finding a "completely simplified" solution to a problem. Fractions introduce us to the first criteria for something being simplified.

¹² Multiplication by a fancy version of 1 is an application of the identity property of multiplication, which we will study in more detail in chapter 5.

Simplified rational numbers #1

Fractions should be simplified to **lowest terms**, meaning that the numerator and denominator of the fraction are **relatively prime** integers.

Two integers are said to be relatively prime (or coprime) if they have no common factors other than 1. So, the fraction $\frac{21}{34}$ is in lowest terms, since the factors of 21 are $\{1, 3, 7, 21\}$ and the factors of 34 are $\{1, 2, 17, 34\}$. They have no factors in common, other than 1.

On the other hand, $\frac{18}{84}$ is not in lowest terms. Both 18 and 84 are even, and so both the numerator and denominator are divisible by at least 2. To write this fraction in lowest terms, we “undo” fraction multiplication and search for some fancy ones that we can eliminate:

$$\frac{18}{84} = \frac{2 \cdot 3 \cdot 3}{2 \cdot 2 \cdot 3 \cdot 7} = \frac{2}{2} \cdot \frac{3}{3} \cdot \frac{3}{2 \cdot 7} = 1 \cdot 1 \cdot \frac{3}{14} = \frac{3}{14}$$

Once we know how and why this works, we can take a shortcut and “cancel” common factors from the numerator and denominator:

$$\frac{18}{84} = \frac{2 \cdot 3 \cdot 3}{2 \cdot 2 \cdot 3 \cdot 7} = \frac{\cancel{2} \cdot 3 \cdot 3}{\cancel{2} \cdot 2 \cdot \cancel{3} \cdot 7} = \frac{3}{2 \cdot 7}$$

Simplify before you multiply

We can also use the “fancy one” to save ourselves some work! What if we have to multiply:

$$\frac{1}{2} \cdot \frac{2}{3} \cdot \frac{3}{4} \cdot \frac{4}{5} \cdot \frac{5}{6}$$

Let’s write the product out “the long way” before we actually do any computations. The numerator of the product will be the product of the numerators, and likewise for the denominator:

$$\frac{1}{2} \cdot \frac{2}{3} \cdot \frac{3}{4} \cdot \frac{4}{5} \cdot \frac{5}{6} = \frac{1 \cdot 2 \cdot 3 \cdot 4 \cdot 5}{2 \cdot 3 \cdot 4 \cdot 5 \cdot 6}$$

Then, a common factor in the numerator and denominator is like multiplication by 1, so we can make some simplifications:

$$\frac{1}{2} \cdot \frac{2}{3} \cdot \frac{3}{4} \cdot \frac{4}{5} \cdot \frac{5}{6} = \frac{1 \cdot 2 \cdot 3 \cdot 4 \cdot 5}{2 \cdot 3 \cdot 4 \cdot 5 \cdot 6} = \frac{1 \cdot \cancel{2} \cdot \cancel{3} \cdot \cancel{4} \cdot \cancel{5}}{\cancel{2} \cdot \cancel{3} \cdot \cancel{4} \cdot \cancel{5} \cdot 6} = \frac{1}{6}$$

The alternative would have been to multiply all of those numbers together by hand and then simplify the fraction to lowest terms¹³, which is a bunch more work. It pays to be clever!

¹³ That would have been $\frac{120}{720}$.

Example 1.6

Multiply: (a) $\frac{3}{4} \cdot \frac{7}{16}$ (b) $\frac{15}{8} \cdot \frac{4}{7} \cdot \frac{14}{3} \cdot \frac{1}{5}$

Problem (a) is nothing special. No common factors are shared between numerators and denominators, so we simply multiply as usual.

$$\frac{3}{4} \cdot \frac{7}{16} = \frac{3 \cdot 7}{4 \cdot 16} = \frac{21}{64}$$

Since there were no common factors to cancel before multiplying, we know the product is already in lowest terms.

To tackle (b), a first helpful step is to factor the individual numerators and denominators to expose all of the factors that are going to be in the product.

$$\frac{3 \cdot 5}{2 \cdot 2 \cdot 2} \cdot \frac{2 \cdot 2}{7} \cdot \frac{2 \cdot 7}{3} \cdot \frac{1}{5}$$

Then, a common factor in the numerator and denominator is like multiplication by 1, so common factors can be crossed out. Look at what happens in this case!

$$\frac{\cancel{3} \cdot \cancel{5}}{\cancel{2} \cdot \cancel{2} \cdot 2} \cdot \frac{\cancel{2} \cdot \cancel{2}}{\cancel{7}} \cdot \frac{\cancel{2} \cdot \cancel{7}}{\cancel{3}} \cdot \frac{1}{\cancel{5}} = 1$$

Note that all of the factors cancel out in the denominator. A common mistake at this point is to replace the “empty” denominator with 0 – but remember, a common factor in the numerator and denominator is like a factor of 1, which is the fraction $\frac{1}{1}$, not $\frac{1}{0}$, so there’s always a “phantom one” lurking there even if we don’t write it.

1.3.3 Adding and subtracting rational numbers

Knut Krumbli would tell you that bringing together a herd of 8 goats and a herd of 11 goats results in a herd of 19 goats. It’s goat herding, not rocket science.

But, combining a herd of 4 goats and a flock of 13 sheep doesn’t really give us 17 of anything until we can find some shared characteristic that is common to both groups. For instance, we could say we have a group of “17 mammals” or “17 quadrupeds”.

So it goes with fractions. When our fractions have a shared name (a common denominator) we can total up how many things we have with that name: 8 thirds and 11 thirds makes 19 thirds. In math symbols, we write

$$\frac{8}{3} + \frac{11}{3} = \frac{19}{3}$$

On the other hand, when our quantities *don’t* share a common unit (a common denominator), we have to find one before we can add in a meaningful way.

Given a fraction, we can use multiplication by a “fancy one” to generate a new fraction that has the same value but a different, perhaps more helpful, denominator.

Some people like to try and find the *least common denominator*, but that’s not strictly necessary. Any common denominator will do. In fact, a guaranteed common denominator of any two fractions is the *product of their denominators*.

Adding and subtracting rational numbers

To add rational numbers, we must have a common denominator, for example the product of the original denominators. Then we add the numerators, and keep the common denominator.

$$\frac{a}{b} + \frac{c}{d} = \frac{a \cdot d + b \cdot c}{b \cdot d}$$

To subtract rational numbers, change the subtraction problem to an “addition of the opposite” problem and then follow the algorithm for addition.

Here’s the derivation of fraction addition in a bit more detail. Note how we use multiplication by 1 to find a common denominator $b \cdot d$:

$$\frac{a}{b} + \frac{c}{d} = \left(\frac{a}{b} \cdot 1\right) + \left(1 \cdot \frac{c}{d}\right) = \left(\frac{a}{b} \cdot \frac{d}{d}\right) + \left(\frac{b}{b} \cdot \frac{c}{d}\right) = \frac{a \cdot d}{b \cdot d} + \frac{b \cdot c}{b \cdot d} = \frac{a \cdot d + b \cdot c}{b \cdot d}$$

Negative fractions

Where should we put the negative sign when we have a negative fraction? Does it matter? Consider the following three possibilities. Are they all equivalent?

$$-\frac{3}{4} \quad \frac{-3}{4} \quad \frac{3}{-4}$$

A fraction is a way of writing a division problem. If Knut’s four children share three bowls of lingonberries equally, then each child will get $3 \div 4 = \frac{3}{4}$ of a bowl of berries.¹⁴ The fraction $\frac{3}{4}$ is just another way of writing $3 \div 4$. So, all three of the fractions above have the same value.

In the first example, the whole fraction has been negated. In the second and third examples, the numbers have opposite signs and so the quotient will be negative. In other words it actually doesn’t matter where we put the negative sign. We can put it where it is most convenient for the problem. Very often, that’s in the numerator of the fraction.

¹⁴ Lingonberries are a popular fruit in Scandinavia and throughout northern, central, and eastern Europe. The berries are quite tart, and so they are usually mixed with sugar and preserved as jam or compote. In Sweden and Norway, reindeer is traditionally served with gravy and lingonberry sauce. Yes, Scandinavians eat reindeer.

Mixed numbers

Improper fractions have a numerator that is greater than or equal (in absolute value) to their denominator, like $\frac{5}{3}$ or $\frac{-84}{16}$. Improper fractions have been scorned by many elementary school mathematics teachers, who instead prefer mixed numbers: $1\frac{2}{3}$ or $-5\frac{1}{4}$. But, improper fractions are often much easier to work with than mixed numbers.¹⁵

Simplified rational numbers #2

Simplified **improper fractions** are preferred over **mixed numbers** and decimals. Only convert to a mixed number or decimal when the context (or the directions) require it.

We usually prefer exact fraction answers over decimal approximations. Writing the decimal $\frac{10}{7}$, for instance, is much preferred over 1.43, and it's better even than the exact answer $1.\overline{428571}$ (yep, that's a big chunk o' repeating decimal).

But, be sure to read questions carefully! There are exceptions to these rules. When working in a real-world context, a certain number format may make more sense. For example, when solving a problem about money, the answer \$3.50 makes a lot more sense than $\$2\frac{7}{2}$. Likewise, the answer " $1\frac{1}{3}$ pounds of cheese" is better than " $\frac{4}{3}$ pounds of cheese". In ambiguous cases, we will make it clear what number format is preferred.

When faced with mixed numbers in a problem, we have to be careful. When adding, we can convert all mixed numbers to improper fractions, or work with them "as is". Subtracting with mixed numbers is tricky, however, because we may have to handle regrouping. Multiplication is even trickier.

Since we prefer improper fractions as final answers anyway, we recommend converting all mixed numbers to improper fractions before starting computations. To convert a mixed number to an improper fraction, all we have to do is think about the mixed number as an addition problem:

$$3\frac{5}{8} = 3 + \frac{5}{8} = \frac{3}{1} + \frac{5}{8} = \frac{3 \cdot 8 + 1 \cdot 5}{1 \cdot 8} = \frac{24 + 5}{8} = \frac{29}{8}$$

Here we used the fact that any integer has a "phantom one" in its denominator: $3 = \frac{3}{1}$. We don't usually write it, but it's there when we need it.

¹⁵ One situation where improper fractions are superior is when describing the slope of a line, as we will see in chapter 7.

Example 1.7

Compute each of the following:

1. $\frac{3}{4} + \frac{5}{6}$

These fractions do not have a common denominator, so we'll have to find one. We could use their least common denominator (which is 12) or use the product of the denominators (which is 24). Let's use 24:

$$\frac{3}{4} + \frac{5}{6} = \frac{3 \cdot 6 + 4 \cdot 5}{4 \cdot 6} = \frac{18 + 20}{24} = \frac{38}{24} = \frac{19}{12}$$

2. $1\frac{2}{5} - 1\frac{7}{8}$

First, we'll convert to improper fractions and change the subtraction to addition-of-the-opposite, putting the negative sign in the numerator of the fraction. Then we add. In the end, we can adjust the negative sign again:

$$\frac{7}{5} - \frac{15}{8} = \frac{7}{5} + \frac{-15}{8} = \frac{7 \cdot 8 + 5 \cdot -15}{5 \cdot 8} = \frac{56 + -75}{40} = \frac{-19}{40} = -\frac{19}{40}$$

In the startup exploration for this section, $0 < a < b < 1$. Part (a) asks about $a + b$. Since both numbers are positive, their sum is positive, but we don't have enough information to tell whether the sum is greater than 1. If a and b are both less than $\frac{1}{2}$, for example, then their sum will be less than 1. On the other hand, if they are both greater than $\frac{1}{2}$, then their sum will be greater than 1.

Part (b) asks us to consider $a - b$. Since a is less than b , we're subtracting a larger number from a smaller number, and so the answer must be negative.

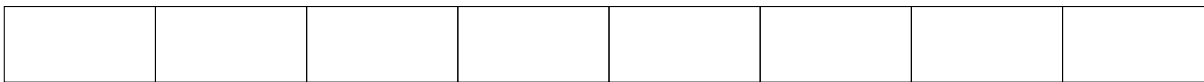
We can reason this out in another way: $a - b$ is the same as $a + (-b)$. The absolute value of b is the same as the absolute value of $-b$. And so this sum will be negative because the negative number is the one with the greater absolute value. In any case, $a - b$ is negative and so we know for sure that it is less than 1.

1.3.4 Dividing rational numbers

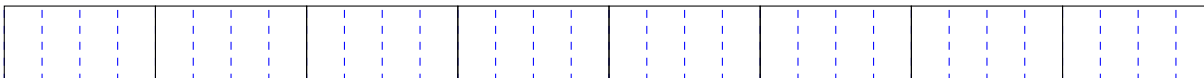
Fraction division may be the most poorly understood operation in all of arithmetic. The algorithm for dividing fractions seems arbitrary, and it's often difficult to judge whether our answers make sense. Let's pause for a moment to think about what fraction division means.

Suppose Jorunn Krumbli, Knut's wife, is making scarves for the goats (Scandinavian winters are chilly). She has 8 meters of burlap, and each scarf requires $\frac{3}{4}$ of a meter. How many scarves can she make?

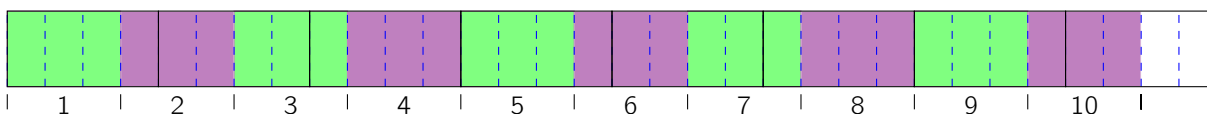
This question is asking us to compute $8 \div \frac{3}{4}$, but let's try to solve the problem by drawing a picture. Suppose this rectangle represents Jorunn's 8 meters of burlap.



To figure out how many pieces of $\frac{3}{4}$ meter are in there, let's first determine the number of pieces of size $\frac{1}{4}$ meter. To do that, we'll break each meter into four pieces. That gives us $8 \cdot 4 = 32$ pieces total.



Now, let's gather these pieces up into groups of three. We'll be able to make 10 whole groups, and we'll have 2 pieces left over. Since we have two pieces, but need three, we have $\frac{2}{3}$ of a group. So, Jorunn can make $10\frac{2}{3}$ scarves for the goats.



Let's retrace our steps. We set out to solve $8 \div \frac{3}{4}$, but in our picture we first multiplied to find the total number of fourths, and then we divided to find how many groups of three we could make. In other words: $8 \div \frac{3}{4}$ must have the same answer as $8 \cdot \frac{4}{3}$. Does that look familiar?

Dividing rational numbers

Dividing by a number is the same as multiplying by the reciprocal of the number. So, we can change a given fraction division problem into an equivalent fraction multiplication problem and then use the rules for fraction multiplication.

$$\frac{a}{b} \div \frac{c}{d} = \frac{a}{b} \cdot \frac{d}{c} = \frac{a \cdot d}{b \cdot c}$$

Recall that the **reciprocal** of a fraction is the fraction that interchanges the numerator and the denominator of the original fraction. The reciprocal of an integer (which is sitting on a phantom 1) is "one over the original integer".

Example 1.8

Compute each of the following:

1. $\frac{5}{6} \div -\frac{3}{4}$

We don't need a common denominator or anything, so we can just jump right in with fraction division. We don't even need to move the negative sign, since we know the answer will be negative.

$$\frac{5}{6} \div -\frac{3}{4} = \frac{5}{6} \cdot -\frac{4}{3} = -\frac{20}{18} = -\frac{10}{9}$$

2. $3\frac{3}{4} \div 5$

First, we'll convert to improper fractions, then we'll carry out fraction division. At the end, we can simplify before we multiply!

$$3\frac{3}{4} \div 5 = \frac{15}{4} \div \frac{5}{1} = \frac{15}{4} \cdot \frac{1}{5} = \frac{3 \cdot 5}{4} \cdot \frac{1}{5} = \frac{3 \cdot \cancel{5}}{4} \cdot \frac{1}{\cancel{5}} = \frac{3}{4}$$

Let's think about this last answer for a second: does it make sense? Look again at the problem we were given. If we interpret this as the question "how many groups of 5 are in $3\frac{3}{4}$?", then we can see that we can't even make one whole group: $3\frac{3}{4}$ is less than 5! So, it makes sense for our answer to be less than 1.

Fractions inside fractions

As if fractions on their own weren't enough, you may soon face the **turducken**¹⁶ of arithmetic, the dreaded "fraction in a fraction". Note that a fraction-in-a-fraction would not be considered simplified since the criteria state that the numerator and denominator must be *integers*.

These creatures can look hard to handle, but don't be intimidated. Recall that a fraction is just a division problem. So, we can rewrite things using a division sign and divide as usual.¹⁷

¹⁶ A "turducken" is a food product where a deboned chicken is stuffed inside a deboned duck, which is then stuffed inside a deboned turkey. In culinary terminology, this is an example of "engastration", a cooking method in which one animal is stuffed inside the gastric cavity of another. . . which is probably yet another phenomenon that you didn't know had a name.

¹⁷ The fancy mathematical name for the \div symbol is **obelus** (plural: obeli). . . another thing that you probably didn't know had a name.

Example 1.9

Simplify each of the following:

1. $\frac{4}{(\frac{1}{3})}$

All we have to do is rewrite the fraction-in-a-fraction as “numerator \div denominator”, and then divide as usual.

$$\frac{4}{(\frac{1}{3})} = 4 \div \frac{1}{3} = \frac{4}{1} \div \frac{1}{3} = \frac{4}{1} \cdot \frac{3}{1} = 4 \cdot 3 = 12$$

2. $\frac{(\frac{2}{5})}{(1\frac{3}{8})}$

Here we rewrite using the obelus, then convert to improper fractions, then divide!

$$\frac{(\frac{2}{5})}{(1\frac{3}{8})} = \frac{2}{5} \div 1\frac{3}{8} = \frac{2}{5} \div \frac{11}{8} = \frac{2}{5} \cdot \frac{8}{11} = \frac{16}{55}$$

In the startup exploration for this section, we have rational numbers a and b where $0 < a < b < 1$. Part (d) asks us to consider $a \div b$. This division problem is another way of writing $\frac{a}{b}$. Since a is less than b , we know that this fraction is a proper fraction (as opposed to an improper fraction), which means it is less than 1. So, $a \div b$ is less than 1.

1.3.5 Evil and Wrong

Everyone makes mistakes, but not all mistakes are created equal. Some mistakes are not just wrong, they are **Evil and Wrong**. “Wrong” because they are mistakes, “evil” because they are subtle, and sneaky, and tempting.

What we mean here is that sometimes we feel drawn to perform certain arithmetic or algebraic maneuvers which seem so logical, so easy, so natural, so tempting... but which in reality are a total trap. For example:

Warning!

Armed with the idea of “simplify before you multiply”, we might want to try and pull this stunt in other places:

$$\frac{2+3}{2+7} = \frac{\cancel{2}+3}{\cancel{2}+7} = \frac{3}{7}$$

Seems logical, right? But so wrong! There is no “simplify before you add” maneuver, tempting though it may be. Such a thing is **Evil and Wrong**.

You may be thinking, “Nah, I’d never do that,” and with numerical expressions like these, you may be right. But later, when faced with variable expressions like

$$\frac{x+3}{x+7},$$

this temptation may come back, in disguise. We’ll draw attention to these **Evil and Wrong** mistakes as we go along because they are so tempting. Beware!

1.4 Order of operations

We turn now to something else that is probably familiar, the **order of operations**. Consider the following problem as we get started:

Who's right?

// startup exploration

The four Krumbli kids each performed the following computation:

$$14 - 6 \cdot 7 + 10$$

Sini got -38 , Siri got -18 , Sten got 66 , and Stig got 136 . Which of them performed the computation correctly? What mistakes did the others make?

The order of operations gives us a standard procedure for simplifying numeric expressions. A **numeric expression** is the algebra way of describing something you may have called just a “math problem” in elementary school. For example, $12 - 5$ is a numeric expression. It is not in its simplest form because we can evaluate $12 - 5$ and write 7 instead.

Simplification rule #1

A numeric expression is completely simplified if all operations have been evaluated and all grouping symbols have been eliminated. The resulting quantity is called the *value* of the expression.

As we go along in algebra, we will learn many rules that maintain “mathematical equivalence”. Expressions are mathematically equivalent if they represent the same quantity. For example, $\frac{4}{8}$ is equivalent to $\frac{1}{2}$, and $12 - 5$ is equivalent to 7 . Our job when we simplify an expression is to ensure that we maintain equivalence from one step to the next. The order of operations is a set of rules for how to do that.

Order of operations

An agreed-upon order for simplifying numeric expressions. The big idea is that “more powerful” operations take priority over “less powerful” operations. When we want to alter the usual rules of precedence, we introduce grouping symbols to make our intentions clear.

When simplifying an expression, we evaluate things in the following order:^a

First: Grouping symbols like (parentheses) and [square brackets], as well as other subtle grouping symbols like absolute value and the vinculum. Here we work from the innermost set of groupers^b to the outermost.

Then: Exponents (later, we will include roots and logarithms at this level of the hierarchy). In the case of a “stack” of exponents, work from the top down.

Then: Multiplication and division. Recall that dividing is the same as “multiplying by the reciprocal”. So, these two operations have the same priority and we work from left to right.

Finally: Addition and subtraction. Recall that “subtracting” is the same as “adding the opposite”. So, these two operations have the same priority, and we work from left to right.

^a PEMDAS is an mnemonic acronym for remembering the order of operations: Parentheses, Exponents, Multiplication, Division, Addition, Subtraction. In Canada it's BEDMAS (B for Brackets). In the UK and Australia it's BIDMAS or BODMAS (for Indices and Orders, which are other words for exponent). A better acronym might be GEMS or PEMA, which group together the pairs of operations that have the same priority (the G is for Grouping symbols).

^b As in “grouping symbols”, not the fish.

Example 1.10

Simplify: $14 - 6 \cdot 7 + 10$

Solution: The correct answer follows the order of operations, like so:

$$\begin{aligned} 14 - 6 \cdot 7 + 10 &= 14 - 42 + 10 && \text{work mult'n before addition or subtraction} \\ &= -28 + 10 && \text{work add'n and sub'n left to right} \\ &= -18 \end{aligned}$$

This is the problem from the startup exploration, so it is Siri who had the right answer.

Students who just memorize a clever mnemonic device might be tempted to do the Addition before the Subtraction, but don't be fooled! Those two operations have the same priority. (Sini made this mistake and got -38 .)

Sten worked the operations straight through from left to right, all at the same priority and got 66. Stig did something very, ahem, “creative”: $(14 - 6) \cdot (7 + 10) = 8 \cdot 17 = 136$.

Notice how we showed the work going down the page, simplifying the problem one step at a time. Some people

prefer to work across the page, and that's okay too. The point is that it's important to show work in an organized way when we solve a problem so that our reasoning and thought process are clear.

The work we show is the road map to our solution. Whoever reads over our work must be able to follow and understand our steps, without having to make any assumptions about what we actually did to reach our answer. It's a bad habit to skip steps, and it's no help to people reading our work to claim that we did a step in our head. Every step we take needs to be written down clearly and neatly.

Abuse of the equal sign

Consider the following work, written out by a student to simplify $8 \cdot 4 + 10$

$$8 \cdot 4 = 32 + 10 = 42 \quad \text{OK or not OK?}$$

The student has reached the correct value, and it may even be clear what the student is thinking: "8 times 4 is 32, plus 10 more makes 42". This, however, is a heinous abuse of the equal sign! Look at what the first part of the work says:

$$8 \cdot 4 = 32 + 10 \quad \text{These are not equal!}$$

One way to avoid this misuse of the equal sign is to write your work going down the page (as shown in example 1.10). If you prefer to write across the page, be sure to write out the whole problem as you perform each simplification:

$$8 \cdot 4 + 10 = 32 + 10 = 42$$

In the next few sections, we'll look more closely at some trickier aspects of the order of operations.

1.4.1 About grouping symbols

As expressions get complicated, we may have grouping symbols inside of other grouping symbols. If we have different symbols (like parentheses with [some bracketed expression] inside), we can more easily see where different groups begin and end. But, we may just find a bunch of parentheses, like in the example below. In that case, we have to be a bit careful about what's being grouped together.

Example 1.11

Simplify: $12 + (3 - (4 - 2) + 5)$

Solution: When faced with multiple grouping symbols, we must start with the innermost set of grouping symbols and evaluate our way to the outermost. Once we have simplified the expression inside a set of grouping symbols down to a single quantity, we can write that quantity without the groupers.

$$\begin{aligned} 12 + (3 - (4 - 2) + 5) &= 12 + (3 - 2 + 5) \\ &= 12 + (1 + 5) \\ &= 12 + 6 \\ &= 18 \end{aligned}$$

The vinculum is a grouping symbol

The **vinculum**, as in the fraction bar, is a grouping symbol. For example, if the task is to simplify a fraction such as:

$$\frac{20 + 2^2 \cdot (14 - 9)}{(2 - 4)^3},$$

then we must think of the expression in the numerator as a group, and likewise for the denominator. In other words, like so:

$$(20 + 2^2 \cdot (14 - 9)) \div ((2 - 4)^3).$$

We have two options for simplifying this. We might keep it in a fraction the whole time, or we could simplify the numerator and denominator separately and then squish them back into a fraction at the end.

Example 1.12

Simplify: $\frac{20 + 2^2 \cdot (14 - 9)}{(2 - 4)^3}$

Solution: Let's dismantle this thing and handle it in two pieces. The numerator works like this:

$$\begin{aligned} 20 + 2^2 \cdot (14 - 9) &= 20 + 2^2 \cdot 5 \\ &= 20 + 4 \cdot 5 \\ &= 20 + 20 \\ &= 40 \end{aligned}$$

The denominator works like this: $(2 - 4)^3 = (-2)^3 = -8$. Then, we can put the pieces back into their original fraction configuration:

$$\frac{20 + 2^2 \cdot (14 - 9)}{(2 - 4)^3} = \frac{40}{-8} = -5$$

1.4.2 About exponents

An expression of the form a^b is read “ a to the power of b ” or “ a to the b^{th} power”. In such an expression, a is called the **base** and b is called the **exponent**. We call the whole thing a **power** of a , since a is the base.

We will get into more detail about exponents later on, but we’ll pause here to mention two key ideas. First, recall that we can think about an exponent as shorthand for a repeated multiplication.¹⁸

$$a^b = \underbrace{a \cdot a \cdot a \cdots a}_{b \text{ times}}$$

That part you probably knew already. This next fact may be new:

Raising to the power zero

For any nonzero number a , the expression $a^0 = 1$. In other words, any nonzero number raised to the power 0 equals 1.

When we say here that a is a “nonzero number” this means, naturally enough, that a cannot be 0. The expression $0^0 \neq 1$. What *does* it equal? That’s a tricky question that will have to wait for later. 0^0 is an unusual mathematical creature!¹⁹

We’ll get into the “hows and whys” of exponents in chapters 10 and 11, and we’ll return to the idea of the zero exponent. In the meantime, here’s an example of how this fact might come in handy.

¹⁸ As with “multiplication is repeated addition”, this interpretation breaks down eventually. Expressions like 5^{-3} and $16^{1/2}$ don’t really translate well into “repeated multiplication”. Don’t panic about the idea of a negative number or a fraction up there in the exponent! All will be revealed in time.

¹⁹ It’s not the only one, either. In 1872, Karl Weierstrass (or Weierstraß, if you prefer the German double-S) shook the foundations of calculus with his mathematical monster, now called the “Weierstraß function”. It’s a bit complicated to get into the details, but it made some other mathematicians very upset. Drama!

Example 1.13

Simplify: $\left(\frac{120 - (24 - 5^2)}{7^2 \cdot 400 \div 6^3}\right)^0$

Solution: If we go on “auto-pilot” we might follow all of the simplification rules, work from the inside to the outside, simplify the numerator, simplify the denominator. . .

But, the expression is raised to the power 0. So the answer is probably 1! We must check that we don’t have 0^0 , but we can use a little number sense to do a quick check of the numerator in the fraction, and see that it will not equal zero. (Can you see why without having to work it all out?)

We must also check that the denominator is not zero. A quick check there shows that it is not zero either. (Can you see why?) So, this one’s easy:

$$\left(\frac{120 - (24 - 5^2)}{7^2 \cdot 400 \div 6^3}\right)^0 = 1$$

The lesson here is to look at the entire problem and plan an efficient solution strategy *before* just jumping in and crunching numbers. In this case, we can save ourselves a lot of work with a bit of careful observation.

Fractions versus exponents

If we have a fraction raised to a power, we must be very mindful of the notation and to what, exactly, the exponent applies.

Example 1.14

Simplify each of the following:

1. $\left(\frac{2}{3}\right)^4$

The parentheses indicate that we are multiplying together 4 copies of the fraction:

$$\left(\frac{2}{3}\right)^4 = \left(\frac{2}{3}\right)\left(\frac{2}{3}\right)\left(\frac{2}{3}\right)\left(\frac{2}{3}\right) = \frac{16}{81}$$

2. $\frac{2^4}{3}$

Remember that the numerator is a group, and so the exponent applies only to the 2, not the whole fraction:

$$\frac{2^4}{3} = \frac{2 \cdot 2 \cdot 2 \cdot 2}{3} = \frac{16}{3}$$

Keep this last example in mind when writing your own expressions, as well. Our notation must match our intentions. What exactly do we intend the exponent to apply to? Do we need to add parentheses to make our intentions clear?

One of the trickiest concepts in Algebra 1

What is the difference between the following three expressions?

$$(-3)^2 \quad - (3^2) \quad - 3^2$$

Tricky, right? This is an important difference that will come back over and over again, and it will look a little different every time.

In the first case, the parentheses make it clear: $(-3)^2$ means “raise negative three to the second power”.

$$(-3)^2 = -3 \cdot -3 = 9$$

As usual, the product of two negatives is positive (Hitler gets hit by a truck). The second case is clear as well. The parentheses indicate that we should simplify the exponent first and then take the opposite of the result:

$$-(3^2) = -(3 \cdot 3) = -(9) = -9$$

The third case, -3^2 , is the tricky one. It looks kind of ambiguous, since there are no parentheses. But, we can interpret -3 as the product $-1 \cdot 3$. In this case, we have

$$-3^2 = -1 \cdot 3^2 = -1 \cdot 9 = -9$$

Now, it's clear what to do even without the parentheses, and that is the key.

Opposites of numbers to a power

The expression $-a^n$, written without parentheses, is equivalent to the parenthesized expression $-(a^n)$. For example: $-3^2 = -(3^2) = -9$.

You might be saying to yourself, “No sweat. I get it.” But (if history and human nature are any indication) you may find yourself tripping over this concept at some point. Confusion around the notation can pop up, for instance, when we’re typing expressions into a graphing calculator.²⁰

1.4.3 About multiplication

In algebra we frequently use x (the letter) to stand for a unknown or variable quantity. So, we never use \times to show multiplication. It would be too confusing to have a mix of letter- x ’s and multiplication- \times ’s in the same expression. Can you imagine trying to decode something like:

$$x \times 3 + 4 \times x = x + 4x \times x \times x \quad \text{Yikes!}$$

No more \times for multiplication!

To show the operation of multiplication, use an asterisk, a dot, or parentheses. Instead of writing 3×4 to represent “3 times 4”, we write: $3 * 4$, or $3 \cdot 4$, or $3(4)$.

Implied operations

A key aspect of algebra will be learning to read the notation and use it correctly when writing expressions and equations. Algebra is very much like a language, and all languages have special rules. For example, in English we sometimes smash up two words as a contraction: instead of writing “do not” we can write “don’t”.

We used contractions (of a kind) in mathematics as well. We don’t usually write a positive sign in front of positive numbers. We don’t usually write the phantom 1 that is in the denominator of an integer.

Another kind of mathematical contraction is the use of an **implied operation**. This arises most often when dealing with multiplication. Here’s an example of how implied operations can sneak into a problem.

²⁰ A note to old-school parents who are used to working with calculators that use reverse Polish notation (RPN) or a stack, and might want to argue that $-3^2 = 9$. When using an RPN calculator we punch in -3 and press enter to push that quantity onto the stack. The number and the negative are both in the stack together, so squaring the entry on the top of the stack means squaring negative three, which is equivalent to $(-3)^2$ and not the same as -3^2 .

Example 1.15

Simplify: $12 - 5(2 + 8)$

Solution: Be on the lookout for implied operations!

$$\begin{aligned} 12 - 5(2 + 8) &= 12 - 5(10) && \text{add in the parentheses first} \\ &= 12 - 50 && \text{that } 5(10) \text{ means mult'n, and must happen first!} \\ &= -38 \end{aligned}$$

A very common mistake in the second step is to do the $12 - 5$ first, instead of the $5(10)$. But that would violate the order of operations!

» Chapter summary «

In this chapter we have thought deeply about fundamental ideas regarding numbers and operations. This exploration has perhaps presented some familiar ideas in a new light. Prepare to see more of this! One of the fun and surprising things about mathematics is that we can often learn something new from looking at a familiar concept in a different way.

Here, we looked closely at two important sets of numbers – the integers and the rational numbers – and at our first example of simplification using the order of operations. Armed with this knowledge, we now venture deeper into the algebraic wilderness. Onward!

A mathematician, like a painter or a poet, is a maker of patterns.

G. H. Hardy, British mathematician

Chapter 2

Sequences

In chapter 1 we reviewed some key ideas from arithmetic. Arithmetic involves the manipulation of numbers via certain operations (like addition and multiplication). In algebra, on the other hand, we often use symbols to replace numbers. This can be disorienting at first, but it is useful because it allows us to speak about relationships and patterns involving numbers, rather than specific numbers.

Given a few numbers that form a pattern, for example, we can use algebraic symbols to describe all of the numbers that fit the given pattern – even though there may be *infinitely many* numbers that fit the pattern! Since patterns lie at the heart of algebra, they are the focus of this chapter.

2.1 Sequences and recursion

Communicating a pattern

// startup exploration

As children, the twins Bob and Yeardleigh Krumbli (great-great-great-grandchildren of Knut and Jorunn) liked to play a kind of guessing game. One would think of a number pattern and try to communicate it to the other using as little information as possible.

Predict the next few numbers in the number pattern shown below.

2, 5, 8, 11, 14, 17, ...

How would you describe this pattern to a partner who could not see it? Could you communicate the pattern *without simply listing the numbers*? What's the minimum amount of information you could give so that your partner could recreate the pattern?

Informally, we call this a number pattern. Mathematically speaking, an ordered list of numbers like this is called a **sequence**. Each of the numbers in the list is called a **term** of the sequence.

Sequences often have patterns within them. Perhaps when thinking about how you'd describe this sequence to a partner, you thought of a rule based on "adding 3". (Do you see how this applies to the given pattern?)

But "add 3" is not enough to recreate the sequence. Consider the sequence: 1, 4, 7, 10, 13, And what about $-10, -7, -4, -1, 2, \dots$? The phrase "add 3" also applies to these sequences, even though they are different from the sequence in the startup exploration.

To distinguish these different sequences, we must include the starting value in our description. We can describe the original sequence clearly and unambiguously by saying something like: "Start with 2, then add 3 to each resulting value."

When describing the pattern of a sequence, we are really describing how to generate the sequence from scratch. To do that, we have to answer these two questions: First, where does the sequence begin? Second, what must we do to the *current* term to find the *next* term of the sequence? This is called the **recursive** description of the sequence.

Recursive

Describes a procedure that is applied over and over again, starting with a number or a geometric figure, to produce a sequence of numbers or figures.

As the definition says, we can start a recursive procedure with a number or a geometric figure. In what comes next, we'll explore sequences by studying geometric figures called **fractals**.

Fractal

A geometric figure that has undergone infinitely many applications of a recursive procedure, and which exhibits the property of self-similarity.

Fractal geometry is often called "the geometry of nature". If we look around the natural world, it is not like we see a lot of perfectly straight lines, rigid rectangles, and regular pentagons. But, the growth of a tree can be described using a recursive procedure: grow towards the sun for a bit, branch off at an angle, repeat. Trees exhibit self-similarity: If we break off a branch from a tree and stick it in the ground, looks just like a little tree!

Clouds, coastlines, mountains, trees, Romanesco broccoli, the folds of your brain, your vascular system, your bronchial tubes, the lining of your small intestines. . . all of these are a kind of fractal.¹ Technically speaking,

¹ In 1968, Hungarian biologist Aristid Lindenmayer developed a method for writing recursive rules that could be used to model the growth of algae. Called "Lindenmayer systems" or "L-systems" today, these methods have been used to model more complex organisms, as well as purely mathematical structures.

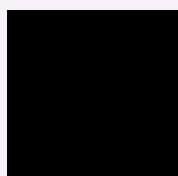
natural fractals only have their recursive procedure applied a limited number of times (we say the procedure has a limited number of *iterations*) so they aren't true mathematical fractals. A mathematical fractal undergoes an infinite number of iterations.²

Our first fractal will be a famous fractal that was first studied by Polish mathematician Wacław Sierpiński.

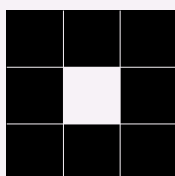
Sierpiński's carpet

// startup exploration

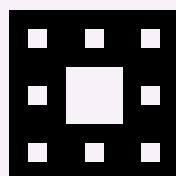
To draw Sierpiński's carpet, we begin with a square called "stage 0". We subdivide this square into nine congruent sub-squares and remove the one in the center. We repeat the process with each of the remaining sub-squares. Stages 0 through 3 of the fractal are shown below.



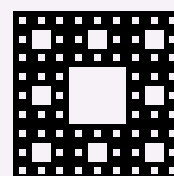
Stage 0



Stage 1



Stage 2



Stage 3

There is one solid square in stage 0, and there are eight (smaller) solid squares in stage 1. How many of the smallest solid squares are there in stage 2? What about stage 3?

2.1.1 Algebra of Sierpiński's carpet

Sierpiński's carpet generates some interesting sequences of numbers. For example, if we consider the number of (smallest) squares at each stage of the fractal. We have one square in stage 0 and eight squares in stage 1.

To create stage 2, we divide each of the eight stage-one squares into 9 pieces and then remove the center square. So each of the eight squares from stage 1 turns into eight new, tiny squares in stage 2. So there are $8 \cdot 8 = 64$ tiny squares in stage 2. To create stage 3, each of these sixty-four tiny squares becomes 8 super-tiny squares, so there are $64 \cdot 8 = 512$ super-tiny squares in stage 3. Together, we have a sequence that begins:

$$1, 8, 64, 512, \dots$$

What is a recursive rule for this sequence? The sequence starts with 1, and then to go from one number to the next, we multiply by 8. So, that's our rule: "start with 1, multiply the previous value by 8". In other words, to find the next term in the sequence, take the previous term (the last term in the sequence that we know) and multiply by 8.

² Fractals may play an interesting role later on in your study of mathematics, for example the Mandelbrot set is a fractal that involves the complex numbers. Do an internet search for "Mandelbrot set" and check out the pictures!

We can use this rule to generate the next few terms of our sequence, but watch out! We quickly end up with a lot of teeny squares, and you don't want to get any on your shoes.

1, 8, 64, 512, 4096, 32 768, 262 144, 2 097 152, 16 777 216, ...

We can always write a recursive rule as a sentence, as we did above. Another way to capture a recursive procedure is using a “now-next” rule, sometimes called a “start-now-next” rule. For example, the now-next rule for the number of squares is: “START = 1, NEXT = NOW · 8”. It's pretty obvious that the first part of the rule says where to start. The second part of the rule says: “To find the *next* number in the pattern, we take the number we have *now* and multiply by 8.”

2.1.2 Recursive rules and formulas

Recursive rules are easy to write in sentence form, and now-next equations are nice and succinct, but there is a more mathematical way. We are going to write what we call a *recursive formula*.

We usually use the letter a with a subscript to represent a specific term of the sequence. So, a_1 represents the first term of the sequence, a_2 represents the second term of the sequence, and a_{98} would represent the 98th term of the sequence.

For example: given the sequence 4, 12, 36, 108, ..., we have:

$$a_1 = 4$$

$$a_2 = 12$$

$$a_3 = 36$$

$$a_4 = 108$$

We can write the recursive rule either as “start with 4, multiply the previous term by 3”, or “START = 4, NEXT = NOW · 3”. Here's how we can translate this into a recursive formula.

“Start with 4” means that the first term of the sequence is 4. We write

$$a_1 = 4,$$

since a_1 represents the first term of the sequence. This is just like “START = 4” in the now-next rule.

To translate “NEXT = NOW · 3”, we use a_n to represent any old term of the sequence. Given that, we write a_{n+1} to represent the *next* term in the sequence. (Can you explain why?) So, we have the recursive step:

$$a_{n+1} = a_n \cdot 3.$$

A note about notation: When multiplying a number and a letter, we usually write the number first and we don't usually write a multiplication symbol in between.³ So, we have created the recursive formula:

$$\begin{aligned}a_1 &= 4 \\ a_{n+1} &= 3a_n\end{aligned}$$

Example 2.1

Write the recursive formula for the sequence 1, 5, 25, 125, 625, . . .

Solution: With a little exploration, we see that the sentence version of this rule is “Start with 1, multiply previous by 5”, and the now-next version is “START = 1, NEXT = NOW · 5”. So, we have the recursive formula: $a_1 = 1$, $a_{n+1} = 5a_n$.

Example 2.2

Write out the first five terms of the sequence generated by each rule.

1. “Start with 128, multiply previous by $\frac{1}{2}$ ”

Solution: The rule states clearly that the first term is 128, no trouble. Then, to find the second term, we multiply the first term by $\frac{1}{2}$, that means $128 \cdot \frac{1}{2} = 64$. To find the third term, we multiply the second term by one-half: $64 \cdot \frac{1}{2} = 32$. We repeat for the next few terms, which gives:

$$128, 64, 32, 16, 8, \dots$$

2. $a_1 = 12$, $a_{n+1} = -2 \cdot a_n$

Solution: The first term is a_1 , and the formula says that's 12. Then, to find a_2 , the second term, we have

$$a_2 = -2 \cdot a_1 = -2 \cdot 12 = -24.$$

We continue to multiply by -2 each step of the way and get:

$$12, -24, 48, -96, 192, \dots$$

³ More on working with letters, or variables, in chapter 3.

Recursive rules and formulas are handy for describing a sequence, but suppose we want to skip around and find random terms of the sequence. In this situation, the recursive rule is the worst possible rule to have!

For example, how could we find the value of a_{1000} , the 1000th term in the sequence, given the rule $a_1 = 4$, $a_{n+1} = 3 \cdot a_n$? The rule tells us that $a_{1000} = 3 \cdot a_{999}$. But, what's a_{999} ?

Well, $a_{999} = 3 \cdot a_{998}$. But, what's a_{998} ?

Hmm. $a_{998} = 3 \cdot a_{997} \dots$ but... oh boy. Can you see the problem here?

If we want to skip around and find random terms in a sequence, it's much easier to use a different kind of formula called an *apparent* or *explicit formula*. More on those in the next section!

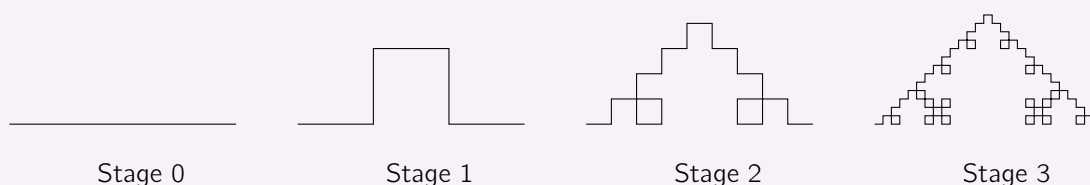
2.2 Geometric sequences

We are going to continue our study of sequences by looking at another fractal. In 1904, Swedish mathematician Helge von Koch first described several variants of a fractal that has since come to be known as a Koch curve.

Koch curve, square version

// startup exploration

We begin in stage 0 with a line segment of length 1. To create stage 1, we alter the segment as follows: cut it into three pieces, and replace the center piece with three sides of a square. We repeat the process for each line segment in the previous figure to create stages 2 and 3.



Write a recursive formula describing the number of segments in each stage of the fractal.

2.2.1 Algebra of the Koch curve

The Koch curves are beautiful things, at once incredibly simple and incredibly complex. As the square-based version above grows, each line segment is replaced by five shorter segments. The recursive rule is “start with 1, multiply the previous value by 5”.

To compute the number of segments in each stage, we might organize our work in a list like this:

1	segment in stage 0
$(1 \cdot 5)$	segments in stage 1
$(1 \cdot 5) \cdot 5$	segments in stage 2
$(1 \cdot 5 \cdot 5) \cdot 5$	segments in stage 3
$(1 \cdot 5 \cdot 5 \cdot 5) \cdot 5$	segments in stage 4
$(1 \cdot 5 \cdot 5 \cdot 5 \cdot 5) \cdot 5$	segments in stage 5

We can use a bit of shorthand, and write this repeated multiplication using an exponent.

$$\begin{aligned}
 1 &= 5^0 && \text{segment in stage 0} \\
 1 \cdot 5 &= 5^1 && \text{segments in stage 1} \\
 1 \cdot 5 \cdot 5 &= 5^2 && \text{segments in stage 2} \\
 1 \cdot 5 \cdot 5 \cdot 5 &= 5^3 && \text{segments in stage 3} \\
 1 \cdot 5 \cdot 5 \cdot 5 \cdot 5 &= 5^4 && \text{segments in stage 4} \\
 1 \cdot 5 \cdot 5 \cdot 5 \cdot 5 \cdot 5 &= 5^5 && \text{segments in stage 5}
 \end{aligned}$$

Notice that the exponent is equal to the stage number. This “5 to a power” notation works even for stage 0, since $5^0 = 1$.

So, if we want to know how many segments are in stage 8 of the fractal, we can use this pattern to predict that there will be $1 \cdot 5^8$ segments. If we let x represent the stage number, then stage x of the fractal will have 5^x segments.

We have discovered a way of calculating the number of segments that is *not* recursive, because it doesn’t rely on our knowing any of the previous terms! Instead, to produce the value of a certain term, all we need is the *number of the term*. We can compute the number of line segments in stage x without having to know anything about the stages that came before it.

2.2.2 Explicit formulas for sequences

In our discussion of fractals, we have always described the first image as “stage 0” of the fractal. But, when we write out a sequence, the first term is, well, the *first* term (not the *zeroth* term).⁴ In other words, the same pattern of values may have a slightly different numbering, depending on whether we’re describing stages of a fractal or terms in a sequence.

Value	1	5	25	125
Fractal Stage Number	0	1	2	3
Sequence Term Number	1	2	3	4

So, if we want to write a recursive formula for the terms of a sequence, we have to make a little adjustment:

$$\begin{aligned}
 a_1 &= 1 && = 5^0 \\
 a_2 &= 5 && = 5^1 \\
 a_3 &= 25 && = 5^2 \\
 a_4 &= 125 && = 5^3
 \end{aligned}$$

⁴ In some scientific disciplines, it is customary to start counting with zero: for example, in computer science. Jason, one of the authors of the *Algebranomicon*, is a computer scientist by training and thinks this way. Jason also prefers to include 0 as one of the natural numbers. Patty, the other author of the *Algebranomicon*, is a mathematician by training and prefers to start counting at 1.

Can you see the relationship between the subscript and the exponent? If we let a_n represents any term of the sequence, then our rule is:

$$a_n = 5^{n-1}$$

Rules of this kind are called apparent formulas or explicit formulas. One benefit of rules like this is that if we want to know, say, the number of segments in the curve at stage 1904, we can compute simply:

$$a_{1904} = 5^{1903}$$

By the way, this number is enormous. It's more than 1300 digits long!

Example 2.3

Write explicit formulas for these other features of the Koch curve: the length of one segment and the total length of the curve.

Solution: *Length of one segment.* Each segment in a certain stage is one-third the length of the segment in the stage before. So, the sequence generated by the length of one segment in each stage is

$$\left(\frac{1}{3}\right)^0, \left(\frac{1}{3}\right)^1, \left(\frac{1}{3}\right)^2, \dots \quad \text{and so we have} \quad a_n = \left(\frac{1}{3}\right)^{n-1}.$$

Total length of the curve. Since we know the number of segments and the length of each segment, we can multiply to find the total length of the curve. We have

$$\left(\frac{5}{3}\right)^0, \left(\frac{5}{3}\right)^1, \left(\frac{5}{3}\right)^2, \dots \quad \text{and so we have} \quad a_n = \left(\frac{5}{3}\right)^{n-1}.$$

Note that we're putting these fractions in parentheses! Our notation has to match our intentions and in this case we want to show that the *whole fraction* is being raised to a given power.

Example 2.4

What if the stage 0 figure had been a segment of length 7, rather than length 1? How would that change our formula?

Solution: The number of segments would not change, but the length of each segment (and the total length of the curve) would! The new sequence for the length of one segment would be generated as

follows:

$$\begin{aligned}
 7 &= 7 \left(\frac{1}{3}\right)^0 && \text{length of one segment in stage 0} \\
 7 \left(\frac{1}{3}\right) &= 7 \left(\frac{1}{3}\right)^1 && \text{length of one segment in stage 1} \\
 7 \left(\frac{1}{3}\right) \left(\frac{1}{3}\right) &= 7 \left(\frac{1}{3}\right)^2 && \text{length of one segment in stage 2} \\
 7 \left(\frac{1}{3}\right) \left(\frac{1}{3}\right) \left(\frac{1}{3}\right) &= 7 \left(\frac{1}{3}\right)^3 && \text{length of one segment in stage 3}
 \end{aligned}$$

Again we can use an exponent to simplify the repeated multiplication of $\frac{1}{3}$. This generates the sequence

$$7 \left(\frac{1}{3}\right)^0, 7 \left(\frac{1}{3}\right)^1, 7 \left(\frac{1}{3}\right)^2, \dots$$

If we let n represent the term number, then the recursive formula for this sequence is

$$a_n = 7 \left(\frac{1}{3}\right)^{n-1}.$$

2.2.3 Geometric sequences

So far, all of our sequences have had recursive rules like “start with A , multiply the previous term by B ”. Sequences with recursive rules of this type are called **geometric sequences**. Geometric sequences belong to the family of *exponential relationships*, because the apparent formula has a variable in the exponent.

To generate the next term of a geometric sequence, we multiply the previous term by a fixed value. This fixed value is sometimes called, naturally enough, the *constant multiplier*. More often, it is called the **common ratio**.

Geometric sequence

A sequence in which the ratio between each pair of successive terms is constant. The constant ratio is often called the *common ratio*. Geometric sequences are exponential relationships.

Example 2.5

Determine whether or not the sequence 4, 12, 36, 108, ... is a geometric sequence.

Solution: If this is a geometric sequence, then it must have a rule of the form “start with A , multiply the previous term by B ”? Let’s check.

To go from 4 to 12, we multiply by $\frac{12}{4} = 3$.

To go from 12 to 36, we multiply by $\frac{36}{12} = 3$. Looking good so far!

To go from 36 to 108, we multiply by $\frac{108}{36} = 3$. Nice! Based on the four terms given, the sequence is geometric.

Now, look at what we did to determine this: we created ratios of successive terms, and found that they were all the same.

$$\frac{12}{4} = \frac{36}{12} = \frac{108}{36} = 3$$

So, the *common ratio* for this sequence is 3.

Example 2.6

Write recursive and explicit formulas for the geometric sequence 32, 24, 18, $13\frac{1}{2}$, ...

Solution: To get from 32 to 24, our first instinct might be to subtract: $32 - 8 = 24$. But, we’re told in the problem that this is a *geometric* sequence, and that means that the recursive rule involves multiplication, not subtraction.

How can we get from 32 to 24 using multiplication? The constant multiplier must be less than one. (Can you explain why?) We can divide to find what it is:

$$\frac{24}{32} = \frac{3}{4}$$

So, $\frac{3}{4}$ is a good candidate for the constant ratio of the sequence. Let’s check the other terms to see if we’re right. We multiply 24 by $\frac{3}{4}$ to see if that gives us the next term in the sequence:

$$24 \cdot \frac{3}{4} = \frac{24}{1} \cdot \frac{3}{4} = \frac{\cancel{4} \cdot 6}{1} \cdot \frac{3}{\cancel{4}} = \frac{6}{1} \cdot \frac{3}{1} = 18 \quad \text{Check!}$$

Now see if 18 times $\frac{3}{4}$ gives the next term:

$$18 \cdot \frac{3}{4} = \frac{18}{1} \cdot \frac{3}{4} = \frac{\cancel{2} \cdot 9}{1} \cdot \frac{3}{\cancel{2} \cdot 2} = \frac{9}{1} \cdot \frac{3}{2} = \frac{27}{2} = 13\frac{1}{2} \quad \text{Check!}$$

So, we have found the correct constant multiplier based on the information we were given. The recursive formula is

$$a_1 = 32, \quad a_{n+1} = \frac{3}{4} \cdot a_n,$$

and the explicit formula is

$$a_n = 32 \cdot \left(\frac{3}{4}\right)^{n-1}.$$

If we look back over the explicit rules for the sequences in this section, we might notice that the formulas have a formula of their own! In other words, the apparent rule for a geometric sequence always has a certain structure, which we summarize here.

Apparent formula for a geometric sequence

Given a geometric sequence with first term a_1 and common ratio r , in other words, a sequence of the form

$$a_1, \quad a_1 r, \quad a_1 r^2, \quad a_1 r^3, \quad \dots$$

The apparent or explicit formula for the sequence is:

$$a_n = a_1 r^{n-1}.$$

2.3 Arithmetic sequences

Not all sequences are geometric sequences, of course. Let's explore some other types of sequences.

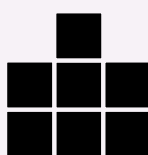
Tile pattern #1

// startup exploration

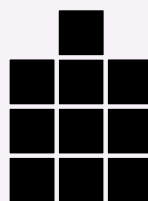
The pictures below represent stages 1, 2, 3, and 4 for a pattern of square tiles.



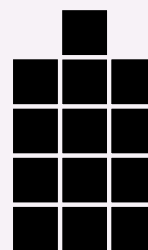
Stage 1



Stage 2



Stage 3



Stage 4

Draw pictures representing stages 5 and 6 in the pattern. Write a sentence or two to describe the pattern in the pictures. What would the stage 0 figure look like?

Write out the sequence for the number of tiles at each stage (starting with stage 1). Write a recursive rule to describe your sequence. How is this rule different from the rules in section 2.2?

The number of tiles in each stage of the pattern creates the sequence

$$4, 7, 10, 13, 16, 19, \dots$$

We might write recursive rules that go something like “start with 4, add 3 to the previous term”, or “START = 4, NEXT = NOW + 3”. The fact that we’re adding in the rule is a clear difference from the rules we saw when studying geometric sequences.

Sequences like these are called **arithmetic sequences**.⁵ Instead of having a common ratio, these sequences have a **common difference**. Arithmetic sequences belong to the family of *linear relationships*.

⁵ A word about pronunciation. The branch of mathematics that deals with calculations and operations on numbers is called “arithmetic”. When used as a noun in this way, the word is pronounced with the emphasis on the second syllable: *a·RITH·me·tic*. The sequences we’re talking about in this section are “arithmetic sequences”. When the word is used as an adjective, the emphasis is on the third syllable: *a·rith·ME·tic*.

Arithmetic sequence

A sequence in which the difference between each pair of successive terms is constant. The constant difference is called the *common difference*, usually denoted d . Arithmetic sequences are linear relationships.

Example 2.7

Verify that the given sequence is arithmetic and write a recursive formula for it: 12, 17, 22, 27, ...

Solution: In order for a sequence to be arithmetic, we must add the same quantity as we go from term to term. We can check this by subtracting (which is why the thing we add is called a “common difference”). So, let’s check:

$$17 - 12 = 5$$

$$22 - 17 = 5$$

$$27 - 22 = 5$$

Check! This is an arithmetic sequence with a common difference of 5.

To write the recursive formula we know that the common difference is added to the current term in order to find the next term. We also know the first term. So:

$$a_1 = 12, \quad a_{n+1} = a_n + 5$$

is the recursive formula for the sequence.

2.3.1 Explicit formulas for arithmetic sequences

Of course, we can write an apparent or explicit formula (that is, a non-recursive formula) for an arithmetic sequence. Consider the sequence from the startup exploration: 4, 7, 10, 13, ... We know where each of the terms come from:

$$a_1 = 4$$

$$a_2 = (4) + 3$$

$$a_3 = (4 + 3) + 3$$

$$a_4 = (4 + 3 + 3) + 3$$

$$a_5 = (4 + 3 + 3 + 3) + 3$$

Notice the repeated addition of 3. This is a case where we can reinterpret repeated addition as multiplication:

$$\begin{aligned}
 a_1 &= 4 & = 4 + 3 \cdot 0 \\
 a_2 &= 4 + 3 & = 4 + 3 \cdot 1 \\
 a_3 &= 4 + 3 + 3 & = 4 + 3 \cdot 2 \\
 a_4 &= 4 + 3 + 3 + 3 & = 4 + 3 \cdot 3 \\
 a_5 &= 4 + 3 + 3 + 3 + 3 & = 4 + 3 \cdot 4
 \end{aligned}$$

Notice now that these multiplications are 3 times “one less than the stage number”! Therefore, we can write

$$a_n = 4 + 3(n - 1)$$

As we did with geometric sequences, we can summarize the general form of the explicit rule for an arithmetic sequence:

Apparent formula for an arithmetic sequence

Given an arithmetic sequence with first term a_1 and common difference d , in other words, a sequence of the form

$$a_1, a_1 + d, a_1 + 2d, a_1 + 3d, \dots$$

The apparent or explicit formula for the sequence is:

$$a_n = a_1 + (n - 1) * d.$$

2.3.2 Using stage zero

There is another way to write the apparent rule for an arithmetic sequence. We can use this approach when we know (or can find) the “zeroth” term. Then, we interpret the stage 1 figure not as the *start*, but rather as though we are joining a sequence that is “already in progress”.

For example, in the tile sequence from the startup exploration, to find the stage 0 figure we have to “back up a step”. Since the pattern goes forward by adding 3, to back up one step we must subtract 3. So, the stage 0 figure is just 1 square tile.

Stage	Value	Start with stage 1?		Start with stage 0?	
1	4	4	$= 4 + 3(0)$	$1 + 3$	$= 1 + 3(1)$
2	7	$4 + 3$	$= 4 + 3(1)$	$1 + 3 + 3$	$= 1 + 3(2)$
3	10	$4 + 3 + 3$	$= 4 + 3(2)$	$1 + 3 + 3 + 3$	$= 1 + 3(3)$
4	13	$4 + 3 + 3 + 3$	$= 4 + 3(3)$	$1 + 3 + 3 + 3 + 3$	$= 1 + 3(4)$

One benefit of this new rule is that we find ourselves multiplying the constant difference by the term number itself (as opposed to multiplying by one less than the term number). In other words, we can write the apparent rule as follows:

Apparent formula for an arithmetic sequence (zero version)

Given an arithmetic sequence with first term a_1 and common difference d , we can write the apparent or explicit formula for the sequence is

$$a_n = a_0 + n * d$$

Where a_0 represents the “zeroth” term of the sequence (the term that comes before the first term).

In later chapters, we will explore in more detail the connections between the “stage 1 version” and the “stage 0 version” of the rule for arithmetic sequences, and we will learn techniques for writing “stage 0 versions” of the rules for geometric sequences.

Example 2.8

Write a stage zero version of the explicit rule for the arithmetic sequence: 43, 35, 27, 19, ...

Solution: This sequence is decreasing, so we must be adding a negative number in the rule. In other words, the common difference must be negative. Subtracting neighboring terms, we can find that the common difference is -8 .

To write a zero-based rule, we have to know the zeroth term, and to find that we have to back up from the first term. So, we have $a_0 = 43 - -8 = 43 + 8 = 51$. This value makes sense: Since the sequence is decreasing, the zero term should be larger than the first term.

Knowing the common difference and the zero term, we can write a zero-based explicit rule:

$$\begin{aligned} a_n &= a_0 + n * d \\ &= 51 + n * -8 \\ &= 51 - 8n \end{aligned}$$

2.4 Other types of sequences

Tile pattern #2

// startup exploration

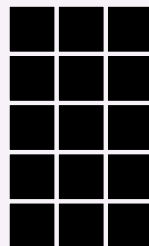
The pictures below represent stages 1, 2, 3, and 4 for a new pattern of square tiles.



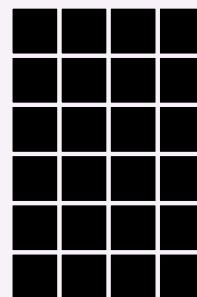
Stage 1



Stage 2



Stage 3



Stage 4

Draw pictures representing stages 5 and 6 in the pattern. Write a sentence or two to describe the pattern in the pictures. What would the Stage 0 figure look like?

Write out the sequence for the number of tiles at each stage (starting with stage 1). Write a recursive rule to describe your sequence. How is this rule different from the rules in the last few sections?

These sequences are a bit harder to work with! The figures in the startup exploration generate the sequence:

3, 8, 15, 24, 35, 48, ...

Is this sequence geometric? Let's check for a common ratio: the ratio between the first two terms is $\frac{8}{3}$, and the ratio between the next two terms is $\frac{15}{8}$. Those are different ratios, since if we write them with a common denominator, we have $\frac{8}{3} = \frac{64}{24}$ and $\frac{15}{8} = \frac{45}{24}$. So, the sequence is *not geometric*.

Is the sequence arithmetic? Let's check for a common difference:

$$8 - 3 = 5$$

$$15 - 8 = 7$$

$$24 - 15 = 9$$

$$35 - 24 = 11$$

$$48 - 35 = 13$$

The sequence does not have a common difference, so it is *not arithmetic*. But take a look at those differences! The differences have a pattern of their own: They go up by 2 every time. In other words, the differences form

an arithmetic sequence! It's a sequence within a sequence! The turducken of sequences!⁶

To describe this sequence with a recursive rule, we'll need to give the starting value, as usual: "start with 3". Then, we must describe the pattern in the differences: in this case, we're adding consecutive odd numbers (starting with 5). So, one way to express this recursive rule is "start with 3, add consecutive odd numbers (starting with 5) to the previous term". Note that we kind of sneak in two starting places: one for the start of the sequence (3, in this case) and one for the start of the sequence of numbers that are being added on (5, in this case). Tricky!

Sequences that exhibit this pattern are called *quadratic sequences* and they belong to the family of *quadratic relationships*. We'll study quadratic relationships in depth starting in chapter 12.

Example 2.9

Verify that the given sequence is quadratic, and write a recursive rule: 1, 4, 10, 19, 31, 46, . . .

Solution: In order for a sequence to be quadratic the differences between successive terms must form an arithmetic sequence. Let's check:

$$4 - 1 = 3$$

$$10 - 4 = 6$$

$$19 - 10 = 9$$

$$31 - 19 = 12$$

$$46 - 31 = 15$$

The differences are: 3, 6, 9, 12, 15, . . . , and that's an arithmetic sequence with common difference 3. So, yes, the original sequence is quadratic.

Now let's try to write a recursive rule (in sentences). Clearly, we start with 1. Then, we add consecutive multiples of three, starting with 3. So, our rule is "start with 1, add consecutive multiples of three (starting with 3) to the previous term".

At this point, our goal is just to recognize that these sequences are neither arithmetic nor geometric, but follow a different kind of pattern. Writing the formulas for them can be quite challenging – but our brains grow when we stretch them around new ideas! Let's give it a shot.

⁶ This sequence-in-a-sequence stuff can get pretty involved. Here, we found an arithmetic sequence in the *differences* between the terms in our quadratic sequence. But why not build the sequence 1, 4, 12, 27, 51, . . . , in which the sequence of differences is our quadratic sequence! Of course we could keep building sequences like this for as long as we wanted. This isn't just a turducken, it's a *rôti sans pareil*! That's French for "roast without equal", a dish that which calls for 17 different birds, each one stuffed into the body cavity of the next. In the years since the dish was first proposed in 1807 by the French gastronome Grimod de La Reynière, several of the birds called for in the recipe have become endangered species.

2.4.1 (;;; Recursive formulas for quadratic sequences

Extension sections

Sections marked with the Cthulhu (;;; emoticon, like this one, are extension sections that might be a bit more intense than the norm. We encourage you to explore the concepts, but don't feel discouraged if you find the material challenging.

Your math brain grows when you think about hard questions, so that kind of thinking is valuable, even if the concepts aren't completely clear right away. As your algebra skills develop over time, you can hopefully return to these extension sections with more confidence.

In the last section, we looked at the sequence, which came from a rectangular pattern of tiles:

$$3, 8, 15, 24, 35, \dots$$

We wrote the recursive rule in sentences: “start with 3, add consecutive odd numbers (starting with 5) to the previous term”. Can we translate this into a recursive formula?

The first step is easy: $a_1 = 3$. Hooray for small victories!

In order to describe the recursive step, we need to describe the sequence of differences: 5, 7, 9, 11, \dots . Since this is an arithmetic sequence, we know how to write its explicit rule. Let's use the symbol b , so we don't get our sequences confused. Then this sequence is $b_n = 5 + (n-1) \cdot 2$ or, if we use a zero-based rule, $b_n = 3 + n \cdot 2$.

Let's try and put these together:

$$\begin{array}{rclcl} a_1 & = & 3 & & \\ a_2 & = & 8 & = a_1 + 5 & = a_1 + b_1 \\ a_3 & = & 15 & = a_2 + 7 & = a_2 + b_2 \\ a_4 & = & 24 & = a_3 + 9 & = a_3 + b_3 \\ a_5 & = & 35 & = a_4 + 11 & = a_4 + b_4 \end{array}$$

So, our recursive step is that $a_{n+1} = a_n + b_n$. Since we have an explicit formula for the b_n 's, we can replace that part with their explicit rule! Altogether we have:

$$a_1 = 3, \quad a_{n+1} = a_n + 3 + n \cdot 2$$

How can we check to see if we're right? One way is to use the rule to try and re-generate the sequence. Our rule states that $a_1 = 3$. To find a_2 , we can use the rule with $n = 1$ and $n + 1 = 2$:

$$a_2 = a_1 + 3 + 1 \cdot 2 = 3 + 3 + 1 \cdot 2 = 3 + 3 + 2 = 8.$$

Then, we can take one step forward and apply the rule again. Now, $n = 2$ and $n + 1 = 3$:

$$a_3 = a_2 + 3 + 2 \cdot 2 = 8 + 3 + 2 \cdot 2 = 8 + 3 + 4 = 15.$$

Let's go one more step and try $n = 3$ and $n + 1 = 4$:

$$a_4 = a_3 + 3 + 3 \cdot 2 = 15 + 3 + 3 \cdot 2 = 15 + 3 + 6 = 24.$$

Phew! It pays to be patient when working out a convoluted rule like this, but in the end, we can see that our rule is behaving as intended!

Example 2.10

Write a recursive rule for the quadratic sequence: 1, 4, 9, 16, 25, . . .

Solution: A bit of tinkering leads us to the rule “start with 1, add consecutive odd numbers (starting with 3) to the previous term”. So $a_1 = 1$.

How do we write the apparent formula for the odd number pattern? The common difference is 2, and the pattern starts at 3, so $b_n = 3 + (n - 1) \cdot 2$ is the apparent formula for the differences.

Putting the pieces together:

$$\begin{aligned} a_1 &= 1 \\ a_2 &= 4 = a_1 + 3 = a_1 + b_1 \\ a_3 &= 9 = a_2 + 5 = a_2 + b_2 \\ a_4 &= 16 = a_3 + 7 = a_3 + b_3 \\ a_5 &= 25 = a_4 + 9 = a_4 + b_4 \end{aligned}$$

So, again, we have $a_{n+1} = a_n + b_n$. Then, we can replace the b_n with the apparent formula we created for the sequence of differences:

$$a_1 = 1, \quad a_{n+1} = a_n + 3 + (n - 1) \cdot 2.$$

If we would rather use a zero-based rule for the pattern in the differences, we could write:

$$a_1 = 1, \quad a_{n+1} = a_n + 1 + n \cdot 2.$$

Note that even though these two rules look quite different, they are equivalent ways of describing the sequence. In later chapters, we will learn techniques that will help us to explain why these two different-looking rules give us the same result.

2.4.2 (,;) Explicit formulas for quadratic sequences

It seems only proper to discuss a method for writing a non-recursive formula for a quadratic sequence.

There are, in fact, multiple methods for writing rules like this. There is a way that requires knowledge of calculus, there is a method that uses a *system of equations* (more on those in chapter 8), there is the not-so-efficient method of guess and check, and so on. Most of these require knowledge of the structure of a quadratic relationship which (seeing as how we're only here in chapter 2) we haven't discussed yet.

But, there is a clever approach that requires a bit of pattern-hunting and detective work. It doesn't always work out nicely, but it's the approach we'll explore here to get a feel for things.

Let us once again consider the sequence 1, 4, 9, 16, 25, You might have recognized these numbers are the **perfect squares**. That name comes from the idea that we can view these numbers as the areas of squares, as shown in fig. 2.1. The first number is the area of a 1-by-1 square, the second is the area of a 2-by-2 square, then a 3-by-3 square, and so on. Knowing this, we can write any term of the sequence: $a_n = n \cdot n$.

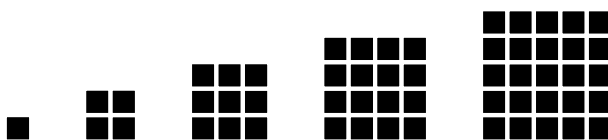


Figure 2.1: Perfect squares

With other quadratic sequences, we can sometimes crack the code if we think about areas of rectangles. We'll look for pairs of integers that could give us the areas we're after, and then look for arithmetic sequences among those integers. The next thing you know, we'll have a non-recursive formula for the quadratic!

Let's go back to the other example we've been studying, the sequence 3, 8, 15, 24, 35, This sequence came from the tile pattern at the start of section 2.4. Go back and take another look at those pictures. What do you notice?

Stage	Number of Squares	Dimensions of Rectangle
1	3	1×3
2	8	2×4
3	15	3×5
4	24	4×6

In stage n , the size of the rectangle is n units tall and $(n + 2)$ units wide! So, we can use those two values to write an explicit formula:

$$a_n = n \cdot (n + 2)$$

How cool is that?

Example 2.11

Write a non-recursive formula to generate the sequence 6, 12, 20, 30, 42,

Solution: It's not obvious how rectangles are related to these numbers, but if we assume we're looking for rectangles with integer side lengths, then there are a limited number of options.

For example, if the first number represents the area of a rectangle with integer side lengths, then it could be either a 1×6 rectangle or a 2×3 rectangle. Let's organize the different options in a table:

Stage	Value	Possible Rectangles
1	6	1×6 or 2×3
2	12	1×12 or 2×6 or 3×4
3	20	1×20 or 2×10 or 4×5
4	30	1×30 or 2×15 or 3×10 or 5×6
5	42	1×42 or 2×21 or 3×14 or 6×7

Now comes the detective work. We are looking for patterns in the factors as they progress through the terms. We've highlighted the key patterns below.

Stage	Value	Possible Rectangles
1	6	1×6 or 2×3
2	12	1×12 or 2×6 or 3×4
3	20	1×20 or 2×10 or 4×5
4	30	1×30 or 2×15 or 3×10 or 5×6
5	42	1×42 or 2×21 or 3×14 or 6×7

Notice that the first set of factors (in yellow) form the arithmetic sequence 2, 3, 4, 5, . . . , and the second set (in green) form the arithmetic sequence 3, 4, 5, 6,

The yellow sequence is always one more than the term number. The green sequence is always two more than the term number. So, we have our explicit formula!

$$a_n = (n + 1) \cdot (n + 2)$$

If these last two sections felt a bit overwhelming, don't worry. After we have some more algebraic tools in our toolbox, we'll return to quadratic relationships and describe them in more detail.

» Chapter summary «

In this chapter we explored and extended the idea of a number pattern with the result that we can now identify and describe certain kinds of patterns in detail. Already we have met the three key “families” of mathematical

relationships that are at the heart of Algebra 1.

The linear family, represented in this chapter by the idea of an arithmetic sequence, will become our focus starting in chapter 5. The exponential family, represented in this chapter by the geometric sequences, will be our focus in chapters 10 and 11. Quadratic sequences and the quadratic family will be in the spotlight starting in chapter chapter 12.

Before we get into the individual families, though, we have a few more tools to add to our algebraic toolbox. Onward!

There is a magic in graphs. The profile of a curve reveals in a flash a whole situation – the life history of an epidemic, a panic, or an era of prosperity. The curve informs the mind, awakens the imagination, convinces.

Henry D. Hubbard, US National Bureau of Standards

Chapter 3

Graphs and data

In chapter 2 we investigated pictures patterns (fractals by Koch and Sierpiński, and patterns made of square tiles) and used those pictures to generate sequences of numbers. We begin this chapter with a discussion of another way to represent our sequences of numbers visually: by making a coordinate graph. Then, we will extend these ideas to create visual representations of other mathematical objects, and of scientific (and other) data.

3.1 Coordinate graphing

Figure 3.1 summarizes the familiar landmarks of the **coordinate plane**. We see the horizontal **x-axis** and the vertical **y-axis**. Using the two axes as number lines, we can locate **ordered pairs** of numbers using the convention (x, y) . The points $(7, -2)$ and $(-6, 3)$ have been plotted as examples. The point $(0, 0)$ where the two axes meet, is a special point called the **origin**.

The axes chop the plane into four regions called **quadrants**, which are numbered starting in the upper right and moving counter-clockwise (as shown in the figure). The signs in parentheses indicate the signs of the x - and y -coordinates in each quadrant. The x -coordinates of the points in Quadrants I and IV are positive, while points in Quadrants II and III have x -coordinates that are negative. Points with positive y -coordinates lie in Quadrants I and II, while points with negative y -coordinates fall in Quadrants III and IV.¹

¹ A point that lies on an axis doesn't actually lie in any of the four quadrants. So, the point $(4, 0)$ lives on the positive x -axis, but neither in Quadrant I nor in Quadrant IV.

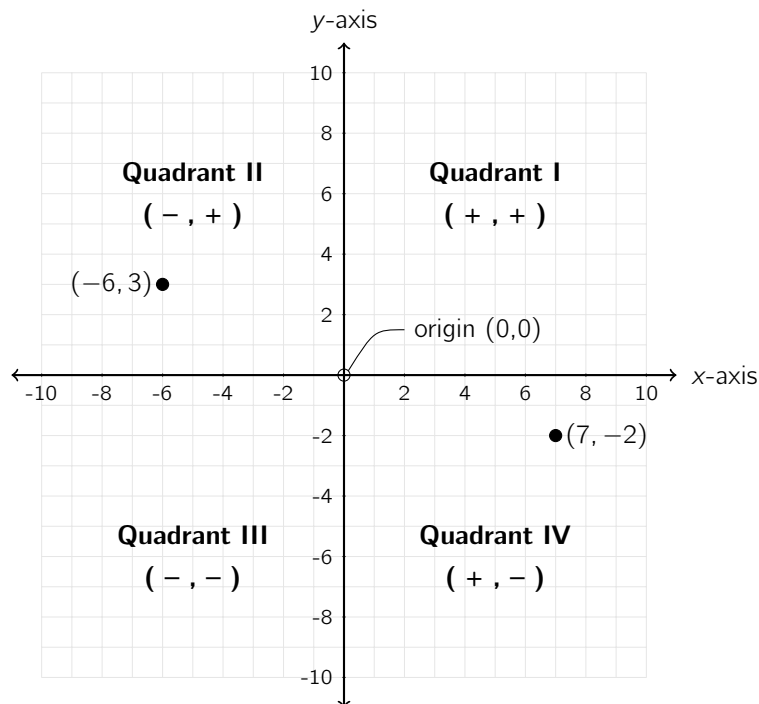


Figure 3.1: The coordinate plane and its important landmarks

Pizza intersections

// startup exploration

Bob's friend Melvin "Hambone" Jones once worked delivering pizzas in his hometown of Euclid, Ohio. The town has streets running north-south and east-west.^a

Hambone is currently parked at the intersection we will call $(0, 0)$. If he drives one block east, he will arrive at the intersection $(1, 0)$. If he then turns right and drives one block south, he will arrive at the intersection $(1, -1)$.

Starting from $(0, 0)$, describe all the intersections that Hambone can reach by driving a total distance of exactly 10 blocks.

^a Euclid, Ohio is also the home of the Polka Hall of Fame, though that has nothing to do with this problem. Also, Euclid isn't laid out in a grid as this problem implies, though it should be, given its name.

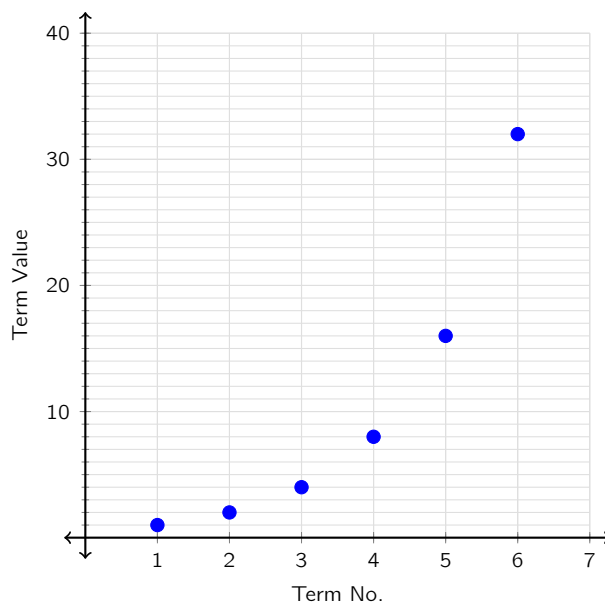
3.1.1 Graphing a sequence

Our goal is to make a visual representation of a sequence on the coordinate plane.

"But," you may be asking yourself, "a sequence is just a list of numbers. How do we make a coordinate graph, which needs coordinate pairs of numbers? It takes *two numbers* to make a *pair*!"

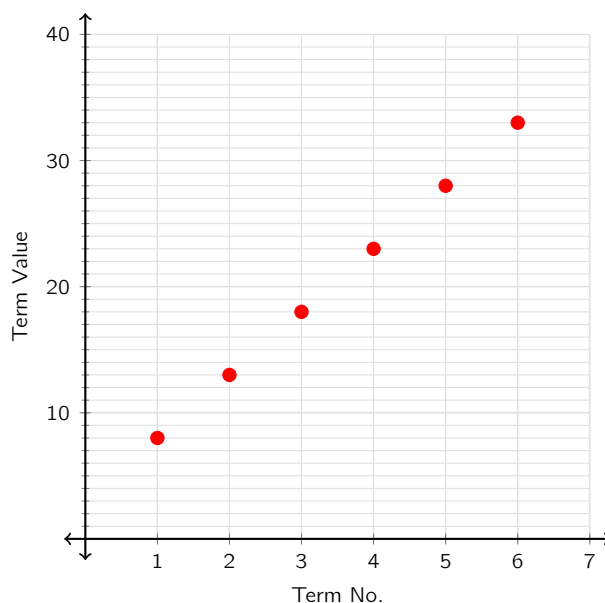
To graph a sequence, we use the term number as the x -coordinate and the term's value as the y -coordinate. Consider the sequence $1, 2, 4, 8, 16, 32, \dots$. The first term is 1, the second term is 2, and so on. We can organize this in a table, and write out the ordered pairs. Then, we can plot those ordered pairs and see a visual representation of our sequence!

Term No.	Term Value	Coord. Pair
x	y	(x, y)
1	1	(1, 1)
2	2	(2, 2)
3	4	(3, 4)
4	8	(4, 8)
5	16	(5, 16)
6	32	(6, 32)



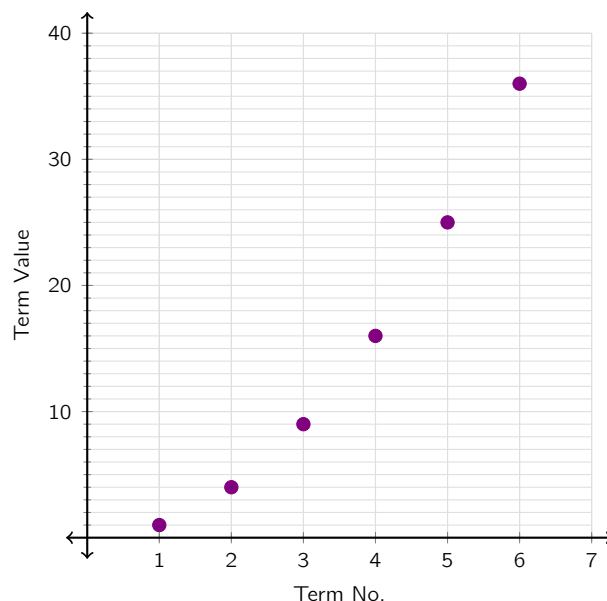
Recall that that sequence above, where the terms have a constant ratio, is called a geometric sequence. Let's look at an example of an arithmetic sequence, in which the terms have a constant difference. For example, consider the sequence $8, 13, 18, 23, 28, 33, \dots$. This sequence produces the following table and graph:

Term No.	Term Value	Coord. Pair
x	y	(x, y)
1	8	(1, 8)
2	13	(2, 13)
3	18	(3, 18)
4	23	(4, 23)
5	28	(5, 28)
6	33	(6, 33)



Finally, let's see what a quadratic pattern looks like. For example, the perfect squares form a quadratic sequence $1, 4, 9, 16, 25, 36, \dots$. Their table and graph go like this:

Term No.	Term Value	Coord. Pair
x	y	(x, y)
1	1	(1, 1)
2	4	(2, 4)
3	9	(3, 9)
4	16	(4, 16)
5	25	(5, 25)
6	36	(6, 36)



Take a moment to compare the graphs of these three sequences. How are they alike? How are they different?

3.1.2 Features of the graph of a sequence

Note that since the term number is always greater than 0, our graphs only show the positive part of the x -axis. We'll soon see that this is an artificial limitation that doesn't apply to most situations: the negative part of the x -axis is just as important as the positive part.

Note also that we haven't connected the dots. A sequence has a first term and a second term, but no one-and-a-halfth term. So, we shouldn't have any points with x -coordinates in between the natural numbers. We'll soon see that this is an artificial limitation, too. Fractional x -values are appropriate for most rules.

If we change our point of view and allow the input to be any real number, we will have turned our sequence into a mathematical relationship called a **function**.² We'll discuss the details of what functions are (and what they are not) in chapter 4.

² Actually, a sequence is *already* a function. The distinction we make here has to do with what kinds of input values are allowed. Later, once we have some additional concepts under our belts, we'll talk about a sequence as "a function whose domain is the natural numbers".

3.2 Algebraic expressions

Don't take all year

// startup exploration

Find three natural numbers x , y , and z which satisfy the equation $28x + 30y + 31z = 365$. Can you find more than one set of numbers x, y, z that satisfy the equation?

In chapter 1 we used the order of operations to simplify numeric expressions, which are made up of numbers and arithmetic operators. For example,

$$3 \cdot 4 - 8(4^2 - 1)$$

is a numeric expression. It contains only numbers and operators and, in the end, it simplifies down to a single number.³ An **algebraic expression** on the other hand can contain letters in addition to numbers and operators, for example

$$3x - 5y + 18.$$

These letters, called **variables**, stand in for numbers that we don't know or which may change.

Variable

A representation of a value that can change. In algebra, variables are often represented by letters. We usually use letters from the Latin alphabet (a, b, c, d, \dots), but we sometimes also use other symbols, such as letters from the Greek alphabet ($\alpha, \beta, \gamma, \delta, \dots$).

Algebraic expression

A symbolic representation of mathematical operations that can involve both numbers and variables.

3.2.1 Numbers and variables

In your mathematical career so far, you have probably worked with letters that stand in for numbers. Recall the formula for the circumference of a circle:

$$C = \pi d.$$

³ Spoiler alert: It's -108 .

This formula explains the relationship between d , which stands in for the diameter of some circle, and C , which stands for the circumference of that circle. The letters d and C are variables. They stand in for numbers that can change, depending on which circle we're talking about.⁴

Once we start to introduce letters into our expressions, we have to discuss some standard notation and terminology. As we have seen, we don't usually write any multiplication symbol when multiplying a number by a variable. Rather than writing $3 \cdot x$ or $3(x)$, we can write $3x$ without anything in between.

An algebraic expression that is built using only multiplication (or division) is called a **term**. For example, $3x$ and $\frac{1}{2}m$ are terms. On the other hand, the expression $3x + 2y$ is not a term because it includes addition. In fact, this expression is the sum of two terms.

Term

An algebraic expression that represents only multiplication and division between variables and constants (numbers).

When we have the product of a number and a variable, like $3x$ or $-11g$, the number part is called the **coefficient** of the term. So, the coefficient of $3x$ is 3, and -11 is the coefficient of $-11x$. If we have a variable all alone without a number attached, like y or w , then we picture a "phantom 1" lurking there as the coefficient: y is the same as $1y$, and $1w$ is the same as w .

Coefficient

The numerical factor in a term with a variable. If no number is explicitly written, the coefficient is understood to be 1.

3.2.2 Evaluating algebraic expressions

A variable is a "placeholder" that stands in for a number. We can only determine the value of an algebraic expression if we know what numbers the different variables represent.

Consider the expression $3x$. If we know that x represents 15, then we can **evaluate** the expression $3x$ in the case that $x = 15$. In that case, it must be that $3x = 3(15) = 45$.

⁴ Note that π is *not* a variable. We use a (Greek) letter in this case not because the value of π might change, but because it's an irrational number that is impossible to write out in full. We often use letters to stand in for mathematical objects that are inconvenient, sometimes impossible, to write down in another way: e , i , ϕ , and \aleph_0 each has a special mathematical meaning.

When we evaluate an algebraic expression, we substitute in values for its variables, and then simplify the resulting numeric expression using the order of operations. It is a really good habit always to use parentheses when substituting numeric values for variables. This can avoid confusion about negative numbers!

Example 3.1

Evaluate the expressions (a) $6x + 4$, and (b) $x^2 - 5$ for the x values 3, -1 , and $\frac{1}{2}$.

Solution:

- (a) To evaluate the expression $6x + 4$ for the given values of x , we simply substitute and follow the order of operations.

When $x = 3$:

$$\begin{aligned} 6x + 4 &= 6(3) + 4 \\ &= 18 + 4 \\ &= 22 \end{aligned}$$

When $x = -1$:

$$\begin{aligned} 6x + 4 &= 6(-1) + 4 \\ &= -6 + 4 \\ &= -2 \end{aligned}$$

When $x = \frac{1}{2}$:

$$\begin{aligned} 6x + 4 &= 6\left(\frac{1}{2}\right) + 4 \\ &= 3 + 4 \\ &= 7 \end{aligned}$$

- (b) We do the same in order to evaluate $x^2 - 5$ for the given x values.

When $x = 3$:

$$\begin{aligned} x^2 - 5 &= (3)^2 - 5 \\ &= 9 - 5 \\ &= 4 \end{aligned}$$

When $x = -1$:

$$\begin{aligned} x^2 - 5 &= (-1)^2 - 5 \\ &= 1 - 5 \\ &= -4 \end{aligned}$$

When $x = \frac{1}{2}$:

$$\begin{aligned} x^2 - 5 &= \left(\frac{1}{2}\right)^2 - 5 \\ &= \frac{1}{4} - 5 \\ &= -\frac{19}{4} \end{aligned}$$

Note how the parentheses help out when $x = -1$. Without those parentheses, we would have run the risk of making the most common mistake in Algebra 1! Remember the difference between $(-1)^2$ and -1^2 .

3.3 Graphing a function

The graphs of sequences that we created earlier were quite limited. Since sequences use only the natural numbers as input values, the only points we had available to plot were the points where $x = 1, 2, 3, 4, \dots$. But now, knowing how to evaluate algebraic expressions, we can create more complete graphs by choosing a wider variety of x -values.

Extending our sequences

// startup exploration

Write a zero-based explicit rule for the arithmetic sequence shown below (this is the second example from section 3.1). First write the rule in terms of n and a_n , then translate your rules into a graphable format in terms of x and y .

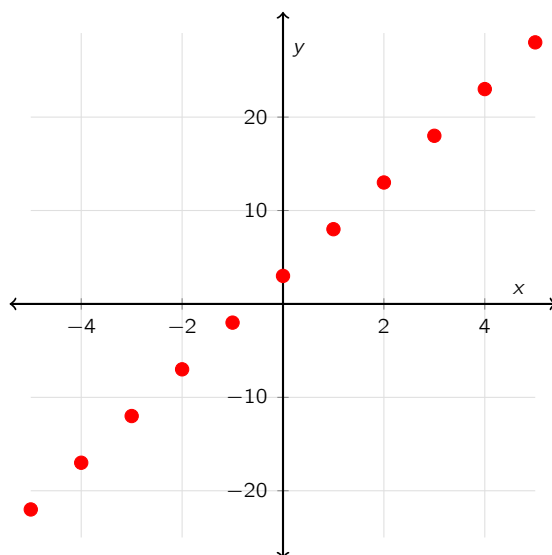
8, 13, 18, 23, 28, \dots

The given sequence represents the y -values of the rule for the x -values 1, 2, 3, 4, and 5. Evaluate your rule for the x -values 0, -1, -2, -3, -4, and -5. Then, plot these 11 points on a coordinate grid.

The rule for the sequence in the startup exploration is $a_n = 5n + 3$, or in terms of x and y , we have the rule $y = 5x + 3$. To create the coordinate graph, we can substitute the different x -values into the rule and compute the y -values. The middle column in the table below is our “process column” in which we substitute an x -value and compute the corresponding y -value.

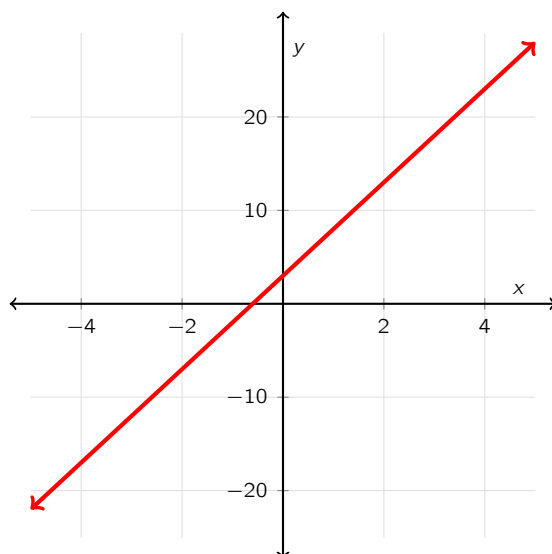
x	$y = 5x + 3$	(x, y)
0	$y = 5(0) + 3 = 0 + 3 = 3$	$(0, 3)$
-1	$y = 5(-1) + 3 = -5 + 3 = -2$	$(-1, -2)$
-2	$y = 5(-2) + 3 = -10 + 3 = -7$	$(-2, -7)$
-3	$y = 5(-3) + 3 = -15 + 3 = -12$	$(-3, -12)$
-4	$y = 5(-4) + 3 = -20 + 3 = -17$	$(-4, -17)$
-5	$y = 5(-5) + 3 = -25 + 3 = -22$	$(-5, -22)$

In the end, we generate 6 new coordinate pairs, which we can graph alongside the five points that we were given.



Can you anticipate the location of the point that we plot when $x = \frac{1}{3}$? What about when $x = -\frac{5}{2}$? What would our graph look like if we used all of the points in \mathbb{R} as the x -values?

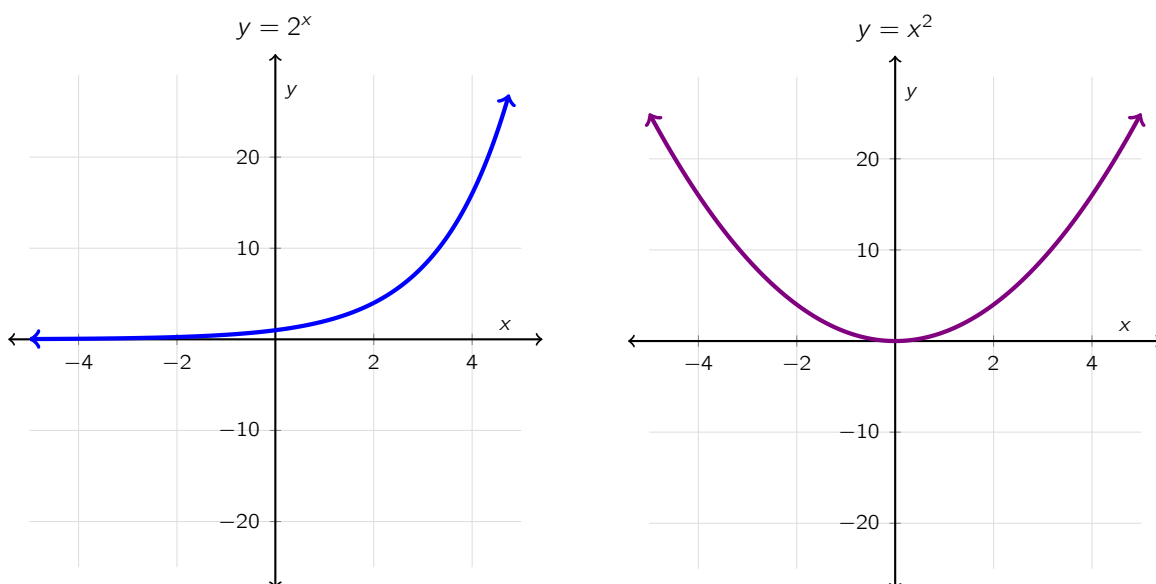
When we use all of the real numbers as input to the rule $y = 5x + 3$, the resulting graph is a straight line. Of course, it would be impossible to actually plot *all* of the points (there are infinitely many of them), but the pattern holds true, and so we can replace the dots with a continuous line.



The rules below correspond to the other two sequences we studied in section 3.1. Under each rule is the graph that is created when we plot the rule over \mathbb{R} .

Are you surprised by these two graphs? Back in section 3.1, the graphs of these two sequence looked similar. But, we were only looking at the first quadrant! Their graphs are very different for negative values of x .

The moral of the story is that when we are asked to graph an equation by hand, we need to use a variety of different x -values, including negative numbers and fractions. Often, a problem will clearly indicate exactly what values to use.



3.3.1 Criteria for high quality graphs

In algebra we make a distinction between “sketching” and “graphing”. A sketch is just a quick drawing and it doesn’t need to be super accurate. A sketch can be drawn on notebook paper (or scribbled on a napkin).

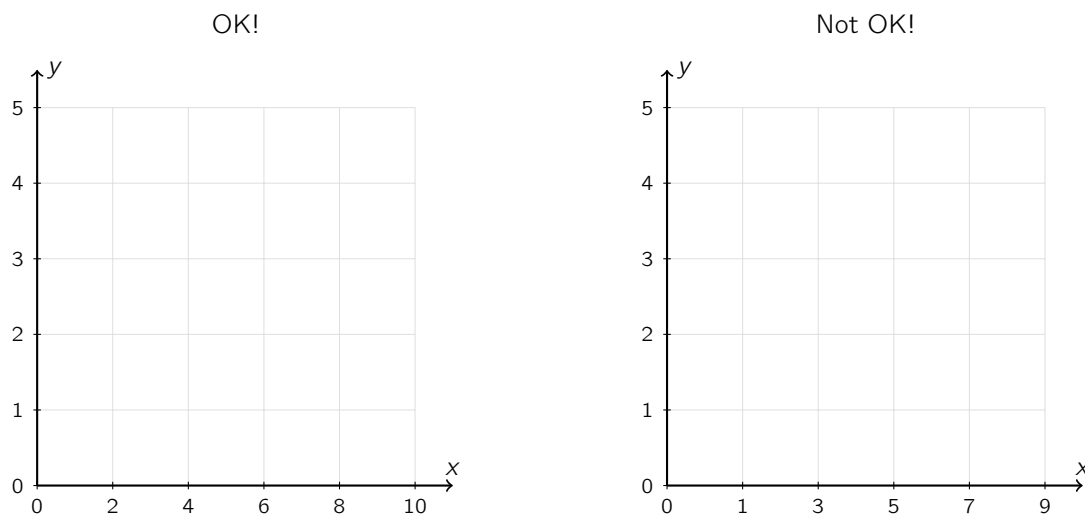
On the other hand, a proper graph is meant to communicate something to the viewer. A graph must be accurately drawn on graph paper. The most important features of whatever we’re graphing – a sequence, a function, a plot of a data set – must be clear, accurate, and neatly represented. Here are some guidelines for creating high-quality graphs.

High quality graphs should be drawn on graph paper. An individual graph does not have to take up an entire piece of graph paper, although it should be drawn large enough to be easily understood.

Use a ruler or straightedge to make straight lines, in particular the coordinate axes or any linear data. Plus, we should draw arrowheads on anything that continues forever. This includes axes, lines, and curves (note the arrowheads in the graphs above).

Be mindful when choosing a scale for the axes. Choose a scale that fits the data and ensures that your graph shows all important features of the curve. The origin does not necessarily need to be in the center of the grid. For example, if we are plotting data that includes only positive values, then we only need the first quadrant. The origin in this case might be in the lower left-hand corner of the graph.

Scales for the x - and y -axes can be different, and in some cases must be different, so this can take a little bit of planning. The scale must be the same for the whole length of an axis, and cannot have any “jumps” or “breaks”. Changes in the scale will distort the shape of the graph, which defeats the purpose! Compare the two grids below. The grid on the left is fine, but the grid on the right includes a common mistake. Can you identify the problem?



High quality graphs are clearly labeled. The scale should be indicated on the axes, and the axes should be labeled. For graphs of equations, this might mean simply labeling the axes x and y . When plotting data, include more informative names like “time” or “distance”.

A final question to consider regarding data points: to connect or not to connect? In the next section, we will discuss this question in some detail. But we’ve already seen some important pieces of this puzzle. When we have a sequence, or other data that skips values, we should not connect the points.

If we do want to connect the points, say, when graphing a function by plotting some sample points, we should connect with a smooth curve and *not* individual line segments. If we use a straight line to connect points, then we are telling our audience that the data in between the points is linear, which may not be the case! Go back and have a look at the graph of $y = x^2$. It doesn’t come to a point at the bottom. Rather it’s a smooth curve that passes through the origin.

3.4 Patterns in data

The graphs that we have been working with so far have been very orderly. Technical and scientific data, however, are not always so tidy. Data can be noisy, messy, and incomplete. We will need some tools that can help us to see and describe patterns that may (or may not) exist in experimental data.

Water consumption

// startup exploration

The graph shown in fig. 3.2 depicts water consumption in Edmonton, capital of the Canadian province of Alberta, during the gold medal men's ice hockey game at the 2010 Winter Olympics in Vancouver. The game was played between Canada and United States.

Water consumption during the game is shown in blue, while data from the same time period on the previous day is shown in green.

Write down anything you notice or wonder about the data presented in this graph.

A few notes: Ice hockey games are played in three 20-minute periods with breaks in between. In this particular game, the score was tied at the end of regular play. Canada scored the winning goal in overtime and was then awarded the gold medal.

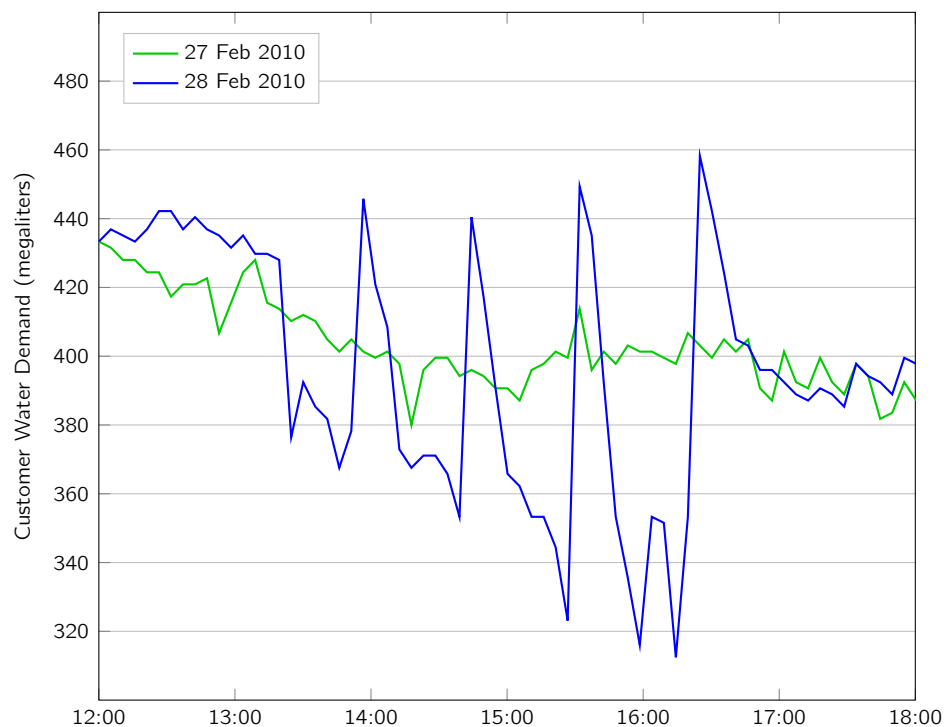


Figure 3.2: Water consumption in Edmonton on 27 and 28 February 2010 (source: EPCOR)

In the graph of Edmonton water consumption, the amount of water being used varies depending on the time of day (not the other way around). We say that “time of day” is the *independent variable* and “water demand” is the *dependent variable*.

Independent variable

A variable whose values affect the values of another variable. In a graph of the relationship between two variables, the quantity represented on the horizontal axis (the x-axis) usually represents the independent variable.

Dependent variable

A variable whose values depend on the values of another variable. In a graph of the relationship between two variables, the quantity represented on the vertical axis (the y-axis) usually represents the dependent variable.

3.4.1 Correlation

We can compare just about any two quantities. One way to do this is with a graph called a **scatter plot**.

Scatter plot

A graph that relates data of two different sets. The two sets of data are displayed as ordered pairs.

Suppose we wish to compare, say, the height and weight for all of the players on the top two 2010 Olympic men’s ice hockey teams. Let’s make a graph comparing every player’s height and weight as ordered pairs: (height, weight). This graph is shown on the left in section 3.4.1.⁵

While we’re at it, let’s make another comparison. The graph on the right in section 3.4.1 plots every player’s height and jersey number as ordered pairs (height, jersey number). What do you notice about these two graphs? How are they the same? How are they different?

Notice that the graph of weight and height has a clear upward slant. This seems reasonable: the taller someone is, the more we might expect that person to weigh. There are some data points that don’t fit the trend, but

⁵ Data from the International Olympic Committee, as reported in Wikipedia.

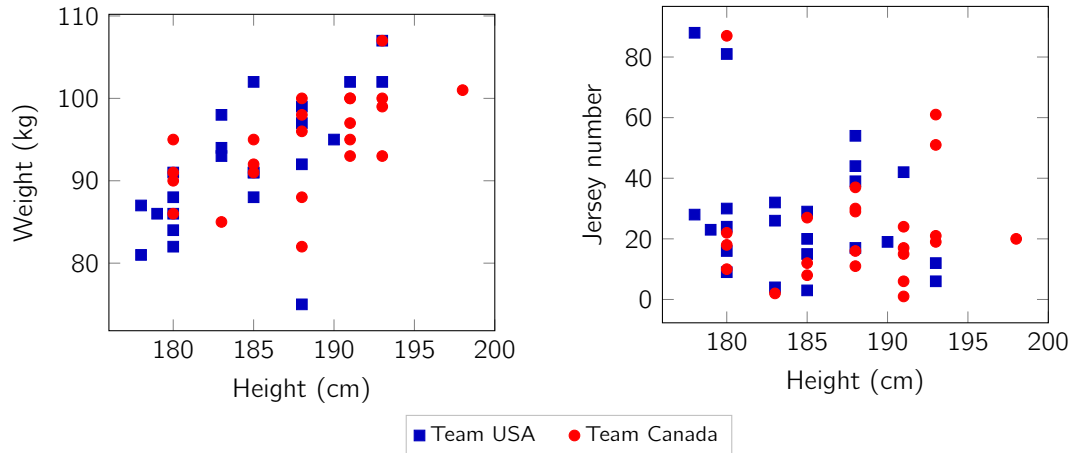


Figure 3.3: Comparing the Canadian and US men’s ice hockey teams (2010 Winter Olympics).

generally speaking the data points are increasing as we look toward the right on the graph. We say that this data shows a positive **correlation**.

Correlation

A trend between two sets of data, as seen in a scatter plot. A trend can show positive, negative, or no correlation. Positive correlation shows an **increasing** trend in data. Negative correlation shows a **decreasing** trend in data.

The graph of jersey number versus height is more of a blob. There’s no clear trend in this data, and so we say that it shows *no correlation*. This makes sense, too: there’s no logical connection between a player’s height and the number they wear on their shirt.

A third possibility would be data which shows a negative correlation, meaning that the data are decreasing as we look towards the right on the graph. Can you imagine two variables that might show a negative correlation when compared on a scatter plot?

Graphing experimental data on a scatter plot helps us to see if there is a relationship between variables. If there is, a pattern will emerge in the graph. The points will fall (approximately) in a line or a curve and will have a correlation. A key thing to remember when it comes to looking at data is that “correlation does not imply causation”. In other words: If we see that two variables are correlated, we might be tempted to assume that the change in one variable *causes* the change in the other. This is sometimes true, but not always.

For example, it seems reasonable to believe that a change in height will cause a change in weight. But, there is data that shows a positive correlation between “consumption of mozzarella cheese per person” and “number of civil engineering doctorates awarded”. This has to be a coincidence! There’s no (good) reason to think that changing one of these variables would cause a change in the other one.⁶

⁶ This fact is courtesy of the website Spurious Correlations, which has many graphs of interesting and ridiculous data that show

3.4.2 Continuous and discrete data

When we drew the graph of a sequence, we didn't connect the dots. A sequence has a first term and a second term, but no one-and-a-halfth term. The x -values have no "in-betweens".

Similarly, imagine a graph showing "time" as the independent variable and "number of hockey players on the ice" as the dependent variable. In this case, the y -values would have no "in-betweens". There could be 11 players or 12 players on the ice, but never 11.5 players. This is called **discrete data**.

Discrete data

Data for which it doesn't make sense for measurements to exist between given data points. Discrete data often involves *counting items*, such as the number of cars in a parking lot over time.

On the other hand, when we started to picture the graph of a rule that could accept any real number as input, we drew a continuous line on the graph. The graph of water consumption in Edmonton, is jagged, spiky, and irregular – but it's a continuous line. We can measure how much water has been used at any point in time, and we can measure the amount of water in fractions of a unit. We call this **continuous data**.

Continuous data

Data that has no holes, gaps, or breaks. Continuous data often involves *measuring some value* where measurements exist (and may change) between data points. For example, a person's height over time.

Consider a data collection scenario in which we want to graph "Bob's distance from home at any given time (in kilometers)". Bob is always a certain distance away from home (perhaps 0 km, if he is at home), and he could be any distance (even fractions of a kilometer). So, this is continuous data and our graph should be an unbroken line.

Now consider the scenario in which we graph "number of customers in line at the cheese counter at Middle Market." Although there is always a certain number of people in line (maybe zero), there can never be 5.7 people in line. We must count people, and so this data is discrete. Our graph would have to "jump" from 5 people to 6 people without going through the in-between values.

correlation but not causation. Sometimes, it's fun to try to "explain" the correlation. Perhaps in this case the increase in mozzarella cheese consumption is due to an increase in takeout pizza demand, which leads to more pizza delivery shops, which in turn requires more roads, which means we need more engineers to design them.

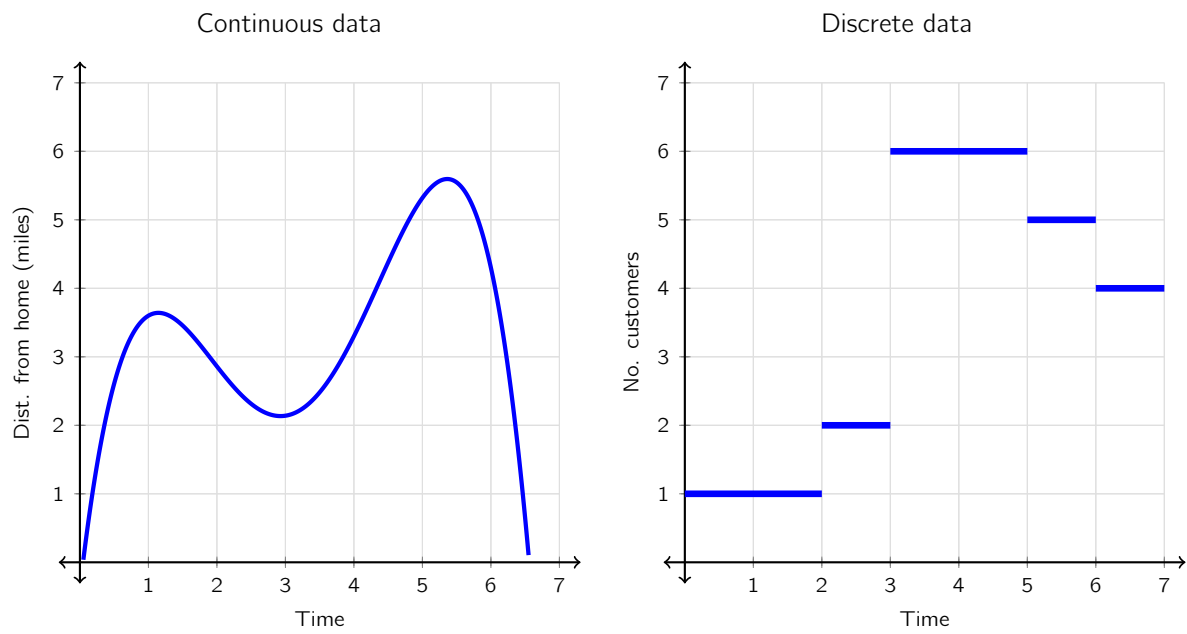


Figure 3.4: Comparing graphs of continuous and discrete data.

3.5 Interpreting graphs

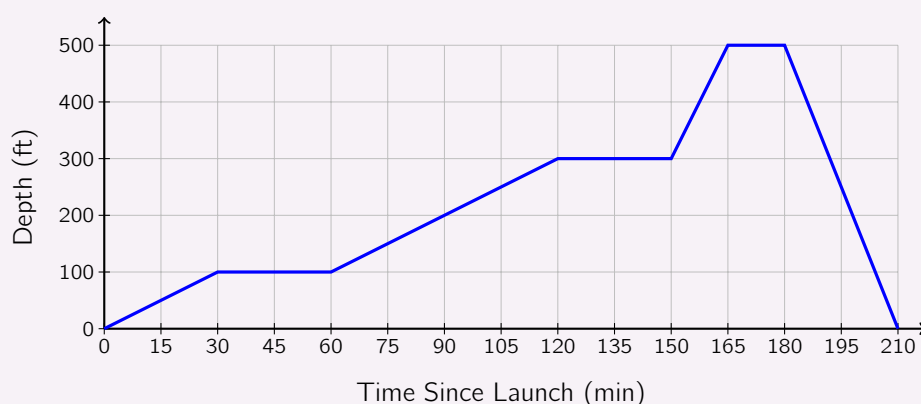
In this section, our goal is to hone our skills at understanding what a graph is *communicating*.

Yearleigh's submersible

// startup exploration

Always questing after the most delicious ingredients, Yearleigh buys an underwater submersible vehicle so that she can hunt the ocean floor for interesting sea plants. The graph below shows the depth of Yearleigh's submersible over time.

Study the graph. What can you tell about what's happening? Write a short paragraph telling, in words, the same story that the graph is telling visually.



At the beginning of the trip, Yearleigh's submersible dives a total of 100 feet in the first 30 minutes. At the end of the trip, it returns to the surface, rising 500 feet in the last 30 minutes. This tells us that the depth of the submersible was *changing much faster* at the end of the journey compared to the beginning.

Note that the graph reflects this: the line is quite steep at the end of the trip and not so steep at the beginning. The steeper the line, the faster the dependent variable is changing with respect to the independent variable.⁷

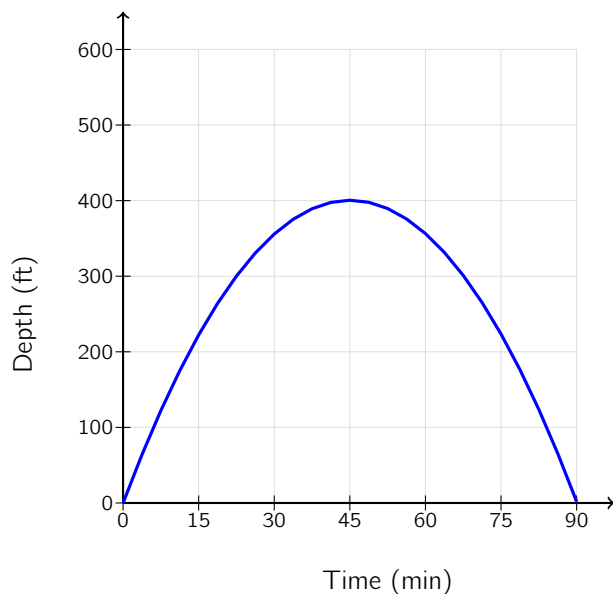
Plus, we can tell from the graph when the submersible is getting deeper (the depth is increasing; the line shows a positive trend) and when it is getting shallower (the depth is decreasing over time; the line shows a negative trend).

Notice that other parts of the graph are horizontal, for example between 30 and 60 minutes. This tells us that the depth of the submersible stayed the same during that time. Of course Yearleigh could still be moving around below the surface, but she remains at a constant depth, neither diving nor surfacing.

⁷ We call this the "rate of change", and it will become an important focus of our work in chapter 7.

3.5.1 Interpreting curves

Straight lines are fine, but what if the graph shows curved lines? Consider this graph showing the submersible's depth over time. What story would we tell about this graph?



One way to get a feel for what's happening is to imagine leaning a ruler or pencil against the rounded edge of a soda can. Then picture the ruler rolling along the side of the can, and how the angle of the ruler will change as it rolls.

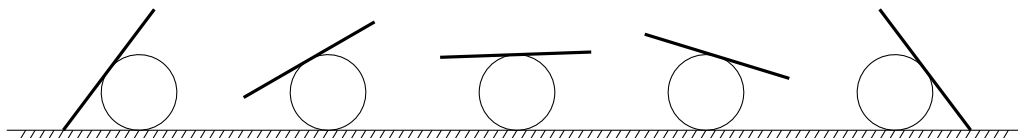
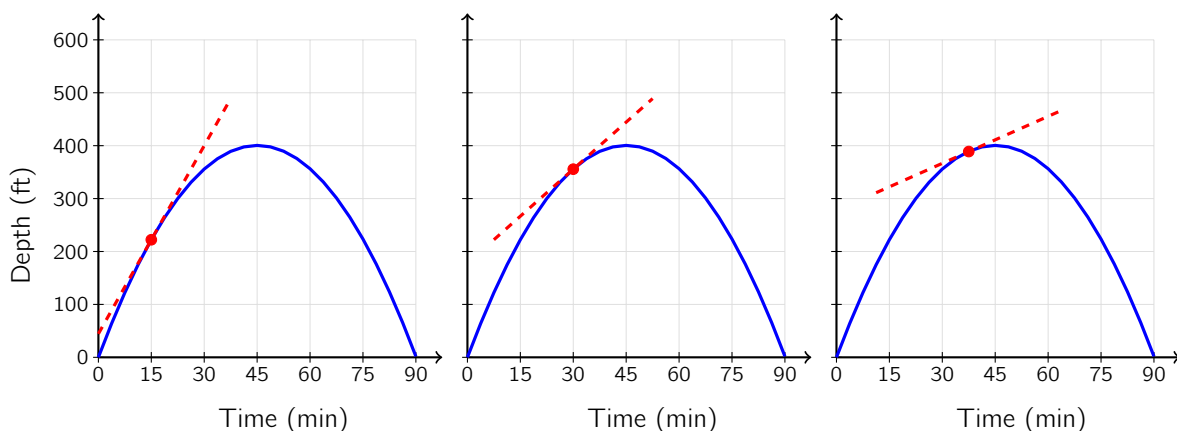


Figure 3.5: Pencil rolling along the side of a can

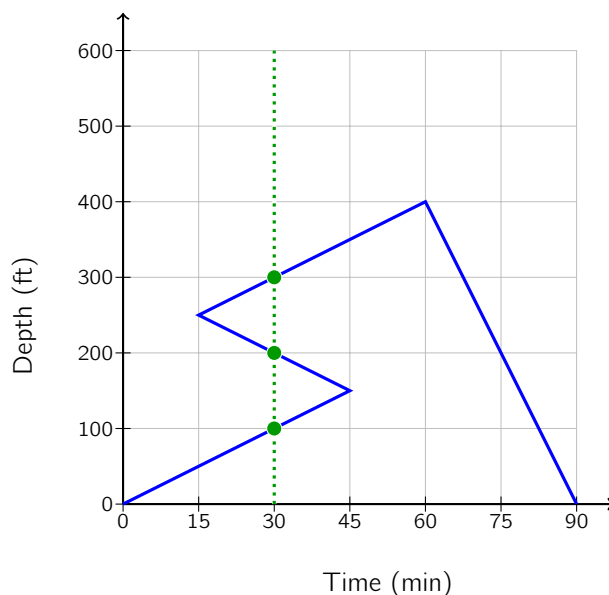
Now picture a straight line rolling along the surface of the curved line in the graph above. The straight line approximates the curved line at the point where they touch.



We can see that the submersible starts out diving at a fairly high rate. It gradually slows its rate of descent until it eventually stops diving. Then it gradually accelerates as it returns to the surface.⁸

3.5.2 Impossible situations

Consider this graph showing the submersible's depth over time. What's going on here?



This graph is a problem, given the context of “depth of the submersible over time”. Consider this question: How deep is the submersible 30 minutes after launch?

According to the graph, the submersible is 100 feet deep... and 200 feet deep... *and* 300 feet deep... all at the same time! That’s impossible!

This kind of problem could pop up in other places. For example, in a distance-time graph, points that line up vertically mean that something is in more than one place at one instant in time. Though we might wish reality were different, nothing can be in two (or more) places at the same time.

Graphs like this – where several y -values stack up vertically over the same x -value – violate a certain requirement that we will learn about in the next chapter, as we delve into the important mathematical idea of a *function*.

⁸ Believe it or not, this idea – approximating a curved line with a straight line – is one of the fundamental motivating ideas in calculus. As you work to interpret these curved graphs, you’re growing your calculus brain, right here in Algebra 1. How cool is that?

» Chapter summary «

In this chapter we began with plotting points to create the graph of a sequence. Then, we built upon these ideas to create the graph of function. We then turned our attention from making graphs to understanding graphs that are given to us.

These two skills – creating a graph from a given rule, and extracting information from a given graph – are both key mathematical skills that we will develop in this course. At the heart of both skills lies the connection between a function's rule and its graph. This connection, and more about the concept of a function, are discussed in the next chapter. Onward!

On two occasions I have been asked, “Pray, Mr. Babbage, if you put into the machine wrong figures, will the right answers come out?” I am not able rightly to apprehend the kind of confusion of ideas that could provoke such a question.

Charles Babbage, English inventor of the first mechanical computer

Chapter 4

Fundamentals of functions

We have seen a number of relationships so far: the relationship between distance and time of a moving object, for example, or the relationship between a number's position in a sequence and that number's value. A *function* is a certain kind of relationship that is of key importance to us in algebra and, indeed, throughout mathematics. In this chapter we'll get an overview of functions and their different representations. Then, in the rest of the course, we will look closely at three specific kinds of functions.

4.1 Mathematical relationships

Number machines

// startup exploration

A number machine accepts numbers as input and produces numbers as output. When machine A receives the number 8 as input, it produces the output 28. When this machine receives -12 as input, the output is -42 .

Machine B has a different mechanism for producing output values. When machine B receives the number 8 as input, it outputs 15. When it receives -12 as input, the output is 23.

What do you suppose each machine will output when given the number -7 as input? Write a sentence explaining how you believe each machine works.

The number machines described in the startup exploration relate certain input values to certain output values. Mathematically speaking, a number machine defines a *relation*.

Relation

A **relation** defines how certain numbers (or other objects) are connected to other numbers (or objects). We can think of a relation as a collection of ordered pairs (x, y) , meaning “ x is related to y ”.

We might think of a number machine as a set of pairs of numbers: (input, output). Machine A in the startup exploration is defined by the pairs $(8, 28)$ and $(-12, -42)$. We sometimes say that “machine A maps 8 to 28” and that “-12 is mapped to -42”.

One possible explanation is that the machine multiplies the input value by 3.5 to produce the output value. In this case, we’d expect the input -7 to produce the output -24.5 . In other words, $(-7, -24.5)$ is a third point associated with machine A.

For machine B we are given the points $(8, 15)$ and $(-12, 23)$. One, somewhat complicated, explanation is that the machine doubles the input, takes the absolute value of the result, and then subtracts 1. In this case, we’d have $(-7, 13)$ as another input-output pair for machine B.¹

We can define relations that compare things other than numbers. For example, we might define the “name-the-mother” relation that accepts a person as input and which gives that person’s biological mother as output. Or, we might define the “name-the-pet” relation, which accepts a human as input and outputs other animals.

Note that there is a difference between these two relations we have just defined. Not all people own pets, so some inputs to the name-the-pet machine have no output. Plus, some people own multiple pets, so some inputs to the name-the-pet machine have multiple outputs. On the other hand, the name-the-mother machine is guaranteed to always produce a unique output: every person has a biological mother, and it is impossible for a person to have more than one biological mother!²

We saw something similar in section 3.5, where certain graphs could be drawn in a way that would suggest an impossible situation. For example, we drew a graph (shown in fig. 4.1) that suggested Yearleigh’s submersible could at three different depths simultaneously. This is as ridiculous as the idea that a person could have more than one biological mother!

Relationships which follow this very natural rule – any given input produces a unique output – have a special status in algebra. They’re called **functions**.

¹ There are other rules that we could write based on the two data points we were given, and perhaps you thought of different rules than the ones we describe here. Can you find alternative explanations for how each machine works? Be creative!

² We don’t deny that a family with two mothers is still a family! It is true that a person can have multiple mothers in a legal or emotional sense. . . but a person always has exactly one *biological* mother.

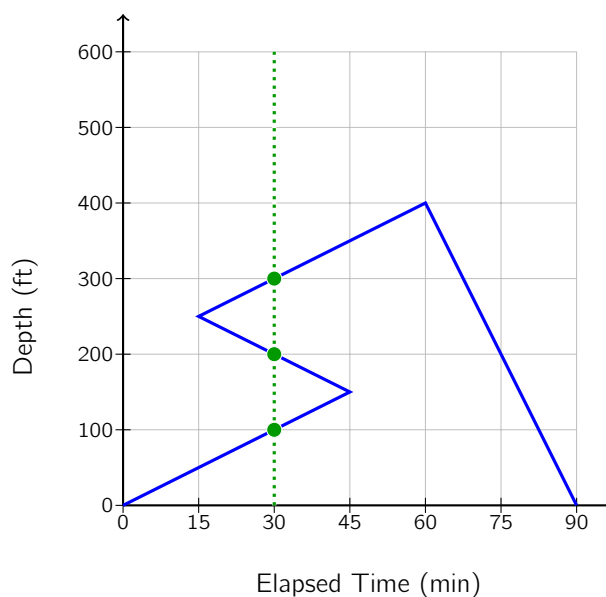


Figure 4.1: A graph of depth over time? Impossible!

Function

A **function** is a special type of relation in which the ordered pairs have the following property: each x -value is paired with one, and only one, y -value.

The “one and only one” requirement is what makes a function special. The graph in fig. 4.1 violates this requirement: certain x -values have more than one associated y -value (look at x values between 15 and 45 minutes). So, this graph does not depict a function.

The sequences we studied in chapter 2 are functions: it would be silly to suggest that a sequence could have more than one value as its third term. Depending on how we think about number machines, we might expect that a machine would act predictably and always produce the same output for a given input. A number machine that behaves in this way defines a function.

Representing relations

There are several ways to represent a mathematical relation. We might represent a relation using:

- A graph. This representation gives us a visual of the relationship between input values and output values.
- A table of values. This representation gives us an organized list of input values and their corresponding output values. (Related forms include a *mapping diagram* or just a collection of ordered pairs.)
- An equation. This representation gives us a rule for turning input values into output values, the way a number machine might do.

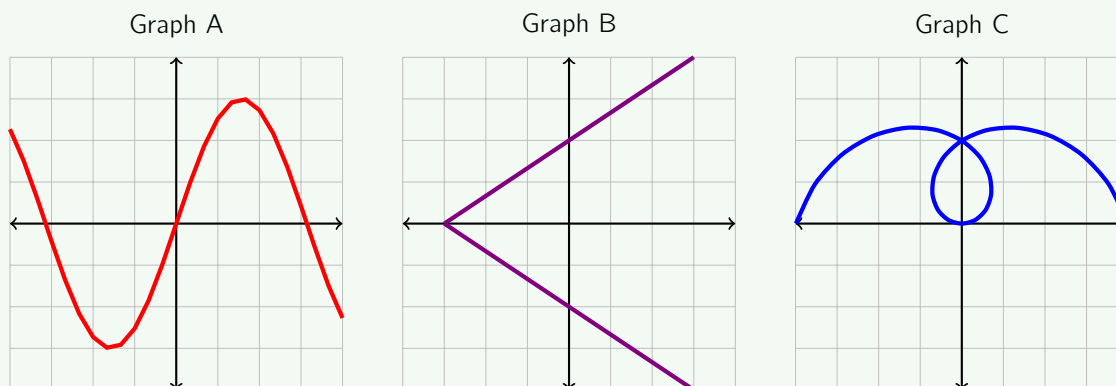
Since functions hold a special position in the universe of all relations, our first task is to determine whether a given mathematical relationship is, in fact, a function. To do this we have to check it against the definition of what it means to be a function. In other words, we must make sure each x -value corresponds to one and only one y -value.

4.1.1 Graphs of functions

In order for a graph to be a function it must avoid the situation in which multiple y -values stack up above a particular x -value. A handy way of remembering this requirement is that the graph of a function passes the **vertical line test**. The vertical line test states that for the graph of a function, any vertical line drawn on the graph will intersect the graph *exactly once*.

Example 4.1

Which of the graphs below, if any, represents a function? How do you know?

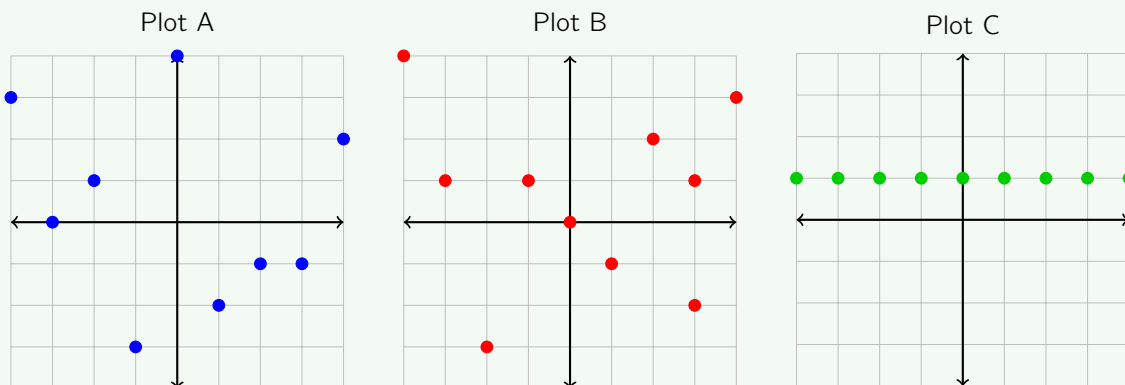


Solution: Only graph A represents a function. We can draw vertical lines on graphs B and C which intersect the graph at more than one point. This means that those x -values have more than one y -value, violating the definition of function.

Note that in graph C, certain sections of the graph pass the vertical line test. There is no partial credit, though. A graph fails the vertical line test if any vertical line (even just one) crosses the graph in more than one point. This happens also with the graph in fig. 4.1. That graph fails the vertical line test for some vertical lines. Therefore the relationship depicted is not a function.

Example 4.2

Which of the scatter plots below, if any, represents a function? How do you know? You may have to look closely!



Solution: Plots A and C both pass the vertical line test, and therefore represent functions. Plot B fails the vertical line test because it fails for the vertical line at $x = 3$.

Note that we don't care if different x -values map to the same y . For example, in plot C, all of the x -values map to the same y -value. That's OK. In other words, a function *can* fail the "horizontal line test" without penalty.³

For example, the name-the-mother function does not pass the horizontal line test because two different inputs (people) can have the same output (mother). In fact, it is not at all uncommon for two different people to have the same mother: they're called siblings!

4.1.2 Functions from points

A scatter plot gives a visual representation of a collection of ordered pairs. If the ordered pairs are presented in an alternative format – say, a table of values – we can use the same reasoning to determine whether the given relation is a function.

³ The horizontal line test *is* a thing, and passing or failing the horizontal line test does have a mathematical meaning. A discussion of this will have to wait until later, however.

Example 4.3

Which of the data sets below, if any, represents a function? How do you know?

Set A

x	y
-3	4
-2	6
-1	4
0	2
1	0
2	-2
3	0

Set B

x	y
2	1
1	2
-1	7
4	3
3	4
1	0
0	6

Set C

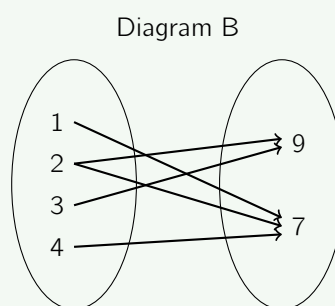
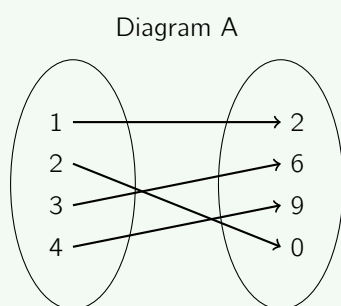
$$\left\{ (0, 4); (-1, 6); (2, 3); (0, 4); (1, 5); (-2, 5) \right\}$$

Solution: Set A is a function. There are no repeated x-values! Set B is not a function. The x-value 1 appears twice and with two different related y-values (0 and 2). This violates the definition of function. Set C is a function. The x value 0 appears twice, but it is mapped to the same value in both cases.

A **mapping diagram** is a similar way of representing a relationship between specific input and output values. Here we draw two regions (represented by a box or an oval). In one region, we list the input values, in the other region we list the output values. Then we draw arrows to show which input maps to which output.

Example 4.4

Which of the mapping diagrams below, if any, represents a function? How do you know?



Solution: The first mapping is a function. Each x-value maps to exactly one y-value. The second mapping is not a function: the value 2 maps to both 7 and 9.

Note again that the problem in diagram 2 is *not* that multiple arrows point to each of the output numbers. It's OK for multiple arrows to go *toward* a certain output value. The problem is that the input value 2 has multiple arrows *leaving* and going to different targets (both 7 and 9).

4.1.3 Equations of functions

The “number machine” metaphor is often helpful when thinking about a function. A number (an x -value) goes into the machine, the machine’s rule works on the number, and another number comes out of the machine (the y -value). Most of the time, when we have a particular rule and a particular input, we get a single, predictable output.

For example, the equation $y = 3x$ defines a function. We can substitute in 5 for x . The function machine multiplies 5 by 3 and outputs 15. This generates the ordered pair $(5, 15)$ and we say that “5 maps to 15”. If we choose other values of x , we can generate more ordered pairs.

Not every equation defines a function, though. In order to be a function, every input value must produce one and only one output value. In Algebra 2 we will study many interesting and important mathematical relationships that are not functions. We will learn, for instance, how to write the equation for a circle – and a circle is not a function. (Can you explain why not?)

Our focus in Algebra 1, though, will be on three main “families” of functions, and our rules will all be of the form “ y equals some expression in terms of x ”.

4.1.4 Function notation

We can thank Swiss mathematician Leonhard Euler for the concept of a function.⁴ He was the first to coin the term “function” and was the first to use a handy way of writing a function, called *function notation*.

Up until now, we have used “ $y = \text{something in terms of } x$ ” to define all of our rules. Euler’s function notation looks a bit different, but it is quite helpful in certain situations. The generic form of function notation is “ $f(x) = \text{something in terms of } x$ ”. In this case, we simply replace y with the notation $f(x)$, which is read aloud as “ f of x ” or “ f as a function of x ”.

In its generic form, the letter f is used to name the function. We use the letter f , of course, because it stands for “function”. And as usual, we use x to represent the independent variable.

What’s nice about this notation is that it allows us great flexibility to use other letters or symbols that can be more descriptive. For example, if we wanted to describe how the height of a bouncing ball changes over time, we might want to let the variable t represent time and then name the function h for height.

So, we could write a function “ $h(t) = \text{some expression in terms of } t$ ”. That’s read “ h of t ”, which does a pretty good job of capturing the idea that we’re describing *height* in terms of *time*.

We might write a function like $d(t) = 60t$ which described distance traveled d as a function of time t . This is exactly the same equation as $y = 60x$, but we’ve changed the names of things to better match the context.

⁴ The surname Euler is German, and so it is pronounced *OIL · er* (and *not* *YOO · ler*).

Evaluation using function notation

Function notation has another convenience. We can use the notation to indicate a specific choice of value for the independent variable. Earlier, we said: “Evaluate the function $y = x^2 - 4$ when $x = 3$.” This was fine, but function notation gives us a way to shorten these instructions.

For example, if we have $f(x) = x^2 - 4$ then we might want to evaluate $f(3)$, which is said aloud “ f of three”. The x in the function notation was replaced by 3, which is exactly what it means to evaluate the function at that value! The notation is telling us to input 3 for x and find the output value. So if $f(x) = x^2 - 4$, then

$$f(3) = (3)^2 - 4 = 9 - 4 = 5.$$

We will use function notation and $y =$ notation through this course. The flexibility of function notation means that it is the preferred way of writing functions in higher level mathematics courses.

4.2 Domain and range

Machines (meaning real, mechanical devices) are physical things that have physical limitations. Number machines (meaning functions) often have similar operational restrictions.

The ins and outs of functions

// startup exploration

Consider the function machines below.

$$f(x) = |x| \qquad g(x) = \frac{1}{x}$$

What sorts of values are allowed as input to each of these machines? What sorts of values will each one produce as output?

A blender is a machine that has limited operating parameters. Anything we put into the blender comes out “blended” in some way. But, of course, there are some things we can’t put in a blender: a piano, for example.⁵ Similarly, we wouldn’t put bread into a blender, hit the button, and expect toast to pop out. That kind of output is associated with a different machine.

4.2.1 Regarding input values

An important component of the definition of any function is a specification about what sorts of numbers it can accept as input. In the language of mathematics, we call the set of all “legal” input values the *domain* of the function.

Domain

The set of all allowed input values to a function. This is the set of all allowed x -values or all possible values of the independent variable.

We have two functions in the startup exploration. The first function, $f(x) = |x|$, is very welcoming: any real number can be used as input to this number machine: whole numbers, fractions, positive numbers, negative numbers, it doesn’t matter. The other function is a little more finicky.

⁵ We’ll pause here to note that “blended” is defined rather loosely: smoothies come out “blended”, ice comes out “crushed”, chickpeas come out “puréed”, magazines come out “shredded”. . . . We consider all of these to be a kind of “blended”. Speaking of magazines, we’ll also mention that there are things which one *can* put into a blender, but *shouldn’t*. Magazines are one example. Another example: your hand.

In chapter 1, we discussed the illegality of division by zero. So, if we try to put 0 into the function $g(x) = \frac{1}{x}$, then we're going to have a problem: $g(0)$ – that's the function g evaluated at $x = 0$ – is undefined. So, our user manual for the function g has to include a note that only numbers other than 0 are allowed as input. We say that the function $g(x)$ has a *limited* or *restricted domain*.

There are other circumstances where we will want or need to limit the domain. For example, a domain can be limited by the context of a problem we are solving. If we write an equation that shows distance as a function of time, it usually makes sense to limit the domain to only include positive time-values. If we have a function which maps “number of pizza toppings” to “price of the pizza”, then there is (probably) an upper limit on the domain: there are only so many pizza toppings available.⁶

A sequence is a function whose domain is limited to the set of natural numbers (the positive integers). We saw this in chapters 2 and 3: a sequence can have a first term and a second term, but no one-and-a-halfth term. We exclude from the domain of a sequence all numbers except $1, 2, 3, 4, \dots$.

We can also define a specific domain for convenience. On an assignment, for example, we might be asked to evaluate the function $y = 4x + 5$ for the x -values $-1, 0$, and 1 . We did something like this in example 3.1. This is another case of limiting the domain.

4.2.2 Regarding output values

In addition to knowing about input values, it is often important for us to know what sorts of numbers we can expect as output from a certain function. We call the set of all output values the *range* of the function.

Range

The set of all output values from a function. This is the set of all y -values or all possible values of the dependent variable.

The function $f(x) = |x|$ can accept any number as input, but only non-negative numbers ever come out. We say that the range of f is non-negative real numbers.

In the case of $g(x) = \frac{1}{x}$, notice that the only way for a fraction to equal 0 is if the numerator is equal to zero. No matter what nonzero input value we put into the function g , we will never get 0 as the output! The range of g is nonzero real numbers.

⁶ Did you know that the so-called “Hawaiian pizza”, a pizza topped with pieces of ham and pineapple, is not actually a Hawaiian invention? Several news organizations agree that this pizza originated in Ontario, Canada.

Piecewise definition

An interesting way to control the output of a function is to give it different definitions depending on the input. For example, suppose we define a function like so:

$$f(x) = \begin{cases} x & \text{if } x \text{ is positive} \\ -x & \text{if } x \text{ is negative} \\ 0 & \text{if } x \text{ is equal to } 0 \end{cases}$$

This unusual notation gives us several definitions at once, depending on which of the “if” statements applies to a given input value x .

If the input x is positive, then the output is simply x itself. If the input is negative, then the output is the opposite of x (so again, we get a positive value). Finally, in the third case, if the input is 0, then we get 0 as the output. Do you recognize this function? This case-by-case definition is really just a complicated way of saying $f(x) = |x|$.

Here’s a more interesting example. Suppose we define a function whose domain is the non-negative integers as follows:

$$z(n) = \begin{cases} \frac{n}{2} & \text{if } n \text{ is even} \\ -\frac{n+1}{2} & \text{if } n \text{ is odd} \end{cases}$$

This function produces different output depending on whether the input value is even or odd. Let’s try the first few input values to see what happens. What happens when we input 0? Zero is an even number, so the first case applies and:

$$z(0) = \frac{0}{2} = 0.$$

One is an odd number, so the second case applies:

$$z(1) = -\frac{1+1}{2} = -\frac{2}{2} = -1$$

Two is even, and three is odd. So we have:

$$z(2) = \frac{2}{2} = 1 \quad \text{and} \quad z(3) = -\frac{3+1}{2} = -\frac{4}{2} = -2.$$

This function produces the sequence:

$$0, -1, 1, -2, 2, -3, 3, -4, 4, \dots$$

which is a list of all the integers. In other words: this machine turns the natural numbers \mathbb{N} (plus zero) into the set of integers \mathbb{Z} .⁷

⁷ This function suggests a shocking idea. Most people would say that there are more integers than there are natural numbers.

After all, since the integers include all of the natural numbers *and* all of the opposites of the natural numbers, it must be that the set of integers has *twice as many members*. . . right? Nope! The set of natural numbers and the set of integers are *exactly the same size*. Think of it this way: the function z defined above pairs up natural numbers and integers. If I give you a natural number (input), can you tell me what integer (output) it partners up with? If I give you an integer (output), can you tell me what natural number (input) it partners up with? If so, then the two sets match up exactly and neither set has any leftovers. They must therefore be the same size! Mind blown!

4.3 Three families of Algebra 1

The majority of Algebra 1 will focus on the in-depth study of three key families of functions: the linear family, the exponential family, and the quadratic family. We've encountered each of these families before, and now we will recap all that we've learned about them. Though this is our first "formal introduction", you might be surprised to discover how much you already know about the families!

4.3.1 Parent functions

For a given function to be considered a member of a certain family, it must share the characteristics of other members of that family as seen in its graph, its equation, and the patterns it exhibits in a table of values. These features are shown in their most basic form by the *parent function* of each family.

The linear parent is $y = x$ and the quadratic parent is $y = x^2$. In this course, we will consider $y = 2^x$ to be the parent function for the exponential family, though other choices are possible.⁸

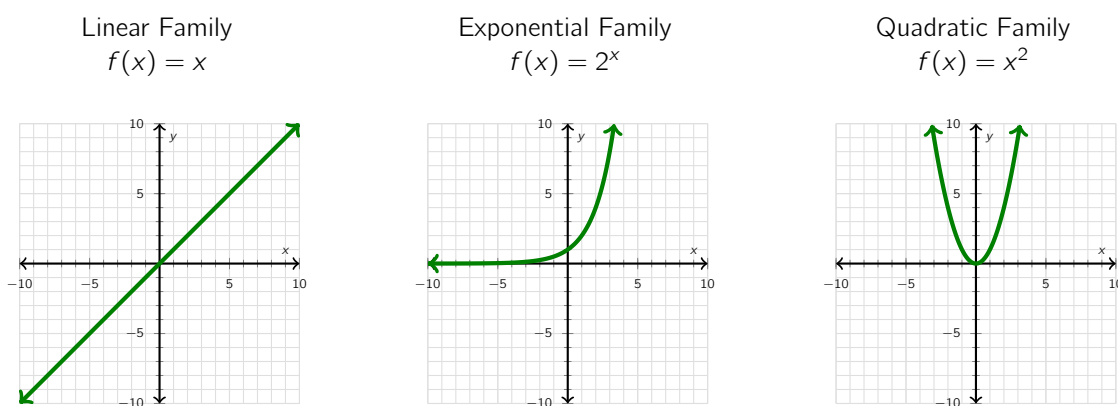


Figure 4.2: The parent functions of the three families

Compare the equations for each of the parent functions: How are they the same? How are they different? Compare the graphs: How are they the same? How are they different? When presented with a new equation of graph, what features might we look for to decide which family it belongs to?⁹

In section 4.1 we learned techniques for identifying whether a given relation is a function. One of our tasks going forward will be to learn techniques for distinguishing the function families from one another.

⁸ A more natural choice for the exponential parent is the function $y = e^x$, where e is Euler's number (yes, that's the same Euler as the function notation guy). Euler's number has many important connections to the family of exponential functions, but those will have to wait until later. By the way e , like π , is an irrational number: $e \approx 2.71828\,18284\,59045\,23536\,02875\dots$

⁹ These are not the only three families of functions out there, of course. Other families of functions will become the subject of study in Algebra 2 (and beyond).

4.3.2 Transforming the parents

As we have seen many times already, a graph can be a particularly helpful way to see, and therefore understand, a mathematical relationship. One of the skills we'll develop as we go forward in this course is the ability to connect features of the equation of a certain function to features of its graph.

Over time and with practice, we'll get better at picturing the graph of a function in our mind's eye without having to draw the graph on paper. Drawing graphs will still be informative, of course, but some features of the graph will start to “jump out at us” simply from looking at the equation of the function.

Not far from the tree

// startup exploration

Study the three sets of animations depicted in figs. 4.3 to 4.5. Each set of images shows a way in which we can change the parent equation and produce a change in its graph.

Figure 4.3 shows what happens when we multiply the parent function by a constant value. Figure 4.4 shows what happens when we add a constant value to the parent. Figure 4.5 shows what happens when we add a constant value to the argument (the x -value) before that value is used as input to the parent function.

Study the animations^a or experiment with the interactive versions at algebranomicon.org. Then write a sentence or two summarizing what is happening in each case.

^a The images will only appear animated when this PDF document is viewed using Adobe Reader.

4.3.3 Fundamentals of linear functions

The first family we will study in depth is the family of linear functions. Here we summarize the key features of this family, all of which will return in future chapters.

We will learn several different forms for the equation of a linear function. The feature they all share is that the variable x has no exponent, or rather the highest power of the variable x is 1. For example, the parent of this family, $y = x$, is the same as $y = x^1$. That's a “phantom 1” there in the exponent, and we don't usually write it.

The graph of a linear function is a straight, non-vertical line. (Can you explain why we have to exclude vertical lines from this family of functions?) We sometimes also exclude horizontal lines from this family: a function of the form “ $y = \text{a number}$ ” is a horizontal line, for example $y = 6$.¹⁰

¹⁰ In fact, horizontal lines can be considered their own family: the family of *constant functions*.

Figure 4.3: Comparing the parent function $f(x)$ with offspring $a \cdot f(x)$.

Figure 4.4: Comparing the parent function $f(x)$ with offspring $f(x) + c$.

Figure 4.5: Comparing the parent function $f(x)$ with offspring $f(x + b)$.

Recall that arithmetic sequences are in the family of linear functions. Arithmetic sequences have a rule that involves repeatedly adding on a constant difference at each step. So, this is the feature we see in the table of a linear function: when the x -values increase by a fixed amount, the y -values increase by a (possibly different) fixed amount.

There are two key transformations of the parent function. Multiplying by a constant value, as in $y = a \cdot x$ changes the “steepness” of the graph. Certain a values cause the graph to flip over the x -axis and show a trend that goes in the opposite direction (we get a graph that has been reflected over the x -axis). Adding a value as in $y = x + c$ shifts the graph upwards or downwards. Viewed from a different perspective, we could also say that this shifts the graph left or right. Up-down shifts and left-right shifts look the same for a straight line. The distinction between these two transformations is more clear in the other families.

4.3.4 Fundamentals of exponential functions

The second family of functions we will explore in this course is the exponential family. In fact, we will begin our study of this family in Algebra 1 and then return to it again in Algebra 2.

An exponential equation is one in which the variable x appears as the exponent (hence the name). The graph of an exponential function is a smooth, curving shape that we sometimes describe as “J-like”. This is only an approximation, though, since an exponential graph may sometimes look more like a backwards J or an upside-down J.

Exponentials have the interesting property that they get almost vertical on one side and almost horizontal on the other side. For example, one side of the parent function, $y = 2^x$, gets “infinitely large” while on the other side we have tinier and tinier fractions that get closer and closer to zero... but never actually disappear! Mathematically speaking, we say that the function approaches an **asymptote** at the flat horizontal portion. The parent function has a horizontal asymptote at $y = 0$; other functions in this family may have their asymptote at a different position.

Recall that geometric sequences are in the family of exponential functions. Geometric sequences have a rule that involves repeatedly multiplying by a constant ratio. This is the feature we look for in a table.

Multiplying the parent by a value, as in $y = a \cdot 2^x$, changes the steepness of the graph and (if a is negative) causes it to be reflected over the x -axis.

Adding a value to the parent function, $y = 2^x + c$, causes a shift up or down. (Can you predict where the function’s asymptote will be in this case?) If we add to the argument, $y = 2^{(x+b)}$, we get a shift left or right.

4.3.5 Fundamentals of quadratic functions

We will study the quadratic family last in this course. In fact, it will be our gateway into the study of techniques that will become helpful as we learn about other families and types of functions in Algebra 2.

The key feature of a quadratic equation is that the highest exponent on the variable x is 2. The equation may also contain an x term, but that's the same as x^1 , so the highest power of x is still 2!

The graph of a quadratic function is a smooth U-like graph called a **parabola**. Of course, under certain circumstances, the U opens downwards.

The transformations of the parent function are perhaps more clear with this family. Multiplying by a constant, $y = a \cdot x^2$, changes the “steepness” of the graph. And, in the case that a is negative, controls whether the graph opens upwards or downwards.

Adding to the parent function $y = x^2 + c$ produces a clear vertical shift, and adding to the argument $y = (x + b)^2$ produces a clear horizontal shift. Interestingly, the horizontal shifts appear to be “backwards” compared to what we might expect. Adding to the argument shifts the graph to the left, whereas subtracting from the argument (or adding a negative value) shifts the graph to the right. (Can you spot this unusual behavior in the linear and exponential families, too?)

Recall from our study of sequences that quadratic patterns show a constant second difference. In other words, they have a recursive rule that states: “Start with A , add to each previous term the members of some arithmetic sequence.” This constant second difference will be the feature that we look for in a data table.

4.3.6 Domain and range of the three families

It's quite easy to identify the domains of the three main families of Algebra 1. Linear, exponential, and quadratic functions all have “all real numbers” as their domain. We can always use any number as input. This is convenient!

The range of a linear function is also pretty simple: it is all real numbers. As we saw as we transformed linear functions, the a (non-horizontal) line will eventually stretch up (or down) as high (or as low) as you want to go on the y -axis.

The ranges for exponential and quadratic functions are a little more complicated, but we can see what is happening by examining their graphs. Notice, for example, that the graphs of the parents $y = 2^x$ and $y = x^2$ never dip below the x -axis. This feature of the exponential and quadratic families will always be retained, in some form, when we transform these equations: the range will always be limited by some minimum or maximum value.

The quadratic parent passes through the origin, and so the range of $y = x^2$ is all real numbers *greater than or equal to* zero. We might write that y is any real number such that $y \geq 0$. In fact, every parabola has a point, called its **vertex**, that is either its lowest point (if the parabola opens upwards) or its highest point (if the parabola opens downwards). We will learn more about identifying and describing the vertex of a parabola when we study the quadratic family in in depth.

When it comes to the exponential parent, we mentioned that it has an asymptote at $y = 0$. This means that the function approaches, but never reaches, that value. So, the range of $y = 2^x$ is all real numbers *strictly greater than* zero. We might write that y is any real number such that $y > 0$. We will learn more about what's going on with this asymptote when we study the exponential family.

» Chapter summary «

In this chapter we have brought together a number of different concepts. Our work with sequences, graphs, functions, and transformations has been organized under three key headings: the three families that we will study closely in Algebra 1.

With the next chapter we begin to focus on the linear family, the most logical place to start. In one sense, the linear family is the easiest family to understand, and so it makes for a good “training level”. At the same time, the linear function is an important structural concept in mathematics that gives rise to a whole branch of study called, appropriately enough, *linear algebra*.

In any case, studying the linear family will give us the chance to learn a whole collection of new algebraic tools and techniques that we will use throughout the course. Onward!

Mathematics is the art of giving the same name to different things.

Henri Poincaré, French mathematician

Chapter 5

Solving equations

One of the most important concepts in mathematics is the equation: a statement that two quantities are the same. We use equations in mathematics, of course, but equations are used in all sorts of places: in biology, chemistry, and physics laboratories. . . in the offices of business leaders, economists, and politicians. . . in hospitals, in schools, and in homes.

In this chapter we will learn the key ideas about writing and solving equations, and we will progress through the various challenges associated with linear equations.

5.1 Equivalence

Equivalence is an important concept in algebra and, indeed, throughout mathematics. The kernel of the idea is that we can manipulate mathematical objects in ways that change their appearance without changing their value.

Never the tween shall meet?

// startup exploration

Bob thinks: “Since $\frac{3}{7}$ and $\frac{4}{7}$ are ‘right next to each other’ on the number line, there must be no fractions in between them.” Use what you know about equivalent fractions to find a fraction in between $\frac{3}{7}$ and $\frac{4}{7}$.

Given any two fractions $\frac{a}{b}$ and $\frac{c}{d}$, describe a procedure for finding a fraction that lies between them on the number line. Does your procedure work for the fractions $\frac{3}{5}$ and $\frac{5}{8}$.

We’ve studied equivalence before and know, for instance, that multiplication by 1 is an operation which maintains equivalence. When we multiply some quantity by 1 – even a very fancy version of 1 – the quantity remains

unchanged. This is the idea behind equivalent fractions:

$$\frac{3}{7} = \frac{3}{7} \cdot 1 = \frac{3}{7} \cdot \frac{18}{18} = \frac{54}{126}$$

We also know the order of operations, which gives us a process for simplifying numerical expressions. The order of operations guarantees we always have exactly the same quantity we started with, even though it may look different.

As we progress through this chapter (and beyond), we will learn more tools that we can use to simplify algebraic expressions and solve equations. We will have to be careful to use the tools correctly, though, so that the maneuvers we perform are sure to maintain equivalence. The devil is in the details, so it is a good habit to pay attention and proceed carefully.

The first key piece of the equivalence puzzle is something that probably seems obvious:

Substitution

We may replace one quantity with another quantity that we know has the same value.

This is a property about numbers that we use all the time, for example when we replace $4+3$ with 7 . We do a series of substitutions when we solve a problem using the order of operations. Given $2 + 3 \cdot 6$ we first substitute 18 for $3 \cdot 6$, giving $2 + 18$. Then we substitute 20 for $2 + 18$.

5.1.1 Field axioms

An **axiom** is a statement that is accepted as true without proof. The real numbers \mathbb{R} are built upon several axioms, called the **field axioms**, which are the properties and laws that make arithmetic and algebra work.¹

Axiom: Closure properties

When we add two real numbers, their sum is a real number. Mathematically speaking, we say that *the real numbers are closed under the operation of addition*.

Similarly, when we multiply two real numbers, their product is a real number. We say that *the real numbers are closed under the operation of multiplication*.

¹ A *field* is a mathematical object that pairs up a set of numbers with some operations on those numbers. The number system we know so well – the real numbers \mathbb{R} , along with the operations of addition and multiplication – are a field. But many other fields exist! If you continue to study algebra in college, the focus gradually shifts from studying familiar fields like the real numbers, to studying the properties of fields in general.

This might seem obvious, but we saw in Chapter 1 that closure is not guaranteed for all sets and all operations. For example, the integers are not closed under the operation of division.

Axiom: Identity properties	
<p>The identity property of addition states that for any real number a,</p> $0 + a = a + 0 = a.$ <p>The number 0 is called the additive identity.</p>	<p>The identity property of multiplication states that for any real number a,</p> $1 \cdot a = a \cdot 1 = a.$ <p>The number 1 is called the multiplicative identity.</p>

As we saw above, the identity property of multiplication is the axiom that allows us to create equivalent fractions.

Axiom: Inverse properties	
<p>The inverse property of addition states that for any real number a, there exists a real number $-a$ such that</p> $a + -a = -a + a = 0.$ <p>The number $-a$ is called the additive inverse of a. Often we will call it the opposite of a.</p>	<p>The inverse property of multiplication states that for any nonzero real number a, there exists a real number $1/a$ such that</p> $a \cdot \frac{1}{a} = \frac{1}{a} \cdot a = 1.$ <p>The number $1/a$ is called the multiplicative inverse of a. Often we will call it the reciprocal of a.</p>

Note that when discussing the multiplicative inverse of a , this axiom states that a has to be nonzero. In ordinary arithmetic, division by zero is undefined, and expressions such as $5 \div 0$ and $\frac{5}{0}$ have no meaning.

Tangent: Explaining division by zero

// startup exploration

Consider the following two equations. How are they related?

$$72 \div 8 = \square \quad \text{and} \quad \square \cdot 8 = 72$$

Now consider these two equations. How are they related?

$$5 \div 0 = \square \quad \text{and} \quad \square \cdot 0 = 5$$

The first pair of equations demonstrates the relationship between multiplication and division, which tells us that the same number goes in both boxes. Whatever number makes the first equation true, must also make the second equation true.^a

This second pair of equations is related in the same way. But, there's a problem with the multiplication sentence $\square \cdot 0 = 5$. *Anything times zero is zero*. So, no number exists that could make that multiplication sentence true! Therefore, there can be no answer in the related division sentence.

At this point, troublemakers always like to ask about $0 \div 0 = \square$ because that sentence, they say, is related to the multiplication sentence $\square \cdot 0 = 0$, and *every number in the universe* makes that sentence true. They're right, of course. We say that $a \div 0$ is "undefined" when $a \neq 0$, whereas $0 \div 0$ is called an "indeterminate form". It doesn't really matter as far as we're concerned, though. Dividing *anything* by zero means that we don't get a clear answer. It's mathematically off-limits.

^a Spoiler alert: it's 9.

Axiom: Commutative properties

The **commutative property of addition** states that for real numbers a and b :

$$a + b = b + a.$$

The **commutative property of multiplication** states that for real numbers a and b :

$$a \cdot b = b \cdot a.$$

The commutative properties allow us to rearrange the order of things when we add or multiply. For instance, when adding up a string of numbers, it's sometimes helpful to "make a ten":

$$\begin{aligned} 12 + 6 + 8 & \quad \text{grouping up the 12 and the 8 would be easier than adding left-to-right} \\ = 12 + 8 + 6 & \quad \text{commutative property of addition says } 6 + 8 = 8 + 6, \text{ so we can swap them} \\ = 20 + 6 \\ = 26 \end{aligned}$$

Note that this property is for *addition* and *multiplication*, but not subtraction or division! To move things around in, say, a subtraction problem, we must use the definition of subtraction as "addition of the opposite" to create an equivalent addition problem.

$$6 - 5 \neq 5 - 6 \quad \text{but} \quad 6 + (-5) = (-5) + 6$$

Axiom: Associative properties

The **associative property of addition** states that for real numbers a , b , and c :

$$(a + b) + c = a + (b + c).$$

The **associative property of multiplication** states that for real numbers a , b , and c :

$$(a \cdot b) \cdot c = a \cdot (b \cdot c).$$

The associative properties allow us to move around certain parentheses and regroup addition and multiplication. Here, the numbers don't move; the parentheses do.

$$\begin{aligned} &5 \cdot (8 \cdot 7) && \text{order of operations would force us to do what's inside the parentheses first} \\ = &(5 \cdot 8) \cdot 7 && \text{associative property of multiplication says we can move the parentheses} \\ = &40 \cdot 7 \\ = &280 \end{aligned}$$

A few words of warning: We can only use this property around addition or multiplication. Subtraction, for example, would have to be transformed into addition of the opposite.

$$(8 - 4) + 3 \neq 8 - (4 + 3) \quad \text{but} \quad (8 + -4) + 3 = 8 + (-4 + 3)$$

Also, we can't apply the associative property when there are different operations involved:

$$6 \cdot (5 + 4) \neq (6 \cdot 5) + 4$$

Here's a sneaky one, can you explain why these two equations are not equivalent?

$$5 + 6(3 + 4) \neq (5 + 6)3 + 4$$

To handle situations like these last two, we need the final field axiom:

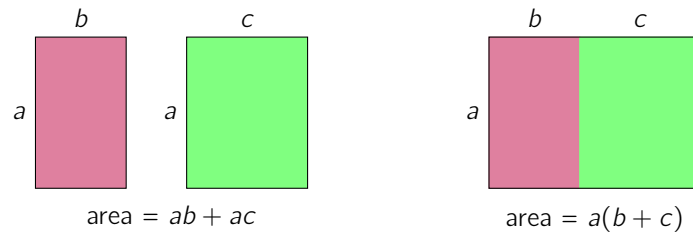
Axiom: Distributive property

The **distributive property** states that for real numbers a , b , and c :

$$a \cdot (b + c) = a \cdot b + a \cdot c.$$

This property explains how multiplication and addition are related to each other. To get a feel for this, suppose we have two rectangles: one that is a units tall and b units wide, and a second rectangle that is a units tall and c units wide. Together, their combined area is $a \cdot b + a \cdot c$.

But, since the two rectangles have the same height, we could glue them together perfectly and create one conjoined rectangle with height a and width $(b + c)$. The area of this rectangle is $a \cdot (b + c)$. But of course, the area hasn't changed, so these two different ways of expressing the area are equal.



Another way to think about the distributive property is to think of taking the a (in this case) and sprinkling it over the parentheses so that it becomes a part of each of the terms inside.

$$a(b + c) = ab + ac$$

Both the rectangle-area metaphor and the sprinkling metaphor can be extended to include expressions like

$$a(b + c + d) = ab + ac + ad,$$

and to expressions with any number of terms inside the parentheses.

5.2 Simplified algebraic expressions

The axioms are a set of tools that we can use to manipulate algebraic objects such that the result is equivalent to the original. That's great... but why would we ever need to do that?

Clarity and simplicity are important when it comes to communicating mathematical ideas. The fraction $\frac{16}{24}$ is fine, but the simplified fraction $\frac{2}{3}$ communicates the same information more clearly. For example, it's much easier to visualize a circle with $\frac{2}{3}$ shaded than it is to visualize a circle with $\frac{16}{24}$ shaded.

If a movie starts at 8 o'clock, we would all prefer to see that simple number rather than have to compute $(4 + 5) \cdot 2 - 10$ o'clock (yes, even algebra teachers feel this way). A simpler form will almost always be easier to understand.

So, as we did with rational numbers and arithmetic expressions, we will define what it means for an algebraic expression to be "simplified". Recall from our work in section 3.2 that an **algebraic expression** could be a single term, or the sum or difference of terms. A **term** could be a number, or a variable, or the product/quotient of numbers and variables.

Criteria for simplified algebraic expressions

An algebraic expression is considered completely simplified if...

1. Numerical expressions have been evaluated
2. Redundant negative signs have been rewritten
3. Terms have been arranged in order of decreasing degree, with coefficients written first
4. Grouping symbols have been eliminated (whenever possible)
5. Different variable terms appear at most once

Several of these criteria are old news. Criteria #1 just means that we have to simplify numerical expressions using the order of operations: the expression $2 \cdot 3 \cdot x$ is not simplified until we evaluate " $2 \cdot 3$ " and write "6" instead. The equivalent expression $6 \cdot x$ or $6x$ is simplified.

Criteria #2 is also something we know how to handle. In an expression like $x + -6$, where there are two signs in a row, we can simplify using the definition of subtraction: $x + -6 = x - 6$.

Criteria #3 is just a bit of cosmetics. We say that the **degree of a term** is the power to which the variable is raised in a variable term. So, if we have an expression that includes a variable raised to different powers, it's

often helpful to see the terms arranged in order from greatest degree to smallest degree. For example, instead of:

$$2x + 3x^4 + 7 - 6x^3 - 18x^2$$

it's often nicer to write:

$$3x^4 - 6x^3 - 18x^2 + 2x + 7$$

Note that when there's a variable with no exponent written, as in the $2x$ here, we think of it as having a "phantom 1" in the exponent: $2x = 2x^1$. Also note that the term 7 has no variable part at all. In this case, we picture a different kind of "phantom 1": $7 = 7 \cdot 1 = 7x^0$. If we make these substitutions, we can really see how the terms are arranged by degree:

$$3x^4 - 6x^3 - 18x^2 + 2x^1 + 7x^0$$

Also note that this criteria asks us to write the coefficients first. So, $2x$ is preferred over $x2$. This is, primarily, because it might be tricky to tell the difference between $x2$ and x^2 . It would be especially confusing in a term has both a coefficient and an exponent: x^43 looks a bit too much like x^{43} , whereas $3x^4$ is much clearer.

The only criteria left are criteria #4 and criteria #5. These are more interesting, and so each gets its own section.

5.2.1 Distributive property

To eliminate grouping symbols, as expressed by criteria #4, we can rely on the field axioms for help.² The expression $(x + 2) + 2$ is not simplified, but this can be fixed easily enough using the associative property of addition.

$$(x + 2) + 2 = x + (2 + 2) = x + 4$$

The expression $3(x + 4)$ is not simplified either, and in this case the distributive property comes to the rescue. Get sprinkling!

$$3(x + 4) = 3 \cdot x + 3 \cdot 4 = 3x + 12$$

There are subtleties to using the distributive property. Distribution and negative numbers can lead to some easy-to-miss mistakes. We have to be on the lookout, so study the following examples carefully.

² Some special grouping symbols, like the vinculum, can remain in the final expression. Generally, however, we must get rid of grouping symbols that do not double as another operation.

Example 5.1

Simplify: (a) $-3(2x - 4)$ and (b) $8 - 5(6x - 9)$

Solution: In problem (a) we distribute a negative number. Notice what happens with the signs of the terms in the result.

$$\begin{aligned}
 & -3(2x - 4) \\
 &= -3(2x + -4) && \text{change subtraction to addition of the opposite} \\
 &= -3 \cdot 2x + -3 \cdot -4 && \text{distributive property} \\
 &= (-3 \cdot 2)x + -3 \cdot -4 && \text{commutative property of multiplication} \\
 &= -6x + 12 && \text{substitution (in both terms)}
 \end{aligned}$$

The first thing we did here was to change the subtraction to addition of the opposite. This was helpful because – look! – we end up with *positive 12* in the final expression. That might otherwise have been an easy thing to miss.

In problem (b) notice that we have an implied operation that will require us to apply the distributive property. In other words, the first step is *not* to do $8 - 5$! As with the last problem, we'll start by changing to all-addition.

$$\begin{aligned}
 & 8 - 5(6x - 9) \\
 &= 8 + -5(6x + -9) && \text{change subtraction to addition of the opposite} \\
 &= 8 + -5 \cdot 6x + -5 \cdot -9 && \text{distributive property} \\
 &= 8 + -30x + 45 && \text{substitution} \\
 &= -30x + 8 + 45 && \text{commutative property of addition} \\
 &= -30x + 53 && \text{substitution}
 \end{aligned}$$

Again here, we get a term $-5 \cdot -9 = 45$. Changing to all-addition before distributing is a helpful technique for getting these signs right.

Example 5.2

Simplify: $2x - (x - 4)$

Solution: This one is subtle and sneaky! The negative sign is stuck on the parentheses, which means we'll have to "distribute the negative sign" or – more accurately – we again have an implied operation

hiding in there as a “phantom one”. This expression is the same as $2x - 1(x - 4)$.

Observe:

$2x - (x - 4)$	
$= 2x - 1(x - 4)$	rewrite to expose the “phantom one”
$= 2x + -1(x + -4)$	change subtraction to addition of the opposite
$= 2x + -1x + -1 \cdot -4$	distributive property
$= 1x + 4$	substitution
$= x + 4$	rewrite to hide the “phantom one”

Did you try this problem on your own before reading the solution? Did you get $x - 4$? If so, don’t be too upset: you’re in good company. This is one of the most common mistakes in Algebra 1. Always be on the lookout when you see the distributive property mixed up with subtraction and negative signs.

5.2.2 Combining like terms

Expression sort

// startup exploration

Take a moment to sort the following mathematical objects into different groups, creating whatever categories make the most sense to you. After sorting, describe how you made your decisions. What features define each of the groups you have created?

$17x$	$2xy$	$-4x^2$	8	$0.5xy^2$	$-y$	$-15z$
-3	$8x^2y$	$4y$	xy^2	xyz	$11z$	x
$4xy^2$	$-2x^2y$	$3xyz$	x^2	$-6y$	$4x$	$3x^2$

One way to sort the terms above is to group **like terms** together. Like terms have the same variable factors raised to the same powers. For example, $17x$ and $4x$ are like terms. Also, x^2 and $-4x^2$ are like terms. But, $17x$ and $11z$ are *not* like terms since they have different variable parts. Though it may seem confusing at first, x and x^2 are *not* like terms. They both have x ’s, but those x ’s are raised to different powers.

Explaining the startup exploration

The terms in the startup exploration can be grouped using many different schemes. The categorization below is based on the concept of grouping “like terms” together.

- Terms with no variable: -3 and 8
- Terms with x as the variable: x , $4x$, and $17x$
- Terms with x^2 as the variable: x^2 , $3x^2$, and $-4x^2$
- Terms with y as the variable: $-y$, $4y$, and $-6y$
- Terms with z as the variable: $-15z$ and $11z$
- Terms with xy as the variable: $2xy$
- Terms with x^2y as the variable: $-2x^2y$ and $8x^2y$
- Terms with xy^2 as the variable: $0.5xy^2$, xy^2 , and $4xy^2$
- Terms with xyz as the variable: xyz and $3xyz$

Just as 3 goats and 2 goats combine to make 5 goats, so too do $3x$ and $2x$ combine to make $5x$. Similarly, $8x^2 + x^2 = 9x^2$ (in this case, x^2 really means $1x^2$: there's a phantom 1 lurking there as the coefficient).

The expression $3xy^2 + 3x^2y + 3x^2y^2$ is simplified, since the variables and their corresponding powers do not match (look closely!).

The process of uniting like terms under a single coefficient is called **combining like terms**.³ This is not technically a property in itself, but a “shortcut” for a somewhat longer chain of events:

$$\begin{array}{ll}
 3x + 2x = x(3 + 2) & \text{the distributive property, in reverse} \\
 = x(5) & \text{substitution} \\
 = 5x & \text{commutative property of multiplication}
 \end{array}$$

At first, it's best to commute the terms and group them up. Then we can add up the coefficients of the like terms. We'll see this approach, and a shortcut, in the examples below.

³ Also sometimes called *collecting like terms*, or *gathering like terms*, and sometimes abbreviated *CLT*.

Example 5.4

Simplify: $3x^2 + 2x - 2 + x^3 + 4x^2 - 3x - 2x^3 + 8$

Solution: To make sure we keep track of the signs, we'll convert to all-addition and then move the terms around to put like terms next to each other.

$$\begin{aligned}
 & 3x^2 + 2x - 2 + x^3 + 4x^2 - 3x - 2x^3 + 8 \\
 = & 3x^2 + 2x + -2 + x^3 + 4x^2 + -3x + -2x^3 + 8 && \text{definition of subtraction} \\
 = & x^3 + -2x^3 + 3x^2 + 4x^2 + 2x + -3x + -2 + 8 && \text{commutative property of addition} \\
 = & -1x^3 + 7x^2 + -x + 6 && \text{combine like terms} \\
 = & x^3 + 7x^2 - x + 6 && \text{simplify signs}
 \end{aligned}$$

Example 5.5

Simplify $3x - 4y + 2 - 5x + 7y - 6 + 8y$

Solution: There is no commutative property of subtraction, remember, so we can rewrite all subtraction into adding the opposite, as in the previous example. A shortcut is to remember that *the sign goes with the term* when you rearrange. Really, this is why we rewrite subtraction in the first place! If we're careful, though, we can avoid the rewriting step.

If we look at the given expression, we see that there are three "types" of terms that need combining: terms with x , terms with y and constant terms (numbers).

Let's pick a type of term (let's choose ones with x) and group them up – signs and all! In the given expression, we have $3x - 5x$, so we have $-2x$ in all.

Pick another type of term and repeat as necessary! If we look at the y terms, we have $-4y + 7y + 8y$ so that's $11y$ in all. Looking at the constant terms, we have $2 - 6 = -4$.

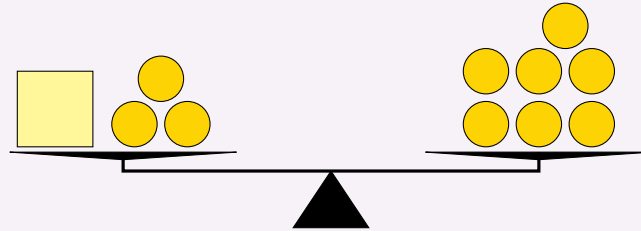
Putting it together, we have $-2x + 11y + -4$ or (after simplifying signs) $-2x + 11y - 4$.

5.3 Properties of equality (POEs)

Cheese scale

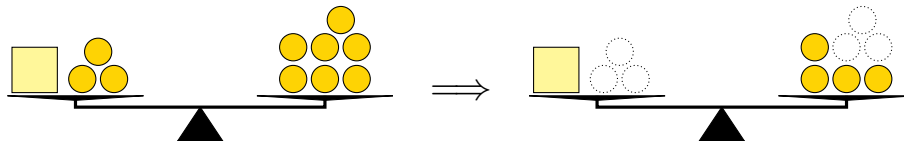
// startup exploration

Bob put a block of edam cheese and three wheels of gouda on one side of a two-pan balance. On the other pan he put seven wheels of gouda. He discovered that the scale was in perfect balance.



If all of the wheels of gouda are identical in weight, how many wheels will balance the block of edam?

If Bob removes a wheel of gouda from the right-hand pan, it will tip the scales. . . unless he also removes an equivalent amount of gouda from the left-hand pan. In particular, the scale will still be in balance after Bob removes three wheels from each pan of the balance.⁴



Taking three wheels from each side of the balance is helpful because it means that we will have gotten the block of edam alone on one side. From there it's easy to see that the block weighs the same as four wheels.

The idea of maintaining balance is the heart of the algebra of solving equations.

Equation

An **equation** is a number sentence which states that two algebraic expressions are equal.

⁴ Edam and gouda are traditional Dutch cheeses. Every summer the city of Alkmaar, in the Netherlands, hosts an open-air cheese market. Cheese merchants demonstrate how cheese was bought and sold in medieval times. It operates on Waagplein, which means "weighing square" in Dutch, because once the buyer and seller agreed on a price, the huge wheels of cheese were taken to the scales by official "cheese carriers", and then weighed very carefully to determine their value. The cheese carriers' guild was established in 1593.

To create an algebraic version of the balance problem, we need a way to represent the weight of the block of edam. In other words, we need a way to write down the weight of the block *before we know what the weight actually is*.

Unknown

An **unknown** is a quantity whose value is not known. We usually represent an unknown using a letter.

Both variables and unknowns are represented by letters though, technically speaking, an unknown is different from a variable. A variable's value can vary or change, whereas an unknown has a fixed value (or set of values) that must be determined.

Let's use B to stand for the unknown weight of the block, and let's assume that each wheel weighs 1 unit. Then, we can translate our balance problem into algebraic symbols:

$$B + 3 = 7.$$

If we take three (wheels) from each side of the balance, we have

$$B = 4$$

which tells us that B , the weight of the block, is four units. Our key task here was transforming the first equation into the second equation.

In section 5.1 we used the field axioms to simplify individual algebraic expressions. An equation is a math sentence with an expression on each side of the equal sign. So, we need some rules that will allow us to manipulate two expressions at the same time, but in such a way that their equality remains intact.

5.3.1 Maintaining equality

An equation states that the two pans of a balance (the expressions on either side of the equal sign) are in alignment. If we were to, say, add five to the expression on one side of the equation, then we would tip the scales out of balance. . . unless we also add five to the expression on other side of the equation.

This idea, and others like it, are captured mathematically in a set of rules called the *properties of equality*, or *POEs* for short.

Properties of equality: Addition and subtraction

The **addition property of equality** (APOE) states that for all real numbers a , b , and c

$$\text{if } a = b, \text{ then } a + c = b + c.$$

The **subtraction property of equality** (SPOE) states that for all real numbers a , b , and c

$$\text{if } a = b, \text{ then } a - c = b - c.$$

Note that since “subtraction” is the same as “addition of the opposite”, SPOE is really just APOE in disguise.

Properties of equality: Multiplication and division

The **multiplication property of equality** (MPOE) states that for all real numbers a , b , and c , where $c \neq 0$

$$\text{if } a = b, \text{ then } a \cdot c = b \cdot c.$$

The **division property of equality** (DPOE) states that for all real numbers a , b , and c , where $c \neq 0$

$$\text{if } a = b, \text{ then } \frac{a}{c} = \frac{b}{c}.$$

Here again, note that since “division” is the same as “multiplication by the reciprocal”, DPOE is really just MPOE in disguise. Note too that c must not equal 0 when using these rules. Using 0 with DPOE must certainly be avoided, lest we be flummoxed upon division by zero.⁵

Let’s use the properties of equality to solve the problem of Bob’s cheese balance.

Example 5.6

Bob put a block of edam cheese and three wheels of gouda on one side of a two-pan balance. On the other pan he put seven wheels of gouda. He discovered that the scale was in perfect balance. If all of the wheels of gouda are identical in weight, how many wheels will balance the block of edam?

Solution: Let B represent the weight of the block of edam. Then, the given information suggests the

⁵ Why must we be cautious about multiplying both sides of an equation by zero? In the end we’ll get $0 = 0$, which is a true statement. What’s the problem? Note that multiplication by 0 is *irreversible*. If we multiply by 2 (or any nonzero number) we can reverse or “undo” this operation by dividing by 2. But, since division by 0 is not allowed, we cannot undo multiplication by 0. Later, we will see that irreversible operations can lead us to “fake” solutions to our equations, so-called *extraneous solutions*.

equation $B + 3 = 7$. We want to isolate B .

$$B + 3 = 7$$

$$B + 3 - 3 = 7 - 3 \quad \text{SPOE: subtract 3 from both sides of the equation}$$

$$B = 4$$

Note that in our solution above, we started out by defining what variable we would use to represent our unknown. Then, we wrote an equation based on the information given in the problem.

Our goal was to *isolate* B or to get B by itself on one side of the equation. Since 3 had been added to B , we had to “undo” this. So, we subtracted 3 from both sides using SPOE. We showed our work going down the page, one step at a time. The individual equations were aligned at the equal sign for clarity.

Inverse operations (a.k.a. opposite operations)

Operations that will “undo” one another. For example, addition and subtraction are inverse operations. Multiplication and division are inverse operations.

When we solve an equation, we use inverse operations to undo the order of operations. Once we get a hang of the rules, solving equations can become a game. Plus, like the best video games, equation-solving has different levels of challenge so that the game stays interesting even as you get better and better at it.

5.4 Level 1 and 2 linear equations

5.4.1 Level 1 equations

Level 1 linear equations are the simplest kind to solve because it takes only one step to isolate the variable. The example we saw in the last section was a Level 1 linear equation.

Example 5.7

Consider each of the equations below. What step should we perform to isolate the variable?

$$x \div 4 = 21$$

$$w - 9 = 50$$

$$y + 13 = -24$$

$$8x = 72$$

Partial solution: In the first equation we must multiply both sides of the equation by 4 (MPOE).

$$x \div 4 = 21$$

$$\frac{x}{4} = 21 \quad \text{definition of division}$$

$$\frac{x}{4} \cdot 4 = 21 \cdot 4 \quad \text{MPOE: multiply both sides of the equation by 4}$$

$$x = 84$$

In the second equation we should add 9 to both sides (APOE). In the third equation, we should subtract 13 (SPOE). Alternatively, we can think of this as APOE in which we add -13 to both sides.

In the last example, we should divide both sides of the equation by 8 (DPOE). Or, we can think of this as MPOE in which we multiply both sides by $\frac{1}{8}$.

5.4.2 Level 2 equations

Here's an example of a level 2 equation. How is it different from Level 1?

$$3x + 16 = 43$$

Notice that in this equation, two things were done to the unknown x . First it was multiplied by 3, and then 16 was added to the result. To isolate the variable, the rule of thumb is to undo whatever was done to the unknown. Usually this means that we will have to undo the order of operations in reverse. Most of the time, we can rely on the “underpants analogy”.

Underpants analogy

When you get dressed for school in the morning, you put your underpants on *before* you put on your jeans. When you get undressed to go to bed, you do the *opposite* of each step, and you do the steps in the *opposite order*. In other words: you have to take your jeans off first, before you can take your underpants off.^a

^a Try doing it the other way around, we dare you. Later.

To undo the order of operations, we have to work upside-down and backwards:

$$3x + 16 = 43$$

$$3x + 16 - 16 = 43 - 16$$

$$3x = 27$$

$$\frac{3x}{3} = \frac{27}{3}$$

$$x = 9$$

SPOE: subtract 16 from both sides of the equation

simplify and substitute

DPOE: divide both sides of the equation by 3

simplify and substitute

In the end we have isolated x , and so this final equation tells us that 9 is a solution to the original equation.

Solution and solution set

A **solution** to an equation is any value that will make the equation true.

Sometimes an equation will have more than one solution. The set of all solutions to an equation are called a **solution set**. To write a solution set, we use set notation and write $S = \{\text{list of solutions}\}$.

As we go through the process of solving an equation, each application of a Property of Equality produces an equation that is equivalent to the original.

Equivalent equations

Equations that have the same solution set.

5.4.3 Showing and checking work

As with problems that involved the order of operations, it is often best to show your work going down the page as we have done in the earlier sections. It is also very helpful to anyone reading your work if you write beside

each step the property that you used to justify turning one equation into an equivalent equation.

A note on vocabulary: we *simplify an expression* and we *solve an equation*. Even though the work we show might be very similar, the goal of our steps is quite different.

Having found a solution, it is a great habit to check the solution by substituting it back into the original equation. In the example above, we had the equation $3x + 16 = 43$ and found that this was equivalent to the equation $x = 9$. To check our solution, we plug 9 back in for x in the original equation:

$3x + 16 = 43$	original equation
$3(9) + 16 \stackrel{?}{=} 43$	substitute in 9 for x
$27 + 16 \stackrel{?}{=} 43$	carry out the order of operations
$43 \checkmark = 43$	check!

When we substitute in 9, we put little question marks over our equal sign, since we are checking the solution – we’re not sure whether the two expressions are equal yet! In the end, when we discover that the left-hand side is equal to the right-hand side, we replace our question mark with a check mark of confirmation!

The key here is to realize that, ultimately, the process we use to solve an equation yields *candidates* for the value of the unknown. We are not guaranteed that all of these candidates actually satisfy the original equation, until we check them to know for sure.⁶

Once we have verified that our solution satisfies the original equation, we should write our answer in set notation. The equation $x = 5$ isn’t our solution, but an equivalent equation – in fact, it’s the simplest equation that has the same solution as the original!

In other words, the equation solving process creates simpler and simpler equations that all have the same solution. Eventually, we get to the an equation that shows the solution clearly. We use set notation to record our solution because it is a notation that suggests the solution works for all of the intermediate equations.

Plus, solution set notation works nicely with all types of equations. Soon we will have equations with multiple solutions (even infinitely many!) and the $x =$ notation starts to become quite awkward.

So, in our ongoing example, we write $\mathcal{S} = \{5\}$. Yes, a mathematical “set” may contain just a single element!⁷ The alternative to using solution set notation is to write the sentence “The solution to the equation is 5.”

⁶ There will come a time when we will do everything correctly and, still, some of our solution-candidates will have to be rejected! Get in the habit of checking solutions now!

⁷ This is another situation in which we find a conflict between everyday language and the language of mathematics. Lots of words have a special meaning when used in a mathematical context: similar, odd, mean, product... What other examples can you think of?

5.5 Level 3 linear equations: Multiple steps

Level 1 equations are called, naturally enough, *one-step equations*, and Level 2 equations are called – you guessed it! – two-step equations. In Level 3, we increase complexity by combining the tasks of expression-simplifying and equation-solving.

Linear level 3

// startup exploration

Determine the value x , given the equation

$$2x + 7 - 5x + 12 = 15.$$

Here we find an equation that must be simplified using the field axioms and related tools. Let's start out by trying to simplify the left-hand side.

$$2x + 7 - 5x + 12 = 15$$

$$2x + 7 + -5x + 12 = 15$$

rewrite as addition, by def'n of subtraction

$$2x + -5x + 7 + 12 = 15$$

commutative property of addition

$$-3x + 19 = 15$$

combine like terms – this is now a Level 2 equation!

$$-3x + 19 - 19 = 15 - 19$$

SPOE

$$-3x = -4$$

$$\frac{-3x}{-3} = \frac{-4}{-3}$$

DPOE

$$x = \frac{4}{3}$$

simplify

Notice that after a few lines, we had turned the Level 3 equation into a Level 2 equation. We know how to solve those! So, our goal will be to try to turn Level 3 equations into Level 1 or Level 2 equations. But, since there are more steps in the process, there are more chances we can make a mistake. So, it's a good habit to remember to check our answer at the end.

We can check our answer for the last problem by substituting our solution back into the original equation and

simplifying using the order of operations.

$2x + 7 - 5x + 12 = 15$	original equation
$2\left(\frac{4}{3}\right) + 7 - 5\left(\frac{4}{3}\right) + 12 \stackrel{?}{=} 15$	substitute in our solution $x = \frac{4}{3}$
$2\left(\frac{4}{3}\right) + 7 + -5\left(\frac{4}{3}\right) + 12 \stackrel{?}{=} 15$	rewrite as addition, by def'n of subtraction
$\frac{8}{3} + 7 + \frac{-20}{3} + 12 \stackrel{?}{=} 15$	multiply fractions
$\frac{8}{3} + \frac{21}{3} + \frac{-20}{3} + \frac{36}{3} \stackrel{?}{=} 15$	rewrite left-hand side with common denominator
$\frac{45}{3} \stackrel{?}{=} 15$	add fractions
$15 \stackrel{\checkmark}{=} 15$	boom!

Writing our solution as a solution set, we have $\mathcal{S} = \{15\}$.

Example 5.8

Determine the value of w , given the equation

$$7w + 2(-3w + 1) = 12.$$

Solution: Our goal is to try and simplify the left-hand side so that it looks like a Level 1 or Level 2 equation.

$7w + 2(-3w + 1) = 12$	
$7w + -6w + 2 = 12$	distributive property
$1w + 2 = 12$	combine like terms – that's a Level 1 equation!
$w + 2 - 2 = 12 - 2$	SPOE
$w = 10$	

Let's check our solution:

$7w + 2(-3w + 1) = 12$	original equation
$7(10) + 2(-3(10) + 1) \stackrel{?}{=} 12$	substitute in our solution $w = 10$
$70 + 2(-30 + 1) \stackrel{?}{=} 12$	carry out the order of operations on the left-hand side
$70 + 2(-29) \stackrel{?}{=} 12$	
$70 + -58 \stackrel{?}{=} 12$	
$12 \checkmark = 12$	

We're going to look at one more example and discuss two alternative ways to approach it.

Example 5.9

Determine the value of g , given the equation

$$-4(2g - 7) = 36.$$

Solution: Approach #1. Those parentheses are just asking to be simplified using the distributive property. Be mindful of the signs, though!

$-4(2g - 7) = 36$	
$-8g + 28 = 36$	distributive property – that's a Level 2 equation!
$-8g + 28 - 28 = 36 - 28$	SPOE
$-8g = 8$	
$\frac{-8g}{-8} = \frac{8}{-8}$	DPOE
$g = -1$	

Approach #2. Let's solve the same equation again. This time, notice that we can divide both sides by

-4 as the first step, which eliminates the need for the distributive property.

$$-4(2g - 7) = 36$$

$$\frac{-4(2g - 7)}{-4} = \frac{36}{-4} \quad \text{DPOE}$$

$$2g - 7 = -9 \quad \text{simplify – that's a Level 2 equation!}$$

$$2g - 7 + 7 = -9 + 7 \quad \text{APOE}$$

$$2g = -2$$

$$\frac{2g}{2} = \frac{-2}{2} \quad \text{DPOE}$$

$$g = -1$$

The two different approaches taken in the previous example are not always available. Compare the solutions to examples 5.8 and 5.9. Can we avoid the distributive property in example 5.8? Why or why not?

5.5.1 More on checking work

Every step we take in solving an equation generates an equivalent equation that has the same solution set. If we make a mistake, we accidentally create an equation with a different solution set. Knowing this can help us check our answers on the more complex equations (ones that have more steps).

Example 5.10

Solve for x , given the equation $2x - 3(4x - 1) = -53$.

Let's say we mess up the signs when carrying out the distributive property – a very common mistake.

$$2x - 3(4x - 1) = -53 \quad \text{Line 1.}$$

$$2x - 12x - 3 = -53 \quad \text{Line 2. (That's a mistake!)}$$

$$-10x - 3 = -53 \quad \text{Line 3.}$$

$$-10x = -50 \quad \text{Line 4.}$$

$$x = 5 \quad \text{Line 5.}$$

Now, we go back to check our work by substituting the solution into the original equation. The answer doesn't work in line 1, so we know it is not the correct solution.

$$2(5) - 3(4(5) - 4) = -53 \implies -38 \neq -53 \quad (\text{Something went wrong!})$$

But, our answer *does* work in line 2 (and also lines 3, 4, and 5).

$$2(5) - 12(5) - 3 = -53 \implies -53 = -53 \quad (\text{That works!})$$

This is the clue helps us pinpoint the location of the mistake. It tells us the mistake must have happened when transforming line 1 into line 2.

If, on a different problem, we find that our solution doesn't work for steps 1, 2, or 3, but *does work* in step 4, then it means that our mistake must have happened when transforming step 3 into step 4!

5.6 Level 4 linear equations: Variables on both sides

In Level 4, we add a wrinkle which can lead to some very unusual results.

Linear level 4

// startup exploration

Determine the value of x given the equation

$$13x - 5x - 5 = x + 7 + x.$$

The goal is the same: to isolate the variable on one side of the equal sign. We begin by simplifying each side.

$$13x - 5x - 5 = x + 7 + x$$

$$8x - 5 = 2x + 7$$

combine like terms on both sides

$$8x - 5 - 2x = 2x + 7 - 2x$$

SPOE: subtract $2x$ from both sides – clever!

$$6x - 5 = 7$$

combine like terms again – that's a Level 2 equation!

$$6x = 12$$

APOE: add 5 to both sides

$$x = 2$$

DPOE: divide both sides by 6

In this example, we subtracted $2x$ from both sides of the equation. This might seem like a tricky move, but it's a completely legal application of SPOE. We are allowed to do the same thing to both sides of the balanced equation, even if that means subtracting an unknown amount from both sides.⁸

Let's see what happens if we isolate the unknown on *the other side of the equal sign*. (It had better give us the same answer!)

$$13x - 5x - 5 = x + 7 + x$$

same equation as before

$$8x - 5 = 2x + 7$$

combine like terms on both sides, as before

$$8x - 5 - 8x = 2x + 7 - 8x$$

SPOE: subtract $8x$ from both sides

$$-5 = -6x + 7$$

combine like terms

$$-12 = -6x$$

SPOE: subtract 7 from both sides

$$2 = x$$

DPOE: divide both sides by -6

⁸ Of course, you have to subtract the same unknown amount from both sides. It's no fair to subtract x from one side of an equation and y from the other side. This will put the equation out of balance unless we know that these two unknown amounts are equal... that is, unless we know that $x = y$. Caution: can you think of a situation in which it might be dangerous to apply a POE using an unknown? Think about what might happen if we use DPOE to divide both sides by an unknown value x .

So, it doesn't matter which side we choose to eliminate the unknown. As long as we apply APOE or SPOE correctly, the solution will be the same. One strategy is to choose the side that will result in a *positive coefficient* for the variable. This isn't required, but it helps avoid the chances of losing a negative sign along the way. (Maybe that's happened to you!)⁹

Plus, if we don't like the side of the equation that the unknown is on, we can always rewrite the equation. For example if we have the equation $3 = 4 + x$ and we really want the x on the left-hand side, we can just rewrite the equation as $4 + x = 3$. Those are equivalent equations!¹⁰

Example 5.11

Determine the value of x given the equation

$$3(x - 2) + 3x = 4x - 6.$$

Solution: Let's jump right in and start simplifying the left-hand side.

$3(x - 2) + 3x = 4x - 6$	
$3x - 6 + 3x = 4x - 6$	distributive property
$6x - 6 = 4x - 6$	combine like terms
$2x - 6 = -6$	SPOE: subtract $4x$ from both sides
$2x = 0$	APOE: add 6 to both sides
$x = 0$	DPOE: divide both sides by 2

Yes, we can get 0 as the solution to an equation. We can always check to be sure:

$3(x - 2) + 3x = 4x - 6$	original equation
$3(0 - 2) + 3(0) \stackrel{?}{=} 4(0) - 6$	substitute in our solution $x = 0$
$3(-2) + 0 \stackrel{?}{=} -6$	simplify
$-6 \stackrel{\checkmark}{=} -6$	

⁹ Napoleon ordered the creation of the first modern lost-and-found office in Paris in 1805, although lost-property systems have existed in Japan since the early 700s.

¹⁰ Even these obvious properties have been given names. For example, this is a fact about equations: if $a = b$, then $b = a$. To describe this, we say that equality is *symmetric*. An even more obvious statement is that $a = a$ (we say that equality is *reflexive*). A third property is a little more interesting: if $a = b$ and $b = c$, then $a = c$ (we say that equality is *transitive*).

5.6.1 Special cases

There are some things to be careful about with when it comes to solving Level 4 linear equations. Here is an example of something strange that can happen when we have unknowns on both sides of the equal sign.

Example 5.12

Determine the value of x given the equation

$$12x + 36 = 2(6x - 10).$$

Solution: We proceed as we have done in earlier problems:

$$12x + 36 = 2(6x - 10)$$

$$12x + 36 = 12x - 20 \quad \text{distributive property}$$

$$12x + 36 - 12x = 12x - 20 - 12x \quad \text{SPOE: subtract } 12x \text{ from both sides}$$

$$36 = -20 \quad \text{huh?}$$

In this case the variable term disappears completely from both sides of the equation. And then, what is left is obviously *not equal*. In a sense, the original “equation” is not an equation at all. The two given expressions are not equal, and never will be.

This means that there is no value of x what will ever work to make the equation true. We say that this equation *has no solution*.

As unsatisfying as it might sound, it is possible to have an equation that does not have a solution.¹¹

How do you write down the solution to a problem that has no solution? This is not quite as philosophical as it sounds. We can, of course, write the words “no solution”. Or, we could write a solution set that contains no numbers, $\mathcal{S} = \{ \}$.

Or, there is a third option. You may or may not be surprised to learn that mathematicians have invented a symbol for just such an occasion. We use the mathematical symbol \emptyset , called the “empty set” or “null set”, as a way to show that that our solution set is empty.¹² We write $\mathcal{S} = \emptyset$ to mean “the solution set is empty”.

¹¹ Equations without solutions are very important in the history of mathematics. The French mathematician Pierre de Fermat made a conjecture in 1637 stating that a certain equation had no integer solutions. His claim, which came to be known as Fermat’s Last Theorem, wasn’t officially proven to be correct for more than 350 years. British mathematician Andrew Wiles published the first complete proof of Fermat’s Last Theorem in 1995. The theorem states that the equation $a^n + b^n = c^n$ has no integer solutions for a , b , and c when the exponent n is a natural number greater than 2.

¹² The symbol \emptyset , which was introduced by French mathematician André Weil in 1939, is a letter in the Norwegian alphabet.

Notice that we don't write the curly braces around the empty set. In other words, we write $\mathcal{S} = \emptyset$ and not $\mathcal{S} = \{\emptyset\}$. The former says " \mathcal{S} is the empty set", which is what we mean when we have an equation with no solutions. The latter says " \mathcal{S} is the set which contains the empty set" which – believe it or not – is not the same thing.¹³

For the record, the empty set is not the same as zero. So, an equation can have the solution set $\mathcal{S} = \{0\}$, like example 5.11. This is different from an equation having the solution set $\mathcal{S} = \{ \}$, like example 5.12.

If you look back at the reason why the last example has no solution, you may be wondering about whether there is another possibility. Study the following example.

Example 5.13

Determine the value of x given the equation

$$5x - 3(x + 2) = 2x - 6.$$

Solution: Let's go for it:

$$5x - 3(x + 2) = 2x - 6$$

$$5x - 3x - 6 = 2x - 6$$

distributive property, watch the signs

$$2x - 6 = 2x - 6$$

combine like terms – something is fishy already. . .

$$2x - 6 - 2x = 2x - 6 - 2x$$

SPOE: subtract $2x$ from both sides

$$-6 = -6$$

but, of course!

In this case the variable term disappears again, and the leftovers are *obviously equal*. This is the opposite of what happened earlier. If that example had no solutions, this one has all of them. We can replace x with literally any real number and the equation will be true. That means that we have a solution set that contains every real number there is.

We say that this equation's solutions are *all real numbers*.

In this example, our solution set is the set of real numbers. We can write this out in words, "all real numbers", or use the shorthand symbol that stands for the real numbers and write, $\mathcal{S} = \mathbb{R}$. Notice no curly braces are used here either.

¹³ The empty set is not the same as zero, and it is not the same as nothing. It is a set with nothing inside it, like an empty bag. So, writing $\mathcal{S} = \emptyset$ is stating that the solution set is like an empty bag. Extending this metaphor, $\{\emptyset\}$ is a bag with an empty bag in it and, therefore, a set which contains the empty set is not empty. Om.

Checking answers in special cases

In the case that we find “all real numbers” as the solution to an equation, we know that every number is going to work. So to check, we might pick a couple of different numbers (easy ones, like 0 and 1) and substitute those into the original equation. Any number that you choose should satisfy the equation.

If we find that an equation has “no solutions”, then any number we use to check will fail to satisfy the equation. Not very informative. In this case, it may be best just to check back over our work to make sure that we did each of the steps correctly.

5.7 Level 5 linear equations: Absolute value

Over the last few sections we have added complexity to the equation-solving picture. We have learned properties that can be used to “undo” whatever has been “done” to the variable. Sometimes this has led to situations where the equation has no solutions or infinitely many solutions. We turn now to equations that involve the absolute value of an unknown.¹⁴

Linear level 5

// startup exploration

Recall the definition of absolute value. Determine the value of x given the equation

$$|x| = 6.$$

Recall that the absolute value of a number is that number’s distance away from zero on the number line. Distances are always positive: -6 is 6 units away from 0, and so $|-6| = 6$. Of course, $|6| = 6$ as well. So, the equation above has two solutions $x = 6$ or -6 .¹⁵ In solution set notation, we write $\mathcal{S} = \{6, -6\}$.

The presence of the absolute value gives us the possibility of having two solutions, since whatever is inside the absolute value bars could be either the positive or the negative version of the value.

Here’s an extension to the idea. Determine the value of x given the equation

$$|2x| = 16$$

This is telling us that the number “ $2x$ ” is 16 units away from zero on the number line. In other words

$$2x = 16 \text{ or } -16.$$

This is really two equations written at once: $2x = 16$ and $2x = -16$. We can apply DPOE and divide everything by 2, which will isolate x . That is, we have

$$x = 8 \text{ or } -8.$$

The following two examples carry this idea a bit further.

¹⁴ We’re discussing absolute value equations here in the chapter on linear equations, though we put “linear” in quotes. Absolute value equations of the kind that we’re discussing here are linear-like enough that they fit here. Other types of absolute value equations are possible, though we won’t get into the full range of variations in this course.

¹⁵ Note, we mean *or* and not *and*. Both 6 and -6 are solutions to the equation, but we must write $x = 6$ or -6 , since x can only take on one value at a time.

Example 5.14

Determine the value of x given the equation

$$|-3x + 9| = 12.$$

Solution: The stuff inside the absolute value bars can equal 12 or -12 . So, this means, we have two equations:

$-3x + 9 = 12$	<u>or</u>	$-3x + 9 = -12$	
$-3x = 3$		$-3x = -21$	SPOE: subtract 9 throughout
$x = -1$		$x = 7$	DPOE: divide by -3 throughout

This equation has solutions $\mathcal{S} = \{-1, 7\}$.

Note that the steps we took for solving were the same for both equations (in this case, first we used SPOE, and then DPOE). So, we can streamline our work a bit, as shown in the next example.

The next example also has things going on outside of the absolute value symbols. For these types of problems, we want to isolate the absolute value expression first (using the properties of equality), then separate into two equations (if necessary).¹⁶

Example 5.15

Determine the value of x given the equation: $3|4 - 4x| = 48$.

Solution: We use the POEs at first, and then split into two equations when necessary.

$3 4 - 4x = 48$	
$ 4 - 4x = 16$	DPOE, to isolate the absolute value expression
$4 - 4x = 16$ <u>or</u> -16	definition of absolute value
$-4x = 12$ <u>or</u> -20	SPOE: subtract 4 throughout
$x = -3$ <u>or</u> 5	DPOE: divide by -4 throughout

¹⁶ In case you are wondering, we won't get into equations that include multiple absolute value expressions. . . although you might enjoy exploring an equation of this kind, just for fun. For example, what values of x satisfy this equation? $|x + 2| = 2|x| - 1$

So, this equation has solutions $\mathcal{S} = \{-3, 5\}$.

5.7.1 Special considerations

A few sneaky things can pop up when it comes to absolute value equations. Recall that the only number that has absolute value zero is zero itself: $|0| = 0$. So, not all absolute value equations have *two* solutions. Consider

$$|3x + 51| = 0.$$

We can't split this into two different equations. All we have is

$$3x + 51 = 0,$$

and that's a Level 2 linear equation with just a single solution, $\mathcal{S} = \{-17\}$. To demonstrate another tricky aspect, consider the equation

$$|4x| + 16 = 5.$$

No problem! We begin by subtracting 16 from both sides:

$$|4x| = -11$$

But, the absolute value of a number can never be negative. So, this equation has no solution. In other words, we have $\mathcal{S} = \emptyset$.

Don't be too quick with the "no solutions" talk, though. If we encounter the equation

$$-3|x + 2| = -12,$$

we may see the -12 and assume there is no solution. But recall that we must isolate the absolute value expression first. To do that we will divide both sides by -3 and then we will have

$$|x + 2| = 4,$$

an equivalent equation that surely has a solution (two solutions, in fact). So, don't jump to conclusions!

Warning!

Some folks like to use the notation "plus/minus" in situations that involve absolute value. For example, writing " ± 6 " with the little stacked-up plus and minus signs to stand for "6 or -6 ". We recommend avoiding \pm notation when solving equations.

Here's why. See if you can spot the error in the following work.

$$|x - 5| = 12$$

$$x - 5 = \pm 12$$

$$x = \pm 17 \quad \text{APOE: add 5 to both sides}$$

And so, $\mathcal{S} = \{17, -17\}$. This work looks reasonable, but let's check the two solutions.

$$\begin{aligned} |17 - 5| &\stackrel{?}{=} 12 \implies |12| \stackrel{?}{=} 12 \implies 12 \checkmark = 12 \\ |-17 - 5| &\stackrel{?}{=} 12 \implies |-22| \stackrel{?}{=} 12 \implies 22 \neq 12 \end{aligned}$$

So, using the \pm notation has led us to get one correct solution and one incorrect solution for the equation! This is why we recommend either writing out the word “or” explicitly or breaking into two separate equations.

(By the way, the correct solution is $\mathcal{S} = \{17, -7\}$. Try working out the problem again using a more reliable technique. Can you see exactly where the work using \pm went astray?)

5.8 Applications of equation solving

Now that we know how to handle different types of equations, let's explore a few applications. First, we'll discuss writing (and then solving) equations based on problem situations. Then, we'll see how to apply the properties of equality to manipulate formulas from science and other disciplines. We'll close this chapter with an extension section about explaining some fundamental ideas using the field axioms.

5.8.1 Writing equations to solve word problems

There are lots of ways to approach problems like this: educated guessing-and-checking, wishful thinking, making an organized list, and so on. One way is to model the situation with (and then solve) an equation. Converting a word problem into an equation is an important skill for your mathematical toolbox.

Trick or treat

// startup exploration

On Halloween night, Hildegard and Ingvar (Bob and Yearleigh's parents) had 200 candies in their trick-or-treat bowl. They gave 5 candies to each kid who came to their door, and at the end of the night they had 10 candies left. How many kids came trick-or-treating?

The key in word problems like this is first to identify the unknown in the problem. We choose a variable to stand in for the unknown and write an equation based on the other information given in the problem.

In our example, the unknown is “how many kids came trick-or-treating”, so let's use k to represent the number of kids. Hildegard and Ingvar gave 5 candies to every kid, so they gave away $5k$ candies. They started with 200 and gave away $5k$. Giving away suggests subtraction, so they were left with $200 - 5k$ candies at the end of the night. The problem tells us that they had 10 candies left over, so:

$$200 - 5k = 10.$$

That's a Level 2 equation! We'll leave it for you to solve and put the solution in the footnote.¹⁷

Here's a second example. “Consecutive integer problems” are a clever type of mystery number problem that might at first appear impossible to solve. There doesn't seem to be enough information! For example: The sum of three consecutive integers is 39. What are the numbers?

What does it mean to have three consecutive integers? Consecutive means “in an unbroken sequence”, so three consecutive integers are three integers in a row, such as 4, 5, 6.

¹⁷ Solving this equation using the POEs leads to the result $k = 38$. So, the answer to the question is that 38 kids came trick-or-treating at the Krumbli house on Halloween night.

Let N represent the smallest of the three consecutive integers. Then, the next integer is $N + 1$, and the integer after that is $N + 2$. The problem says their sum is 39. So we have:

$$\begin{aligned} N + (N + 1) + (N + 2) &= 39 \\ 3N + 3 &= 39 && \text{combine like terms} \\ 3N &= 36 && \text{SPOE} \\ N &= 12 && \text{DPOE} \end{aligned}$$

At this point, many students would be tempted to draw a box around their answer and call it a day. After all, we solved the equation, right? But, look back at the question: it asks us to find the *numbers*, plural. Our answer to the question should list all three of the numbers!

What did we actually compute when we solved the equation? We chose N to represent the smallest of the three consecutive integers, and that's 12. So, the full answer is that the three integers are 12, 13, and 14. (A quick check shows that these three do, in fact, add up to 39.)

The moral of the story here is always to check that we are answering the question that has been asked. This step is a helpful way to avoid giving an irrelevant or incomplete solution.

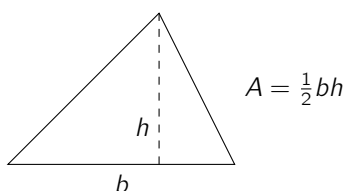
Finally, when it comes to stating the solution to a word problem, we don't usually use set notation. A handy rule of thumb is to think about how we'd say the answer out loud to someone. When we solve an equation with no context, we'd probably say: "The solution is 4." Set notation is appropriate, since the answer is just a number, $S = \{4\}$.

On the other hand, if we were answering a word problem that asked how many pounds of cream cheese Bob ate, we'd say: "Bob ate 4 pounds of cream cheese." When the context of the problem implies a unit on the answer (pounds, dollars, goats, miles per hour), we write the answer and the unit rather than use solution set notation: 4 pounds of cream cheese.

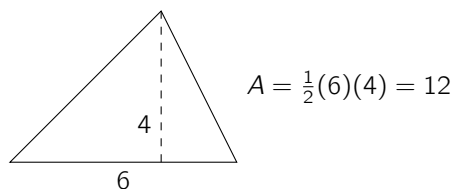
5.8.2 Transforming formulas

In all of the examples in this chapter so far, we have solved equations and found the value of an unknown. In this section, we'll discuss a different use of the field axioms and the properties of equality.

In the past, you may have learned a formula for the area of a triangle. If we have a triangle with base of length b and height h , then its area A has a tidy little formula.



If we know the base and height of a specific triangle, we can find its area.



But what if we are given the area and the base length: Could we work backwards to find the height? Could we “transform” the area formula so that it computes h instead of A ?

Transforming a formula means to isolate a specific variable, even though the other side of the equal sign might not simplify down to a single number. The only difference between this and what we have been doing before now is that the equations may have multiple variables (and not unknowns).

Example 5.16

Transform the triangle area formula, $A = \frac{1}{2}bh$, to isolate h .

Solution: We'll start with the given formula, and then apply the properties of equality to get h by itself on one side of the equal sign.

$$A = \frac{1}{2}bh$$

$$2 \cdot A = 2 \cdot \frac{1}{2}bh \quad \text{MPOE, to eliminate the fraction}$$

$$2A = bh \quad \text{simplify on the righthand side}$$

$$\frac{2A}{b} = \frac{bh}{b} \quad \text{DPOE, to isolate } h$$

$$\frac{2A}{b} = h \quad \text{simplify on the righthand side: mission accomplished!}$$

So, our transformed formula expresses the height of a triangle in terms of its area and the length of a base:

$$h = \frac{2A}{b}$$

This is a skill used in science (especially physics) and higher levels of mathematics. Generally, we have to study the equation we're given and find the variable we are trying to isolate. Then, think about what has been done to that variable which must be undone. We undo these using the POEs.

Even though our answers won't be just a number, they should be simplified as much as possible. For instance

we must avoid fractions-in-fractions. Parentheses, too, can lead to tricky situations. It often pays to formulate a plan before rushing to work.

Example 5.17

Transform the equation $P = 2(w + d)$ to isolate w .

Solution: Sometimes we have to distribute, other times it is easier to avoid it all together. Compare the two options below.

Option 1: Avoiding the distributive property

$$\begin{aligned} P &= 2(w + d) \\ \frac{P}{2} &= w + d && \text{DPOE} \\ \frac{P}{2} - d &= w && \text{SPOE} \end{aligned}$$

Option 2: Distributive property first

$$\begin{aligned} P &= 2(w + d) \\ P &= 2w + 2d && \text{distributive property} \\ P - 2d &= 2w && \text{SPOE} \\ \frac{P - 2d}{2} &= w && \text{DPOE} \\ \frac{P}{2} - \frac{2d}{2} &= w && \text{undo fraction subtraction} \\ \frac{P}{2} - d &= w && \text{write second fraction in lowest terms} \end{aligned}$$

In the previous example, the second approach is a bit more work. The distributive property can't always be avoided though, as we saw back in example 5.8. So, it pays to be observant and do a bit of thinking before we start crunching the numbers.

5.8.3 (;,;) Proving foundational results using the field axioms

As we mentioned in section 5.1, an axiom is a statement that is accepted without proof. In some sense, we “take it for granted” that adding 0 to any number doesn’t change the number: $a + 0 = a$ for any real number a .

Similarly, there's a field axiom about multiplication by 1. But, that there is no axiom about multiplication by 0. Why not? When thinking about the field axioms, an interesting thing to consider is why some rules are omitted.¹⁸

Surprisingly, we can use the axioms given in section 5.1 to explain other basic results of arithmetic. In other words: our axioms are the fundamental building blocks. We can use them to build complex structures (like Level 5 equations) or simple structures. Sometimes, we discover that we can build really simple – even obvious-looking – mathematical statements from *even simpler statements*!

Multiplication by zero

We know that $a \cdot 0 = 0$, but there's no axiom stating this. That's because we can explain multiplication by zero using the axioms that we already have. Consider this argument, where we state the property that we used to transform each line:

$0 + 0 = 0$	identity property of addition
$a \cdot (0 + 0) = a \cdot 0$	MPOE: multiply both sides by a
$a \cdot 0 + a \cdot 0 = a \cdot 0$	distributive property on left-hand side
$a \cdot 0 = 0$	SPOE: subtract $a \cdot 0$ from both sides

In the end, we used only the given axioms and the properties of equality to demonstrate that any number times 0 is zero. We don't need an axiom for this rule, because we can *build it* from the given set of axioms!

Multiplication and negative numbers: Property #1

We are given an axiom stating that $1 \cdot a = a$ (that's the identity property of multiplication). But, we have no axiom stating that $-1 \cdot a = -a$. It turns out that we don't need one:

$1 + -1 = 0$	inverse property of addition
$a \cdot (1 + -1) = a \cdot 0$	MPOE: multiply both sides by a
$a \cdot 1 + a \cdot -1 = a \cdot 0$	distributive property on left-hand side
$a + a \cdot -1 = 0$	simplifications: $a \cdot 1 = a$ and $a \cdot 0 = 0$
$a \cdot -1 = -a$	SPOE: subtract a from both sides

We could apply the commutative property of multiplication on the left-hand side, if we want, to have the new rule $-1 \cdot a = -a$.

¹⁸ Remember that the Cthulhu icon (⌘;) indicates that this is an extension section. Don't feel bad if some of the ideas feel like swimming in the deep end! Stick with it!

Multiplication and negative numbers: Property #2

Another fundamental result that seems obvious is that if we take the “double negative” of a number, we get back to the original number. In other words, the opposite of the opposite of a is a itself: $-(-a) = a$.

Consider this: we know that $a + -a = 0$ (that’s the inverse property of addition). We know that $-a$ also has an additive inverse: $-a + -(-a) = 0$. Since these are both equal to 0, they are equal to each other. So:

$$\begin{array}{ll} a + -a = -a + -(-a) & \text{both equal 0, so they equal each other} \\ a = -(-a) & \text{SPOE: subtract } -a \text{ from both sides!} \end{array}$$

Multiplication and negative numbers: Property #3

We can use the axioms to explain why the product of a negative number and a positive number is negative. Let a and b be positive numbers:

$$\begin{array}{ll} a \cdot 0 = 0 & \text{multiplication by zero} \\ a \cdot (b + -b) = 0 & \text{rewrite 0 using the additive inverse property} \\ a \cdot b + a \cdot -b = 0 & \text{distributive property} \\ a \cdot -b = -(a \cdot b) & \text{SPOE: subtract } a \cdot b \text{ from both sides} \end{array}$$

So, if we have “the opposite of a ” times b , the result is the opposite of the product of a and b .

Multiplication and negative numbers: Property #4

We can use property #3 to prove that the product of two negative numbers is positive. Let a and b be positive numbers:

$$\begin{array}{ll} -a \cdot 0 = 0 & \text{multiplication by zero} \\ -a \cdot (b + -b) = 0 & \text{rewrite 0 using the additive inverse property} \\ -a \cdot b + -a \cdot -b = 0 & \text{distributive property} \\ -(a \cdot b) + -a \cdot -b = 0 & \text{property \#3} \\ -a \cdot -b = a \cdot b & \text{APOE: add } a \cdot b \text{ to both sides!} \end{array}$$

So, the two products are the same! The product of two negative numbers is the same as the product of their opposites (that is, their positive partners).

Mind blown!

We've been taking the properties of equality for granted, but *even the properties of equality* can be derived from the field axioms. The POEs are the properties that allow us to make a cancellation, for example taking

$$x + 4 = 10$$

and thinking about it as

$$x + 4 = 6 + 4.$$

We can cancel the 4 from both sides and get $x = 6$ (we call this SPOE, the subtraction property of equality). Suppose we have real numbers a , b , and c , and we know that $a + c = b + c$. Then, consider this chain of reasoning:

$a = a + 0$	additive identity property
$= a + (c + -c)$	rewrite 0 using the additive inverse property
$= (a + c) + -c$	associative property of addition
$= (b + c) + -c$	substitution: we assumed that $a + c = b + c$
$= b + (c + -c)$	associative property of addition
$= b + 0$	additive inverse property
$= b$	additive identity property

What does this tell us? Notice that we started our equation with a and ended up with b . This means $a = b$. Altogether, we have shown that if we know $a + c = b + c$, then it must be that $a = b$. That's one of our POEs!

In the end, we have been able to use the field axioms to prove that other fundamental properties hold true. Don't worry if it feels overwhelming at first. This section has gotten into some pretty abstract territory! Let these ideas sink in and read through this section again later. It takes time for our brains to incorporate challenging ideas.

» Chapter summary «

In this chapter we learned the super-important *field axioms* and *properties of equality*. We used the axioms to simplify algebraic expressions, and we used the POEs in a number of creative ways to solve the increasing challenges associated with linear equations. In future chapters we will take on other kinds of expressions and equations.

But let's not get ahead of ourselves! We've only just dipped our toes into the linear pool! There's lots more to learn about the family of linear functions. We will continue our study of this family by bringing our new equation-solving skills to a familiar context: solving proportions. Onward!

Give me the place to stand, and I shall move
the earth.

*Archimedes, Ancient Greek mathematician and
philosopher*

Chapter 6

Proportional reasoning

A key concept in pre-algebra mathematics is proportionality. The ideas of ratio, proportion, and percent are probably familiar to you, and you likely have already solved proportions in your mathematical career. In this chapter we will apply the tools and techniques we learned in chapter 5 to some of these familiar contexts. Then we will take proportionality to the “algebra level” by exploring the concepts of direct and inverse variation.

6.1 Proportions as equations

Mathematicians, like Yeardleigh’s team of gourmet chefs, won’t use tools, techniques, or ingredients unless they know exactly where they come from. This is the attitude we adopt, for the most part, in algebra. This point of view, however, may upset some students’ mathematical status quo.

Case in point: solving proportions. Many of us have been taught a mysterious method called “cross multiplication” as a means of solving a proportion. Unless we can explain it’s inner workings – Where does it come from? Why does it give us the correct answer? – then the technique must be considered off-limits.

We’ll explain so-called cross-multiplication in this chapter and explore alternative (easier) methods of handling proportions.

Hamster and superhamster

// startup exploration

Genetic engineers at YeardleighCorp have genetically engineered superhamsters that weigh 15 ounces. Tests indicate that one superhamster can carry a 40-ounce packet of food on its back. If a human could carry the same amount as a superhamster, relative to body mass, how many pounds could a 120-pound teenager carry?

We have mentioned that the study of mathematical relationships is a central idea of algebra. One helpful skill for studying relationships is the ability to make sound mathematical comparisons.

Ratio

A comparison of two quantities, often expressed as a fraction.

There are two main types of ratios: part-to-part ratios and part-to-whole ratios. For example, “number of girls in class to number of boys in class” is a part-to-part ratio, whereas “number of girls in class to number of students in class” is a part-to-whole ratio.

Proportion

An equation stating that two ratios are equal.

Proportions are just equations, and so we don’t need any special techniques like “cross-multiplication” to solve them. We can use the trusty properties of equality, just as we do with any other equation.

Example 6.1

Solve for x : $\frac{x}{24} = \frac{44}{96}$

Solution: Our suggestion is not to think of the left side as a fraction at all. Think of it as x divided by 24. That’s just a Level 1 linear equation! We need to get rid of the “divide by 24” and so we multiply both sides by 24.

$$\begin{aligned}\frac{x}{24} &= \frac{44}{96} \\ 24 \cdot \frac{x}{24} &= 24 \cdot \frac{44}{96} && \text{MPOE: multiply both sides by 24} \\ \cancel{24} \cdot \frac{x}{\cancel{24}} &= \cancel{24} \cdot \frac{44}{\cancel{24} \cdot 4} && \text{simplify} \\ x &= 11\end{aligned}$$

After a bit of fraction-simplifying, we have isolated x with just one application of MPOE.

We call this approach “clear the denominator”. Of course, it’s not the only way to solve a proportion, and not always the easiest way.¹

6.1.1 Adding a bit more complexity

In their most basic form, proportions have the unknown in the numerator of one of the ratios. These are just Level 1 (one-step) equations. All we need to do to solve them is multiply both sides of the equation by the denominator of the unknown (that’s MPOE in action).

But what if the unknown is in the denominator of a fraction?

Example 6.2

The startup exploration problem says that a 15-ounce superhamster can carry a 40-ounce packet of food on its back. At that rate of strength, how many pounds could a 120-pound teenager carry?

Solution: Let’s write a proportion comparing the ratio of weight to carrying capacity, and let us use P to represent the amount that the teenage can carry, in pounds. Then we have:

$$\frac{15 \text{ ounces}}{40 \text{ ounces}} = \frac{120 \text{ pounds}}{P \text{ pounds}}$$

We wrote these ratios as “weight/carrying capacity”, but of course we could also have written the ratios as “carrying capacity/weight”. So, one solution approach is to take the reciprocal of both sides. That will put the unknown conveniently in the numerator!

$$\frac{40 \text{ ounces}}{15 \text{ ounces}} = \frac{P \text{ pounds}}{120 \text{ pounds}}$$

$$120 \cdot \frac{40}{15} = 120 \cdot \frac{P}{120}$$

$$320 = P$$

So, a 120-pound teenage of superhamster strength could carry 320 pounds of food. For comparison, according to the US census, the average American ate approximately 257 pounds of fruit in 2009 (including fresh fruit, dried fruit, and fruit juice). So this teenager could carry 25% more fruit than they would typically eat in a year.

¹ You may have learned other approaches for solving proportions, and that’s a good thing! One very handy approach is to scale up (or down) one (or both) of the ratios so that they have a common numerator or denominator: another application of the identity property of multiplication!

Very often in mathematical comparisons, it is helpful to create a comparison “per unit” of some quantity: miles *per hour*, dollars *per pound*, and so on. These “per unit” comparisons mean “per *one* unit”. For example, if Bob buys 3 pounds of Swiss cheese for \$19.50, then we can find an equivalent ratio

$$\frac{19.50}{3} = \frac{6.50}{1}$$

and see that the cheese costs “6.50 dollars per (one) pound”.

Unit rate

A **unit rate** is a ratio in which the value of one of the quantities is 1. For example “miles per hour”, meaning “miles traveled in *one* hour”.

Before we move on, the maneuver we made in the last example — taking the reciprocal of both sides — bears another moment of reflection.

Tangent: Explaining reciprocal of both sides

Is “take the reciprocal of both sides” a property of equality? In other words, if we know that $a = b$, do we know for sure that $\frac{1}{a} = \frac{1}{b}$ is always true?

The fact that a and b end up in the denominator of a fraction means that we must require that a and b are non-zero at the start. If a and b are both non-zero, then we can use DPOE with both numbers.

$a = b$ assume that this is true, and that a and b are non-zero

$\frac{a}{a} = \frac{b}{a}$ DPOE: divide both sides by a

$1 = \frac{b}{a}$ simplify

$\frac{1}{b} = \frac{b}{a \cdot b}$ DPOE: divide both sides by b

$\frac{1}{b} = \frac{\cancel{b}}{a \cdot \cancel{b}}$ cancel the common factor

$\frac{1}{b} = \frac{1}{a}$ voilà!

So, in the end, if two numbers (both nonzero) are equal, then we know that their reciprocals are equal as well. “Take the reciprocal of both sides” is a property of equality.

Of course, we might be faced with numerators and denominators that are more complicated. Here's an example with variables on both sides of the equation.

Example 6.3

Solve for x : $\frac{x+1}{2} = \frac{x+2}{5}$

Solution: Our approach is to use MPOE to clear *both of the denominators*.

$$\begin{aligned} \frac{x+1}{2} &= \frac{x+2}{5} \\ 2 \cdot 5 \cdot \left(\frac{x+1}{2} \right) &= 2 \cdot 5 \cdot \left(\frac{x+2}{5} \right) && \text{MPOE two times!} \\ 5(x+1) &= 2(x+2) \\ 5x+5 &= 2x+4 && \text{distributive property} \\ 3x &= 1 && \text{SPOE two times!} \\ x &= \frac{1}{3} && \text{DPOE} \end{aligned}$$

To write our answer in set notation, we have $\mathcal{S} = \left\{ \frac{1}{3} \right\}$.

The previous example may give you some insight into how cross-multiplication works!

Warning!

Solve: $\frac{x+2}{5} = \frac{7}{6}$

There may be a temptation to subtract 2 from both sides in the first step.

$$\frac{x+2-2}{5} = \frac{7-2}{6} \quad \Rightarrow \quad \frac{x}{5} = \frac{5}{6}$$

But, to do so would be **Evil and Wrong!** The $x+2$ is grouped together by the fraction bar (vinculum) and that little group is divided by 5. Before we can manipulate the group, we have to undo the division by 5. Then, the 2 is free to be subtracted.

Note that this is the same kind of structure as

$$4(x+2) = 16$$

Here we have to divide both sides by 4, before we can use SPOE on the $x + 2$. For some reason the temptation to subtract first is stronger when in fraction form (maybe because the grouping isn't obvious without the parentheses). In any case: beware!

6.2 Applications of proportional reasoning

Proportional reasoning encompasses a key set of skills that are important to have in your problem-solving toolbox. In this section we'll discuss a few of the classic ways that ratio and proportion show up in problem-solving contexts.

6.2.1 Percent

Spam “less” sodium

// startup exploration

Bob is counting cans of Spam in his pantry. For every 3 cans of “low-sodium” Spam, he has 5 cans of Spam with the usual (higher) amount of sodium. What percent of the Spam in his pantry is low-sodium?

Break the word “percent” into its component parts, “per cent”, and think about what each part means. The word “cent” indicates 100 (there are 100 cents in a dollar and 100 years in a *century*), while the word “per” indicates that we are making a comparison (earning \$10 *per* hour at a job means you earn \$10 for every hour that you work).

Putting the words back together, percent means “per 100” or “for every 100”. When we write a percent as 75%, or say out loud “seventy-five percent”, we really mean “75 per 100” or “75 out of 100”. We could write the decimal 0.75, or the fraction $\frac{75}{100}$.

The fraction $\frac{75}{100}$ isn't in lowest terms, so we could write instead:

$$\frac{3}{4} = \frac{75}{100}$$

Proportion is the heart of percent, and percent problems are all really about the percent proportion:

$$\frac{\text{part}}{\text{whole}} = \frac{\text{percent}}{100}$$

So given a percent question, we just have to ask ourselves which pieces of the percent proportion we know and which we are trying to find.

Example 6.4

Write proportions that could be used to solve each of the following:

1. What number is 10% of 20?
2. What percent of 75 is 50?
3. 15 is 25% of what number?

Partial solution:

$$1. \quad \frac{x}{20} = \frac{10}{100}$$

We're given the percent (10%) and the whole (20). The word "of" is one hint that 20 is the whole. We're looking for the percent, so that's where we'll write the unknown (x , or whatever variable you like). Solve the proportion for the unknown, and we'll have our answer.

$$2. \quad \frac{50}{75} = \frac{x}{100}$$

Here we are asked for the percent. The other information has to be untangled a bit, but we can see that 75 is the whole and 50 is the part.

$$3. \quad \frac{15}{x} = \frac{25}{100}$$

We're given the percent again, only this time the question asks "of what number", indicating that we're seeking the whole. The unknown ends up in the denominator, but that's no problem if we remember to reciprocalize!

In Bob's case, he has 3 cans of "low-sodium" Spam for every 5 cans of "regular sodium". That's a part-to-part comparison. To write that as a part-to-whole ratio, note that 3 cans are "low-sodium" for every 8 cans that he counts. We setup the percent proportion to find the percent of cans that were "low-sodium":

$$\frac{3 \text{ less sodium cans}}{8 \text{ cans}} = \frac{p \text{ "low-sodium" cans}}{100 \text{ cans}}$$

Which means 37.5% of the cans in Bob's pantry are "low-sodium".

6.2.2 Probability

Questions about chance and probability are very often just proportional reasoning questions in fancy clothes. We won't get into much detail about probability in this course except to hit some of the highlights.

1d12

// startup exploration

Suppose we were to roll a standard 12-sided die (with faces showing the numbers 1–12) 1000 times. How many times would we expect to roll a prime number?

There are 12 numbers on the die, and these numbers comprise the **sample space** of the experiment:

$$\{1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12\}.$$

Five of the numbers in the sample space are prime numbers $\{2, 3, 5, 7, 11\}$. So, the probability of rolling a prime number on a single toss of the die is $\frac{5}{12}$.

Sample space

The sample space of a probability experiment is the set of all possible outcomes of the experiment.

Generally speaking, the probability of a certain desired outcome occurring in an experiment is the ratio

$$\frac{\text{number of ways the desired outcome can occur}}{\text{number of outcomes in the sample space}}$$

This probability $\frac{5}{12}$ tells us that if we threw the die 12 times and wrote down the numbers, we'd expect to see approximately 5 prime numbers among our data. To predict how many prime numbers we might see when throwing the die 1000 times, we set up a proportion. Let p be the number of primes we expect in 1000 trials.

$$\frac{5 \text{ primes}}{12 \text{ trials}} = \frac{p \text{ primes}}{1000 \text{ trials}}$$

Solving for p , we see that $p \approx 417$. So, we predict 417-ish primes in 1000 trials.²

2d6

// startup exploration

Suppose we were to roll *two* standard 6-sided dice and add the two numbers. If we perform this experiment 1000 times, how many times would we expect to roll a prime number?

Determining the sample space for this experiment requires a bit of thinking. It might be tempting to simply list all of the possible sums — the lowest is 2, the highest is 24 — but these sums are not all equally likely.

² Of course, nothing is guaranteed when it comes to probability experiments. It could be that we get 1000 prime numbers! It's not very likely that this could happen, in fact it's practically impossible. In smaller experiments — throwing the die 50 times, say — it's quite likely that we'd get many more (or many fewer) than the 20 or so primes we would expect.

There's only one way to roll a 2, but there are multiple ways to roll a 9. A clever strategy here is to organize the outcomes in a table like the one shown in table 6.1.

	1	2	3	4	5	6
1	2	3	4	5	6	7
2	3	4	5	6	7	8
3	4	5	6	7	8	9
4	5	6	7	8	9	10
5	6	7	8	9	10	11
6	7	8	9	10	11	12

Table 6.1: Sample space for rolling two 6-sided dice.

From the table we can see that there are, for instance, four ways to roll a 9. With a bit of counting in the table, we can see that there are 15 cells that contain a prime number (out of 36 cells total). So, the probability of rolling a prime number in this experiment is $\frac{15}{36}$. We set up and solve our proportion:

$$\frac{15}{36} = \frac{p}{1000} \implies p \approx 417.$$

In 1000 trials, then, we would expect to see approximately 417 prime numbers when rolling two 6-sided dice and finding their sum.³

The examples above discuss *theoretical probability*. We were working in a “perfect world” scenario and basing our predictions and conclusions on properties of the dice. The alternative would have been to conduct an experiment and then use the data from that experiment to predict the outcome of future experiments. We call that *experimental probability*, and an example from that department will close out this section.

6.2.3 Experimental data

In an ongoing project about goat safety, cryptozoologists at YeardleighCorp Labs have proposed a study of the hunting behavior of the chupacabra.⁴ The “scientists” plan to release a chupacabra into the hills around YeardleighCorp Labs and then study the impact on the local wild goat population.

Before the study can begin, the scientists need to estimate the existing population of goats. It's clearly impossible to actually count all of the wild goats, so to make an estimate the biologists propose to use the *capture-recapture* (or *capture-mark-recapture*) process.

³ Isn't it interesting that the probability of rolling a prime number on a single 12-sided die is the same as rolling a prime number on a pair of 6-sided dice? Is this a coincidence, or is there a mathematical connection between these two probabilities that suggests why they are the same?

⁴ Cryptozoology (which comes from the Greek “crypto”, meaning “hidden”, and “zoology”, the study of animals) is a pseudoscience (the prefix “pseudo” means “false”) based on the study of animals that have not been proven to exist. The “animals” in question are called “cryptids”. Famous cryptids include the Loch Ness monster, bigfoot, the yeti, and the chupacabra. The chupacabra, Spanish for “goat sucker” is, allegedly, a creature that attacks and drinks the blood of livestock, especially goats.

First, they will capture a sample of goats and mark them with ear tags. Then these tagged goats will be released and allowed to mingle with all the untagged goats. The biologists will wait a few weeks for the goats to mix thoroughly, then capture a second sample of goats.

The assumption is that the proportion of “tagged goats to total goats *in the second sample*” is the same as the proportion of “tagged goats to total goats *in the study area*”. (A second assumption is that the population remains constant during this process.) Solving the resulting proportion will give an estimate for the number of goats in the study area.

Example 6.5

Suppose this team of cryptozoologists initially captures and tags 25 goats, then releases them. Several weeks later, they capture a second sample of 100 goats, 6 of which have tags. What is a reasonable estimate for the size of the goat population?

Solution: Our assumption is that the ratio of *tagged goats* to *all goats* represented by the second sample is equivalent to the ratio for the whole population. Originally, 25 goats were tagged. Then the second sample of 100 contained 6 tagged goats. So, we can estimate the total population as follows:

$$\frac{6 \text{ tagged in second sample}}{100 \text{ total in second sample}} = \frac{25 \text{ tagged in population}}{g \text{ total in population}}$$

Solving the proportion:

$$\frac{6}{100} = \frac{25}{g}$$

$$\frac{100}{6} = \frac{g}{25}$$

take reciprocal of both sides

$$\frac{100}{6} \cdot 25 = \frac{g}{25} \cdot 25 \quad \text{MPOE}$$

$$417 \approx g$$

Based on the data, a reasonable estimate is that there are approximately 417 goats in the study area. Recall that the symbol \approx means “approximately equal to”.

There is a great deal more that could be said about probability, some of which you have probably seen before. There are other ways to represent a sample space (like a tree diagram), there are other kinds of experiments in which certain events influence other events (you may recall that it sometimes matters whether or not you “put the marble back into the bag”). . . but for now, we just wish to mention the aspects of probability include a taste of proportional reasoning.

We now turn to another application of proportional reasoning, and our first in-depth look at a certain kind of linear function.

6.3 Direct variation

Rupee exchange

// startup exploration

In her search for delicious new flavors, Yearleigh travels to India to learn about traditional cooking techniques and flavor combinations. In preparing for her visit, she looks up the currency exchange rate and learns that 1 US dollar (USD) is equal to approximately 60 Indian rupees (INR).

Make a table and a graph showing how many rupees Yearleigh will get in exchange for 100, 200, 300, 400, 500 USD.

How many rupees will Yearleigh get for exchanging 250 USD? How much is 22 500 INR worth in USD?

6.3.1 Variation and direct variation

When scientists begin to understand a phenomenon, like a law of physics, it's often quite helpful to understand it in terms of a relationship: how one quantity changes with respect to another.

For example: The greater the mass of an object, the greater the gravitational force acting on that object.⁵ We say that the force varies directly with mass. There is an equation that describes this relationship, but the key idea is that the force acting on an object due to gravity is directly proportional to the object's mass.

Direct variation

The variables x and y are said to be **directly proportional** if their values have a constant ratio. In other words, $\frac{y}{x} = k$, where k is a constant called the **constant of variation**.

An equation of the form $y = kx$ is called a **direct variation**. The quantities x and y are directly proportional (can you see how this equation is related to the equation in the previous paragraph?).

Since we have a proportional relationship, we can solve direct variation problems using our knowledge of proportions. We will also see that the graphs, tables, and equations for direct variations share special characteristics.

⁵ When the masses of the two objects are very different (for example, if we are comparing your mass to the mass of the Earth), the larger object's mass dominates the interaction. The equation describing this relationship is part of Isaac Newton's "Law of Universal Gravitation". It states that any two bodies in the universe attract each other with a force that is directly proportional to the product of their masses and inversely proportional to the square of the distance between them.

This also means that all of the proportion problems we solved earlier can be solved by looking at them as linear functions. If we want to estimate how many goats live in the chupacabra study area, we can model the situation using a linear equation and find the answer as a point on the line.

Let's look back at Yearleigh's currency exchange. The rate "60 rupees per dollar" is constant. The number of Indian rupees Yearleigh receives is directly proportional to the number of US dollars she exchanges. The two quantities are in direct variation.

If we let y represent the number of rupees, and x represent the number of dollars, then we have the equation

$$y = 60x,$$

or we can write this equation to highlight the ratio

$$\frac{y}{x} = 60$$

and see that the constant of variation in this scenario is the exchange rate of 60 rupees per dollar.

6.3.2 Direct variation: a type of linear function

The family of direct variations is a subset of the family of linear functions, in the same way that the integers \mathbb{Z} are a subset of the rational numbers \mathbb{Q} .

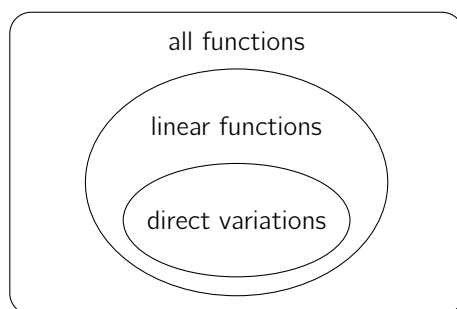


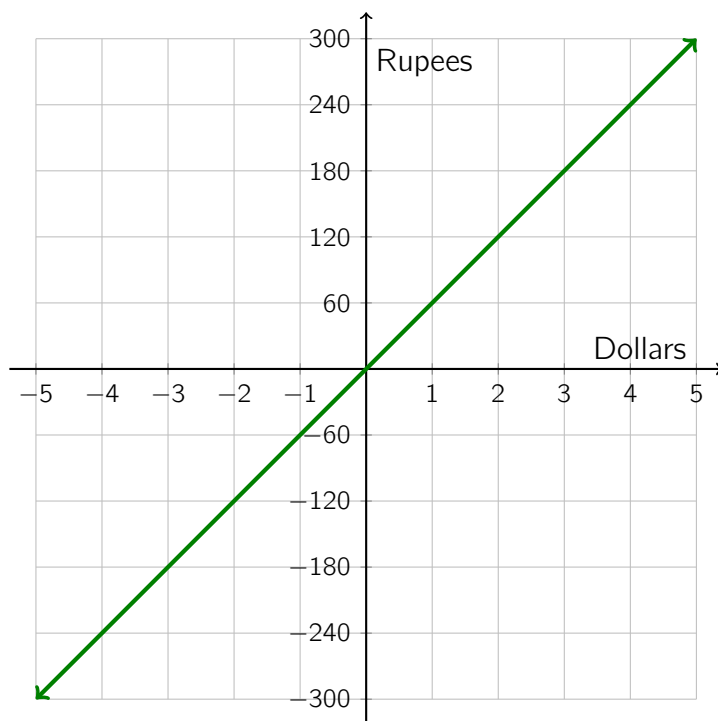
Figure 6.1: Relationship between the sets of direct variations and linear functions.

Every direct variation is a linear function, but not every linear function is a direct variation.⁶ What makes direct variation "special"?

⁶ A diagram like the one shown in section 6.3.2 is called a *Venn diagram*, named after the English philosopher John Venn. Venn diagrams are helpful for visualizing relationships between sets. In some cases, certain sets are completely contained within one another. In other cases, sets overlap on partially. How might we represent the relationships between the sets \mathbb{N} , \mathbb{Z} , \mathbb{Q} , and \mathbb{R} from chapter 1? How might we draw a Venn diagram to show the relationships between the sets of parallelograms, rectangles, rhombuses, and squares?

Graph of direct variation

This is the easiest way to spot a direct variation. It must be a straight line that goes through the origin. That's it, really! Since every direct variation is of the form $y = kx$, it's quite easy to see that when $x = 0$, we have $y = k \cdot 0 = 0$. So the point $(0, 0)$ is always on the graph of a direct variation. This makes sense in the context of the startup exploration: if Yearleigh exchanges 0 dollars, she gets back 0 rupees.

**Equation of direct variation**

The equation of a direct variation will be of the form $y = kx$. It seems simple enough, but the trickiest thing that might come up here is if we encounter an equation which, at first glance, doesn't appear to be a direct variation but which, with a little manipulation (using the POEs), can be transformed into a direct variation.

Example 6.6

Which of the following rules show that x and y are directly proportional? If a rule is a direct variation, state the constant of variation.

(a) $\frac{y}{x} = \frac{1}{2}$

(b) $2y = 8x$

(c) $y + 7 = 2x + 4$

(d) $3y + 1 = x + 1$

Solution: Equations (a), (b), and (d) represent direct variation; equation (c) does not. Equation (a) is clearly a direct variation with $k = \frac{1}{2}$. With equation (b) we can use DPOE to divide both sides by 2. The result is an equation of the form $y = 4x$, and that's another direct variation with $k = 4$.

With equation (c) we can use SPOE to get y by itself, but the result is $y = 2x - 3$, and this is not a direct variation. When we isolate y with equation (d) — first using SPOE and then DPOE — we have $y = \frac{1}{3}x$. That's a direct variation with $k = \frac{1}{3}$.

The lesson here is: if we can transform an equation using the POEs into an equation of the form $y = kx$, then we have a direct variation.

Data table of direct variation

Up until this point, we have studied techniques for identifying a table of data (or a sequence) as linear, exponential, or quadratic. We've always been given tables that count sequentially through x -values $(0, 1, 2, 3, \dots)$. As we study the different types of functions in more depth, we will be able to identify functions from tables which aren't sequential or which jump around.

For directly proportional relationships, we should be able to recognize not just that they are “linear”, but that they are specifically “direct variation”. The test for a direct variation goes back to the defining characteristic of a direct variation: the fact that $\frac{y}{x}$ is constant.

Example 6.7

Do the following data tables show direct variation? If so, write the equation for the direct variation.

(a)

x	y
-7	21
9	-27
13	-39

(b)

x	y
40	10
16	64
-20	-5

Solution: Table (a) shows a direct variation, since for each line of the table, we have the same ratio:

$$\frac{y}{x} = \frac{21}{-7} = \frac{-27}{9} = \frac{-39}{13} = -3$$

The equation for this direct variation is $\frac{y}{x} = -3$ or $y = -3x$.

Table (b) does not represent a direct variation. The first and last rows have the same ratio, but the ratio

of the middle row is different:

$$\frac{10}{40} = \frac{-5}{-20} = \frac{1}{4} \quad \text{but this is different from} \quad \frac{64}{16} = 4$$

The ratio $\frac{y}{x}$ has to be the same for all points, otherwise we do not have a direct variation. This is one of the things that makes a direct variation special. A table can still be linear and not be a direct variation: there are lots of straight lines that *do not* go through the origin! In chapter 7 we will learn a different way to determine whether a table of data represents a (generic) linear function.

Here's one more example of the kind of question associated with direct variation.

Example 6.8

Find the missing value in each case: (a) A direct variation includes the points $(16, -2)$ and $(x, 4)$. Determine the value of x . (b) A different direct variation includes the points $(6, 30)$ and $(-10, y)$. Determine the value of y .

Solution: We are told that we have a direct variation, and so we know that the two points must share the same ratio. We can therefore set up a proportion to solve part (a):

$$\frac{-2}{16} = \frac{4}{x}$$

and solve to find the value of x . We find that $x = -32$.

Let's use a different approach to solve part (b). Knowing one point is enough to write the equation for the direct variation. We know that the point $(6, 30)$ is on the direct variation, and so

$$k = \frac{y}{x} = \frac{30}{6} = 5.$$

Therefore, we have the equation $y = 5x$. We can then substitute -10 for x and solve to discover that $y = -50$.

A final note: Technically, we can write all of these ratio with x over y and find the answers we are looking for. The reason we stick so closely to “ y over x ” is the relationship that this quantity has to the unit rate in the problem. Plus, we want the notation that we use with direct variations to be consistent with the notation we use with linear functions generally. For example, we will soon meet a very important ratio called *rate of change* in which we set the change in the independent variable over the change in the dependent variable.

6.4 Inverse variation

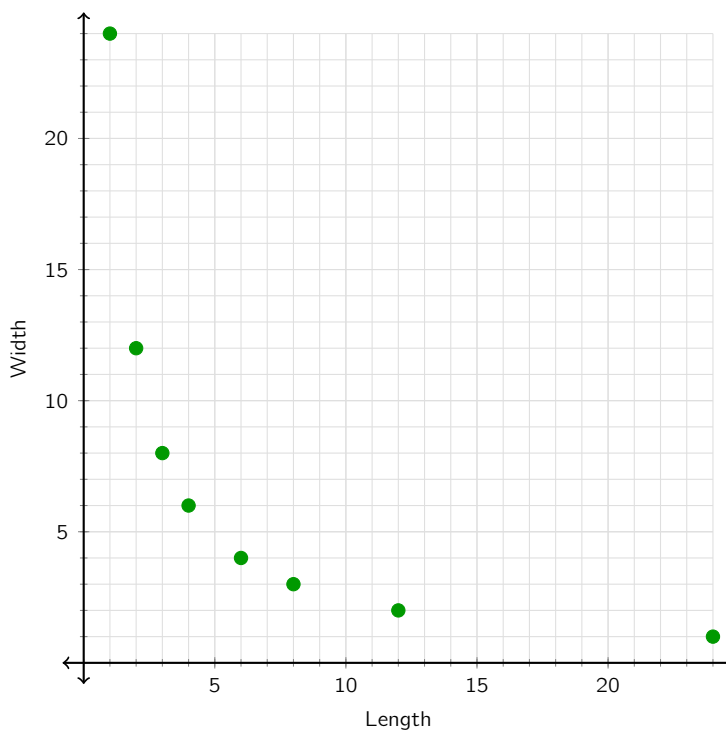
Herb garden

// startup exploration

François has enough mulch to plant an herb garden covering 24 square meters. Plus, he has a whole barn full of snap-together fencing units that are each 1-meter long. How many different herb gardens can François create if he wants to use all the mulch and surround the garden with a rectangular fence?

To solve François's problem, we need to find a rectangle that has natural number side lengths and whose area is 24 square meters. One candidate is the long, skinny rectangle that is 1 meter by 24 meters. Of course, that's not the only one! We list the possible rectangles below and plot a graph of the data.

Length (x)	Width (y)
1	24
2	12
3	8
4	6
6	4
8	3
12	2
24	1



The first thing that should jump out about this graph is that it is definitely not linear! The J-like shape suggests that it could be an exponential relationship. . . but is it? What could we do to determine whether or not this is an exponential relationship?

Full disclosure

Even though we're beginning our discussion of linear functions, inverse variation is *not* a linear relationship. We discuss it here just as a way to compare and contrast inverse and direct variation.

Inverse variation

The variables x and y are said to be *inversely proportional* if their values have a constant product. In other words, $xy = k$, where k is a constant called the *constant of variation*.

An equation of the form $y = \frac{k}{x}$ is called an *inverse variation*. The quantities x and y are inversely proportional.

In an inverse variation (where the variables are inversely proportional), the most important thing to remember is the fact that $x \cdot y$ is constant. For every point (x, y) the product of x and y is always the same number.

When we force the product of two numbers to be the same, certain patterns arise. Most notably, if the value of one of the quantities goes up, the other one has to go down.

Data table of inverse variation

If we step away from context and think purely about data, the test for whether some data shows an inverse variation is whether $xy = k$, some constant value.

Example 6.9

Do the tables below show inverse variation? If so, write the equation for the direct variation.

(a)

x	y
2.5	20
-10	-5
8	5.5

(b)

x	y
3.5	4
1	14
-7	-2

Solution: Table (a) does not represent an inverse variation. The first and second rows of the table have the same product (50), but the last row has a different product (44). Table (b) does represent an inverse variation: all three rows have the same product (14). So, table (b) has equation $xy = 14$.

Example 6.10

Find the missing values in the table, given that the data represent an inverse variation.

(a)

x	y
3	a
9	4
2	b
c	8

Solution: The second row of the table gives us the clue we need: $9 \cdot 4 = 36$, so that must be the constant of variation for this inverse variation. Then, it is simply a matter of finding the partner of each value that multiplies to get 36.

In the first row, $3a = 36$ implies that $a = 12$. In the third row, $2b = 36$ implies that $b = 18$. In the fourth row, $8c = 36$ implies that $c = 4.5$.

Equation of inverse variation

We have seen that the equation for an inverse variation is either $yx = k$ or $y = \frac{k}{x}$. Writing the equation for an inverse variation brings out two important features that we must be aware of. Consider the equation

$$y = \frac{k}{x}$$

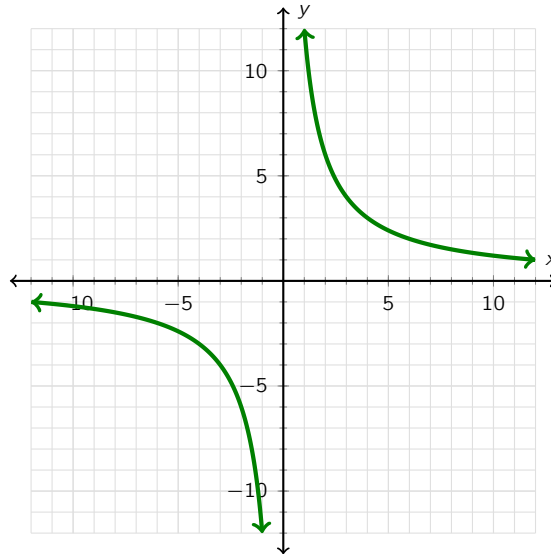
First, what happens when $x = 0$? Yikes! Remember, division by zero is undefined, and so 0 is an illegal value for x . We say that 0 must be *restricted from the domain* of the function.

Second, what happens when $y = 0$. . . or is that even possible? Note that as x gets larger and larger, the value of the fraction $1/x$ gets smaller and smaller (closer and closer to zero). But, no matter how big x gets, $1/x$ will never be *equal* to zero. It will get close to zero. Super close! Impossibly close! But it will never equal zero.

We can see this unusual behavior reflected in the graph.

Graph of inverse variation

Not being able to have an x or y value of zero does very interesting things to the graphs. Consider the inverse variation $xy = 12$ or $y = \frac{12}{x}$:



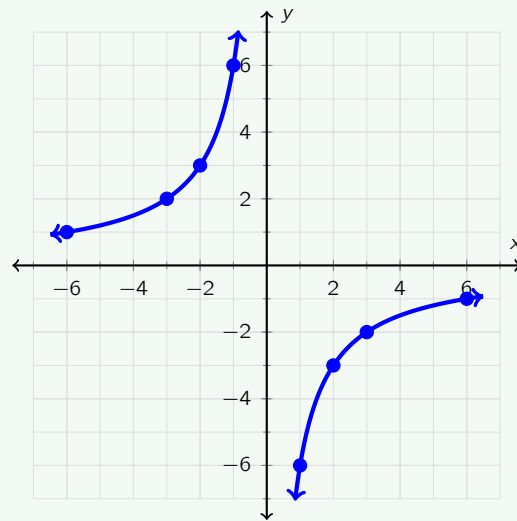
Notice a few things: First, we can see from the graph that this is not an exponential relationship. Although the portion of the graph in the first quadrant might look exponential, when we plug in negative input values we get different behavior that we'd expect from an exponential relationship.

Secondly, the graph is in two disconnected pieces, and since it never crosses the x - or y -axis, it appears entirely in the first and third quadrants. It appears in those particular quadrants because of the defining feature of inverse variation: x and y have a constant product.

This graph is the picture of all points whose coordinates have a product of (positive) 12. The only way to get a positive product is to either multiply two positive numbers, or two negative numbers — exactly the description of the coordinates in the first and third quadrants! What do you suppose the graph of $y = \frac{-12}{x}$ might look like?

Example 6.11

Write the equation for the inverse variation pictured in the graph below.



Solution: Since we know that this is an inverse variation, all we need is to find the coordinates of one point. Some are plotted for us, so that's good news. If we use the point $(-2, 3)$, we can see that $k = xy = -2 \cdot 3 = -6$. Once we have found the constant of variation, we can easily write the equation: $xy = -6$ or $y = \frac{-6}{x}$.

About the “family” of inverse variation

You may have noticed that inverse variation doesn't seem to fit into any of the function families we have looked at so far. This is true. Inverse variation doesn't belong to linear, quadratic, exponential families of functions, and so we won't do much more with inverse variation for quite a while.

Inverse variation is actually a member of the super-family of functions called *rational functions*, which is more of a focus in Algebra 2. This rational super-family has a number of interesting and surprising features, and they can take on a whole variety of bizarre shapes.

Inverse variation graphs do have something in common with exponential functions: asymptotes. As we saw, inverse variations get impossibly close to both the x - and y -axis, but never reach them. Generally, these graphs have two asymptotes, one horizontal, one vertical.

» Chapter summary «

In this chapter we started in familiar territory: writing and solving proportions. We quickly progressed into a more “algebraic zone” by extending the idea of proportionality to the idea of variation.

We also saw how a directly proportional relationship gives rise to a linear function: the direct variation. We'll take this idea and run with it in the next chapter: Why should we be limited to straight lines that go through the origin when we can learn to describe straight lines in general? Onward!

A curve does not exist in its full power until contrasted with a straight line.

Robert Henri, American painter

Chapter 7

Linear functions

So far we have learned quite a bit about the family of linear functions. From our work with sequences we learned how to tell if a sequence was arithmetic. We can write recursive rules for arithmetic sequences, and we can change those recursive rules into explicit formulas.

We turned arithmetic sequences into tables of data, and when we graphed some linear functions, we saw that all were straight lines. We saw a preview of how different transformations might change the way its graph looks.

Most recently we learned about a specific type of linear function, the direct variation. In this chapter we will bring together all of the pieces we have learned previously into a single coherent, and very important, concept.

7.1 Rate of change and slope

CCC

// startup exploration

Bob enrolls at Cheeseville Community College, and the admissions counselor there gives him the following table estimating the cost to attend. The total cost of a semester at CCC is made up of two expenses: (a) the fixed fees, which are the same for every student every semester (technology fee, printing allowance, and so on), and (b) the tuition rate, which varies per credit hour.

Most courses at CCC are either 3 or 4 credit hours per semester. The table summarizes several combinations of 3- and 4-hour courses.

No. Credit Hours	3	4	6	7	10
Cost (dollars)	501	654	960	1113	1572

What can you learn from the table? How much is tuition per credit hour? How much does CCC estimate Bob will pay in fees?

7.1.1 Unit rates in data

Since we're fresh from a discussion of proportional reasoning, perhaps our first guess might be that this is a direct variation. Unfortunately, a quick check of the ratios $\frac{y}{x}$ shoots that idea down:

$$\frac{501}{3} = 167 \quad \text{but} \quad \frac{654}{4} = 163.5$$

Maybe we have an arithmetic sequence? If so, then the terms of the “cost” sequence will have a constant difference. But, if we take the differences between neighboring terms, we have different values:

$$654 - 501 = 153 \quad \text{but} \quad 906 - 654 = 306$$

But, hang on. Look at the corresponding “number of credit hours”. These numbers don’t increase by the same amount at every step: the first step goes from 3 to 4 (an increase of 1), but the second step goes from 4 to 6 (an increase of 2). This is unlike what we saw for arithmetic sequences, in which the term number increased steadily, by 1 unit every step. We’re going to have to be a bit more clever to account for this!

Notice that when “credit hours” increases by 1, “cost” increases by \$153. And then, when “credit hours” increases by 2, “cost” increases by \$306 – the increase in both values has doubled! This even works at the end of the table: When the number of hours jumps by 3 and the cost jumps by \$459, the rate “dollars per 1 credit hour” says the same:

$$\frac{153}{1} = \frac{306}{2} = \frac{459}{3} = \frac{\text{change in the total cost}}{\text{change in the number of hours}}$$

So, even though the table skips over some numbers, the *change* in the cost still corresponds to the *change* in the number of hours. So, the tuition rate at CCC must be \$153 per credit hour.

Example 7.1

Does the data in the table below represent someone moving at a constant speed? If so, what is the speed?

Time (sec)	Distance (m)
5	17.5
12	42.0
31	108.5

Solution: Comparing the first two data points, we see that the change in distance is $42.0 - 17.5 = 24.5$

meters, and that this occurs over $12 - 5 = 7$ seconds. This ratio “24.5 meters per 7 seconds” corresponds to the unit rate “3.5 meters per second”.

Comparing the second pair of data points, we see that a change in distance of $108.5 - 42.0 = 66.5$ meters happens over $31 - 12 = 19$ seconds. The ratio “66.5 meters per 19 seconds” corresponds to the unit rate “3.5 meters per second”.

So, yes, this table does represent movement at a constant speed of 3.5 meters per second.

7.1.2 Rate of change

We want to be able to tell if *any* data set is linear. The examples of speed and hourly tuition begin to suggest a general approach. We are seeking a constant **rate of change**.

Rate of change

Rate of change is a measurement of how quickly the dependent variable changes relative to a one-unit change in the independent variable.

To compute the rate of change of y with respect to x , we want to “unitize” a change in y (the dependent variable) with respect to the corresponding change in x (the independent variable).

To write this out using mathematical notation, we use an abbreviation for the idea of “change”: the Greek letter delta, Δ . This makes the definition of rate of change look like the following:

$$R.O.C. = \frac{\Delta \text{ dependent variable}}{\Delta \text{ independent variable}} = \frac{\Delta y}{\Delta x}$$

The symbol “ Δy ” is pronounced “delta y ” or “change in y ”. We sometimes say that rate of change is “delta y over delta x ”.

Speed is an example of a rate of change. It is a measurement of how distance changes relative to a one-unit change in time: meters per (one) second. Bob’s tuition at Cheeseville Community College is another example that measures how his total expense changes relative to a unit-change in the number of credit hours he takes: dollars per (one) hour.

7.1.3 Slope

Rate of change is a very important concept in mathematics, and it is a general term that doesn’t just apply to straight lines. Different types of functions have different patterns that emerge in their rates of change. The study of rate of change led to the development of calculus!

But, the rate of change of a straight line is special: it's always the same. When it comes to linear relationships, the key is knowing that *the rate of change for linear data is constant*. We have a special term for the rate of change in a linear relationship.

Slope

Slope is a measurement of the steepness of a line.

For a line, the terms “slope” and “rate of change” describe the same thing: steepness. When we have a curve (like the graph of an exponential or quadratic function), we don't usually talk about its “slope” because the steepness changes, as we saw in section 3.5. We can use the term “rate of change” with non-linear functions, as we will see in later chapters.

In mathematical notation, the letter most often used to abbreviate slope is m .¹ So, we can update our notation from earlier:

$$m = \frac{\Delta y}{\Delta x}$$

Slope from data

As we have seen in the earlier examples, finding slope from a data table is rather easy. We find the change between two y -values and find the change between the corresponding x -values. When we unitize this ratio – in other words, when we divide Δy by Δx – we've got slope. Woot!

This works even if the data is not given in table form. You may have heard the old saying, “The shortest distance between two points is a straight line.” Well... what's the slope of that line?

Example 7.2

Find the slope of the line connecting the points $(11, 4)$ and $(16, 29)$.

Solution: The change in the y -values is $\Delta y = 29 - 4 = 25$, and the change in the x -values is $\Delta x = 16 - 11 = 5$. So, the slope of the line between these two points is

$$m = \frac{\Delta y}{\Delta x} = \frac{25}{5} = 5.$$

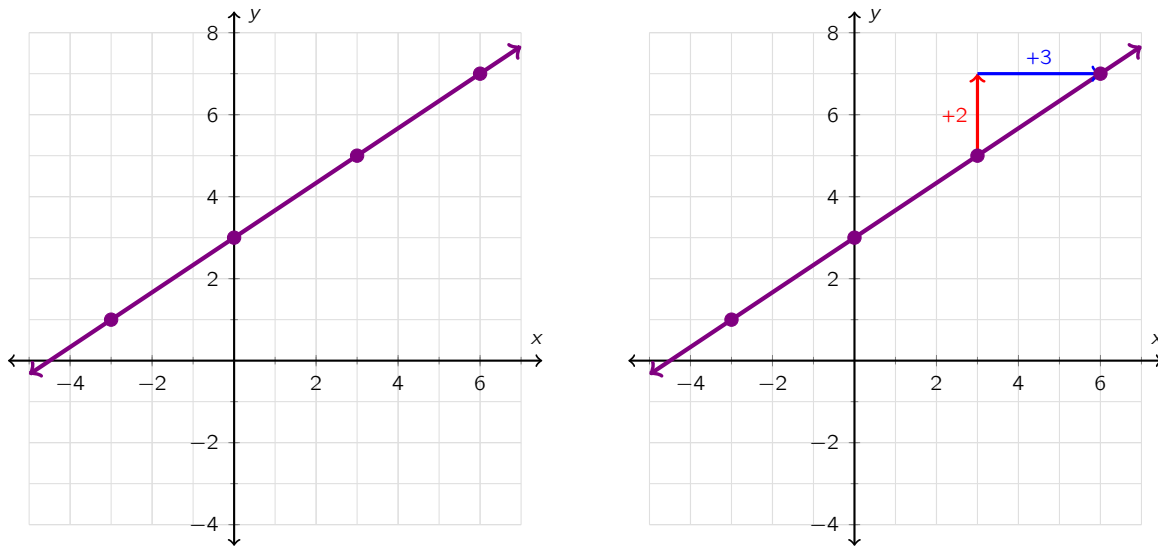
¹ “Why m ?” we bet you're wondering. Many people have looked into this question, digging back through mathematical writing to try and learn where this abbreviation comes from. It seems that the earliest recorded use of m for slope was in 1757 by Italian mathematician Vincenzo Riccati. But his work includes no explanation of *why* the letter m was used.

So, the line connecting the points (11, 4) and (16, 29) has a slope of 5.

Slope from a graph: The slope triangle

Suppose we're given the graph of a line and asked to find the slope. We could create a data table using points on the line and find the slope from that data. Or, we could inspect the graph directly using a *slope triangle*. A slope triangle is a right triangle that connects two points with a single vertical move and a single horizontal move.

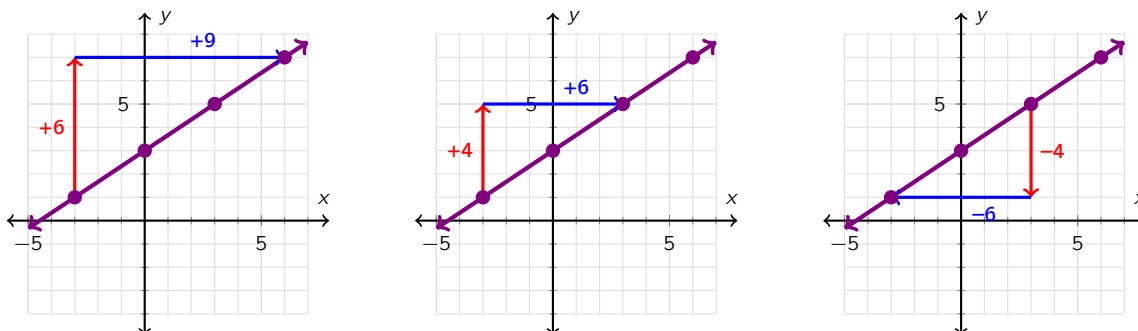
Here is the graph of a line with some points marked. To find the slope of the line, we draw a slope triangle.



the red arrow shows the vertical change. It starts at 5 and goes to 7, a vertical distance of $7 - 5 = 2$ units. The blue arrow shows the horizontal change. It starts at 3 and goes to 6, a horizontal change of $6 - 3 = 3$ units. Now we can calculate slope

$$m = \frac{\text{vertical change}}{\text{horizontal change}} = \frac{2}{3}.$$

Since the rate of change is constant for a straight line, it doesn't matter which two points we choose, or which point we choose as the starting place. Each of the slope triangles below can be used to compute slope of this line!



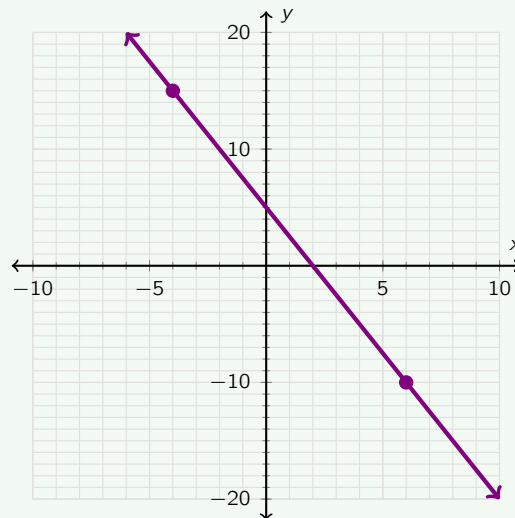
Again, we don't have to use a slope triangle. We can just convert at least two of the points on the graph into a table and then use the techniques for finding slope from a data set.

There are two commonly mixed up things about slope. First: reciprocating the ratio. Remember it's change in y over change in x , not the other way around.

Second, folks sometimes have problems with sign. A rule of thumb: if the data has a positive correlation, the slope is positive. If the data has a negative correlation, the slope is negative. It's a good habit always to double-check the sign of a slope to verify that it makes sense in context.

Example 7.3

Find the slope of the line depicted in the following graph.



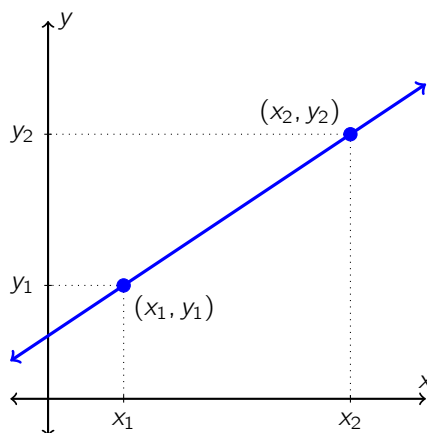
Solution: This problem exhibits two key features: the slope of the line will be negative (since the line decreases as we move to the right), and we're going to have to be careful about the scale on the axes! We might see a vertical change of one "tick mark" on the axis, but note that one tick mark is a change of 5 units. Similarly, one tick on the x -axis is a change of two units.

If we start at the point on the left and move to the point on the right, we have a horizontal change of +10 units in the horizontal direction and -25 units in the vertical direction (a negative vertical change means moving downwards). So, the slope is

$$m = \frac{-25}{10} = \frac{-5}{2} = -\frac{5}{2}$$

7.1.4 Slope formula

We always use the same basic approach to compute slope, and so we can derive a formula for slope. Study the picture below, showing a line through two “generic points” called (x_1, y_1) and (x_2, y_2) .



The vertical change between the two points is $y_2 - y_1$, and the horizontal change is $x_2 - x_1$. So we can compute the slope between these two points using our familiar formula

$$m = \frac{\Delta y}{\Delta x} = \frac{y_2 - y_1}{x_2 - x_1}.$$

The trickiest part about using this formula is being consistent about which point we call (x_1, y_1) and which point we call (x_2, y_2) . It doesn't matter which point is which – but we have to be consistent.

Tangent: Order of the points

That last sentence says “it doesn't matter which point is which.” Why is that? Explain why it doesn't matter which point we call “point 1” and which we call “point 2”. In other words, if we have two points called (x_1, y_1) and (x_2, y_2) , explain why

$$\frac{y_2 - y_1}{x_2 - x_1} = \frac{y_1 - y_2}{x_1 - x_2}.$$

(Hint: Experiment with some numbers first, then try to extend your observations to the two “generic” points given.)

7.1.5 Horizontal and vertical lines

If we think about how slope is calculated – vertical change over horizontal change – a horizontal line has *no vertical change*, or a vertical change of zero. If we pick two points on a horizontal line and calculate slope, we would get 0 divided by some number, which always equals zero! So the slope of a horizontal line is always 0.

This makes sense if we remember the slope is a measure of the steepness: a horizontal line is flat, and that means no “steepness” at all!

If we pick two points on a vertical line and try to calculate slope, we will get a horizontal change of 0. So we'll have some number divided by 0... and that's off limits. We say that the slope of a vertical line is *undefined*, since division by 0 is undefined.

7.1.6 Using slope

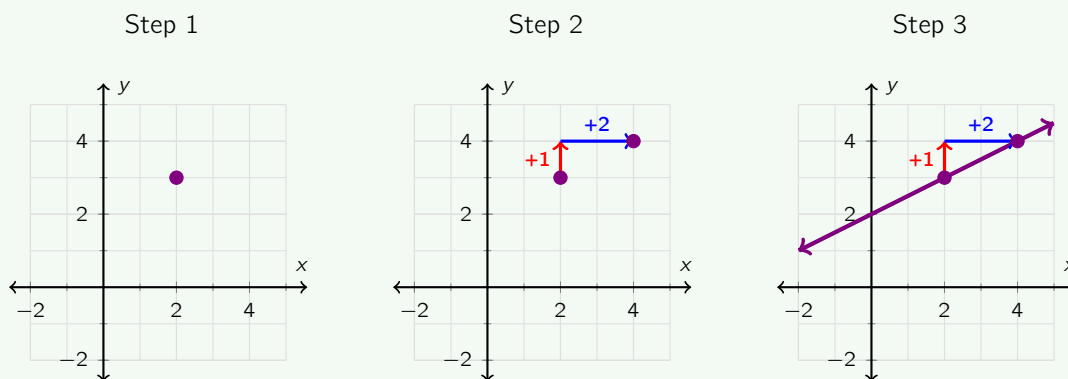
Before, when we were graphing a line we had to take some domain values, like the integers between 3 and 3, plug them in to the equation, and plot every single point, one at a time. Now we have an alternative!

All we really need to graph a line is two points. If we have one point on the line and the *slope*, we can find another point on the line.

Example 7.4

Graph the line that goes through $(2, 3)$ and has a slope of $\frac{1}{2}$.

Solution: Step 1: Plot the given point. Step 2: The slope $\frac{1}{2}$ tells us how to find another point on the line: the numerator is the change in y , so we move up 1 unit. The denominator is the change in x , so we move 2 units to the right. Now we have a second point: $(4, 4)$. Step 3: Connect the dots and you have your graph!



By the way, perhaps now you see why we've been such fans of improper fractions. Seriously, which form is easier to use when creating a graph: a line with slope $\frac{4}{3}$ or a line with slope 1.33333...?

Colinearity

Points that lie on the same line are said to be *colinear*. If we are asked to determine whether two points are colinear... well, that's easy! Any two points will define a line, so any pair of points will be colinear.

But what about three points? How could we determine whether three points are colinear – in other words, whether they all lie on the same line?

Example 7.5

Are the points $(4, 1)$; $(-1, 5)$; and $(1, 2)$ colinear? Why or why not?

Solution: The slope between the first pair of points is:

$$\frac{y_2 - y_1}{x_2 - x_1} = \frac{1 - 5}{4 - -1} = \frac{-4}{5} = -\frac{4}{5}.$$

The slope between the second pair of points

$$\frac{y_2 - y_1}{x_2 - x_1} = \frac{5 - 2}{-1 - 1} = \frac{3}{-2} = -\frac{3}{2}.$$

Those two slopes are not equal, so: nope! These points won't line up in a single line! They are not colinear.

7.2 Point-slope form

We are going to learn three different ways to write equations for lines: the point-slope form, the slope-intercept form, and the standard form. Each form tells you something a little different about the line, and each one has its own set of pros and cons. Over the next few sections we will learn about these different forms, what they tell us about lines, what they are good for, and how to convert from one form to another.

One point, one slope

// startup exploration

A line with slope 2 passes through the point $(4, 3)$. Name three other points that also lie on this line. Can you find a point on the line that lies in the second quadrant? The third quadrant? The fourth quadrant?

The first form that we will study is called *point-slope form*. As the name suggests, all we need to write an equation in this form is the slope of the line and a point on the line.

In the startup exploration, we are asked to find different points on the line that has slope 2 and passes through the point $(4, 3)$. Finding other points isn't that hard. We saw how to do this in the last example in section 7.1.² It is more interesting to consider a generic point (x, y) on this line.

We now have two points on the line $(4, 3)$ and (x, y) , and we know the slope is 2. We can arrange all of this information using the slope formula:

$$2 = \frac{y - 3}{x - 4}.$$

There should be something familiar about this: we've got an equation with a y and an x in it... so this is starting to look like a graphable equation. All we have to do is transform this equation to isolate y . Let's do that!

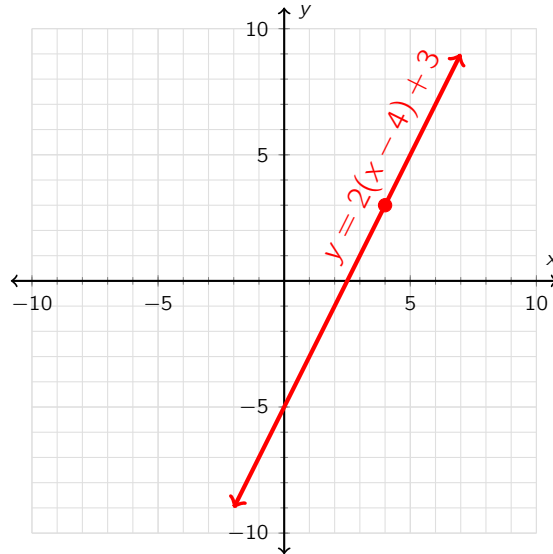
$$2 = \frac{y - 3}{x - 4}$$

$$2(x - 4) = y - 3 \quad \text{MPOE: multiply both sides by } (x - 4)$$

$$2(x - 4) + 3 = y \quad \text{APOE: add 3 to both sides}$$

If we flip this around to write it in a more familiar format, we have $y = 2(x - 4) + 3$. That's an equation we can graph! If we plot some points (or pull out our favorite piece of graphing technology) we can graph this equation:

² Well, the line doesn't pass through the second quadrant, so finding a point that meets that requirement is impossible. Something to ponder: are the following statements true or false? (a) A linear function can pass through at most 3 of the 4 quadrants. (b) A linear function will always pass through exactly 3 of the 4 quadrants.



What do you know! Graphing this equation gives us exactly the line we were looking for: a line with slope 2 that goes through the point $(4, 3)$. So the equation $y = 2(x - 4) + 3$ describes the line. Can you see the slope and the given point hiding there in the equation?

7.2.1 Deriving point-slope form

Let's make things even more generic. Suppose we have a line with slope m that goes through the specific point (x_1, y_1) . Then let (x, y) be any other random point on the line. We can summarize this using the slope formula, and then transform our equation into a graphable, $y = \text{something}$ format:

$$m = \frac{y - y_1}{x - x_1}$$

$$m(x - x_1) = y - y_1$$

$$m(x - x_1) + y_1 = y$$

MPOE: multiply both sides by $(x - x_1)$

APOE: add y_1 to both sides

This last line is what we call the **point-slope form** of a line.

Point-slope form

The form of a linear equation that uses the slope and any point on the line. It is written either $y - y_1 = m(x - x_1)$ or $y = m(x - x_1) + y_1$, where m is the slope of the line and (x_1, y_1) is a point on the line.

7.2.2 Using point-slope form

The trickiest part about the point-slope form – as you might anticipate – is handling the plusses and minuses. It can be tricky to remember the $-x_1$ and the $+y_1$, so be on the lookout and **always** check your signs.

Example 7.6

State the slope of each of the following lines and name one point on each line.

$$(a) \ y = -3(x - 1) - 4 \quad (b) \ y = 5 - 8(x + 4) \quad (c) \ y - 8 = 4(x - 2)$$

Solution: Let us compare the given equation (a) to the generic point-slope formula

$$y = -3(x - 1) - 4 \quad \Longleftrightarrow \quad y = m(x - x_1) + y_1,$$

We can see that the slope is -3 . The equation has “ -4 ” whereas the formula wants “ $+y_1$ ”, so we change the rule to reflect addition outside the parentheses: $y = -3(x - 1) - 4 = -3(x - 1) + (-4)$. Now we can see that a point on the line is $(1, -4)$.

With (b), note that things are a bit out of order, but we can commute the terms (taking the sign along, remember!) and have $y = -8(x + 4) + 5$. Now, we have addition in the parentheses, whereas the rule wants subtraction in there. We can make that adjustment easily enough: $y = -8(x - (-4)) + 5$. From this version, we can see that the slope is -8 and a point is $(-4, 5)$.

The definition of point-slope forms gives us this form as an alternative (check the box above)! So, the slope of this line is 4 and a point on the line is $(2, 8)$.

Back in chapter 3, we had to pick some values for the domain, substitute them into the equation, find the values of the range, and plot each point. In section 7.1, we learned that we really just need a point and a slope. This information is ready to be extracted from the point-slope form!

Example 7.7

Graph (by hand) the line $y = 3(x - 4) - 5$.

Partial solution: We can see that the slope is 3 and that a point on the line is $(4, -5)$. With this information we find a second point, and then graph the line!

How would we go about writing an equation in point-slope form for data given to us in a table or as a collection of points?

Example 7.8

Write the equation of the line that goes through the points $(6, 10)$ and $(-1, -4)$.

Solution: First, we find the slope of this line:

$$m = \frac{y_2 - y_1}{x_2 - x_1} = \frac{10 - (-4)}{6 - (-1)} = \frac{14}{7} = 2.$$

Then, we can use either of the two points as the “point” in point-slope form:

$$y = 2(x - 6) + 10 \quad \text{or} \quad y = 2(x + 1) - 4$$

7.2.3 Relating point-slope to earlier concepts

There is a way to look at the point-slope form that relates to transformations in a plane. In section 6.3, we studied direct variations of the form $y = kx$. This is a straight line through the origin, and k is the slope. So let's rewrite this as $y = mx$.

If we want to move the graph of the direct variation up or down, we add something to the equation. For example, if we want to move the graph 12 units up, the equation would become $y = mx + 12$. If we want to move down 7 units, the equation would become $y = mx + -7$. In general, if we want to move the graph vertically y_1 units the equation would become $y = mx + y_1$.

If we want to move the graph of the direct variation to the right 5 units, we would replace x with $(x - 5)$ and our new equation would be $y = m(x - 5)$. If we want to move the graph to the left 8 units, we have the equation $y = m(x + 8)$. In general, if we wanted to move the graph horizontally x_1 units, the equation would become $y = m(x - x_1)$.

If we want to cause *both* a horizontal and vertical shift, we would do both transformations to get the point-slope form: $y = m(x - x_1) + y_1$. We can think of this transformation as “moving the origin” of the direct variation from $(0, 0)$ to (x_1, y_1) .

Yet another way to look at point-slope form is to think back to the rules we wrote for linear sequences back in section 2.3. Given a sequence like 4, 7, 10, 13, \dots , we wrote the rule

$$a_n = 4 + 3(n - 1).$$

Can you see the connection to point slope form? In this rule the slope is 3 and the point is $(1, 4)$. The slope is telling us the constant difference: the amount that we add on for each step forward in the sequence. The point $(1, 4)$ is telling us that the first term in the sequence is 4.

Back in section 2.3 we also wrote zero-based rules for arithmetic sequences. What would be the zero-based rule for the sequence given above? These zero-based rules are at the core of the form that we will study in the next section.

7.3 Slope-intercept form

Point-slope form is handy and quite easy to write in many cases. But it has those parentheses in it which means that we could simplify it even further. We explore that possibility in this section.

Two become one

// startup exploration

Graph the line that goes through the points $(-1, 7)$ and $(2, -2)$. Find the slope through these two points and write two equations in point-slope form for the line.

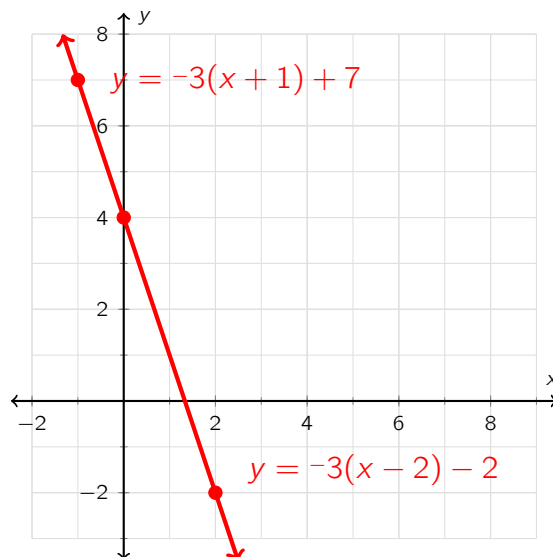
Use your toolbox of equivalence rules to eliminate the parentheses and rewrite each of your point-slope equations in simplest form. What happens? How is your new equation reflected in the graph?

Point-slope form is the easiest form for algebra students to use when writing the equation of a line. Once we know the slope of the line we only need one point on the line – any point at all – to write the equation.

However, point-slope form is a pain for algebra *teachers* to grade. Why? Well, how many points are there on a line? Each one of those points can be used to write an equation for that line, which means there are infinitely many different point-slope equations for a single line! If a student gets creative about what point to use, the teacher may have to spend valuable time to determine whether the student's answer is correct or not.³

One way to simplify things would be to designate a kind of “standard point” that we always use at *the point* when writing a point-slope form equation. Luckily, there's a natural point to set as the standard.

The startup exploration asks us to graph and write two different equations for the given line.



³ Don't get any ideas.

We can eliminate the parentheses using the distributive property, and then we can combine like terms as needed. If we use the point $(-1, 7)$, then we have:

$$\begin{aligned}
 y &= -3(x + 1) + 7 \\
 &= -3x + -3(1) + 7 && \text{distributive property} \\
 &= -3x + -3 + 7 && \text{substitution} \\
 &= -3x + 4 && \text{substitution/combining like terms}
 \end{aligned}$$

If we use the point $(2, -2)$, then we have:

$$\begin{aligned}
 y &= -3(x - 2) - 2 \\
 &= -3(x + -2) + -2 && \text{adjust negative signs, for safety} \\
 &= -3x + -3(-2) + -2 && \text{distributive property} \\
 &= -3x + 6 + -2 && \text{substitution} \\
 &= -3x + 4 && \text{substitution/combining like terms}
 \end{aligned}$$

After this simplification process, the two different point-slope form equations have turned into the *same equation*! It turns out that this always happens: if we simplify two different point-slope equations for the same line (by getting rid of the parentheses and combining like terms), we will always end up with the same equation. This simplified version is called the **slope-intercept form**.

Slope-intercept form

The **slope-intercept form** of a linear equation is an equation of the form

$$y = mx + b.$$

The value m is the slope of the line, and b is the value at which the line crosses the y -axis. The point $(0, b)$ is called the y -intercept of the line.

Note from the definition that the equation $y = -3x + 4$ contains two pieces of information. It tells us the slope of the line is -3 , and it tells us that the line passes through the point $(0, 4)$. We can verify this in the graph above.

Since a linear function has a constant slope and can only have one y -intercept (two straight lines can't cross more than once!), a linear function has only one slope-intercept equation.

Graphing a line in slope-intercept form is exactly the same as graphing a line in point-slope form. In fact, it's easier! Figuring out the "point" in point-slope form can lead to some challenges with positive and negative numbers. With slope-intercept form, the "point" is much easier to identify. Then, we can use the slope to find a second point and graph the line.

7.3.1 Writing equations in slope-intercept form

Suppose we are given the slope of a line and a point on the line that is *not* the y -intercept. There are two primary ways to get to an equation for a line in slope-intercept form.

Method 1: Simplify point-slope form. This method works just as we saw earlier. We use the slope and the point to write an equation in point-slope form, which we then simplify using the distributive property and combining like terms. The result is slope-intercept form.

Method 2: The “old school” method for folks who don’t like (or never learned) point-slope form. In this approach, we use the slope and the x and y coordinates of the point to write an equation that we can use to calculate b .

Example 7.9

A line has a slope of 2 and passes through the point $(3, 7)$. Write the slope-intercept equation of the line.

Solution: We’re given $m = 2$, so we can start writing an equation in slope-intercept form: $y = 2x + b$. We don’t know b yet, but we do know a pair (x, y) that is on the line! So, we plug in the x and y values from the point: $x = 3$ and $y = 7$. In other words:

$$7 = 2(3) + b$$

We can now solve for b by subtracting 6 from both sides (SPOE). We find that $b = 1$. So, the slope-intercept equation is $y = 2x + 1$.

These two methods are equally valid mathematically, so pick the approach that you understand the best and feel most confident about.

Example 7.10

Write an equation in slope-intercept form for the line containing the following data points.

x	y
5	12
6	8
7	4
8	0

Solution: If a table of data ever includes the y -intercept – the point with x -coordinate 0 – then our job

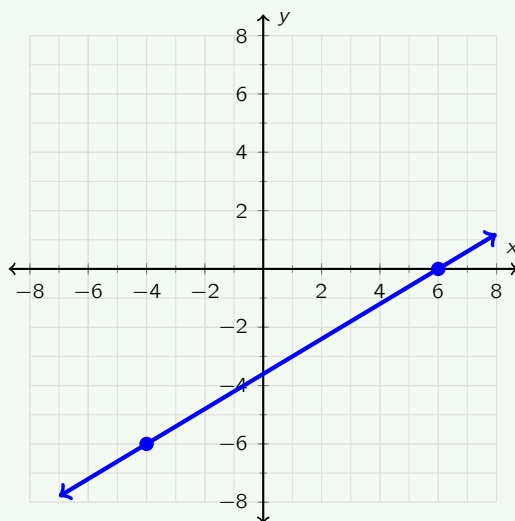
is quite easy. In this case, though, we don't have that. To write the equation, we could use one of the two methods above, or we could do a bit of detective work and find the y -intercept using the pattern in the table.

Notice that the y -values decrease by 4 as x increases by 1. That means the slope of the line is -4 . It also means we can work our way backwards in the table to the point where $x = 0$ and $y = 32$. So, our equation is $y = -4x + 32$.

These equations should look familiar. We saw equations just like these when we were studying arithmetic sequences! The value of the sequence at stage 0 was just the y -intercept. The common difference (the value we added at each step) was the slope!

Example 7.11

Write an equation in slope-intercept form for the line pictured in the graph below.



Solution: If a graph has a clear y -intercept then our job is quite easy. In this case, though, we don't have that. We might try to guess about the point – maybe it's -3.5 ...ish? Would you bet on that? – but we need another approach to be sure.

So, we'll use a slope triangle to find the slope between the two given points $\frac{6}{10} = \frac{3}{5}$. Then, we'll write an equation in point-slope form using the point $(6, 0)$: $y = \frac{3}{5}(x - 6) + 0$. Simplifying using the distributive property and combining like terms gives:

$$y = \frac{3}{5}x - \frac{18}{5},$$

and $-\frac{18}{5} = -3.6$, which is close to -3.5 , but not the same!

7.3.2 (,,:) Equivalence of the point-slope and slope-intercept forms

In this extension section, we'll look in more detail at a claim we made earlier in this chapter. If the content here feels like too much, take a break and come back to it later. Your brain keeps working on things even when you're not thinking about them consciously!

In section 7.3 we made the claim that *always get the same answer* when we start with an equation in point-slope form and we convert to an equation in slope-intercept form (using the distributive property and combining like terms).

Have you started to become suspicious of claims like this? Have you started to wonder, "How do we know that we *always* get the same answer?" This is a good question to ask. After all there are infinitely many different point-slope equations we could write. Surely we're not going to try and convert them all and check that we get the same answer every time.

Here's one way to think about this comparison. We have a line. Since we are going to assume we have a linear function, we know that this line is not vertical. So, that means it must cross over the y -axis at some point $(0, b)$. Suppose (x_1, y_1) is another point on that line.

This line has a slope that we can represent using the slope formula. We'll use the y -intercept as "point 2":

$$m = \frac{b - y_1}{0 - x_1} = \frac{b - y_1}{-x_1}$$

Trust us on this next bit: it will be helpful. We're going to use the POEs to isolate b so that we can make a comparison.

$$m = \frac{b - y_1}{-x_1}$$

$$-mx_1 = b - y_1 \quad \text{MPOE: multiply both sides by } -x_1$$

$$-mx_1 + y_1 = b \quad \text{APOE: add } y_1 \text{ to both sides}$$

We're going to come back to this last line in a moment.

Our goal is to show that point-slope form and slope-intercept form are equivalent. So, let's write the point-slope form of the equation using the generic point (x_1, y_1) and try to convert it to slope-intercept form.

$$y = m(x - x_1) + y_1 \quad \text{point-slope form}$$

$$y = m(x + -x_1) + y_1 \quad \text{rewrite subtraction}$$

$$y = mx + -mx_1 + y_1 \quad \text{distributive property}$$

$$y = mx + -mx_1 + y_1 \quad \text{the expression } (mx_1 + y_1) \text{ is equal to } b$$

$$y = mx + b \quad \text{slope-intercept form}$$

So, in the end, we have shown that if we have a line with y -intercept $(0, b)$, then the point-slope equation for the line – using any point on the line! – can be converted into the slope-intercept equation of the line.

7.4 Standard form

So far, we know a lot about writing equations for lines. We can find the slope of a line, we can write equations in point-slope form, and we can convert point-slope form into slope-intercept form. We now look at a final linear form, the so-called **standard form**.

Goats and chickens

// startup exploration

The chickens got into the goat pen on François's farm! When François asked Hermine how many of each animal were mixed up, she coyly replied that there were a total of 44 feet in the pen.

If the pen contains only (anatomically normal) goats and chickens, how many of each animal could there be inside? How many different combinations are possible?

We know that an anatomically normal goat has 4 feet, and so if the pen contains x goats, they account for $4x$ feet. Similarly, an anatomically normal chicken has 2 feet. If there are y chickens in the pen they account for $2y$ feet. Finally, we know the total number of feet. So, we have the equation

$$4x + 2y = 44.$$

We can tinker with a little guess and check to find combinations that work. For instance there could be 8 goats and 6 chickens, since $4(8) + 2(6) = 44$. If you find a few combinations, you might start to notice some patterns in the solutions (of which there are 12).

We'll come back to problems like "goats and chickens" in the future. For now, we'll simply point out that our equation above is an example of a linear equation in standard form.

Standard form

If A , B , and C are integers where A and B are not both zero, then

$$Ax + By = C$$

is a linear equation in standard form.

Note that this is not a " $y =$ " form. This has its challenges and its benefits. In some situations, standard form equations require a bit of work before they are very useful.⁴ For example, A , B , and C don't tell us much about

⁴ Standard form: the fixer-upper form of a linear equation.

the line, especially when compared to point-slope form and slope-intercept form (in which important features of the line are visible right there in the equation).

A benefit of this different format is that any line – whether it is a function or not – has a standard form equation. The definition says that A and B can't *both* be zero (at the same time), but one or the other of them could be zero. This is how we get equations for vertical and horizontal lines. Vertical lines are of the form $Ax = C$ and horizontal lines are of the form $By = C$.

Example 7.12

Are the following equations written in standard form? If not, why not?

(a) $3y = 4x + 2$

(b) $-2x + 6y = 17$

(c) $3x = 15$

(d) $\frac{1}{2}x + 5y = -12$

(e) $2x - 4(y - 2) = 8$

(f) $10 = 3x - 3y$

Solution: Lines (a), (d), and (e) are *not* in standard form. To be in standard form, the x term and the y term must be on the same side of the equal sign; line (a) violates this rule. Standard form states that the coefficients must be integers; line (d) violates this requirement. Standard form has no parentheses; line (e) violates this.

Lines (b), (c), and (f) are in standard form. Note that (c) has $B = 0$, but that's OK; this is the equation for a vertical line. Line (f) is written with the variables on the right-hand side, but otherwise it meets the requirements.

7.4.1 Converting to and from standard form

If we are given a line in some other form, then we can convert to standard form using the properties of equality. This is an application of “transforming formulas” that we saw in section 5.8.

Example 7.13

Convert the lines $y = \frac{3}{2}x - \frac{4}{3}$ and $y = \frac{1}{4}(x - 8) + 2$ to standard form.

Solution: The first line is in slope-intercept form, but there's a fraction in there which is no good for standard form.

$$y = \frac{3}{2}x - \frac{4}{3}$$

$$6y = 9x - 8 \quad \text{multiply through by 6 to eliminate fractions}$$

$$-9x + 6y = -8 \quad \text{put } x \text{ and } y \text{ on the same side}$$

The second line is in point-slope form, and we can start out by doing the distributive property.

$$y = \frac{1}{4}(x - 8) + 2$$

$$y = \frac{1}{4}x - 2 + 2 \quad \text{distributive property}$$

$$y = \frac{1}{4}x \quad \text{combine like terms}$$

$$4y = x \quad \text{multiply through by 4 to eliminate fractions}$$

$$-x + 4y = 0 \quad \text{put } x \text{ and } y \text{ on the same side}$$

Of course, we can convert in the other direction as well, if needed – at least to slope-intercept form (to point-slope form, not so much).

Example 7.14

Convert the line $6x - 14y = 21$ to slope-intercept form.

Solution: All we need to do is transform this equation to isolate y .

$$6x - 14y = 21$$

$$-14y = -6x + 21 \quad \text{SPOE}$$

$$y = \frac{-6}{-14}x + \frac{21}{-14} \quad \text{DPOE}$$

$$y = \frac{3}{7}x - \frac{3}{2} \quad \text{simplify fractions}$$

7.4.2 Simplified standard form

It turns out that we can generate infinitely many equivalent equations in standard form just by multiplying both sides of a given form by an integer. For example, if we have the line $3x - y = 5$, we might through by 5 and

get $15x - 5y = 25$. Or we could multiply through by -1 and get $-3x + y = -5$. All of these lines are the same because we've applied a property of equality (MPOE) to create them.

Since there are so many possible ways to write an equation in standard form, we have a "standardized" version of it. Standardized standard form sounds like an oxymoron, so we call it "simplified standard form".

There are two additional criteria for simplified standard form.⁵ In addition to the rules given in the definition above, we require that A is non-negative (it must be greater than or equal to zero). We also require that A , B , and C share no common factors between them other than 1.

The second of these new rules is the trickier one. The line $10x + 6y = 12$ violates this second requirement since each term is divisible by 2. To simplify this, we divide all of the terms by 2 and have an equation that is in simplified standard form: $5x + 3y = 6$. Note that here 3 and 6 share a common factor, but we're OK since the only factor that *all three* numbers share is 1.

The line we wrote to describe the startup exploration $4x + 2y = 44$ does not meet the requirements for simplified standard form. We must divide through by 2 to fix this: $2x + y = 22$ (does this new equation have an interpretation in the context of the problem?).

Example 7.15

Write an equation, in simplified standard form, for the line that goes through the points $(-3, 6)$ and $(7, 12)$.

Solution: We'll start out as usual, finding the slope between these points:

$$m = \frac{\Delta y}{\Delta x} = \frac{12 - 6}{7 - (-3)} = \frac{6}{10} = \frac{3}{5}$$

Then, we can use either point to write a line in point-slope form:

$$y = \frac{3}{5}(x + 3) + 6.$$

Since we need integer coefficients we can multiply through by 5 at the start. This will avoid having to

⁵ These rules are just a convention, and they are somewhat arbitrary. In future math courses, many of these restrictions about standard form, for example what kinds of coefficients are allowed, may be relaxed.

carry out the distributive property with that fraction.

$$\begin{array}{ll}
 y = \frac{3}{5}(x + 3) + 6 & \\
 5y = 3(x + 3) + 30 & \text{multiply through by 5} \\
 5y = 3x + 9 + 30 & \text{distributive property} \\
 5y = 3x + 39 & \text{combine like terms} \\
 -3x + 5y = 39 & \text{SPOE: subtract } 3x \text{ from both sides} \\
 3x - 5y = -39 & \text{multiply through by } -1
 \end{array}$$

This last step is required because simplified standard form requires that the coefficient of the x -term be non-negative.

7.4.3 Graphing standard form

To graph standard form on most graphing calculators, the only thing we can do is convert to slope-intercept form by solving for y . Many online tools (like Desmos) allow graphing in standard form without any transforming required on our part.

When graphing by hand, there is a clever alternative technique!

Example 7.16

Graph the line $3x - 2y = 12$.

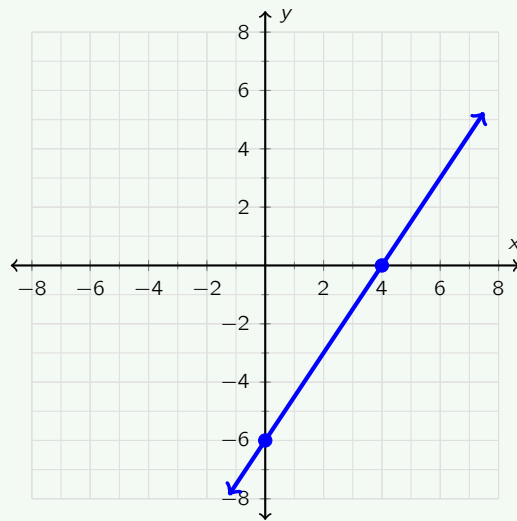
Solution: We can always plot some points by substituting in some values for x or y . We could choose random values, but here's a clever idea: plugging in zero will eliminate a term and simplify our calculations! If we let $x = 0$, then:

$$3(0) - 2y = 12 \quad \implies \quad -2y = 12$$

That's a one-step equation, and $y = -6$. If we plug in $y = 0$, we have:

$$3x - 2(0) = 12 \quad \implies \quad 3x = 12,$$

and so $x = 4$. So, we have found two points on the line: the y -intercept $(0, -6)$, and the x -intercept $(4, 0)$. That's enough to draw the graph!



The only time this approach doesn't work is when the y -intercept and the x -intercept are the same. What does that sort of a line look like? How can we spot such a line from its standard form equation?

Speaking of x -intercepts, they haven't gotten much of our attention so far. Our work with linear equations has highlighted the importance of the y -intercept. We'll learn much more about the significance of x -intercepts in future chapters.

7.5 Parallel and perpendicular lines

This section is dedicated to a geometric connection to linear functions: parallel and perpendicular lines. Our goals for now are to learn how to tell if two lines are parallel or perpendicular and how to write equations for a line that is parallel or perpendicular to a given line. In the future (like in a geometry class), you might use concepts from this lesson to prove some geometric concepts algebraically.

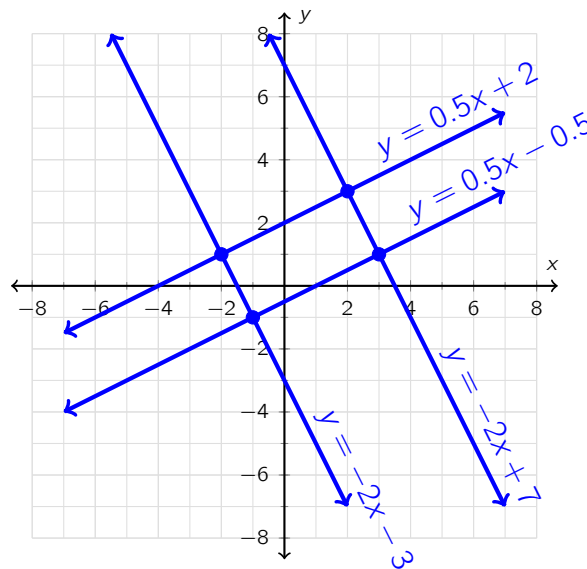
Graphical rectangle

// startup exploration

The four points below form the vertices of a rectangle. Graph the rectangle and write four equations in slope-intercept form for the lines containing the four sides of the rectangle.

$$(2, 3) \quad (3, 1) \quad (-1, -1) \quad (-2, 1)$$

We won't go into the details of how to write these equations. Those details are explained in earlier sections. The key thing to study in this example is how the lines are related to each other.



7.5.1 Parallel lines

Parallel lines never intersect. They exist in the same plane and stay the same distance apart through infinity... and beyond! In the rectangle we drew as part of the startup exploration, opposite sides are parallel. Examine the slopes of each pair of parallel lines. What do you notice?

Algebraically, parallel lines have the *same slope* and *different y-intercepts*. Note that it is not sufficient to just look at whether two lines have the same slope. If two lines have the same slope and also have the same y-intercept, then they aren't actually *two lines*. Those are the same line. . . and you can't be parallel to yourself!⁶

In our rectangle, opposite sides of the rectangle are parallel. For example, the two lines that have positive slope:

$$y = \frac{1}{2}x + 2 \quad \text{and} \quad y = \frac{1}{2}x - \frac{1}{2}.$$

Notice that the slopes of these two lines are the same and that their y-intercepts are different. The same is true for the other pair of lines.

Example 7.17

Which of these lines, if any, are parallel to each other?

(a) $y = \frac{3}{4}x + 5$

(b) $y = \frac{3}{4}(x + 8) - 1$

(c) $y = \frac{3}{4}(x - 12) + 1$

(d) $4x - 3y = 9$

Solution: Parallel lines will have the same slope and different y-intercepts, so we should extract this information. We'll have to do the a bit of work first, converting all of the lines to slope-intercept form:

(a) $y = \frac{3}{4}x + 5$

(b) $y = \frac{3}{4}x + 5$

(c) $y = \frac{3}{4}x - 8$

(d) $y = \frac{4}{3}x - 3$

From this we can see that (a) and (b) are actually the same line, and that this line is parallel to line (c). Line (d) has a different slope, so it is not parallel to any of the others.

Example 7.18

Write an equation for the line that contains $(-2, 1)$ and is parallel to the line $y = -3(x - 4) + 7$.

Solution: Parallel lines have the same slope, so the slope of our line must be -3 , just like the slope of the given line. Since it doesn't specify which form to use, we can write our answer in point-slope form:

$$y = -3(x + 2) + 1.$$

⁶ Don't start with the interpretive dance about this. You can't be parallel to yourself, no matter how you hold your arms.

The amount of work we will have to do for questions like the ones in this chapter will depend on the format the equations take. We will have to work a little harder if the lines are given in standard form or when we are asked to write our answer in standard form.

Example 7.19

Write an equation *in simplified standard form* for the line that contains $(2, -1)$ and is parallel to the line $5x + 3y = 2$.

Solution: First, we need to remember that we're after a point and a slope to write the equation of our line. The line we are given is in standard form, so we'll have to convert it to slope-intercept by solving for y .

Using the techniques from section 7.4, we convert $5x + 3y = 2$ into

$$y = -\frac{5}{3}x + \frac{2}{3},$$

and we can see that the slope of the line is $-\frac{5}{3}$.

Our line will have the same slope as the given line (since our line is parallel to it), and we know a point on the line. So, we can write an equation for our line in point-slope form:

$$y = -\frac{5}{3}(x - 2) - 1.$$

But, we are asked for an equation in standard form, so we have to convert.

First we distribute and combine like terms, getting a line in slope-intercept form:

$$y = -\frac{5}{3}x + \frac{7}{3}.$$

Then, we get rid of the fractions and rearrange into standard form:

$$5x + 3y = 7 \quad \text{Phew!}$$

And now for some good news: There is a clever shortcut that we can use when searching for a standard form equation for a line parallel to a line given in standard form.

You may have noticed that when we convert a standard form equation into a slope-intercept equation, we are doing the same procedure every time:

$Ax + By = C$	standard form
$By = -Ax + C$	SPOE: subtract Ax from both sides
$y = -\frac{A}{B}x + \frac{C}{B}$	DPOE: divide through by B to isolate y

This means that the slope of the line is always $-\frac{A}{B}$. That's a recipe for finding the slope of a line in standard form.

Since A and B are the only parts of the standard form equation that relate to slope, and given that parallel lines have the *same slope*, we can reason that all parallel lines have the same A and B values in standard form.

Look at the previous example: the given line and our answer have the same A and B values, but with different C values! In other words, any line parallel to $5x + 3y = 2$ is going to look like $5x + 3y = C$. In the example, we have to ensure that our line goes through the point $(2, -1)$, so we just need to figure out what C value will accomplish that goal.

How can we figure out what C value will make the line $5x + 3y = C$ go through the given point? We can substitute in these values for x and y and solve the equation to identify C ! Observe:

$$\begin{array}{ll} 5x + 3y = C & \text{our almost-there equation} \\ 5(2) + 3(-1) = C & \text{substitute in a known point on the line} \\ 7 = C & \text{simplify the left-hand side using the order of operations} \end{array}$$

So, our standard form equation is $5x + 3y = 7$.

7.5.2 Perpendicular lines

Perpendicular lines intersect at right angles (that is, 90° angles). In the rectangle we drew in the startup exploration, adjacent sides are perpendicular. Examine the slopes of each pair of perpendicular lines. What do you notice?

In our rectangle, the lines that meet in the second quadrant are the lines

$$y = \frac{1}{2}x + 2 \quad \text{and} \quad y = -2x - 3.$$

Notice that the slopes of these two lines are both opposites and reciprocals. This is true in general: perpendicular lines have slopes that are *both opposites and reciprocals* of each other. This means that the slopes of perpendicular lines will always have a product of 1.

Example 7.20

Which of these lines are perpendicular to each other?

- | | |
|-----------------------------------|-----------------------------------|
| (a) $y = \frac{3}{4}x + 5$ | (b) $y = -\frac{3}{4}(x + 8) - 1$ |
| (c) $y = \frac{4}{3}(x - 12) + 1$ | (d) $4x + 3y = 9$ |

Solution: The key is to look for slopes with opposite signs and then to look for reciprocals. First, we ought to convert line (d) into a form that reveals its slope:

$$(d) y = -\frac{4}{3}x + 3$$

This means that line (a) is perpendicular to line (d), since $\frac{3}{4} \cdot -\frac{4}{3} = -1$. Also, line (b) is perpendicular to line (c) since $-\frac{3}{4} \cdot \frac{4}{3} = -1$.

Example 7.21

Write an equation for the line that contains $(-2, 1)$ and is perpendicular to the line $y = -3(x - 4) + 7$.

Solution: The given slope is -3 , the perpendicular slope must be the opposite, reciprocal slope: $\frac{1}{3}$. The point is $(-2, 1)$, and since the format of the equation was not specified, we can answer in point-slope form:

$$y = \frac{1}{3}(x + 2) + 1.$$

Again, the amount of work we will have to do for questions like the ones in this chapter will depend on the format the equations take.

Example 7.22

Write an equation *in simplified standard form* for the line that contains $(2, -1)$ and is perpendicular to the line $5x + 3y = 2$.

Partial solution: The line we seek is $3x - 5y = 11$. Above, we discussed a clever shortcut for working with parallel lines in standard form. Can you invent a similar shortcut for working with perpendicular lines in standard form? Use the given solution to test out your ideas!

» Chapter summary «

In this chapter we learned a great deal about the behavior of linear functions. We derived several methods for finding the slope of a line, and we extended this idea by deriving a number of different forms for writing a linear function.

We closed the chapter by observing how two lines might interact: We used the slopes of lines to determine whether those lines were parallel, perpendicular, or something in between. In the next chapter we continue to study the interactions between two lines and learn a further set of equation-solving tools in the process. Onward!

The property of the straight line as being the shortest connection between two points can be transferred to curved surfaces. . . ; on the surface of the sphere the great circles play the role of the shortest line of connection. . . . Yet. . . all great circles of the sphere intersect and therefore there are no parallels among these “straight lines”.

Hans Reichenbach, German philosopher of science

Chapter 8

Linear systems

Many important natural and manmade phenomena can be represented using mathematical functions. We may have to work quite hard to figure out the best way to “mathematize” a real-world context, but once we do we can very often learn something about the real-world from studying our mathematical representation.

In this chapter we will learn a set of techniques for understanding how functions interact. We will continue our focus on linear functions, though many of the ideas can be applied to other situations. In the last section, we’ll explore some classic applications of the tools and techniques that we learn in the earlier sections.

8.1 Mathematical modeling

Mathematics has many applications to the sciences, medicine, business, and economics. Often, though, we have to “translate” some problem in the real world so that it can be interpreted mathematically. It is usually necessary to break down the situation into a set of variables and equations, then look at them all together and see how they interact with each other. Doing this is called **mathematical modeling**.

We can model weather patterns, transportation infrastructure, airline schedules, missile guidance systems, factory efficiency, the financials of a business – and many other phenomena – in this way.

I ♥ Cheeseville Zoo

// startup exploration

After YeardleighCorp took over and renovated the failing Cheeseville Zoo, Yeardleigh launched a marketing campaign to announce the grand reopening. She spent \$270 on a t-shirt design. For production, she found a supplier who sold her blank cotton shirts in a variety of sizes and colors for \$3.00 per shirt and a printer who would print the logo in full color for \$1.50 per shirt. If Yeardleigh sold the shirts for

\$12.00 each, how many would she have to sell before she started to make a profit on t-shirt sales?

8.1.1 Simultaneous equations

One approach to mathematical modeling is to model problems using multiple functions and then explore how those functions relate to one another.

System of equations

A set of two or more equations with the same variables is called a **system of equations**. They are also sometimes called *simultaneous equations*.

In the startup exploration, the amount of money Yearleigh earns will depend on how many t-shirts she buys and sells. So, suppose we let x , the independent variable, represent the number of t-shirts. If we let y , the dependent variable, represent the amount of money, then we can write two equations: one that describes the how much the t-shirts cost, and another that describes how much money the t-shirts generate for the company.

Yearleigh earns \$12.00 for every t-shirt she sells, so one equation involved with the scenario is

$$y = 12.00x \quad \text{Yearleigh's income equation.}$$

She has already spent \$270 on her design, and every t-shirt she prints costs an additional \$4.50 (that's \$3.00 for the shirt and \$1.50 for printing). So, another equation in the scenario is

$$y = 270 + 4.50x \quad \text{Yearleigh's cost equation.}$$

We now have two linear equations. Observe what happens when we graph these two equations on the same set of axes (fig. 8.1). What does this graph tell us about the problem?

On the graph in the figure, the red line is Yearleigh's cost equation and the green line is Yearleigh's income equation. The x -coordinate shows the independent variable (number of shirts), and the y -coordinate shows the dependent variable (amount of money spent or earned).

Notice that these two lines intersect (in other words: they cross over one another). Does this intersection point have any special significance?

The intersection shows that for a certain number of t-shirts, cost and income are the same. The points to the left of the intersection show where cost is greater than income, points to the right of the intersection show where income is greater than cost. So that intersection point is really important for the problem!

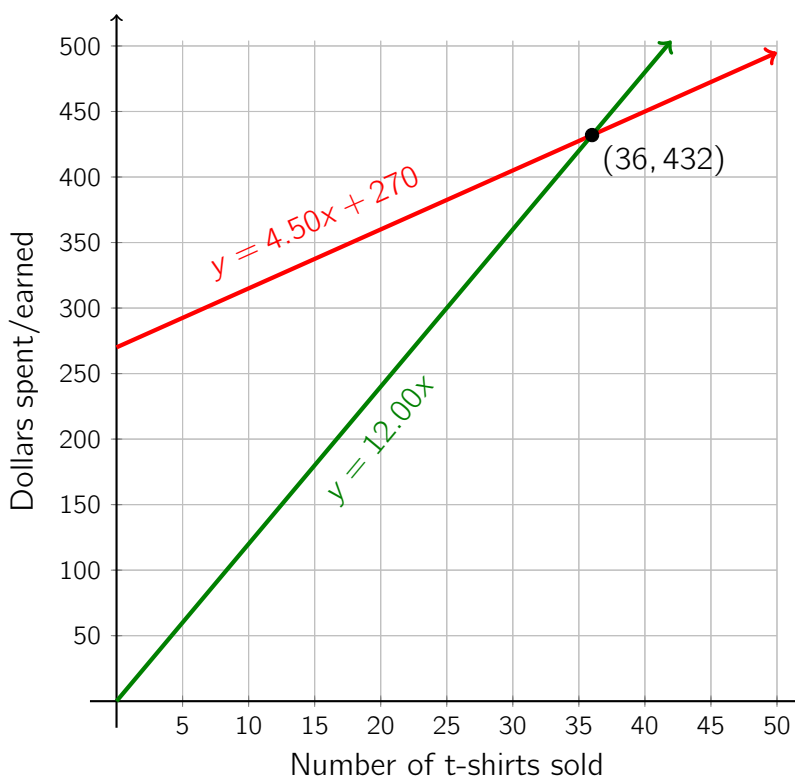


Figure 8.1: Graph showing both Yeardeigh's cost and income equations.

The coordinates of the intersection point are (36, 432). So, if Yeardeigh sells 36 t-shirts, she will make 432 in income. . . and it will also cost her exactly that much to produce the t-shirts. If she sells more than 36 t-shirts, she will earn more than it costs to make the shirts.

We have found Yeardeigh's *break-even point*, a very important idea in economics and business. It is the point where a company's costs and income are the same. Before this point, the company will be losing money. After this point, the company will start to make a profit. In the context of our problem, Yeardeigh will make a profit on t-shirt sales only if she sells more than 36 shirts.

Writing a system of equations

When we have two (or more) equations that we want to treat as a system, we usually write them stacked vertically, and we very often add a big curly brace (usually on the left-hand side only) to show that they are meant to be considered a group. In the case of the startup exploration, we write:

$$\begin{cases} y = 12.00x \\ y = 4.50x + 270 \end{cases}$$

8.1.2 Solutions to systems

Imagining intersections**// startup exploration**

Imagine two lines drawn in the coordinate plane. In the previous section, we saw that these lines might intersect at a single point. Is that the only possibility? Could they intersect in more than one point? Could they not intersect at all?

What about if we have three lines in the plane? Describe the ways they might (or might not) intersect.

What about a parabola and a line? What about two parabolas?

In the startup exploration, we are visualizing the possible solutions to various systems of equations.

Solution to a system of equations

The set of all points that are common to all equations in the system. Graphically, these are the points where all of the graphs of the functions intersect. Note: a system of equations may have no solution.

We will focus in this course on systems of two linear equations, and there are three ways in which two lines in the plane might interact:

- The lines might intersect at a single point. Such a system has a unique solution.
- The lines might be parallel and not intersect at all. In this case, the system has no solution.
- The lines might overlap completely (that is, the “two lines” might actually be the same line). In this case, the system has infinitely many solutions: every point one line is also on the other line – they’re the same line!

Since we have straight lines, these are the only possibilities. For example, a system of linear equations cannot have “exactly two solutions”. (Can you explain why not?) Later (in Algebra 2, for instance), systems will get more complex. Then, we won’t have to limit ourselves to linear equations! A system of two quadratic equations, for example, might have no solution, one solution, two solutions, or infinitely many solutions. (Can you picture how each of those might happen? Look back at some of the graphs we drew in chapter 3 and chapter 4.)

8.1.3 Checking a solution

Suppose we have a point that we think might be a solution to a system of equations. How can we check?

Example 8.1

Determine whether or not the point $(-1, 5)$ is a solution to each of the given systems.

$$\text{System A: } \begin{cases} x + y = 4 \\ x = -1 \end{cases} \quad \text{System B: } \begin{cases} y = -x + 4 \\ y = -\frac{1}{5}x \end{cases}$$

To check whether the point is a solution to System A, we'll substitute x and y into both equations. If the point makes both equations true, then the point is a solution to the system.

Checking System A, we have:

$$\begin{array}{rclcl} x + y = 4 & \text{and} & x = -1 \\ (-1) + 5 \stackrel{?}{=} 4 & & -1 \stackrel{\checkmark}{=} -1 \\ 4 \stackrel{\checkmark}{=} 4 & & \end{array}$$

The given point makes both of the equations true, so yes, the point $(-1, 5)$ is a solution to System A.

Checking System B, we have:

$$\begin{array}{rclcl} y = -x + 4 & \text{and} & y = -\frac{1}{5}x \\ 5 \stackrel{?}{=} -(-1) + 4 & & 5 \stackrel{?}{=} -\frac{1}{5}(-1) \\ 5 \stackrel{\checkmark}{=} 5 & & 5 \neq \frac{1}{5} \end{array}$$

The given point works for the first equation but not for the second. To be a solution, the point has to satisfy both equations, so no, the point $(-1, 5)$ is not a solution to System B.

8.1.4 Writing solutions

When we are solving equations (meaning one equation at a time) we can use set notation to record our solutions. Since the solution to a system involves two numbers, we have to be mindful about how we write our answers.

Two cases are pretty straightforward. If the system has a unique point as its solution, we can write that point as the solution set. For example, the solution to System A in the previous example is $\mathcal{S} = \{(-1, 5)\}$. Note that we've just enclosed the point $(-1, 5)$ inside the curly braces to show that the set includes that point.

The second easy case is when a system has no solution. Then, we can use the same approach for when a single equation has no solution: $\mathcal{S} = \{ \}$ or $\mathcal{S} = \emptyset$.

The trickier case is when two lines overlap completely. We can't say that the solutions in the case are "all real

numbers" (as we said for a single equation). In this case we say that $\mathcal{S} = \{\text{all the points on the line } y = 2x\}$ or write out the sentence.¹

8.1.5 Techniques for solving linear systems

Over the next few sections, we will study three different techniques for solving systems of linear equations. A fourth technique, which we only mention here, will be an important topic in Algebra 2.

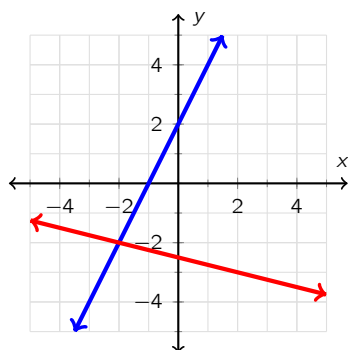
1. **Graphing the system.** To solve a system by graphing, we graph the equations on the same set of coordinate axes and look for the intersection point. If we are lucky, we'll get "nice" points with integers for their coordinates. Graphing with technology works a little better, since trace functionality can give us good estimates for the coordinates if they are not integers.
2. **Collapsing the system by substitution.** To collapse a system is to do something algebraically that fuses two equations into one. If done correctly, we can turn two equations with two variables into one equation in one variable! One method uses the property of substitution to achieve this.
3. **Collapsing the system by elimination.** This is an alternative approach to collapsing a system, that is based on a mathematical structure called a matrix (although we won't go into detail about matrix operations).
4. **Modeling the system as a matrix.** A *matrix* (plural: matrices) is a rectangular arrangement of numbers. Matrix operations are good for solving systems with more than two variables, which we won't see in Algebra 1. We will learn much more about matrices (in general) in Algebra 2, along with a number of ways of using them to solve systems.

The different approaches have pros and cons, and you may find that you prefer a particular approach. Generally speaking, you should solve problems using whatever techniques make the most sense to you, and some assignments will be open and allow you to choose your own method. Other assignments may ask you to use a specific technique, though, so it's important to read directions carefully.

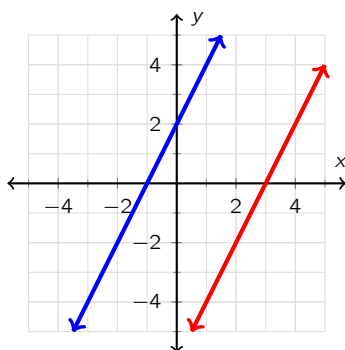
¹ An alternative is to use set notation with a fancy notation extension. We write $\mathcal{S} = \{(x, y) \mid x, y \in \mathbb{R} \text{ and } y = 2x\}$. This is called *set builder notation*, and it uses the vertical bar to mean "such that". So, this math sentence reads " \mathcal{S} is the set of all points (x, y) such that x and y are real numbers and $y = 2x$."

8.2 Solving by graphing

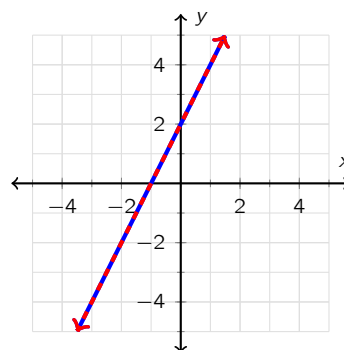
We solved the startup exploration at the start of section 8.1 by graphing two equations on the same set of axes. That's the idea behind this solution technique. We graph the two lines given in our system on the same set of axes, and then we look to see how the two lines interact.



One solution.



No solution:
parallel lines.



Infinitely many solutions:
the two lines are the same.

We graphed quite a few lines in the previous chapter, and all of those skills come into play now. When graphing by hand, it's always a good idea to use graph paper but remember that the graph itself is not the answer. The graph is the work that helps us to find the solution to the system. This means that we don't always need to create a "high quality graph". Of course, the graph must be of high-enough quality that you can accurately find the intersection point!

Solving by graphing can be made easier with the right technology, though technology also imposes some restrictions. If you graph using a calculator or an online graphing tool, the equation may have to be in a $y =$ form (either slope-intercept form or point-slope form). Also, technology might not help if the answers don't come out "nice". It may be hard to identify some rational and irrational numbers from the decimal that we get as output from the calculator!

These restrictions are avoided with the algebraic methods we will learn for solving a system (in the upcoming sections). Still, graphing will often come in handy as one good way to check our answers.

8.3 Solving by substitution

Ugly graphing

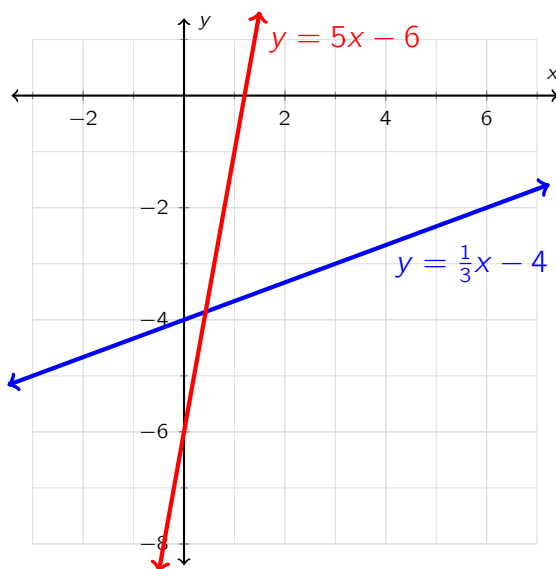
// startup exploration

Solve the following system by graphing.

$$\begin{cases} y = \frac{1}{3}x - 4 \\ y = 5x - 6 \end{cases}$$

What are the challenges that arise when solving this system graphically?

The challenge we uncover in the startup exploration is that the intersection point does not have integer coordinates. The x -coordinate is between 0 and 1, and the y -coordinate is between -3 and -4 , but beyond that it is quite hard to find an exact intersection, even using graphing technology (since the decimal values of the intersection point are not all that “nice”).



In this section we will explore the first of two algebraic methods for solving systems of equations. These algebraic approaches will avoid the challenges demonstrated by the startup exploration. The first algebraic technique is the *substitution method*.

We've been substituting numbers for variables for a long time. For example, given $f(x) = 3x + 1$, we know how to find $f(5)$ by substituting 5 for x in the equation:

$$f(x) = 3x + 1 \implies f(5) = 3(5) + 1 = 15 + 1 = 16$$

The clever idea behind solving a system using substitution is that rather than just replacing variables with letters, we can substitute one algebraic expression for another.

Substitution method

A method for solving a system of equations that involves solving one of the equations for one variable and substituting the resulting expression into a different equation.

Example 8.2

Solve the following system using substitution.

$$\begin{cases} y = 3x \\ y = 2x - 4 \end{cases}$$

Solution: We are looking for a point where the two equations have the same x value and the same y value. So, let's assume that the y 's in these two equations are equal. If the two y 's are equal, then it means " $3x$ " is equal to " $2x - 4$ " and that is a Level 4 linear equation (one with variables on both sides)!

$$3x = 2x - 4 \quad \text{substitution}$$

$$x = -4 \quad \text{SPOE: subtract } 2x \text{ from both sides}$$

So, if the two y values are equal, as we assumed they were, then it must be that $x = -4$. We just found the x -coordinate of the intersection point! How can we find the y -coordinate?

Notice that the first equation says $y = 3x$. So, if $x = -4$, it must be that $y = 3x = 3(-4) = -12$. So, the solution to the system is the point $(-4, -12)$. Or we can write $\mathcal{S} = \{(-4, -12)\}$.

To check our work, we can either graph the system or plug the x and y values into the other equation to verify that equality holds.

$$y = 2x - 4 \quad \text{the second equation we were given}$$

$$-12 \stackrel{?}{=} 2(-4) - 4 \quad \text{substitute our proposed solution}$$

$$-12 \stackrel{?}{=} -8 - 4 \quad \text{simplify right-hand side}$$

$$-12 \stackrel{\checkmark}{=} -12$$

When we assume that the two y 's were equal – as they must be if we have found a solution to the system – then substitution means that we can replace one y with the other.

We can carry this out for the problem in the startup exploration. Since we have two equations in $y =$ format, and since any point that satisfies the system has a y -coordinate that satisfies both equations, we can set the

equations equal to one another:

$$\frac{1}{3}x - 4 = 5x - 6 \quad \text{substitution}$$

$$2 = \frac{14}{3}x \quad \text{SPOE: variable terms to the right, constants to the left}$$

$$\frac{6}{14} = x \quad \text{MPOE}$$

$$\frac{3}{7} = x \quad \text{simplify the fraction}$$

Aha! Even though this x -value is inconvenient to graph by hand (since it's not an integer) or using technology (since it's not a very tidy decimal)... it emerges with no trouble using substitution.

To find the y -coordinate, we use our newly-discovered x -coordinate and one of the two original equations. Why not pick the second equation to minimize the fractions?

$$y = 5x - 6 \quad \text{the second equation}$$

$$y = 5\left(\frac{3}{7}\right) - 6 \quad \text{substitute in our new } x\text{-value}$$

$$y = \frac{15}{7} - 6 \quad \text{multiply fractions}$$

$$y = \frac{15}{7} - \frac{42}{7} \quad \text{write with a common denominator}$$

$$y = \frac{-27}{7} \quad \text{simplify the fraction}$$

To check our work, we can substitute our candidate values for x and y back into the first equation. Note that since we used the second equation already to figure out the answer, we use the first equation to check it.

$$y = \frac{1}{3}x - 4 \quad \text{the first equation}$$

$$\frac{-27}{7} \stackrel{?}{=} \frac{1}{3}\left(\frac{3}{7}\right) - 4 \quad \text{substitute in our candidate solution}$$

$$\frac{-27}{7} \stackrel{?}{=} \frac{3}{21} - 4 \quad \text{multiply fractions}$$

$$\frac{-27}{7} \stackrel{?}{=} \frac{3}{21} - \frac{84}{21} \quad \text{write with a common denominator}$$

$$\frac{-27}{7} \stackrel{?}{=} \frac{-81}{21} \quad \text{subtract fractions}$$

$$\frac{-27}{7} \stackrel{\checkmark}{=} \frac{-27}{7} \quad \text{simplify the right-hand side}$$

So, we have found our solution $\mathcal{S} = \left\{\left(\frac{3}{7}, -\frac{27}{7}\right)\right\}$.

In the two examples we have seen so far, both equations have been in $y =$ format. This is not a requirement for using substitution, as the following two examples will demonstrate.

Example 8.3

Solve the following system using substitution.

$$\begin{cases} y = 3x - 5 \\ x - y = 4 \end{cases}$$

Solution: Here, we have one equation in point-intercept form and one equation in standard form. No problem! We simply substitute y from the first equation for y in the second equation. It's a good idea to use parentheses when doing the substitution: watch what happens with the negative signs in this example.

$$\begin{array}{ll} x - y = 4 & \text{start with the second equation} \\ x - (3x - 5) = 4 & \text{substitute } y \text{ from first equation, using parentheses} \\ x - 3x + 5 = 4 & \text{distributive property} \\ -2x + 5 = 4 & \text{combine like terms} \\ -2x = -1 & \text{SPOE} \\ x = \frac{1}{2} & \text{DPOE} \end{array}$$

This tells us that if the y 's are equal, then it must be that $x = \frac{1}{2}$. To find the y -coordinate, we substitute the x value we just found back into one of the original equations. Let's choose the second equation, which seems a bit easier:

$$\begin{array}{ll} x - y = 4 & \text{the second equation} \\ \frac{1}{2} - y = 4 & \text{substitute } x = \frac{1}{2}, \text{ which we just computed} \\ -y = \frac{7}{2} & \text{SPOE} \\ y = -\frac{7}{2} & \text{MPOE} \end{array}$$

So, $\mathcal{S} = \left\{ \left(\frac{1}{2}, -\frac{7}{2} \right) \right\}$. To check we can plug these values back into the other equation (the first equation, in this case, since we used the second equation to find y).

$$\begin{array}{ll} y = 3x - 5 & \text{the first equation} \\ -\frac{7}{2} \stackrel{?}{=} 3 \left(\frac{1}{2} \right) - 5 & \text{substitute our proposed solution} \\ -\frac{7}{2} \stackrel{?}{=} \frac{3}{2} - 5 & \text{simplify on the right-hand side} \\ -\frac{7}{2} \stackrel{?}{=} \frac{3}{2} - \frac{10}{2} \\ -\frac{7}{2} \stackrel{\checkmark}{=} \frac{-7}{2} \end{array}$$

Notice that after our substitution step, we've been getting linear equations like the ones we saw in chapter 5.

In the first example in this section, we got an equation with the variable x on both sides. Recall that when equations have variables on both sides, we sometimes find that the equation had one unique solution, sometimes "no solution", and sometimes "infinitely many solutions". Do you see how these correspond to the different possible solutions to a linear system?

The key to the substitution method is that we need at least one equation to be in $x =$ or $y =$ form. What if neither equation is in this form? We can use our skills at transforming formulas to create what we want!

Example 8.4

Solve the system by substitution.

$$\begin{cases} 3x - 12y = 15 \\ x - 4y = 7 \end{cases}$$

Solution: We could transform either equation, but notice that there's less work to do in the second equation, since the x term has coefficient 1. So, let's transform the second equation to $x = 7 + 4y$ using APOE. Then, we can substitute that expression into the first equation.

$$\begin{array}{ll} 3x - 12y = 15 & \text{the first equation} \\ 3(7 + 4y) - 12y = 15 & \text{substitute our expression for } x \\ 21 + 12y - 12y = 15 & \text{distributive property on the left-hand side} \\ 21 = 15 & \text{D'oh.} \end{array}$$

We have arrived in one of those impossible situations that tells us the equation we created has no solution. In this case, it means that the system has no solution. In other words, these lines were parallel all along.
 $\mathcal{S} = \emptyset$

That was the first "special case". The other special case is when the two lines completely overlap. What do you suppose will happen when we solve the system that lets us know that we're in this special case?

8.4 Solving by elimination

Ugly substitution

// startup exploration

Study the following system of linear equations.

$$\begin{cases} 4x + 3y = 22 \\ 3x + 5y = 11 \end{cases}$$

What challenges would you face if you tried to solve this system by graphing? What challenges would you face if you tried to solve it by substitution?

The equations in the startup exploration would require quite a bit of work to graph accurately by hand (even using our x - and y -intercept strategy for graphing lines in standard form), and we'd likewise have to do some work before one could be substituted into the other. The second algebraic approach to solving a system, called the **elimination method**, addresses these concerns.

This technique is particularly helpful when the given equations are both in standard form. The premise behind elimination is that we use the properties of equality to, say, add two equations together. Recall that the addition property of equality says that we can add the same quantity to both sides of an equation. If we know two things are equal, we can add one thing to one side of an equation and the other thing to the other side, and still maintain equality.

Example 8.5

Solve the following system using elimination.

$$\begin{cases} y = 3 \\ 5x - y = 12 \end{cases}$$

Solution: The first equation tells us that y is equal to 3. Of course, we could add y to both sides of an equation... or we could add 3 to both sides of an equation... but since we know y and 3 are equal we can add y to one side of an equation and 3 to the other side of that equation!

In other words, we can add y to the left-hand side of the second equation we are given while simultaneously add 3 to the right-hand side. Since we know $y = 3$, we know equality will be maintained! Watch what

happens:

$$\begin{array}{ll}
 5x - y = 12 & \text{the second equation} \\
 5x - y + y = 12 + 3 & \text{add on the first equation} \\
 5x = 15 & \text{combine like terms: } y \text{ has been eliminated!} \\
 x = 3 & \text{DPOE}
 \end{array}$$

Having found x , and having y given in the original system, we have $\mathcal{S} = \{(3, 3)\}$.

Elimination method

A method for solving a system of equations that involves adding two equations to eliminate a variable.

The method is called “elimination” because it eliminates a variable.²

Example 8.6

Solve the following system by elimination.

$$\begin{cases} 2x + 2y = 7 \\ 3x - 2y = 8 \end{cases}$$

Solution: Notice how the coefficients on the y terms are opposites. If we add these two equations together, the y 's will be eliminated:

$$\begin{array}{r}
 2x + 2y = 7 \\
 + \quad 3x - 2y = 8 \\
 \hline
 5x + 0y = 15 \\
 5x = 15 \\
 x = 3
 \end{array}$$

We then substitute $x = 3$ back into one of the equations to find $y = \frac{1}{2}$. So, $\mathcal{S} = \{(3, \frac{1}{2})\}$.

The key to the elimination approach is to “add the equations together” so that one variable vanishes, just as the y disappeared in the third step of the previous example. In order to facilitate this elimination, one or both of the equations must sometimes be multiplied by a constant before we add them (an application of MPOE).

² This name is also related to a technique from the study of matrices called *Gaussian elimination*.

This approach is required to handle the startup exploration. Recall that in that problem we were asked to solve the following system by elimination.

$$\begin{cases} 4x + 3y = 22 \\ 3x + 5y = 11 \end{cases}$$

Adding these two equations won't eliminate any variables, as we have encountered in the earlier examples. But, with a few clever applications on MPOE, we can put ourselves into exactly the situation that we desire.

Let's make it our goal to eliminate x . What could we do to these equations so that we have equivalent equations but ones in which the x terms have opposite coefficients?

One way (but not the only way) to accomplish this is to multiply the first equation by 3 and multiply the second equation by -4 .

$$\begin{array}{rcl} 4x + 3y = 22 & \xrightarrow{\text{multiply through by 3}} & 12x + 9y = 66 \\ 3x + 5y = 11 & \xrightarrow{\text{multiply through by -4}} & -12x - 20y = -44 \end{array}$$

Then, we add equations:

$$\begin{array}{r} 12x + 9y = 66 \\ + \quad -12x - 20y = -44 \\ \hline -11y = 22 \\ y = -2 \end{array}$$

Substituting $y = -2$ into the second of the original equations, we get $3x + 5(-2) = 11$. This implies that $x = 7$. So, $\mathcal{S} = \{(7, -2)\}$.

8.5 Applications of systems

“A train leaves New York traveling south at 180 miles per hour...” If you were to read that sentence out loud in a room full of adults, you might cause some folks to start sweating and mumbling nervously to themselves. People all over the world, for generations, have been scarred by algebra word problems. Word problems have a bad reputation.

Yes, word problems require us to ponder a bit, decipher the information we’re given, and stumble around as we figure out how to proceed. But this is the process of problem solving! The algebra skills we have learned so far, along with a little determination and creativity, will serve us well as we tackle these algebra classics.

We discuss several kinds of problems below. Our general strategy is this: We’ll think about the unknown quantities and relevant facts given in the problem. We’ll choose variables to represent the unknowns and model the situation using equations. (Drawing pictures and making charts might help!) Then, we’ll solve the equations and check to ensure our answer makes sense.

A final thought before we get started: When we encounter a new problem it can be helpful to think back over our earlier work. If we have solved a similar problem before, that can often give us ideas about a good place to start when faced with something new.

8.5.1 Two unknowns, two facts

The first examples we did with substitution and elimination are of this type. We are given two unknowns and two different facts about them. The facts usually have different units which hint at the two equations we need to write for our model.

Example 8.7

Ticket prices at the Cheeseville Zoo were quite a bit cheaper years ago. In 1995, the zoo charged \$8 for each adult ticket and \$4 for each child ticket. On one particularly fateful day, the zoo sold 3505 tickets and collected exactly \$18,940. How many of each type of ticket were sold on that day?

The two unknowns in this problem are the number of adult tickets and the number of child tickets. The two facts are the total number of tickets sold and the total amount of money that those tickets cost. Let’s use A to represent the number of adult tickets and C to represent the number of child tickets that were sold on this fateful day at the zoo.

The fact about the total number of tickets sold tells us that one equation is

$$A + C = 3505.$$

We add up the numbers of adult tickets and child tickets to find the total number of tickets that were sold. Using the other fact about the total cost of the tickets, we write the equation

$$8A + 4C = 18940.$$

In the second equation, the term $8A$ is the amount of money that the zoo makes from the adult tickets alone, since each adult ticket costs \$8. The term $4C$ is the total cost of all the child tickets. The sum of these two values is the total amount of money that was collected.

We can now solve this system using our favorite method. It looks like a good candidate for elimination! (Why?)

8.5.2 Mixture problems

Mixture problems ask us to combine things that have different properties to create something new, with its own unique properties. Sometimes, we have to do a bit of clever thinking in order to determine information that is not directly stated in the problem. One helpful approach is to organize the given facts into a table.

Example 8.8

On this fateful day at the Cheeseville Zoo, zookeeper Florence Erlenmeyer was mixing ripe bananas and roasted crickets to make her secret-recipe monkey chow.^a Bananas cost \$3.50 per kilogram and crickets cost \$4.75 per kilogram. How many kilograms of each ingredient did Florence mix together if the result was 20 kilograms of monkey chow worth \$4.00 per kilogram?

^a Seven out of eight monkeys prefer Florence the zookeeper's secret-recipe monkey chow to the leading commercial brand of monkey chow, "Primate Crunch".

Our unknowns are the number of kilograms of bananas (let's call that B) and the number of kilograms of roasted crickets (let's call that C). We are given the prices of each ingredient per kilogram, but nothing about a total cost. Finding the total cost isn't that challenging though: all we have to do is multiply!

	Amount (kg)	Unit price (\$)	Total price (\$)
Bananas	B	3.50	$3.50B$
Crickets	C	4.75	$4.75C$
Mixture	20	4.00	80

The last row of the table shows that if Florence the zookeeper makes 20 kg of monkey chow worth \$4.00 per kilogram, then she has $20 \cdot 4.00 = \$80.00$ worth of monkey chow. The other rows in the table are using this same multiplication even though they each include a variable.

We now have our system of equations:

$$\begin{cases} B + C = 20 \\ 3.50B + 4.75C = 80 \end{cases}$$

Once again, this looks like a job for elimination. (Why?)

8.5.3 Work problems

Very often in a work problem, a group of workers split up a job. Working on their own, it might take different workers different amounts of time to finish the job. Our task might be to figure out how long the job will take if several workers cooperate.

A helpful way to think about problems like this is to try and identify the “work rate” of the various participants. If it takes someone 3 days to paint an apartment, their work rate for painting is “one third of an apartment per day”.

Example 8.9

For quality assurance purposes, Florence the zookeeper preferred to buy live crickets and toast them herself. By hand she could roast a crate of live crickets in 4 hours. She had recently invested in a Toast-O-Matic™ cricket roaster which could toast a crate of crickets in 2 hours. How long would it take to toast three crates of crickets if both Florence and the machine work simultaneously?

If we think of the “job” as “toasting one crate of crickets”, then Florence’s work rate is “1 job in 4 hours”, or “ $\frac{1}{4}$ job per hour”. Similarly, the Toast-O-Matic’s work rate is “ $\frac{1}{2}$ job per hour”.

The unknown in this scenario is the amount of time that it will take them to finish the job if they work together. Let’s use t to represent this unknown amount of time. If our two workers each work for t hours, Florence will complete $\frac{1}{4}t$ jobs and the Toast-O-Matic will complete $\frac{1}{2}t$ jobs.

	Work rate	Work time	Jobs done
Florence	$\frac{1}{4}$	t	$\frac{1}{4}t$
Toast-O-Matic	$\frac{1}{2}$	t	$\frac{1}{2}t$

We want to know how long it takes them to toast three crates of crickets. In other words, how long it takes them to complete “3 jobs”. This gives us the equation:

$$\frac{t}{4} + \frac{t}{2} = 3$$

Note that when we approached the problem in this way, we get a single linear equation rather than a system!

8.5.4 Rate, time, distance

“Rate, time, distance” problems (RTD for short) come in a variety of flavors. Sometimes the participants are moving in the same direction, sometimes in opposite directions. Sometimes one of the participants has a head

start, other times the participants begin moving at the same moment. Sometimes the participants start in the same location, sometimes they start in different locations. . .

These variations can combine to produce some confusing problems, but don't panic! Many times, drawing a simple picture at the start will help us to understand how the problem is set up. If there is important information that is not directly stated in the problem, a picture can help us to find it.

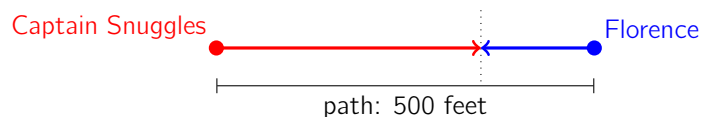
Example 8.10

Florence the zookeeper's office was 500 feet from the entrance to the Monkey Island attraction. On this fateful morning, she loaded a fresh batch of monkey chow into a wheelbarrow and began walking on the path to Monkey Island at a speed of 4 feet per second. At precisely the same moment, Captain Snuggles, the zoo's 350-pound gorilla, escaped from Monkey Island and began running toward the zookeeper along the same path at a speed of 26 feet per second. How long did it take Florence and Captain Snuggles to meet along the path?

In this problem, we have two participants who start moving at the same time but in different locations. They move toward each other (at different speeds) until they meet in the middle somewhere.

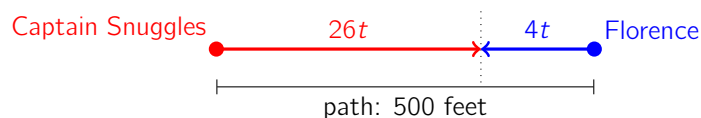
A common mistake with this type of situation is assuming that the travelers will meet at exactly the half-way point. This will not always be true! In this case, for instance Captain Snuggles moves much faster than Florence the zookeeper, and so they will meet closer to Florence's end of the path.

We might represent this with a very simple drawing:



The problem asks a question about time, so let's use t to represent the time (in seconds) it takes for the two travelers to meet.

What else do we know from the problem? We have the speeds that both of the characters are moving.³ Florence moves at a speed of 4 feet per second, so in t seconds she can travel $4t$ feet. Captain Snuggles can travel $26t$ feet in t seconds. So, we can update our picture:



³ We are also given the weight of the gorilla, although this is just a distraction and not important to the problem.

From the drawing we can see that the combined distance that the two characters travel must be exactly 500 feet! So, we have our equation:

$$26t + 4t = 500$$

Again, we've got just a single linear equation and not a system. Who would have thought that such a complicated-looking problem would have a relatively simple mathematical representation!

A trickier RTD problem

Example 8.11

Florence the zookeeper, sadly, found herself on the business end of a hungry gorilla. Luckily, this occurred just outside the security hut, where the zoo's security squad witnessed the terrifying (but non-fatal) mauling. When they sounded the alarm, Captain Snuggles fled toward the zoo's entrance. The security squad gave chase along the same route 30 seconds later, after arming their tranquilizer guns and mounting their bicycles.

Captain Snuggles ran at 18 miles per hour. The security team cycled at 25 miles per hour, which was just fast enough: they reached the zoo entrance at the same moment as the gorilla and tranquilized him into submission. How much time elapsed between the sounding of the alarm and the tranquilizing of the gorilla? How far was the security hut from the zoo entrance?

Since the security squad started 30 seconds later, they have 30 seconds less travel time than Captain Snuggles. If we let t represent the amount of time that Captain Snuggles is running (in seconds), then the security squad is chasing for $t - 30$ seconds.

Even though Captain Snuggles has a head start, the security squad reaches the entrance at the same time as the gorilla because of their greater speed. So, both parties travel the same distance d .

Now, we have to be careful about units! The speeds in this problem are given in *miles per hour*, and the time difference is given in *seconds*. So, we could be careful to keep our units aligned. Let's use a little dimensional analysis (from chapter 6) to convert the speeds to "feet per second".⁴

$$\begin{array}{lcl} \text{Captain Snuggles:} & \frac{18 \text{ miles}}{\text{hour}} \cdot \frac{5280 \text{ feet}}{\text{mile}} \cdot \frac{1 \text{ hour}}{3600 \text{ seconds}} & = \frac{26.4 \text{ feet}}{\text{second}} \\ \text{Security Squad:} & \frac{25 \text{ miles}}{\text{hour}} \cdot \frac{5280 \text{ feet}}{\text{mile}} \cdot \frac{1 \text{ hour}}{3600 \text{ seconds}} & = \frac{36.\bar{6} \text{ feet}}{\text{second}} \end{array}$$

⁴ Alternatively, we could let t represent time in minutes, and then convert the given speeds to distance *per minute*. In that case, the security team would be in pursuit for $t = \frac{1}{2}$ minutes, since 30 seconds is half of one minute.

If we let d represent the distance (in feet) from the security hut to the zoo entrance, then we can write an equation to describe the motion of Captain Snuggles:

$$d = 26.4t.$$

We can also model the motion of the security squad. Our formula shows both the different speed, and the fact that they start 30 seconds late:

$$d = 36.\bar{6}(t - 30).$$

These two equations comprise our system. Since we know that the security team and the gorilla arrive at the zoo at the same moment, we know that they travel the same amount of distance. In other words, d is the same for both travelers. Substitution looks like a good choice for solving this system! (Can you explain why?)

8.5.5 Motion in a current

Have you ever tried to walk *up* the *down* escalator at the mall? (Be honest!) When you walk down the down escalator – that is, when you walk in the same direction that the escalator is moving – the motion of the escalator helps you to go faster than if you were just walking down regular stairs. On the other hand, if you try walk in the wrong direction on an escalator, the motion of the escalator will make you go slower than you would walk on regular stairs.

This is the idea of motion in a current. When you go “with the current” (or “downstream”, if you’re floating on a river) the current adds to your usual rate of motion. Going “against the current” (or “upstream”), the current reduces your usual rate of motion.

Example 8.12

Bob and Yearleigh were at the Cheeseville Zoo on a school field trip at the time of the mauling. When the alarm sounded, they both ran towards the escalators leading up the side of Mount Ploom toward the aviary. Yearleigh, of course, ran up the up escalator. Bob, in a panic (and, well, because he’s Bob), ran up the down escalator.

Yearleigh reached the top in 19 seconds and Bob reached the top 23 seconds later. If the escalators are each 213 feet long, how fast were the twins running? How fast were the escalators moving? (Assume that the twins run at the same speed, and the two escalators move at the same speed.)

In this problem, we don’t know either of the speeds. So, let’s use r to represent the running speed of the twins and e to represent the moving speed of the escalators.

When Yearleigh runs with the help of the escalator, she attains a speed of $r + e$. Bob, on the other hand, has his overall speed reduced by the movement of the escalator and moves at a speed of $r - e$.

	Rate (ft/sec)	Time (sec)	Distance (ft)
Yearleigh	$r + e$	19	213
Bob	$r - e$	42	213

Be sure to read carefully! Notice that the problem says Bob reaches the top 23 seconds *later* than Yearleigh. That means the trip takes him $19 + 23 = 42$ seconds.

Given this setup, our system is:

$$\begin{cases} 19(r + e) = 213 \\ 42(r - e) = 213 \end{cases}$$

To solve this, we'll need the distributive property and the elimination method!

» Chapter summary «

We learned a number of techniques in this section for writing and solving systems of linear equations – graphing, substitution, and elimination – and we now have all three approaches in our algebraic toolbox. Each of the techniques has its advantages and disadvantages, but together they make a powerful collection.

In the next chapter we will broaden our algebraic vocabulary, so to speak, by dropping the requirement that we always talk about *equations*, or *systems of equations*. Equality is important (in mathematics and in life), but we all know that things are not always in perfect balance. Onward!

Meanwhile, in Cheeseville...

Florence the zookeeper made a full recovery, and Captain Snuggles successfully completed an anger management course and was reintroduced to the Cheeseville Zoo community. Unfortunately, the day's events were a public relations nightmare for the zoo. Annual attendance declined exponentially until 2003, at which point the zoo was purchased by YearleighCorp.

The tremendous power of the simplex method is a constant surprise to me.

George Dantzig, American mathematician

Chapter 9

Inequalities

Until now we have focused almost exclusively on the concept of equality, for example we have learned to solve linear *equations*. But *equality* is not the relationship shared by most numbers. For example 10 and 20 do not have the “equality relationship”: 10 is less than 20! There is a relationship of *inequality*.

How do we describe mathematically unequal relationships? What rules govern these relationships? How to we represent such relationships visually? These questions are at the heart of this chapter. By the end, we will have developed the simple idea of “inequality” into a sophisticated tool for finding an optimal solution in certain complex situations where many choices are possible (the *simplex method* mentioned in the quote above).

9.1 Equations versus inequalities

Honest senators?

// startup exploration

Imagine that the following statements are released from a government watchdog group: “(1) There exists at least one honest senator. (2) Given any two senators, at least one of them is dishonest.”

If both of these statements are true, what can we say about our 100 senators? Determine the number of honest senators or, explain why we don’t have enough information to determine exactly how many are honest.

An equation states that two expressions are equal and thus includes an equal sign “=”. When we solve an equation (by using one of the properties to simplify it), we create a series of equivalent equations until we can determine the solution set. Recall that the solution set is the collection of all the numbers that make the original equation “true”.

Inequalities express an unequal relationship and thus include a mathematical symbol of inequality. Rather than “ x is equal to 5” we might have “ x is less than 5”. To solve an inequality is to find all of the values that make it true. But whereas an equation may have only one solution, there could be many values that make an inequality true. Perhaps infinitely many! All of these are part of the solution set.

9.1.1 Inequality symbols

Equations all include the equal sign, but we have more options for expressing inequality:

\neq	not equal to
$<$	less than
$>$	greater than
\leq	less than or equal to
\geq	greater than or equal to

Using these symbols we can write true and false statements. The statement $4 < 5$ is true statement, and the statement $3 > 10$ is false.¹ Note that the statement $12 < 12$ (12 is less than 12) is false. In contrast, the statement $12 \leq 12$ (12 is less than or equal to 12) is true, because the “or equal to”.

The math sentence $x < 5$ is an inequality that has infinitely many solutions. For example, we can replace x with 4, π , 3, 2, 1.889, 1, $\frac{1}{2}$, 0, -6 . . . in fact, every number x that is less than the “endpoint” 5 is included in the solution set! But again, 5 itself is not included in the solution set for this inequality.

Because of their behavior around the endpoints, the symbols $<$ and $>$ are said to be “exclusive inequalities”, meaning they do not include the endpoint as a solution: $5 < 5$ is a false statement. These are often described as “strictly less than” or “strictly greater than”. In contrast, the symbols \leq and \geq are said to be “inclusive” because they do include the end point as a solution: $5 \leq 5$ is a true statement.²

¹ Remember your elementary school days: the symbols are like hungry alligators who open their mouths to eat the larger number. Nom, nom, nom.

² Another way to describe these “or equal to” inequalities is using the terms “at most” or describe “less than or equal to”, and “at least” to describe “greater than or equal to”. For example, if I ask you to give me “at least 5 dollars”, then I’m asking you for “greater than or equal to 5 dollars”.

9.2 One-variable inequalities

Wire triangle

// startup exploration

A nine-inch piece of wire is bent at two points such that its ends come together to form a triangle. If the bending points must be on the inch marks, how many possible choices of bending points are there?

Consider the equation $3x = 12$. This equation is true when x has the value 4 and we say the solution set $S = \{4\}$. Now consider the inequality $3x \leq 12$. What values of x make this inequality true? The solution set will be the set of all of those values.

This means that the solution set contains “all real numbers less than or equal to 4” (notice that 4 is included in the solution set). To write this in set notation, we introduce a few new symbols:

$$S = \{x \in \mathbb{R} \mid x \leq 4\}$$

The symbols in this set notation can all be translated into regular English. The symbol \in , which looks like a little stylized ‘e’, means “is an element of”. So, the phrase $x \in \mathbb{R}$ means “ x is an element of \mathbb{R} ”, or in other words: “ x is a real number”.

The little vertical line has a meaning, too! We can think of the bar as saying “such that” or “for which”. So what we have is a math sentence that reads “ S is the set of real numbers x such that x is less than or equal to 4”, or a bit more succinctly “all real numbers less than or equal to 4”.

9.2.1 Graphing one-variable inequalities

Instead of writing the set notation, another way to show the solution to a one-variable inequality is to draw a graph of the solution. These graphs aren’t like the two-dimensional graphs of linear functions. Since we have just one variable, we’ll have a one-dimensional graph! Rather than graphing on the coordinate plane, we draw these graphs on a number line.

These graphs may seem pretty basic, but stay tuned! In Algebra 2 we will solve “combined” and “polynomial” inequalities and we won’t be able to figure out the answers without graphs like these!

Example 9.1

Draw a graph of $x \leq 4$.

Solution: First, we'll show the graph, and then we'll explain what decisions we made in drawing it.



We first found the boundary (which is 4) and plotted it on the number line using a filled in dot. Then we drew an arrow (a ray) pointing to the left because we are interested in values that are less than 4, in addition to 4 itself. We also placed the zero on the number line, for reference.

Contrast the previous example with the next example, and you will probably have a good idea about how to graph any one-variable inequality!

Example 9.2

Draw a graph of $x > -2$.

Solution: Here's the graph. What features are different compared to the previous example?



First, we found the boundary (it's -2) and plotted it on the number line, along with 0. The boundary is excluded from the solution set, so we used an hollow circle to show that -2 itself is not part of the solution. Then we used an arrow (a ray) pointing to the right because we are interested in values that are greater than -2 .

The following comments may not be necessary, given the previous two examples, but we'll summarize them here for clarity. There are two main things to consider when graphing a one-variable inequality.

Consideration #1: Boundary. What is the greatest or least possible value of the solution, and is that point included in the solution set?

If we have an inclusive inequality (\leq , \geq , and $=$), then the boundary point is included in the solution set. On the graph, we use a filled in circle or “closed” endpoint to show that the boundary is part of the solution. We call the boundary point an **inclusive boundary**.

If, on the other hand, we have an exclusive inequality ($<$, $>$, and \neq), then the boundary point is not included in the solution set. On the graph we draw a hollow or “open” endpoint to show that the boundary is not part of the solution. We call the boundary an **exclusive boundary**.

Consideration #2: Direction. Are the solutions greater than the boundary (stretching to the right, towards positive infinity) or are they less than the boundary (stretching to the left towards negative infinity)? We draw a heavy arrow in that direction.

We always label the boundary point, and also place the zero on the number line for reference.

9.2.2 (;,;) Interval notation

In addition to set notation and a number line graph, there is a third way to describe an inequality. In many higher level math classes (like Algebra 2 and, later, Calculus), we use a technique called *interval notation*. With interval notation, we describe the endpoints of a region on the number line.

Inclusive boundaries are denoted with square brackets: [and]. Exclusive boundaries are denoted using round brackets: (or). Lines that extend forever have no endpoint, so we use positive infinity ∞ or negative infinity $-\infty$ to indicate that the set extends forever in the positive (or negative) direction.

In the first example above, we have $x \leq 4$. To write this in interval notation, we write $x \in (-\infty, 4]$. To write the solution to the second example, $x > -2$, in interval notation, we write $x \in (-2, \infty)$.

Try drawing the graph corresponding to $x \in [8, \infty)$? What about $x \in (-\infty, -10)$? What do you suppose the graph of $x \in [-3, 7)$ might look like?

9.3 Solving inequalities

When we solve an equation using the POEs, we create a series of equivalent equations until we can determine the solution set. The big question of this section is: Can we do something similar for inequalities? What does it mean to have *equivalent inequalities*?

Equivalent inequalities is not an oxymoron.³ It just means that two inequality statements have the same solution set! For example, $3x < 12$ and $x < 4$ are equivalent because the exact same values of x will work in both statements. They have the same graph!

But, how can we transform an inequality into an equivalent inequality? Can we use the same properties as we used for solving equations? In essence: are there Properties of Inequalities like there are Properties of Equality (POEs)?

Strictly speaking, an inequality doesn't only state that two numbers are not-equal.⁴ An inequality really shows that two numbers appear in a certain *order on the number line*. If a certain inequality states a "less than" relationship, we want to find properties that maintain that.

Since we are looking for rules which "maintain order", we will call these rules the "Properties of Order", or POOs.⁵

9.3.1 Properties of order (POOs)

So what are the POOs, and do they work like POEs? Yes... almost.

Suppose we start with an inequality that we know is true, say, $5 < 6$. What operations can we perform to this inequality that maintain its ordered relationship?

Could we add (or subtract) the same things on both sides and maintain order? To take just one example, adding 3 to both sides gives $5 + 3 < 6 + 3$, which maintains order since $8 < 9$. If we now add -8 to both sides we have $0 < 1$, and order is again maintained.

In fact, adding the same thing to both sides does nothing to change their relative order on the number line. Adding simply translates the original inequality to the right or left on the number line. (Note that translations to the left happen when adding a negative number, which is the same as subtracting!)

³ An oxymoron is a figure of speech that brings together two things that appear to be contradictory, like the term "jumbo shrimp". The titles of the movie "Night of the Living Dead" and the song "Alone Together" contain oxymorons.

⁴ Okay, yes, this is exactly what the \neq symbol does. But, we're not going to do much with the \neq symbol. Solving problems with \neq isn't usually very interesting, as we'll see later in this chapter.

⁵ We considered using "POI" to stand for "Property of Inequality", but POO is both more mathematically accurate, and a lot funnier.

Properties of order: Addition and subtraction

The **addition property of order** (APOO) states that for all real numbers a , b , and c

$$\text{if } a < b, \text{ then } a + c < b + c$$

The **subtraction property of order** (SPOO) states that for all real numbers a , b , and c

$$\text{if } a < b, \text{ then } a - c < b - c$$

Note! Rather than write out the same sentence with all the different inequality symbols, we simply note here that APOO and SPOO hold for $a \leq b$, $a > b$, and $a \geq b$.

What about multiplication and division? If we take our sample inequality $5 < 6$ and multiply both sides by 4, we have $5 \cdot 4 < 6 \cdot 4$, which maintains order, since $20 < 24$. If we now divide both sides by 2 we have $10 < 12$, and order is maintained.

This is encouraging! But... we have to be a bit cautious. Notice that we have not explored multiplication by a negative number. Observe that if we start with $5 < 6$ and multiply both sides by -4 , we get -20 and -24 and $-20 > -24$ — not the other way around. The order of the inequality has changed!

There's a geometric interpretation to this: Multiplication (and division) perform a scaling of the points away from (or towards) zero. Multiplying by a positive number performs that scaling on the same side of zero as the original numbers. But, multiplying by a negative number first reflects our inequality about zero. That reflection changes the order of the numbers!

Properties of order: Multiplication and division

The **multiplication property of order** (MPOO) states that for all real numbers a , b , and c :

$$\text{if } a < b \text{ and } c > 0, \text{ then } ac < bc$$

$$\text{if } a < b \text{ and } c < 0, \text{ then } ac > bc$$

The **division property of order** (DPOO) states that for all real numbers a , b , and c :

$$\text{if } a < b \text{ and } c > 0, \text{ then } \frac{a}{c} < \frac{b}{c}$$

$$\text{if } a < b \text{ and } c < 0, \text{ then } \frac{a}{c} > \frac{b}{c}$$

Note that, as before, MPOO and DPOO behave the same for all of the inequality symbols, not just $<$.

For the most part, we can solve an inequality just like we solve an equation. The addition and subtraction POOs are no different than the POEs. However, we have to be a bit more careful with multiplication and division. When multiplying or dividing both sides of an inequality by a negative number, you must change the direction of the inequality.

Example 9.3

Solve and graph: $-3x < 21$.

Solution: We can divide both sides by -3 and change the direction of the inequality: $x > -7$. To check, we might try plugging in values that are on both sides of the boundary (like -8 and -6), into the original inequality:

$-3 \cdot -8 = 24$ and $24 > 21$, so -8 *does not* make the original inequality true. The solution set cannot include this value.

$-3 \cdot -6 = 18$ and $18 < 21$, so -6 *does* make the original inequality true. The solution set must include this value, so the solutions must extend from the boundary (-7) toward to the right (hitting -6 and all the points toward positive infinity).

**9.3.2 What about the field axioms?**

Recall that the field axioms were our other tools for simplifying equations, but they were used on one side of an equation only. Since they are not performed on both sides, there's no risk of them messing with order.

So, we can distribute and combine like terms without it having any effect on the truth of the inequality. We can use them to simplify, just like we did when we were solving equations

Example 9.4

Solve and state the solution set: $4x + 3 - 2(3x + 1) > 13$.

Solution: We simplify on the left-hand side first, and then apply the POOs.

$$\begin{array}{ll}
 4x + 3 - 2(3x + 1) > 13 & \\
 4x + 3 - 6x - 2 > 13 & \text{distributive property (check your signs!)} \\
 -2x + 1 > 13 & \text{combine like terms} \\
 -2x > 12 & \text{APOO} \\
 x < -6 & \text{MPOO with a negative number}
 \end{array}$$

In set notation, we write $\mathcal{S} = \{x \in \mathbb{R} \mid x < -6\}$.

9.3.3 Special case solutions

As with equations, inequalities with variables on both sides can have no solution or all real numbers as its solution.⁶ The variables may vanish, as they sometimes do, but now we have to look at the inequality that is left over to determine whether it is always true or always false. If the remaining inequality is a true statement, then our solution set is *all real numbers*. If the remaining inequality is false, then our solution set is empty.

Example 9.5

Solve and state the solution set: $6x - 2 > 2(3x + 1)$.

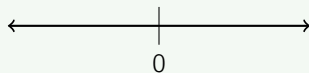
$$6x - 2 > 2(3x + 1)$$

$$6x - 2 > 6x + 2 \quad \text{distributive property}$$

$$-2 > 2 \quad \text{subtract } 6x \text{ from both sides (APOO)}$$

When we examine the remaining inequality, we see that it's false: -2 is not greater than 2 . Therefore, the original inequality has no solution. To write the answer in set notation: $\mathcal{S} = \emptyset$.

The graph of “no solution” is just an empty number line. In that case, we can either write “no graph” or “empty graph” or just draw a blank number line. Boring, perhaps, but mathematically accurate.



The mirror-image situation, as you might expect, is demonstrated by the following example.

Example 9.6

Solve and state the solution set: $5x + 4 \geq 2x + 3x - 1$.

$$5x + 4 \geq 2x + 3x - 1$$

$$5x + 4 \geq 5x - 1 \quad \text{combine like terms}$$

$$4 \geq -1 \quad \text{subtract } 5x \text{ from both sides (APOO)}$$

This time when we interpret the remaining inequality, we see that it's true: 4 is greater than or equal to -1 . Therefore, every real number is a solution to the original inequality. To write the answer in set notation: $\mathcal{S} = \mathbb{R}$.

⁶ We can't simply say “infinitely many solutions”, since an inequality like $x < 0$ has infinitely many solutions.

The graph of all real numbers is a graph with the entire number line colored in.



9.3.4 Not equal

We haven't talked much about inequalities that use \neq , but that's because they're pretty uninteresting. For example: Solve $8x + 3 \neq 19$. We can subtract 3 from both sides, then divide both sides by 8, yielding: $x \neq 2$. So, the only number that *doesn't* work in this inequality is when $x = 2$.

In set notation, we write: $\mathcal{S} = \{x \in \mathbb{R} \mid x \neq 2\}$. The graph is the whole number line, except that single point:



9.4 Two-variable inequalities

Pick some points

// startup exploration

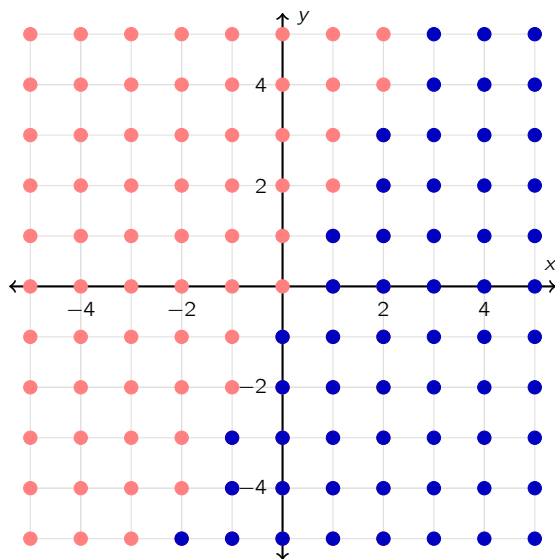
Consider the two-variable inequality $y < 2x$. If we let $x = 5$ and $y = 1$, we have a true statement since $1 < 2 \cdot 5$, that is to say $1 < 10$. The point $(5, 1)$ satisfies our inequality.

On the other hand, the point $(0, 1)$ does not satisfy our inequality. If we let $x = 0$ and $y = 1$, our inequality doesn't work: $1 < 2 \cdot 0$ is false.

Find some other points that satisfy this inequality, and some other points that do not. Plot these points on a graph (perhaps using different colors). What patterns do you notice?

Having spent all of that time talking about one-variable inequalities, we can now get on with something more interesting: turning a linear equation like $y = 2x$ into a two-variable inequality like $y < 2x$.

In the startup exploration, we set out to find some points that either satisfy or fail to satisfy this inequality. If we plot some points on a graph – dark blue for “satisfy”, pale red for “fails to satisfy” – an interesting pattern emerges:



We could carry out this procedure with any inequality, say $y \leq -2x + 6$, by checking a few random points, say

$(4, 5)$ and $(-3, 2)$, to see if they satisfy the inequality .

$$\begin{array}{ll}
 y \leq -2x + 6 & \text{original inequality} \\
 5 \stackrel{?}{\leq} -2(4) + 6 & \text{substitute candidate point} \\
 5 \stackrel{?}{\leq} -8 + 6 & \\
 5 \not\leq -2 &
 \end{array}$$

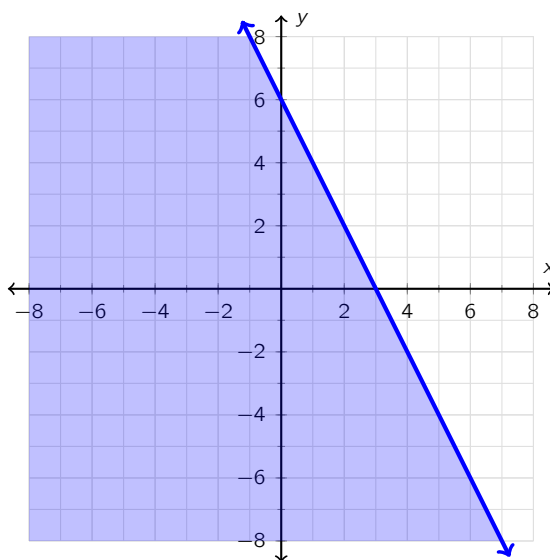
The final inequality is false, so the answer is no, the point $(4, 5)$ is not part of the solution set to the inequality $y \leq -2x + 6$. If we test the other point, $(-3, 2)$, we have:

$$\begin{array}{ll}
 y \leq -2x + 6 & \text{original inequality} \\
 2 \stackrel{?}{\leq} -2(-3) + 6 & \text{substitute candidate point} \\
 2 \stackrel{?}{\leq} 6 + 6 & \\
 2 \leq 12 &
 \end{array}$$

This statement is true, and so the point $(-3, 2)$ is part of the solution set. But our work above suggests that this probably isn't the *only* point in the solution set. What other points might be included?

Note that the inequality is inclusive, so any point on the line $y = -2x + 6$ must be part of the solution set (that's the line *y equals* $-2x + 6$).

In addition, the solution set must include any point that is "less than" that line. In other words, any point on the same side of the line as $(-3, 2)$, the point which we checked above, and which we *know* is in the solution set. So, our solution must look like the graph below.



So, graphs of two-variable inequalities share some of the features of one-variable inequalities. We will have a boundary, but rather than a dot or endpoint, it will be a line (or a curve). We will shade a section of the graph to show all the solutions, but this will be a whole region of the plane, rather than a ray on the number line.

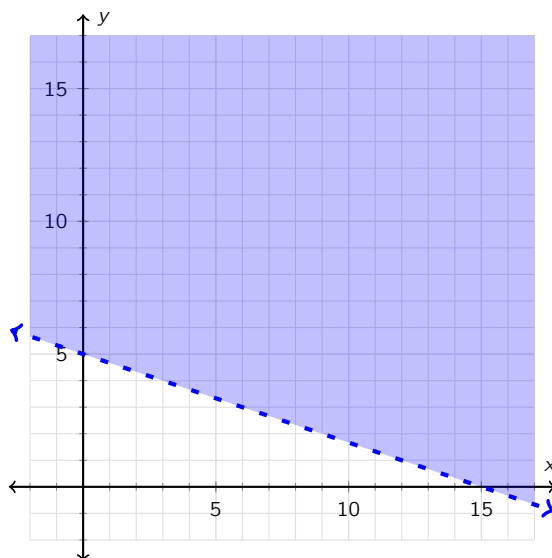
Let's work through another example in detail: graphing the inequality

$$x + 3y > 15.$$

Here we have a line in standard form. An easy way to graph standard form is to find the x - and y -intercepts. When we plug in $x = 0$, we have $y = 5$. When we plug in $y = 0$, we have $x = 15$. So, we can draw our graph by connecting the points $(0, 5)$ and $(15, 0)$.

Before we draw the line, though, note that we have an exclusive inequality: “greater than”. For one-variable inequalities, we drew a hollow dots which allowed us to show where the boundary was without actually including the boundary itself. In the graph of a two-variable inequality, we will draw a dotted or dashed line to show that the line itself is not included in the solution set.

Since our inequality is “greater than” it seems reasonable to shade the region above the line. The resulting graph is shown below.



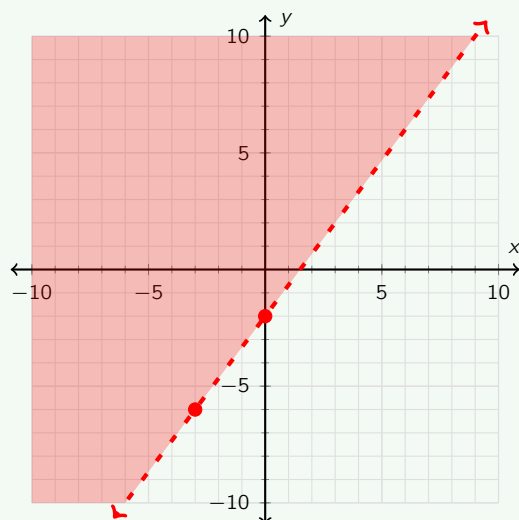
We can check that we've shaded the correct side by testing a point. A convenient point to test is the origin — all of that multiplication by zero makes this easy! If we plug in $x = 0$ and $y = 0$ to our equation, we have

$$\begin{array}{ll} x + 3y > 15 & \text{original inequality} \\ 0 + 3(0) \stackrel{?}{>} 15 & \text{substitute candidate point} \\ 0 \not> 15 & \end{array}$$

We have a false statement in the end, so the origin is not included in the solution set. This means the solution set must lie on the other side of the line. Our graph agrees with this.

Example 9.7

Write the inequality pictured in the graph below.



Solution: We can use the two given points to find that the slope of the line is $\frac{4}{3}$ (perhaps using a slope triangle). We're also given the y -intercept, so we can write the equation for the boundary line: $y = \frac{4}{3}x - 2$.

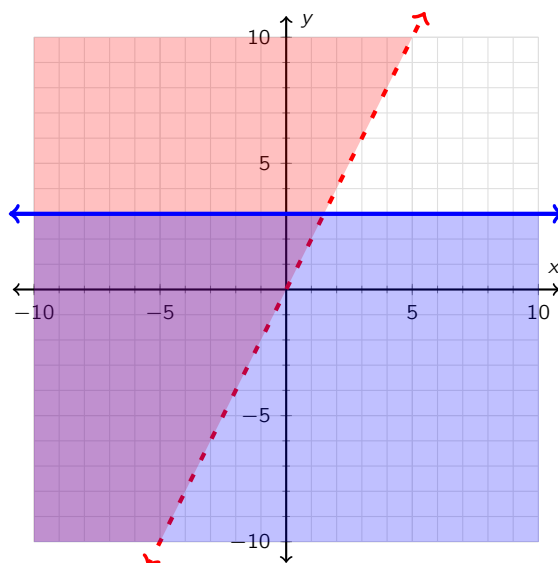
Since the line is dotted, we know that we're dealing with an exclusive inequality, and the shading lines above the line, so this graph depicts the inequality

$$y > \frac{4}{3}x - 2.$$

Finally: Remember to keep in mind the criteria for high-quality graphs (graphs should be done on graph paper, we should be mindful about how we place and scale the axes, and so on.)

9.5 Systems of inequalities

By now you can perhaps anticipate where this is heading! Here you see a graph of a system of inequalities. The solution to the system is the set of points that the graphs share. In other words, the region where they “overlap”.



System of inequalities

A set of two or more inequalities with the same variables.

Solution to a system of inequalities

The set of all points common to each inequality in a system. Graphically, the region where the graphs overlap.

The only way to find the solution to a system of inequalities is to graph the system! That is, to graph both inequalities on the same set of axes. There is really no meaningful set notation for the solution; the solution is the graph. In that visual we can see the region of the plane that contains the solution set, boundaries and all.

Graphing tips

Systems of inequalities can be quite fun to graph. We can use different colors for each inequality, for example yellow and blue colored pencils. Where the graphs overlap, we'll get green!

If we're using just a regular pencil and have only one color choice, we might shade one inequality with horizontal lines and shade the other inequality with vertical lines. The region where we see the checkerboard shading is the solution set.

In any case, it will be important to make sure the solution region is clear. If the colors or shading are muddled and hard to interpret, consider using an arrow or label to identify the solution set.

Example 9.8

Write the system pictured in the graph at the beginning of this section.

Solution: The horizontal line goes through the point $(0, 3)$. It is a solid line and shaded below, so that means we have the inequality "less than or equal to". So, the blue graph depicts $y \leq 3$.

The red graph is a direct variation (a straight line through the origin) through the point $(1, 2)$. So the equation for the line is $y = 2x$. This line is dashed and shaded above, so we have the inequality $y > 2x$.

So, together we have the system of inequalities

$$\begin{cases} y \leq 3 \\ y > 2x \end{cases}.$$

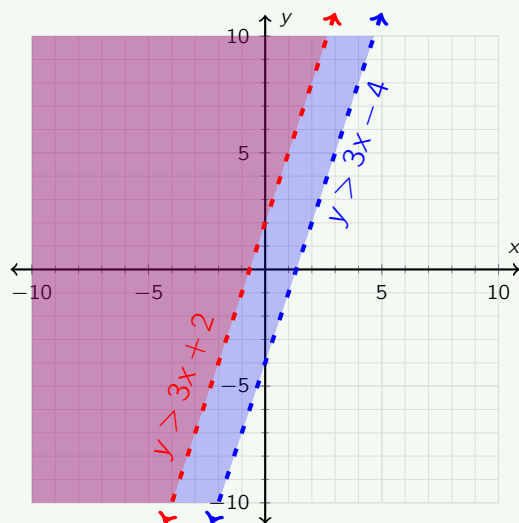
9.5.1 Special case systems

We have been graphing lines for a while now, and shading on one side of a line doesn't usually add that much of a challenge. Sometimes, however, we encounter some unexpected images.

Example 9.9

Example. Graph the system $\begin{cases} y > 3x - 4 \\ y > 3x + 2 \end{cases}$

Solution: Before we start graphing, note that the lines are parallel! Our first instinct might be to say that this system has "no solution". That would be correct, if we were dealing with a system of *equations*, but here we have inequalities! Let's take a look at the graph.



Note that the red region is entirely covered over by the blue region! So the overlapping region is, in fact, the red inequality $y > 3x + 2$.

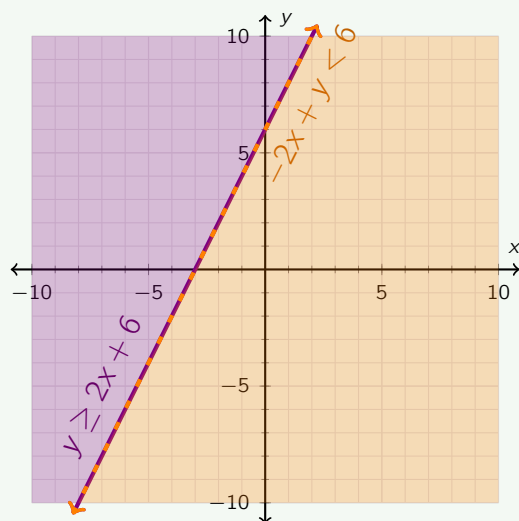
As the previous example shows, a pair of parallel lines can create a pair of inequalities which overlap. But, could the shading have gone differently? There are four different scenarios that can arise when turning two parallel lines into inequalities (one of this is shown in the example). Can you draw pictures of the other three scenarios? How does the solution set look in each case?

Example 9.10

Graph the system:

$$\begin{cases} y \geq 2x + 6 \\ -2x + y < 6 \end{cases}$$

Solution: Let's fast-forward to the graph and see what's happening here.



It might not have looked like it at first, but we have two different ways of expressing the same line! Both inequalities have the same boundary, but they are shaded in opposite directions, so those regions do not overlap anywhere except the boundary itself.

Since the boundary is exclusive for one of the inequalities, the boundary cannot be part of the solution set. So in this case, there is no part of the graph where the two shaded regions overlap. This system of inequalities has no solution!

What are the other cases that might arise when we have two inequalities that share the same boundary line? Can you draw graphs that express these different possibilities?

9.5.2 Checking a solution

We jumped right in to graphing a system, but suppose we want simply to check to see if a given point is a solution to a given system of inequalities?

Example 9.11

Determine whether the point $(-2, 1)$ is a solution to the system:

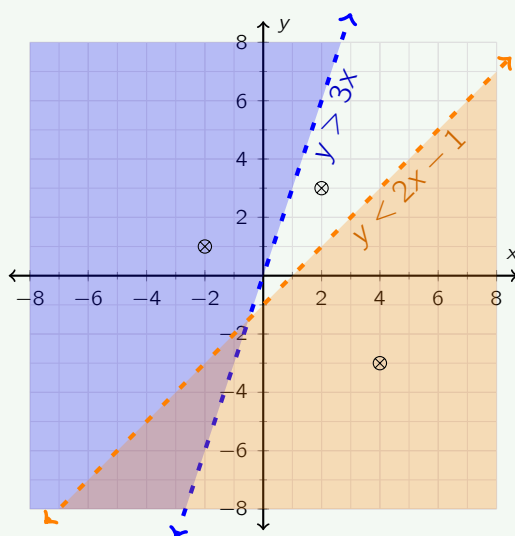
$$\begin{cases} y > 3x \\ y < 2x - 1 \end{cases}.$$

Solution: To answer this question, we don't have to graph. We can just substitute the point into each inequality and check to see whether it makes both true. Let's check $(-2, 1)$:

$$\begin{array}{ll} y > 3x & \text{and} \quad y < 2x - 1 \\ 1 \stackrel{?}{>} 3(-2) & 1 \stackrel{?}{<} 2(-2) - 1 \\ 1 \checkmark > -6 & 1 \not< -5 \end{array}$$

The point satisfies the first inequality, but not the second. So, $(-2, 1)$ is not a solution to the system.

Graphing is not required to answer this question, but it may be helpful to see a visual. Note where the given point lies on the graph. Predict what would happen if we tested the non-solution $(2, 3)$. Do the same for the non-solution $(4, -3)$.



9.6 Application: Linear programming

Evil vegan appliances

// startup exploration

A subsidiary of YearleighCorp, Evil Vegan Appliances, manufactures solar-powered soymilk makers. They manufacture 2 types: a large capacity model for commercial use, and a smaller one for home use. Since these soymilk makers delicate technology, the factory can only hand-make a total of 16 machines per day.

In order to keep demand up — and because she enjoys toying with the emotions of her customers — Yearleigh decides to restrict production in another way: She decides that they should build no more than 10 commercial models and no more than 12 family models per day, just to keep everyone wanting more.

If Evil Vegan Appliances makes \$75 dollar profit on each family size soy milk maker and \$100 profit on each commercial model, how many of each type should they build each day to maximize profit? What will their maximum daily profit be in this case?

Businesses want to maximize their profits and minimize their costs. At the same time, businesses have constraints on the availability and cost of resources. They have to hire workers, buy materials, build production facilities, and so on. A key challenge for any business is to navigate the various constraints so that they can configure their operations in the optimal way.

Optimization

Maximizing or minimizing a quantity, given a set constraints.

There are multiple techniques for optimization. The approach we will learn here was first called the *simplex method*, and has now come to be called **linear programming**.⁷ Real-world linear programming problems have many variables, sometimes numbering in the millions. Since we're only in Algebra 1, we'll stick with just a few variables.

9.6.1 Process of linear programming

To begin the process of modeling the startup exploration, our first goal is to identify what quantity we are trying to optimize, and what variables are at play in the scenario.

⁷ More on the interesting history of linear programming at the end of this section!

In our case, the quantity that we wish to optimize is *profit*, and the variables are the number of commercial machines and number of family machines that can be made per day.

We use the variables and information from the problem to write what is called the *objective function*. This is the equation we use to calculate the quantity we want to optimize.

Let P represent the total profit, let C represent the number of commercial machines that can be made per day, and let F represent the number of family machines that can be made per day. The company makes \$100 per commercial machine and \$75 per family machine, so our objective function (profit function) is

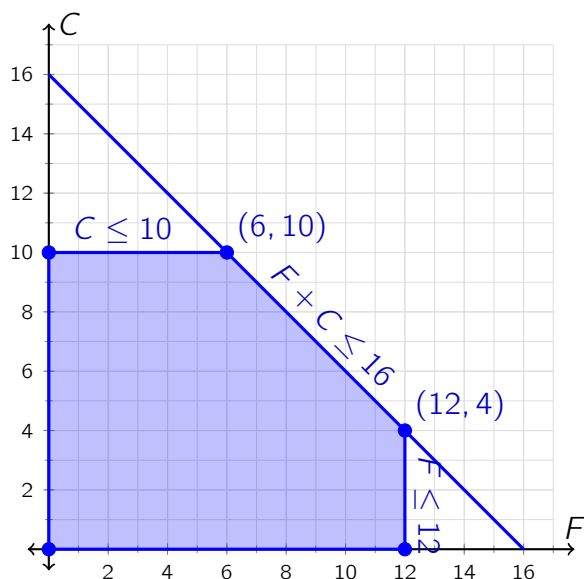
$$P = 100C + 75F.$$

We then model the constraints as a system of inequalities. The constraints on our system include: Per day they can make at most 16 machines (of any kind). In particular, no more than 10 commercial machines, and no more than 12 family machines per day. There are also natural “minimum” constraints at zero, since the factory can’t make a negative number of machines.

Now, we combine our variables with our constraints to write a system of inequalities:

$$\begin{cases} F + C \leq 16 & \text{limit on the total number of machines} \\ C \leq 10 & \text{limit on the number of commercial machines} \\ F \leq 12 & \text{limit on the number of family machines} \\ C \geq 0 & \text{floor on the number of commercial machines} \\ F \geq 0 & \text{floor on the number of family machines} \end{cases}$$

We graph all five of these inequalities on the same set of axes. F and C aren’t really dependent or independent variables, so it doesn’t matter which we place on which axis. We’ll put F on the horizontal axis.



The graph is a polygon, and every point inside this region represents a number of family size machines and a number of commercial size machines that the factory could possibly make, given the limits on production. Now, the task is to figure out which point earns the most profit.

The key insight here is that we should find the intersection points on the system. These intersection points show us where we utilize multiple resources to their fullest extent. It turns out that one of these intersection points will be the point where we optimize the objective.⁸

The intersection points are $(0,0)$; $(0,10)$; $(12,0)$; $(6,10)$; and $(12,4)$. Those coordinates are of the form (F,C) . It's important to keep track of what the numbers represent!

Now, we plug these coordinates into our objective function for total profit, $P = 100C + 75F$. One of these points will give the maximum profit and one will give the minimum profit.

- The point $(0,0)$ means 0 Family, 0 Commercial: $P = 75(0) + 100(0) = \$0$
- The point $(0,10)$ means 0 Family, 10 Commercial: $P = 75(0) + 100(10) = \$1000$
- The point $(12,0)$ means 12 Family, 0 Commercial: $P = 75(12) + 100(0) = \$900$
- The point $(6,10)$ means 6 Family, 10 Commercial: $P = 75(6) + 100(10) = \$1450$
- The point $(12,4)$ means 12 Family, 4 Commercial: $P = 75(12) + 100(4) = \$1300$

So, now we have our solution: To maximize daily profit, the factory needs to make 6 family and 10 commercial machines per day, for a maximum daily profit of \$1450.

Discussion

The first vertex $(0,0)$ is kind of a silly point to test. You might not be surprised to see that it's the scenario that creates the minimum profit.

The points $(0,10)$ and $(12,0)$ have the company maximizing only one constraint — the number of a certain type of machine they can make — but completely ignoring the fact that they can make a total of 16 machines per day.

The other two vertices show the company making all 16 machines they can make per day, and maximizing one of the other constraints. We can't maximize all three constraints at the same time. (Can you explain why not?)

Finally, our solution makes sense because it seems reasonable that we would want to find the solution that maximizes overall production, but also maximizes production of the machines that make more money for the company!

⁸ See "History of Linear Programming" below to learn a little about the people who created and proved that this process works.

In a real linear programming problem, there would likely be many more constraints. For example, the maximum number of machines per day might depend on the number of employees we have, the amount of time each employee can work, how productive those workers are during that time, their salaries, how much money there is in the payroll account, how many parts we need, how many parts we have in stock, how much it costs to make new parts, how long it will take to make them, . . . well, this could go on for a while. At any rate: businesses work quite hard to figure out challenges like this!

9.6.2 History of linear programming

A Soviet mathematician named Leonid Kantorovich was the first to use linear programming in 1939. During World War II he was helping the Soviet army minimize their costs while at the same time maximize losses for their enemies, Nazi Germany. His technique was effective and the Soviets kept it a secret. Because of his work, Kantorovich won a Nobel Prize in Economics, the only Soviet economist to ever win in that field.

In 1947, George Dantzig, an American mathematician, created and published his own algorithm for linear programming. The example he used was a problem that would otherwise have taken vast amounts of computing time to figure out because the number of possibilities to account for was more than the number of particles in the universe! In 1947, they did not have computers that could handle that much data. Rather than test each of the vast number of possibilities, Dantzig's "simplex" method for linear programming took just a few moments and a few calculations.

A funny side-note about Dantzig: one time he was late to a college statistics class at Berkeley. He saw two problems written on the board. He just thought they were really hard homework problems, so he solved them and turned them into his professor.

It turns out that the two problems were actually "open problems" in mathematics. An open problem is a problem in mathematics that no one has been able to solve yet. His professor spent the beginning portion of the class discussing these problems, but since Dantzig was late, he missed all of that.

Also in 1947, another American mathematician, John von Neumann, who contributed quite a bit to mathematics and physics, developed another way to approach linear programming called the "theory of duality". There are other mathematicians who advanced the process over the years. Now, every business and government agency uses linear programming to determine the best way to use their resources.

» Chapter summary «

We have reached the summit of our exploration of linear functions! We have learned to solve linear equations, and systems of linear equations. We can write linear functions in a variety of different forms. Plus, we have taken our understanding of equations and extended to a whole new world: one- and two-variables inequalities, and systems of two-variable inequalities.

Take a moment to look back over all the ground we have covered in chapters 5 to 9. These accomplishments are something to be proud of!

In the next chapter, we shift our focus to the family of exponential functions. Much of what we have learned so far will be helpful: the field axioms and the POEs will be by our side from now until the end, for instance. Of course, we will also be faced with unfamiliar situations that will require us to master a new collection of properties and methods. This is both a challenge and an opportunity! Onward!

The greatest shortcoming of the human race is our inability to understand the exponential function.

Albert A. Bartlett, American physicist

Chapter 10

Exponential functions

We first saw exponential relationships back in chapter 2, when we studied geometric sequences. In this chapter we will look more closely at exponential functions, but we will be forced to leave some questions unanswered.

Our focus now will be on the fundamentals of exponentials. Then, exponential functions will return in Algebra 2. In that course we will learn more techniques for representing and manipulating exponential functions, and learn tools for solving exponential functions.

10.1 Exponential relationships

Super seaweed

// startup exploration

YeadleighCorp genetic engineers have developed a new kind of fast-growing, nutrient-rich seaweed for the turtle pens. The length of a strand of this seaweed doubles every day.

Scientists measure a particular strand of seaweed to be 15 centimeters long. How long will this strand be in 1 day? 2 days? 3 days? How long was this strand 1 day ago? 2 days ago? 3 days ago?

We are told that the seaweed doubles in length each day, so after one day the strand will be 30 cm long. Continuing the doubling pattern, the strand will be 60 cm after two days, and then 120 cm after three days.

We can analyze this doubling pattern using the skills we learned in chapter 2. The recursive rule for the length of the strand is: “start with 15 centimeters, and multiply the previous value by 2”. If we think of the starting length as the “stage 0” length, then we can write a rule for the pattern:

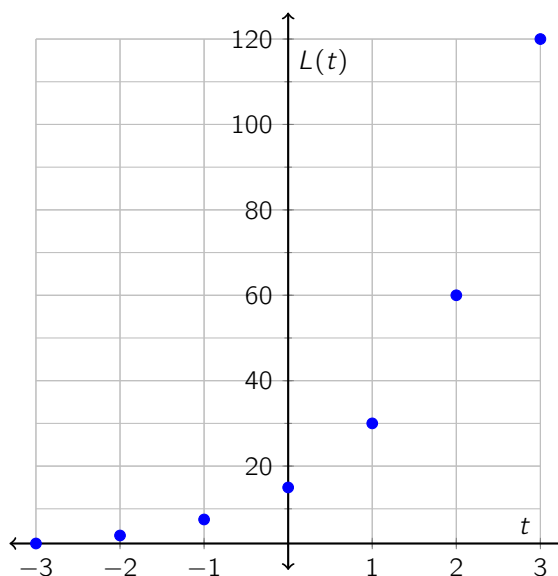
$$L(t) = 15 \cdot 2^t,$$

where $L(t)$ represents the length of the seaweed strand in centimeters after t days.

To make this doubling pattern go “backwards in time”, we must divide by 2. One day ago the strand must have been 7.5 cm long, since doubling this value get us to the original measurement of 15 cm. Going further, the strand was 3.75 cm long two days ago, and 1.875 cm long three days ago.

We can organize our data in a table, and make a scatter plot. Note that we use negative numbers to represent days in the past (days before day 0, on which we took the starting measurement).

Time (days) t	Length (cm) $L(t)$
−3	1.875
−2	3.75
−1	7.5
0	15
1	30
2	60
3	120



From these three representations, it is clear that we have an exponential relationship on our hands: The rule has the independent variable used as an exponent, the table shows repeated multiplication, the graph as that smooth J-like shape that we saw in chapter 4.

A new wrinkle that comes up here, though, is the idea of using negative numbers as the exponent. If we’re right about our rule, for example, then it must be that

$$15 \cdot 2^{-2} = 3.75,$$

which is not an obvious statement! Let’s look more closely at this idea.

10.1.1 Negative exponents

Recall that when we talk about exponents and exponential expressions, we’re talking about expressions of the form: a^b . This is read “ a to the power of b ” or “ a to the b^{th} power”. In such an expression, a is called the *base* and b is called the *exponent*. Since a is the base, we call the whole thing a *power of a* .

One interpretation of this expression, as we have seen, is as repeated multiplication:

$$a^b = \underbrace{a \cdot a \cdot a \cdots a}_{b \text{ times}}$$

This representation is a bit limited — it really only works when the exponent b is a positive integer — but it is still helpful for gaining some intuition about how exponents work.

For clarity, we will name the different forms of the exponential expression. We will use the term **exponential form** to describe a^b , and **expanded form** to describe the form where the multiplication is written out. So, for example, 12^4 is in exponent form, and

$$12 \cdot 12 \cdot 12 \cdot 12$$

is the same number written in expanded form. Of course, if the base is a number, we can compute this value:

$$12^4 = 12 \cdot 12 \cdot 12 \cdot 12 = 20\,736.$$

This last value (the number) is in “standard form” or “decimal form”.¹ Expressions like a^b have no decimal form until we know what numbers a and b represent.

10.1.2 Extending repeated multiplication

Recall that in section 1.4 we mentioned a rule for using zero as an exponent:

Zero exponent

For any nonzero real number a ,

$$a^0 = 1.$$

Our goal now is to extend this idea to include exponents that are negative integers. We’ll do this informally by looking at exponential patterns. We’ll start with a table that just shows the non-negative integral powers of 5:

5^5	$= 1 \cdot 5 \cdot 5 \cdot 5 \cdot 5 \cdot 5$	$= 3125$
5^4	$= 1 \cdot 5 \cdot 5 \cdot 5 \cdot 5$	$= 625$
5^3	$= 1 \cdot 5 \cdot 5 \cdot 5$	$= 125$
5^2	$= 1 \cdot 5 \cdot 5$	$= 25$
5^1	$= 1 \cdot 5$	$= 5$
5^0	$= 1$	$= 1$

Moving “upwards” in this table involves multiplication by 5. Each new row that we pile on top of the table would add 1 to the exponent and put a new factor of 5 in the middle column. If moving upwards in the table means multiplication by 5, then moving “downwards” in the table must mean division by 5, or in other words, multiplication by $\frac{1}{5}$ (the reciprocal of 5).

¹ Note that there are many things in algebra that are called “standard form”. We have to look at context clues to determine which one we need!

So, we can extend the table downwards, as shown below. We take away 1 from the exponent in each step, and introduce repeated multiplication by $\frac{1}{5}$.

5^5	$= 1 \cdot 5 \cdot 5 \cdot 5 \cdot 5 \cdot 5$	$= 3125$
5^4	$= 1 \cdot 5 \cdot 5 \cdot 5 \cdot 5$	$= 625$
5^3	$= 1 \cdot 5 \cdot 5 \cdot 5$	$= 125$
5^2	$= 1 \cdot 5 \cdot 5$	$= 25$
5^1	$= 1 \cdot 5$	$= 5$
5^0	$= 1$	$= 1$
5^{-1}	$= 1 \cdot \frac{1}{5}$	$= \frac{1}{5}$
5^{-2}	$= 1 \cdot \frac{1}{5} \cdot \frac{1}{5}$	$= \frac{1}{25}$
5^{-3}	$= 1 \cdot \frac{1}{5} \cdot \frac{1}{5} \cdot \frac{1}{5}$	$= \frac{1}{125}$
5^{-4}	$= 1 \cdot \frac{1}{5} \cdot \frac{1}{5} \cdot \frac{1}{5} \cdot \frac{1}{5}$	$= \frac{1}{625}$
5^{-5}	$= 1 \cdot \frac{1}{5} \cdot \frac{1}{5} \cdot \frac{1}{5} \cdot \frac{1}{5} \cdot \frac{1}{5}$	$= \frac{1}{3125}$

There are a number of important patterns to see in this table. For example, the pattern “divide by 5 as we move downwards in the table” helps us to explain why $5^0 = 1$. Also note which numbers in the table are reciprocals of one another. This suggests a general rule for working with negative exponents.

Negative exponents

For any nonzero real number a , and any positive integer n ,

$$a^{-n} = \frac{1}{a^n}.$$

In other words, a^{-n} is the reciprocal of a^n (the base raised to the opposite power).

One aspect of negative exponents that can be confusing is that the negative exponent does not change the sign of the evaluated expression. For example, all of the numbers in the table above are positive numbers, even though they have a negative number in the exponent!

Also note that this rule explicitly states that a must be a *nonzero* real number. Can you explain why is this important? Why can't we raise 0 to a negative exponent?

Example 10.1

Write each of the following in standard form (as a fraction in lowest terms).

(a) 36^{-1} (b) 2^{-6} (c) $\left(\frac{3}{4}\right)^{-2}$

Solution: In each case, we can take the reciprocal of the number and raise it to the positive version of

the power, so

$$36^{-1} = \left(\frac{1}{36}\right)^1 = \frac{1}{36}$$

The next two require a bit more computation, but the idea is the same. Problem (b) becomes:

$$2^{-6} = \left(\frac{1}{2}\right)^6 = \left(\frac{1}{2}\right)\left(\frac{1}{2}\right)\left(\frac{1}{2}\right)\left(\frac{1}{2}\right)\left(\frac{1}{2}\right)\left(\frac{1}{2}\right) = \frac{1}{64}$$

And problem (c) becomes:

$$\left(\frac{3}{4}\right)^{-2} = \left(\frac{4}{3}\right)^2 = \left(\frac{4}{3}\right)\left(\frac{4}{3}\right) = \frac{16}{9}$$

It pays to take time when working on problems like this. Consider the next examples carefully!

Example 10.2

Rewrite each of the following expressions in an equivalent form that uses only positive exponents.

$$(a) \ 7m^{-2} \qquad (b) \ \frac{3^{-1}}{4}$$

Solution: We might make a mistake if we rush into these expressions too quickly. Recall that we have to be clear about exactly what an exponent is attached to. In the first expression, the exponent applies only to the m , not the whole expression (there are really two operations here — what are they?). And so,

$$7m^{-2} = 7\left(\frac{1}{m}\right)^2 = \left(\frac{7}{1}\right)\left(\frac{1}{m}\right)\left(\frac{1}{m}\right) = \frac{7}{m^2}$$

In the second expression, the exponent applies only to the 3 in the numerator, and not the whole fraction (there's a subtle grouping symbol here — what is it?). This expression becomes:

$$\frac{3^{-1}}{4} = \frac{\frac{1}{3}}{4} = \frac{1}{3} \div 4 = \left(\frac{1}{3}\right)\left(\frac{1}{4}\right) = \frac{1}{12}$$

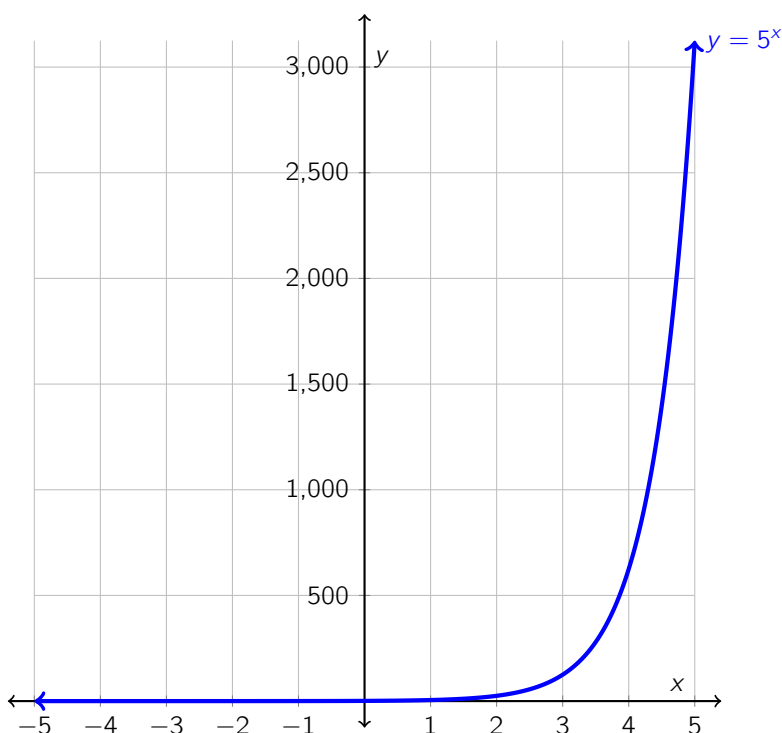
The second expression went through a rather ugly, fraction-in-a-fraction phase, but don't worry: we'll soon learn some simplification rules that will help us make quick work of expressions like these.

10.1.3 Smaller and smaller

An interesting consequence of this negative exponent business is that we get a progression of numbers that get smaller and smaller, but never disappear. For example, in the sequence of powers of 5 above, as n gets larger, the number 5^n grows huge! By the time $n = 10$, the number 5^{10} is nearly 10 million.

This means that 5^{-10} is 1-over-something-close-to-10-million. That number is incredibly small. But note that it is still a positive number. As n gets larger and larger 5^{-n} gets closer and closer to zero, but never actually reaches zero. There is always a tiny fraction which makes the number 5^{-n} a wee bit greater than zero.

If we look at the graph of $y = 5^x$, we see that the curve flattens out as we look to the left, along the negative x -axis. The line approaches 0 but never reaches it. To capture this idea, we say the graph has a *horizontal asymptote* at $x = 0$.²



We can also see how steep the curve gets as we go along the positive x -axis. An important aspect of understanding exponential relationships is understanding this dual nature: incredibly steep on one end, incredibly flat on the other.

Exploring Zeno's paradox

// startup exploration

A paradox is a statement from which we can follow a logical path of reasoning to an impossible or self-contradictory end.

The ancient Greek philosopher Zeno proposed a set of paradoxes about motion. One of these says:

A moving object must arrive at the half-way point on its path before it reaches its goal.

² We'll learn more about asymptotes, and other unusual graph behavior, at the very end of this course and in Algebra 2.

This seems like an obvious statement, right? If you shoot an arrow at a target, the arrow must go half of the distance between you and the target before it goes the whole way to the target!

Let's call the half-way point "point A" and suppose the arrow is there. Now consider this: before the arrow can fly to the target from point A, it has to go half of the remaining distance. So, the arrow arrives at a point three-fourths of the way along the path. Let's call this point B.

Now, before the arrow can go from point B to the target, it has to go half-way to the target. Now it's seven-eighths of the way there.

Before it can cover the remaining distance it has to cover half of the remaining distance. . . and this is always true. So, Zeno says, the result is that the arrow never actually reaches the target: there's always a little but more to go!

In fact, it's even worse! Before the arrow can go halfway to the target, the arrow has to go one-fourth of the way. But before it can do that, it has to go one-eighth of the way. In fact, before the arrow can cover any distance, it has to cover half of that distance. . . and so the arrow never moves in the first place!

Zeno says that movement is impossible, since the above statement about "going half-way" is true. Of course, we know that movement is possible! We do it all the time! This contradiction is what makes the paradox.

Many resolutions to the paradox have been proposed over the years, but it wasn't until the 19th century — around 2000 years after Zeno lived — that mathematics had a clear and rigorous solution to the paradox.


What do you think about these ideas?

10.2 Exponential growth


Cantor set

// startup exploration


German mathematician Georg Cantor suggested a way to construct a very interesting set of numbers. We start in Stage 0 with a line segment covering the interval from 0 to 1 on the number line. The recursive procedure is to remove the middle third from each of the line segments in the previous step. The first few iterations are shown.




Stage 0



Stage 1



Stage 2



Stage 3

Let $f(x)$ represent the number of segments in stage x of this construction. Write an equation for $f(x)$. How many segments are in stage 10 of the figure? What is the first stage to contain more than 100,000 segments?

Bonus puzzler. The set of points from the original line segment that are not deleted at any step of the (infinite) process is called the *Cantor ternary set*. Can you give an example of a point that is in this set? What can we say about the points that survive this process?

10.2.1 Modeling exponential growth

The number of segments in the construction of the Cantor set doubles each time, so the pattern follows the rule “start with 1, multiply the previous value by 2”. We can write out the repeated multiplication in expanded form, or use an exponent.

Stage No.	No. of Segments		
0	1	1	$1 \cdot 2^0$
1	2	$1 \cdot 2$	$1 \cdot 2^1$
2	4	$1 \cdot 2 \cdot 2$	$1 \cdot 2^2$
3	8	$1 \cdot 2 \cdot 2 \cdot 2$	$1 \cdot 2^3$
4	16	$1 \cdot 2 \cdot 2 \cdot 2 \cdot 2$	$1 \cdot 2^4$

If we compare the first and last columns in the table, we can see that the explicit rule for the number of segments depends only on the stage number. We have the rule

$$f(x) = 1 \cdot 2^x.$$

Now that we have an equation, it's quite easy to find the number of segments in stage 10 of the figure:

$$f(10) = 1 \cdot 2^{10} = 1024.$$

To determine the stage in which the figure first has more than 100,000 segments, we need to solve the following equation for x :

$$100,000 = 2^x$$

Hmm. This is a bit of a problem, since we don't have a POE that helps us to get x out of the exponent.³ For now, we'll have to settle for doing a bit of detective work to find this value.

One approach is to use technology. With a graphing calculator or other tool that will display a table of values when given a rule, we can punch in the rule and search through the table for the first row in which the number of segments exceeds 100,000.

If we're working by hand (or working with a regular four-function calculator) we can extend our table manually, or adopt an educated guess-check-revise strategy. In this case, we work forward from $f(10) = 1024$. We don't have to hunt for very long:

$$f(11) = 2048$$

$$f(12) = 4096$$

$$f(13) = 8192$$

$$f(14) = 16\,384$$

$$f(15) = 32\,768$$

$$f(16) = 65\,536$$

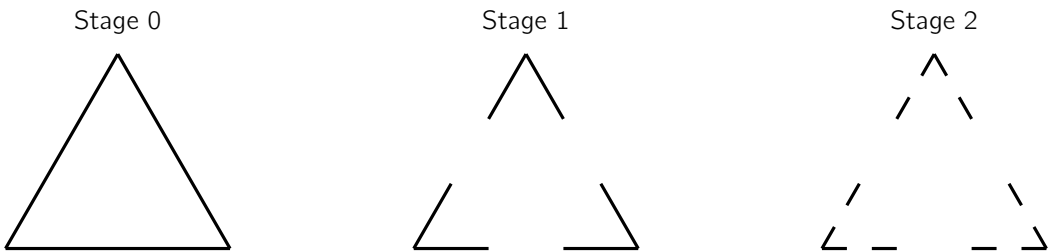
$$f(17) = 131\,072$$

So, stage 17 is the first stage to contain more than 100,000 segments. Are you surprised by this? The figure grows from around 1000 segments in stage 10 to more than 100 times that number of segments less than 10 stages later! Surprising results like this are part of what make exponential relationships so much different from linear relationships.

³ We don't have such a POE yet. There are POEs that will help us in situations like this, but they're going to have to wait until Algebra 2, when we will learn about the concept of a *logarithm*.

10.2.2 Extending the growth model

Suppose we build an equilateral triangle in which each side is a Cantor segment. Then we let $g(x)$ represent the number of segments in stage x of this new figure. Let's try to answer the same questions as above: how many segments in stage 10, and in what stage will the number of segments first exceed 100,000?



The only difference here is that we've changed the initial number of segments. Now, we have the rule "start with 3, multiply the previous value by 2". The updated table is shown below.

Stage No.	No. of Segments		
0	3	3	$3 \cdot 2^0$
1	6	$3 \cdot 2$	$3 \cdot 2^1$
2	12	$3 \cdot 2 \cdot 2$	$3 \cdot 2^2$
3	24	$3 \cdot 2 \cdot 2 \cdot 2$	$3 \cdot 2^3$
4	48	$3 \cdot 2 \cdot 2 \cdot 2 \cdot 2$	$3 \cdot 2^4$

Notice how we only need to change one part of the first rule that we wrote:

$$g(x) = 3 \cdot 2^x$$

To find the number of segments in stage 10, we compute $g(10) = 3 \cdot 2^{10} = 3072$. Then, we do some detective work to find the stage number in which we first exceed 100,000 segments. Stage 15 is close with $g(15) = 98\,304$, but it is stage 16 that first breaks the barrier: $g(16) = 196\,608$.

10.2.3 Exponential equations

If we summarize our findings from the explorations above, we have the "general form" of an exponential function.

General form of the exponential function

A equation of the form

$$f(x) = a \cdot b^x$$

is an exponential function if $a \neq 0$, $b \neq 1$, and $b > 0$. The constant value a is the *initial value* of the function (when $x = 0$). The constant value b is called the *constant multiplier*.

Note that there are some constraints placed on the “legal” values of a and b . Why do we need these restrictions?

Consider what happens when $a = 0$. In this scenario, the function $f(x) = ab^x$ degenerates to $f(x) = 0$. That isn’t an exponential function, it’s a horizontal line along the x -axis! Similarly, if $b = 1$, then $f(x) = ab^x$ degenerates into $f(x) = a$. That’s a horizontal line through the point $(0, a)$!

What about the restriction $b > 0$? Consider what happens when b is negative, for example $f(x) = (-2)^x$. When x is an even integer, $f(x)$ is positive. When x is an odd integer, $f(x)$ is negative. This back-and-forth means that the graph doesn’t have the same smooth curving shape that we associate with the exponential family. Plus, we haven’t even talked about values of x that are not integers! Those have all sorts of other problems!

When b is greater than 1, we have a situation like the ones we have seen in this chapter, called *exponential growth*. Note that b could also be a positive fraction: numbers between 0 and 1 are allowed. In the case that b is a fraction, the exponential behaves a bit differently. We’ll study this situation, called *exponential decay*, in section 10.5.

10.3 Percent change

Earlier explorations have involved an amount doubling as time went by. But what if we have a quantity that grows at a different rate? For example, the United Nations estimates that the population of the Earth will grow by approximately 0.77% per year between now and 2050. How do our rules have to change to accomodate an arbitrary rate of growth?

Hermine's choice

// startup exploration

Hermine is shopping for a mother's day gift. The vase she wants to buy is on sale for 25% off. She also has a 10% off coupon that can be applied to any item in the store. The store manager agrees that she can use both discounts, and gives her two options: *Option 1*: Take 25% off the full price of the vase, then apply the coupon to the discounted price. *Option 2*: Apply the 10% coupon to the full price of the vase, and then take 25% off the discounted price.

Which option should Hermine choose? Why?

Later, Hermine told Bob this story. He said, "Wow! 35% off! You got a great deal!" How would you respond to Bob?

In chapter 6, we discussed the percent proportion:

$$\frac{\text{part}}{\text{whole}} = \frac{\text{percent}}{100}.$$

A key use of percents, as we saw, is to make clear and accurate mathematical comparisons. One comparison that is often illuminating is to compare a certain quantity with itself. If the quantity has increased or decreased over time, we may want to know *by what percent* has it changed?

To calculate a percent change, we compare the amount that a quantity has changed to its starting value. In other words,

$$\frac{\text{amount of change}}{\text{original amount}} = \frac{\text{percent change}}{100}.$$

If we wish to isolate percent change, we can multiply both sides by 100.

Of course, a value may grow larger or smaller, and so we must remember to be clear and describe the percent change as either a "percent increase" or "percent decrease".

Example 10.3

Yardleigh gives Hermine \$300 worth of YardleighCorp stock as a birthday gift. After one month, the value of YardleighCorp stock had increased by 15%. How much were Hermine's shares worth at that point?

After two months, Hermine's shares are worth \$240. What is the percent change from her original \$300 gift? What is the percent change from the value of her shares after one month?

Solution: To compute the value after one month, we find %15 of the value: $300(0.15) = 45$, and then we add this on to find the new total value. So, after one month Hermine's shares are worth $300 + 45 = 345$ dollars.

Since the stock value goes down in the second month, we're looking for percent decrease. The drop from the original value is $300 - 240 = 60$ dollars. We compare this to the original price to find the percent change from that price:

$$\frac{60}{300} = 0.2 = 20\%$$

We can make the same comparison with the price after one month. The amount of the change is $345 - 240 = 105$ dollars, so the percent change from that price is:

$$\frac{105}{345} \approx 0.304 \approx 30\%$$

So Hermine's shares have undergone a 20% decrease from their starting value, and a 30% decrease from their price one month ago.

Note that to calculate the percent increase we found 15% of 300 and then added that on to the original 300. So the new amount is

$$300 + 300(0.15).$$

We can apply the distributive property to this and get

$$300(1 + 0.15) = 300(1.15).$$

If we interpret this as a percent question in its own right, we are saying that the new value is 115% of the original value. This might seem like an unusual statement, but it makes sense: we have 100% of the amount (all of what we started with) and then we add on an additional 15%.

Now consider the case of a percent decrease, such as Hermine's discount coupon in the startup exploration. If Hermine wants to buy a vase that costs \$75 and has a 10% off coupon, we can find 10% of the value of the vase and subtract that from the original value:

$$75 - 75(0.10)$$

Applying the distributive property as we did above, we have:

$$75 - 75(0.10) = 75(1 - 0.10) = 75(0.90)$$

This makes sense, too: if Hermine gets 10% off the price of the vase, then she is only paying 90% of the original price. Armed with these ideas, we can determine which option Hermine should choose in the startup exploration.

Explaining the startup exploration

Solution: It may have been helpful to make up a price for the vase (a convenient number, like \$100), but we'll use a variable here. Let V represent the price of the vase.

If Hermine chooses option 1, she will first get a 25% discount. In other words, the price of the vase will be reduced to 75% of its original value and Hermine would pay $(0.75)V$. The next step is to take this value and apply the 10% coupon. If we take 10% off, that's like paying for 90% of the vase. So in the end the new price is:

$$(0.90)((0.75)V) = ((0.90)(0.75))V = (0.675)V$$

Hermine pays for 67.5% of the vase, which is a 32.5% discount.

If Hermine chooses option 2, she gets the 10% discount first. This makes her cost 90% of the original price. Then we take 75% of this value. So option 2 means the new price would be:

$$(0.75)((0.90)V) = ((0.75)(0.90))V = (0.675)V,$$

which is exactly the same as option 1.

So it doesn't matter which option Hermine chooses. She will pay the same amount for the vase either way. In response to Bob's comment, this is not a 35% reduction in the price, but a 32.5% reduction.

Note how the associative and commutative properties of multiplication play a role in this problem! The field axioms come to the rescue again!

10.3.1 Using growth rate

In section 10.1, we studied a kind of fast-growing seaweed under development at YeardleighCorp labs. Suppose that YeardleighCorp genetic engineers have a different kind seaweed which grows at a rate of 25% per day.

Scientists measure a particular strand to be 6 centimeters long. How long will the strand be in 1 week? How long was the strand 1 week ago? When will this strand be 100 meters long?

To find the length of the strand after one day, we compute $6(1.25) = 7.5$ cm. The “1.25” here means that we’ve got 100% of the original length of the strand, plus an extra 25%. On day 2, the strand will grow an additional 25% and be $7.5(1.25) = 9.375$ cm long.

To find the length day-by-day, we could use the recursive rule “start with 6 centimeters, multiply the previous value by 125%”. We can model this as an exponential equation. If $L(t)$ represents the length of the seaweed in centimeters after t days, then we have

$$L(t) = 6(1.25)^t$$

To find the length of the strand after one week (7 days), we compute

$$L(7) = 6(1.25)^7 \approx 28.610 \text{ cm}$$

To find the length one week ago, we compute

$$L(-7) = 6(1.25)^{-7} \approx 1.258 \text{ cm}$$

How long before the strand is 100 m long? Our rule gives the length in centimeters, so we want to know when the strand is 10,000 cm long.⁴ If we do a bit of detective work, perhaps using some graphing technology to examine a table of values, we can discover:

$$f(33) \approx 9466.331 \text{ cm}$$

$$f(34) \approx 11\,832.914 \text{ cm}$$

So this strand of seaweed reaches 10,000 cm in length sometime between day 33 and day 34. For the record: this kind of approximation is a perfectly acceptable way to answer for this type of question, in Algebra 1.⁵

You may be suspicious about the fact that we predict a 6 cm piece of seaweed will grow to be longer than an American football field in a little more than one month. A note on realism: Exponential functions assume unlimited growth. But, in the real world, this type of growth cannot be sustained: a piece of seaweed cannot grow to be 10 miles long, a flock of geese cannot grow to include millions of members. Constraints (like the availability of food or space) restrict unlimited growth. The kind of function that models the way things *really* grow is a bit too complicated for Algebra 1. For our purposes, we will assume that the creatures and businesses we study have unlimited resources and can grow without bound.

We summarize our approach here, and present a second way to express the equation for exponential growth.

⁴ Recall that there are 100 centimeters in one meter.

⁵ If you’re feeling detail-oriented, you can use your calculator to narrow down the answer by adjusting the table settings to that x -values increase by 0.1 or 0.01, rather than by 1 whole unit. If you want to go there, you can actually narrow down the interval considerably (if we assume the seaweed grows uniformly, then the actual value is around 33.246 days). This is not required: an approximation between two integers is fine for now.

Exponential growth, given growth rate

A equation of the form

$$f(x) = a \cdot (1 + r)^x$$

is an exponential function where a is the initial value, and r is the growth rate (the percent increase each “generation”).

If we compare this form of the equation to the previous form, $y = ab^x$, we can see that $b = (1 + r)$. This value is called the *growth factor*. Be mindful that “growth rate” and “growth factor” are different the “growth factor” is “1 + growth rate”.

10.4 Personal finance

This material in this section is quite possibly the most important thing we will discuss in all of Algebra 1! The topics here apply not just to students who plan on studying science, technology, engineering, mathematics (or other “mathematically intense” disciplines). The concept of personal finance is important for anyone who plans to have a job or a bank account, anyone who wants to own a car or a home.

10.4.1 Interest

Cars are expensive. Most people don’t have enough money lying around to just go out and buy a car. Instead, most people enter in to an agreement with a bank. The bank agrees to lend you the money you need to buy the car, and you agree to pay back that money over time.

Of course, a bank is a business and banks want to make money. So, they charge you for the “convenience” of being able to get all the money you need, all at once. This means that the bank requires you to pay back *more* than you borrowed. This extra amount you will have to pay is called **interest**.

Interest is usually calculated as a percentage of the amount you borrowed. The percentage is called the **interest rate**. The interest rate is usually given as an annual percentage rate, or APR (you might have heard this term in a car commercial).

There are two kinds of interest in the world: good interest and bad interest.⁶ The “extra money” that you have to pay to the bank is “bad interest”. On the other hand, some savings accounts earn interest, meaning that every so often the bank will add on a percentage of however much money is in the account. This is “good interest”: free money!

As a future consumer, it is important to learn about how interest works.⁷ A good rule of thumb is to try and minimize “bad interest” which works *against* you, and to try and get as much “good interest” as possible working *for* you.

10.4.2 Simple interest

The simplest way to calculate interest is called, naturally enough, **simple interest**. It is calculated based only on the original amount borrowed or invested, called the **principal amount**, or just *principal*.

⁶ There are 10 kinds of people in the world: People who understand binary numbers and people who don’t.

⁷ When you go to the bank to get a car loan, you won’t have to do all the math yourself. The bank will do the interest calculations up front and tell you how much you should expect to pay. But: credit cards are different, and — in the opinion of the authors — much more sinister.

Simple interest

We calculate simple interest using the formula

$$I = Prt,$$

where I represents the interest earned, P represents the principal (the initial amount invested or borrowed), r represents the annual percentage rate, and t represents the time (in years).

The formula $I = Prt$ only calculates the interest, meaning the extra amount that we add on. The formula doesn't tell us the total amount that we own (or owe). To calculate the total amount, we can create an alternative formula for A the total amount:

$$A = P + Prt.$$

Let's take an example: Suppose Bob wants to buy a car and needs to borrow \$20,000. The bank offers Bob 6.5% APR on a five-year loan. In other words: at the end of each year, 6.5% of \$20,000 will be added to the amount Bob owes, and Bob agrees to pay back the total amount in 5 years. Table 10.1 shows a schedule of how Bob's balance changes over the five years.

No. Years (x)	Calculations	Total Amount (y)
0	—	20000
1	$20000 + 20000 (0.065) (1)$	21300
2	$20000 + 20000 (0.065) (2)$	22600
3	$20000 + 20000 (0.065) (3)$	23600
4	$20000 + 20000 (0.065) (4)$	25200
5	$20000 + 20000 (0.065) (5)$	26500

Table 10.1: Calculating simple interest of 6.5% on \$20000.

In the end, Bob will sign a contract agreeing to pay back a total of \$26,500, which is \$6500 more than the amount he wants to borrow. This is the “extra money” that goes to the bank.

Notice that the total goes up by the same amount each year. This makes sense, because the calculation for each year is done based on the original 20,000. The growth in the total amount is linear.

Let's compare this to the evil-genius that is *compound interest*.

10.4.3 Compound interest

The key idea of compound interest is that it isn't calculated based just on the principal amount, but based on the principal amount *plus all of the interest that has been earned up to that point*.

Let's see how this impacts Bob. Suppose his \$20,000 car loan has an APR of 6.5% "compounded annually". Each time we calculate his interest, we will need to add on 6.5% of the previous year's total amount, not just the principal amount.

No. Years (x)	Calculations	Total Amount (y)
0	—	20000.00
1	$20000.00 (1.065)$	21300.00
2	$21300.00 (1.065)$	22684.50
3	$22684.50 (1.065)$	24158.99
4	$24158.99 (1.065)$	25729.33
5	$25729.33 (1.065)$	27401.74

Table 10.2: Calculating (annual) compound interest of 6.5% on \$20000.

In the end, Bob will sign a contract agreeing to pay a total of \$27,401.74. That's \$7401 more than he borrowed, and \$901.74 more than he would pay in the simple interest scenario! That's the incentive banks have to use compound interest!

Notice that to calculate the amount we owe in the next year, we take the current year and multiply by 1.065. In other words, we have the rule $\text{NEXT} = \text{NOW} \cdot 1.065$, and that's an exponential relationship!

Over the 5 years of his car loan, Bob pays an extra \$900 or so compared to simple interest. That's bad news, but it might not hurt Bob's bottom line that much: it's just \$15 extra per month. This is part of the evil genius of compound interest: the difference isn't huge in the short run. If there were a huge difference, then no one would agree to pay compound interest! We would all protest and boycott organizations that used it!

Those that know how to use compound interest, however, know that they can change what is known as the "period of compounding." This changes how often interest is calculated. Rather than calculating interest once a year, we break up the annual percentage into smaller pieces and calculate the interest more often.

Let's return to the story of Bob and his car loan, and suppose that interest is compounded semi-annually. That means that we take the 6.5% APR, cut it into two pieces, and calculate interest twice a year.

No. Years (x)	Calculations	Total Amount (y)
0	—	20000.00
1	$20000(1.065/2)^2$	21321.13
2	$20000(1.065/2)^4$	22729.52
3	$20000(1.065/2)^6$	24230.95
4	$20000(1.065/2)^8$	25831.55
5	$20000(1.065/2)^{10}$	27537.89

Table 10.3: Calculating (semi-annual) compound interest of 6.5% on \$20000.

The exponents here show that in one year we have gone through 2 compounding periods, in two years we have gone through 4 compounding periods, and so on. Generally, if we compound interest semi-annually, we will have $2t$ compounding periods in t years.

Notice that the difference is just a little bit higher at the end of 5 years: around \$136. Now, let's look at what happens when we change to monthly compounding, meaning we divide the interest rate into 12 pieces, but that we recompute interest 12 times per year.

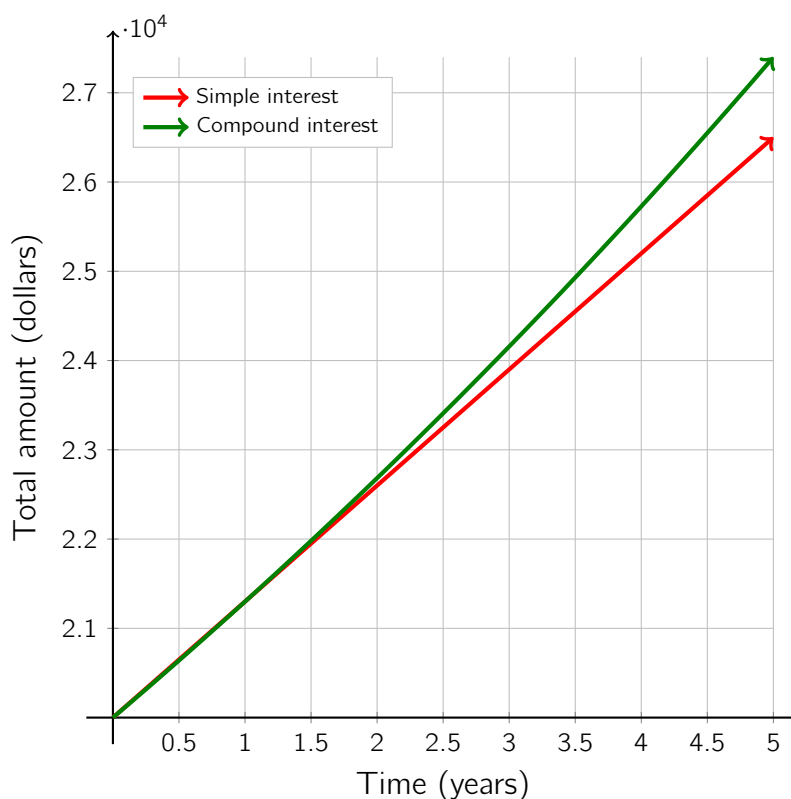


Figure 10.1: Comparing 6.5% simple and compound interest over 5 years.

After 5 years, Bob will owe

$$20000(1.065/12)^{5 \cdot 12} = 27656.65$$

which is slightly more than anything we've seen so far. What happens if we change to daily compounding? In this case the interest rate will be divided into 365 pieces (let's ignore leap years) and we will calculate interest 365 times per year.⁸ After 5 years, Bob will owe

$$20000(1.065/365)^{5 \cdot 365} = 27679.81$$

Once again, notice that the difference isn't that great compared to the other calculations we've made. But, it's higher than anything we've computed so far. Bob would end up owing the bank \$7679.81 more than he borrowed!

We can distill these calculations into a formula for compound interest.

⁸ Daily compound interest is how credit card companies charge interest — seriously!

Compound interest

We calculate compound interest using the formula

$$A = P \left(1 + \frac{r}{n} \right)^{nt}$$

where A represents the total amount, P represents the principal, r represents the annual interest rate, t represents the time in years, and n represents the number of compounding periods per year.

For example in the scenario of “monthly compounding”, $n = 12$ and so we use the equation

$$A = P \left(1 + \frac{r}{12} \right)^{12t}$$

to compute interest after t years. Note that we have $\frac{r}{12}$, meaning we’ve divided the rate into 12 pieces, and $12t$ in the exponent meaning that in each year t we have 12 compounding periods.

Credit cards and compound interest

Credit card companies charge compound interest on your balance, and interest is compounded daily. Plus, credit cards often charge a higher interest rate than other types of financial products (car loans and student loans generally charge a much lower rate of interest).

These two things — high interest rates and daily compounding — combine to make it very hard to pay off large amounts of credit card debt. It’s troubling that the average American consumer owes several thousand dollars in credit card debt, with American consumers together owing billions of dollars to credit card companies.

10.4.4 (;,;) Continuous compounding

At this point, someone usually wonders whether a bank could make an unlimited amount of money using this technique. What happens if a bank were so evil that they compounded interest every minute of every day? What about every second of every day? What about every millisecond?

Interestingly, there is a limit to the amount of money that a bank can rake in using compound interest. Notice that the amounts we have seen — under annual, semi-annual, monthly, and daily compounding — have all been increasing, but that they have been increasing by less and less. The increases are sort of “trailing off” and, over time, the change between compounding schemes gets closer and closer to zero.

Exploring continuous compounding

// startup exploration

Imagine that we invest a principal amount P at an evil bank which charges 100% interest. Experiment with the amount we would owe after one year given different compounding periods. What is the maximum that we would be expected to pay back after one year?

This scenario — continuous compounding for one year at 100% interest — gives rise to a very interesting number which will be of great importance in future mathematics courses, including Algebra 2.⁹

⁹ Actually, this number has already been mentioned in the *Algebramicon*! Go on a little scavenger hunt through the footnotes in chapter 4 and see what you can find!

10.5 Exponential decay

No pressure

// startup exploration

Patrons to the Cheeseville Zoo can take a hot air balloon ride over the park. They have to fly high enough to be at a safe distance above the most dangerous areas of the zoo: Lake Mascarpone (the reptile exhibit) and NAME THE ISLAND (the primate enclosure).

Of course humans can't fly too high: atmospheric pressure (the pressure of air around you) decreases as altitude increases, by about 6% for every 500 meters. Above a certain altitude, people need breathing aids if they aren't in a pressurized environment (like an airplane).

Mountain climbers who scale Mount Everest say that "altitude sickness" sets in around 2500 m, and they say that climbing about 8000 m is entering the "death zone" (a little morbid, but an effective deterrent).

What percentage of normal air pressure is safe for humans? Specifically: at what percentage does a person risk altitude sickness? At what percentage does a person risk death?

At an altitude of 0 meters (sea level), we experience 100% of normal atmospheric pressure. When we rise to a height of 500 meters, the atmospheric pressure decreases by 12%. In other words, we are experiencing approximately 88% of normal air pressure at that altitude.

Rising to 1000 feet, we reduce by an additional 12%. To calculate the new atmospheric pressure value, we could find 12% of 88%, and subtract this from the original 88%:

$$88 - 88(0.12) = 88 - 10.56 = 77.44$$

Or, we could think about the percent remaining and find 88% of the previous value 88%:

$$1 \cdot 0.88 \cdot 0.88 = 1 \cdot (0.88)^2 = .7744 = 77.44\%$$

This second approach tells us that we've got an exponential relationship. This time things are a little different because the numbers are decreasing, but the general idea is the same.

If we let $P(a)$ represent the atmospheric pressure at altitude a , then our relationship is

$$P(a) = 1 \cdot (0.88)^a$$

Note that the altitude value goes in steps of 500 meters, so when $a = 2$ that represents an altitude of $2 \cdot 500 = 1000$ meters.

To calculate the pressure at 2500 meters, we should notice that this is 5 "steps" of 500 meters each. So, $a = 5$

and the pressure at that altitude is:

$$(P(5) = 1 \cdot (0.88)^5 \approx 0.528$$

According to our model, altitude sickness sets in at around 52.8% of normal atmospheric pressure. This number seems overly precise; it's probably better to say "around 50%".

The death zone is at 8000 meters, or when $a = 16$ in our model. The pressure at that altitude is:

$$P(16) = 1 \cdot (0.88)^{16} \approx 0.129$$

Our model predicts that atmospheric pressure of around 13% is life-threatening.¹⁰

10.5.1 Exponentially decreasing

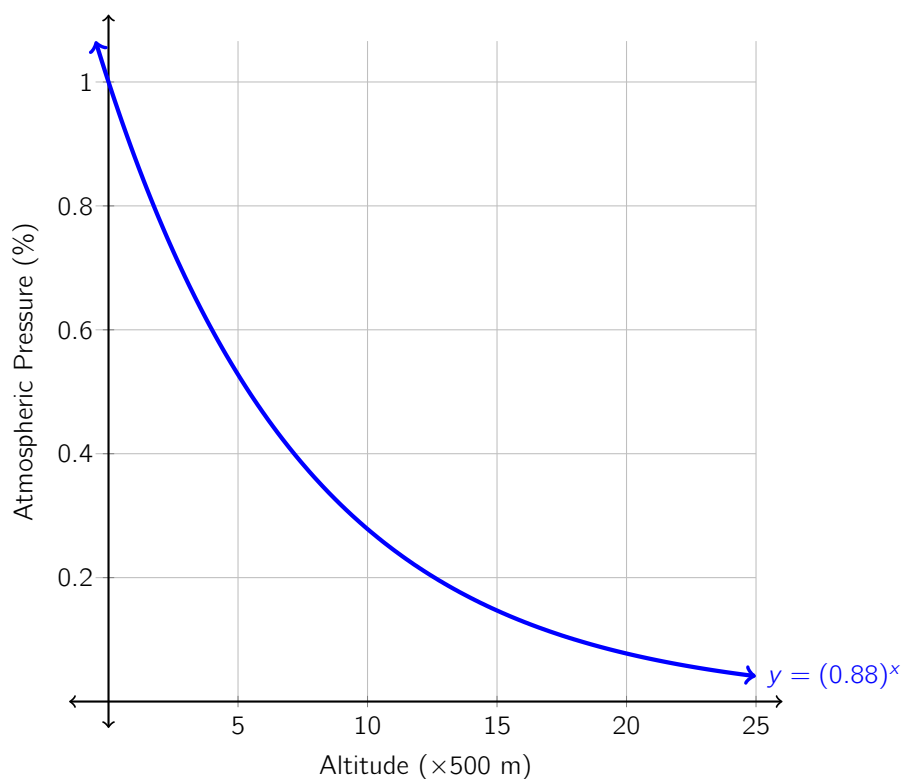
Exponential decay describes a relationship which decreases exponentially. These have the same general formula for exponential relationships — we have an initial value a and a constant multiplier b — but now the constant multiplier is a fraction between 0 and 1.

Exponential decay

An equation of the form $f(x) = a \cdot b^x$ represents exponential decay when $a \neq 0$ and $0 < b < 1$.

The graph of our atmospheric pressure example shows the familiar J-like shape, only backwards: a decreasing trend.

¹⁰ For reference, most commercial airlines cruise at an altitude of around 12,000 m and pressurize their cabin to approximate an altitude of around 2100 m.



As with exponential growth, we distinguish between *decay factor* and *decay rate*. The decay rate is the percent by which some value decreases. It tells us the amount that has “disappeared”. The decay factor is another word for the constant multiplier. It will tell us what to multiply by to get what is “left over”.

If air pressure decreases by 12% per step change in elevation, then the decay rate is 12%. This means that 88% of the pressure “remains”, and so this is the decay factor. The relationship between these two values is straight forward: $\text{decay factor} = 1 - \text{decay rate}$. One way to write an equation for exponential decay uses the decay rate.

Definition

Exponential decay is described by an equation of the form

$$f(x) = a(1 - r)^x$$

where a represents the initial value, and r represents the decay rate (a percent, written as a decimal).

Example 10.5

In the late 1990's, the Cheeseville Zoo saw some hard times. After a record high attendance of 1,000,000 visitors in 1995, attendance declined at a rate of 26% per year. Approximatley how many people attended the zoo in 2000? In what year did attendance first drop below 100,000?

Solution: A decay rate of 26%, means we can write a rule using the decay factor $1 - 0.26 = 0.74$ or 74%. So, let's create a rule

$$A(t) = 1\,000\,000(0.74)^t$$

to represent the attendance after t years (where 1995 is "year 0").

In this model, the year 2000 is when $t = 5$. In that year, the attendance was:

$$A(5) = 1\,000\,000(0.74)^5 \approx 221\,900$$

To find when attendance first dropped below 100,000, we have to do some hunting. We find the two values:

$$A(7) \approx 121\,513$$

$$A(8) \approx 89\,919$$

This tells us that "year 8", meaning calendar year 2003, was the first time attendance dropped below 100,000 visitors.

10.5.2 Linear versus exponential change

Before this chapter, we spent most of our time discussing linear equations. We have seen situations that can be represented quite effectively using linear functions. For example, exchanging money between two different currencies, the amount a person might earn from working a certain number of hours at a particular pay rate, or how much distance a person can cover when moving at a constant speed.

As we have seen, though, many important natural phenomena don't follow straight lines. Fundamental aspects of biology and physics are build around relationships that follow curved lines: population growth over time, the decrease of atmospheric pressure with altitude.

It's sometimes hard to appreciate the difference between a curved line and a straight line, and it might not seem like all that big of a deal. After all, a straight line could keep pace with an exponential curve, if the line were steep enough. . . right?

To compare linear and exponential growth, we close this chapter with a famous fable first told by the Persian poet Ferdowsi.

Exploring the chessboard fable

// startup exploration

Many centuries ago, so the story goes, the inventor of the game of chess showed his invention to the king. The king was so delighted by the new game that he told the inventor that he could have whatever he liked as compensation.

The inventor proposed that the king give him one grain of wheat for the first square of the chessboard, two grains of wheat for the second square, four grains of wheat for the third square, and so on, doubling the number of grains for each of the 64 squares on the board.

The king was shocked that the inventor should ask such a low price and offered instead to pay *one million grains of wheat* for every single square of the chessboard! The inventor refused, saying he much preferred his original proposal.

The king agreed to pay the inventor's price, chuckling to himself that the foolish man had asked for so little.

But when the time came for the king to pay the debt, he was no longer chuckling. What happened?

Suppose we use some algebra to compare the different proposals. The king is offering to pay 1 million grains of rice every for every square on the chessboard. We can represent his proposal with the equation

$$f(x) = 1\,000\,000\,x,$$

where $f(x)$ is number of grains of wheat that the inventor earns from x squares on the chessboard. This is a straight line with slope of one million. The enormous slope means that the line is so steep it would be indistinguishable from a vertical line on most graphs.

The inventor's proposal is a relatively tame exponential function: start with 1, double the previous value. That is the function

$$g(x) = 2^x,$$

where $g(x)$ represents the number of grains on square x of the chessboard. Note that $f(x)$ and $g(x)$ are measuring different things: f gives us a running total after x squares, whereas g gives us the amount sitting on the single square x (to find the total number of grains, we'd have to add up these values).

At first glance, there hardly seems to be a competition between these two: the super-steep linear function should dominate the mild exponential function. . . But have a close look at the graphs in section 10.5.2, which show the behavior of these graphs across 8, 16, 32, and 64 squares of the chessboard. What are these graphs communicating?

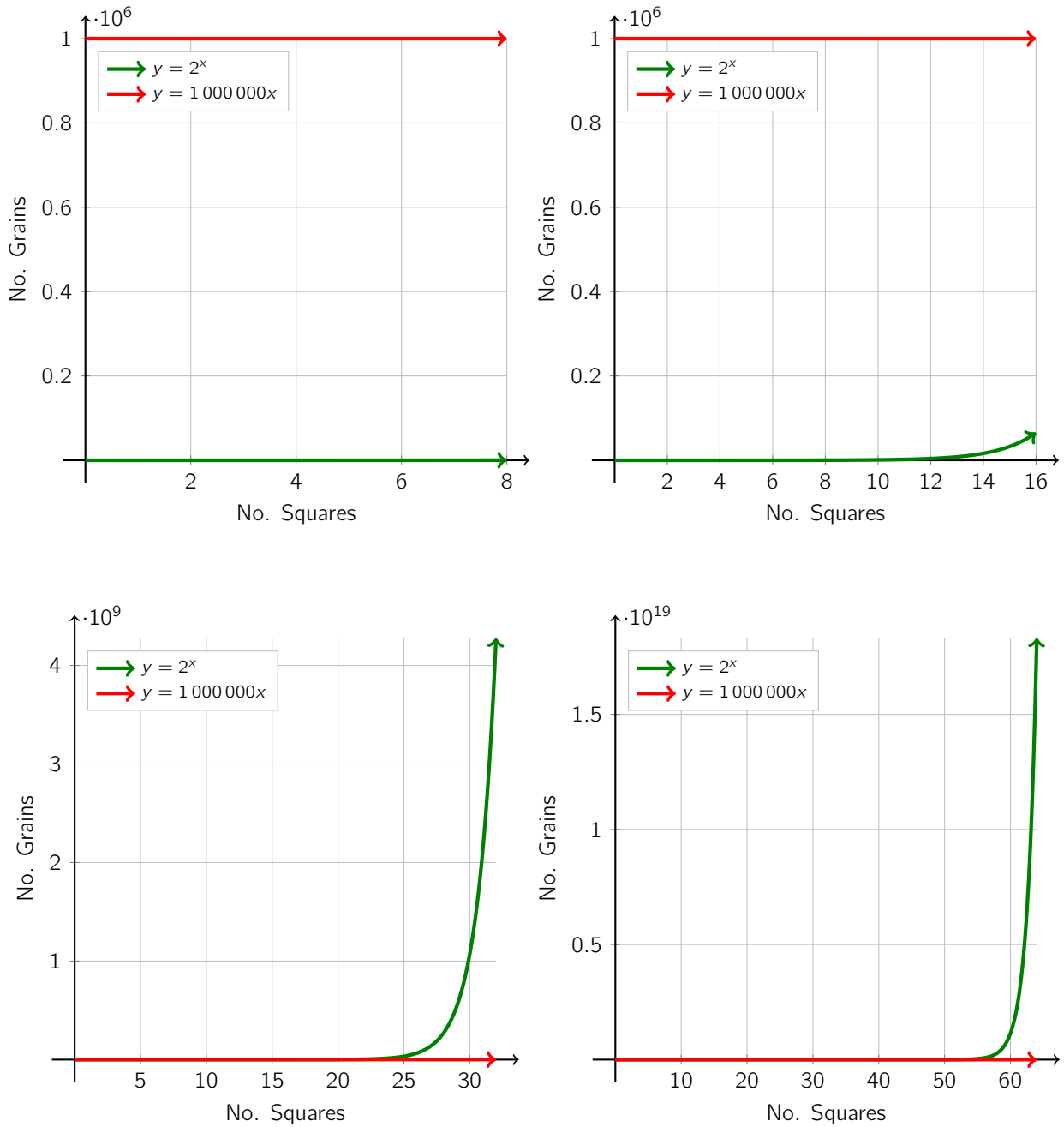


Figure 10.2: Growth in the chessboard fable: 8, 16, 32, and 64 squares.

» Chapter summary «

In this chapter we uncovered some of the surprising differences between the linear and exponential families. We saw examples of exponential growth and decay, including the very practical application of percent change and exponential growth in the context of personal finance.

In the course of our work in this chapter, we have encountered expressions that we have been content to leave “unsimplified”. For example, we have been okay with leaving the parentheses alone in an expression like $(1 + r)^t$, whereas in earlier chapters, parentheses were often taboo.

How do we eliminate the parentheses from an expression like this? In the next chapter, we will expand our set of equivalence rules to handle the simplification of algebraic expressions that involve exponents. Onward!

To divide a cube into two other cubes, a fourth power or in general any power whatever into two powers of the same denomination above the second is impossible, and I have assuredly found an admirable proof of this, but the margin is too narrow to contain it.

Pierre de Fermat, French lawyer and mathematician

Chapter 11

Exponential expressions

We have quite a collection of tools in our algebraic toolbox. Until now, though, we haven't needed many rules for simplifying exponents. We know how exponents behave under the order of operations, of course, but in that case we are simplifying a numerical expression. Things are pretty straightforward when our expression involves only numbers, and becomes a single number in its simplest form.

Our task in this chapter will be to derive a set of rules that can be used to simplify expressions where numbers, variables, and exponents all get tangled up together.

11.1 Fundamentals of exponents

When we talk about exponents and exponential expressions, we're talking about expressions of the form: a^b . This is read " a to the power of b " or " a to the b^{th} power". In such an expression, a is called the **base** and b is called the **exponent**. We call the whole thing a **power** of a , since a is the base.

One interpretation of this expression, as we have seen, is a repeated multiplication:

$$a^b = \underbrace{a \cdot a \cdot a \cdots a}_{b \text{ times}}$$

We say that the left-hand expression is written in **exponential form**, and that the right-hand side is written in **expanded form**.

Reminder: One of the trickiest concepts in Algebra 1

We first saw this tricky little item back in chapter 1, and haven't really had to deal with it again until now. Recall the difference between the following three expressions:

$$(-2)^4 \quad \text{and} \quad -2^4$$

If we write these out in expanded form we have

$$(-2)^4 = (-2)(-2)(-2)(-2) = 16.$$

On the other hand,

$$2^4 = 2 \cdot 2 \cdot 2 \cdot 2 = 16.$$

Be careful not to jump to conclusions, though! Compare these two:

$$-2^3 = -2 \cdot 2 \cdot 2 = -8 \quad \text{and} \quad (-2)^3 = (-2)(-2)(-2) = -8$$

Whenever we are asked to simplify an exponential expression, a good habit of mind is to ask ourselves, "What, exactly, does the exponent apply to?"

11.2 Simplified algebraic expressions: Exponents edition

In earlier chapters we learned properties that could be used to simplify linear expressions. The criteria we learned in chapter 5 for “simplified algebraic expressions” still apply, but now that we are dealing with exponents, there are a few new criteria that we need to add. We will summarize the criteria here and then learn the details over the next few sections.

Simplified exponential expressions

An exponential expression is considered completely simplified if. . .

1. It contains only positive exponents
2. Bases appear at most once per term
3. Powers with a numeric base are written in standard form (if that’s reasonable)
4. It contains no explicit grouping symbols

Criteria #1 asks for only positive exponents. In section 10.1.1, we saw how to use the reciprocal to rewrite a negative exponent using a positive exponents. We’ll see more about negative exponents in section 11.4.

Criteria #2 prohibits a term to be written in the form $x^2 \cdot x^3$, since this is a term that has the same base (x) appearing twice. In a way, we’re looking for a kind of “combining like terms” that will allow us to write this as x raised to a single exponent. Properties of this kind will be the main focus in this chapter.

Criteria #3 asks us to evaluate expressions like 2^3 and write 8 instead. . . if that’s reasonable. We have already seen that exponential expressions can yield some enormous numbers, so it’s better to write 6^{12} than 2 176 782 336. Use your best judgement here: a number that’s larger than, say, five digits should probably be left in exponential form. It’s also best to continue our convention of using simplified improper fractions instead of decimals.

Criteria #4 is already in effect: explicit grouping symbols have been banned from simplified algebraic expressions since chapter 5. We mention it again here because some special care is required when it comes to eliminating groupers that are tangled up with exponents.

Remember: Order of operations

The expression $2 \cdot 3^x$ is completely simplified. We might be tempted to write 6^x , but this would be **Evil and Wrong**. Since exponents must be evaluated before multiplication, it is impossible to fully evaluate this expression, We

don't know what x is, and so it is impossible to resolve the expression 3^x , and so we cannot do the multiplication. So, we stop here.

Exponent properties

In the next few sections we will derive the properties for simplifying exponential expressions and achieving simplified exponential form.¹ With each new property, work out the initial examples and think about how to state a general simplification rule.

¹ These aren't all of the definitions and properties for exponents, but they are the key ones we need for now. We will learn a few more tools in later chapters when we look at quadratics, polynomials, and radicals. There are also a few saved for Algebra 2, when we will see non-integer exponents and logarithms.

11.3 Product properties

11.3.1 Product rule

Derivation #1

// startup exploration

Use expanded form to write each of the expressions below as a single base raised to a single exponent.

$$x^4 \cdot x^3 \quad \text{and} \quad z^5 \cdot z^8$$

Make a conjecture about a general rule for simplifications of this type.

This shortcut is called the product rule, since it applies to the situation in which we are finding the product of two exponential expressions with the same base.

We can write the first example out in expanded form, and then rewrite the long string of x 's under a single exponent:

$$x^4 \cdot x^3 = x \cdot x \cdot x \cdot x \cdot x \cdot x \cdot x = x^7$$

The same goes for the z 's in the second expression:

$$z^5 \cdot z^8 = z \cdot z \cdot z \cdot z \cdot z \cdot z \cdot z \cdot z \cdot z \cdot z \cdot z = z^{13}$$

The new exponent will be the sum of the original exponents!

Product rule of exponents

For any nonzero real number a and any integers m and n ,

$$a^m \cdot a^n = a^{m+n}.$$

The intuition behind this rule is that $a^m \cdot a^n$ means we have m factors of a multiplied by n factors of a , and so we have $m + n$ factors of a in all. That means a^{m+n} . This is pretty straightforward! There's really no need to "memorize" this rule. We can always re-derive it if we forget exactly how it works.

Example 11.1

Simplify $3b^2 \cdot 4b^5$.

Solution: In this example, we will need to use the commutative property of multiplication to move the coefficients together and the variable parts together. In other words:

$$3b^2 \cdot 4b^5 = 3 \cdot 4 \cdot b^2 \cdot b^5 = 12b^{2+5} = 12b^7.$$

11.3.2 Power rule**Derivation #2**

// startup exploration

Use expanded form to write each of the expressions below as a single base raised to a single exponent.

$$(x^3)^4 \quad \text{and} \quad (z^5)^8$$

Make a conjecture about a general rule for simplifications of this type.

This simplification is called the power rule, since it applies to situations in which we have a power which is itself being raised to a power.

Writing the first expression in expanded form gives us a product. From there we can apply the product rule that we learned in the last section:

$$(x^3)^4 = x^3 \cdot x^3 \cdot x^3 \cdot x^3 = x^{3+3+3+3} = x^{4 \cdot 3} = x^{12}$$

Notice that we get a repeated addition in the exponent, which we can rewrite as a multiplication. In the second expression, we have:

$$(z^5)^8 = z^5 \cdot z^5 \cdot z^5 \cdot z^5 \cdot z^5 \cdot z^5 \cdot z^5 \cdot z^5 = z^{8 \cdot 5} = z^{40}$$

The new exponent is the product of the original exponents!

Power rule of exponents

For any nonzero real number a and any integers m and n ,

$$(a^m)^n = a^{n \cdot m}$$

In this case, the intuition is that the outermost exponent means we will multiply together n factors of a^m . Since each of those factors has m factors of a , we have “ n groups of m factors of a ” in all, which is another way of saying $n \cdot m$ factors of a .

Example 11.2

11.3.3 Product to a power

Derivation #3

// startup exploration

Use expanded form to write each of the expressions below without parentheses.

$$(xy)^4 \quad \text{and} \quad (abc)^8$$

Make a conjecture about a general rule for simplifications of this type.

In this case, we are taking the power of a product so we call this shortcut “power of a product” or “product to a power”. We can write out the first expression in expanded form, and then use the commutative property to rearrange the factors:

$$(xy)^4 = (xy)(xy)(xy)(xy) = xy \, xy \, xy \, xy = xxxx \, yyyy = x^4 y^4$$

The second example works just the same way:

$$\begin{aligned} (abc)^8 &= (abc)(abc)(abc)(abc)(abc)(abc)(abc)(abc) \\ &= abc \, abc \, abc \, abc \, abc \, abc \, abc \, abc \\ &= aaaaaaaaaa \, bbbbbbbb \, cccccccc \\ &= a^8 b^8 c^8 \end{aligned}$$

We can remove the parentheses by attaching the exponent to each of the factors inside!

Product to a power rule

For any nonzero real numbers a and b and any integer m ,

$$(ab)^m = a^m b^m$$

The idea is that the exponent gives us m copies of the product, which means m copies of each of the factors in that product. The commutative property of multiplication allows us to move these factors around and regroup them.

Note that the thing in the parentheses is a product, *not a sum*. We'll learn how to handle sums raised to powers, for example $(x + y)^2$, in section 11.5.

Example 11.3

—

11.4 Quotient properties

The rules in the last section were all based on multiplication. We'll look at some division-related laws, which behave in a very similar way (since multiplication and division are inverse operations). We pull these rules out into their own section since they bring us into closer contact with negative exponents, which take some getting used to.

11.4.1 Simplifying negative exponents

Derivation #4

// startup exploration

Use the definition of a negative exponent to write each of the expressions using only positive exponents.

$$\frac{a^{-2}}{b} \quad \text{and} \quad \frac{2x^3y^{-2}}{m^{-5}q^2}$$

Make a conjecture about a general rule for simplifications of this type.

We'll use the definition of negative exponents to work these out, but sometimes that can involve numerous steps. For example in the first expression, we have:

$$\frac{a^{-2}}{b} = \frac{\frac{1}{a^2}}{b} = \frac{1}{a^2} \div b = \frac{1}{a^2} \cdot \frac{1}{b} = \frac{1}{a^2b}$$

After going through all of that work, we simply end up with a^2 in the denominator. In the second example, we can see that the 2, the x^3 , and the q^2 have positive exponents already and are going to stay right where they are. Let's extract the expressions with negative exponents and treat those separately:

$$\frac{2x^3y^{-2}}{m^{-5}q^2} = \frac{2x^3}{q^2} \cdot \frac{y^{-2}}{1} \cdot \frac{1}{m^{-5}} = \frac{2x^3}{q^2} \cdot \frac{1}{y^2} \cdot \frac{m^5}{1} = \frac{2x^3m^5}{q^2y^2}$$

When we see a negative exponent, we can move that base to the other part of the fraction and change the sign of the exponent.

Beware of the headless fraction

We may write a number without a denominator. The number 41, for example, has no denominator — or rather, it has a phantom 1 in the denominator which we don't write unless we need to.

We cannot write a number without a numerator. For example, if we are asked to simplify the expression g^{-6} , we can use the shortcut and put g^6 in the other part of the fraction, but now we have an empty numerator. . .

$$g^{-6} = \frac{\quad}{g^6} \quad \leftarrow \text{ what goes here?}$$

We can't leave the numerator empty because then we don't have a number that makes sense. "Empty numerator" might imply "numerator = 0", but that would make the whole fraction equal to 0, which would be wrong: if $g \neq 0$, then $\frac{1}{g} \neq 0$.

It must be that the numerator is 1. Indeed this makes sense, for we can introduce a phantom 1 before we apply the shortcut:

$$g^{-6} = 1 \cdot g^{-6} = \frac{1}{g^6}$$

Notice that in the first example of derivation #4, we filled the empty numerator with 1, just as we did here.

11.4.2 Quotient rule

Derivation #5

// startup exploration

Use expanded form to write each of the expressions below as a single base raised to a single exponent.

$$\frac{m^6}{m^2} \quad \text{and} \quad \frac{z^5}{z^8}$$

Make a conjecture about a general rule for simplifications of this type.

This simplification is called the quotient rule, since it applies when we are finding the quotient of two exponential expressions with the same base.

When we write out the first expression in expanded form, we have some common factors of x which "cancel" from the numerator and denominator.

$$\frac{m^6}{m^2} = \frac{m \ m \ m \ m \ m \ m}{m \ m} = \frac{m \ m \ m \ m \cancel{m} \cancel{m}}{\cancel{m} \cancel{m}} = \frac{m \ m \ m \ m}{1} = m^4$$

Here, we replace the empty denominator with a soon-to-be phantom 1, which promptly vanishes. In the second example, the leftovers are in the denominator.

$$\frac{z^5}{z^8} = \frac{z \ z \ z \ z \ z}{z \ z \ z \ z \ z \ z \ z \ z} = \frac{\cancel{z} \ \cancel{z} \ \cancel{z} \ \cancel{z} \ \cancel{z}}{\cancel{z} \ \cancel{z} \ \cancel{z} \ \cancel{z} \ z \ z \ z} = \frac{1}{z^3} = z^{-3}$$

The new exponent is the difference of the original exponents: numerator minus denominator!

Quotient rule

For any nonzero real number a and any integers m and n ,

$$\frac{a^m}{a^n} = a^{m-n}$$

The intuition here is that when we have a quotient involving two expressions with the same base, there are going to be common factors in the numerator and denominator. Some of these will “cancel”, and in fact either the whole numerator or the whole denominator (or both!) will be reduced simply to the number 1.

Since multiplication and division are inverse operations, this rule can be seen as a version of the product rule if we allow negative exponents:

$$\frac{a^m}{a^n} = a^m \cdot a^{-n} = a^{m+(-n)} = a^{m-n}$$

This rule also helps us to understand the zero exponent. We know that anything divided by itself is 1, and so for instance:

$$\frac{a^m}{a^m} = 1 \quad \text{since a number divided by itself is 1}$$

If we interpret this using the quotient rule, we have:

$$\frac{a^m}{a^m} = a^{m-m} = a^0 \quad \text{application of the quotient rule}$$

This means we have two different interpretations of the same quantity, and so these two interpretations must be equal. In other words, it must be that:

$$a^0 = 1.$$

11.4.3 Quotient to a power**Derivation #6**

// startup exploration

Use expanded form to write each of the expressions below without parentheses.

$$\left(\frac{x}{y}\right)^2 \quad \text{and} \quad \left(\frac{ab}{c}\right)^7$$

Make a conjecture about a general rule for simplifications of this type.

By now, you may be getting the hang of these simplification properties! In this case, we have

$$\left(\frac{x}{y}\right)^2 = \left(\frac{x}{y}\right)\left(\frac{x}{y}\right) = \frac{x \cdot x}{y \cdot y} = \frac{x^2}{y^2}$$

The second example is longer, but works the same:

$$\begin{aligned}
 \left(\frac{ab}{c}\right)^7 &= \left(\frac{ab}{c}\right) \left(\frac{ab}{c}\right) \left(\frac{ab}{c}\right) \left(\frac{ab}{c}\right) \left(\frac{ab}{c}\right) \left(\frac{ab}{c}\right) \left(\frac{ab}{c}\right) \\
 &= \frac{ab\ ab\ ab\ ab\ ab\ ab\ ab}{c\ c\ c\ c\ c\ c\ c} \\
 &= \frac{(ab)^7}{c^7} \\
 &= \frac{a^7 b^7}{c^7}
 \end{aligned}$$

In general, when we have a quotient raised to a power, we can apply the exponent to both the numerator and denominator. To eliminate the parentheses entirely in the second example, we had to use the power of a product property to simplify the numerator.

Quotient to a power rule

For any nonzero real numbers a and b and any integer m ,

$$\left(\frac{a}{b}\right)^m = \frac{a^m}{b^m}$$

The exponent means that we have m copies of the fraction $\frac{a}{b}$ which, when we multiply, become a single fraction with m copies of a in the numerator and m copies of b in the denominator.

11.4.4 Combo problems

The second example in derivation #6 required us first to simplify the power of a quotient, and then the power of a product. Smashup problems like this require some careful dissection. Consider the following nasty-looking expression

$$\frac{(5x^3)(3x^{-2})}{30x^{12}}.$$

With a monster like this, it is sometimes hard to know how to begin. A good strategy is to proceed one step at a time, carefully applying the properties from this section to try and make this slightly simpler with each step.

For instance, we might start by multiplying in the numerator. We'll have a coefficient ($5 \cdot 3 = 15$), and we'll use the product property to multiply the variables ($x^3 \cdot x^{-2} = x^{3+(-2)} = x^1$). This gives us:

$$\frac{(5x^3)(3x^{-2})}{30x^{12}} = \frac{15x}{30x^{12}}.$$

Now, we'll need the quotient property to simplify the variables. The coefficients also share a common factor of 15. This will eliminate the numerator entirely!

$$\frac{(5x^3)(3x^{-2})}{30x^{12}} = \frac{15x}{30x^{12}} = \frac{1}{2x^{11}}.$$

That final answer might not look all that "simple", but it meets the requirements of a simplified exponential expression, and so we're done. Try the following examples, and then study the solutions carefully.

Example 11.4

Simplify each of the following.

$$\left(\frac{x^3y^{-2}}{x^{-1}y^5}\right)^3 \quad \text{and} \quad \frac{8x^2y}{y^2} \div \frac{16xy^2}{(4y^2)^2}$$

Solution: We recommend simplifying what's inside those parentheses before dealing with the outermost exponent. To simplify the insides, we'll need to move some negative exponents, and then apply the product property.

$$\left(\frac{x^3y^{-2}}{x^{-1}y^5}\right)^3 = \left(\frac{x^3x^1}{y^5y^2}\right)^3 = \left(\frac{x^{3+1}}{y^{5+2}}\right)^3 = \left(\frac{x^4}{y^7}\right)^3 = \frac{x^{4 \cdot 3}}{y^{7 \cdot 3}} = \frac{x^{12}}{y^{21}}$$

The last thing we did was apply the power of a quotient to get our final answer. To tackle the second problem, recall that division is like multiplication by the reciprocal.

$$\frac{8x^2y}{y^2} \div \frac{16xy^2}{(4y^2)^2} = \frac{8x^2y}{y^2} \cdot \frac{(4y^2)^2}{16xy^2}$$

Then, before we can apply any of the product properties, we need to fix the product of a power, which will in turn require the power property: $(4y^2)^2 = 4^2 \cdot (y^2)^2 = 16y^{2 \cdot 2} = 16y^4$. From here, we can simplify the individual fractions, then apply the product property to multiply them, and then apply the quotient property (again) to simplify the result.

$$\begin{aligned} \frac{8x^2y}{y^2} \div \frac{16xy^2}{(4y^2)^2} &= \frac{8x^2y}{y^2} \cdot \frac{(4y^2)^2}{16xy^2} \\ &= \frac{8x^2y}{y^2} \cdot \frac{16y^4}{16xy^2} \\ &= \frac{8x^2}{y} \cdot \frac{y^2}{x} \\ &= \frac{8x^2y^2}{yx} \\ &= 8xy \end{aligned}$$

Some of the simplification steps we took above could have been completed in a different order. There are multiple paths to the most simplified form!

11.5 Raising sums to powers

Check this work!

// startup exploration

The following is a very common error:

$$(a + b)^2 = a^2 + b^2 \quad (\text{nice try, but wrong}).$$

This equation looks reasonable, but it is not true most of the time. Find three different pairs of values a and b which demonstrate that this equation does not work in general. Can you find any “lucky” values of a and b that do satisfy this equation?

As we see in the startup exploration, $(x + 3)^2 \neq x^2 + 3^2$. The temptation to “sprinkle” an exponent over a sum — using a kind of exponent distributive property — is **Evil and Wrong**. We *will* use the distributive property to work out what it means to raise a sum to a power, but there’s more to it than a simple sprinkle.

We learned the distributive property in section 5.1, and it has come in quite handy since then, for example when converting the point-slope form of a line to slope-intercept form. Recall that the distributive property tells us that for real numbers a , b , and c :

$$a \cdot (b + c) = a \cdot b + a \cdot c$$

But suppose “ a ” isn’t just a number, but an expression. For example, given the expression

$$(x + 3)^2 = (x + 3)(x + 3)$$

we can use the distributive property (twice, in fact) to simplify the right-hand side. Observe:

$$\begin{aligned} (x + 3)^2 &= (x + 3)(x + 3) \\ &= (x + 3) \cdot x + (x + 3) \cdot 3 && \text{distribute the expression } (x+3) \\ &= x \cdot x + 3 \cdot x + x \cdot 3 + 3 \cdot 3 && \text{distribute in each of the resulting expressions} \\ &= x^2 + 3x + 3x + 9 && \text{simplify} \\ &= x^2 + 6x + 9 && \text{combine like terms} \end{aligned}$$

We do a kind of “double distribution”: first with one set of parentheses over the other, and then distributing over the results.

Example 11.5

Simplify $(3x - 5)^2$.

Solution: The exponent means that we're multiplying this expression by itself, so we will carry out "double distribution":

$$\begin{aligned}(3x - 5)^2 &= (3x - 5)(3x - 5) \\ &= 3x(3x - 5) - 5(3x - 5) \quad \text{distribute in each of the resulting expressions}\end{aligned}$$

So, in the end, $(3x - 5)^2 = 9x^2 - 30x + 25$.

11.5.1 Old school multiplication

There can be a lot of stuff to keep track of in these multiplications. A technique that can be helpful is to set up the multiplication like an "old school" (literally) multiplication problem from back in your elementary school days.

Recall how to multiply 43×43 via long "multiplication". We interpret "43" in a more expanded form as "40+3", then we break this multiplication into two steps. We first multiply 3×43 to get the first *partial product* of 129. Then, we multiply 40×43 to get the second partial product 1720.

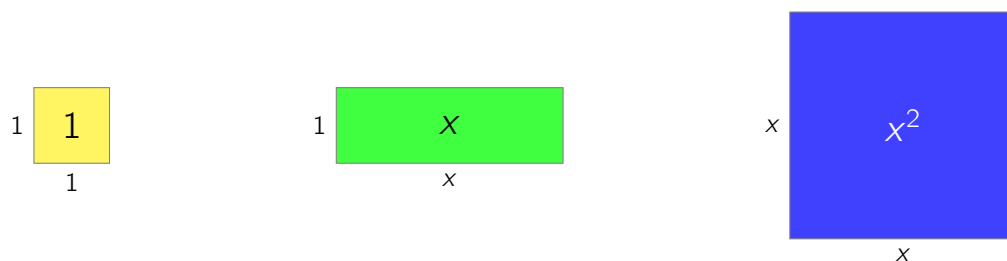
$$\begin{array}{r} 43 \\ \times 43 \\ \hline 129 \quad (3 \times 43) \\ + 1720 \quad (40 \times 43) \\ \hline 1849 \end{array}$$

We can use this helpful organization scheme to multiply $(4x + 3)(4x + 3)$.

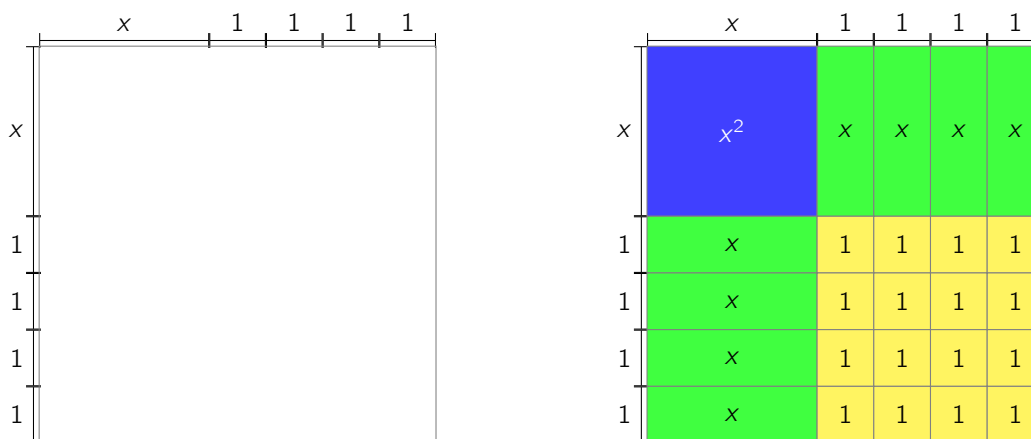
$$\begin{array}{r} 4x + 3 \\ \times 4x + 3 \\ \hline 12x + 9 \quad 3(4x + 3) \\ + 16x^2 + 12x \quad 4x(4x + 3) \\ \hline + 16x^2 + 24x + 9 \end{array}$$

11.5.2 Algebra tiles

Another helpful visual may be to draw “algebra tiles”.² In a set of algebra tiles, we pick a short length to represent 1 and a long length to represent x . Then, we create a set rectangular tiles:



To multiply, say, $(x + 4)^2$, we create a rectangle (a square, actually) that is “ $x+4$ ” units on each side, and we fill in the area with our tiles.



The area is filled in with one x^2 -tile, eight x -tiles, and sixteen 1-tiles. So, we have a total area of $x^2 + 8x + 16$.

11.5.3 Beyond squares

Of course, the methods above all work for multiplying other expressions as well — not just something times itself. Can you extend the models above to find the product of, say,

$$(3x + 4)(2x + 3)$$

When we multiply, we get the product $6x^2 + 17x + 12$. How can we get to this product using the “double distribution” approach? What about the old-school method? What would the algebra tiles area model look like?

² Sometimes drawing a picture is enough. It can also be helpful, though, to actually manipulate physical or virtual tiles. For example you could cut a set of algebra tiles from cardstock, or do an internet search for an online algebra tiles simulator.

We can also raise an expression to a power other than 2. For instance, we could raise an expression to the third power, as in the example below. There's a lot going on here, read through the work slowly and see if you can understand each step.

$$\begin{aligned}
 (2x + 5)^3 &= (2x + 5)(2x + 5)(2x + 5) \\
 &= (4x^2 + 20x + 25)(2x + 5) && \text{multiply the first two terms} \\
 &= 2x(4x^2 + 20x + 25) + 5(4x^2 + 20x + 25) && \text{distribute the three-term expression!} \\
 &= \underline{8x^3 + 40x^2 + 50x} + \underline{20x^2 + 100x + 125} && \text{distribute twice more} \\
 &= 8x^3 + 60x^2 + 150x + 125 && \text{combine like terms}
 \end{aligned}$$

» Chapter summary «

Here ends our work with exponential functions. It's been a quick tour of the highlights of exponents and exponential relationships, but our algebra toolbox has expanded quite a bit. As we have mentioned, exponential expressions will make a glorious return in Algebra 2.

We turn now to the family of quadratic functions. This family offers us a great deal to think about – both abstract and contextual. Our exploration of quadratics will yield an impressive collection of techniques for manipulating abstract algebraic expressions. Plus, we will use quadratic functions to problems from physics about the motion of falling and flying objects. Onward!

We must say that there are as many squares as there are numbers.

Galileo Galilei, Italian physicist and astronomer

Chapter 12

Quadratic equations

We saw quadratic patterns for the first time back in chapter 2, when we learned about quadratic sequences. We worked a bit with equations for quadratic data, but we haven't seen quadratic equations since then.¹

In chapter 5 we learned a set of tools in for solving linear equations. Many of those tools will be helpful here, but in this chapter, we'll also see what makes quadratic equations different from those we've studied until now. Our main task will be to develop techniques that will help us to overcome the trickier aspects of quadratic equations.

12.1 Challenges to solving quadratics

A problem with quadratics

// startup exploration

Consider the following equation. Do we know what we need to know to solve this equation, without any guesswork? In other words: Do we have POEs, axioms, or other properties that will allow us to isolate x ?

$$x^2 + 2x - 15 = 20$$

If you can, solve the equation. Otherwise, identify where you get stuck. What new tools would be helpful for solving this equation?

¹ By the way, now might be a good time to have a second look at the final sections of chapter 2. Those are (;;;) sections, and so they may have been a bit challenging the first time around. You have a lot more algebra skills at this point in the course, though, and those sections might not be as daunting now!

Our task in this chapter is to be able to solve quadratic equations like the one given in the startup exploration. We won't give the details for how to solve this equation yet – we'll develop those ideas over the next few sections. For now, we will simply point out a few features.

Using what we know so far, it's impossible to isolate x . We can make some progress:

$$x^2 + 2x - 15 = 20$$

$$x^2 + 2x = 35 \quad \text{APOE, to get all the numbers to the right-hand side}$$

But now what? We can't combine like terms on the left-hand side, and subtracting anything from that side would give us x 's on both sides of the equation (making things worse). We might try undoing the distributive property on the left-hand side. That would give us

$$x(x + 2) = 35.$$

But this doesn't seem to be much of an improvement either. Using DPOE on the left-hand side to isolate x would move $(x + 2)$ to the right hand side of the equation (and vice versa). It seems that we'll need some other techniques to help us out of this situation.

If we're allowed to just solve the equation by making a clever observation, we might notice that

$$5(7) = 5(5 + 2) = 35,$$

and so $x = 5$ is a solution to the equation! That's progress. But, it might not be obvious that -7 is *also* a solution to the equation, since

$$-7(-7 + 2) = -7(-5) = 35.$$

Plus with a different equation, it might not be quite so easy to see a solution just by inspection. What to do with $x(x + 2) = 29$, for example? Our new techniques should help us overcome these challenges.

12.1.1 Rectangles and squares

As we saw in chapter 2, quadratic sequences can be related to rectangles and squares. Our very first quadratic sequence was related to rectangles (fig. 12.1), and so are the familiar perfect squares (fig. 12.2).

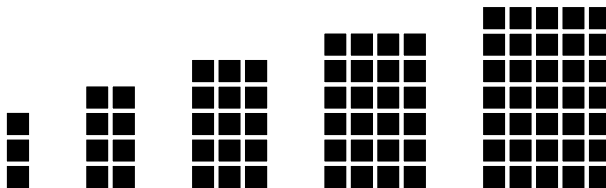


Figure 12.1: Some “rectangular numbers”: 3, 8, 15, 24, 35, ...

The perfect squares have a straightforward formula. If we let $f(x)$ represent the area of figure x , then the

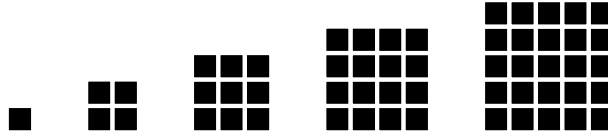


Figure 12.2: The perfect squares: 1, 4, 9, 16, 25, ...

perfect squares are represented by the formula

$$f(x) = x^2.$$

This is the parent function of the quadratic family.

The rectangles in fig. 12.1 also have a formula. If we let $g(x)$ represent the area of figure x , then we might notice that figure x is x units wide and $(x + 2)$ units tall. So these rectangles are represented by the formula

$$g(x) = x(x + 2).$$

We can simplify this formula using the distributive property:

$$g(x) = x(x + 2) = x^2 + 2x$$

We have learned that the highest degree term in a quadratic equation is an x^2 term. The connection between a quadratic rule and rectangles will help us when it comes to solving quadratic equations. In fact, we will solve these equations using the beautiful symmetry of the square.

12.2 Level 1 and 2 quadratics

As we did with linear equations, we'll treat quadratic equations like a game. Level 1 is the easiest kind of quadratic equation to solve, so this is where we'll start. Then, we'll keep things interesting as we increase our skill level by adding challenge and complexity along the way.

12.2.1 Level 1 quadratics

Quadratic level 1

// startup exploration

Determine the value of x given the equation: $x^2 = 64$.

Talk about starting with the easy stuff. We're looking for a number x which, when multiplied by itself, is equal to 64. In other words, we are looking for the "square root of 64". Clearly, $x = 8$ is a solution to this equation. But we can't be too hasty! Notice that $x = -8$ is also a solution, since $(-8)^2 = (-8)(-8) = 64$. So, this equation has two solutions:

$$x = 8 \text{ or } -8$$

If we prefer to write our answer in solution set notation, we have

$$\mathcal{S} = \{8, -8\}.$$

So, Level 1 quadratics are pretty easy: we simply take the square root of both sides of the equation.² We might have a hard time if the constant value is not a perfect square, as in $x^2 = 12$. But, we'll learn more about handling the square roots of non-perfect-squares soon enough (in chapter 13, to be precise).

This seems like a good time to suggest that it may come in handy to memorize the first 25 or so perfect squares, for quick recognition when they come up in a problem.

$1^2 = 1$	$2^2 = 4$	$3^2 = 9$	$4^2 = 16$	$5^2 = 25$
$6^2 = 36$	$7^2 = 49$	$8^2 = 64$	$9^2 = 81$	$10^2 = 100$
$11^2 = 121$	$12^2 = 144$	$13^2 = 169$	$14^2 = 196$	$15^2 = 225$
$16^2 = 256$	$17^2 = 289$	$18^2 = 324$	$19^2 = 361$	$20^2 = 400$
$21^2 = 441$	$22^2 = 484$	$23^2 = 529$	$24^2 = 576$	$25^2 = 625$

² More soon on square roots, including why and under what circumstances "square root of both sides" is, in fact, a property of equality.

Finally, note that zero is also a perfect square, since $0^2 = 0$. Zero, in fact, is the only number that has only one square root. Whereas both 3 and -3 are square roots of 9, the only square root of 0 is 0.

Finally, consider the square root of a negative number, say, -25 . This is a bit of a problem. We're meant to find the number that when multiplied by itself gives -25 as the result, but that's not going to work. Since we have a negative product, the two factors must have opposite signs! The best we can do is $5 \cdot -5 = -25$ or $-5 \cdot 5 = -25$, and in these cases we're not multiplying a number times itself: 5 and -5 are different numbers!

So, there is no real number equal to the square root of -25 . Of course, -25 isn't special, the same argument applies to any negative number. The moral of the story is this: If, as we go about solving a quadratic equation, we come to the point where we need to take the square root of a negative number, we have to stop and say that the equation has "no real number" as its solution. $\mathcal{S} = \emptyset$.

12.2.2 Level 2 quadratics

Quadratic level 2

// startup exploration

Determine the value of w given the equation: $(w + 3)^2 = 16$.

Here, we're told that something-squared is 16. Well, that means that the something in question must either be 4 or negative 4. That is to say, $(w + 3)$ is either 4 or -4 . So, this equation is actually two equations at once. We have:

$$w + 3 = 4 \quad \text{or} \quad w + 3 = -4$$

Solving these equations (using SPOE in both cases), we have that w must be either 1 or -7 . So, those are our two solutions: $\mathcal{S} = \{1, -7\}$.

We can record this work in a down-the-page format, like so:

$$(w + 3)^2 = 16$$

$$w + 3 = 4 \quad \text{or} \quad -4$$

square root of both sides

$$w = 1 \quad \text{or} \quad -7$$

SPOE: subtract 3 throughout

A few things to note: First, we can't subtract 3 from both sides as the very first step. The parentheses require that we undo the exponent first. Second, beware the use of \pm . It may be tempting to use this shorthand notation and write

$$w + 3 = \pm 4,$$

but then it is also tempting to subtract 3 and write

$$w = \pm 1. \quad \text{Nope!}$$

We recommend splitting into two equations, or writing out the two solutions explicitly using the word “or”.

Example 12.1

Determine the value of a given the equation: $(a - 1)^2 - 2 = 23$.

Solution: This equation requires an extra step, but it's quickly transformed into an equation like the one from the startup exploration.

$$(a - 1)^2 - 2 = 23$$

$$(a - 1)^2 = 25$$

$$a - 1 = 5 \text{ or } -5$$

$$w = 6 \text{ or } -4$$

APOE

square root of both sides

APOE: add 1 throughout

In the end, we have $S = \{6, -4\}$.

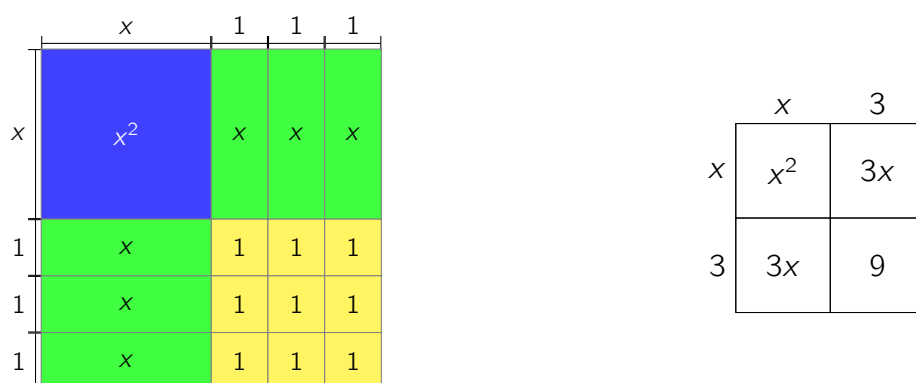
12.3 Level 3 quadratics: Quadrangle method

Quadratic level 3

// startup exploration

Use the sum to a power property (from section 11.5) to write the expression $(x+3)^2$ without parentheses. Then, determine the value of x given the equation: $x^2 + 6x + 9 = 49$.

To expand $(x+3)^2$, we might think of using algebra tiles to fill in a square with side length $(x+3)$. We used algebra tiles to solve this kind of problem back in section 11.5. Or, we could just “sketch” the algebra tiles diagram (shown below, on the right) and calculate the areas of the four regions.



Note how we have simplified the picture in the “sketch” version. For example, rather than draw three 3-unit-by- x -unit rectangles, we simply write the area of the rectangle $3x$. In the lower right-hand region, we write the area 9 rather than draw nine yellow squares.

We’ll call this simplified sketch a *quadrangle diagram*, and we’ll call the method we are developing here the *quadrangle method*.³ The quadrangle diagram can help us to see the relationship between the two expressions in the startup exploration:

$$(x+3)^2 = x^2 + 6x + 9.$$

Using this fact, we can solve the given equation:

$$x^2 + 6x + 9 = 49$$

$$(x+3)^2 = 49$$

$$x+3 = 7 \text{ or } -7$$

$$x = 4 \text{ or } -10$$

based on the quadrangle diagram

square root of both sides

SPOE: subtract 3 throughout

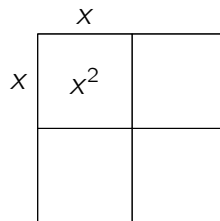
³ The term *quadrangle* is a synonym for *rectangle*. History suggests that the term *quadratic* is related to the term *quadrangle* since this method of using square diagrams is so helpful in solving quadratic equations.

Can you see the clever trick that we used here? We rewrote an expanded expression as a something-squared expression, and then we solved it like a Level 2 quadratic!

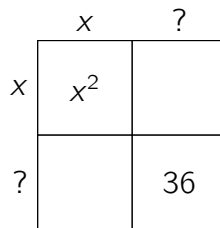
Let's work through another example in detail, for example, solving the Level 3 equation

$$x^2 + 12x + 36 = 144.$$

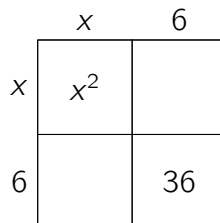
Our goal is to write the left-hand side in a something-squared form. To do that, we'll begin by drawing an empty quadrangle diagram and filling in the bits that we know. For example, we know that the upper left-hand box must contain x^2 , and so the sides of that square must each be x units long.



We know that the 36 will appear in the lower right-hand box. But how could we label the side lengths here? There are lots of combinations of numbers that multiply together to get 36...



Remember that the goal is to get an expression of the form something-*squared*, and so we should always strive to create a quadrangle that is, in fact, a square. This means that we must choose 6 and 6 as the side lengths. If we chose another pair of factors, like 2 and 18, we would get the proper product in the lower right-hand region, but our overall diagram would no longer be a square.



To fill in the remaining regions, we multiply the dimensions of each. In both cases we get $6x$. This is good news! Together these make $12x$ (which is what we have in the original, expanded expression). Plus, the two regions contain the same value, and so the beautiful symmetry of the square is preserved.

	x	6
x	x^2	$6x$
6	$6x$	36

So, we have rewritten our expression as a something-squared expression:

$$x^2 + 12x + 36 = (x + 6)^2.$$

We can use this to solve the equation that we were given:

$$x^2 + 12x + 36 = 144$$

$$(x + 6)^2 = 144$$

$$x + 6 = 12 \text{ or } -12$$

$$x = 6 \text{ or } -18$$

based on the quadrangle diagram

square root of both sides

SPOE

Example 12.2

Determine the value of n given the equation: $n^2 - 16n + 64 = 1$.

Solution: Let's build the quadrangle diagram. The upper left-hand corner contains n^2 , as above. To maintain symmetry, we should split the $-16n$ exactly in half. Don't worry about the negative coefficient, we can work with that. We should be on guard though for sign-related issues.

	n
n	n^2
	$-8n$
	$-8n$

This tells us that the remaining portion of the square's side length must be -8 . Let's not worry too much about the fact that negative distances are impossible. . . the idea still works. This implies that the remaining region contains 64 (note, that positive 64). This agrees with the equation we were given.

	n	-8
n	n^2	$-8n$
-8	$-8n$	64

Now, we can solve the equation:

$$n^2 - 16n + 64 = 1$$

$$(n - 8)^2 = 1$$

$$n - 8 = 1 \text{ or } -1$$

$$n = 9 \text{ or } 7$$

based on the quadrangle diagram

square root of both sides

SPOE

So, we have $\mathcal{S} = \{7, 9\}$. We can check our work by substituting our proposed solutions back into the original equation. First, we'll test $n = 7$:

$$n^2 - 16n + 64 = 1$$

$$(7)^2 - 16(7) + 64 \stackrel{?}{=} 1$$

$$49 - 112 + 64 \stackrel{?}{=} 1$$

$$1 \stackrel{\checkmark}{=} 1$$

First, we'll check $n = 9$:

$$n^2 - 16n + 64 = 1$$

$$(9)^2 - 16(9) + 64 \stackrel{?}{=} 1$$

$$81 - 144 + 64 \stackrel{?}{=} 1$$

$$1 \stackrel{\checkmark}{=} 1$$

12.4 Level 4 quadratics: Adjust the constant term

Quadratic level 4

// startup exploration

Determine the value of x given the equation: $x^2 + 8x + 15 = 99$.

Let's try to build the quadrangle diagram. The upper left-hand corner contains x^2 , as before. To maintain symmetry, we should split the $8x$ exactly in half. This tells us that the large square must be $(x + 4)$ units on a side.

	x	4
x	x^2	$4x$
4	$4x$	

The problem is that, according to our quadrangle diagram, the lower right-hand corner should be 16 . . . but the equation we are given tells us to put 15 in that space. Now what? Our equation does not represent a complete square!

	x	4
x	x^2	$4x$
4	$4x$	15

POEs to the rescue! Why not add 1 to both sides of the given equation to complete the square? This will give us the number we want on the left-hand side and, since we add 1 to both sides, we have an equivalent equation. So, instead of solving the equation

$$x^2 + 8x + 15 = 99,$$

we will add one to both sides and solve the equation

$$x^2 + 8x + 16 = 100.$$

Now, we can draw the quadrangle diagram exactly as we wanted.

	x	4
x	x^2	$4x$
4	$4x$	16

And now that we have a proper quadrangle diagram, we can solve the equation! Here's the full process:

$$x^2 + 8x + 15 = 99$$

$$x^2 + 8x + 16 = 100$$

$$(x + 4)^2 = 100$$

$$x + 4 = 10 \text{ or } -10$$

$$x = 6 \text{ or } -14$$

APOE: add 1 to both sides

based on the quadrangle diagram

square root of both sides

SPOE

This is a clever application of the POEs. Rather than use the properties to eliminate terms from one side of the equation, we can use the properties to change one side into a particular, more-helpful form.⁴

Example 12.3

Determine the value of x given the equation: $x^2 - 6x + 11 = 27$.

Solution: When we start the quadrangle diagram, we split the $-6x$ as usual. This means that the square has sides of length $(x - 3)$. This, in turn, implies that the lower right-hand region should be 9. Our equation has 11 as its constant term: not what we want.

	x	-3
x	x^2	$-3x$
-3	$-3x$	9

To fix, this we can subtract 2 to each side of our equation. This will give us a constant term of 9, which

⁴ In fact, this is exactly what we've been doing all along: changing one side of an equation into a form that is more helpful. In this chapter we are expanding our notion of what it means for a change to be "helpful".

is what we need to make our quadrangle diagram work. So, we have:

$$x^2 - 6x + 11 = 27$$

$$x^2 - 6x + 9 = 25$$

$$(x - 3)^2 = 25$$

$$x - 3 = 5 \text{ or } -5$$

$$x = 8 \text{ or } -2$$

SPOE: subtract 2 from both sides

based on the quadrangle diagram

square root of both sides

SPOE

So, our solutions are $\mathcal{S} = \{8, -2\}$.

12.5 Level 5 quadratics: Adjust the linear term

The last few levels have had “polite” middle terms, which have split evenly into two pieces. What happens if we get a linear term with an odd coefficient?

Quadratic level 5

// startup exploration

Determine the value of x given the equation: $x^2 + 3x + 1 = 5$.

If we jump right in and try the quadrangle method, we start to get into fraction territory. If we want to split the $3x$ term exactly in half, then each piece would be $\frac{3}{2}$. Then, the quadrangle method would predict $\frac{9}{4}$ in the lower-right corner.

	x	$\frac{3}{2}$
x	x^2	$\frac{3}{2}x$
$\frac{3}{2}$	$\frac{3}{2}x$	$\frac{9}{4}$

The lower-right corner isn't what we have in our equation, so we could use APOE and add $\frac{5}{4}$ to adjust both sides. . . Hmm. Not pretty. To be clear, the quadrangle method will not let us down: if we keep going with the fractions, we will arrive at the correct answer! But, perhaps there is an alternative approach that avoids the fractions.

Here's a clever idea: We could use MPOE and multiply through by 2. This would give us a linear term with an even coefficient! In other words:

$$x^2 + 3x + 1 = 5 \xrightarrow{\text{multiply through by 2}} 2x^2 + 6x + 2 = 10$$

This fixes our odd coefficient problem, since now we can break the $6x$ up into two sets of $3x$. But, what do we do with that $2x^2$? We can't use x and $2x$ as the side lengths, for although that gives us the correct product, we would no longer have a square.

	$?$	3
$?$	$2x^2$	$3x$
3	$3x$	

Now, here's a *really* clever idea. Let's multiply through by 2 *again*. In other words, we will multiply the original equation by 4:

$$x^2 + 3x + 1 = 5 \quad \xrightarrow{\text{multiply through by 4}} \quad 4x^2 + 12x + 4 = 20$$

We still have an even linear coefficient, and now we can write $4x^2$ as $2x$ times $2x$. Note that we have to take that factor of 2 into account when we're figuring out the other dimensions of the square. Study our new diagram closely and be sure you understand where each of the labels comes from.

	2x	3
2x	4x ²	6x
3	6x	9

We now find ourselves in a Level 4 situation: our equation has 4 as the constant term, whereas the quadrangle diagram predicts 9 as the constant term. No problem! We can add 5 to both sides, and continue the process as we did with Level 4 quadratics. Here's a summary of the whole process:

$x^2 + 3x + 1 = 5$	original equation
$4x^2 + 12x + 4 = 20$	MPOE: multiply both sides by 4
$4x^2 + 12x + 9 = 25$	APOE: add 5 to both sides
$(2x + 3)^2 = 25$	based on the quadrangle diagram
$2x + 3 = 5$ <u>or</u> -5	square root of both sides
$2x = 2$ <u>or</u> -8	SPOE
$x = 1$ <u>or</u> -4	DPOE

Let's pause and review. When faced with an odd coefficient for the x term, our strategy is to multiply through by 4. This will give us an even coefficient for the x term, and at the same time keep the coefficient of the x^2 term in a state where it can be written as something times itself: $4x^2 = 2x \cdot 2x$. After we do this, we'll have a Level 4 quadratic on our hands, and we can apply techniques for handling those.

Example 12.4

Determine the value of x given the equation: $x^2 - 5x + 12 = 62$.

Solution: Faced with an odd linear coefficient, we multiply through by 4. This gives us the revised equation

$$4x^2 - 20x + 48 = 248.$$

We set up the quadrangle diagram and see whether our equation represents a complete square.

	$2x$	-5
$2x$	$4x^2$	$-10x$
-5	$-10x$	25

Our (revised) equation does not make a proper square: the constant term in the equation is 48, but the quadrangle diagram predicts 25. We can fix this problem using SPOE: subtract 23 from both sides:

$x^2 - 5x + 12 = 62$	original equation
$4x^2 - 20x + 48 = 248$	MPOE: multiply both sides by 4
$4x^2 - 20x + 25 = 225$	SPOE: subtract 23 from to both sides
$(2x - 5)^2 = 225$	based on the quadrangle diagram
$2x - 5 = 15 \text{ or } -15$	square root of both sides
$2x = 20 \text{ or } -10$	SPOE
$x = 10 \text{ or } -5$	DPOE

So in the end, we have solutions $\mathcal{S} = \{10, -5\}$.

12.6 Level 6 quadratics: Adjust the quadratic term

We've arrived at the highest level of the quadratic equation challenge! Until now, all of our equations have started with just x^2 . What if the coefficient of the leading term is something other than 1?

Quadratic level 6

// startup exploration

Determine the value of x given the equation: $3x^2 + 8x + 1 = 12$.

What shall we do in this scenario? A clever idea is to use DPOE and divide through by 3. That would make the leading coefficient 1, as in the earlier problems. The downside is that most of the other numbers turn into fractions.

$$3x^2 + 8x + 1 = 12 \quad \xrightarrow{\text{divide through by 3}} \quad x^2 + \frac{8}{3}x + \frac{1}{3} = 4$$

The quadrangle method will absolutely work on an equation like this, but perhaps we'd prefer an approach that avoided all the fractions.

If scaling the equation down doesn't help, why not try to scale it up? Could we multiply through by some helpful value? Recall that the goal will be to write the x^2 term as something times itself. Since there's already a 3 there, we can solve our problem if we multiply through by 3.

$$3x^2 + 8x + 1 = 12 \quad \xrightarrow{\text{multiply through by 3}} \quad 9x^2 + 24x + 3 = 36$$

Note that now we are in a good position, since $9x^2 = 3x \cdot 3x$. So, let's fill out our quadrangle diagram. We have an even coefficient on the linear term, so we can split that evenly. Note that we have to take all the coefficients into account when completing the diagram. For example, when figuring out the dimensions of a box containing $12x$.

	3x	4
3x	9x ²	12x
4	12x	16

The quadrangle diagram predicts 16 as the constant term, and so our equation is not a complete square. We

can use APOE to fix that, adding 13 to both sides. Here's how it goes:

$3x^2 + 8x + 1 = 12$	original equation
$9x^2 + 24x + 3 = 36$	MPOE: multiply through by 3
$9x^2 + 24x + 16 = 49$	APOE: add 13 to both sides
$(3x + 4)^2 = 49$	based on the quadrangle diagram
$3x + 4 = 7 \text{ or } -7$	square root of both sides
$3x = 3 \text{ or } -11$	SPOE
$x = 1 \text{ or } -\frac{11}{3}$	DPOE

In summary, we multiplied through by the coefficient of the x^2 term, which gave us a coefficient that was a perfect square. In the final example for this chapter, we put it all together.

Example 12.5

Determine the value of x given the equation $-5x^2 - x + 18 = 0$.

Solution: Since we have a leading coefficient that is not a perfect square, we multiply through by that coefficient, -5 in this case. This gives us

$$25x^2 + 5x - 90 = 0$$

(careful with the negative signs). This is an improvement, but we have a linear coefficient that is odd. So, we multiply by 4 to fix that:

$$100x^2 + 20x - 360 = 0$$

Notice that multiplying through by 4 (a perfect square) gives us a leading coefficient that is still a perfect square. This is because the product of two perfect squares is itself a perfect square! (Can you prove that this statement is always true using the properties of exponents?)

We now have a revised equation that we can bring to the quadrangle method.

	10x	1
10x	100x ²	10x
1	10x	1

To get our constant terms to agree, we must add 361 to both sides of our revised equation. This gives us

$$10x^2 + 10x + 1 = 361.$$

And from here, we can complete a familiar process.

$$10x^2 + 10x + 1 = 316$$

$$(10x + 1)^2 = 316$$

based on the quadrangle diagram

$$10x + 1 = 19 \text{ or } -19$$

square root of both sides

$$10x = 18 \text{ or } -20$$

SPOE

$$x = \frac{18}{10} \text{ or } -2$$

DPOE

After we simplify our one fraction answer, we have a final solution: $\mathcal{S} = \{\frac{9}{5}, -2\}$.

12.6.1 (,;) Quadratic formula

Now that we have the quadrangle method at our disposal, we can tackle any quadratic equation that is thrown at us. We might be tempted to really get generic, and solve the all at once.

Consider a completely generic quadratic equation of the form

$$ax^2 + bx + c = 0.$$

This quadratic has three coefficients: a is the coefficient of the quadratic term, b is the coefficient of the linear term, and c is the constant term. We've set it equal to zero for reasons that will be made clear in a few chapters.⁵

What happens if we apply the quadrangle method to this totally generic equation? We don't really know anything about a , so to be safe, let's multiply the whole equation by a .

$$a^2x^2 + abx + ac = 0.$$

This will ensure that the leading term is a perfect square, and that's what we need for the box method. Now, the linear coefficient is ab and this might be an odd number for all we know. So, we multiply through by 4, as we have often done above.

$$4a^2x^2 + 4abx + 4c = 0.$$

This is not a pretty sight, but we'll try to make it work with the quadrangle method.

⁵ Sorry for the mystery, but all will be revealed soon. In the meantime, note that we can always turn a quadratic equation into this form. If we have an equation that does not equal zero, for example $ax^2 + bx + c = d$ where d is not zero, then we can use SPOE to get 0 on the right-hand side: $ax^2 + bx + c - d = 0$. Then, we pull a little algebraic sleight of hand, and replace $c - d$ with just c . This might look like an illegal move, but in the general form of a quadratic, " c " just stands for a constant value. Since " $c-d$ " is a constant value, we pretend that this was the constant value we meant all along.

	$2ax$	b
$2ax$	$4a^2x^2$	$2abx$
b	$2abx$	b^2

Much of this looks workable, but what about the lower right-hand corner? The quadrangle method predicts b^2 , but our equation has $4c$. Well, we'll do what we usually do: we'll use the POEs to turn our left-hand side into the form we want. So:

$$4a^2x^2 + 4abx + 4c = 0$$

$$4a^2x^2 + 4abx = -4c$$

$$4a^2x^2 + 4abx + b^2 = b^2 - 4c$$

SPOE: subtract $4c$ from both sides

APOE: add b^2 to both sides

It might seem like things are getting worse, but now the left-hand side is what we need to use our quadrangle diagram. Let's go!

$$(2ax + b)^2 = b^2 - 4c$$

$$2ax + b = \pm \sqrt{b^2 - 4c}$$

$$2ax = \pm \sqrt{b^2 - 4c} - b$$

$$x = \frac{\pm \sqrt{b^2 - 4c} - b}{2a}$$

$$x = \frac{-b \pm \sqrt{b^2 - 4c}}{2a}$$

based on the quadrangle diagram

square root of both sides

SPOE: subtract b from both sides

DPOE: divide both sides by $2a$

rearranging the numerator

That last rearrangement in the numerator is because it looks a little better, and to make it clear that the $-b$ isn't underneath the radical along with that other stuff. Also, it's tradition. This formula is a famous thing, and traditionally appears in this form.

» Chapter summary «

The properties of equality which we learned before this chapter weren't enough to solve any random quadratic equation. To overcome the quadratic obstacle, we learned new techniques for solving quadratic equations. These new techniques are based on the symmetry of a square, and we learned various techniques for adjusting a quadratic equation so that we can use our new quadrangle method.

We have, however, avoided the necessity of finding the square root of a number that was not a perfect square. This was a helpful simplification – it allowed us to focus on the process of equation solving – but not all quadratics will have answers that come out so “nice”. We turn our attention to square roots in the next chapter.

There is geometry in the humming of the strings, there is music in the spacing of the spheres.

Pythagoras, Ancient Greek philosopher

Chapter 13

Radical expressions

In the last chapter, we found integer solutions to nearly all of the equations that we studied. This was good for understanding the workings of quadratic equations, but not all equations will necessarily be so “polite”. The main focus of our work in this chapter is around understanding more about all of those square roots that don’t come out evenly. We begin by looking more closely at the sets \mathbb{Q} and \mathbb{R} .

13.1 Real numbers

Share the cheese

// startup exploration

Middle Market sells mini-wheels of cheese for snacking. Mini-wheels can be sold one at a time, or in boxes of 10. The cheese arrives at the warehouse in crates of 10 boxes (containing 100 wheels of cheese in total).

The warehouse workers need to divide their stock of mini-wheels up among three trucks, each of which will deliver to a local Middle Market branch. The warehouse has a total of 13 crates, 7 boxes, and 9 mini-wheels in stock.

The manager allows the workers to open crates or boxes, if needed, to divide the supply evenly. Describe a process for dividing up the cheese that requires opening the minimum number of boxes.

In chapter 1, we made some comments which might, at first, appear contradictory. On the one hand, we saw that the set of real numbers, \mathbb{R} , includes every possible decimal number. We also saw that the set of rational numbers, \mathbb{Q} , includes “terminating decimals and repeating decimals”.

On the other hand, we know that \mathbb{Q} is the set of fractions, meaning those numbers that can be expressed in the form

$$\frac{a}{b} \text{ where } a \text{ and } b \text{ are integers, and } b \text{ is not zero.}$$

These statements raise a few questions. What is the relationship between “fraction” and “terminating or repeating decimal”? What can we say about decimal numbers that are neither terminating nor repeating?

13.1.1 Fractions into decimals

Recall that a fraction is simply a division problem in disguise. If we execute the division problem, we can easily turn a fraction into a decimal. It's especially easy if we have a calculator handy. . . otherwise, we're in for some long division.

Long division

Don't worry if your long division is a bit rusty, just take another look at the startup exploration. The warehouse workers must divide 1379 mini-wheels of cheese among the three trucks (that's $1379 \div 3$) but they must do this with a minimum amount of regrouping.

One solution is to put 4 crates on each of the three trucks. This takes care of 12 crates (1200 mini-wheels in all), but leaves one crate. They have no choice but to open this crate and treat it as 10 boxes. Of course they already had 7 boxes in stock, so now they have 17 boxes in all. They can put 5 boxes on each truck (accounting for 15 boxes, or 150 mini-wheels), but they will have 2 boxes left over. They open these two boxes, revealing 20 mini-wheels. They add these to the 9 mini-wheels they had already, giving 29 mini-wheels in all. Each truck gets 9 of these (using up 27), and they have 2 left over.

So, in the end: each truck gets 4 crates, 5 boxes, and 9 mini-wheels – that's 459 mini-wheels in all – and there are 2 left behind. Now have a look at the long division for this problem: can you spot each of the steps that we took above in the work below?

$$\begin{array}{r}
 459 \\
 3 \overline{) 1379} \\
 \underline{12} \\
 17 \\
 \underline{15} \\
 29 \\
 \underline{27} \\
 2
 \end{array}$$

In the cheese example, it makes sense to stop here with a remainder of 2. In general, though, we could continue the process of division and create a number that extends to the right of the decimal point.

Example 13.1

Convert $\frac{3}{4}$ and $\frac{1}{6}$ into their decimal representations.

Solution: Recall that the fraction three-fourths is equivalent to the division problem $3 \div 4$. To do this by long division, we put the dividend (that's 3) inside the "division house" and leave the divisor (that's 4) outside.

$$\begin{array}{r} 0.75 \\ 4 \overline{) 3.00} \\ \underline{28} \\ 20 \\ \underline{20} \\ 0 \end{array}$$

So the decimal representation of $\frac{3}{4} = 0.75$ (you may have known that already). To tackle one-sixth, we note that it is equivalent to $1 \div 6$. The long division starts out like this:

$$\begin{array}{r} 0.166 \\ 6 \overline{) 1.000} \\ \underline{06} \\ 40 \\ \underline{36} \\ 40 \\ \underline{36} \\ 4 \end{array}$$

We might as well stop here, though, because we're stuck in a loop! The 6's in the answer are going to repeat forever. (Can you see why?) So, the decimal representation of $\frac{1}{6} = 0.1\overline{6}$.

We say that 0.75 is a terminating decimal, because the process of long division stops with a remainder of zero. On the other hand, $0.1\overline{6}$ is called a repeating decimal because the long division process gets stuck in a loop. Note that we've used a vinculum over the 6 to indicate which digits repeat.

When it comes to a division problem like $1 \div 6$ on a calculator, the display will likely show 0.166666667, where the 6's repeat for a while and are followed by a 7. Don't be fooled by this 7: the 6's really do go on forever! That 7 is the calculator rounding up. Always be skeptical about the rightmost digit on your calculator screen.

A bold claim

This process of division leads us to make a pretty bold claim: every rational number can be represented *either* as a terminating decimal or a repeating decimal. How can we be sure that every crazy fraction, for instance $\frac{19}{81}$, either terminates or repeats?

Let's think about how long division works: the "subtraction step" in particular. Here's how the process of long division starts out for $\frac{3}{4}$:

$$\begin{array}{r} .7 \\ 4 \overline{) 3.00} \\ \underline{28} \\ 2 \end{array}$$

In the subtraction step, we get 2 as the remainder, and so we know that we have to keep dividing. If we ever get the remainder 0, then we know that we're done with division. This is what happens eventually with $\frac{3}{4}$. On the other hand, if we ever get a remainder that we've gotten before, then we know that we're stuck in a loop. This is what happened with $\frac{1}{6}$.

Now here's a simple yet profound idea: the remainder is always less than the divisor. When dividing by 4, the remainder has to be less than 4. When dividing by 6, the remainder has to be less than 6. (Can you explain why that is? For example, when dividing mini-wheels of cheese among three trucks, could the workers have seven mini-wheels of cheese left over?)

The result is that we have a limited number of choices for the remainder. When dividing by 4, the remainder can only be 0, 1, 2, or 3. When dividing by 6, the remainder can only be : 0, 1, 2, 3, 4, or 5.

Having a limited number of choices means that eventually we have to recycle one of those remainders! We can't go on forever without either using the remainder 0 (in which case the decimal terminates) or reusing one of the nonzero remainders (in which case the decimal repeats).

Even when dividing something ugly like $1903 \div 8167$, the remainders in the subtraction steps will always be less than 8167. We might have to divide for a long time, but we know it can't carry on forever. Eventually we'll either use the remainder 0, or reuse a remainder we've used already. So the fraction $\frac{1903}{8167}$ has a decimal representation that either terminates or repeats.

Our argument applies to any denominator, and so to any rational number. Therefore, it's true that every rational number has a decimal representation that either terminates or repeats! Have we blown your mind yet? If not, stay tuned.

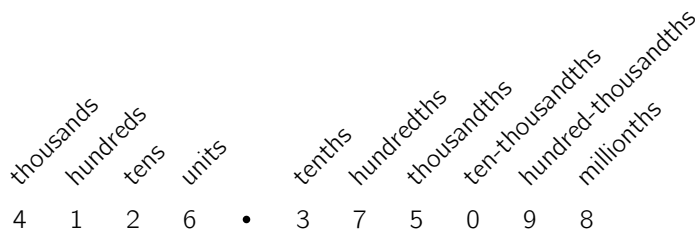
13.1.2 Decimals into fractions

What about the other way around? Does every terminating-or-repeating decimal have a corresponding fraction representation?

Terminating decimals into fractions

Consider a terminating decimal like 0.375. Can we turn this decimal into a fraction, meaning a ratio of two integers? If so, how?

Recall the notion of *place value*, and how the individual digits in a number are each standing in some “place” that is named after a power of ten.¹ The key to turning a terminating decimal into a fraction is recalling how to read a decimal using its place value.



To read the decimal 0.375, we can say “zero point three seven five”, but this isn’t very helpful. Instead, we read the number using place value and say “three hundred seventy-five *thousandths*”. Now, if someone were to say that number aloud, it sounds just like the fraction

$$\frac{375}{1000}$$

In fact, this decimal number and this fraction represent exactly the same value. Of course, the fraction isn’t in simplest form yet, but that’s easy to fix:

$$0.375 = \frac{375}{1000} = \frac{3}{8}$$

We have accomplished the goal of turning a terminating decimal into a fraction. The technique is simply to read the decimal aloud using its place value, and then write down the fraction we hear.

Repeating decimals into fractions

The “read the number with its place value” technique won’t work for repeating decimals. (Why not?) Instead, we’ll use some clever applications of the techniques we learned when solving equations.

Suppose we try to write the repeating decimal $0.\overline{4}$ as a fraction. Let’s give this number a name so that we can do some algebraic manipulations.

$$x = 0.\overline{4}$$

¹ We really are blowing the cobwebs off of some old mathematics in this chapter: Long division! Place value! It goes to show that even simple mathematical ideas can have deep and meaningful consequences.

Our goal will be to find an alternative way of writing x . To do that, we're going to make two clever moves.

The first clever move is to use MPOE: we will multiply both sides of this equation by 10. Multiplying $0.\overline{4}$ by 10 moves the decimal point one place to the right. But remember, the 4's repeat *forever*, so there are *still infinitely many* 4's to the right of the decimal point! We have:

$$10x = 4.\overline{4}$$

The second clever move is to use an idea from when we were solving systems of equations: the elimination method. Watch what happens when we subtract the first equation we wrote from the second equation:

$$\begin{array}{r} 10x = 4.\overline{4} \\ - \quad x = 0.\overline{4} \\ \hline 9x = 4.0 \end{array}$$

Notice that the two numbers on the righthand side of our equations are exactly four units apart. In other words: the infinitely long tail of 4's disappears when we subtract! Now all we have to do is use DPOE to isolate x :

$$9x = 4 \quad \Rightarrow \quad x = \frac{4}{9}$$

If you have a calculator handy, you can perform this division and see that we have accomplished the goal of turning our repeating decimal into a fraction:

$$0.\overline{4} = \frac{4}{9}$$

This process is sometimes called *killing the tail*, since our goal is to subtract two different decimal forms that have the same repeating part, thereby eliminating the infinitely long tail of digits.²

Example 13.2

Convert the repeating decimal $0.\overline{63}$ to its decimal representation.

Solution: We'll kill the tail again, but note that we have two digits after the decimal which repeat. This will require a slight adjustment. We'll start as we did before, by assigning an algebraic name to our number:

$$x = 0.6363\dots$$

² An alternative technique, which we'll discuss in Algebra 2, involves turning the repeating decimal into an infinitely long sequence of ever-decreasing numbers, and then finding the sum of that sequence. Yes, we can – in certain circumstances – add up infinitely many numbers. This is just one of the amazing things that awaits you in Algebra 2!

If we multiply both sides by 10, we'll have

$$10x = 6.3636 \dots$$

which is also a repeating decimal, but with a *different* repeating tail. We could work with this, but it's a bit easier to multiply by 10 again (in other words, to multiply the original equation by 100):

$$100x = 63.6363 \dots$$

Now we have an equation in which the number on the righthand side has exactly the same tail as in the original equation. So, we subtract:

$$\begin{array}{r} 100x = 63.\overline{63} \\ - \quad x = 0.\overline{63} \\ \hline 99x = 63.00 \end{array}$$

We divide both sides by 99, and then simplify our fraction to lowest terms. In the end, we have:

$$0.\overline{63} = \frac{63}{99} = \frac{7}{11}$$

The moral of the story is that we may have to adjust our method and choose the “just right” powers of 10. Consider how we might use kill the tail to turn $0.1\overline{6}$ back into $\frac{1}{6}$? (Note that the 6 repeats in the decimal form, but the 1 does not.)

Let's pause to reflect. In the first part of this section, we explained why every fraction can be written as either a terminating or repeating decimal. We can make this conversion using long division. Then we went on to show the reverse: that every terminating decimal can be written as a fraction (by reading it with its place value) and every repeating decimal can be written as a fraction (by killing the tail).

Armed with these tools, we might get the idea that *every decimal* number can be turned into a fraction. Unfortunately (or fortunately, depending on how you look at it), this is not the case.

13.1.3 Existence of irrational numbers

The **irrational numbers** are all of the real numbers that are not rational numbers. In other words, those decimal numbers that cannot be expressed as either a terminating or repeating decimal.

Back in chapter 1 we gave an example of such a number:

$$0.10110111011110111110\dots$$

This number clearly has a pattern. We might explain it by saying: “After the decimal point write one, then zero, the 2 ones, then zero, then 3 ones, then zero, and so on, always writing 1 more one than you did the last time.”

The problem is that it does not terminate (our pattern will continue forever), but it doesn't repeat either. The strings of 1's get longer and longer. There is never a set of always-repeating digits to group under a vinculum.

This single number is enough to prove that irrational numbers exist. Of course, there are lots of them. The famous number π is irrational, and in some sense is even more diabolical.

$$\pi \approx 3.1415926535\ 8979323846\ 2643383279\ 502884197\ 6939937510\ 5820974944\ 5923078164 \dots$$

This number doesn't even have a pattern that we can use to describe it (as far as we know). The digits go on infinitely, and come in a random sequence.

You may be wondering, "How do we know π is irrational?" After all, it may be clear why the first number with the ones and zeros is irrational, but how do we know for sure that π never terminates and never repeats?

Unfortunately, explaining the irrationality of π requires a bit more mathematics than we can get into here. However, we have learned enough to prove that certain other numbers are irrational. More on that at the end of section 13.2.

13.2 Square roots

We have already worked quite a bit with exponents, but every exponent so far has been an integer. Could we have a rational number as an exponent? If so, what would it mean?

Half power

// startup exploration

Consider the expression

$$9^{1/2},$$

that is, “nine to the power one-half”. What are some possible interpretations of this?

Use a calculator to explore what happens when we raise certain numbers to the one-half power (start with the natural numbers between 1 and 20). What patterns do you notice? What conjectures do you have about what’s happening?

Suppose we let $x = 9^{1/2}$. We’d like to find an alternative way of expressing x that uses only integer exponents. One approach is to multiply each side of this equation by itself. Then we’d have:

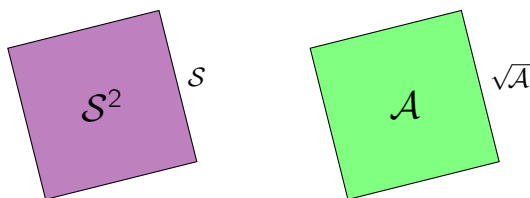
$$\begin{aligned} x &= 9^{1/2} \\ x \cdot x &= 9^{1/2} \cdot 9^{1/2} && \text{multiply each side by itself} \\ x \cdot x &= 9^{(1/2) + (1/2)} && \text{product rule for exponents} \\ x \cdot x &= 9^1 \\ x \cdot x &= 9 \end{aligned}$$

So, x is the number that when multiplied by itself gives 9 as the result. That could be either 3 or -3 , since $3 \cdot 3 = -3 \cdot -3 = 9$. Using some vocabulary that we already know: we say that $9^{1/2}$ is a *square root* of 9.

Square root

A **square root** of a is a number b such that $b \cdot b = a$.

The reason this is called a “square root” has to do with the geometric interpretation of this operation. If we have a square of side length S , then the area of the square is S^2 . Conversely, if we have a square with area A , then the *square root of A* gives us the side length of the square.



Positive numbers have two square roots: one positive, one negative. Both 3 and -3 are square roots of 9, since $(3)(3) = 9$ and $(-3)(-3) = 9$. Zero is the only number that has exactly one square root: the square root of 0 is 0. Negative numbers cause some problems when it comes to square roots (stay tuned for more on that).

Most of the time, we'll be working with the positive square root of a number, called the **principal square root**. For example, the principal square root of 9 is 3 and the principal square root of 25 is 5.

The checkmark-ish symbol we use to denote the square root of a is called the **radical**: \sqrt{a} . This symbol means "the principal square root of a ". So, $\sqrt{9} = 3$ and $\sqrt{25} = 5$. So, although -3 is a square root of 9, it is incorrect to write $\sqrt{9} = -3$ or $\sqrt{9} = \pm 3$. This isn't quite rise to the level of **Evil and Wrong**, but it's not right.³

When simplifying an expression or solving an equation, we won't always give both square roots. A good guideline as we go along is to pay close attention and use the notation given in the problem. For example:

$$\pm\sqrt{100} = \pm 10 \quad \text{and} \quad -\sqrt{36} = -6$$

In these cases, the notation specifically indicates that we want both the positive and negative square root, or only the negative square root.

13.2.1 Imaginary numbers

Consider the expression $\sqrt{-16}$. This is a bit of a problem. We're meant to find the principal square root of -16 , but the closest we can get is $4 \cdot -4 = -16$. It's true that 4 and -4 have the same absolute value, but they're not the same number, which means we're not *squaring* anything. So, there is no real number equal to $\sqrt{-16}$. Of course, -16 isn't special, the same argument applies to any negative number.

Square root of a negative number

If $a < 0$, then there is no real number equal to \sqrt{a} .

Note that this *doesn't* mean that -16 doesn't have any square roots. It simply means that the square roots are not real numbers. The square roots of -16 are members of a set called the *complex numbers* \mathbb{C} , but not members of the real numbers \mathbb{R} .

To build up the complex numbers, we introduce the so-called imaginary unit i which has the property that $i^2 = -1$. Remember, the domain of Algebra 1 is the set real numbers. But, the domain of Algebra 2 and beyond is the set of complex numbers. So, if you're intrigued about imaginary numbers, just hang in there.

³ An algebra note from the future! When we simplify the expression $\sqrt{m^2}$, the result is $|m|$, the absolute value of m . Do you see why? The value that comes out of the radical must be positive, because that notation gives the principal square root.

Our work in Algebra 1 will bring us mostly in contact with *square roots*, but other roots are possible. For example a “cube root” of a number a is a number b such that $b \cdot b \cdot b = a$. We write $\sqrt[3]{a}$ to denote the cube root of a . Negative numbers under the cube root symbol are no problem:

$$\sqrt[3]{-8} = -2 \quad \text{since} \quad -2 \cdot -2 \cdot -2 = -8$$

13.2.2 (,;) Irrationality of the square root of two

We have shown that irrational numbers exist. Consider the mathematical argument below, which explains why the square root of 2 is irrational.

Step 1. Assume (for the moment) that $\sqrt{2}$ is, in fact, rational. In other words, that it can be written as the ratio of two integers. Then, we can write

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

where a and b are relatively prime. That is to say, our fraction is in lowest terms.

Step 2. Square both sides of the equation above.

$$2 = \frac{a^2}{b^2}$$

Step 3. Multiply both sides of this equation by b^2 .

$$2b^2 = a^2$$

This means that a^2 is an even number. If a^2 is an even number, then a must be an even number.

Step 4. If a is an even number, then it is divisible by 2. In other words, there is an integer m such that $a = 2m$.

Step 5. Substitute $2m$ for a in the equation from Step 3, and simplify.

$$2b^2 = (2m)^2$$

$$2b^2 = 4m^2$$

$$b^2 = 2m^2$$

This means that b^2 is an even number. If b^2 is an even number, then b is an even number.

Step 6. If both a and b are even numbers, then the original fraction $\frac{a}{b}$ was not in lowest terms, which was our assumption! This contradiction shows that our original assumption cannot be true.

Thus, $\sqrt{2}$ cannot be written as a fraction. In other words, $\sqrt{2}$ is an irrational number.

Thoughts to chew on

Steps 3 and 5 both include assertions about numbers being even. How do we know when a number is even? Specifically, how do we know that a^2 and b^2 are even?

Steps 3 and 5 both go on to say something like “if a^2 is even, then a is even”. How do we know this is true? Under what circumstances will the square of a number be even or odd?

Step 6 argues that $\frac{a}{b}$ is not in lowest terms. How do we know this is true?

This is an example of *proof by contradiction*. We assume that some statement is true, then show that this assumption leads to some kind of impossible situation. The impossibility means we have to reject the original assumption. What was our original assumption in this proof? What is the contradiction that results from that assumption?⁴

⁴ Another famous proof by contradiction is Euclid’s proof that there are infinitely many prime numbers. Have a look in appendix A!

13.3 Simplified radical form

When we take the square root of a perfect square we get an integer as the answer. But, things are not so easy when taking the square root of a number that is not a perfect square. In fact, the square root of a non-square natural number will be an irrational number, like $\sqrt{2}$.

The *exact value* of an irrational number can only be represented using some kind of symbol, like π or $\sqrt{2}$. Writing out a decimal value – no matter how many decimals you write down – will always be an approximation. So, it's a good habit of mind to think "should I be giving an exact answer to this problem, or is a decimal approximation good enough". Very often, the context (or the directions) will make this choice clear.

To help us standardize the way we write radical expressions, we all agree to comply with *simplified radical form*.

Simplified radical form

A radical expression is considered completely simplified if...

1. Like radical terms have been combined.
2. The expression under the radical has no perfect square factors other than 1.
3. There are no fractions under the radical.
4. There are no radicals in the denominator of a fraction.

Over the next few sections, we will discuss each of these criteria and the algebraic manipulations that we can use to make sure our expressions comply. The first criteria is quite straightforward, so let's get right to it.

13.3.1 Like radical terms

Criteria #1 states that like radical terms must be combined. We combine radical terms as we do variable terms. For example, we are very familiar with the simplification

$$x + x = 2x.$$

We combine radical terms in exactly the same way:

$$\sqrt{5} + \sqrt{5} = 2 \cdot \sqrt{5} = 2\sqrt{5}.$$

Note, in particular, that the sum here is not $\sqrt{10}$. Similarly, $3x + 4x = 7x$ and so with radicals, we have $3\sqrt{21} + 4\sqrt{21} = 7\sqrt{21}$. When we have a multiplication of a number times a radical, we can omit the multiplication symbol.

Warning!

When we say like radical terms “can be combined”, don’t go thinking you can add the numbers under the radical. To add the values like this is **Evil and Wrong**.

$$\sqrt{3} + \sqrt{3} \neq \sqrt{6}$$

Since a radical is like an exponent, this is the equivalent of saying $(a+b)^2 = a^2 + b^2$ which, by now, we know is not generally the case. You can’t sprinkle that exponent across the sum!

While we’re at it, don’t get any ideas about splitting the radical-of-a-sum into the sum-of-radicals. This, too, is **Evil and Wrong**.

$$\sqrt{2+14} \neq \sqrt{2} + \sqrt{14}$$

13.3.2 Product properties of radicals

Building blocks

// startup exploration

The number 1 is the *additive building block* of the natural numbers. In other words: If we want to “build” any natural number using only addition, the only number we need is the number 1. Every natural number can be written as the sum of a bunch of 1’s.

What are the *multiplicative building blocks* of the natural numbers? Note that we have to say *blocks* (plural) since the number 1 is not enough: multiplying together a bunch of 1’s always gives us 1 as the product. What is the smallest collection of natural numbers that we need in order to build the rest using only multiplication?

The second criteria for simplified radical form states that the expression under the radical may have no perfect square factors other than 1. This may seem strangely worded. It clearly handles the idea that there should be no perfect squares under the radical, and that makes sense. Expressions like $\sqrt{4}$ and $\sqrt{25}$ can pretty obviously be simplified.

But, this criteria also catches expressions like $\sqrt{24}$ and $\sqrt{50}$ because those numbers, neither of which is a perfect square, each have a perfect square as a factor: 24 has 4 as a factor, and 50 has 25 as a factor.

How can we simplify an expression like $\sqrt{50}$ so that it has no perfect square factors under the radical? For help, we turn to:

Product rule of radicals

For any $a \geq 0$ and $b \geq 0$,

$$\sqrt{ab} = \sqrt{a} \cdot \sqrt{b}.$$

Note: In Algebra 1 we only use the square root version of this property, though in fact it applies to radicals of any degree: cube roots, fourth roots, and so on.

This property looks an awful lot like the product rule for exponents, which makes sense since here we are undoing the power of a product rule, where the power is the exponent one-half!

Example 13.3

Express $\sqrt{50}$ in simplified radical form.

Solution: We know $\sqrt{50}$ is not yet in simplified radical form because 50 is divisible by a perfect square, $50 = 25 \cdot 2$. We apply the multiplication property of radicals like so:

$$\begin{aligned}\sqrt{50} &= \sqrt{25 \cdot 2} && \text{rewrite 50 to show its perfect square factor} \\ &= \sqrt{25} \cdot \sqrt{2} && \text{product rule of radicals} \\ &= 5 \cdot \sqrt{2} && \text{simplify the square root of a perfect square}\end{aligned}$$

So, $\sqrt{50} = 5\sqrt{2}$. These two expressions are equal, but only the second expression satisfies the criteria of simplified radical form.

13.3.3 Different approaches to simplifying

There are a number of ways to go about applying this property to simplify expressions. Use whatever approach makes the most sense to you! Here are some alternatives, though you might find a different approach that fits you better. In any case, it will probably be helpful to learn a variety of methods. Depending on the problem, some methods may be easier to use than others.

For example, let's examine different ways to get $\sqrt{108}$ into simplified radical form.

Strategy 1: Largest square factor

In this strategy, we find the largest perfect square factor and simplify it using the product rule for radicals. We might notice that $108 = 3 \times 36$:

$$\sqrt{108} = \sqrt{36 \cdot 3} = \sqrt{36} \cdot \sqrt{3} = 6\sqrt{3}$$

Strategy 2: One square at a time

It might not be obvious what the largest perfect square is, so in this strategy we look for *any* perfect square factor and work one square at a time. For instance, we might notice that 108 is divisible by 9 (how can we quickly spot divisibility by 9?). Then:

$$\sqrt{108} = \sqrt{9 \cdot 12} = \sqrt{9} \cdot \sqrt{12} = 3 \cdot \sqrt{12} = 3 \cdot \sqrt{4 \cdot 3} = 3 \cdot \sqrt{4} \cdot \sqrt{3} = 3 \cdot 2 \cdot \sqrt{3} = 6\sqrt{3}$$

In this approach, we have to keep checking to see whether the number under the radical is “square-free” or not. After our first simplification, we have 12 under the radical. But 12 has 4 as a factor, so we have to do another simplification step.

This process might take a little longer, but it is sometimes easier to identify smaller perfect square factors and chip away at the problem, than it is to identify the largest perfect square factor and finish the problem in a single step.

Strategy 3: Sniper method

The idea here is to write the *prime factorization* of the number under the radical, and then look for pairs of factors.⁵ The factorization of $108 = 2 \cdot 2 \cdot 3 \cdot 3 \cdot 3$, so:

$$\sqrt{108} = \sqrt{2 \cdot 2 \cdot 3 \cdot 3 \cdot 3} = 2 \cdot 3 \cdot \sqrt{3} = 6\sqrt{3}$$

We’ve given this strategy the memorable (though perhaps gruesome) name *the sniper method*. Think of the radical as a prison. There are snipers outside and any number that tries to escape needs to have a decoy. A single factor of 2 is stuck inside for life, but if the 2 has a partner (that is, if there’s a $2 \cdot 2$ under the radical), then 2 can make a break for it!

But, only one of the partners survives the jailbreak. The snipers take out the decoy. In the example above, one 2 makes it out, and so does one 3. The final factor of 3 is partnerless, and left trapped inside its radical prison.

Example 13.4

TO DO.

⁵ One way to break a number down into primes is using a factor tree. For a handy, if unusually-formatted, list of prime numbers, see appendix A.

13.3.4 Quotient properties of radicals

Criteria #3 and #4 for simplified radical form are both pretty antiquated. They came about in the pre-calculator days when folks had to do a lot more calculation by hand and use large data tables to approximate radical values. Yet, these last two properties are still considered “standard” for simplified radical form.

Rules were made to be broken, though, and there will be times when it’s OK to break away from these criteria (#4 especially). But we’ll burn that bridge when we come to it. For now, all four criteria are in effect.

Both of these have to do with interactions between radicals and fractions. Criteria #3 disallows fractions under the radical, and criteria #4 forbids radicals in the denominator of a fraction.

To tackle Criteria #3, for example when faced with expressions like

$$\sqrt{\frac{4}{49}} \quad \text{or} \quad \sqrt{\frac{24}{25}},$$

we turn to:

Quotient rule of radicals

For any $a \geq 0$ and $b \geq 0$,

$$\sqrt{\frac{a}{b}} = \frac{\sqrt{a}}{\sqrt{b}}.$$

Again, this property applies to radicals of any degree (though for now we’ll focus on square roots). And again, this property is just like the quotient rule for exponents, but with a rational exponent.

Example 13.5

Express $\sqrt{\frac{4}{49}}$ and $\sqrt{\frac{24}{25}}$ in simplified radical form.

Solution: Here we have a fairly clear application of the rule:

$$\begin{aligned} \sqrt{\frac{4}{49}} &= \frac{\sqrt{4}}{\sqrt{49}} && \text{quotient rule of radicals} \\ &= \frac{2}{7} && \text{simplify square roots} \end{aligned}$$

In the second example, the numerator doesn’t contain in a perfect square, so we must apply the product

rule.

$$\begin{aligned}
 \sqrt{\frac{24}{25}} &= \frac{\sqrt{24}}{\sqrt{25}} && \text{quotient rule of radicals} \\
 &= \frac{\sqrt{24}}{5} && \text{simplify denominator} \\
 &= \frac{\sqrt{4 \cdot 6}}{5} && \text{product rule in the numerator} \\
 &= \frac{\sqrt{4} \cdot \sqrt{6}}{5} \\
 &= \frac{2\sqrt{6}}{5}
 \end{aligned}$$

13.3.5 Rationalizing the denominator

When simplifying an expression using the division property, we may encounter something like the following:

$$\sqrt{\frac{9}{2}} = \frac{\sqrt{9}}{\sqrt{2}} = \frac{3}{\sqrt{2}}$$

Back in the pre-calculator days, this led to criteria #4, no radicals in the denominator of a fraction. After all, long division is bad enough to carry out by hand. Why not try to put yourself into a situation that makes division as easy and accurate as possible, and avoid dividing by a big ugly decimal?

When we have a radical in the denominator of a fraction we have an *irrational denominator*. Our goal is to fix this by creating an equivalent fraction with a *rational denominator*. The process of making this translation is called **rationalizing the denominator**.

We will employ the trusty *identity property of multiplication*. Remember, multiplying a number by a fancy 1 does not change the value of the number. The trick will be to choose the way our version of 1 looks. We are going to choose a fancy version of 1 that when multiplied by our irrational denominator gives us a rational number (in fact, an integer).

Study the following examples:

Example 13.6

Write $\frac{3}{\sqrt{2}}$ in simplified radical form.

Solution: Note the clever use of multiplication by a fancy version of 1.

$$\begin{aligned}
 \frac{3}{\sqrt{2}} &= \frac{3}{\sqrt{2}} \cdot 1 && \text{identity property of multiplication} \\
 &= \frac{3}{\sqrt{2}} \cdot \frac{\sqrt{2}}{\sqrt{2}} && \text{substitute a fancy version of 1} \\
 &= \frac{3\sqrt{2}}{\sqrt{2} \cdot \sqrt{2}} && \text{multiply fractions} \\
 &= \frac{3\sqrt{2}}{2} && \text{definition of square root (in the denominator)}
 \end{aligned}$$

Note that we chose as our fancy 1 *exactly what we needed* to make the denominator of our fraction turn into an integer. This might seem like cheating, but it's a completely legal move, algebraically speaking.

Be sure to pay close attention. Sometimes we can use the division property in reverse to get rid of radicals in the denominator. We'll work the next example in two different ways to show the comparison.

Example 13.7

Write $\frac{\sqrt{84}}{\sqrt{6}}$ in simplified radical form.

Solution: First, we'll rationalize the denominator using a fancy version of 1.

$$\begin{aligned}
 \frac{\sqrt{84}}{\sqrt{6}} &= \frac{\sqrt{84}}{\sqrt{6}} \cdot 1 && \text{identity property of multiplication} \\
 &= \frac{\sqrt{84}}{\sqrt{6}} \cdot \frac{\sqrt{6}}{\sqrt{6}} && \text{substitute a fancy version of 1} \\
 &= \frac{\sqrt{84} \cdot \sqrt{6}}{\sqrt{6} \cdot \sqrt{6}} && \text{multiply fractions} \\
 &= \frac{\sqrt{84 \cdot 6}}{6} && \text{product rule for radicals} \\
 &= \frac{\sqrt{2 \cdot 2 \cdot 3 \cdot 7 \cdot 3 \cdot 2}}{6} && \text{simplify numerator using the sniper method} \\
 &= \frac{2 \cdot 3 \sqrt{2 \cdot 7}}{6} \\
 &= \frac{6\sqrt{14}}{6} \\
 &= \sqrt{14}
 \end{aligned}$$

Now, an alternative approach: We'll use the division property of radicals in reverse first, and then simplify

the fraction under the radical.

$$\begin{aligned}\frac{\sqrt{84}}{\sqrt{6}} &= \sqrt{\frac{84}{6}} && \text{division property of radicals} \\ &= \sqrt{\frac{14}{1}} && \text{simplify the fraction} \\ &= \sqrt{14} && \text{Voilà.}\end{aligned}$$

The second approach is much easier in this case. It pays to work smart and do a bit of planning before charging ahead with an algorithm blindly. It may not always be this easy, though. Under what circumstances will we be able to use the kind of shortcut?

Warning!

Answers with fractions must be simplified, but folks sometimes get overly aggressive with the simplification. Consider the following:

$$\frac{2}{\sqrt{6}} = \frac{2}{\sqrt{6}} \cdot \frac{\sqrt{6}}{\sqrt{6}} = \frac{2\sqrt{6}}{\sqrt{6} \cdot \sqrt{6}} = \frac{2\sqrt{6}}{6}$$

At this point, we can do one more simplification:

$$\frac{2\sqrt{6}}{6} = \frac{\sqrt{6}}{3} \quad \text{Yes!}$$

But we might be tempted to try and simplify even more:

$$\frac{\sqrt{6}}{3} = \frac{\sqrt{2}}{1} \quad \text{No!}$$

It's tempting, but we can't simplify using things *under* the radical and things *outside* the radical. That 6 under the radical cannot cancel with the 3 outside! To attempt such a simplification is **Evil and Wrong**.

Example 13.8

Determine the value of x given the equation: $2x^2 + 4x - 3 = 40$.

Solution: We'll use everything that we learned in chapter 12! Our first step is to ensure that the first term is a perfect square, so we multiply through by 2.

$$4x^2 + 8x - 6 = 80$$

This also gives us an even linear coefficient, so it looks like we're ready for the quadrangle method.

	2x	2
2x	4x ²	4x
2	4x	4

The quadrangle method predicts 4 as the constant term, but our equation has -6 . APOE to the rescue:

$$4x^2 + 8x - 6 = 80$$

$$4x^2 + 8x + 4 = 90$$

$$(2x + 2)^2 = 90$$

$$2x + 2 = \pm\sqrt{90}$$

$$2x = -2 \pm \sqrt{90}$$

$$x = \frac{-2 \pm \sqrt{90}}{2}$$

APOE: add 10 to both sides

based on the quadrangle diagram

square root of both sides

SPOE

DPOE

Almost there! Remember, we need to simplify our radicals! We can use the sniper method in this case:

$$\sqrt{90} = \sqrt{3 \cdot 3 \cdot 2 \cdot 5} = 3\sqrt{2 \cdot 5} = 3\sqrt{10}.$$

And so in the end, we have

$$\mathcal{S} = \left\{ \frac{-2 \pm 3\sqrt{10}}{2} \right\}$$

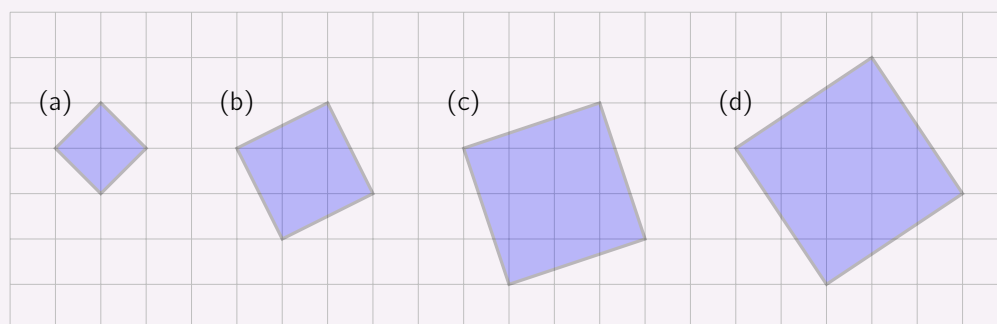
We'll admit that these answers ain't pretty (note that there are two answers there!), but they are the values that satisfy our original quadratic equation.

13.4 Coordinate geometry

Squarea

// startup exploration

As we saw in section 13.2, a square with area \mathcal{A} has side length $\sqrt{\mathcal{A}}$. Consider the figures below.



What is the area of each square? What is the side length of each square? Can you draw a square with side length $\sqrt{8}$? What about $\sqrt{13}$?

By the way, how do we know that each of these figures is, in fact, a square? (Hint: Think back to the slopes of parallel and perpendicular lines. What argument can we make for why the diagonal segments must be the same length in each figure?)

As an application of radicals and radical expressions, which are closely connected to the side lengths of squares, it's natural to discuss concepts from geometry. We'll begin with one of the most famous and important statements in mathematics.

13.4.1 Pythagorean theorem

It's a good bet that have seen the Pythagorean Theorem before, and that you will see it in every high school mathematics class you take, and many of the mathematics classed you take in college. In fact, the Pythagorean theorem is a foundational piece of an entire branch of mathematics based on the properties of triangles called *trigonometry*.⁶

⁶ Trigonometry begins with the study of triangles. *Trigon* is another way of saying a *triangle* – in fact it might be a better way of naming that shape! Most of the other polygons we know (pentagons, hexagons, octagons) have that *-gon* suffix, and the prefix *tri-* means “three” (as in tricycle).

Pythagorean theorem

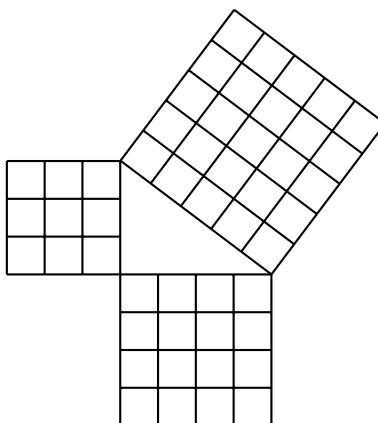
The sum of the squares of the lengths of the **legs** of a right triangle is equal to the square of the length of the **hypotenuse**.

In other words, if a and b represent the lengths of the legs (the perpendicular sides) of a right triangle, and c represents the length of the hypotenuse (the longest side, opposite the right angle), then

$$a^2 + b^2 = c^2.$$

The theorem is named after Greek philosopher and mathematician Pythagoras of Samos, who lived around 570–495 BCE.⁷ However, there is substantial evidence that the theorem was known to many different cultures from many different time periods. There is evidence, for instance, that the ancient Babylonians knew about Pythagorean triples (see the next section) more than 1000 years BCE.

Let's jump in with a famous right triangle: one with legs of length 3 and 4, and with hypotenuse of length 5. If we draw squares on the sides of the triangle, we can see that the sums of the areas of the two smaller squares (9+16) is exactly equal to the area of the largest square (25).



This relationship holds true for all right triangles as does its *converse*. The converse of the theorem states that if we have three numbers that satisfy the Pythagorean theorem, then we know that they must form the sides of a right triangle.

You might be asking yourself, “How do we know that the theorem is true for *every right triangle ever*?” Those who are curious might enjoy exploring this question at the end of this section.

13.4.2 Pythagorean triples

Pythagorean triple

A *Pythagorean triple* is a set of three positive integers satisfying the Pythagorean theorem.

There are infinitely many Pythagorean triples, and it is quite handy to know a few by heart. (This is because they are used quite a bit in problems, and those delightful standardized tests we all know and love.) Here are a few common Pythagorean triples:

$$(3, 4, 5) \qquad (5, 12, 13) \qquad (7, 24, 25) \qquad (9, 40, 41) \qquad (8, 15, 17)$$

The benefit of memorizing a few of these is that, if you see that a right triangle that has leg lengths of 7 and 24, you know that the hypotenuse has length 25 without having to do any calculations.

The triples above are called **primitive Pythagorean triples** because the three values are relatively prime. You can generate new Pythagorean triples by scaling up a triple that you know. For example $(6, 8, 10)$ is a triple formed by scaling up the $(3, 4, 5)$ triple by a factor of 2.

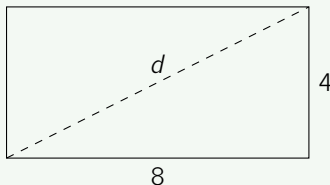
Find-the-missing-side problems

The classic application of the Pythagorean theorem is to find a missing side length, either a leg or a hypotenuse, and you may have solved problems like this before. Now that we have some algebra skills, however, our answers should be given in simplified radical form! No more decimal approximations (unless the directions state otherwise)!

Example 13.9

Find the length of the diagonal of a 4-by-8 rectangle.

Solution: This problem might not, at first, seem to have anything to do with the Pythagorean theorem. The theorem, after all, is about right triangles, and this question is about a rectangle! Drawing a picture helps to reveal the connection:



If we let d represent the length of the diagonal, then we can see that it is the hypotenuse of a right triangle with legs of length 4 and 8. So:

$$d^2 = 4^2 + 8^2$$

Pythagorean theorem

$$d^2 = 16 + 64$$

$$d^2 = 80$$

$$d = \sqrt{80}$$

square root of both sides

$$d = \sqrt{2 \cdot 2 \cdot 2 \cdot 2 \cdot 5}$$

simplify using the sniper method

$$d = 2 \cdot 2 \cdot \sqrt{5}$$

$$d = 4\sqrt{5}$$

So, a 4-by-8 rectangle has a diagonal which is $4\sqrt{5}$ units long.

To get a feel for whether this answer is reasonable, we could find a decimal approximation using a calculator (it's about 9 units long, which seems OK), or we could reason as follows. We know that $\sqrt{4} = 2$, and so $\sqrt{5}$ must be a bit more than 2. So, $4\sqrt{5}$ must be a bit more than 8. This seems reasonable.

On the other hand, if we had gotten $4\sqrt{8}$ we might reason that $\sqrt{9} = 3$ and so $4\sqrt{8}$ must be a bit less than $4\sqrt{9} = 4 \cdot 3 = 12$. This is too long for the diagonal of a 4-by-8 rectangle, since the two sides together are only 12 units long in total!

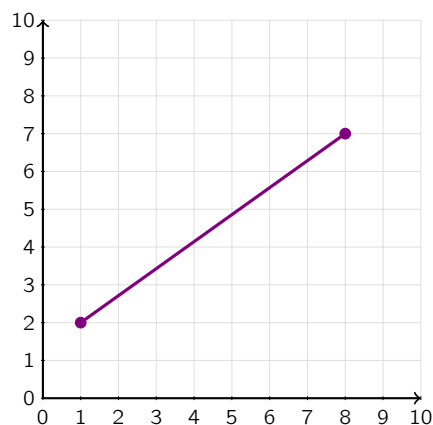
13.4.3 Distance formula

Grid distance

// startup exploration

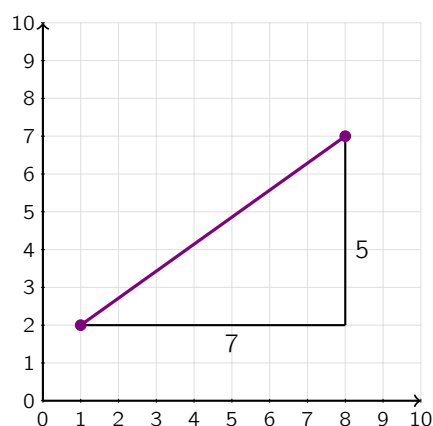
Find the length of the line segment connecting the points $(1, 2)$ and $(6, 7)$.

One important application of the Pythagorean theorem is called the **distance formula**. It is a formula that we can use to calculate the distance between two points on a coordinate grid. In the startup exploration, we have the segment pictured below:



If we think of the line segment into the hypotenuse of a right triangle, then we can use the Pythagorean theorem! The legs are the vertical and horizontal distances between the points, as in a slope triangle.

Draw the triangle, find the horizontal and vertical distances, then apply the Pythagorean theorem.



The horizontal distance is 7 units, and the vertical distance is 5 units. Those are the legs of the right triangle. Then, we use the theorem to find the length of the hypotenuse:

$$a^2 + b^2 = c^2$$

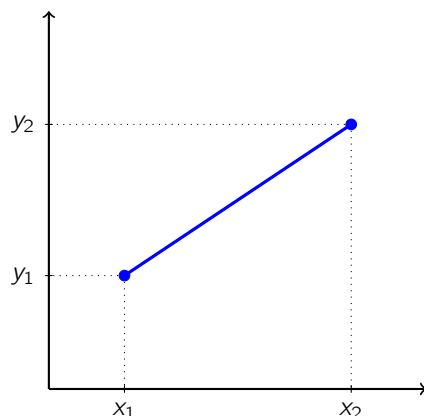
$$5^2 + 7^2 = c^2$$

$$25 + 49 = c^2$$

$$74 = c^2$$

$$\sqrt{74} = c$$

Since the process is the same every time, we can generalize to find the distance between any two points (x_1, y_1) and (x_2, y_2) in the plane.



Distance formula

Given two points in the plane (x_1, y_1) and (x_2, y_2) , the length d of the line segment connecting the points is given by the formula:

$$d = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}.$$

This might look like a bunch of alphabet soup, much harder to remember than the Pythagorean theorem. But remember: this *is* the Pythagorean theorem! If you forget the formula, don't panic! Just remember that the line segment between the two points is the hypotenuse of a right triangle.

Example 13.10

Find the distance between the points $(5, 12)$ and $(-4, -2)$.

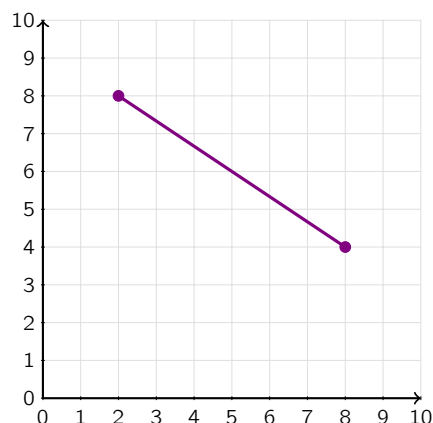
Solution: We will use the distance formula, but note that we have both subtraction and negative numbers. Watch those minus signs!

$$\begin{aligned} d &= \sqrt{(5 - (-4))^2 + (12 - (-2))^2} \\ &= \sqrt{9^2 + 14^2} \\ &= \sqrt{81 + 196} \\ &= \sqrt{277} \end{aligned}$$

Since 277 is prime, we know our answer complies with simplified radical form, so we're all done. The distance between the two points is $\sqrt{277}$ units.

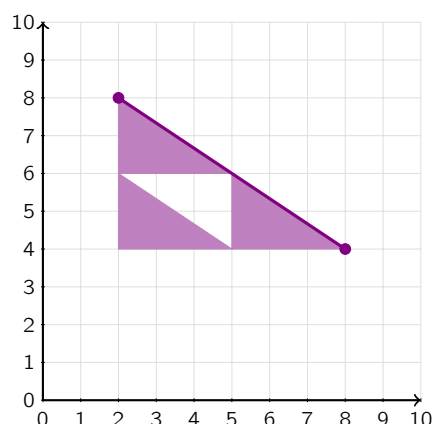
13.4.4 Midpoint formula

When working with the distance formula, there is a related formula for finding the coordinates of *midpoint* of a given line segment. Suppose we wanted to find the midpoint of the line segment connecting $(2, 8)$ and $(8, 4)$.



Well, it sure looks like the midpoint of this segment is the point $(5, 6)$... but can we be sure?

One way to explain this is by drawing in the right triangle, and then chopping that triangle into four congruent sub-triangles. (Remember our work with the Sierpiński triangle ages ago?)



This diagram suggests that the x -coordinate of the midpoint of the hypotenuse is exactly halfway between the x -coordinates of the legs. The same goes for the y -coordinate. So, to find the coordinates of the midpoint, all we have to do is average the coordinates of the endpoints! Once again, you don't have to memorize this formula if you remember where it comes from!

Midpoint formula

Given a line segment with endpoints (x_1, y_1) and (x_2, y_2) , the coordinates of the midpoint of the segment are

$$\left(\frac{x_1 + x_2}{2}, \frac{y_1 + y_2}{2} \right)$$

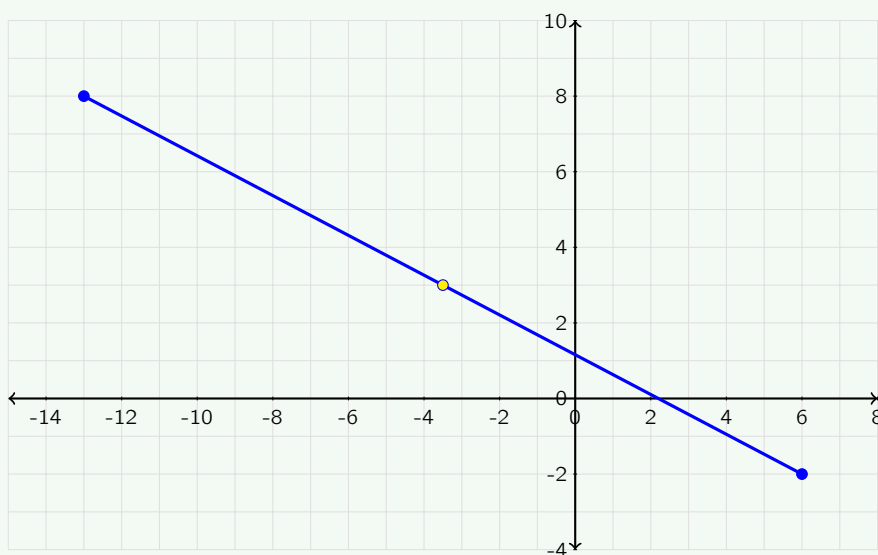
Example 13.11

Find the midpoint of the segment connecting the points $(-13, 8)$ and $(6, -2)$.

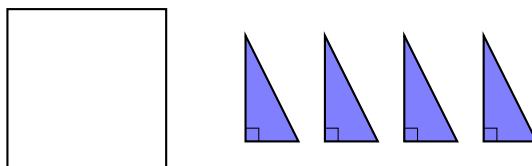
Solution: Let's calculate the midpoint first, and then check our answer by drawing a picture. The formula is pretty straightforward. The midpoint should be located at:

$$\left(\frac{-13 + 6}{2}, \frac{8 + -2}{2} \right) = \left(\frac{-7}{2}, \frac{6}{2} \right) = (-3.5, 3)$$

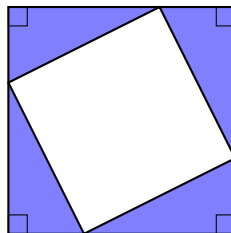
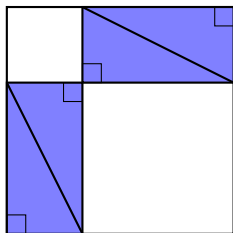
Here's a graph to check that our answer is reasonable:

**13.4.5 (;,;) Proof of the Pythagorean theorem**

Imagine a pizza box and four identical slices of pizza. . . where the pizza slices are right triangles. (Not typical for pizza slices, we know, but go with it.)



Consider two different arrangements of the same four pizza slices inside the same pizza box. The slices are arranged so that they fit inside without overlapping, and they lay flat on the bottom of the box.



Step 1: Consider the image on the left: Write an expression for the area of the box that is left *uncovered*. You may find it helpful to label the sides of the triangles, or the sides of the box (or both).

You may be tempted to say that the uncovered regions are squares. How do we know – for sure – that these two regions are actually *squares*, as opposed to some other quadrilateral?

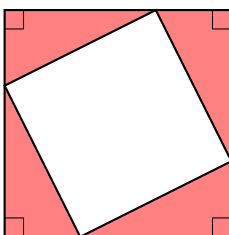
Step 2: Consider the image on the right: Write an expression for the area of the box that is left *uncovered*. Use the same labels you assigned when studying the other picture.

Again, you may be tempted to say that the uncovered region is a square. How do we know that it is a square (and not, say, a rhombus)?

Step 3: The area that is uncovered by the pizza slices is the same. (How do we know this?) What does this tell us about the two expressions for the uncovered area?

Alternative approach

In fact, we can explain the theorem using only the right-hand diagram. Let a and b be the lengths of the legs of one of the triangles, and let c be the length of the hypotenuse.



Step 1: Express the sides of the largest square (the pizza box itself) in terms of a and b .

Step 2: Express the area of the largest square in terms of a and b . (Hint: You'll need the sum to a power rule.)

Step 3: The area of the largest square can also be expressed at the sum of the areas of the four triangles plus the area of the tiled square. Express the area of the largest square in this way (in terms of a , b , and c).

Step 4: We now have two ways of expressing the area of the largest square. What happens when set them equal to each other?

Generalizing

All the figures in this discussion have been drawn using a particular square and a particular triangle. Explain why these arguments are proof that the Pythagorean theorem is true for any right triangle, not just the specific triangle that is pictured in these diagrams.

» Chapter summary «

Our algebra toolbox grows ever larger! We can now handle the simplification of numerical expressions that include square roots. This, combined with our work from the previous chapter, means that we can handle even the most “impolite” of quadratic equations. Plus, along the way we have discovered some fundamental results from geometry.

These skills are valuable, without a doubt, but we haven’t looked closely at quadratic functions in detail. These are the focus of the next chapter. Onward!

Nature creates curved lines while humans
create straight lines.

Hideki Yukawa, Japanese theoretical physicist

Chapter 14

Quadratic functions

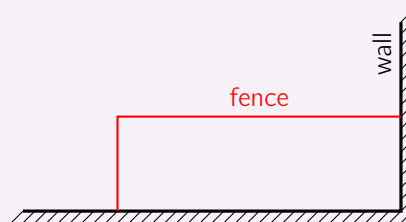
In the last few chapters, we have studied quadratic equations, the methods for solving them, and the various algebraic rules for manipulating the kinds of answers we get (square roots). We turn now to quadratic relationships and situations that can be represented using quadratic functions. We'll look more closely at the equations and graphs of quadratic functions, and look closely at one of the classic applications from physics.

14.1 Quadratic Relationships

Compost Heap

// startup exploration

Uncle François wants to make a compost heap in the corner of his vegetable garden. He plans to enclose a rectangular area using the two existing walls of the garden, plus 8 yards of leftover fencing.



If he wants to enclose the largest possible area, how should he arrange the fence? What is the maximum area he can enclose?

A good way to start this problem might be simply to get our hands dirty and see what happens.¹ If François

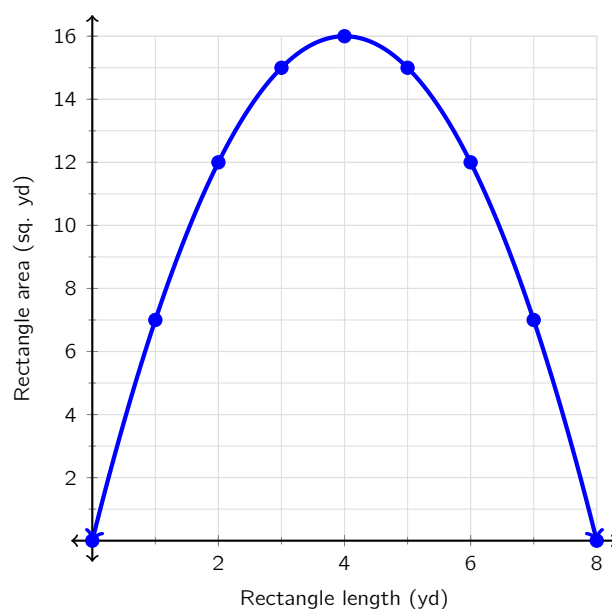
¹ After all, we are talking about a compost heap.

has 8 yards of fencing, he could fence in a rectangle that is, say, 2 meters long and 6 meters wide. That would give him a compost heap with an area of 12 square yards. Can he enclose more area than this?

To enumerate all of François's options, suppose we set up a data table that shows the dimensions of the rectangle and its area.

Length (yd)	Width (yd)	Area (sq. yd)
0	8	0
1	7	7
2	6	12
3	5	15
4	4	16
5	3	15
6	2	12
7	1	7
8	0	0

Aha! There are some patterns to be found here. It certainly seems that the 4-by-4 rectangle (the square, that is) encloses the most area. Indeed, if we plot a graph of length versus area, we have another representation of this maximum value.



Note that we connected the dots in our graph. Under the given constraints, a reasonable length for the rectangle is any real numbers between 0 and 8. Of course, if the length is 0, then it's not really a rectangle...so we should exclude the value $x = 0$ as an option. The same goes for $x = 8$ (since in this case the width of the rectangle is zero). So, this is the domain of our quadratic function in the context of the problem: x must be greater than zero and less than 8, or $0 < x < 8$.

Note that the range of our function includes all y values between 0 and 16 square yards. In fact given the domain, we have $0 < y \leq 16$, since the area can be 16, but can never actually be 0.

This graph demonstrates a number of important properties about quadratic functions: We get a smooth U-shaped graph that is perfectly symmetrical, and which has an extreme value (a maximum value in this case).² These aspects of the graph all have names.

Parabola

The U-shaped graph of a quadratic function.

Vertex

The extreme (lowest or highest) point on a parabola.

Axis of symmetry

The imaginary line that passes through the vertex and divides the parabola into two symmetric halves.

We also see a familiar pattern in the sequence of area values: a constant second difference.

Area values	0	7	12	15	16	15	12	7	0
Differences		7	5	3	1	-1	-3	-5	-7
Second differences			-2	-2	-2	-2	-2	-2	-2

Recall that we can't simply rely on the "increases then decreases" pattern (or the reverse) to determine whether a sequence or a data table is quadratic. We must look closely at the pattern and find that constant second difference.

14.1.1 A rectangular rule

What equation describes our rectangle area function? In our graph we are displaying area as a function of length. We know how to calculate the area of a rectangle: that's length times width. The trick in this case is to notice that the length (ℓ) and width (w) of the rectangle always add up to 8 (the fixed length of fence we have to work with). So:

$$\ell + w = 8 \quad \text{or} \quad \ell = 8 - w$$

² There is an interesting geometric definition of a parabola. First note that a circle is the set of all points that are a fixed distance from a given point. The "given point" is called the center of the circle, and the "fixed distance" is called the radius of the circle. A parabola is the set of all points which are the same distance from a fixed point (called the *focus*) and a fixed line (called the *directrix*). We'll see more definitions like this later, when we study the so-called *conic sections* in Algebra 2.

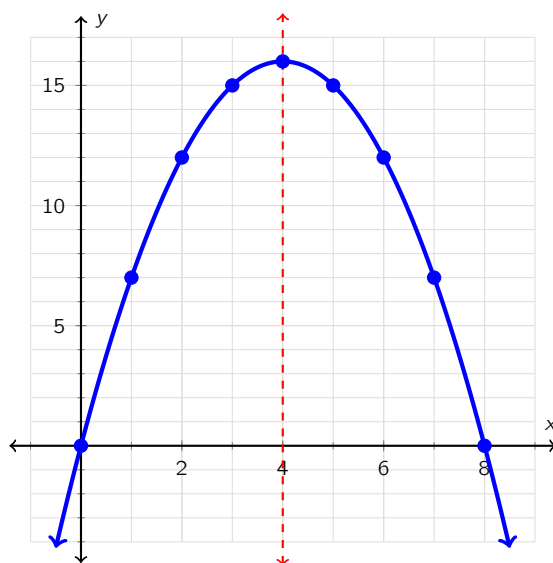
So, we can write a rule for $A(\ell)$, the area of the rectangle as a function of its length: $A(\ell) = \ell(8 - \ell)$. Of course we can always use the distributive property to eliminate those parentheses:

$$A(\ell) = \ell(8 - \ell) = 8\ell - \ell^2$$

So, we have two different forms of our quadratic equation! In fact, there is a third form that is also important. In the next section (and then again in the next chapter), we will look more closely at quadratic rules. In the meantime, note that we did something a bit like we did with linear functions. Back in chapter 7, we turned point-slope form into slope-intercept form using the distributive property and combining like terms. Part of our goal now will be to understand the techniques that allow us to convert between the different forms of a quadratic equation.

14.1.2 Powerful symmetry

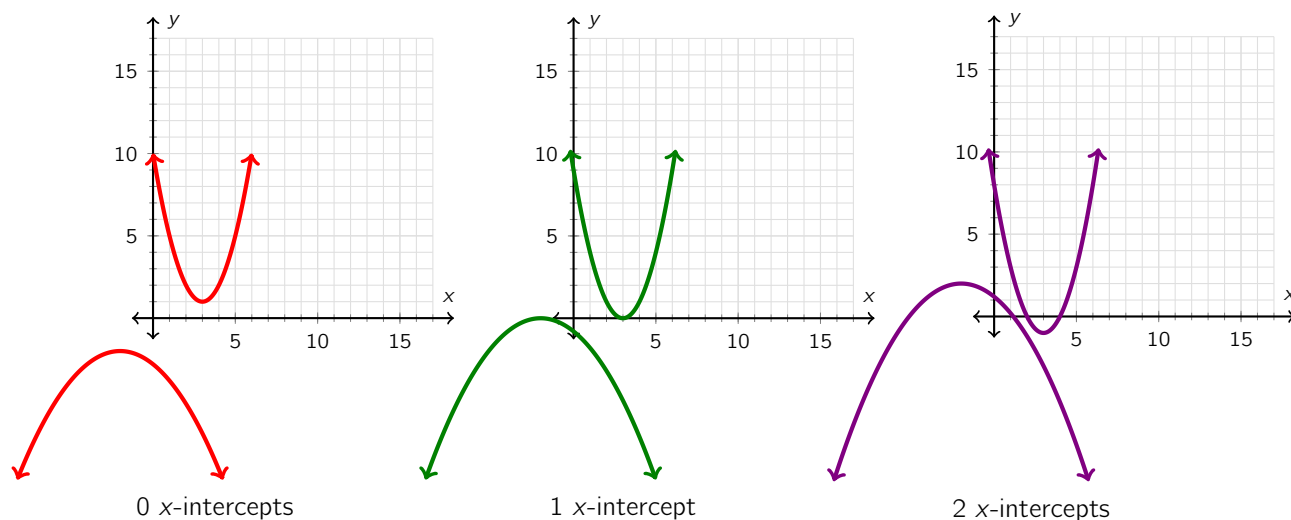
The symmetrical qualities of a quadratic function will be very helpful to us. We have already seen how the symmetry of the square can help us to solve quadratic equations (via the quadrangle method), and we will use the symmetry of the parabola, too.



Note that the line of symmetry, the line $x = 4$ in this case, cuts the parabola into two identical halves. This graph has two x -intercepts that are equidistant from the line of symmetry. In fact, every point (other than the vertex) has a buddy on the other side of the line of symmetry: a point that is at the same y -value, and an equal distance from the line of symmetry.

Speaking of intercepts

Every quadratic function has one and only one y -intercept. The situation is slightly more complicated for x -intercepts. The graph above has two x -intercepts. Can you picture the other possibilities?



A parabola may have 0, 1, or 2 x -intercepts. If there are no x -intercepts, then the graph lies entirely above or below the x -axis. If there is only one x -intercept, then it must be the vertex! Otherwise every point comes with a buddy at the same height.

A final thought on fixed perimeters

You may have noticed that the rectangle which maximized the area of the compost heap was, in fact, a square. This is true in general: given a fixed perimeter, many rectangles can be drawn which have that perimeter, and the one with the largest area will be the square.

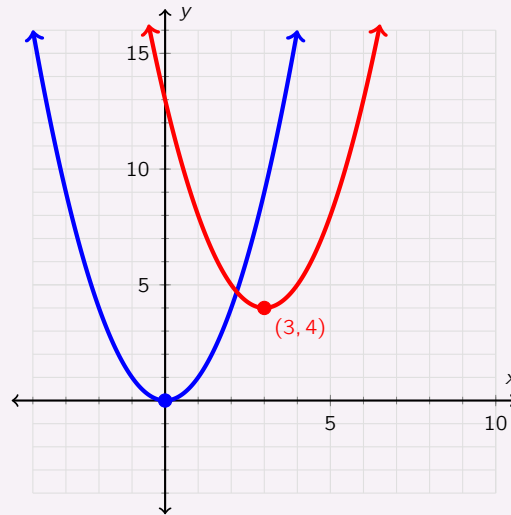
What if we lift the restriction that the shape be a rectangle? Given a fixed perimeter, could we create a hexagon with area larger than our square? What shape should we make to get the maximum possible area of all possible shapes?

14.2 Vertex form

Move that curve

// startup exploration

The graph below shows the parent of the quadratic family, $y = x^2$, and a offspring function that has the same shape but is shifted so that the vertex is at the point $(3, 4)$.



Look back over our discussion of transformations (back in section 4.3), and write an equation for the offspring graph. Tinkering around with some some graphing technology may be helpful!

With some careful observation and a bit of tinkering, we can move the parent parabola to its new position using the rule:

$$y - 4 = (x - 3)^2 \quad \text{or} \quad y = (x - 3)^2 + 4$$

The “ $y - 4$ ” in the rule controls the up-and-down movement of the parabola, shifting it 4 units upwards in this case. The “ $x - 3$ ” inside the parentheses is the component that controls left-and-right movement. The values 3 and 4 are related to the location of the vertex of the parabola. Naturally enough, an equation like this is called the **vertex form**.

Vertex form of a quadratic function

An equation of the form

$$y = a(x - h)^2 + k,$$

where $a \neq 0$, is a quadratic function in vertex form. The point (h, k) is the vertex of the parabola.

Note the similarity of this form to the point-slope form of a line:

$$\text{point-slope form of a linear function: } y = m(x - x_1) + y_1$$

$$\text{vertex form of a quadratic function: } y = a(x - h)^2 + k$$

14.2.1 Graphing quadratics in vertex form

To graph a quadratic function give its rule, we could make a table of values that has many points. In the past, we've plotted 7 data points to make a graph: for example using the integers between -3 and 3 as the x -values.

When working with linear equations, though, we learned that we need just two points to graph a line. This cuts down on the work! There's good news here, too: every parabola can be created with exactly 3 points. Our goal as we graph is to find a few strategic points to plot so that we can sketch a reasonable curve efficiently.

Good candidates for "strategic points" include: the vertex, the y -intercept, and the x -intercepts (if they exist). We will also take advantage of the fact that parabolas are symmetric.

If we are given a line in vertex form, then the location of the vertex (and the line of symmetry) is right there in the equation. That's a huge benefit of vertex form!

For example, given the function $y = \frac{2}{3}(x - 3)^2 + 2$, we can read the location of the vertex directly from the rule: $(3, 2)$. This also tells us that the line of symmetry (which passes through the vertex) must be the vertical line $x = 3$.

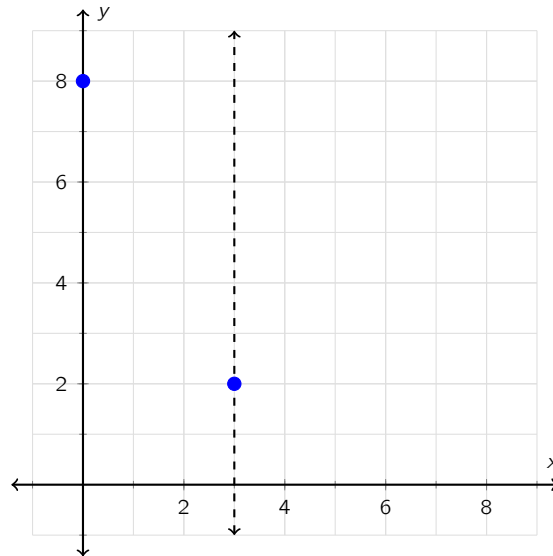
What else can we learn? The coefficient 3 in front tells us that this parabola will open upwards (since 3 is positive), and that it will be narrower than $y = x^2$ (the parent function for this family).

Plus, since the parabola opens upwards and has its vertex above the x -axis, this parabola must be one of those that has no x -intercepts. We really can learn a lot just from looking at the equation!

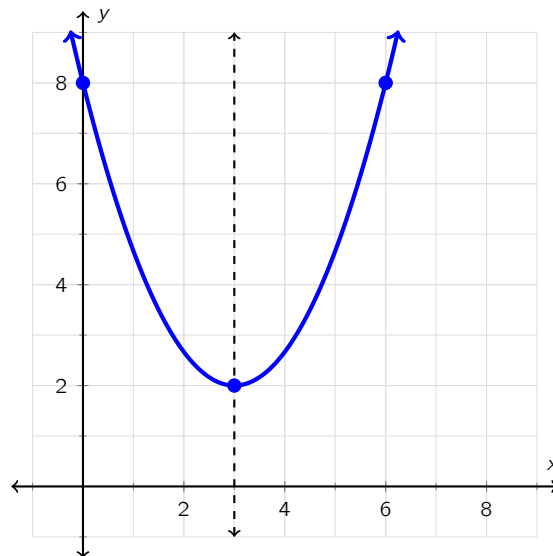
We still have only one point, though, and we need three points to plot our graph. How can we find the y -intercept? This is the value of the function when $x = 0$, so we can evaluate the function at that point and see what we get!

$$\begin{aligned} y &= \frac{2}{3}(x - 3)^2 + 2 && \text{original function} \\ y &= \frac{2}{3}(0 - 3)^2 + 2 && \text{let } x = 0 \text{ to find the } y\text{-intercept} \\ &= \frac{2}{3}(-3)^2 + 2 \\ &= \frac{2}{3}(9) + 2 \\ &= 6 + 2 \\ &= 8 \end{aligned}$$

So, the y -intercept is the point $(0, 8)$. We now have two points that we can plot on our graph. How do we find the third?



Here's where the symmetry of the parabola comes in handy. The point that is symmetrical to the y -intercept will be the same horizontal distance from the line of symmetry, and at the same height as the y -intercept. So, the point we want is $(6, 8)$ – it is at the same height (8 units), and three units away, but on the other side of the line of symmetry. We can plot this point, and graph the parabola!



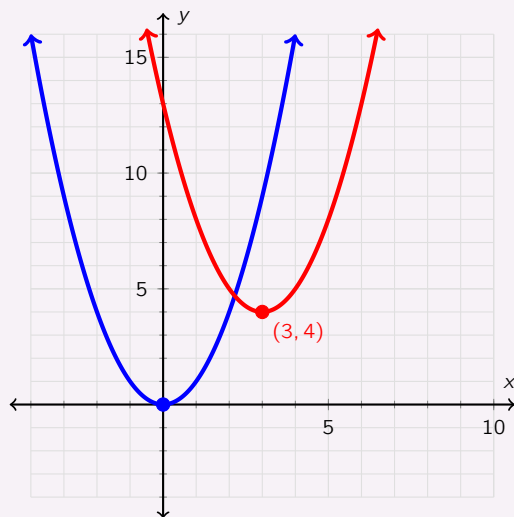
Now, if we find that the vertex lies on the y -axis, then we can't use the y -intercept as our second point (because the vertex *is* the y -intercept). We'll address graphing parabolas like this in the next section.

14.3 Standard form of a quadratic

Rename that curve

// startup exploration

In the last section, we wrote a rule for shifting the parent of the quadratic family, $y = x^2$, such that the vertex was at the point $(3, 4)$.



The rule we wrote was $y = (x - 3)^2 + 4$. Use everything you know about simplifying algebraic expressions to rewrite this rule in its simplest form. Use graphing technology to verify that your simplified rule and the original rule produce the same graph.

The first problem with the rule in the startup exploration is that patenthesized expression. To eliminate this, we can use the distributive property (raising a sum to a power). After that, we'll need to combine some like terms:

$y = (x - 3)^2 + 4$	the original equation
$= (x - 3)(x - 3) + 4$	expand the exponent
$= x(x - 3) - 3(x - 3) + 4$	distribute one $(x - 3)$ over the other
$= x^2 - 3x - 3x + 9 + 4$	distribute again (watch the signs!)
$= x^2 - 6x + 13$	combine like terms

Now we have a new equation that is equivalent to the original: an equation in what we call *standard form*.

Standard form of a quadratic function

An equation of the form

$$y = ax^2 + bx + c,$$

where $a \neq 0$, is a quadratic function in standard form.

In some sense, this equation has obscured key things about the parabola. It is not obvious, for instance, where the vertex is. But, something new has been revealed. What does that “+13” tell us about the graph?

Compare this form to the slope-intercept form of a line:

$$\text{slope-intercept form of a linear function: } y = mx + b$$

$$\text{standard form of a quadratic function: } y = ax^2 + bx + c$$

Note that the constant term in each case is the y -intercept of the graph. This makes sense, for if we evaluate each function at $x = 0$ (that's the y -axis), then any term with an x in it disappears and we are left only with the constant term. To be clear in the case of our standard-form quadratic: when $x = 0$, we have

$$y = a \cdot 0^2 + b \cdot 0 + c = 0 + 0 + c = c.$$

14.3.1 Graphing standard form

Vertex form is nice for graphing, since the vertex is right there in the equation. When tasked with graphing an equation in standard form, our strategy will be much the same... though we'll have to do a bit more work. Our first goal will be to find the vertex.

Axis of symmetry

As with all things parabolic, symmetry is our friend. We know that the point $(0, c)$ is on the parabola: that's the y -intercept. If we assume that this point is not the vertex of the parabola, then we know there is a second point on the parabola, also at height c . What is the x -coordinate of this other point?

We can figure that out by solving an equation. We want to know what x values lead to $y = c$, so that is like solving the equation:

$$ax^2 + bx + c = c \quad \text{or, by SPOE} \quad ax^2 + bx = 0$$

We will soon learn a quick way to solve this equation, but in the meantime, we can use the quadrangle method! We multiply through by a (to make the quadratic term a perfect square) and then by 4 (to ensure that the linear term has an even coefficient). This gives us:

$$4a^2x^2 + 4abx = 0$$

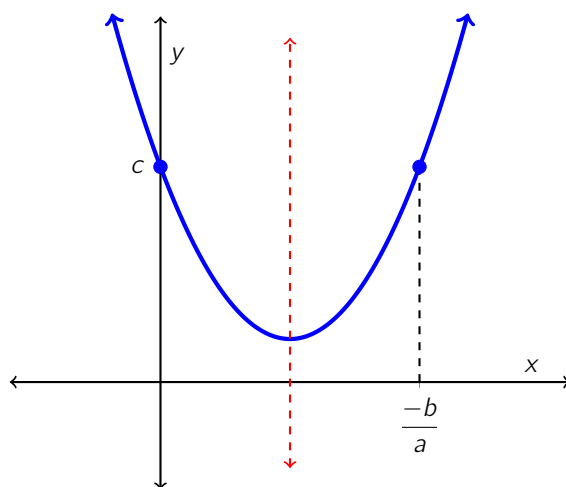
Arranging in the quadrangle gives us:

	$2ax$	b
$2ax$	$4a^2x^2$	$2abx$
b	$2abx$	b^2

We must add b^2 to both sides of our equation, to accomodate the quadrangle diagram. From here, it's clear sailing:

$$\begin{aligned}
 4a^2x^2 + 4abx &= 0 \\
 4a^2x^2 + 4abx + b^2 &= b^2 && \text{APOE} \\
 (2ax + b)^2 &= b^2 && \text{based on the quadrangle method} \\
 2ax + b &= b \text{ or } -b && \text{square root of both sides} \\
 2ax &= 0 \text{ or } -2b && \text{SPOE} \\
 x &= 0 \text{ or } \frac{-b}{a}
 \end{aligned}$$

So, we have discovered two x -values for which the quadratic equation produces $y = c$. Those values are $x = 0$ (we knew this one already, but it's nice that our work confirms our expectations!), and $x = \frac{-b}{a}$. What does this tell us?



We know that the line of symmetry must be exactly halfway between these two points. In other words, exactly halfway between 0 and $\frac{-b}{a}$. We just have to divide that distance in half and we have found an equation for the line of symmetry!

Axis of symmetry

Given a quadratic equation of the form $y = ax^2 + bx + c$, the axis of symmetry for the parabola is the vertical line at

$$x = \frac{-b}{2a}.$$

Once we have found the axis of symmetry, we know the x -coordinate of the vertex. We can then evaluate the function at this point to find the y -coordinate.

Example 14.1

Locate the vertex of the parabola given by $y = 2x^2 + 12x + 11$.

Solution: We can use our equation for the line of symmetry to start us out. In this case, we have $a = 2$ and $b = 12$, the this parabola is symmetric about the line

$$x = \frac{-b}{2a} = \frac{-12}{2 \cdot 2} = \frac{-12}{4} = -3.$$

So the x -coordinate of the vertex is -3 . We plug this value into our equation to determine the corresponding y -coordinate:

$$\begin{aligned} y &= 2x^2 + 12x + 11 \\ &= 2(-3)^2 + 12(-3) + 11 \\ &= 18 - 36 + 11 \\ &= -7 \end{aligned}$$

So, we have located the vertex! It is the point $(-3, -7)$.

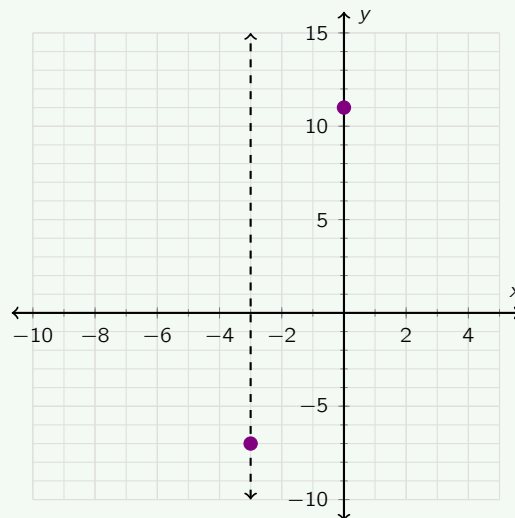
Once we have found the vertex, we have one point on our parabola. To graph it, we must find two more. If the vertex is not on the y -axis, then this job isn't too hard because we get a second point for free: the y -intercept!

Once we plot the vertex and the y -intercept, we can find and plot the point that is symmetrical to the y -intercept on the other side of the axis of symmetry.

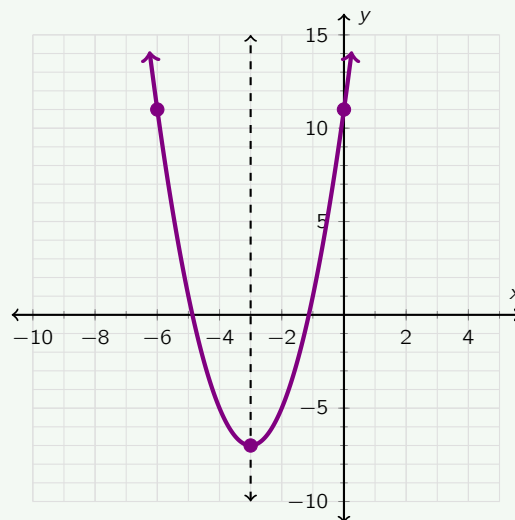
Example 14.2

Graph the parabola given by $y = 2x^2 + 12x + 11$.

Solution: In the previous example, we found the vertex of this parabola to be $(-3, -7)$. We can also see from the equation that the y -intercept is $(0, 11)$. We can plot these points quite easily:



Now, we're on the hunt for the point that is symmetrical to the y -intercept. This point will be at the same height ($y = 11$) but on the other side of the line of symmetry. The line of symmetry is three units from the y -axis, so the point we seek must be three units away on the other side of the line of symmetry. That means we want the point $(-6, 11)$. We can plot this point, and graph the parabola!



As we mentioned above, we may find that the vertex of our parabola lies on the y -axis. In this case we can't use the y -intercept as our second point (because the vertex and the y -intercept are *the same point*). So in

this case, we are forced to choose another x -value and compute its corresponding y -value. Since we have the luxury of choosing our x -value it makes sense to try and choose a value that makes our calculations easy, but that's up to us to decide.

Can we predict what the equation of such a quadratic will look like? If the axis of symmetry is the vertical line $x = 0$, then the x -coordinate of the vertex is 0. So in vertex form, $h = 0$. This means that we have:

$$y = a(x - h)^2 + k = a(x - 0)^2 + k = ax^2 + k$$

So, if the quadratic function has no linear term (no bx term), then its line of symmetry is y -axis.

Example 14.3

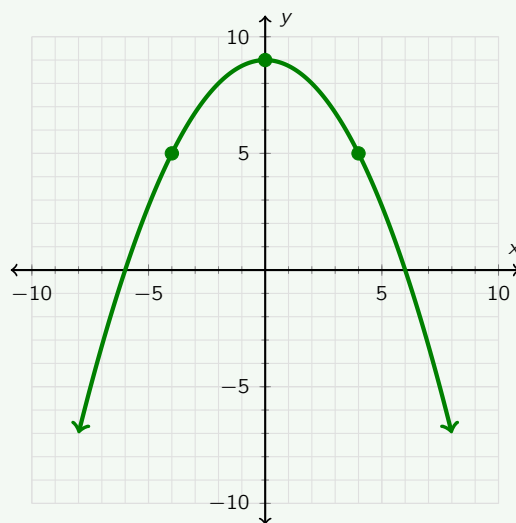
Graph the parabola $f(x) = -\frac{1}{4}x^2 + 9$.

Solution: This quadratic has no linear term, the line of symmetry is the y -axis. The y -intercept is 9, and the coefficient a is negative, so the parabola opens downward. But now it's up to us to find another two points.

Suppose we choose $x = 4$, which might help to eliminate that fraction. In that case, we have:

$$f(4) = -\frac{1}{4}(4)^2 + 9 = -\frac{1}{4}(16) + 9 = -4 + 9 = 5.$$

So the point $(4, 5)$ is on the graph. Because of symmetry, this means that the point $(-4, 5)$ is also on the graph. We have our three points!



In the previous example, we have another option: we could find the x -intercepts, which occur when $y = 0$.

Here's how we'd do it for the example:

$$y = -\frac{1}{4}x^2 + 9 \quad \text{original equation}$$

$$0 = -\frac{1}{4}x^2 + 9 \quad \text{let } y = 0 \text{ to find the } x\text{-intercepts}$$

$$-9 = -\frac{1}{4}x^2 \quad \text{SPOE}$$

$$36 = x^2 \quad \text{MPOE}$$

This means that x must be 6 or -6 . In other words, the x -intercepts are the points $(6, 0)$ and $(-6, 0)$, and this conclusion is shown on the graph that we drew above!

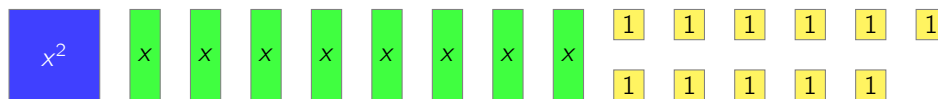
Although the numbers worked out nicely in this case, finding x -intercepts can require some clever thinking, and some new algebraic tools. It will be a topic we investigate more closely in chapter 16.

14.3.2 Standard to vertex form: Completing the square

Converting from vertex form to standard form is easy: we simply expand the squared expression using the distributive property, and then combine like terms. Converting in the other direction is harder, although we have already seen a technique that is helpful: the quadrangle method.

Note that an equation in vertex form has a square inside it: the $(x - h)^2$ term is a square with sides of length $(x - h)$. Can we manipulate a standard-form quadratic to find a hidden square?

Let's take a specific example: consider the equation $y = x^2 + 8x + 11$. Does this expression make a square? One way to visualize this is to imagine our trusty algebra tiles. The given quadratic expression in this case can be represented by the algebra tiles:



If we try to arrange these tiles into a square, we're a few tiles short. What do we need to *complete the square*?

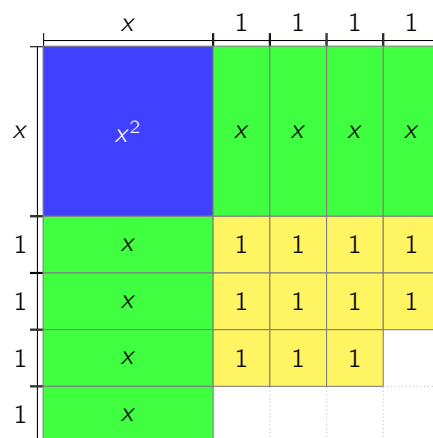
If only we had 5 more 1-unit tiles in our diagram! Well, we can accomplish this by changing our original equation using the POEs, just like we did in the quadrangle method:

$$y = x^2 + 8x + 11 \quad \text{original equation}$$

$$y + 5 = x^2 + 8x + 16 \quad \text{APOE: add 5 to both sides}$$

$$y + 5 = (x + 4)^2 \quad \text{based on the quadrangle diagram}$$

$$y = (x + 4)^2 - 5 \quad \text{SPOE, to isolate } y$$



In the end, we have converted our standard form equation into an equivalent vertex form equation! This method is called **completing the square**, a name inspired by the goal of creating a complete square (quadrangle) diagram.

Example 14.4

Convert the following equation to vertex form: $y = 2x^2 + 12x + 21$.

Solution: If our goal is to have a square, remember, the quadratic term must be a perfect square, and the linear term must be easily divided into two pieces. In this case, our quadratic term needs some adjusting. We'll multiply through by 2, giving us:

$$2y = 2x^2 + 12x + 21$$

Some standard-form quadratics do not yield so easily to the process of completing the square. So we will return to this method and learn more about it in chapter 16.

14.3.3 A third option: Factored form

In addition to vertex form and standard form, there is a third form of a quadratic function called *factored form*. Each of these forms tells us something about the graph.

Vertex form displays the vertex very clearly. Standard form displays the y -intercept very clearly. Factored form displays the x -intercepts of the parabolas very clearly... but, as we saw above, not every parabola has x -intercepts. So, not every quadratic function has a factored form.

Factored form

An equation of the form

$$y = a(x - p)(x - q),$$

where $a \neq 0$, is a quadratic equation in factored form. The points $(p, 0)$ and $(q, 0)$ are the x -intercepts of the related quadratic equation. The x -intercepts are sometimes also called the *roots* or the *zeroes* of the function.

This is the third topic in a row that we've just briefly mentioned. We'll get much deeper into the details of factored form, x -intercepts, and completing the square in chapter 16.

14.4 Application: Projectile motion

Projectile motion is a classic application of quadratic functions, and we'll study both the mathematics and the science in this section.

Rocket Launch

// startup exploration

When an object is launched vertically, it flies straight up and then falls straight down. Height data for the flight of one of Hermine's model rockets is given below.

Time (sec)	Height (feet)
0	0
1	224
2	384
3	480
4	512
5	480
6	384
7	224
8	0

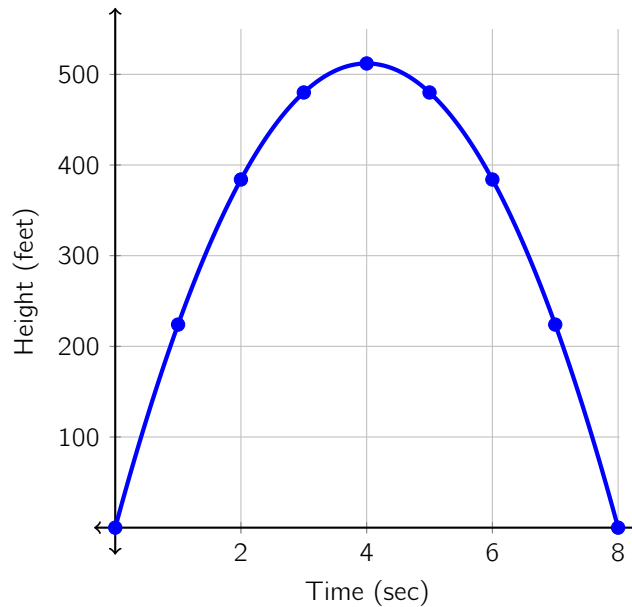
Study the table and make a graph of the data to see what you can learn. Then, write a few sentences describing the flight of Hermine's rocket in as much detail as you can.

From the data itself we can see the up-and-down pattern. Plus, we can see a nice symmetry: The object starts at a height of 0 meters (so it starts on the ground) and lands on the ground 8 seconds later. One second after launch it is at the same height as 1 second before it lands. This mirror image pattern continues until the object reaches its maximum height 4 seconds after launch.

A graph of the data shows the same up-and-down pattern, and the same symmetry: we could fold this graph along the vertical line at $x = 4$ and the two halves would match up exactly.

These features all suggest that we may have a quadratic pattern.³ Scientifically speaking, Hermine's rocket is a *projectile*, and quadratic functions are the underpinning of the physics that describes the trajectory of a projectile.

³ The title of this chapter might also have been a hint.



Projectile

Any object that moves through the air or through space acted on only by the force of gravity.

We use projectiles in games: throwing a baseball, basketball, or dart. We use projectiles as weapons: shooting an arrow or a cannonball, throwing a spear, or launching something from a catapult.

Not all flying things are projectiles. Airplanes, helicopters, and birds use energy to keep themselves in the air, and they can use that energy to steer and change direction. Projectiles are not “powered flight”, they are launched and the allowed to follow their path, acted on only by gravity.

Trajectory

An imaginary tracing of a projectile’s position as it moves through space.

Hermine’s rocket was launched, but we can use the term “launch” rather loosely. An item that has been dropped is, according to our definition, a projectile. This is the first kind of motion we’ll investigate closely, beginning with a discussion about gravity.

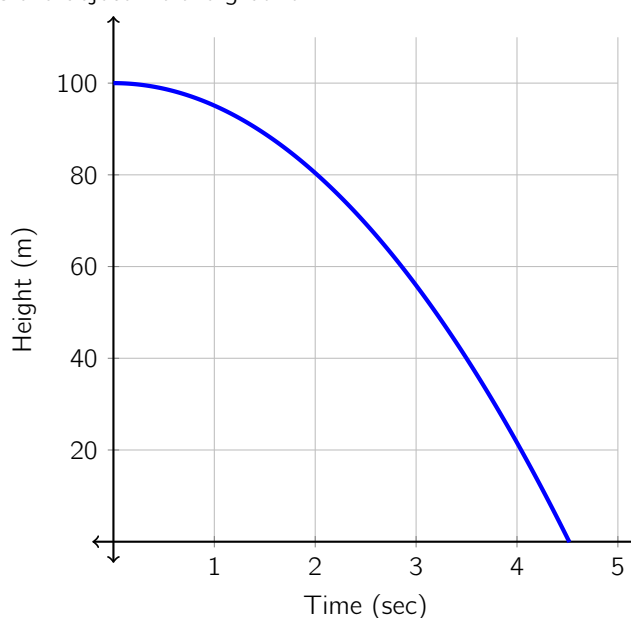
14.4.1 Falling Objects

An object that has been dropped and falls through space is one example of a projectile. Gravity causes the object to *accelerate* towards the ground. In other words, falling objects *increase in velocity* (speed) as time goes by.

The acceleration of gravity is -9.8 meters per second per second, or -32 feet per second per second. No, that's not a typo: we're talking about meters per second *per second*. Acceleration is a change in speed. Speed might be measured in "meters per second", and so if an object is accelerating its speed might be changing at a rate of 9.8 meters per second, *per second*. We sometimes abbreviate "meters per second per second" as m/s^2 .

The acceleration of gravity is a negative quantity because acceleration has both a magnitude and a direction, and that direction is downwards (toward the center of the Earth). Gravity will slow down an object that is thrown upwards, and pulls dropped objects towards the ground.

This graph shows the height of an object that has been dropped over time. Notice how the curve of the graph gets steeper the longer it is falling. What can you discern from the graph? How high up was the object when it was dropped? When does the object hit the ground?



A quick comment about a common misconception regarding this graph. It may look at though this graph represents an object that was thrown outwards from the top of a cliff. In other words, the graph might appear to be the path of an object flying through the air both horizontally and vertically. But that is not what this graph represents! What we have here is a graph of height over time, meaning vertical motion only.⁴

The motion of a falling object can be described using a quadratic equation.

⁴ It is possible to draw the kind of graph that shows an object's movement in both the horizontal and vertical dimensions simultaneously. We can define so-called *parametric functions* to draw such graphs and describe such motion... but that must wait for a different time. Just be aware of what these graphs are showing, and what they are not.

Falling objects

The height, as a function of time, of an object falling under the force of gravity is given by

$$h(t) = \frac{1}{2}gt^2 + h_0,$$

where g is the acceleration due to gravity and h_0 is the initial height of the object (the height from which it was dropped)

Note that we have two choices for the g in our equation, depending on whether we are working in feet or in meters.

Example 14.5

Mr. Campbell, middle school science teacher and the world's leading expert on the aerodynamics of Spam, drops a can of Spam off a bridge that is 80 feet high. When will the Spam hit the ground?

Solution: "The Spam hits the ground" is the same as saying "the height of the can of Spam is 0 feet".^a So, we want to know at what time t does $h(t) = 0$?

We can plug in some values, using $-32 \frac{\text{ft}}{\text{sec}^2}$ as the acceleration of gravity, and using $h_0 = 80$ as the initial height of the Spam.

$$h(t) = \frac{1}{2}(-32)t^2 + 80$$

$$0 = \frac{1}{2}(-32)t^2 + 80 \quad \text{let } h(t) = 0 \text{ to find when the can is on the ground}$$

$$0 = -16t^2 + 80$$

$$-80 = -16t^2 \quad \text{SPOE}$$

$$5 = t^2 \quad \text{DPOE}$$

When we take the square root of both sides, we get $t = \sqrt{5}$ or $-\sqrt{5}$.

We can discard the negative value as a possibility, since "negative time" values don't make any sense in this context. So, we have $t = \sqrt{5}$. In this case, a decimal approximation is appropriate, so we can say the can hits the ground approximately 2.24 seconds after being dropped.

^a OK, maybe these aren't exactly the same. A can of Spam resting on the ground would be a much less spectacular sight than a can of Spam hitting the ground after being dropped from a bridge... But they are the same *mathematically speaking*.

When we solve a quadratic equation like this, we often get both a positive and a negative value. But the negative solution usually doesn't make sense in the context of the problem. In the previous example, that would

be a point in time *before* the Spam was dropped.

14.4.2 Horizontal Launch

Dropping objects is boring! Let's throw something! But, we're going to add a constraint: we will throw things horizontally, parallel to the ground.

Objects that are thrown or launched horizontally travel in two dimensions: both horizontally and vertically. However, the vertical and horizontal motion are independent. Gravity pulls the object downwards, but has no influence on the projectile's horizontal movement.

To describe the vertical part of the projectile's motion, we use the equation for a falling object. To describe the horizontal motion, use the usual equation $d(t) = r \cdot t$, where r is the object's initial horizontal velocity. Notice how we have separated the two aspects of the projectile's motion: the vertical component is just like dropping, the horizontal component is just like movement along the ground.

In other words, if we need to find out how long it will take a horizontally launched object to hit the ground, we treat it just like a falling object. The only thing that is different is that a horizontally launched object will land further away from the launch site than an object that is dropped.

Example 14.6

Mr. Campbell fires a can of Spam horizontally from the 120-meter high cliff using his Spam Cannon (patent pending). The Spam has an initial horizontal velocity of 68 meters per second. When will the Spam hit the ground? How far from the base of the cliff will it land?

Solution: The vertical motion of the Spam is governed by our equation from earlier. So, we can calculate the time it takes to hit the ground in the same way as we did above. (Note, however, that we're using meters in this example!)

$$\begin{aligned}
 h(t) &= \frac{1}{2}(-9.8)t^2 + 120 \\
 0 &= \frac{1}{2}(-9.8)t^2 + 120 && \text{let } h(t) = 0 \text{ to find when the can is on the ground} \\
 0 &= -4.9t^2 + 120 \\
 -120 &= -4.9t^2 && \text{SPOE} \\
 24.49 &\approx t^2
 \end{aligned}$$

So, $t \approx 4.95$ seconds (again, we discard the negative square root).

So we know when the can hits the ground. But where does it land? The can's horizontal motion is completely separate from its vertical motion. It flies horizontally at a speed of 68 meters per second for

4.95 seconds (the time until it lands).

$$d \approx 68(4.95) \approx 336.6 \text{ m}$$

14.4.3 Vertical Launch

Objects that are thrown or launched vertically travel in only one dimension. The rising and falling components of their motion are symmetrical. We saw an example of this vertical motion at the start of this chapter with Hermine's model rocket.

This equation is just like the falling object equation, but it has a linear " $v_0 t$ " added to it in which v_0 is the initial vertical velocity.

Vertical motion

The height, as a function of time, of an object launched vertically and travelling under the force of gravity is given by

$$h(t) = \frac{1}{2}gt^2 + v_0 t + h_0,$$

where g is the acceleration due to gravity, v_0 is the initial vertical velocity, and h_0 is the initial height of the object.

Note that when an object is dropped, the initial velocity is zero. So the falling object equation is, in fact, a specific case of this equation in which $v_0 = 0$. Note, too, that if an object is launched from the ground, then $h_0 = 0$.

Example 14.7

Mr. Campbell fires a can of Spam vertically from the ground with an initial velocity of 96 ft/s. When will the Spam reach its maximum height? What is the maximum height of the can of Spam?

Solution: To find the time at which the Spam reaches its maximum height, think about the shape of the graph. It is a parabola, and so the maximum height will be the vertex. The time at which this occurs is the x-coordinate of the vertex, which is given by the line of symmetry.

The equation which models this situation is:

$$h(t) = \frac{1}{2}(-32)t^2 + 96t + 0 = -16t^2 + 96t,$$

and the line of symmetry for this parabola is at:

$$x = \frac{-b}{a2} = \frac{-96}{2(-16)} = 3.$$

So, the can of Spam reaches its maximum height after 3 seconds. Due to the symmetry of the parabola, we now also know that the Spam lands 3 second later (after flying for a total of 6 seconds).

To find the maximum height, we want to know the height after 3 seconds, $h(3)$.

$$h(3) = -16(3)^2 + 96(3) = 144$$

So the maximum height of the Spam is 144 ft.

As a final example, consider the somewhat more complicated scenario below.

Example 14.8

Mr. Campbell fires a can of Spam vertically from the top of a 240-foot tower with initial velocity of 128 ft/s. As it falls, the Spam drifts slightly, misses the launch site, and falls all the way to the base of the tower. What is its total flight time?

Solution: We can model this using the same equation as above. In this case, we have:

$$h(t) = \frac{1}{2}(-32)t^2 + 128t + 240 = -16t^2 + 128t + 240.$$

Since we want to know what the Spam reaches the ground, we want to find t such that $h(t) = 0$. So, we're out to solve the equation:

$$-16t^2 + 128t + 240 = 0$$

This looks like a job for the quadrangle method!

Our quadratic term must be a perfect square, and it's nearly there. We should first multiply through by -1 , giving: $16t^2 - 128t - 240 = 0$. From here, we can make the quadrangle diagram.

	$4t$	-16
$4t$	$16t^2$	$-64t$
-16	$-64t$	256

These numbers aren't pretty, but we won't let that stop us.

$$16t^2 - 128t - 240 = 0$$

$$16t^2 - 128t + 256 = 496$$

$$(4t - 16)^2 = 496$$

$$4t - 16 = \sqrt{496} \text{ or } -\sqrt{496}$$

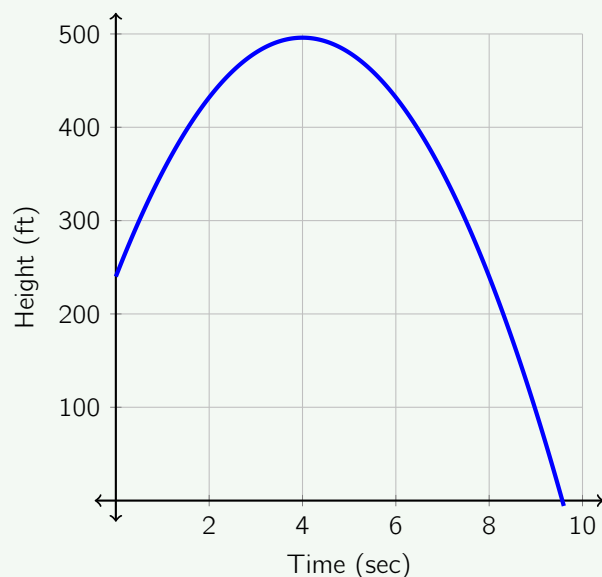
APOE: add 496 to both sides

based on the quadrangle diagram

Yuck. Perhaps it's time to resort to some decimal approximations: $\sqrt{496} \approx 22.27$. After applying APOE and DPOE, and discarding the negative solution, we have

$$t \approx 9.57 \text{ seconds.}$$

A graph helps us to see that our answer is reasonable.



On launching at an angle

You may be wondering about why we haven't discussed launching a projectile at an angle, the way you might do if you were throwing a ball to a partner in a game of catch.

In this scenario, the initial velocity of the projectile is at an angle, and that means it has both a horizontal and a vertical velocity at the same time. The challenge, then, is to somehow resolve the single “diagonal” velocity into a vertical component (for the vertical motion piece) and a horizontal component (for the horizontal motion piece). This requires a bit more mathematics than we’re ready to get into now (trigonometry, to be precise). Rest assured, though, that we will get there eventually!

» Chapter summary «

In this chapter we saw that, like linear functions, quadratic functions can be written in more than one form. We worked with both vertex form and standard form in this chapter, and later we will study a third form (called intercept form).

Each of these forms has certain advantages and highlights certain information about the parabola. The vertex form highlights the vertex (the minimum or maximum point), while the standard form highlights the y -intercept. Regardless of the form we are given, we learned how to create a graph by finding key points, and taking good advantage of the fact that quadratic graphs are symmetrical.

Projectile motion can be described using quadratic functions, and so this mathematical idea is of key importance in the field of physics. We studied an equation that can be used to model the height of an object that has been thrown or dropped as a function of time.

We have some unanswered questions, though: for example, regarding the factored form of the quadratic. The next two chapters will help us tie up these loose ends. Onward!

It is India that gave us the ingenious method of expressing all numbers by means of ten symbols, each symbol receiving a value of position as well as an absolute value; a profound and important idea which appears so simple to us now that we ignore its true merit.

Pierre-Simon Laplace, French mathematician

Chapter 15

Polynomials

Don't be scared by the complicated-sounding name: we have actually been working with polynomials throughout the course! In this chapter, we will put some new names on some old concepts, and extend those ideas into a more general form. Then we'll start with the foundations and learn the rules for performing the standard operations with polynomials: adding, subtracting, multiplying, dividing, factoring, and raising to a power.

15.1 Building Polynomials

What qualifies?

// startup exploration

Before reading the official definition of a polynomial, consider the two lists below.

Polynomials	Not Polynomials
$2x + 4$	$2^x + 4$
$3x$	$\sqrt{3x}$
$1.2y^2$	$1.2y^{-2}$
$5x^2 + 3x - 7$	$ 5x + 3 - 7$
$-2w^3 - 8w^2 + w$	$-2w^3 - 8x^2 + y$
$\frac{1}{4}x + 15$	$\frac{4}{x} + 15$

Based on this information, conjecture about what it means to be a polynomial. Which of the following (if any) are polynomials?

$6x - 8 + 7x^{-1}$

150

$x^2 + x^4 - 3x^3 + 8 - x$

Recall that a **term** is a number, a variable, or the product (or quotient) of numbers and variables. The following expressions are all terms:

$$4 \quad 3x \quad 14x^6$$

We have been dealing with terms throughout this course. We know how to *combine like terms*. When we perform the distributive property, we are multiplying one term with each of several terms.

Remember that terms are separated by addition and subtraction, and that an **algebraic expression** is the sum or difference of terms. An expression can be a single term all by itself: since x and 0 are both terms, $x + 0$ (which is the same as x) is an expression.

A **polynomial** is a special type of algebraic expression.

Polynomial

A algebraic expression which contains only a single variable, in which the coefficients are all rational numbers, and the degree of each term is a non-negative integer.

Recall that the **degree of a term** is the exponent on the variable of a term. It is usually quite easy to identify the degree of a term: the degree of $5x^3$ is 3. Remember, though, that some special situations arise.

If we have a variable with no exponent, we imagine a phantom 1 as the exponent. So for example, the term $18x$ is of degree 1, since $18x = 18x^1$. If we have a number with no variable, we can imagine a phantom 1 of a different kind: a variable raised to the zero power. So, the term 10 is of degree 0, since $10 = 10x^0$.

This means that a polynomial can always be written in the form

$$a_0x^0 + a_1x^1 + a_2x^2 + a_3x^3 + \cdots + a_nx^n,$$

where this sum stops at some positive integer n , and in which the a -values (the coefficients) are rational numbers, some of which might be zero. Usually, we prefer to write a polynomial with the terms in order of decreasing degree, that is:

$$a_nx^n + \cdots + a_3x^3 + a_2x^2 + a_1x^1 + a_0x^0$$

Not all algebraic expressions are polynomials. A polynomial may not have a variable in the exponent, so exponential functions are not polynomials. A polynomial may not include a variable raised to a non-integer exponent (so, no variables under radicals), nor a negative exponent (so, variables in a denominator). Also, the variable may not be enclosed in an absolute value expression.

Look back over the list of polynomials and non-polynomials in the startup exploration. The expressions in the list of polynomials all meet these requirements. The non-polynomials all fail on at least one point (can you

identify the point of failure for each one?). Of the three expressions we are asked to classify, the first one is not a polynomial, since it includes the variable x raised to a negative exponent.

The other two expressions are polynomials. The number 150 is a polynomial since it is of the form $150x^0$. The third expression given meets the definition of a polynomial, and can be rewritten as follows:

$$x^4 - 3x^3 + x^2 - x + 8$$

Degree of a Polynomial

The degree of the highest-degree term in the polynomial.

The degree of a polynomial is the highest exponent in the polynomial. This is very important because the highest degree has the biggest impact on the value of the expression. If we were to put a “ $y =$ ” in front of the polynomial to turn it into a function, the highest degree term will tell us which family it belongs in. For example, every member of the quadratic family is a polynomial of degree 2.

Standard form of a polynomial

A polynomial is written in standard form if the degree of the terms decreases from left to right, and no two terms have the same degree.

This is probably common sense by this point in the course. Basically, standard form is simplified (like terms are combined and there are no parentheses), and in addition, we make be sure to write the terms in a “reasonable” order. The highest degree term comes first, then the rest follow in order of decreasing degree. A polynomial is still a polynomial otherwise, but this format makes polynomials easier to understand and communicate.

15.1.1 Naming polynomials

In the life sciences, biological classification is the practice of grouping organisms into species, and then grouping various species together into groups, and then grouping those groups together, and so on, to produce a systematic classification of living things. Such a classification is important so that biologists can then be clear when communicating to one another.

For example, there are many different types of fruit fly, and so the term “fruit fly” is ambiguous. In English, we make a distinction between the so-called “common fruit fly” and the “Asian fruit fly”. Biologists in Korea, however, might not use these terms in the same way (is the Asian fruit fly more common in Korea than the common fruit fly?). Biologists have therefore agreed to use a different, more universal, system for describing

organisms so that they can distinguish between *Drosophila melanogaster* (what we call the common fruit fly) and *Drosophila suzukii* (what we call the Asian fruit fly).

The Latinized names for organisms identify the genus and species. In mathematics, we classify polynomials using Latin and Greek names that identify the degree and the number of terms.

The first name is by degree, and this is related to the name you would give to that family of functions. We worked closely with two examples in this course so far. A polynomial of degree 1 is called a *linear* polynomial, and a polynomial of degree 2 is called a *quadratic* polynomial.

Degree	Name
0	Constant
1	Linear
2	Quadratic
3	Cubic
4	Quartic
5	Quintic

Table 15.1: List of polynomial names by degree.

For polynomials with degree higher than 5, we usually say “polynomial of degree n ”, or “ n th degree polynomial”. For example $3x^8 + 4$ is an eighth degree polynomial, or a polynomial of degree 8.

When naming a polynomial, the second name tells us how many terms it has.

No. Terms	Name
1	Monomial
2	Binomial
3	Trinomial

Table 15.2: List of polynomial names by number of terms.

For polynomials with more than three terms, we say “polynomial with n terms” or “an n -term polynomial”. For example,

$$11x^8 + x^5 + x^4 - 3x^3 + 5x^2 - 3$$

is a polynomial with 6 terms, or a 6-term polynomial.

Note that some of these prefixes are familiar. *Tri-* means “three” (as in triangle and tricycle), and *bi-* means “two” (as in bicycle or bisexual). There are less familiar word parts, too: *Mono-* means one (a monopoly is when a single company is the only one offering a particular product or service). The suffix *-nomial* means “names” (this is the same root as in the words nominate and denominator).

When giving the full name of a polynomial, we include both the name by degree and the name by the number of terms. Here are some examples:

Polynomial	Name
$14x^3$	cubic monomial
$-1.2x^2$	quadratic monomial
18	constant monomial
$7x - 2$	linear binomial
$2x^2 - 4x + 8$	quadratic trinomial
$3x^3 + 2x - 11$	cubic trinomial
$x^4 + 3$	quartic binomial

Table 15.3: Examples of polynomial names.

15.2 Adding, subtracting, and multiplying polynomials

The first few operations with polynomials are related to things we have seen before.

15.2.1 Polynomial addition and subtraction

To add or subtract polynomials, we simply combine like terms – nothing new here! The only change is that we might encounter more terms than before, and perhaps terms of higher degree.

Example 15.1

Add: $(3x^3 + x^2 + 6x - 8) + (4x^2 + 2x - 5)$

Solution: To add polynomials, we simply combine like terms. Recall that to be like terms, two terms must have identical variable parts. So, in this case, the sum of these two polynomials is:

$$3x^3 + 5x^2 + 8x - 13$$

Be sure to read carefully. A common mistake in the above example is to combine the first term of each polynomial – but that's no good, since we are adding polynomials of different degree!

Example 15.2

Subtract: $(7x^3 - 3x + 1) - (x^3 - 4x^2 - 2)$

Solution: When subtracting we'll have to distribute the negative, so remember to check the signs. In this case, we'll have:

$$\begin{aligned}(7x^3 - 3x + 1) - (x^3 - 4x^2 - 2) &= (7x^3 - 3x + 1) + (-x^3 + 4x^2 + 2) \\ &= (7x^3 - 3x + 1) + (-x^3 + 4x^2 + 2) \\ &= 6x^3 + 4x^2 - 3x + 3\end{aligned}$$

The key with subtraction is to remember that every term of that second polynomial is being subtracted.

When showing your work, it is sometimes easier to keep track of everything if we set it up like “old school” addition or subtraction problem. To do this, we think of the degree of the term like the place value of a digit. If there is a degree missing, we can insert a term with coefficient 0. Here's how example 15.1 would look:

$$\begin{array}{r}
 3x^3 + x^2 + 6x - 8 \\
 + 0x^3 + 4x^2 + 2x - 5 \\
 \hline
 3x^3 + 5x^2 + 8x - 13
 \end{array}$$

Consider putting the subtraction in parenthesis as a reminder that every term needs to be subtracted. Here's example 15.2 presented in this format:

$$\begin{array}{r}
 7x^3 + 0x^2 - 3x + 1 \\
 - (x^3 - 4x^2 + 0x - 2) \\
 \hline
 6x^3 + 4x^2 - 3x + 3
 \end{array}$$

15.2.2 Polynomial multiplication

We have already studied the key ideas of polynomial multiplication, back in section 11.5 when we learned how to raise a sum to a power. The key idea is to use the distributive property.

For example, given the task of multiplying the linear binomials $(x + 1)(x + 2)$, we treat the $(x + 1)$ as a single entity and distribute that over the two terms of $(x + 2)$.

$$\begin{array}{ll}
 (x + 1)(x + 2) = (x + 1)x + (x + 1)2 & \text{distribute the } (x + 1) \\
 = (x + 1)x + (x + 1)2 & \text{identify two new distribution opportunities} \\
 = x \cdot x + 1 \cdot x + x \cdot 2 + 1 \cdot 2 & \text{distribute twice more} \\
 = x^2 + x + 2x + 2 & \text{simplify} \\
 = x^2 + 3x + 2 & \text{combine like terms}
 \end{array}$$

We can also set things up using an old school style, as we saw back then, and this representation makes the connections more clear. Consider the two multiplication problems below: one with four-digit numbers, and one with four-term polynomials.

$$\begin{array}{r}
 5 \ 6 \ 8 \ 2 \\
 \times 3 \ 9 \ 4 \ 7 \\
 \hline
 \end{array}
 \quad \text{and} \quad
 \begin{array}{r}
 5x^3 + 6x^2 + 8x + 2 \\
 \times 3x^3 + 9x^2 + 4x + 7 \\
 \hline
 \end{array}$$

The process of multiplication is the same in these two cases: in fact, if we substitute $x = 10$, we have *exactly the same problem*, since we can rewrite a number in expanded form using its place value, for example:

$$\begin{aligned}
 5682 &= 5000 + 600 + 80 + 2 \\
 &= 5 \cdot 1000 + 6 \cdot 100 + 8 \cdot 10 + 2 \\
 &= 5 \cdot 10^3 + 6 \cdot 10^2 + 8 \cdot 10^1 + 2 \cdot 10^0
 \end{aligned}$$

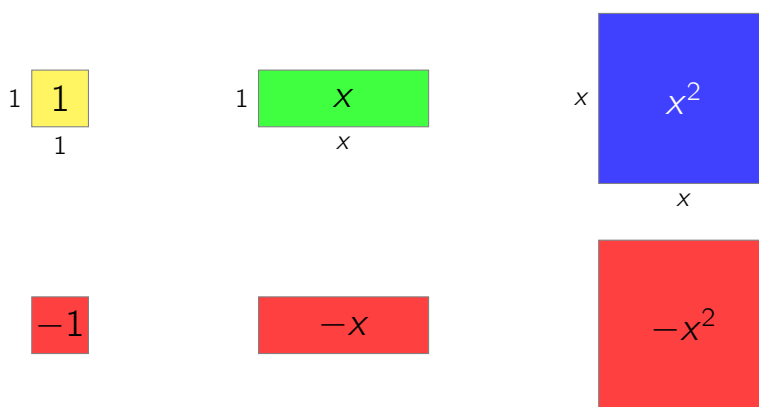
This last line is the same as the polynomial $5x^3 + 6x^2 + 8x + 2$, evaluated at $x = 10$!

In any case, each term in one polynomial will be multiplied by each term in the other polynomial. It's a pretty tedious exercise to multiply two four-digit numbers by hand, and the same goes for multiplying two cubic polynomials. So, instead, let's look at something more interesting.

15.2.3 Return of the algebra tiles

In section 11.5 we saw a visual representation called algebra tiles. Algebra tiles can be used to represent polynomials, although they are limited to polynomials of degree 0, 1, or 2.

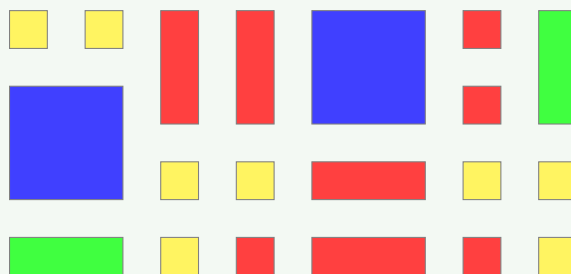
Recall that the basic set of tiles includes three colored tiles that represent 1, x , and x^2 . To represent negative numbers, we also have tiles of the same size, all colored red.



We can use the tiles to show a picture of a polynomial, and to show how addition and subtraction work, but they are really helpful for showing the key ideas around multiplication. Let's start with the basics.

Example 15.3

What polynomial is represented below?



Solution: A red tile will cancel out a non-red tile of the same size, since the two add up to zero. After

removing all of the “zero pairs”, we just count up what’s left over. In the end, we’ll have:

$$2x^2 - 2x + 4.$$

If we have two piles of tiles at the start, we can scrape them into one big pile, remove all the zero pairs, and then count up what we have. This is polynomial addition! To subtract polynomials using tiles, we would negate all the tiles in the polynomial that’s being subtracted (exchanging red tiles for their non-red versions, and vice versa), and then add. Nothing to it!

Representing multiplication with algebra tiles is quite handy. The length and width are the factors and the product is the area. For instance, if we want to multiply $(x + 1)(x + 2)$, we set up a rectangle that has width $x + 1$ and length $x + 2$. We fill in the rectangle with algebra tiles and then interpret the result.



Figure 15.1: Multiplication using algebra tiles: $(x + 1)(x + 2) = x^2 + 3x + 2$

So, we have the product $(x + 1)(x + 2) = x^2 + 3x + 2$.

It may seem silly to manipulate these rectangular pictures one we know how to use the distributive property, but recall how helpful the square diagrams were in solving quadratic equations. Rectangular pictures will help us to visualize the “undoing” of multiplication in a section or two. Never underestimate the power of a picture!

An algebraic F-word

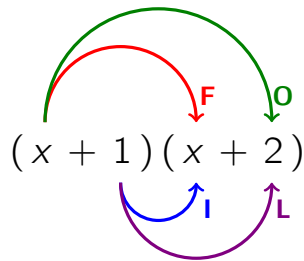
At this point, many algebra textbooks introduce the term *FOIL*.¹ The authors of the *Algebramicon* are not fans of FOIL, and we mention it at this point only to dispell the misconception that we don’t know what it is.

FOIL is an acronym that stands for “First, Outer, Inner, Last”. It is meant to be a mnemonic device for remembering how to multiply two binomials.

The problem with FOIL is that it only works with *two binomials*. It’s a special-case formula! If we’re going to learn a technique for multiplication, we should focus on a technique that works for everything: the distributive property always works. FOIL does not.

¹ This term is also popular among online tutorial videos, math-help websites, private tutors, and parents.

F	First	Multiply the first terms of the binomials
O	Outer	Multiply the outer terms of the binomials
I	Inner	Multiply the inner terms of the binomials
L	Last	Multiply the last terms of the binomials



15.3 Dividing polynomials

Polynomials represent real numbers, and since we can divide real numbers, we can divide polynomials. We won't focus on polynomial division here – it will be a topic of more discussion in Algebra 2. We discuss the idea now simply to “complete the set” and show how all four of the basic operations work with polynomials.

Polynomial division works just like long division of numbers. Recall from your elementary school days (or from section 13.1) that to do the division $643 \div 12$ we proceed like so:

$$\begin{array}{r}
 53 \\
 12 \overline{) 643} \\
 \underline{- 60} \\
 43 \\
 \underline{- 36} \\
 7
 \end{array}$$

In the end, we can make 53 whole groups of 12 and we are left with a remainder of 7. In other words: we can't make another group of 12 because we have only 7 left over. We can express this as a fraction since we have $\frac{7}{12}$ of another group. So we have:

$$643 \div 12 = \frac{643}{12} = 53 \frac{7}{12}$$

Generally speaking, we have:

$$\frac{\text{dividend}}{\text{divisor}} = \text{quotient} + \frac{\text{remainder}}{\text{divisor}}$$

Of course, if the remainder is zero then it means that the divisor is a factor of the dividend. Both of these situations can arise when dividing polynomials. Later, in Algebra 2, we will return to polynomial division and discover what we can learn from these remainders when they exist.

We find ourselves in a particularly powerful and important situation, though, when one polynomial is a factor of the other. This is the focus of chapter 16, and we will spend quite a bit of time learning the essential skills around *factoring* polynomials.

Before we get into that special situation, let's look at some examples of polynomial division. Suppose we wish to divide $x^3 - 5x^2 + 4x - 2$ by $x - 2$.

First, we have to set up the division. We should check and make sure the polynomials completely simplified and that their terms are arranged by decreasing degree. Then, we set up the division:

$$x - 2 \overline{) x^3 - 5x^2 + 4x - 2}$$

To divide we look at the first term in the divisor and ask: "What do we have to multiply that term by to get the first term in the dividend?" In this case, we have to multiply x (the first term in the divisor) by x^2 to get x^3 (the first term in the dividend). So, x^2 becomes the first term in the quotient.

$$\begin{array}{r} x^2 \\ x-2 \overline{) x^3 - 5x^2 + 4x - 2} \end{array}$$

Then, we follow the algorithm of long division. Multiply the entire divisor by x^2 , line up these terms with their counterparts under the dividend and subtract. Watch out for those negative signs!

$$\begin{array}{r} x^2 \\ x-2 \overline{) x^3 - 5x^2 + 4x - 2} \\ \underline{-(x^3 - 2x^2)} \\ -3x^2 \end{array}$$

Bring down the extra terms, and we have a new polynomial whose degree is less than the original. So, we start again with the division step and determine what we have to multiply the divisor by so that we match the leading term in the new polynomial. We can stop when the degree of the remainder is less than the degree of the divisor. Here's the full process in this case:

$$\begin{array}{r} x^2 - 3x - 2 \\ x-2 \overline{) x^3 - 5x^2 + 4x - 2} \\ \underline{-(x^3 - 2x^2)} \\ -3x^2 + 4x - 2 \\ \underline{-(-3x^2 + 6x)} \\ -2x - 2 \\ \underline{-(-2x + 4)} \\ -6 \end{array}$$

Therefore the quotient is $x^2 - 3x - 2$ and the remainder is -6 . Written in its final form, we have:

$$\frac{x^3 - 5x^2 + 4x - 2}{x - 2} = x^2 - 3x - 2 + \frac{-6}{x - 2}$$

To check our solution, we can multiply the quotient by the divisor, then add the remainder. That should give us the dividend back again!

That's a lot of alphabet soup and, as we mentioned above, we'll come back to generic polynomial division in Algebra 2. For the remainder of our time in Algebra 1, we will concern ourselves with the important situation in which one polynomial is a factor of another. Consider this final example as we make the transition.

Solution: In this case, we find that there are terms “missing” from the dividend. So, we must add in a linear term with a zero coefficient: $0x$. Then, we can go about long division as usual:

No remainder means that we have found two polynomials which multiply together to produce the other:

Check our solution by carrying out the multiplication of these two linear binomials. What happens?

In this chapter we learned the key ideas around describing and operating with polynomials. Much of this was old news: Addition and subtraction of polynomials is simply an application of combining like terms. Multiplication of polynomials, though it takes some careful bookkeeping, is just an application of the distributive property.

We also discussed a bit about polynomial division, but only a bit. Generic polynomial division will return in Algebra 2. In the meantime, we will look at the important – and surprisingly useful – situation in which one polynomial is a factor of another. Onward!

Ten less a thing, multiplied by itself, is a hundred plus a square less twenty things.

Muḥammad ibn Mūsā al-Khwārizmī, Persian mathematician

Chapter 16

Factoring

When we *factor* a number, we re-write it in a way that reveals its multiplicative structure. For example, we can factor 78 and rewrite it as $2 \cdot 39$ or as $2 \cdot 3 \cdot 13$. Indeed, you have probably been finding the *prime factorization* of numbers for a while now.

Factoring a polynomial has the same goal: to rewrite the polynomial as an equivalent multiplication problem. We will learn several important techniques for factoring, but we begin with the end in mind: an example of how factoring is helpful for solving equations.

16.1 Quadratic equations, revisited

Solutions and factors

// startup exploration

Solve the following quadratic equation using the quadrangle method that we studied in chapter 12:

$$x^2 + 3x + 2 = 0$$

Then, look back at section 15.2.3 (figure?), which shows a way of writing this quadratic trinomial as the product of two binomials.

$$(x + 1)(x + 2) = x^2 + 3x + 2$$

What relationships, if any, are there between these factors and the solutions to the given equation?

We solve the quadratic $x^2 + 3x + 2 = 0$ using the quadrangle method, beginning with the observation that the linear term has an odd coefficient. So, we multiply through by 4: $4x^2 + 12x + 8 = 0$. From here, we can draw the quadrangle diagram:

	2x	3
2x	4x ²	6x
3	6x	9

The quadrangle diagram predicts a constant term of 9, but we only have 8. So, we add 1 to both sides. Here's the full work:

$$\begin{array}{ll}
 x^2 + 3x + 2 = 0 & \\
 4x^2 + 12x + 8 = 0 & \text{multiply through by 4} \\
 4x^2 + 12x + 9 = 1 & \text{add 1 to both sides} \\
 (2x + 3)^2 = 1 & \text{based on the quadrangle diagram} \\
 2x + 3 = 1 \text{ or } -1 & \text{square root of both sides} \\
 2x = -2 \text{ or } -4 & \text{subtract 3 throughout} \\
 x = -1 \text{ or } -2 & \text{divide by 2 throughout}
 \end{array}$$

So, we have solutions -1 and -2 . Now consider the factorization $(x + 1)(x + 2)$. Isn't it interesting that the constants in the two factors are the opposites of the two solutions? It turns out that this connection is not a coincidence.

16.1.1 Zero product property

To explain the connection, we will turn to one of the most obvious statements in elementary mathematics. It is a statement so simple that you might be surprised that it could ever serve any truly helpful purpose.

Zero product property

If $a \cdot b = 0$, then either $a = 0$ or $b = 0$ (or both). In other words: if the product of two numbers is 0, then at least one of those numbers must be zero.

Ridiculously simple, right? But watch it in action! Consider again the original quadratic equation we were given in the startup exploration:

$$x^2 + 3x + 2 = 0.$$

We have seen that we can factor this quadratic expression into the product of two binomials. In other words, we may write:

$$(x + 1)(x + 2) = 0.$$

Here we have the product of two expressions, and that product equals 0. The zero product property tells us that one or the other of those factors must be equal to zero! In other words, either $(x + 1) = 0$ or $(x + 2) = 0$.

But if $x + 1 = 0$, then $x = -1$. So -1 is a solution to the equation. On the other hand, if $x + 2 = 0$, then $x = -2$. So -2 is also a solution to the equation! We just found the two solutions to our quadratic without the quadrangle method!

Behold the power of factoring! When tasked with solving a quadratic equation (or any polynomial equation, really), we can sometimes attack it using factoring and the zero product property. This approach doesn't always work, and in those cases we must rely on other techniques (like the quadrangle method). But, factoring will give us an elegant way to solve equations, and teach us other things as well, for example about the graphs of polynomial functions.

In the next several sections we will progress through the levels of factoring challenge, revisit the quadrangle diagram in a rectangular format, and fill our toolbox with an impressive collection of tools and techniques.

16.2 Factoring level 1: Greatest common factor

The greatest common factor, or GCF, of two integers is the largest number (also an integer) that divides evenly into both of the given integers. One way to find the greatest common factor of a pair of numbers is to list all of the factors of each and then hunt through both lists to find the largest number they share.

For example, suppose we wish to find the greatest common factor of 36 and 60. We list all the factors of each: Searching through the lists show us that 12 is the largest number they have in common, so 12 is the greatest

36	1, 2, 3, 4, 6, 9, 12, 18, 36
60	1, 2, 3, 4, 5, 6, 10, 12, 15, 20, 30, 60

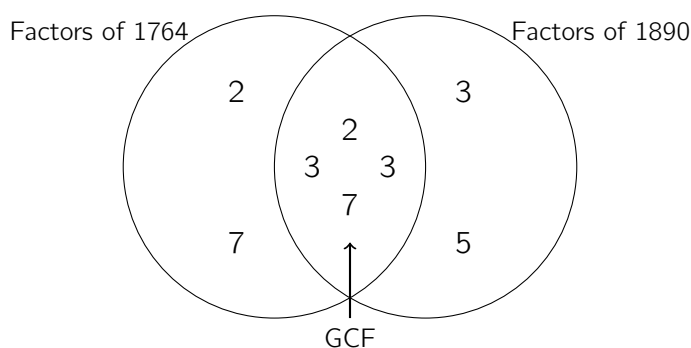
common factor of 36 and 60. We sometimes write $\text{GCF}(36, 60) = 12$.

The list-all-the-factors approach is fine for small numbers, but quickly becomes tedious if the numbers get big. For example, if we want to find the GCF of 1764 and 1890. An alternative method is to find the prime factorizations of the two numbers and compare them using a Venn diagram.

The prime factorizations of our two numbers are

$$1764 = 2 \cdot 2 \cdot 3 \cdot 3 \cdot 7 \cdot 7 \quad \text{and} \quad 1890 = 2 \cdot 3 \cdot 3 \cdot 3 \cdot 5 \cdot 7.$$

We arrange their prime factors in a Venn diagram like the one below, placing the prime factors of each number inside the appropriate circle. If the two numbers share a factor, we place that factor (once only) in the overlapping region.



The factors in the overlapping region are those shared by both of the numbers. The product of all the numbers in the overlap is the GCF. In this case, we have:

$$\text{GCF}(1764, 1890) = 2 \cdot 3 \cdot 3 \cdot 7 = 126.$$

We will use this Venn diagram approach to tackle our first polynomial factoring challenge.

16.2.1 GCF for polynomials

Undoing the Distributive Property by Factoring Out the GCF

This is what we are familiar with. This is how we simplify by distributing.

Simplify: $3x(2x + 5)$ $3x \cdot 2x + 3x \cdot 5$ $6x^2 + 15x$

Now factoring takes the $6x^2 + 12x$ (the product) and rewrites it so that you have the original, un-simplified multiplication problem. There is a lot of logic involved. You have to find the Greatest Common Factor of the terms of a polynomial. To find the GCF of a polynomial $6x^2 + 15x$, you have to look at each term. What do you notice about each of the terms? You should notice that both terms have x and both terms are multiples of 3. This gives you a clue to what some common factors are. . . 3 and x . Now, to find the greatest common factor. You can do this systematically and prime factorize each term.

$6x^2$: $2 \cdot 3 \cdot x \cdot x$ $15x$: $3 \cdot 5 \cdot x$

Circle the common factors. All of the common terms multiply together to form the GCF. What do you notice between the GCF and the previous problem? The GCF is that number that got sprinkled during the distribution. This means that if you can find the GCF, you can undo the distributive property and factor a polynomial.

Example 1

Find the GCF of $6x^4 + 4x^3 + 8x^2$ Solution: What do you notice about each of the terms? You may be able to find the GCF that way, or, if you can't, prime factorize each term. $6x^4$: $2 \cdot 3 \cdot x \cdot x \cdot x \cdot x$ $4x^3$: $2 \cdot 2 \cdot x \cdot x \cdot x$ $8x^2$: $2 \cdot 2 \cdot 2 \cdot x \cdot x$

All of the terms have an x^2 and all of them are even, so $2x^2$ is the GCF. Example 2

Find the GCF of each of the following polynomials. 1. $12x^4 + 18x^3$ 2. $32y^4 - 16y^2$ 3. $-4y^2 - 8y - 12$ Solution: 1. $6x^3$ 2. $16y^2$ 3. 4 or -4 (-4 technically isn't the greatest but it is sometimes convenient to use the negative factor to change all of the signs)

Factoring out the GCF

Technically, when you undo the distributive property you are "factoring out a monomial term." Now, if you can find the GCF, this is really easy. The GCF is the thing being sprinkled. The stuff left over after the GCF is taken out goes in parenthesis. It is the thing that gets the sprinkling. Now, the cool thing, is that this is so easy to check. If you want to see that you factored correctly, all you have to do is redistribute.

We will use $6x^3 + 4x^2 + 8x$. Start by finding the GCF, which we know is $2x$ from example 1. The GCF is the monomial that goes outside of the () in the distribution problem., whatever factors are left over from the prime factorization makes up the polynomial that goes inside of the (). $2x(3x^2 + 2x + 4)$. Now, check to see if they are equivalent. Distribute the $2x$ or just graph both on your graphing calculator.

If you factor out the GCFs of the polynomials in example 2, you get the following:

1. $6x^3(2x + 3)$ 2. $16y^2(2y^2 - 1)$ 3. $4(-y^2 - 2y - 3)$ or $-4(y^2 + 2y + 3)$

Often, when you are asked to factor a polynomial, the first thing you should look for is a GCF that can be factored out of the problem. Actually, one of the most missed things on the assessment over factoring is forgetting to factor out a GCF.

16.3 Factoring level 2: Monic quadratics

This is how to factor using algebra tiles. We are basically taking the multiplication and doing it in reverse. Instead of starting with the dimensions, we are going to start with the tiles and arrange them into a rectangle. The length and width of the rectangle will be the factors we are looking for. How can these tiles be arranged into a rectangle?

First, you must know the polynomial being represented by the tiles. Then, experiment and turn them into a single rectangle. After that, find the length and width of the rectangle. Remember that the “ x^2 ” tiles have to be in the top left corner and the unit tiles have to form a rectangle in the lower right corner. The “ x ” tiles fill in the spaces.

Now all you have to do is determine what the factors are that create this rectangle. Look along the left side. It is made up of $1x$ and 1 unit, so the binomial is $(x+1)$. Finally look along the top. It is made up of $1x$ and 2 units, so the binomial is $(x+2)$. Therefore the factored form of $x^2 + 3x + 2 = (x+1)(x+2)$. This is easy to check. Just multiply the binomials back out or use a graphing calculator.

The tiles are manipulatives that are used to show the area model of a polynomial. It is just as easy to realize that a quadratic trinomial results from the product of two binomials from experience. The multiplication of 2 binomials results in 4 terms, two of which are combined in the final step. Instead of drawing the individual tiles, you can just draw a rectangle and break it up into 4 sections. The top left section represents the quadratic term, the bottom right is the section containing the constant term. The other two sections are the two that combined to form the linear term. If you use this instead, you have to figure out how to break up the linear term. There is only one combinations that is going to work. In addition, once you figure out how to break it up and have your rectangle drawn, you find the GCF of each column to find one factor. Find the GCF of each row to find the other factor.

Quadratic Term Part of linear term Part of linear term Constant Term

Of course, we want to be able to do this without the tiles and without drawing rectangles. In essence we are anti-foiling, or anti-double distributing. Think about where all of the terms in the quadratic come from in terms where/when during the process of multiplication.

1. Where does quadratic term come from? 2. Where does constant term come from? 3. Where does linear term come from?

Knowing the answers to the 3 questions above, you can factor using logic.

1. The quadratic term comes from multiplying x by x . So we know our factors have to start $(x \dots)(x \dots)$ 2. Since all of the terms are positive, we know our factors have to be $(x + \dots)(x + \dots)$ 3. We also know that the factors have to look like $(x + \text{some number})(x + \text{some other number})$ 4. So what could those numbers be? Well, we know they multiply together to give us the “ c ” term. Since the c in this case is 2, we know the two numbers have to be 1 and 2. 5. It looks like our factors are going to be $(x+1)$ and $(x+2)$, but we have to be sure. The

way to check is to look at the “bx” term. We know that comes from multiplying the “outer” and “inner” terms together and adding those products together. Let’s check to see if this is going to work. The “outer” product is $2x$, the inner is $1x$, the sum is $3x$. That is what we were given originally, so $x^2 + 3x + 2 = (x+1)(x+2)$

If you want to factor without the tiles, you have to “guess and check”. Here’s an organized way to do just that. It is, of course, based on the above sequence of logical steps.

1. Find the factors of the “c” term. 2. The factors that add up to the “b” term are the correct ones. 3. Check the signs to make sure they will multiply correctly! 4. Check your answer!

That is the actual procedure we are going to use to factor quadratic trinomials of the form $x^2 + bx + c$. Remember to incorporate the sign of the factors of c . The final step is to always check to make sure everything works right. It is really easy to pick the wrong factors or the wrong sign on the factors. I’ll go through the reasoning only on some of these, but give the solution to all of them. Remember, for this special case, I am looking for factors of c that add to b .

Example 1

Factor:

1. $x^2 + 7x + 12$ 2. $x^2 + 8x + 12$ 3. $x^2 + 2x - 3$ 4. $x^2 - 6x + 8$ 5. $x^2 + x - 12$ 6. $x^2 - 3x - 10$ 7. $x^2 - 8x + 15$
8. $x^2 - 3x - 18$ 9. $x^2 - 3x + 2$ 10. $x^2 - 10x + 21$

Solution: 1. We need to look for factors of $+12$ that add to 7 . This means they both have to be positive. The possible factor pairs are $(1, 12)$, $(2, 6)$, and $(3, 4)$. Which of these pairs add up to 7 ? That would be 3 and 4 . It seems that $(x + 3)(x + 4)$ might be the factors. They add up to 7 and multiply out to 12 . Time to check $(x+3)(x+4) = x^2 + 4x + 3x + 12 = x^2 + 7x + 12$. 2. $(x + 6)(x+2)$ 3. We need to look for factors of -3 that add to 2 . This means that one is negative and one is positive. Also that the positive one is “bigger.” The possible factor pairs are $(1, -3)$ and $(-1, 3)$. Which of these add up to $+2$? $(-1, 3)$. $(x - 1)(x+3)$ looks to be the answer. Time to check $(x-1)(x+3) = x^2 + 3x - x - 3 = x^2 + 2x - 3$ 4. We need to look for factors of $+8$ that add to -6 . This means that they both are negative. The possible factor pairs are $(-1, -8)$, $(-2, -4)$. Which of these add up to -6 ? $(-2, -4)$. $(x - 2)(x - 4)$ might be the solution. Check: $(x-2)(x-4) = x^2 - 4x - 2x + 8 = x^2 - 6x + 8$ 5. $(x + 4)(x - 3)$ 6. $(x-5)(x+2)$ 7. $(x - 3)(x-5)$ 8. $(x - 6)(x + 3)$ 9. $(x - 2)(x-1)$ 10. $(x - 7)(x-3)$

16.4 Factoring levels 3 and 4: Two special cases

16.4.1 Level 3: Difference of squares

Factoring level 3

// startup exploration

The polynomials below don't look like Level 2 factoring problems, since they have only two terms. How can we change these into Level 2 problems? What happens when we factor them using Level 2 techniques?

$$x^2 - 9$$

$$x^2 - 100$$

$$x^2 - 36$$

We can extend each of the binomials in the startup exploration by adding a linear term with coefficient zero. Then we'd have, for example,

$$x^2 + 0x - 9$$

To factor this, we are looking for two numbers that multiply to -9 and add up to 0 . The natural choice is 3 and -3 . So, we have the factorization:

$$x^2 - 9 = (x + 3)(x - 3)$$

Following this chain of reasoning, we have:

$$x^2 - 100 = (x + 10)(x - 10) \quad \text{and} \quad x^2 - 36 = (x + 6)(x - 6)$$

Polynomials of this form are called the *difference of squares* since we have some perfect square minus a different perfect square. The resulting binomial factors are almost identical, but with different signs on their second term. They have to be of opposite sign for that bx term to disappear from the final product!

Not the sum of squares!

Important note: this is not the sum of squares! For example, we can't factor the following polynomial in the same way:

$$x^2 + 100$$

Folks can be tempted to see the sum of squares and write out a factorization, but this does not work the same way! In fact, expressions like this are never factorable

in Algebra 1. (Can you explain why not?)

When modeling the difference of squares with algebra tiles, we have a picture that looks like fig. 16.1. The green x tiles and the red $-x$ tiles zero one another out, leaving the x^2 tile and a collection of -1 tiles.

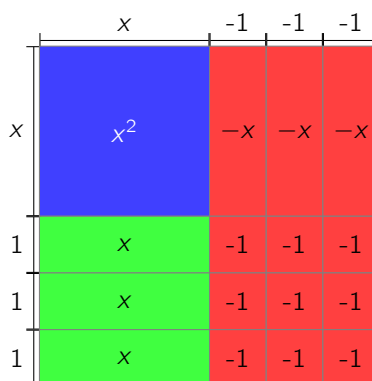


Figure 16.1: Difference of squares: $x^2 - 9 = (x + 3)(x - 3)$

The general form of a difference of squares, in which a and b can be any terms, is given below. Thinking of a and b as generic terms will let us expand this technique a bit further.

Difference of squares

$$a^2 - b^2 = (a + b)(a - b).$$

Beyond monics

If the quadratic is monic, then its leading term is a perfect square. After all $x^2 = 1x^2 = x \cdot x$. But we can have other leading terms that are also perfect squares. In fact, as we saw in chapter 12, the quadratic term is a perfect square if its coefficient is a perfect square. So, for example:

$$16x^2 = 4x \cdot 4x$$

Given this idea, we can factor quadratics like:

$$16x^2 - 25 = (4x + 5)(4x - 5)$$

So, we have a technique now for handling non-monic quadratics, at least in certain cases.

Beyond quadratics

We are going to keep our main focus on quadratics, but it is worth mentioning that we have a technique for factoring certain higher-degree polynomials. How might we tackle this one?

$$x^4 - 16$$

In this case, 16 is definitely a perfect square, and so is $x^4 = (x^2)^2$. So, this is the difference of squares, and we have:

$$x^4 - 16 = (x^2 + 4)(x^2 - 4)$$

We just factored a quartic polynomial, and that's something! But wait a minute – the second factor here is again a difference of squares. So, we can factor even more. The simplest form is:

$$x^4 - 16 = (x^2 + 4)(x^2 - 4) = (x^2 + 4)(x + 2)(x - 2)$$

Beware at this point! Don't go "factor crazy" and try and factor that $x^2 + 4$. Remember: that's the sum of squares and we can't factor that (yet).

We like to call this a Goldilocks problem. Remember "Goldilocks and the Three Bears", the bedtime story about breaking and entering? In the story, Goldilocks goes through the bears' home messing with everything until she finds something that's "just right". The first bowl of porridge is too hot, the second bowl is too cold, but the third is "just right". The first bed is too hard, the second is too soft, but the third is "just right".¹

We call polynomials like $x^4 - 16$ "Goldilocks polynomials" because there's a temptation to factor too little... and a temptation to factor too much. We have to know what's "just right".

16.4.2 Level 4: Perfect square trinomials

Factoring level 4

// startup exploration

Carry out the following multiplications. What patterns do you notice that relate the factors to the product?

$$(x + 5)^2$$

$$(x - 4)^2$$

$$(3x - 5)x^2$$

A trinomial that is formed by squaring a binomial is called, naturally enough, a **perfect square trinomial**. In the exponential unit (section 11.5), we called this raising a sum to a power. We used the quadrangle method to

¹ In the original versions of the Goldilocks story, the person who trespasses in the bears' home was an old woman – a foul-mouthed criminal, in fact – which, perhaps, makes a lot more sense given her actions in the story.

great benefit when solving quadratic equations (chapter 12). The diagram that we drew was a kind of abstract algebra tiles picture in which all the tiles together formed a large square.

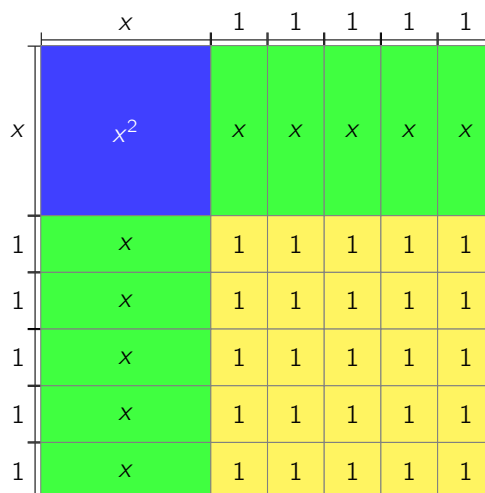


Figure 16.2: Perfect square trinomial: $(x + 5)^2 = x^2 + 10x + 25$

Perfect square trinomials have some special features which make them easier to multiply and easier to factor. Notice that the first and last terms of the trinomial are both perfect squares. Note also that the bx term is twice the product of the terms in the binomial that we are squaring!

Perfect square trinomial

$$(a + b)^2 = (a + b)(a + b) = a^2 + ab + ba + b^2 = a^2 + 2ab + b^2$$

$$(a - b)^2 = (a - b)(a - b) = a^2 - ab - ba + b^2 = a^2 - 2ab + b^2$$

As with the difference of squares, we can use this model to factor non-monic polynomials. The key is knowing when we have a perfect square trinomial and when we don't. The first, and easiest, thing to check is that the quadratic term and the constant term are both perfect squares. From there, we can imagine a candidate factorization – but we still have to check the linear term!

Example 16.1

Determining whether each of the following are perfect square trinomials. If so, write the factorization.

$$x^2 - 12x + 36$$

$$9x^2 + 34x + 25$$

$$64x^2 + 48x + 9$$

Solution: In the first polynomial, x^2 is a perfect square and 36 is a perfect square. So, we might conjecture that the factorization is $(x + 6)$. We have to check the linear term, which must be double the

product of the terms in our binomial: $2 \cdot x \cdot 6 = 12x$, which is close. To get the signs to work out, we can rely on karmic multiplication: $(-6)^2 = 36$. So, our factorization is:

$$x^2 - 12x + 36 = (x - 6)^2$$

The first and last terms of the second example should lead us to conjecture that the factorization is $(3x + 5)^2$. But, if this were true then the middle term would be $2 \cdot 3x \cdot 5 = 30x$. We have $34x$, and so the given trinomial is not a perfect square.

In the third example, we conjecture that the factorization is $(8x + 3)^2$. This would, in turn, require the linear term to be: $2 \cdot 8x \cdot 3 = 48x$. Good news, that's what we have! So, the factorization is:

$$64x^2 + 48x + 9 = (8x + 3)^2$$

The second polynomial in the previous example *is factorable*, however it is not a perfect square. More on how to tackle problems like that one in the next section!

16.5 Factoring level 5: Factoring by grouping

This is a general form of factoring that will allow us to factor just about anything that can be factored. To get the hang of the mechanics, we will start by factoring expressions that look a bit abstract. Then, we'll apply this technique in context.

Factoring level 5

// startup exploration

Use the “multiple distribution method” to simplify $(a + b)(c + d)$.

Based on the patterns you notice after doing the multiplication above, rewrite the following expression as the product of two binomials.

$$pw + ps + hw + hs$$

Factoring by grouping is simply the multiple distribution method in reverse. We can multiply the two binomials in the startup exploration by first distributing the $(c + d)$ term over the other term. We would then have:

$$a(c + d) + b(c + d).$$

After distributing again, we have:

$$ac + ad + bc + bd.$$

We can visualise this using an abstract algebra tiles diagram, like so:

	a	b
c	ac	bc
d	ad	bd

How is the structure of the expression above related to this four-term expression?

$$pw + ps + hw + hs$$

Suppose we arrange the four terms in a diagram like the one above, what can we learn?

pw	hw
ps	hs

Among these four terms, we have several pairs that share a GCF. In fact, each of the columns and rows share a GCF! If we factor out those GCFs and write them as the column headings, we have factored the expression: $(p + h)(w + s)$!

	p	h
w	pw	hw
s	ps	hs

Another way to think about what's happening is start again with the expression

$$pw + ps + hw + hs$$

and to notice that the terms $pw + ps$ have p as a common factor, and the terms $hw + hs$ have h as a common factor. If we factor out these GCFs, then we have:

$$p(w + s) + h(w + s)$$

This is an expression with two terms, and they share the common factor $(w + s)$. So, we can factor this out as their GCF:

$$(p + h)(w + s)$$

Note: Sometimes you will have to rearrange the factors to find the correct groups. You may also have to factor out a negative to get the leftovers to be identical.

Example 1

Factor completely.

1. $ad + ac - d - c$ 2. $3x^2 - 4x - 6x + 8$ 3. $xy + 3y + 4x + 12$ 4. $3x(x - 4) + 2(x - 4)$ 5. $y^2(2x + 5) - (2x + 5)$

Solution:

1. The two groups are $ad + ac$ and $-d - c$. Factor out "a" from the first group. For the second group you will need to factor out a -1 to change the signs. Whenever you need to change the signs of an expression, you can factor out a negative 1. $\rightarrow (a - 1)(d + c)$ 2. The two groups are $3x^2 - 4x$ and $-6x + 8$. Factor out an "x" from the first group. For the second group, you will need to factor out a -2, not just a two, once again to make the signs correct. $\rightarrow (x - 2)(3x - 4)$ 3. The two groups are $xy + 3y$ and $4x + 12$. Factor out "y" from the first group. For the second group you will need to factor out a 4. $\rightarrow (y + 4)(x + 3)$ 4. The first few steps were already done for us. $\rightarrow (x - 4)(3x + 2)$ 5. The first few steps were already done for us. $\rightarrow (2x + 5)(y^2 - 1)$... but this last one is tricky! First, when one of the groups has "nothing" in front, or only a + or -, that means that there is a "1" there that becomes part of the factor. Second... my answer is "completely factored". You might have noticed that $y^2 - 1$ is a difference of squares so the answer is $\rightarrow (2x + 5)(y + 1)(y - 1)$

Now to apply these ideas to polynomials. Up until this point, we have only been able to factor “special” polynomials that had a certain helpful structure. Factoring by grouping is going to allow us to factor any polynomial, although we might have to do some tinkering first.

Let’s use this new technique on a problem we already know how to solve. We know how to factor $x^2 + 5x + 6$ already (that’s a Level 2 problem). How could we use grouping to factor this?

First of all, we don’t have enough terms: factoring by grouping requires four terms, as though we’re filling in a tile diagram. So, we have to break that middle term up into 2 different terms – we need to “un-combine like terms”. But how?

Our options are to break $5x$ into $x + 4x$, or $2x + 3x$... and only one of them will work. Which one? Let’s try both to find out.

Option 1, we break $5x$ into $x + 4x$ and try to factor by grouping.

$$x^2 + 5x + 6 = x^2 + x + 4x + 6 = x(x + 1) + 2(2x + 3)$$

When we factor the GCF from each pair of terms, the leftovers inside the parentheses are not identical. So, the resulting terms don’t share a GCF and we’re stuck.

Option 2, we break $5x$ into $2x + 3x$ and try again.

$$x^2 + 5x + 6 = x^2 + 2x + 3x + 6 = x(x + 2) + 3(x + 2)$$

Woo hoo! The two terms here share a GCF and so, we can factor it out. In the end, we have:

$$x^2 + 5x + 6 = (x + 2)(x + 3)$$

Of course, we should have known that was coming. Still, we can see how the process works.

Let’s try: $2x^2 + 13x + 15$

If you hate grouping, you don’t have to use it. You may have to spend a lot of time guessing and checking though. If I were to use pure guess and check, I have the following options:

$$(x + 15)(2x + 1) \quad (x + 3)(2x + 5) \quad (x + 5)(2x + 3) \quad (x + 1)(2x + 15)$$

I need to find two linear binomials so that the first term is $2x^2$ (formed by $2x * 1x$) and the last term is 15 (formed by either $1*15$ or $3*5$). I then need to find the combination that will give me a $13x$. This one isn’t too bad because there aren’t very many options. It turns out that $(x+5)(2x+3)$ is the combination that will work. You have to be aware of the sign and check your answer!

If you want something more streamline and systematic than just guess and check, you have to factor by grouping. I just have to figure out what to break $13x$ into to get two groups. By inspection it is obvious that I want to break it into $3x + 10x$. By grouping $2x^2 + 3x + 10x + 15 = x(2x + 3) + 5(2x + 3) = (x + 5)(2x + 3)$.

Now, of course, there is a structured way to find what to break the terms up into. It's not always... "just look at it and figure it out"

The Procedure to factor a quadratic trinomial with grouping. ($ax^2 + bx + c$)

1. Multiply a and c 2. Find the factors of the product "ac" that will add to b. 3. Split "bx" into two terms using the numbers found in step 2. 4. Group 5. Factor out the GCF from each group 6. Factor into binomials

Now, let's see this with an example that would be horrible to use guess and check with: $20x^2 + 7x - 6$

1. $20(-6) = -120$ 2. Option: 1, 120; 2, 60; 3, 40; 5, 24; 6, 20; 8, 15; 10, 12 \rightarrow since it is negative, that means one is positive and one is negative, but they have a difference of +7 so it has to be -8 and 15. 3. $20x^2 - 8x + 15x - 6 \rightarrow$ (it doesn't matter which way you split the terms) 4. $(20x^2 - 8x) + (15x - 6)$ 5. $4x(5x - 2) + 3(5x - 2)$ 6. $(4x + 3)(5x - 2)$

Example 2

Factor each of the following by grouping

1. $6x^2 - 7x - 5$ 2. $10x^2 + 17x + 3$ 3. $8x^2 + 6x - 5$ 4. $6x^2 - 19x + 10$

Solution:

1. $(2x + 1)(3x - 5)$ 2. $(5x + 1)(2x + 3)$ 3. $(2x - 1)(4x + 5)$ 4. $(3x - 2)(2x - 5)$

16.6 Solving equations by factoring

We began this chapter with an example of how factoring can be helpful when it comes to solving polynomial equations. We introduced the zero product property, which – although it seems pretty elementary – is a powerful tool when cleverly applied. (Have a look back at section 16.1 for a refresher.)

Here's a fresh example. Suppose we want to solve the equation

$$x^2 - 3x - 40 = 0.$$

This is a Level 2 factoring problem, and so which can be factored without too much trouble:

$$(x - 8)(x + 5) = 0.$$

Since the product of these two equals zero, it must be that either $x - 8 = 0$ (meaning $x = 8$) or $x + 5 = 0$ (meaning $x = -5$). So, we have found two solutions to the quadratic, and $\mathcal{S} = \{8, -5\}$. In this case, factoring might be easier than applying the quadrangle diagram approach.

Factoring works even when other approaches are possible. For example, it's easy to solve $x^2 = 25$ using the POEs. We know that $x = 5$ or -5 . But this is confirmed by a factoring approach:

$x^2 = 25$	the original equation
$x^2 - 25 = 0$	prepare for the zero product property
$(x + 5)(x - 5) = 0$	factor the difference of squares

This means that either $(x + 5) = 0$, in which case $x = -5$, or $(x - 5) = 0$, in which case $x = 5$. And so again, we have $\mathcal{S} = \{5, -5\}$. In this case, solving using the POEs is easier than factoring.

In some cases it's not clear which approach is easier. Suppose we have the quadratic equation

$$6x^2 - 7x - 1 = 4.$$

We could carry out the quadrangle diagram procedure on this equation, as we have done many times before. Or, we could try to apply what we know about factoring. This technique is the most helpful in conjunction with the zero product property, and so we need first to have an equation equal to zero. So, we first use SPOE, which gives us a trinomial equal to zero.

$$6x^2 - 7x - 5 = 0$$

Can we factor the left-hand side? If so, it will require factoring by grouping (since neither the first nor last term is a perfect square). This means we're hunting for two numbers whose product is -30 and whose sum is -7 . That means -10 and 3 are the numbers we're after! This gives us the factorization:

$$6x^2 - 10x + 3x - 5 = 2x(3x - 5) + 1(3x - 5) = (2x + 1)(3x - 5)$$

And so this means that we have the equation:

$$(2x + 1)(3x - 5) = 0,$$

which means that we have two equations to solve: $2x + 1 = 0$ and $3x - 5 = 0$. Solve these two-step equations gives us our solutions:

$$S = \left\{ -\frac{1}{2}, \frac{5}{3} \right\}$$

In this case, it may have been just as easy (perhaps even easier) to apply the quadrangle method.

Not every polynomial is factorable, and so not every equation can be solved by factoring. And just because an equation can't be solved by factoring doesn't mean that it has no solutions. There are quadratics that have irrational solutions that aren't factorable over integers. Conversely, if a quadratic has no solution, then it is definitely not factorable.

Note for EOC... Sometimes you will see that the EOC will ask you to find "roots" or "zeroes" of a function. That means that you are trying to find x-intercepts. By now, you should have noticed that finding x-intercepts means to set the y to zero and solve the corresponding equation.

16.6.1 Completing the square

16.6.2 Quadratic formula

At the end of our study of quadratic equations, in section 12.6.1 to be precise, we used the quadrangle method to solve a generic quadratic equation. It was an ugly bit of alphabet soup, but it is yet another way to solve a quadratic equation.

The quadratic formula

The solutions to the quadratic equation $ax^2 + bx + c = 0$ (where $a \neq 0$) are given by the formula:

$$x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}.$$

» Chapter summary «

Wow! Think of how far you have come! For one thing, we have reached the apex of our work with quadratics. We have various tools for solving quadratic equations: the quadrangle method, the factoring method, and the quadratic formula. We can graph quadratic functions, and manipulate them in a variety of ways. We can manipulate radical expressions. That's quite a skill set!

Plus, we've seen a glimpse of what lies beyond quadratics. We have tinkered with other polynomials: factoring a quartic and even solving a cubic equation!

If we add this list of skills to everything that came before – everything we learned about linear functions, systems of equations and inequalities, exponential functions – well. . . we have come quite a long way.

The only thing left to say is: Congratulations!

To add:

irreducible

factored form – connect back

x-intercepts – connect back

completing the square – connect back

solve a cubic

Appendix A

Table of prime numbers to 2000

In addition to the numbers 2 and 5, the tables below show all of the prime numbers up to 2000.

H	0 2 1 3 4			
U T	1	3	7	9
0	•	••	••	•
1	••	••	•	•
2	•	••	••	••
3	••	••	••	••
4	••	•	•	••
5	••	•	••	••
6	•	••	••	•
7	••	••	•	••
8	••	••	•	•
9	•	••	••	•

5 7 6 8 9			
1	3	7	9
••	•	••	••
••	•	•	••
••	•	••	••
•	•	•	••
••	••	••	•
•	•	••	••
••	•	•	••
••	••	••	•
•	••	••	••

10 12 11 13 14			
1	3	7	9
••	•	•	••
•	••	••	••
••	••	••	••
••	•	•	••
••	••	••	••
•	•	•	••
••	••	••	••
•	•	•	••
••	••	••	••

15 17 16 18 19			
1	3	7	9
••	•	••	••
••	•	•	•
••	••	•	•
••	••	•	•
•	•	••	••
•	••	•	••
••	••	••	•
••	••	••	••
•	••	••	••

How to use the table

Suppose we want to know whether the number 1439 is prime. Here's how we use the table to determine this:

1. The box on top of each table shows a group of hundreds (that's what the H stands for at the left side). We should think of the number 1439 as having "14" in the hundreds place. The table heading of the third table includes the number 14, so 1439 will appear in this section of the table.

2. The rows of the table indicate the digit in the tens place (that's what the T stands for at the left side). Since 1439 has a 3 in the tens place, we'll use that row of the table.
3. The columns in each table indicate the units place (that's what the U stands for at the left side). We need the column with heading 9, since that's the number in the units place of 1439.
4. We have narrowed our selection down to a single cell in the table. That cell looks like this:



The dots indicate which of the numbers associated with this cell are prime, using the same layout as in the topmost box. This pattern indicates that 1439 is prime. The box also tells us that 1039 is prime, and that 1139, 1239, and 1339 are all composite.

Notice and wonder

Notice that the columns include only the headings 1, 3, 7, and 9. Why do you suppose that is?

Notice that none of the boxes have all five dots. Why do you suppose that is? If we continue the table, do you suppose we will ever find a box with all five dots? Why or why not?

Notice that if a box contains four dots, it's always the four corners. Why do you suppose that is? If we continue the table, do you suppose we will ever find a box with four dots in a different arrangement? Why or why not?

Glossary

Symbols · A · B · C · D · E · F · G · H · I · L · M · N · O · P · Q · R · S · T · U · V · X · Y · Z

Symbols

Δ Delta, the fourth letter of the Greek alphabet. Used to represent change. The symbol Δx is read “delta x ” or “the change in x ”.

A

abscissa The x -coordinate of a point in the coordinate plane.

absolute value For real numbers, it is the distance a number is away from zero on a number line. It is a scalar quantity, meaning it just has a magnitude and no direction (sign). The absolute value of a number is always non-negative. In the order of operations, it works like a grouping symbol. The “absolute value of x ” is denoted $|x|$.

addition property of equality For all real numbers a , b , and c : If $a = b$, then $a + c = b + c$. This axiom is used when solving equations.

addition property of order For all real numbers a , b , and c : If $a > b$, then $a + c > b + c$. This axiom is used when solving inequalities and also applies to inclusive symbols of order.

additive identity The number which, when added to a given number x , leaves x unchanged. In the real number system, 0 is the additive identity. The existence of the additive identity is a **field axiom**.

additive inverse The number which, when added to a given number x , gives a sum of 0, the additive identity. The opposite of a number is its additive inverse. The existence of the additive inverse is a **field axiom**.

algebraic expression A symbolic representation of mathematical operations that can involve both numbers and variables. There is no equal sign in an expression.

algebraic number A number that is the root on a nonzero polynomial equation in one variable with rational coefficients. The set of algebraic numbers is a subset of the real numbers.

arithmetic sequence A sequence where the difference between each pair of successive terms is constant. The constant difference is called the “common difference”, usually denoted d .

associative property of addition For all real numbers a , b , and c : $a + (b + c) = (a + b) + c$. This field axiom allows for the regrouping of longer strings of addition.

associative property of multiplication For all real numbers a , b , and c : $a(bc) = (ab)c$. This field axiom allows for the regrouping of longer strings of multiplication.

asymptote A line that a curve approaches as they both tend towards infinity. There are three types of asymptotes: vertical, horizontal, and oblique (slant). Exponential functions have a horizontal asymptote.

axiom A property or statement that is accepted without proof.

axis One of two perpendicular number lines used to locate points in the coordinate plane. The plural form is “axes”.

axis of symmetry The line about which one can reflect an image onto itself. For example, a parabola has an axis of symmetry. Given the graph of a quadratic function, the axis of symmetry is a vertical line through the vertex. When written in standard form, the equation for the line of symmetry is given by $x = -\frac{b}{2a}$.

B

base (1) For triangles: A side of the triangle. (2) For expressions: A term or expression that is raised to a power.

binomial A polynomial with exactly 2 terms.

boundary For a one-variable **inequality**, the boundary is a point on the number line. Inclusive boundaries are drawn as closed or filled in points, and exclusive boundaries are drawn as open circles. For a two-variable inequality, the boundary is a line or curve. Inclusive boundaries are drawn as solid lines/curves, and exclusive boundaries are lines/curves drawn with a dashed or dotted line. For a linear inequalities, the boundary line separates the plane into two **half-planes**, one of which will contain the solutions to the inequality.

C

Cartesian plane See **coordinate plane**.

closure A set is said to be to “have closure” (or to “be closed”) under an operation performing the operation on members from the set always yields a result that is also a member of the set. The **natural numbers**, for example, are closed under the operation of addition, since the sum of any two natural numbers is itself a natural number. The natural numbers are not closed under the operation of subtraction.

coefficient The numerical factor in a term with a variable. If the number is not explicitly written, the coefficient is understood to be 1.

colinear To be on the same line.

combining like terms A short-cut used to add terms that have exactly the same variables raised to the same exponents.

common difference In an arithmetic sequence, it is the constant difference between successive terms.

common monomial factor A monomial that is a factor of every term in a polynomial expression.

common ratio In an geometric sequence, it is the constant ratio between successive terms.

commutative property of addition For all real numbers a and b : $a + b = b + a$. This field axiom allows for the reordering longer strings of addition.

commutative property of multiplication For all real numbers a and b : $ab = ba$. This field axiom allows for the reordering longer strings of multiplication.

completing the square Using the properties of equality on a quadratic equation to convert one side into a perfect square trinomial. Completing the square can be used as a technique to solve quadratic equations.

complex number A member of the set of numbers that consists of real and imaginary numbers. The set is denoted \mathbb{C} .

compound interest A way to calculate interest based on both the principal amount and any interest already accrued. This type of interest is an exponential relationship. The formula is $A = P \left(1 + \frac{r}{n}\right)^{nt}$, where P is the principal amount, r is the rate of interest, t is the amount of time over which interest is to be computed, and n is the number of compounding periods per unit of time.

constant A value that does not change.

constant function A function whose graph is a horizontal line. It is of the form $f(x) = c$, where c is a constant. Constant functions are polynomial functions of degree zero.

constant multiplier In a sequence that grows or decays exponentially, the number each term is multiplied by to get the next term. Also known as the “common multiplier”, or “common ratio”.

constant of variation The constant ratio in a direct variation or the constant product in an inverse variation. It is designated with the variable k .

constant term A term that includes no variable.

constraint The limitations on the values of the variables in a problem. Equations, inequalities and systems are used model the constraints in real-world situations.

continuous data Data that has no breaks and has measurements that can change between data points. Graphically, the measured data points are connected with lines or curves.

continuous function A function that has no breaks in the domain or range. The graph of a continuous function is a line or curve with no holes, gaps, or vertical asymptotes.

converse of the Pythagorean theorem If a triangle has sides a , b , and c , such that $a^2 + b^2 = c^2$, then the triangle is a right triangle with a hypotenuse of length c .

conversion factor A ratio used to convert measurement from one unit to another.

coordinate plane A plane with a pair of scaled, perpendicular axes allowing one to locate points with ordered pairs and to represent lines and curves by equations. Also known as the Cartesian plane, named for its creator, French philosopher René Descartes.

coprime See [relatively prime](#).

correlation Used in describing data graphed in a scatter plot. It is a trend between two variables. A trend can show positive, negative, or no correlation. Positive correlation shows an [increasing](#) trend in data. Negative correlation shows a [decreasing](#) trend in data.

cubic A function, number, or expression raised to the third power. Called cubic as it relates to the volume of a cube.

D

decay factor In exponential decay, the constant multiplier used to calculate the amount of decay after each unit of time. In the formula $y = (1 - r)^x$, it is the quantity $(1 - r)$. It represents the quantity remaining and is the common multiplier in the exponential relationship.

decay rate In exponential decay, the fraction or percentage by which a population decreases for each unit of time. In the formula $y = (1 - r)^x$, it is the quantity r .

decreasing A function is said to be decreasing if as x increases, y decreases. Lines with negative slopes are decreasing.

degree of a polynomial The degree of the term in a polynomial with the highest degree.

degree of a term The power (exponent) to which the variable is raised in a variable term. If there is no exponent explicitly written on a variable in a term, the term is understood to be of degree 1. The degree of a constant term is zero.

denominator The number or expression below the [vinculum](#) in a [rational number](#) or [rational expression](#). For example, in the number $\frac{5}{2}$, the denominator is 2.

dependent variable A variable whose values depend on the values of another variable. In a graph of the relationship between the two variables, the values on the vertical axis represent the values of the dependent variable. The generic variable used is y .

difference of squares A binomial of the form $a^2 - b^2$.

dimensional analysis A strategy for converting a measurement from one unit to another using multiplication by a string of conversion factors. The key is to include the units with the numbers. It is used often in science.

direct variation In algebra 1, a relationship in which the ratio of two variables is constant. A direct variation has an equation of the form $y = kx$. The quantities represented by x and y are said to be **directly proportional**. The value k is called the **constant of variation**.

directly proportional Used to describe two variables whose values have a constant ratio.

discrete data Data that can only take on certain values. Discrete data usually involves a count of items.

discrete function A function whose domain and range have breaks or are made up of distinct values rather than intervals of real numbers. The graph of a discrete function will have breaks or will be made up of distinct points.

discriminant The expression under the square root in the quadratic formula, used to determine the number and nature of the roots of a quadratic. If a quadratic equation is written in standard form, then the discriminant is $b^2 - 4ac$. If the value of the discriminant is greater than 0, there are two real solutions to the quadratic equation. If it is equal to zero, there is one real solution. If it is less than zero, there are no real solutions to the quadratic equation.

distance formula A formula based on the Pythagorean theorem that uses the coordinates of two points to calculate the distance between the two points. The formula for the distance d between any two points (x_1, y_1) and (x_2, y_2) is $d = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}$.

distributive property For all real numbers a , b , and c : $a(b + c) = ab + ac$. This field axiom allows one to simplify an expression without having to evaluate the sum inside the grouping symbol first.

division property of equality For all real numbers a , b , and c , where $c \neq 0$: if $a = b$, then $\frac{a}{c} = \frac{b}{c}$. This property is a version of the multiplication property of equality. It is used when solving equations.

division property of order For all real numbers a , b , and c where $c > 0$: if $a < b$, then $\frac{a}{c} < \frac{b}{c}$. If, on the other hand, $c < 0$, then $a < b$ implies $\frac{a}{c} > \frac{b}{c}$. This property is a version of the multiplication property of order. It is used when solving inequalities.

domain The set of all input values of a function, or the x -values. In a problem context it is represented by the independent variable.

domain restriction Values that cannot be used in the domain of a function. Radical and rational functions have domain restrictions.

doubling time In exponential growth, the amount of time it takes for a population, or amount, to double in size. It is constant for an exponential relationship.

E

elimination method A method for solving a system of equations that involves adding or subtracting multiples of the equations in order to eliminate a variable. It is based on Gaussian Elimination, a method to solve systems of equations that have been converted into matrices.

equation A statement that says the value of one expression is the same as the value of another expression.

equivalent equations Equations that have the same solution set.

equivalent inequalities Inequalities that have the same solution set.

evaluate To find the value of an expression. If an expression contains variables, values must be substituted for the variable before the expression can be evaluated.

exclusive boundary See [boundary](#).

exclusive inequality See [inequality](#).

expanded form The form of an exponential expression in which repeated multiplication is written out. For example, the right-hand side of the equation $x^3 = x \cdot x \cdot x$, is written in expanded form.

exponent A number or variable written as a small superscript to a number or a variable, called the [base](#), that indicates how many times the base is being used as a factor.

exponential decay A decreasing pattern in which amounts decrease by a constant percent.

exponential equation An equation in which a variable appears in the exponent.

exponential form The form of an expression in which repeated multiplication is written using exponents.

exponential function A function that repeatedly multiplies an initial amount by the same positive number. They can all be modeled using $y = ab^x$ where a is the initial amount and b is the constant multiplier.

exponential growth An increasing pattern in which amounts increase by a constant percent.

extraneous solution An apparent solution of an equation that does not satisfy the original equation. They occur when the transformation of an equation changes the solution set of the original equation, for example squaring both sides of an equation or multiplying by a quantity that can be zero.

F

factor One of the numbers, variables, or expressions multiplied to obtain a product.

factored form The form of an expression when it is written as the product of factors. The factors can be numbers, variables, or expressions. Factored form is not simplified.

factoring The process of rewriting an expression as a product of factors.

family of functions Similar functions that are all transformations of the same parent function.

field axiom One of a set of axioms including closure, identity, inverse, associative, commutative, and distributive properties. Along with a few definitions and properties of equality, they create the foundation upon which algebra is built.

FOIL A mnemonic for remembering the procedure to multiply two binomials. F stands for multiplying the first term in each binomial. O stands for multiplying the outer terms of the binomials. I stands for multiplying the inner terms of the binomials. L stands for multiplying that last term in each binomial.

fractal A geometric figure that has undergone infinite applications of a recursive procedure and which exhibits the property of self-similarity.

function A relation in which there is exactly one output value for each input value. The graph of a function must pass the vertical line test.

function notation A notation in which a function is named by a letter and the input is shown in parenthesis after the function name, generically, $f(x)$, read “ f of x ”. The variables used may be changed to better represent quantities in a problem, for example $d(t)$ may represent distance d as a function of time t . When graphing in the **coordinate plane**, $f(x)$ is another way to write y . When x is replaced by a number, it indicates that one should evaluate the function at that value. The notation was first used by Swiss mathematician Leonhard Euler.

function rule An expression that represents the relationship between the variables of a function.

G

GCF Greatest Common Factor

geometric sequence A sequence where the ratio between each pair of successive terms is constant. The constant ratio is called the “common ratio”, usually denoted r . Geometric sequences are exponential.

growth factor In exponential growth, the constant multiplier used to calculate the amount of growth after each unit of time. In the formula $y = (1 + r)^x$, it is the quantity $(1 + r)$. It is the common multiplier in the exponential relationship.

growth rate In exponential growth, the fraction or percentage by which a population increases for each unit of time. In the formula $y = (1 + r)^x$, it is the quantity r .

H

half-life The time needed for an amount of a substance to exponentially decay to half the original amount. Half-life is constant for an exponential relationship.

half-plane The set of points on a plane that fall on one side of a boundary line. Part of the solution of a linear inequality in two variables is a half-plane.

hypotenuse The side of a right triangle opposite the right angle. It is the longest side of the triangle.

I

identity When solving equations with variables on both sides, identities occur when the equation is true for every value of the variable. The solution set S is written as $S = \mathbb{R}$.

identity property of addition The sum of any number and 0 is that number. For every real number a , $a + 0 = a$ and $0 + a = a$. The existence of the **additive identity** is a **field axiom**.

identity property of multiplication The product of any number and 1 is that number. For every real number a , $a \cdot 1 = a$ and $1 \cdot a = a$. The existence of the multiplicative identity is a **field axiom**.

imaginary number A member of the set of numbers that is created by taking the square root of a negative number. In the set of imaginary numbers, the square root of -1 is represented by the letter i . The set of imaginary numbers is a subset of the complex number system. The sets of real and imaginary numbers are disjoint, meaning they have no common members.

implied operation An operation that is not explicitly written. For example, in $3(x + 4)$ the multiplication between 3 and $(x + 4)$ is an implied operation, since no multiplication symbol is explicitly written in between.

improper fraction A fraction whose **numerator** is greater than its **denominator**. For example, $\frac{5}{2}$ is an improper fraction. A fraction that is not an improper fraction is called a **proper fraction**. See also **mixed number**.

inclusive boundary See **boundary**.

inclusive inequality See **inequality**.

increasing A function is said to be decreasing if as x increases, y increases. Lines with positive slopes are decreasing.

independent variable A variable whose values affect the values of another variable. In a graph of the relationship between the two variables, the values on the horizontal axis represent the values of the dependent variable. The generic variable used is x .

inequality A statement that one quantity is less than or greater than another. An inequality may exclusive or inclusive. The exclusive inequalities are $<$ and $>$, read "less than" and "greater than". The inclusive inequalities are \leq and \geq , read "less than or equal to" and "greater than or equal to".

initial value The starting value of a sequence or exponential function.

integer A member of the set of natural numbers, their opposites, and zero. The set is denoted \mathbb{Z} , and we may write $\mathbb{Z} = \{0, \pm 1, \pm 2, \pm 3, \dots\}$. The integers are a subset of the rational numbers.

intercept The point which a graph intersects one of the axes.

interest A percentage of the balance added to an account at regular time intervals.

interest rate The percentage used to calculate interest.

inverse property of addition For any real number a , there exists a real number $-a$ such that $a + -a = 0$. The number $-a$ is called the **additive inverse** of a . Very often we will call it the **opposite** of a .

inverse property of multiplication For any nonzero real number a , there exists a real number $\frac{1}{a}$ such that $a \cdot \frac{1}{a} = 1$. The number $\frac{1}{a}$ is called the **multiplicative inverse** of a . Very often we will call it the **reciprocal** of a .

inverse variation In algebra 1, a relationship in which the product of two variables is constant. An inverse variation has an equation in the form $xy = k$, or $y = \frac{k}{x}$. The quantities represented by x and y are said to be **inversely proportional**. The value k is called the **constant of variation**.

inversely proportional Used to describe two variables whose values have a constant product.

irrational number A number that cannot be expressed as the ratio of two integers. In decimal form, an irrational number has an infinite number of digits and does not repeat. The set of irrational numbers consist of algebraic and transcendental numbers. The set of irrational numbers is a subset of the real numbers.

irreversible operation An operation performed when solving an equation that changes the solution set of the equation. Multiplying or dividing both sides of an equation by an expression that might equal zero are considered irreversible operations.

L

leg One of the perpendicular sides of a right triangle.

like terms Terms with exactly the same variable factors in a variable expression. The variables and the powers to which the variables are raised must be identical for the terms to be considered like terms.

limited domain The restricted domain of a function. Domains are usually limited in real world contexts. For example, we rarely allow negative values for a variable that represents “time”. For this reason it is often referred to as a reasonable domain.

line of best fit A line used to model a set of data. A line of best fit shows general direction of the data. When hand-drawn, one should have about the same number of data points above and below the line. When using the linear regression tool on the calculator, the correlation coefficient will show how well the line fits the data.

line of symmetry See [axis of symmetry](#).

linear In the shape of a line or represented by a line. In mathematics, a linear equation or expression has variables raised only to the power of 1.

linear function A function characterized by a constant rate of change. The graph of a linear function is a non-vertical line. It is a polynomial of degree one.

linear inequality An inequality of two variables whose boundary is formed by a linear function. It describes a region of the coordinate plane that consists of a boundary line and a half-plane.

linear programming A method to optimize a quantity that uses an objective function to represent the quantity and a system of linear inequalities to represent the constraints on the variables involved. The system of inequalities are graphed to represent a set of feasible solutions and the vertices of the region will describe the optimal amount of the quantity.

linear relationship A relationship that can be represented by a linear function. A linear relationship is characterized by a constant rate of change.

linear term A term of degree 1.

lowest terms The form of a fraction in which the numerator and denominator are [relatively prime](#). A fraction in lowest terms is also called a reduced fraction.

M

mapping diagram A diagram used to determine if a relation is a function. The values of the domain and range are written in circles. Arrows are drawn from the elements of the domain to the corresponding elements of the range. It is a visual that shows how the members of the domain map to the members of the range.

mathematical equivalence The idea that numbers, expressions, equations, functions, or other mathematical objects can be algebraically manipulated, using specific rules, such that their representations and appearance are changed while other fundamental properties remain unchanged.

mathematical modeling Translating a real-world scenario with a given set of constraints into an abstract representation that can be manipulated and studied mathematically. For example, creating a set of variables and equations to solve a [linear programming](#) problem.

maximum The greatest value. In a quadratic function, the vertex will be a maximum if the coefficient of the quadratic term is negative.

midpoint The point on a line segment halfway between the endpoints. The coordinates of the midpoint are found by averaging the abscissas and ordinates of the endpoints.

midpoint formula The formula that can be used to compute the midpoint of a line segment. Given a line segment with endpoints (x_1, y_1) and (x_2, y_2) , the midpoint of the segment has coordinates $\left(\frac{x_1+x_2}{2}, \frac{y_1+y_2}{2}\right)$.

minimum The smallest value. In a quadratic function, the vertex will be a minimum if the coefficient of the quadratic term is positive.

mixed number The sum of a nonzero **integer** and a **proper fraction**. For example $2\frac{3}{5}$ is a mixed number. See also **improper fraction**.

monomial A polynomial with only one term.

multiplication property of equality For all real numbers a , b , and c : if $a = b$ then $ac = bc$. This property is used to solve equations.

multiplication property of order For all real numbers a , b , and c and $c > 0$: if $a < b$ then $ac < bc$. If, on the other hand, $c < 0$, then $a < b$ implies $ac > bc$. This property is used to solve equations.

multiplicative identity The number which, when multiplied by a given number x , leaves x unchanged. In the real number system, 1 is the multiplicative identity. The existence of the multiplicative identity is a **field axiom**.

multiplicative inverse The number which, when multiplied by a given nonzero number x , gives a product of 1, the multiplicative identity. The reciprocal of a number is its multiplicative inverse. The existence of the multiplicative inverse is a **field axiom**.

N

natural number A member of the set $\{1, 2, 3, 4, \dots\}$, denoted \mathbb{N} . Also called the counting numbers. The number 0 is sometimes included as a natural number.

negative correlation See **correlation**.

null set A set that contains no elements. Also called the empty set. Used to show that there is no solution to an equation. Denoted \emptyset or $\{\}$.

numerator The number or expression above the **vinculum** in a **rational number** or **rational expression**. For example, in the number $\frac{5}{2}$, the numerator is 5.

numeric expression An expression containing only numbers and mathematical operations.

O

obelus The division symbol \div .

one-variable data Data that measures only one trait or quantity. A one-variable data set consists of single values (as opposed to ordered pairs) and is graphed on a number line. Compare with: **two-variable data**.

opposite See **additive inverse**.

optimization To maximize or minimize a quantity given constraints. For example a company will want to optimize (maximize) their profits while faced with constraints such as the cost and availability of labor and materials.

order of magnitude A way of expressing the size of an very large or very small number by giving the power of 10 associated with the number.

order of operations The agreed-upon order in which operations are carried out when evaluating an expression.

ordered pair A pair of numbers named in an order that matters. The coordinates of a point are given as an ordered pair in which the first number is the x-coordinate (abscissa) and the second number is the y-coordinate (ordinate).

ordinate The y-coordinate of a point in the coordinate plane.

origin The point where the coordinate axes intersect. In a coordinate plane it has the coordinates (0, 0).

P

parabola The set of all points whose distance from a fixed point (called the focus) is equal to the distance from a fixed line (called the directrix). Also known as the smooth "U" shaped curve of a quadratic function.

parallel lines Lines in the same plane that never intersect. They are always the same distance apart in Euclidean geometry. The slopes of parallel lines are the same.

parent function The most basic form of a function. A parent function can be transformed to create a family of functions.

percent change The percent by which an amount differs from its original amount. It is calculated by taking the amount of the change and dividing it by the original amount.

perfect cube A number that is equal to the cube of an integer, or a polynomial that is equal to the cube of another polynomial.

perfect square A number that is equal to the square of an integer, or a polynomial that is equal to the square of another polynomial.

perfect square trinomial A trinomial generated by squaring a binomial. For example, squaring the binomial $(a + b)$ yields $(a + b)^2 = a^2 + 2ab + b^2$. Thus, $a^2 + 2ab + b^2$ is a perfect square trinomial.

period of compounding The number of times interest is calculated during a year for compound interest. It is represented by n in the **compound interest** formula.

perpendicular lines Lines that intersect at a right angle. The slopes of perpendicular lines are opposites and reciprocals. The slopes of perpendicular lines multiply to -1 .

point-slope form The form of a linear equation that uses the slope and any point on the line. It is written either $y - y_1 = m(x - x_1)$ or $y = m(x - x_1) + y_1$, where m is the slope of the line and (x_1, y_1) is a point on the line. It can be derived from the slope formula and represents the transformation of the line $y = mx$ where a vertical shift of y_1 and a horizontal shift of x_1 has occurred.

polynomial A sum of terms that have positive integer exponents. In algebra 1, all polynomials are in one variable.

positive correlation See [correlation](#).

power An expression of the form a^n is called a power of a .

principal amount The original amount invested in a situation that involves accumulating interest. It is represented by P in the [compound interest](#) and simple interest formulas.

principal square root The positive square root of a number.

proper fraction A fraction whose [numerator](#) is less than its [denominator](#). For example, the fraction $\frac{7}{9}$ is a proper fraction. A fraction that is not a proper fraction is called an [improper fraction](#).

proportion An equation stating that two ratios are equal.

Pythagorean theorem A formula that expresses the relationship between the sides of a right triangle. It states that the sum of the squares of the legs of a right triangle is equal to the square of the [hypotenuse](#).

Q

quadrangle method A technique for solving quadratic equations that is based on the symmetry of the square. When executing the quadrangle method, we often draw a *quadrangle diagram*.

quadrant One of the four regions that a coordinate plane is divided into by the two axes. The quadrants are numbered I, II, III, and IV, starting in the upper right and moving counterclockwise.

quadratic formula The formula used to find the exact solution to any quadratic equation. Given that $ax^2 + bx + c = 0$, the formula states

$$x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$$

It is derived by completing the square on the standard form quadratic equation.

quadratic function A function with an equation of the form $y = ax^2 + bx + c$ where $a \neq 0$. The graph of a quadratic function is a [parabola](#).

quadratic term A term of degree 2.

R

radical The root symbol $\sqrt{}$, used to denote square roots, cube roots, and so on. The symbol $\sqrt[n]{x}$ is read “ n th root of x .” If n is not stated, as in \sqrt{x} , it is understood to be 2 and the radical indicates the square root.

radical expression An expression containing a radical (square root, cube root, or any n th root).

radical function A function where the independent variable is under a radical (square root, cube root, or any n th root).

radioactive decay The process by which an unstable element loses mass with a release of energy, transforming it into a different element or isotope.

range (1) In statistics it is the difference between the greatest value in a data set and the smallest value in a data set. (2) In the study of functions it is the set of all output values of a function. It is represented by the dependent variable.

rate A **ratio** that measures two quantities with different units.

rate of change A measurement of how quickly one quantity changes relative to another quantity. Given values (x_1, y_1) and (x_2, y_2) , the rate of change of y with respect to x is $\frac{\Delta y}{\Delta x} = \frac{y_2 - y_1}{x_2 - x_1}$. The patterns of the rate of change of a set of data can be used to determine what type of data is represented by the pattern. For example, the rate of change of linear data is constant.

ratio A comparison between two quantities, often written in fraction form.

rational expression An expression that can be written as a ratio of two polynomials. The value of the variable cannot make the denominator 0.

rational function A function that is expressed as the ratio of two polynomial expressions. The values of the independent variables that make the denominator zero are restricted from the domain.

rational number A number that can be written as a ratio of two integers $\frac{a}{b}$ where $b \neq 0$. Their decimal forms are either terminating or repeating. The set of rational numbers is denoted \mathbb{Q} . The rational numbers are a subset of the real numbers.

rationalizing the denominator The process of making the denominator of a fraction a rational number without changing the value of the expression. It is used to eliminate a radical from the denominator of a fraction.

real number Denoted \mathbb{R} , the set of real numbers include the integers, rational numbers, and irrational numbers, but not imaginary numbers. This is the number set used in algebra 1. The set is closed under the operations of addition and multiplication. Members can be graphed on the standard number line. The real numbers is a subset of the complex numbers.

reciprocal The multiplicative inverse. The reciprocal of a given number is the number it must be multiplied by to get 1 (the multiplicative identity). To find the reciprocal of a number, we can write the number as a fraction and then invert the fraction. The reciprocal of n is $\frac{1}{n}$.

recursive Describes a procedure that is applied over and over again, starting with a number or a geometric figure, to produce a sequence of numbers or figures. The procedure requires previous entries in the pattern to find subsequent entries.

recursive rule Instructions for producing each stage of a sequence from the previous stage. It must contain a description of “stage 0”, or the starting value.

recursive sequence An ordered list of numbers defined by a starting value and a recursive rule. We generate a recursive sequence by applying the rule to the starting value, then applying the rule to the resulting value, and so on.

relation Any set of ordered pairs.

relatively prime Two numbers are said to be relatively prime (or coprime) if they have no common factors other than 1. For example, 16 and 21 are relatively prime. In contrast, 21 and 24 are not relatively prime, since both numbers are divisible by 3.

repeating decimal A decimal representation of a rational number with a digit or group of digits after the decimal point that repeat infinitely.

root A zero or an x-intercept of a function.

S

sample space The set of all possible outcomes of a probability experiment.

scatter plot A two-variable data display in which values on a horizontal axis represent values of one variable and values on the vertical axis represent values of the other variable. The coordinates of each point represent a pair of data values.

scientific notation A notation in which a number is written as the product of a number greater than or equal to 1 but less than 10, multiplied by an integer power of 10.

sequence A function whose domain is the set of positive integers. A sequence is an ordered list of objects, like numbers. The individual objects are called terms. Unlike a set, order matters, and terms may be repeated.

set An unordered collection of items. Often denoted by listing the elements inside a set of braces.

set notation Using curly braces { and } to designate quantities that belong to a set. Certain sets do not require the use of braces, as they have symbols used to denote them, like the **null set**, the set of **integers**, and the set of **real numbers**.

simple interest Interest calculated using the formula $I = Prt$. The interest is only ever calculated using the initial investment (called the **principal amount**) and show linear growth.

simplified radical form A radical written so that (1) no perfect square factors exist under the radical (2) no fractions are under the radical and (3) there are no radicals in the denominator of the fraction.

simplify Using algebraic laws and properties which maintain equivalence in order to write an answer so that it fits a set of criteria. The criteria depend on what is being simplified.

slope The measurement of the steepness of a line, or the rate of change of a linear relationship. Often denoted m , and referred to as “rise over run.” Given points (x_1, y_1) and (x_2, y_2) , the slope of the line between the points is calculated as $m = \frac{\Delta y}{\Delta x} = \frac{y_2 - y_1}{x_2 - x_1}$.

slope-intercept form The form $y = mx + b$ of a linear equation. The value of m is the slope and the value of b is the y -intercept. It is the simplified version of **point-slope form**.

solution A solution to an equation (or inequality) is any value of the variable (or variables) in the equation (or inequality) that make the equation (or inequality) true. The solution to a system of equations (or inequalities) is the set of all of the points common to all equations in the system. If there is no solution, the system is said to be inconsistent. If there are infinitely many solutions to a system, the system is said to be dependent. If there is a single solution, the system is said to be independent. In a system of two equations in two variables, the solution is the intersection point of the two lines.

solution set The set of values that make an equation, inequality, or system true.

solution set notation One way to denote the solution set to an equation, written as $S = \{ \text{solutions} \}$.

solve To find the solution set of an equation.

square root The square root of a number a , denoted \sqrt{a} , is the number b such that that $b \cdot b = a$. Every positive number has two square roots, a **principal square root** and a negative square root. The set of real numbers is not closed under the operation of square root.

standard form (1) For linear equations, it is an equation of the form $Ax + By = C$, in which A and B are not both 0. (2) For a polynomial, it is an expression written such that it is simplified and the terms are written in decreasing order of degree (highest degree term appears first). (3) For quadratic equations, it is an equation of the form $ax^2 + bx + c$, where $a \neq 0$.

subset A subset is a set that consists entirely of members from another set. If a set A is a subset of a set B , then every item in A is in B .

substitution To replace a quantity with another one that is equivalent.

substitution method A method for solving a system of equations that involves solving one of the equations for one variable and substituting the resulting expression into the other equation. See also: **elimination method**.

subtraction property of equality For all real numbers a , b , and c : if $a = b$ then $a - c = b - c$. This property is a restatement of the **addition property of equality** and is used to solve equations.

subtraction property of order For all real numbers a , b , and c : If $a > b$, then $a - c > b - c$. This property is a restatement of the **addition property of order** and is used to solve inequalities.

system of equations A set of two or more equations with the same variables. The equations act as constraints on the variables.

system of inequalities A set of two or more inequalities with the same variables. The inequalities act as constraints on the variables.

T

term An algebraic expression that represents only multiplication and division between variables and constants.

terminating decimal A decimal number with a finite number of nonzero digits after the decimal point.

transcendental number An irrational number that is not algebraic. The number π is transcendental because it is not the root of a polynomial equation in one variable with rational coefficients.

trinomial A polynomial with exactly three terms.

turducken A food product in which a deboned chicken is stuffed inside a deboned duck, which is then stuffed inside a deboned turkey. In culinary terminology, this is an example of “engastration”, a cooking method in which one animal is stuffed inside the gastric cavity of another.

two-variable data A collection of data that measure two traits or quantities. A two-variable data set consists of pairs of values. Compare with: **one-variable data**.

U

unit rate A ratio in which one of the quantities has the value of 1.

unknown A quantity in an equation whose value is not known. In algebra, letters are often used to represent unknowns.

V

variable A trait or quantity whose value can change, or vary. In algebra, letters are often used to represent variables.

vector A quantity that has both a size (or magnitude) and a direction. Vectors play an important role in physics and engineering, since many physical quantities (such as velocity, acceleration, and force) are best represented using vectors.

vertex Of a parabola, the point where the graph changes direction from increasing to decreasing or from decreasing to increasing.

vertex form A form of a quadratic equation. Given that (h, k) is the vertex, this form is written either as $y - k = a(x - h)^2$ or $y = a(x - h)^2 + k$. It can be derived by completing the square on standard form and represents the transformation of $y = ax^2$ by translation h units horizontally and k units vertically.

vertical line test A method for determining whether a graph on the coordinate plane represents a function. If all possible vertical lines cross the graph only once or not at all, the graph represents a function. If even one vertical line crosses the graph in more than one point, the graph does not represent a function.

vertical motion formula When an object is dropped or launched vertically, its height can be expressed using $h(t) = at^2 + vt + h_0$, where $h(t)$ is the object's height at time t , v is its initial vertical velocity, h_0 is its starting height, and a is the acceleration of gravity. For dropped objects, v is zero. This formula is used in the study of projectile motion.

vinculum A bar used in mathematics to show grouping. Examples of vincula include: the fraction bar (as in $\frac{1}{x+2}$), the bar used to show repeating digits (as in $0.\overline{3}$), and the horizontal bar of a radical (as in $\sqrt{2+5}$).

X

x-axis The horizontal number line on a coordinate graph. The independent variable is drawn on the x-axis.

x-intercept Any point at which a graph intersects the x-axis.

Y

y-axis The vertical number line on a coordinate graph. The dependent variable is drawn on the y-axis.

y-intercept Any point at which a graph intersects the y-axis.

Z

zero Of a function, the values of the independent variable that make the corresponding values of the function equal to zero, also known as the **roots** or x-intercepts of the function.

zero product property Property of real numbers stating that if the product of two or more factors equals zero, then at least one of the factors must equal zero. This property is used along with factoring as a method for solving a polynomial equation.

Bibliography

- [1] Richard G. Brown, Mary P. Dolciani, Robert H. Sorgenfrey, and William L. Cole. *Algebra: Structure and Method, Book 1*. McDougal Littell, 2000.
- [2] Richard G. Brown, Mary P. Dolciani, Robert H. Sorgenfrey, and Robert B. Kane. *Algebra and Trigonometry: Structure and Method, Book 2*. McDougal Littell, 1999.
- [3] Edward B. Burger. *Extending the frontiers of mathematics: Inquiries into proof and argumentation*. Key College Publishing, 2007.
- [4] Edward B. Burger and Michael Starbird. *The heart of mathematics: An invitation to effective thinking*. Key Curriculum Press, 2000.
- [5] Sergey Dorichenko. *A Moscow math circle: Week-by-week problem sets*. American Mathematical Society, 2011.
- [6] Paul A. Foerster. *Algebra and Trigonometry: Functions and Applications*. Addison-Wesley, 2005.
- [7] Timothy Gowers. *Mathematics: A very short introduction*. Oxford University Press, 2002.
- [8] Ken Johnson and Ted Herr. *Problem Solving Strategies: Crossing the River with Dogs and Other Mathematical Adventures*. Key Curriculum Press, 2001.
- [9] Ron Larson and Robert P. Hostetler. *Algebra and Trigonometry*. Cengage Learning, 2003.
- [10] Sandor Lehoczky and Richard Rusczyk. *The Art of Problem Solving, Vol. 1: The Basics*. AoPS Incorporated, 2006.
- [11] Carol Meyer and Tom Sallee. *Make it simpler*. Dale Seymour Publications, 1983.
- [12] Frederick Mosteller. *Fifty challenging problems in probability with solutions*. Dover Publications, 1965.
- [13] Jerald Murdock, Ellen Kamischke, and Eric Kamischke. *Discovering advanced algebra: An investigative approach*. Key Curriculum Press, 2004.

- [14] Jerald Murdock, Ellen Kamischke, and Eric Kamischke. *Discovering algebra: An investigative approach*. Key Curriculum Press, 2007.
- [15] Alfred S. Posamentier and Charles T. Salkind. *Challenging problems in algebra*. Dover Publications, 1996.
- [16] Paula Rozell. *Plotting Pictures: Coordinate Graphing and Number Skills Practice, Grades 5-10*. Pearson Education, 1995.
- [17] Richard Rusczyk and Sandor Lehoczky. *The Art of Problem Solving, Vol. 2: And Beyond*. AoPS Incorporated, 2006.
- [18] Reuben Schadler. *Algebra Problems: One Step Beyond*. Pearson Learning, 1990.
- [19] Harold L. Schoen, editor. *Teaching mathematics through problem solving: Grades 6-12*. National Council of Teachers of Mathematics, 2006.
- [20] Michael Serra. *Discovering geometry: An investigative approach*. Key Curriculum Press, 2003.
- [21] Peter Winkler. *Mathematical puzzles: A Connoisseur's Collection*. A K Peters/CRC Press, 2003.
- [22] Paul Zeitz. *The art and craft of problem solving*. John Wiley and Sons, 2007.

Index

- absolute value, 6
- addition
 - of rational numbers, 17
 - of signed numbers, 7
- addition property of equality, 109, 407
- addition property of order, 214, 408
- additive identity, 97, 399
- additive inverse, 7, 97, 400, 402
- algebraic expression, 61, 101, 357
- apparent formula, 41
- arithmetic sequence, 46
- associative property of addition, 99
- associative property of multiplication, 99
- asymptote, 92
- axiom, 96
- axis of symmetry, 401

- base, 29, 261, 397
- boundary, 397, 399

- closure, 2
- coefficient, 62
- combining like terms, 105
- common difference, 46
- common ratio, 43
- commutative property of addition, 98
- commutative property of multiplication, 98
- completing the square, 345
- complex number, 5
- compound interest, 403, 404
- constant of variation, 146, 396, 400
- continuous data, 71
- coordinate plane, 57, 393, 398
- correlation, 70, 402, 404

- decreasing, 70, 395

- degree of a term, 101, 357
- denominator, 13, 399, 404
- direct variation, 146
- directly proportional, 146, 396
- discrete data, 71
- distance formula, 323
- distributive property, 99
- division
 - by zero, 97
 - of rational numbers, 20
 - of signed numbers, 9
- division property of equality, 109
- division property of order, 214

- elimination method, 198, 407
- equation, 107
- evaluate, 62
- exclusive boundary, 211
- expanded form, 234, 261
- explicit formula, 41
- exponent, 29, 261
- exponential form, 234, 261

- field axiom, 96, 392, 399, 402
- formula
 - explicit, 41
 - recursive, 37
- fractal, 35
- function, 60, 78, 79

- geometric sequence, 43

- half-plane, 393
- hypotenuse, 321, 404

- identity property of addition, 97
- identity property of multiplication, 97

- imaginary number, 1
- implied operation, 32
- improper fraction, 19, 402, 404
- inclusive boundary, 211
- increasing, 70, 395
- inequality, 393, 397, 399
- integer, 3, 402, 406
- interest, 248
- interest rate, 248
- inverse property of addition, 97
- inverse property of multiplication, 97
- inversely proportional, 400
- irrational number, 4, 305

- leg, 321
- like terms, 104
- linear programming, 401
- lowest terms, 16

- mapping diagram, 82
- mathematical modeling, 186
- mixed number, 19, 399
- multiplication
 - of rational numbers, 14
 - of signed numbers, 9
- multiplication property of equality, 109
- multiplication property of order, 214
- multiplicative identity, 97
- multiplicative inverse, 97, 400

- natural number, 2, 393
- null set, 406
- numerator, 13, 399, 404
- numeric expression, 25

- obelus, 22
- one-variable data, 408
- opposite, 7, 97, 400
- order of operations, 25
- ordered pair, 57
- origin, 57

- parabola, 93, 404
- perfect square, 54
- perfect square trinomial, 379
- point-slope form, 166, 407
- polynomial, 357
- power, 29, 261
- principal amount, 248, 406
- principal square root, 308, 407
- proper fraction, 399, 402

- quadrant, 57

- radical, 308
- rate of change, 158
- ratio, 405
- rational expression, 395, 402
- rational number, 3, 395, 402
- rationalizing the denominator, 316
- real number, 4, 406
- reciprocal, 21, 97, 400
- recursive, 35
- recursive formula, 37
- relation, 78
- relatively prime, 16, 395, 401
- root, 409

- sample space, 143
- scatter plot, 69
- sequence, 35
- set, 2
- set notation, 2
- simple interest, 248
- slope, 159
- slope-intercept form, 171
- solution, 112
- solution set, 112
- square root, 307
- standard form, 175
- subset, 3
- subtraction
 - of rational numbers, 17
 - of signed numbers, 8
- subtraction property of equality, 109
- subtraction property of order, 214
- system of equations, 187

term, 35, 62, 101, 357

turducken, 22

two-variable data, 402

unit rate, 138

unknown, 108

variable, 61

vertex, 93

vertex form, 335

vertical line test, 80

vinculum, 13, 28, 395, 402

x-axis, 57

y-axis, 57

zero

division by, 97