EXCERPTS FROM

Precalculus Made Difficult

The whole book can be purchased as a paperback at Amazon.com, or as a pdf at BraverNewMath.com

Seth Braver

South Puget Sound Community College

VECTOR VECTORUM BOOKS

Vector Vectorum Books

Olympia, WA

Cover images created by Seth Braver with Dall-E 2

ISBN-13: 979-8-9881402-4-5

Library of Congress Control Number: 2023912570

The 1st Edition of this book was independently published in 2019.
This 2nd Edition is published by Vector Vectorum Books in 2023.

This book may occasionally be updated with minor changes.

Contents

ONLY THESE SECTIONS are contained in "EXCERPTS from PMD"

The full text can be purchased as a paperback at Amazon.com or as a pdf at BraverNewMath.com

Preface for Teachers

Precalculus Made Difficult is a straightforward text that guides students from the Plains of Mathematical Nowhere (I presuppose only arithmetic and a hazily-recalled encounter with the rudiments of algebra) to the base of Mt. Calculus in just 200 pages. I achieve this concision largely by treating my readers as literate, interested, capable people – people who just happen not to know much mathematics yet.

My choice of topics is mainly conventional, but I've squirreled novelties away in nooks and crannies. A few examples: an honest discussion of why we have an arithmetic of negative numbers (it's not for debts and subzero temperatures); a unit-circle definition of the *tangent* function (in lieu of the logically equivalent but aesthetically unsatisfying sine-over-cosine definition); a clean, symmetric *substitution*-based exposition of transformations in analytic geometry (which holds for graphs of *all* equations, unlike the usual asymmetric treatment, which applies only to graphs of functions); a proof Heron's formula (supplemented by a meditation on Heron's name); a rare excerpt from the *Annals* of Long Shu; dozens of delightful epigraphs; occasional dry humor flying low under the radar; and several story problems of questionable taste.

I am sufficiently immodest to assert that this book will be something of a connoisseur's choice. If you are a teacher who can appreciate the fine-wine qualities of G.F. Simmons' *Precalculus in a Nutshell* or Gelfand and Shen's *Algebra* (while recognizing that neither would work as an actual classroom text), then *Precalculus Made Difficult* may be the right text for you and your students. Unlike Simmons, which is addressed to *calculus* students in need of a quick refresher, or Gelfand and Shen, which is addressed to isolated Russian prodigies in correspondence school, this is an honest-to-God textbook, accessible to the average student who is meeting this material for the first time. I and a few others have used it in classes of perfectly ordinary community college students. Students who take their studies seriously handle it well. Students who do not do not. Such is life.*

On my website, BraverNewMath.com, you can find information about my other books and an email address for those with a burning desire to contact me.

Changes in the 2nd Edition

The changes are small and unobtrusive. I've tidied the book up, quietly improving its exposition in places as I did so, and I have corrected some typos sent to me by sharp-eyed readers of the 1st edition. The only substantial change – the true occasion for issuing this 2nd Edition – is the change of publisher.

I am honored to have been invited to publish *PMD* with the highly exclusive (and notoriously secretive) Vector Vectorum Books. They gambled by publishing my most recent book, *The Dark Art of Linear Algebra*, and pleased with the results, they then expressed interest in my back catalog. Here's to a long and fruitful relationship.

* As the old saying goes, you can lead a horse to water, but you can't make him drink. Less well-known is Dorothy Parker's variant, created when she was asked to use "horticulture" in a (spoken) sentence: "You can lead a whore to culture, but you can't make her think." One of *Precalculus Made Difficult*'s more obscure selling points is that it is the only precalculus book on the market whose prefatory material includes the word "whore".

Preface for Students

Yes, it's made difficult, but *I* didn't make it that way. The difficulty is intrinsic. It is also surmountable. It requires no special intelligence to master the material in this book, but it does require a commitment: a couple of hours a day, every day, for the better part of a year. Fortunately, this material (algebra, coordinate geometry, trigonometry) is genuinely engaging if you approach it in the right way. The right way is to *understand* it. If you understand mathematics, you'll experience the pleasure of feeling its pieces lock logically together in your mind in an aesthetically satisfying way.*

On the other hand, if you approach mathematics – as so many people do – in the *wrong* way (memorizing procedures without understanding *why* they work), then the subject will be sheer tedium. Don't let this happen. The difficulty is intrinsic. The tedium is not. The choice is yours.

Precalculus Made Difficult is meant to be read slowly and carefully. Strive to read the relevant sections in the text before your teacher lectures on them. The lectures will then reinforce what you've understood, and clarify what you haven't. Read with pencil and paper at the ready.† When I omit details, you should supply them. When I use a phrase such as "as you should verify", I am not being facetious. Verify it. Only *after* reading a section should you attempt to solve the problems with which it concludes. Whenever you encounter something in the text that you do not understand (even an individual algebraic step), you should mark the relevant passage and try to clear it up, which may involve discussing it with your classmates or teacher or reviewing earlier material.

Hundreds of thousands of people succeed in learning this material every year. You can be one of them. But it will require hard work, and at times, you may wonder whether it is worthwhile. It is.

Let's begin.

---

* Two relevant quotes:

"The most important thing a student can get from the study of mathematics is the attainment of a higher intellectual level."
- I.M. Gelfand (from the preface to *Functions and Graphs* by Gelfand, Glagoleva, and Schnol).

"Because it gives me chills in the spine."
- Serge Lang, after being asked why he studies mathematics (from Lang's *The Beauty of Doing Mathematics*)

† I recommend that purchasing a print copy so that you can scrawl notes in the margins. Online reading is rarely active reading.

Part I
Algebra

Chapter 1

Fractions and So Forth

Holy Whole Numbers, Fractious Fractions

God created the integers. All the rest is the work of man.
- Leopold Kronecker

Stories embedded deep in Western culture hint of an archaic past when **counting** was a form of magic, a divine prerogative too dangerous for mere humans. Zeus, the most powerful of the Greek gods, cruelly punishes Prometheus for giving humans the godlike knowledge of fire and, in Aeschylus's version, of *numbers*. Yahweh, the ancient Hebrews' "jealous god", punishes King David by killing 70,000 of his men, prompting David's impassioned cry, "I have sinned, and I have done wickedly: but these sheep, what have they done?" Never mind the sheep; what had David done? He had *counted* his men (2 Samuel 24).

Today, counting rarely provokes divine retribution, but the magical aura surrounding the whole numbers (1,2,3,4…) remains potent: We count sheep (not David's) as a charm to overcome insomnia, we "take a deep breath and count to 10" as a spell to allay anger or anxiety, we rate books, films, even pain, on pseudo-scientific scales of 1 to 10. Whole numbers speak to something deep within us. Fractions, alas, do not.

If whole numbers correspond to the divine activity of counting, then fractions – those broken, all-too-human numbers – correspond to the mundane, utilitarian activity of *measurement*. The gods do not seem particularly concerned to keep men from fractions, which is fortunate for the progress of science.

Far too many students fail calculus. They fail not because calculus is difficult (it isn't, really), but because they struggle with algebra. Often, they have cracks in their algebraic foundations running all the way back to arithmetic, particularly the arithmetic of fractions. To understand calculus, one must understand algebra; to understand algebra, one must understand arithmetic. Accordingly, in this first chapter, we'll review some basic algebra and arithmetic, emphasizing *why* the rules for manipulating numbers and algebraic symbols are what they are.

Understanding why such rules hold is every bit as important as understanding how to apply them. Pick up a 1000-page calculus book sometime. Feel its heft. Mere memorization of algorithms will not do, except as a temporary stopgap. The only way to truly learn mathematics is to understand it.

In Plato's *Meno*, Socrates distinguishes *knowledge* from mere *true opinions*. For our purposes, a "true opinion" corresponds to a correctly memorized mathematical rule. "As long as they stay put," says Socrates, "true opinions are fine things and do us much good. Only, they tend not to stay put for long. They're always scampering away from a person's soul. So they're not very valuable until you shackle them by figuring out what *makes* them true." *

Do not settle for true opinions. Strive for knowledge.

In this first chapter, we will simultaneously review numerical and algebraic fractions. To lay some basic groundwork for the algebraic side of fractions, we'll need to devote a few preliminary pages (before we reach fractions themselves) to some foundational algebraic ideas, all rooted in one core property: the so-called "distributive property".

* *Meno*, 98a.

The Distributive Property

The **distributive property** (or more precisely, "the distributive property *of multiplication over addition*") is the fact that multiplying something by a *sum* is equivalent to multiplying that same something by *each individual term in the sum,* then adding the results. For example, the distributive property tells us that

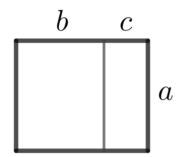
$$5(7+10) = (5 \cdot 7) + (5 \cdot 10) \quad \text{and} \quad (b+c+d)a = ab + ac + ad.$$

When describing such operations in words, we speak of "distributing the 5" (or the a) over the sum. Subtraction is just a special sort of addition (the addition of a negative), so we can distribute multiplication over subtraction, too. Thus, the distributive property guarantees that

$$(x-y-w)z = xz - yz - wz \quad \text{and} \quad 2a(a-b+c) = 2a^2 - 2ab + 2ac.$$

The distributive property is the bedrock on which much algebra is built, as you'll see in the next few pages. It also justifies some simple techniques of mental arithmetic, as you'll see in the exercises below.

Exercises



1. To understand the distributive property *visually*, consider the figure at right. The whole figure's area equals the sum of the areas of the two rectangles it contains. Rewrite the previous sentence as an algebraic equation involving a, b, and c.

[**Hint**: *The whole figure's height is* a. *What is whole figure's width?*]

2. Consider the following technique for mental calculation:

What's 32 times 7?
Let's see, 32 sevens is 30 sevens and then 2 more sevens.
Well, 30 sevens is 210, and 2 sevens is 14.
Thus, 32 times 7 must be 210 plus 14, which is 224.

Explain how the distributive property makes a quiet appearance in that calculation.

3. Using the technique from Exercise 2, compute the following in your head: $64 \cdot 5$, $82 \cdot 4$, $39 \cdot 9$, $6 \cdot 42$.

4. One can mentally calculate a 15% tip as follows:

What's 15% of $32?
Well, 10% of $32 is $3.20.
*Half of **that** (i.e. 5% of the whole) is $1.60,*
So a 15% tip would be $3.20 + $1.60 = $4.80.

Explain how the distributive property is at work here, too.

5. Calculate 15% of the following amounts in your head: $28, $50, $72, $90.

6. Students sometimes get overenthusiastic about distribution and try to distribute where distribution won't work. For example, we can't distribute multiplication over multiplication. To see why, find a *counterexample*. That is, find specific numbers a, b, c for which the expression $a(b \cdot c)$ is **not** equal to $ab \cdot ac$.

7. Now think of some specific counterexamples to demonstrate that

a) Exponents do *not* distribute over addition. (That is, $(a+b)^n \neq a^n + b^n$.)

b) Square roots do *not* distribute over addition. (That is, $\sqrt{a+b} \neq \sqrt{a} + \sqrt{b}$.)

The moral of the story: Multiplication distributes over addition, but not everything distributes over addition!

FOILed Again
(In Praise of Distribution, Part 1)

Early in your first algebra course, you learned how to multiply two "binomials". That is, you learned how to expand $(a+b)(c+d)$ out into the form $ac+ad+bc+bd$. Most algebra teachers summarize the steps in this process with the acronym FOIL (First, Outside, Inside, Last). The acronym is so common that it has become a verb, as in "when we FOIL this out, we obtain…"

All algebra students know how to "FOIL", but surprisingly few know *why* it works. This is a pity, since the explanation is so simple. It involves nothing but the distributive property. The key idea is that we can, when it's convenient to do so, think of a binomial as a *single entity* to be distributed. To emphasize this idea, I'll put a binomial in a grey box when I want to emphasize its unitary nature. Watch carefully:

$$\begin{aligned}(a+b)(c+d) &= \boxed{(a+b)}(c+d)\\ &= \boxed{(a+b)}c + \boxed{(a+b)}d & &\text{(distributing the "box", } (a+b)\text{)}\\ &= ac+bc+ad+bd & &\text{(distributing the } c\text{, and also the } d\text{)}\end{aligned}$$

Thus, under the hood, the mysterious "FOIL" operation is just shorthand for carrying out the distributive property several times. If you understand this, you can easily figure out how to multiply two *tri*nomials, three binomials, etc., without having to wait for someone to drill new acronyms into your head.

Exercises.

8. Explain to someone else why the "FOIL" rule is just a consequence of the distributive property.

9. You probably know that the algebraic expression $3x^2+4x^2$ can be rewritten as $7x^2$. But why is this allowed? It might *feel* obvious, but a feeling isn't an explanation. After all, someone else might *feel* that $(3x^2)(4x^2)$ should be $12x^2$, which is entirely wrong. (Factors don't care about your feelings.) In fact, $3x^2+4x^2$ equals $7x^2$ because of the distributive property. Explain why.

[**Hint:** *This requires some cleverness. Try "undistributing" something from* $3x^2+4x^2$.]

10. While reviewing some algebra, Esau encounters the expression x^2x^3. "Oh, I think I know how to simplify that," says Esau. "You just multiply the exponents so it becomes x^6, right?" Looking up from the delicious red stew he is cooking, Jacob flashes a grin at his brother and replies, "I don't think so. You're supposed to *add* the exponents, Esau. You should end up with x^5." Esau begins to change his answer, but then hesitates and says, "You're always tricking me, Jacob. Are you tricking me now?" To which Jacob replies, "Maybe. Maybe not."

a) Did Jacob give Esau the correct answer?

b) Don't settle for a true opinion. Turn it into knowledge: Explain *why* the correct answer is correct.

c) Simplify each of the following: x^3x^3, xx^4, x^4x^6, $(2x^2)(3x^4)$, $(-3x^3)(2x)(5x)$.

11. Use the distributive property (or equivalently, FOIL, when appropriate) to multiply the following polynomials.

a) $(3x-7)(2x+4)$
b) $(-x+2)(-2x-3)$
c) $(-x^4+3)(-2x^2+6x-1)$
d) $(x^2-2x+3)(-x^2+2x-7)$
e) $(x^3+x^2-x+1)(-x^3+x)$
f) $(x+1)(x+2)(x+3)$

[**Hint for Part F**: *One step at a time*. *First do* $(x+1)(x+2)$. *Then multiply the result by* $(x+3)$.]

g) $(x-1)(x^3+x^2+x+1)$
h) $(x-1)(x^5+x^4+x^3+x^2+x+1)$
i) $(x-1)(x^{99}+x^{98}+x^{97}+\cdots+x^3+x^2+x+1)$. [The "$\cdots$" indicates that the pattern continues.]

Factoring Polynomials
(In Praise of Distribution, Part 2)

"Factoring" is nothing more than distribution in reverse (or "undistribution" as I called it in Exercise 9). For example, why is it true that $3x^2 + 6x - 18 = 3(x^2 + 2x - 6)$? If you read that equation from right to left, you'll see why: It's the good old distributive property that justifies that equals sign.

You may have met (and forgotten) the "difference of squares" identity, $a^2 - b^2 = (a-b)(a+b)$. Reading it from left to right, it is somewhat mysterious. What makes this strange factoring formula true? Why should we believe it? Well, if we read it from right to left, we see how we might *prove* that it's true. We'll just need to multiply the factors on the right and see if we end up with the expression on the left. I'll do this now, putting a binomial in a grey box when we're thinking of it as one block.

Claim. $a^2 - b^2 = (a-b)(a+b)$ for all a and b.

Proof. $(a-b)(a+b) = \boxed{(a-b)}(a+b)$

$= \boxed{(a-b)}a + \boxed{(a-b)}b$ (distributing the $\boxed{(a-b)}$)

$= a^2 - ab + ab - b^2$ (distributing the a, and also the b)

$= a^2 - b^2$, as claimed. ■

Note how the proof boils down to three applications of distributive property – and nothing else. It is no exaggeration to say that the difference of squares formula holds *because of* the distributive property. In fact, the difference of squares formula is just the tip of the factoring iceberg. *Every* polynomial factorization can be justified by "reading right to left", multiplying the factors, and verifying that the result is the original polynomial. Since multiplying those factors is, as we've seen, just a matter of repeatedly distributing, it follows that the entire topic of factoring polynomials is justified by the humble distributive property. There's nothing deep here.

Most algebra textbooks make it seem as though factoring a polynomial is a complicated procedure. It isn't. Almost any polynomial you'll ever need to factor by hand will yield to one (or some combination) of three basic tricks:

1. Pull out a factor common to all the terms.
2. Use the difference of squares formula.
3. Make (and check) educated guesses until you find the right combination.

The first two tricks are entirely mechanical, and require little comment. Here's an example that uses both tricks in turn:

$$2x^2 - 32 = 2(x^2 - 16) = 2(x-4)(x+4).$$

Nothing to it: We factored out the 2, noticed that one of the resulting factors was a difference of squares, and thus applied the difference of squares formula. You should commit the difference of squares formula to memory since it comes in handy so often.

The third trick, making educated guesses, works best for polynomials of the form $ax^2 + bx + c$. For example, suppose we wish to factor $x^2 + x - 12$. Well, *if* it factors, the result might look like this: $x^2 + x - 12 = (x \quad\quad)(x \quad\quad)$. (We'll leave the constant terms blank to keep some play in the joints.) Note that putting those two x's on the right was an *educated* guess; we did it because their product (FOIL's "F") is x^2, which matches one term of our "target polynomial" on the left. To continue our educated guessing, we observe that whatever the factors' *constant* terms are, *their* product (FOIL's "L") must be

-12 to match the target polynomial's constant term. There are many possibilities here, so let's pick one pair of numbers whose product is 12 and try it out: How about $(x - \mathbf{4})(x + \mathbf{3})$? Will this work? We can check by multiplying it out in our heads;* doing so, we find that we end up with the wrong x term: we end up with $-x$ instead of $+x$. Close, but no cigar. What if we just swapped the locations of the positive and negative, like this: $(x + \mathbf{4})(x - \mathbf{3})$? A quick check shows that this *does* work, so our factoring is complete: $x^2 + x - 12 = (x + 4)(x - 3)$. And that's all there is to the third trick. With a little practice, which you'll in the exercises below, you'll develop an intuition for making good guesses.

As you can see, factoring efficiently depends upon being able to multiply binomials in your head… which is just FOILing… which is just using the distributive property. Yes, it requires practice and patience, but everything here should be entirely comprehensible.

Exercises.

12. Prove that $(a + b)^2 = a^2 + b^2 + 2ab$.

13. Prove that $(a - b)^2 = a^2 + b^2 - 2ab$.

14. Prove that $(a + b)^3 = a^3 + b^3 + 3a^2b + 3ab^2$.

15. Prove the "difference of cubes" identity: $a^3 - b^3 = (a - b)(a^2 + ab + b^2)$.

16. Factor the following polynomials as much as possible.

a) $x^2 + 6x + 8$ b) $x^2 - 100$ c) $10x^2 + 5x$ d) $x^2 - 7x + 10$

e) $3x^2 - 3x - 6$ f) $x^2 - 25$ g) $-4x^2 - 32x - 64$ h) $-15x^2 - 30x + 45$

i) $x^4 - 16$ [**Hint**: $x^4 = (x^2)^2$] j) $9x^2 - 4$ [**Hint**: *The first term is a square*.] k) $81x^8 - 1$

l) $2x^2 + 5x + 2$ [**Hint**: ($\mathbf{2}x$)(x)] m) $3x^2 - 8x - 3$ n) $4x^2 - 4x - 3$

17. The algebraic identities that you proved in exercises 12 and 13 are used so frequently that you should commit them to memory. From now on, whenever you need to square a binomial, you should apply these identities directly; don't reinvent the wheel each time. For example, to expand $(x + 3)^2$, you should **not** "FOIL it out" (even though doing so will, of course, yield the correct expansion). Rather, you should just mentally apply the identity, which allows you to write $(x + 3)^2 = x^2 + 9 + 6x$ without any further ado.

In this spirit, quickly expand the following:

a) $(x + 5)^2$ b) $(x - 5)^2$ c) $(x + 11)^2$ d) $(x - 12)^2$ e) $(2x + 1)^2$ f) $(3x - 2)^2$

g) $\left(a + \sqrt{2}\right)^2$ [**Hint***:* *$\sqrt{2}$ is, by definition, the number whose square is 2, so the value of $\left(\sqrt{2}\right)^2$ must be…*]

h) $\left(a - \sqrt{2}\right)^2$ i) $(2x + 1/2)^2$ j) $(2a + 3b)^2$ k) $\left(\sqrt{x} - 1\right)^2$

* The mental multiplication is easy: We know the product will include x^2 and 12 as terms (these are FOIL's "F" and "L" terms). The question is whether we'll get the x as well. Looking at our prospective factorization, $(x - 4)(x + 3)$, we see that FOIL's "O" and "I" terms give us $3x - 4x$, which simplifies to $-x$, so we've missed the target, alas. We'll need to try again.

Why Minus Times Minus is Plus
(In Praise of Distribution, Part 3)

Minus times minus is plus.
The reasons for this we need not discuss.
– Ms. Anonymous

Most rules for working with negatives are easy to understand if we think in terms of debits and credits. It's obvious why $(-1)(1)$ should be -1: A debit of $1 incurred one time is obviously a debit of $1. Similarly, it is obvious why $-1+1=0$: Adding a credit of $1 to a debit of $1 results in a net value of 0.

But why should $(-1)(-1)$ be 1? A "debits and credit justification" exists, but not a very good one.* A much better argument – and certainly a more artistic argument – rests on the distributive property.

Claim 1. $(-1)(-1)=1$.
Proof. We'll begin the proof in godlike fashion, by creating something from nothing.

$$\mathbf{0}=(-1)(0)=(-1)(-1+1)=(-1)(-1)+(-1)(1)=(-\mathbf{1})(-\mathbf{1})-\mathbf{1}.^{\dagger}$$

We've established that $0=(-1)(-1)-1$. Thus, whatever $(-1)(-1)$ may be, we've deduced that subtracting 1 from it leaves us with zero. Obviously, the only number with *that* property is 1, so it follows that $(-1)(-1)=1$, as claimed. ■

Since every negative number can be written as (-1) times a *positive* number, we see that, for instance,

$$(-2)(-3)=(-1)(2)(-1)(3)=(2)(3)(-1)(-1)=(2)(3),$$

where that last equals sign is justified by Claim 1. And if we wish to write up a formal proof covering the product of any two negatives, we could do so as follows.

Claim 2. (Minus times minus is *always* plus.) $(-a)(-b)=ab$ for any a and b.
Proof. Let $-a$ and $-b$ represent any two negative numbers. Then

$$(-a)(-b)=(-1)(a)(-1)(b)=ab(-1)(-1)=ab.$$

The last equals sign is justified by Claim 1. ■

Congratulations. You are now among the select few people who know *why* minus times minus is plus! By now, I trust that you've been convinced of the fundamental importance of the distributive property, which, among other things, lets us make rapid mental calculations, explains why we can multiply and factor polynomials, and helps us understand why minus times minus is plus. I'll not dwell on it any longer. It will always be there, functioning smoothly under the hood. Part of the process of learning to think mathematically is to develop an appreciation for how even simple mathematical ideas can have enormous logical ramifications. You'll have another opportunity shortly, when we turn our attention to fractions. But first, some exercises.

---

* A debit of $1 incurred "$-1$ time" is a *credit* of a $1. Hence, $(-1)(-1)=1$. Such is the argument. Alas, it involves the rather unnatural idea of incurring a debit *a negative number of times*. Yes, we can interpret this peculiar notion in such a way that we recover the expected multiplication rule, but it feels like sleight of hand. It's hard to avoid a nagging sense that the argument employs circular reasoning, quietly assuming the very thing that we wanted to demonstrate.

† The second and fourth equals signs are justified by statements in the first paragraph of this page. The third equals sign, the vital one that heralds the appearance of $(-1)(-1)$ on the stage, is justified by the distributive property.

Exercises.

18. You can now understand how to *divide* with negatives, provided you understand division itself.

Quick review: $10 \div 5$ is the number of times 5 "goes into" 10. Phrased differently, $10 \div 5$ asks the question, "5 times *what* is 10?" The answer, obviously, is 2. Similarly, $8 \div (-4)$ asks how many times (-4) goes into 8; in other words, "-4 times *what* is 8?" Of course, the answer is -2.

a) Explain why $(-8) \div (-4)$ is 2.

b) Convince yourself that minus divided by minus must *always* be plus.

c) Explain why $9 \div (-3)$ is -3.

d) Convince yourself that plus divided by minus must *always* be minus.

e) What about minus divided by plus?

19. Explain *why* $10 \div (1/3) = 30$.

20. The expressions $-8/2$, $8/-2$, and $-(8/2)$ are all equal, right? ("Of course!" you shout out, "They're all -4!") And how about $3/-5$, $-3/5$, and $-(3/5)$? They're all equal, too? ("Naturally," I hear you cry, "Each is -0.6.") Let's cut to the chase and express this algebraically:

The expressions $-a/b$, $a/-b$, and $-(a/b)$ are all equal, regardless of what a and b are.

Convince yourself that this is true.

21. Can you divide a nonzero number by zero? If so, what is the result? If not, why not?

22. Can you divide zero by a nonzero number? If so, what is the result? If not, why not?

23. What about $0/0$?

24. Explain to a friend why minus times minus is plus.

Fractions: Two Intuitive Rules That Lead to All the Others

Any fool can know. The point is to understand.
- Albert Einstein

We'll begin our study of fractions with an observation: You get the same amount of dessert by taking *two fifths* of a pie as you do by taking *one* fifth *two times*. Or, translating this statement into symbols,

$$\frac{2}{5} = 2\left(\frac{1}{5}\right).$$

Of course, there was nothing special about 2 or 5 in that last example. We might just as well have noted that *seven eighths* of a pie is the same as *one* eighth taken *seven times*, so $7/8 = 7(1/8)$. Algebra, the science of patterns, allows us to describe the pattern we are seeing here as follows.

Intuitive Fraction Rule 1: $\frac{a}{b} = a\left(\frac{1}{b}\right)$

Now for a second observation. Since 10 tenths make a whole, cutting each tenth into *thirds* would give us 30 equal parts. Thus, each *third of a tenth* is exactly $1/30$ of the whole. Or, in symbols,

$$\left(\frac{1}{3}\right)\left(\frac{1}{10}\right) = \frac{1}{30}.$$

Similarly, an eighth of a half must be a sixteenth, so $(1/8)(1/2) = 1/16$. The algebraic pattern here is:

Intuitive Fraction Rule 2: $\left(\frac{1}{a}\right)\left(\frac{1}{b}\right) = \frac{1}{ab}$

We shall take it as an axiom that our two "intuitive fraction rules" hold for *all* values of a and b.

Now for a surprise: Over the next few pages, we'll see that *every aspect* of fractional arithmetic – that subject that confuses so many people – follows logically from those two simple rules! This should encourage those who find fractions confusing. Along the way, we'll clarify quite a bit of algebra too.

Exercise.

25. The Queen of Sheba tests Solomon with a riddle: "O Great King, we all know the rule for multiplying fractions: Multiply the tops and multiply the bottoms. For example, $(2/3)(4/5) = 8/15$." Solomon nods sagely. "But *why* is this so? The wise men of my land say only that it is the gods' will, but I am told that you, a mortal man, know the explanation. What is it, good King?"

Solomon begins by explaining the two intuitive fraction rules, which the Queen allows are quite intuitive. He then proceeds to show her why $(2/3)(4/5)$ *must* be $8/15$ because of the two rules.

Your problem: Explain how Solomon did it.

Multiplying and Reducing Fractions.

From the two intuitive fraction rules, we can fully explain the well-known rule for multiplying fractions. (And if you've solved exercise 24, you've basically discovered this on your own already.)

> **Multiplication Rule for Fractions.**
>
> To multiply fractions, we multiply the tops and multiply the bottoms. Or in symbols,
>
> $$\frac{a}{b}\cdot\frac{c}{d}=\frac{ac}{bd}$$

Proof. $\frac{a}{b}\cdot\frac{c}{d} = a\left(\frac{1}{b}\right)c\left(\frac{1}{d}\right)$ (applying Intuitive Fraction Rule 1)

$= ac\left(\frac{1}{b}\right)\left(\frac{1}{d}\right)$ (reordering the multiplication)

$= ac\left(\frac{1}{bd}\right)$ (Intuitive Fraction Rule 2)

$= \frac{ac}{bd}$ (Intuitive Fraction Rule 1*) ■

Thus, for example, $(6/7)(2/3) = 12/21$. You probably also know that $12/21$ can be reduced to $4/7$. But can you explain *why* we are allowed to remove that common factor of 3 from its top and bottom? This is no idle question. People who don't understand *why* numerical fractions can be reduced (even if they know *how* to do it) tend to botch the analogous algebraic operation, which we'll discuss in the next section.

We "reduce" a fraction by throwing out factors that are common to its numerator and denominator. An example will show exactly *why* we can do this:

$$\frac{12}{21} = \frac{4\cdot 3}{7\cdot 3} = \frac{4}{7}\cdot\frac{3}{3} = \frac{4}{7}\cdot\mathbf{1} = \frac{4}{7}.^{\dagger}$$

Thus, reducing a fraction is really nothing more than eliminating a hidden factor of 1. Very simple.

Exercises.

26. Reduce the following fractions: $\frac{36}{48}$, $\frac{14}{42}$, $\frac{98}{100}$.

27. Simplify $\frac{208\cdot 144}{12\cdot 104}$ by hand, without first computing $208\cdot 144$ or $12\cdot 104$. Justify each step in your work.

[**Hint**: *Break some of the big numbers on top down to simpler factors; when you do, you'll find the numerator and denominator have some common factors. Once you remove them, you'll be able to proceed more comfortably.*]

28. Explain why $a\left(\frac{b}{c}\right) = \frac{ab}{c}$. (This is another algebraic fact that we use all the time without thinking about it.)

---

* Since equations can be read "backwards" as well as forwards, Intuitive Fraction Rule 1 can be read as $a\left(\frac{1}{b}\right) = \frac{a}{b}$.

† Notice that the second equals sign is justified by the multiplication rule for fractions.

Cancelling Above and Below the Bar

And may there be no moaning of the bar,
When I put out to sea.

- Alfred Lord Tennyson, *Crossing the Bar*

Everyone knows that we can simplify $(ab+a)/a$ by cancelling some a's, but not everyone knows *why*. Consequently, many students produce incorrect "simplifications" such as ab or $b+a$. Those who truly understand algebra never make such mistakes, for they know that cancelling is nothing more than discarding a hidden factor of 1. The logic here is the same as with reducing ordinary numerical fractions.

Example. Simplify $\frac{ab+a}{a}$.

Solution. $\frac{ab+a}{a} = \frac{a(b+1)}{a}$ (factoring a from the numerator)

$= \frac{a}{a} \cdot \frac{b+1}{1}$ (by the multiplication rule for fractions)

$= b+1.$ (because $a/a = 1$) ♦

I repeat: "Cancelling above and below the bar" is merely shorthand for discarding a hidden factor of 1. It is essential that you understand this thoroughly. Whenever you are tempted to "cancel" something from above and below a fraction bar, just ask yourself if you can separate it off as a factor of 1. If you can, you can cancel; if not, you can't cancel. That's all there is to it.

"Cancelling" above and below the fraction bar

If the top and bottom of a fraction have a common **factor**, you can "cancel" it from both places.

You cannot cancel above and below a fraction bar under any other circumstance.

You cannot cancel in other circumstances simply because there is no logical justification for doing so. Mathematics is justified by logic, even if "math" is all too often justified by the teacher's authority.

Exercises.

29. The crucial word in the box above is *factor*. Lest there be any confusion at all, please recall that
*k is called a **factor** of an algebraic expression if the expression can be written in the form $k \times (something)$.*
[For example, $3x^2$ is a *factor* of $6x^3y^2$ because we can write the latter as $3x^2(2xy^2)$.]

a) Is $2a^2b^2$ a factor of $4a^2b^2 - 18a^5b^3$? If so, explain why. If not, explain why not.
b) Provide a definition for a *term* of an algebraic expression similar to the definition of a *factor* above.
c) Can we cancel common *terms* from a fraction's top and bottom? If so, why? If not, give a counterexample.
d) Simplify the following expressions as much as possible:

$$\frac{a^2b}{ab^2}, \qquad \frac{3x+3xy}{6xyz}, \qquad \frac{5a}{5a+10b-15c}, \qquad \frac{c^2-d^2}{c-d}.$$

30. A good mathematical joke: $\frac{1\cancel{6}}{\cancel{6}4} = \frac{1}{4}$. Discuss.

The Old Multiply By 1 Trick

To reduce a fraction, we *remove* a factor of 1. Surprisingly, *introducing* a factor of 1 can also be useful. Multiplying a fraction by a judiciously disguised 1 preserves the fraction's value but changes its *form* to something more convenient. I call this "the old multiply by 1 trick". It nicely captures the spirit of algebra.*

Here's an easy example of it in action: $3/5$ is *how many* twentieths? To answer this question, we observe that multiplying our fraction's bottom by 4 would change it to 20, which is just what we want. Good news: We can get what we want if we also multiply the *top* by 4, because the net effect will be to multiply the fraction by 1, which preserves its value. Thus,

$$\frac{3}{5} = \frac{3}{5} \cdot \frac{\mathbf{4}}{\mathbf{4}} = \frac{12}{20}$$

Such is the old multiply by 1 trick. Now let's see it in a more substantial application.

Dividing Fractions

It's not yours to wonder why,
Just invert and multiply.
- Mr. Anonymous

Everyone knows the rule for dividing fractions, but few know why it works. The explanation, however, is simple if you understand the old multiply by 1 trick. Watch for it in the proof below.

Division Rule for Fractions.

To divide one fraction by another, we "invert and multiply." Or in symbols,

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

Proof. $\frac{\frac{a}{b}}{\frac{c}{d}} = \frac{\frac{a}{b}}{\frac{c}{d}} \cdot \frac{\frac{d}{c}}{\frac{d}{c}}$ (the old multiply by 1 trick)

$= \frac{\frac{ad}{bc}}{1}$ (multiplication rule for fractions)

$= \frac{ad}{bc}$ (dividing by 1 does nothing)

$= \frac{a}{b} \cdot \frac{d}{c}$ (multiplication rule for fractions) ∎

Having proved the theorem, let's consider a typical algebraic example in which we can apply it.

* There is an analogous "add 0" trick, which we use, for example, when *completing the square*, a trick you'll meet in Chapter 3. A good name for an algebra book would be *The Art of Adding Zero and Multiplying by One*.

Problem. Simplify $\dfrac{\frac{a^2-b^2}{c}}{\frac{a-b}{c^2}}$.

Solution.
$$\frac{\frac{a^2-b^2}{c}}{\frac{a-b}{c^2}} = \frac{a^2-b^2}{c}\cdot\frac{c^2}{a-b} \quad \text{(invert and multiply)}$$
$$= \frac{(a^2-b^2)c^2}{c(a-b)} \quad \text{(multiplication rule for fractions)}$$
$$= \frac{(a^2-b^2)c}{(a-b)} \quad \text{(cancelling a factor of } c \text{ from top and bottom)}$$
$$= \frac{(a-b)(a+b)c}{(a-b)} \quad \text{(factoring a difference of squares)}$$
$$= (a+b)c \quad \text{(cancelling a factor of } (a-b) \text{ from top and bottom).} \ ♦$$

Note how each step in our solution was justified by something whose validity we established earlier. Thus, if you've understood all we've done so far, such problems should pose no real difficulties.

Exercises

31. True or False? Explain *why* each true statement is true:

a) $\dfrac{3a+a^2}{3a} = 1+a^2$ b) $\dfrac{3a+a^2}{3a} = \dfrac{3+a}{3}$ c) $\dfrac{6a+6+12}{6} = a+3$ d) $\dfrac{2x+5}{10} = \dfrac{x+5}{5}$

e) $\dfrac{9-x^2}{x^2+3x} = \dfrac{3-x}{x}$ f) $\dfrac{4}{12x+8} = \dfrac{1}{3x+2}$ g) $\dfrac{2b+3c+4d}{2b+5a} = \dfrac{3c+4d}{5a}$ h) $\dfrac{2b+5a}{2b+3c+4d} = \dfrac{5a}{3c+4d}$

i) $\dfrac{a^2b^4c^{19}}{ab^3c^{20}} = \dfrac{ab}{c}$ j) $\dfrac{a^2b^4}{ab^3+ab} = \dfrac{ab^3}{b^2+1}$ k) $\dfrac{(a+b)(c+d)}{ac+ad+bc+bd} = 1$ l) $\dfrac{3x-6ux}{9x^2-12ux} = \dfrac{1-2u}{3x-4u}$

32. The expression $n!$ ("n factorial") represents the product of the first n numbers. (E.g. $4! = 4\cdot 3\cdot 2\cdot 1 = 24$.) Your problem: Simplify the expression $98!/100!$.

33. True or false: $\dfrac{\frac{a}{b}}{c} = \dfrac{a}{\frac{b}{c}}$. [Moral: *In such fractions, it's crucial to draw one fraction bar longer than the other*.]

34. Given any number a, its **reciprocal** is defined to be the number $1/a$. (For example, the reciprocal of 8 is 1/8.) Useful fact: To take the reciprocal of a *fraction*, we just flip the top and bottom. (E.g. the reciprocal of 2/3 is 3/2.)

a) Prove the preceding "useful fact", remember it, and use it whenever such expressions arise from now on.

b) Simplify the following: $\dfrac{1}{\frac{x}{y^2}}$, $\dfrac{1}{\frac{2x+y}{42}}$, $\dfrac{1}{\frac{1}{ab}}$, $\dfrac{\frac{1}{b}}{c}$. [*Careful with that last one. Remember exercise 33.*]

35. Show how to use the multiply by 1 trick to write $\dfrac{-2x}{x-1}$ in the cleaner form $\dfrac{2x}{1-x}$.

36. Simplify the following expressions as much as possible:

a) $\dfrac{x^2-4}{(x-4)(x+4)}$ b) $\dfrac{x^2-16}{(x-4)(x+4)^2}$ c) $\dfrac{\frac{x}{y}}{z}$ d) $\dfrac{\frac{a+b}{c}}{d}$ e) $\dfrac{\frac{a+b}{c}}{cd}$ f) $\dfrac{\frac{\frac{a}{b}}{c}}{bd}\cdot\dfrac{b^2}{a^2}$ g) $\dfrac{1\cdot2\cdot3\cdot4\cdot5\cdot6\cdot7\cdot8\cdot9\cdot10}{4\cdot5\cdot6\cdot7\cdot8\cdot9\cdot10}$

h) $\dfrac{3a+2b}{9a^2-4b^2}$ i) $\dfrac{3x^2-15x}{15x-3x^2}$ j) $\left[\left(\dfrac{ab}{cd}\cdot\dfrac{ac}{bd}\right)\div\dfrac{d^2}{a^2}\right]\dfrac{a^4}{d^2}$ k) $\dfrac{10x^2-10x-60}{5x+10}$ l) $\dfrac{\frac{2x^2-3x-2}{2x+1}}{\frac{x-2}{5}}$.

(A Parenthetical Aside (on Parentheses))

Algebra is generalized arithmetic: Algebraic expressions represent unspecified numbers. Often, we can view the same expression in multiple ways. Consider $3d - c$. We can view this as a representation of one number ($3d - c$), or as the difference of *two* numbers ($3d$ and c), or as a combination of *three* numbers (3, d, and c). Whenever we wish to emphasize that we are thinking of an algebraic expression as a single number – a single package – we do so by enclosing it in parentheses.

This brings us to this brief section's central idea:

> When we *combine* algebraic expressions, we think of each expression as representing **a single number**. Hence, we initially enclose each individual expression in parentheses. We can remove the parentheses in subsequent steps, provided we distribute any negatives or constant factors preceding them.

For example, if we wish to subtract $3d - c$ from $2d + c$, we form the difference $(2d + c) - (3d - c)$. The first set of parentheses isn't preceded by anything that needs to be distributed, so we can drop them when we simplify. But before we remove the second set, we must distribute that pesky negative. Carrying out these simplifications, we obtain:

$$(2d + c) - (3d - c) = 2d + c - 3d + c = 2c - d.$$

Before too long, you will reach the stage in your studies at which you *never* mistakenly omit parentheses. At that point, you'll be able to do much of this mentally, but until then, for your own sake, write it out.

Given a slightly different subtraction such as, say, $(2d + c) - \mathbf{5}(3d - c)$, we'd need to distribute not just the *negative* (i.e. -1) but $-\mathbf{5}$ to the terms in the second set of parentheses. Thus, we'd have

$$(2d + c) - 5(3d - c) = 2d + c - 15d + 5c.$$

Naturally, all that I've written in this section about parentheses applies to other grouping symbols, such as brackets, which we use to mitigate clutter in expressions such as $[3a - (a + b)][(a - b)(a + b)]$.

Exercises

37. Remove all grouping symbols and simplify:

a) $-(a - b) + (a + 2b)$

b) $(2a + b + 3c) - 2(c - a + b)$

c) $3(a - b) - 3(b - a)$

d) $(-2a - b) - [5a - (3a + 3b) - (a - b)]$

e) $x - (x - y + z) + [x - y - (z + y + x)]$

f) $a - \big[a + \big(a - (a + a)\big)\big] - a$

38. Subtract $2x^2 + x - 1$ from $-x^2 + 5x + 1$.

Adding and Subtracting Fractions

To add (or subtract) fractions with the *same denominator*, we just add (or subtract) their numerators, putting the result over their common denominator. This much is clear even to pizza-mad schoolchildren: Suppose a pizza is sliced into 10 equal pieces. A child with one piece ($1/10$ of the pie) who filches two more ($2/10$ of the whole) from inattentive classmates now obviously has $3/10$ of the whole pie.

But what if the fractions have *different denominators*? Mathematicians, being lazy by nature, like to solve new problems by transforming them into old problems that we've already figured out how to solve. Let's be lazy: To add or subtract fractions with *different* denominators, we'll use the "multiply by 1 trick" to transform them into fractions with *the same* denominator. This common denominator will need to be a multiple of both original denominators.* Here's a typical example:

Problem. $\frac{5}{6} - \frac{1}{10}$.

Solution. We need a denominator that is a multiple of both 6 and 10. The smallest such number is 30, so we'll use the "multiply by 1 trick" to change those sixths and tenths into thirtieths:

$$\frac{5}{6} - \frac{1}{10} = \left(\frac{5}{6}\cdot\frac{5}{5}\right) - \left(\frac{1}{10}\cdot\frac{3}{3}\right) \quad \text{(the "multiply by 1 trick")}$$
$$= \frac{25}{30} - \frac{3}{30} \quad \text{(by the multiplication rule for fractions)}$$
$$= \frac{22}{30} \quad \text{(subtracting fractions with the same denominator)}$$
$$= \frac{11}{15} \quad \text{(reducing the fraction).}$$

♦

If you understand that one numerical example, you understand them all. I won't belabor the point.

Algebra is generalized arithmetic, so the rules for numerical fractions work for algebraic fractions, too. For example, to subtract algebraic fractions with the same denominator, we just subtract their numerators. We must, however, take care with parentheses!

Example 1. Subtract $\frac{3d - c}{c\sqrt{b}}$ from $\frac{14d^2 + 2d + c}{c\sqrt{b}}$ and simplify the result (if possible).

Solution.
$$\frac{14d^2 + 2d + c}{c\sqrt{b}} - \frac{3d - c}{c\sqrt{b}} = \frac{(14d^2 + 2d + c) - (3d - c)}{c\sqrt{b}} \quad \text{(note the parentheses!)}$$
$$= \frac{14d^2 + 2d + c - 3d + c}{c\sqrt{b}}$$
$$= \frac{14d^2 - d + 2c}{c\sqrt{b}}.$$

♦

Be sure you understand those parentheses in the first step: When subtracting two fractions with a common denominator, the new numerator is the difference of the two original numerators; each of these was an algebraic expression, so to subtract them, we had to think of each *as a single package.* Consequently, we had to enclose each of them in parentheses. Had we omitted the parentheses, we would have ended up with the wrong numerator, and hence the wrong answer.

* And to simplify our subsequent work, we usually try to find the smallest possible common denominator.

To add or subtract algebraic fractions with different denominators, we'll use – just as you'd expect – the "multiply by 1 trick" to give the fractions a common denominator. For example,

Example 2. Add and simplify: $\frac{3}{2a^2b} + \frac{7}{6ab}$.

Solution. For a common denominator, we need a multiple of $2a^2b$ and $6ab$. The "smallest" such denominator is $6a^2b$. Using the "multiply by 1 trick," we'll convert to this new denominator.

$$\begin{aligned}\frac{3}{2a^2b} + \frac{7}{6ab} &= \left(\frac{3}{2a^2b} \cdot \frac{\mathbf{3}}{\mathbf{3}}\right) + \left(\frac{7}{6ab} \cdot \frac{\boldsymbol{a}}{\boldsymbol{a}}\right) && \text{(the "multiply by 1 trick")} \\ &= \frac{9}{6a^2b} + \frac{7a}{6a^2b} \\ &= \frac{9+7a}{6a^2b}\end{aligned}$$

♦

After you've become comfortable adding and subtracting fractions with different denominators, you need not write out every step. In practice, we usually condense the process as follows:

> **Addition and Subtraction Rule for Fractions.**
>
> We find the denominator (a so-called "common denominator") as follows:
>
> Take any multiple of the given denominators.*
>
> We find the numerator as follows:
>
> Multiply each given numerator by the factor that will turn its denominator into the *common* denominator. These products will be the terms in the new numerator.

If we re-do the previous example using this shortcut addition rule, we'll have somewhat less to write. Taking $6a^2b$ as our common denominator, the addition rule quickly tells us that

$$\frac{3}{2a^2b} + \frac{7}{6ab} = \frac{3(\mathbf{3}) + 7(\boldsymbol{a})}{6a^2b} = \frac{9+7a}{6a^2b}.$$

Beginning students frequently make mistakes when *subtracting* fractions – usually because they omit necessary parentheses. Be especially careful when subtracting.

Example 3. Subtract and simplify: $\frac{3}{x-5} - \frac{2x-1}{x+5}$.

Solution.

$$\begin{aligned}\frac{3}{x-5} - \frac{2x-1}{x+5} &= \frac{[3(\boldsymbol{x+5})] - [(2x-1)(\boldsymbol{x-5})]}{(x-5)(x+5)} && \text{(subtraction rule)} \\ &= \frac{[3x+15] - [2x^2-11x+5]}{(x-5)(x+5)} && \text{(distributing within the brackets)} \\ &= \frac{3x+15-2x^2+11x-5}{x^2-25} && \text{(distributing a minus; difference of squares)} \\ &= \frac{-2x^2+14x+10}{x^2-25} && \text{(combining like terms)}\end{aligned}$$

♦

* Again, using the *smallest* such multiple will make your subsequent work neater.

Exercises

39. Carefully explain why $\frac{7}{6} + \frac{3}{10} = \frac{22}{15}$.

40. Simplify the following expressions.

a) $\left(\frac{\frac{2}{3}+\frac{3}{5}}{\frac{7}{11}-\frac{1}{2}} + \frac{3}{4}\right)\frac{1}{2}$ b) $-\frac{3}{7} - \frac{2}{3}\left(\frac{1}{2} - \frac{2}{5}\right)$ c) $1 - \left[\left(\frac{-\frac{2}{3}}{\frac{8}{9}}\right) \div \left(\frac{2}{3} - \frac{-8}{7}\right)\right]$

41. In Example 2 above, we took $6a^2b$ as our common denominator. Suppose we had used $12a^3b^2$ instead. Would this have changed the result? Work it out that way and see.

42. Express each of the following as a single fraction, and simplify as much as possible:

a) $\frac{3}{5} + \frac{a-3}{5}$ b) $\left(\frac{3}{5}\right)\left(\frac{a-3}{5}\right)$ c) $\frac{a+b}{15b} - \frac{a-b}{15b}$ d) $\frac{5x}{y} + \frac{y}{5x}$

e) $\frac{3}{4ax^2} + \frac{x}{2a}$ f) $\frac{2x-3}{x-2} - \frac{x-4}{x-2}$ g) $\frac{2}{xy} + \frac{3}{yz} + \frac{4}{xz}$ h) $\frac{1}{x} - \frac{2}{x^2} + \frac{3}{x^3} - \frac{4}{x^4}$

i) $\frac{2}{x(x+1)} - \frac{(x-2)}{x(x-1)}$ j) $\frac{3}{a-3} - \frac{2}{a}$ k) $3 - \frac{1}{2x+1}$ l) $1 - \frac{1+x}{1-x}$

m) $\frac{1}{x+h} - \frac{1}{x}$ n) $\frac{1}{x} - \frac{2}{x^2} + \frac{3}{x(x-1)} - \frac{4}{(x-1)^2}$

Last Words and One Nasty Example

You can now add, subtract, multiply, or divide any two algebraic fractions. No lingering mysteries remain; you are fully initiated. Should you ever need to combine seventeen fractions linked in all sorts of intricate arithmetical ways, you have all the knowledge necessary to do it. All you must do is slow down and take each piece in order. It may be tedious, but it shouldn't be difficult.

To illustrate my point, I'll offer one last example. If you have understood everything so far in this chapter, it should pose no conceptual difficulties, even though it is rather involved.

Problem. Simplify the following ugly expression as much as possible: $\dfrac{\frac{2}{x-2} - \frac{x+1}{x+2}}{\frac{3x-7}{x^2-4}} + \dfrac{\frac{2}{x^2}}{-5}$.

Solution. Before we begin hacking away at this overgrown mess, let's consider it from a distance to convince ourselves that, despite its ugliness, this is something we can handle. Begin by observing that the expression has two terms. True, the first one looks nasty ("Ugh! Fractions *in a fraction*!"), until we realize that when we subtract the two fractions in its numerator, they'll obviously combine into a *single* fraction, which we can then divide by the fraction in the left term's denominator. The result of that division will obviously be… a single fraction. Thus, when the dust settles, we'll have rewritten the whole ugly first term as one fraction. All the "fractions within fractions" will be gone. As for the second term, you can probably simplify it in your head, turning it into one fraction, too. Then, all that will remain is to add two ordinary algebraic fractions – an easy task.

Having watched the preliminary dumbshow, let's proceed to the actual details of calculation, confident that, at least in outline, we already know how matters will work out.

The first term's numerator is

$$\frac{2}{x-2} - \frac{x+1}{x+2} = \frac{[2(x+2)] - [(x-2)(x+1)]}{(x-2)(x+2)} \qquad \text{(note the brackets!)}$$

$$= \frac{[2(x+2)] - [(x-2)(x+1)]}{x^2-4} \qquad \text{(difference of squares)}$$

$$= \frac{[2x+4] - [x^2-x-2]}{x^2-4} \qquad \text{(distributing within the brackets)}$$

$$= \frac{2x+4-x^2+x+2}{x^2-4} \qquad \text{(removing brackets, distributing the negative)}$$

$$= \frac{-x^2+3x+6}{x^2-4} \qquad \text{(combining like terms).}$$

Now that we've simplified the first term's numerator, we'll divide it by the first term's denominator. Doing so, we find that the entire first term simplifies to

$$\frac{\frac{-x^2+3x+6}{x^2-4}}{\frac{3x-7}{x^2-4}} = \frac{-x^2+3x+6}{x^2-4} \cdot \frac{x^2-4}{3x-7} \qquad \text{(division rule for fractions)}$$

$$= \frac{(-x^2+3x+6)(x^2-4)}{(x^2-4)(3x-7)} \qquad \text{(multiplication rule for fractions)}$$

$$= \frac{-x^2+3x+6}{3x-7} \qquad \text{(cancelling } (x^2-4) \text{ above and below the bar).}$$

Having reduced the first term to something manageable, we can tackle the original problem:

$$\frac{\frac{2}{x-2}-\frac{x+1}{x+2}}{\frac{3x-7}{x^2-4}}+\frac{\frac{2}{x^2}}{-5}=\frac{-x^2+3x+6}{3x-7}+\frac{\frac{2}{x^2}}{-5}$$ (by our work on the first term above)

$$=\frac{-x^2+3x+6}{3x-7}-\frac{2}{5x^2}$$ (division of fractions, second term)

$$=\frac{[(-x^2+3x+6)(5x^2)]-[2(3x-7)]}{(3x-7)(5x^2)}$$ (subtraction rule for fractions)

$$=\frac{[-5x^4+15x^3+30x^2]-[6x-14]}{15x^3-35x^2}$$ (a great flurry of multiplication)

$$=\frac{-5x^4+15x^3+30x^2-6x+14}{15x^3-35x^2}.$$ (removing brackets upstairs) ♦

There were many steps in that last problem, but each was simple. Occasional mistakes in such problems are inevitable, but bear in mind that there are mistakes and *mistakes*. Accidentally writing $3\times 3=6$ in the midst of a larger problem is wrong, but it presumable isn't a conceptual error. On the other hand, omitting crucial parentheses, forgetting to distribute negatives, or botching the multiply by 1 trick are serious mistakes that most likely stem from conceptual misunderstandings. If you intend to take further mathematics courses, you need to clear up any and all such misunderstandings immediately. Those courses will give you plenty of new material to think about; if you are still struggling with basic algebra at that stage, you won't see the forest for the trees.

Exercises

43. Simplify as much as possible:

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

44. True or false (explain your answers):

a) $-3^2=9$

b) $-x$ always represents a negative number.

c) $(-3)^2=-9$

d) $a-b=-(b-a)$.

e) $\frac{(2x+1)[(3x-7)+(x^2+1)]}{(2x+1)(x^3+8)}=\frac{(3x-7)+(x^2+1)}{x^3+8}$

f) $\frac{(2x+1)(3x-7)+(x^2+1)}{(2x+1)(x^3+8)}=\frac{(3x-7)+(x^2+1)}{x^3+8}$

g) $\frac{a+b}{c+d}=\frac{(a+b)}{(c+d)}$

45. Express as a single fraction – and simplify as much as possible:

a) $\dfrac{3}{x-2}+\dfrac{1}{2-x}$ [**Hint**: *You may find Exercise* 44d *useful*.]

b) $\dfrac{\frac{1}{x+h}-\frac{1}{x}}{h}$

c) $\dfrac{(x+h)^2-x^2}{h}$

d) $\dfrac{(a-b^2)(a+b^2)}{a^2-b^4}\cdot\dfrac{a+\frac{1}{a}}{a}$

e) $\left[\dfrac{5x+4}{x+1}-\dfrac{-3x^2+9x+4}{(x+1)^2}\right]\div\left(\dfrac{4x^3}{(x+1)^2}\right)$

f) $1+\dfrac{1}{x-1}$

g) $\dfrac{1}{1+\frac{1}{x-1}}$

h) $\dfrac{1}{1-\frac{1}{1-\frac{1}{1+\frac{1}{x}}}}$

i) $\dfrac{\frac{1}{a+b}+\frac{1}{a-b}}{\frac{3a}{2}-\frac{4a}{3}}$

j) $\left(\dfrac{x^2}{1+x^2}\cdot\dfrac{(1+x^2)(-x^2)}{x^4}\right)(a-b)$

46. Early in the chapter, I claimed that all of fractional arithmetic follows logically from "two intuitive rules":

$$\frac{a}{b}=a\left(\frac{1}{b}\right)\quad\text{and}\quad\left(\frac{1}{a}\right)\left(\frac{1}{b}\right)=\frac{1}{ab}.$$

Now that you know all the rules of fractional arithmetic, it's worth reexamining my claim. I explicitly pointed out in the text how the multiplication rule for fractions follows directly from the two intuitive rules. What about cancelling factors above and below the bar? Well, we saw that this operation is really just a matter of detaching a hidden factor of 1, like so:

$$\frac{6ab}{3bc}=\frac{3b\cdot 2a}{3b\cdot c}=\left(\frac{3b}{3b}\right)\left(\frac{2a}{c}\right)=1\left(\frac{2a}{c}\right)=\frac{2a}{c}.$$

The "detachment" (at the second equals sign) justified by the multiplication rule… which was built on the two intuitive rules. (The other steps are justified by completely obvious facts such as "multiplying by 1 doesn't change anything" or "anything divided by itself is 1".) Thus cancellation above and below the bar is ultimately a logical consequence of the two intuitive rules.

a) Convince yourself that the "multiply by 1 trick" is ultimately justified by the two intuitive rules.

b) Do the same for the division rule for fractions.

c) Do the same for the addition and subtraction rule for fractions.

d) Congratulate yourself: You've now seen that the whole vexed subject of fractions is actually quite simple, built up from just a few intuitive rules and the logical consequences thereof.

Chapter 6
Transformations in Coordinate Geometry

Transformations

As Gregor Samsa awoke one morning from uneasy dreams,
he found himself transformed...

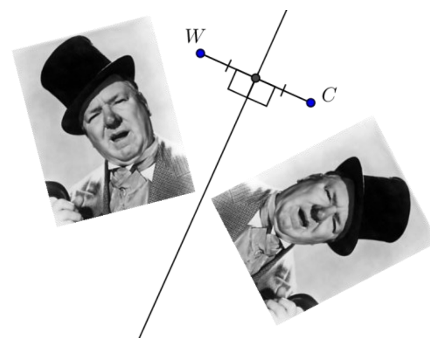
- Franz Kafka, "The Metamorphosis"

Our main theme for this chapter will be understanding relationships between *geometric* transformations [such as shifting a graph to the right by 3 units] and *algebraic* transformations [such as replacing every x in an equation by $(x-3)$]. To keep matters simple, we'll consider just three geometric transformations: reflections, shifts, and stretches.

Reflections

Here sad self-lovers saw in tragic error
Some lovely other or another sky;
In your reversing yet unlying mirror
I saw I was I.

- John Hollander, "At a Forest Pool"



Apart from reveling in Hollander's narcissistic palindrome, our main object in this brief section is to understand what happens on a point-by-point basis when we reflect a geometric object across a line. In the figure at right, a possibly recognizable personage is depicted, along with his reflection across a line.

When we reflect a point across a line, it ends up just as far from the line as it had originally been, but on the opposite side. The segment joining a point and its reflected image (segment WC, for example) is always *perpendicular* to the line of reflection. This is the case, you may observe, with the segment joining the jolly red noses of Mr. Fields and his doppelgänger. The nearer a point lies to the reflecting line, the shorter the distance it will move when reflected. Points *on* the line don't move at all.

Since we will be concerned almost exclusively with reflections over the axes, we shall use the phrase *vertical reflection* to refer specifically to a reflection over the x-axis (such a reflection is "vertical" because it moves points vertically), and *horizontal reflection* for a reflection over the y-axis.

Exercises.

1. Draw the graph of $y = x^2$ and its vertical reflection.

2. Draw the graph of $y = x^2$ and its horizontal reflection.

3. Draw the graph of $y = x^2$ and its reflection across the line $y = 1$.

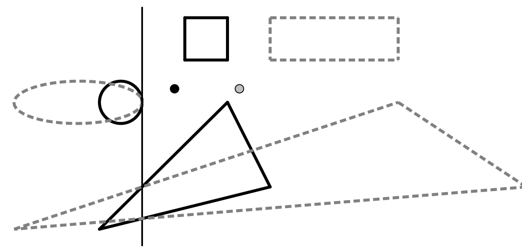
4. Draw the graph of $(x-1)^2 + (y-1)^2 = 4$ and its horizontal reflection.

Stretches

What, will the line stretch out to the crack of doom?

- Macbeth (Act 4, Scene 1)

When we stretch, we stretch relative to some fixed line. Once we've specified the line, a "stretch factor" is needed to specify the stretch's strength: When the stretch factor is k, the stretch sends each point in the plane to a new location exactly k times as far from the fixed line as it had been. Points on the fixed line itself do not move.

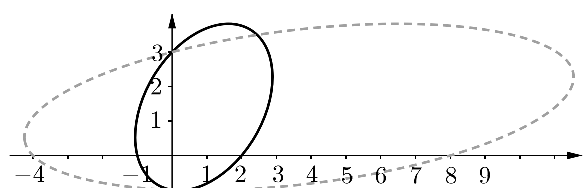


In the figure at right, I've superimposed the "before" state (in solid black) and the "after" state (in dashed grey) of a horizontal stretch relative to the depicted vertical line by a factor of 3. Scrutinize it carefully until you thoroughly understand it. Observe that a stretch, in contrast to a reflection, changes a figure's *shape*. A reflected circle remains circular; a stretched circle does not.

If we stretch a graph by a factor of $1/2$, the distance of each of its points to the fixed line is halved. Such a "stretch" is really a *compression*. This will happen whenever we "stretch" by a factor less than 1. If we view the previous figure with new eyes, seeing the dashed grey figures as the "before" state and the solid black figures as the "after" state, then it represents a horizontal stretch by a factor of $1/3$.

Since we will be concerned almost exclusively with stretches relative to the axes, we'll use the phrase *vertical stretch* to refer specifically to a stretch relative to the x-axis (we call it "vertical" because it moves points vertically), and *horizontal stretch* for a stretch relative to the y-axis.

Exercises.



5. The figure at right shows a horizontally stretched ellipse. What stretch factor will transform the black ellipse into the grey one? What stretch factor will transform the grey ellipse into the black one?

6. Draw the graph of $y = \sqrt{x}$ and then stretch it vertically by a factor of 4. Label some points (with their coordinates) on both graphs.

7. Draw the graph of $y = x^2$ and then stretch it horizontally by a factor of $1/2$. Label some points on both graphs. Could the same effect have been achieved by stretching the original graph *vertically*? If so, then by what factor? If not, why not?

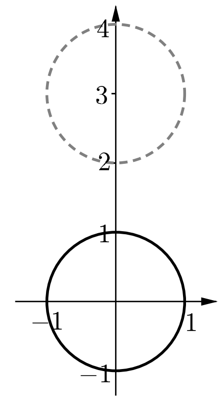
8. Draw the graph of $x^2 + y^2 = 1$ and then stretch it horizontally by a factor of 3. Label some points on both graphs. Could the same effect have been achieved by stretching the original graph *vertically*? If so, by what factor? If not, why not?

9. Draw the graph of $(x-2)^2 + y^2 = 1$ and then stretch it horizontally by a factor of $1/2$. Find the coordinates of the compressed circle's topmost point, and those of its leftmost point.

Shifts

Shift that fat ass, Harry.
But slowly, or you'll swamp the damned boat.

- George Washington.*



Last (and least) are shifts, which are easy to understand: To *shift* a graph, we simply move each of its points a specified distance in a specified direction. For example, in the figure at right, if we shift the solid black circle *vertically* by 3 units, it will occupy the dashed grey circle's position.

We will concern ourselves exclusively with horizontal and vertical shifts, since any shift can be broken into horizontal and vertical components. We'll often let the algebraic sign (+ or –) do the talking for us when we wish to distinguish between right and left, or up and down. For instance, a vertical shift by 5 units will refer to a shift *up*, whereas a vertical shift by -5 units will signify a shift *down*. (Similarly, a negative horizontal shift is a shift to the left.)

Exercises.

10. Draw the graph of $y = |x|$, and then shift it vertically by -2 units. Label some points on both graphs.

11. Draw the graph of $y = \sqrt[3]{x}$, then shift it horizontally by 3 units. Label some points on both graphs.

12. In general, **geometric transformations are non-commutative.** That is, the *order* in which we carry them out often matters a great deal. Convince yourself of this by comparing the results of the following transformations:

a) Begin with the graph of $y = x^2$. Shift it horizontally by 2 units, then reflect it over the y-axis.

b) Begin with the graph of $y = x^2$. Reflect it over the y-axis, then shift it horizontally by 2 units.

13. Geometric transformations sometimes *do* commute. Convince yourself of this by comparing the results of the following transformations:

a) Begin with the graph of $y = x^2$. Shift it horizontally by 2 units, then reflect it over the x-axis.

b) Begin with the graph of $y = x^2$. Reflect it over the x-axis, then shift it horizontally by 2 units.

14. Suppose you must do the following three things to the graph of $y = x^3$, but the order in which to do them isn't specified: shift right by 1, reflect over the y-axis, stretch horizontally by a factor of 2.

a) In how many different orders can these three transformations be applied?

b) How many *different* graphs result from the different orders?

* Washington spoke these words to Henry Knox while entering the vessel in which they and others crossed the icy Delaware River on the night of December 25, 1776 to conduct a surprise attack against a band of German mercenaries.

The Transformation Table

> And it came to pass, as soon as he came nigh unto the camp, that he saw the calf, and the dancing: and Moses' anger waxed hot, and he cast the tables out of his hands, and brake them beneath the mount.
>
> - Exodus 32:19

Now that you understand what reflections, stretches, and shifts are, we can discuss how to use them in coordinate geometry. Because students sometimes find this material difficult, I shall begin at the end. I shall declaim. I shall tell you – with the grave and irrefutable voice of authority, accompanied by the solemn majesty of a table – the precise correspondences between these geometric transformations and their algebraic analogues. Provided you will pause in your revels about the golden calf, I shall deign to show you *how* without telling you *why*.

Of course, the table that I shall show you was not actually handed down to Moses on Mt. Sinai, and you should not, in the long-run, accept it as though it had been. You should demand an explanation. And I will provide one *after* you've developed a feel for using the table. As you'll see, the explanation is not terribly difficult, but it does require some new notation and a slightly new way of thinking about graphs. All this in due time. For now, here is the holy transformation table.

	Horizontal	Vertical
Reflection	Substitute $(-x)$ for each x	Substitute $(-y)$ for each y
Stretch by a factor of k	Substitute $\left(\frac{1}{k}x\right)$ for each x	Substitute $\left(\frac{1}{k}y\right)$ for each y
Shift by k units	Substitute $(x-k)$ for each x	Substitute $(y-k)$ for each y

The geometric transformations are on the table's margins; the algebraic transformations are in its body. Horizontal and vertical transformations correspond, respectively, to substitutions for x and y. There are only three types of substitution, and all are easily memorized. Just as you shouldn't need to consult Moses' tables to remind yourself of their proclamations (*Oh*, *I can't recall… Was it thou* ***shalt*** *commit adultery*?*), you shouldn't need to consult the transformation table to know what it says about horizontal shifts. Memorize the table right away.

Our first example will reaffirm a result we've established by other means, and should therefore give you some confidence that the transformation rules do indeed work as advertised.

Example 1. If we shift the unit circle *up* by 5 units, what will its new equation be?

Solution. According to our transformation table, a vertical shift by 5 units corresponds to a substitution of $(y-5)$ for y. Making this substitution in the equation of the unit circle, $x^2+y^2=1$, we obtain the *shifted* circle's equation, $\boldsymbol{x^2+(y-5)^2=1}$. ♦

* On account of a typesetter's error, a 1631 edition of the Bible published in London actually did contain the commandment "Thou shalt commit adultery". King Charles I was not amused, and a result, the publishers lost their printers' licenses. (Rare book collectors refer to this particular edition as "The Wicked Bible".)

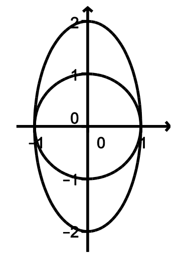
Recall that we use $+$ and $-$ to distinguish between right and left shifts (or between up and down shifts). For example, a shift *down* by 5 units is a vertical shift by -5 units. To shift a graph this way, the transformation table tells to substitute $\big(y-(-5)\big)$, that is, $(\boldsymbol{y+5})$, for each y in its equation.

Example 2. If we shift the unit circle 7 units to the *left*, what will its new equation be?

Solution. The table tells us that shifting left by 7 units corresponds to substituting $(x+7)$ for x. Putting this into the unit circle's equation yields the *shifted* circle's equation: $\boldsymbol{(x+7)^2+y^2=1}$. ♦

Now let's try some stretches.

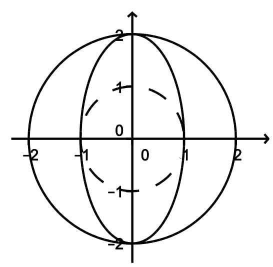
Example 3. Find the equation of the graph that results from stretching the unit circle vertically by a factor of 2.



Solution. Vertically stretching by a factor of 2 corresponds to substituting $(y/2)$ for y. Substituting this into the unit circle's equation yields the equation of the *stretched* circle (called an *ellipse*): $\boldsymbol{x^2+\big(y^2/4\big)=1}$. ♦

Suppose we stretch a circle vertically by some factor (as in the previous example), and then stretch the resulting ellipse horizontally by the same factor. Would the result be a circle again? This seems plausible, but perhaps it isn't so. We need not wonder for long. Coordinate geometry to the rescue!

Example 4. If we stretch the graph we obtained in the previous example horizontally by a factor of 2, will the result be a circle?



Solution. We already found the ellipse's equation in the previous example. If we stretch the ellipse horizontally, its equation changes; the transformation table tells us that we'll get its new equation if we substitute $x/2$ for x in the ellipse's equation. Doing so yields $(x^2/4)+(y^2/4)=1$.

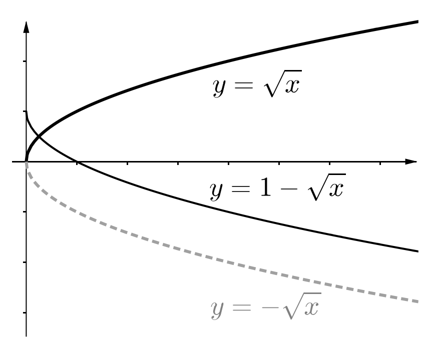
Clearing fractions, this becomes $\boldsymbol{x^2+y^2=4}$, which we recognize. This is indeed the equation a *circle*: the circle of radius 2 centered at the origin. ♦

This last example suggests that we can transform the *unit* circle into *any* circle in four "moves" (or less): two stretches to attain the desired radius, and two shifts to move its center to the desired location.

Example 5. If we reflect the graph of $y=\sqrt{x}$ over the x-axis, then shift the result up by 1 unit, what will the final graph's equation be?

Solution. *Vertical* reflection corresponds to substituting $-y$ for y. Thus, the equation of the reflected graph (dashed, grey at right) is

$$-y=\sqrt{x}.$$



Isolating y gives us an equivalent form,

$$y=-\sqrt{x}.$$

Next, to shift this up 1 unit, we substitute $(y-1)$ for y, obtaining

$$y-1=-\sqrt{x}.$$

Isolating y again, we obtain the final graph's equation:

$$\boldsymbol{y=1-\sqrt{x}}. \quad ♦$$

Exercises.

15. State the algebraic substitution corresponding to each of the following geometric transformations:

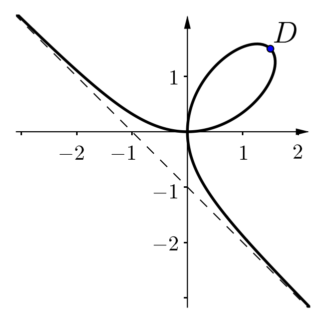
a) shift right by 8 b) shift left by 8 c) shift down by 8 d) shift up by 8 e) shift left by $1/2$
f) vertical stretch by a factor of 6 g) vertical stretch by a factor of $1/3$ h) shift right by 4
i) horizontal stretch by a factor of 10 j) horizontal stretch by a factor of $1/10$
k) horizontal stretch by a factor of 17 l) shift up by 2 m) vertical stretch by a factor of $3/2$
n) vertical stretch by a factor of $3/16$ o) reflection over the x-axis p) shift right by π
q) vertical stretch by a factor of $2/17$ r) reflection over the y-axis s) shift down by 1

16. Draw the graph of $y = 1/x$, then shift it horizontally by 3 units. Label some points on each graph. What is the shifted graph's equation?

17. Draw the ellipse $x^2 + (y^2/4) = 1$ (which we met in example 3), then shift it down by 3 units. Label some points on each graph. What is the shifted ellipse's equation?

18. Draw the graph of $(x-1)^2 + (y-1)^2 = 2$ and its reflection over the y-axis. Find the equation of the reflected graph. Finally, find the coordinates of the points where the reflected graph crosses the axes.
[When simplifying the equation, remember that $(-a-b)^2 = \big(-(a+b)\big)^2 = (a+b)^2$.]

19. The graph at right, called the *Folium of Descartes*, played a small but important role in the history of coordinate geometry – and calculus. (You can read about this online if you are curious.) Its equation is $x^3 + y^3 = 3xy$.



a) The dotted line is not part of the Folium; it is the Folium's *asymptote*: the line it approaches but never touches. Find the asymptote's equation.

b) Draw the reflection of the Folium over the x-axis. What is its equation?

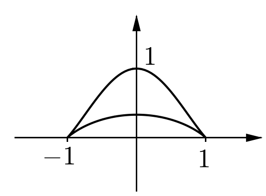
c) Find the coordinates of point D.
[**Hint**: *It is the intersection of the Folium and a certain line through the origin...*]

d) If we were to shift the Folium horizontally until D lies on the y-axis, what would its equation in this new position be?

e) If we were to shift the Folium so as to put D at the origin, what would its equation in its new position be?

20. The graph at right is called a *cocked hat* (or a *bicorn*). Its equation is

$$y^2(1-x^2) = (x^2 + 2y - 1)^2.$$



a) Where does the cocked hat cross the y-axis?

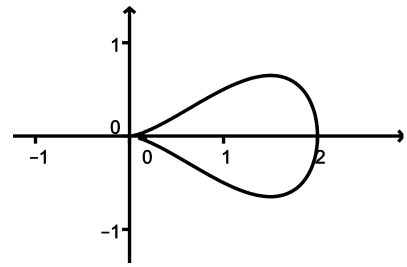
b) The Mad Hatter wants a cocked hat with "corners" lying at $(\pm 2, 0)$, but whose peak still lies at $(0,1)$. Draw it and find its equation.

c) Next, the Mad Hatter creates a cocked hat with corners at $(2,0)$ and $(3,0)$ and peak at $(5/2\,, 1)$. Draw it and find its equation.
[**Hint**: *This will require* ***two*** *successive transformations.*]

d) Finally, being a truly *mad* hatter, he creates an inverted cocked hat, whose corners are at $(-1,0)$ and $(1,0)$ and whose peak – or trough, as the case may be here – is at $(0, -5000)$. Find its equation, and determine precisely where this crazy hat crosses the y-axis.
[**Hint:** *You need not solve an equation to find the points of intersection. Just keep track of the coordinates of the original intersection points as you transform the graph of the original cocked hat*.]

21. The equation of the graph shown at right, a *pear-shaped quartic*, is

$$x^4 - 2x^3 + 4y^2 = 0.$$



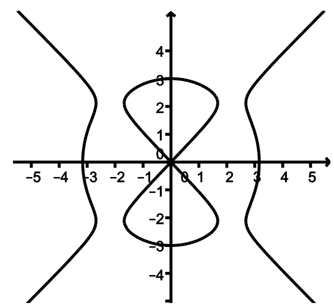
a) Observe that if we substitute $-y$ for y in the curve's equation, the equation itself does not change. This *algebraic* fact corresponds to a *geometric* property of the curve. What is it?

b) Mr. Knickerbocker is writing a computer program in which the pear will appear onscreen with its tip at a given point to alert the user to something important happening there. If the pear is to appear on screen with its tip at (x_0, y_0), what will its equation be? (The pear always remains the same size and always has its tip on the left.)

c) After using the equation you found for him in part (b), Mr. Knickerbocker runs his program and finds a few bugs. For instance, when (x_0, y_0) is near the right hand edge of the screen, the pear is cut off by the screen's edge, and is thus invisible as far as the user is concerned. To fix this, Mr. Knickerbocker wants to revise his program so that pear will sometimes point in the *opposite* direction (i.e. with its tip on the *right*). If the pear is to appear on screen tip rightmost and pointing at (x_0, y_0), what will its equation be?

22. The graph at right is an example of a *Devil's curve*, a class of curves first studied in 1750 by Gabriel Cramer. The equation of this particular specimen is

$$-x^4 + 10x^2 + y^4 - 9y^2 = 0.$$



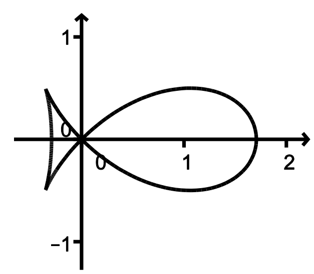
a) Find the points at which this Devil's curve crosses the axes.

b) If we wished to compress the graph vertically so that it crossed the y-axis at $(0, \pm 2)$, what substitution would we have to make in its equation? What would the new equation be?

c) If we wished to compress the result of part (b) horizontally so that it crossed the x-axis at $(\pm 2, 0)$, what substitution would we have to make? What would the new equation be?

23. The equation of the *fish curve* at right is

$$(2x^2 + y^2)^2 - 2\sqrt{2}x(2x^2 - 3y^2) + 2(y^2 - x^2) = 0.$$



a) Find the points at which the fish curve crosses the x-axis.

b) If we were to replace every y in the equation with $(5y)$, what would happen to the fish?

c) If we were to replace every x in the fish's equation with $(-x/2)$, what would happen to the fish? [**Hint**: *You can achieve the same net algebraic effect in two steps – replace each x by $x/2$, then replace each x by $-x$.*]

24. We can compress the transformation table by packing the stretches and reflections together:

	Horizontal	Vertical
Stretch by a factor of k (with a reflection if $k < 0$)	Substitute $\left(\frac{1}{k}x\right)$ for each x	Substitute $\left(\frac{1}{k}y\right)$ for each y
Shift by k units	Substitute $(x - k)$ for each x	Substitute $(y - k)$ for each y

Explain *why* this is valid. After doing so, feel free to use the table in this compressed form.

Preliminaries to the Proof

Now that you know how the transformation rules work, it's high time you understand *why* they work. I will explain this on the next page – after introducing a few preliminary ideas here.

A *function of* ***two*** *variables* is a rule that unambiguously turns *ordered pairs of numbers* into numbers. The notation for a function of two variables is exactly what you'd expect.

Example 1. If $f(x, y) = 2x^2 + y$, find the values of $f(3, -7)$ and $f(-7, 3)$.

Solution. $f(3, -7) = 2(3)^2 - 7 = 11$.

$f(-7, 3) = 2(-7)^2 + 3 = 101$. ♦

Functions of two variables provide the right language for discussing *equations* in two variables, because we can write any equation in two variables in the form $f(x, y) = 0$ simply by pushing all its terms over to the left-hand side.

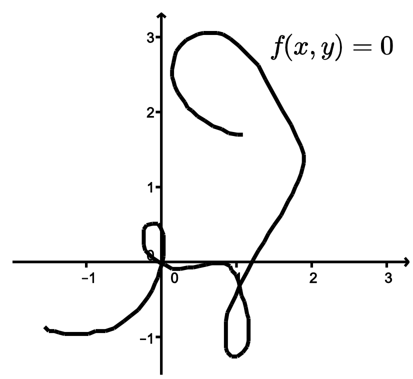
Example 2. Rewrite the equation $x^2 + (y - 1)^2 = 4$ in the form $f(x, y) = 0$.

Solution. Pushing all terms to the left, we obtain an equivalent equation with the required form:

$$\overbrace{x^2 + (y - 1)^2 - 4}^{f(x,y)} = 0$$ ♦

In the previous exercise set, I presented the equations of the fish curve, Devil's curve, and pear-shaped quartic in this $f(x, y) = 0$ form.

There's no real advantage, except arguably a kind of tidiness, gained by putting any *particular* equation in the form $f(x, y) = 0$. However, this form is extremely useful whenever we wish to discuss two-variable equations *in general*. Thus, whenever we wish to discuss the abstract idea of a two-variable equation, or prove a theorem that holds for *all* graphs, we often begin by carelessly drawing a squiggle (such as the one at right) meant to represent any old curve, which we then think of as the graph of some equation $f(x, y) = 0$, with the function f left unspecified.

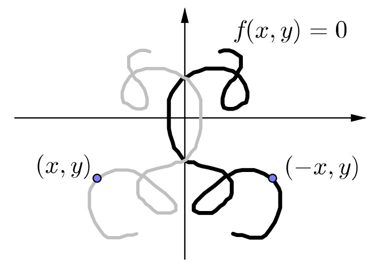


The Explanation at Last

We need to justify all six entries in the transformation table. If you understand the proof of the first entry, you'll understand all the others, as they are all cut from the same logical pattern. This being the case, I'll leave most of the six as exercises for you.

Claim 1. If we reflect a graph over the y-axis, we can obtain the reflected graph's equation by replacing each x the original graph's equation by $-x$.

Proof. Consider an arbitrary graph, such as the black one at right, and reflect it over the y-axis (so that it produces the grey curve). Let $f(x,y)=0$ be the original graph's equation. We must now find the *reflected* graph's equation: an equation that is satisfied by the coordinates of all points on the *reflected* graph.



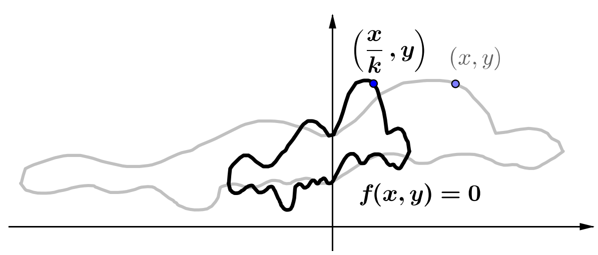
Let (x,y) be a variable point on the grey reflected graph. (Its variability allows it to represent *every* point on the graph.) "Undoing" the reflection would send this point back to $(-x,y)$, and this latter point must lie, of course, on the original black graph. Therefore $(-x,y)$ satisfies that original graph's equation.

That is, for every point (x,y) on the reflected graph, we know that $\boldsymbol{f(-x,y)=0}$.

It follows that this last equation **is** the reflected graph's equation. Happily, we can obtain it from the original graph's equation, $f(x,y)=0$, by simply substituting $-x$ for x. ∎

Claim 2. If we stretch a graph horizontally by a factor of k, we can obtain the reflected graph's equation by replacing each x the original graph's equation by x/k.

Proof. Consider an arbitrary graph, such as the black one in the figure at right, and stretch it horizontally (to produce the grey curve). Let $f(x,y)=0$ be the original graph's equation. We now seek the *stretched* graph's equation: an equation that is satisfied by the coordinates of all points on the *stretched* graph.



Let (x,y) be a variable point on the stretched graph. "Undoing" the stretch would send this point back to $(x/k\,,y)$, and this latter point must, of course, lie on the original black graph. Therefore, $(x/k\,,y)$ satisfies that original graph's equation.

That is, for every point (x,y) on the stretched graph, we know that $\boldsymbol{f(x/k,\ y)=0}$.

It follows that this last equation **is** the stretched graph's equation. We can thus obtain it from the original graph's equation, $f(x,y)=0$, by merely substituting x/k for x. ∎

If you understood these two proofs, you will have no trouble constructing the other four on your own. Essentially, we obtain the transformed curve's equation by "undoing" the transformation algebraically, then substituting the "un-transformed" coordinates back into the original curve's equation.

Exercises.

25. Prove that if we shift a graph horizontally by k units, we can obtain the shifted graph's equation by substituting $(x-k)$ for each x in the original graph's equation.

26. Prove that if we reflect a graph over the x-axis, we can obtain the reflected graph's equation by replacing each y the original graph's equation by $-y$.

27. Prove that if we stretch a graph vertically by a factor of k, then we can obtain the reflected graph's equation by replacing each y the original graph's equation by y/k.

28. Prove that if we shift a graph vertically by k units, we can obtain the shifted graph's equation by replacing each y the original graph's equation by $(y-k)$.

29. In the proof of Claim 1 above, we showed that $f(-x,y)=0$ is satisfied by each point on the reflected graph. Strictly speaking, we should also have shown that $f(-x,y)=0$ is *not* satisfied by any point that *isn't* on the reflected graph. Fill this gap.

[**Hint:** *Let (a,b) be a variable point* **not** *on the reflected graph. Now "undo" the transformation, etc.*]

Right-hand Side Shortcuts for *Functions*

When working specifically with graphs of *functions* (i.e. equations of the form $y=f(x)$), we can carry out transformations by operating directly on the equation's right-hand side. The following "RHS Shortcuts" are equivalent to, but more convenient than, the usual substitutions.*

Vertical Stretch by a factor of k†	Multiply RHS by k
Vertical Shift by k units	Add k to the RHS

To see that these "new" operations are really just shortcuts for familiar substitutions, consider a vertical stretch of the graph of $y=f(x)$. Substituting (y/k) for y gives us $y/k=f(x)$. Algebraic massage turns this into $\boldsymbol{y=kf(x)}$. And voilà! The net algebraic effect is to multiply the original RHS by k, as claimed.‡ The same sort of argument justifies the vertical *shift* shortcut, as you should verify.

Example 1. Roughly speaking, what does the graph of $y=2x^2+3$ look like?

Solution. The graph of $y=x^2$ is the familiar U-shape. From $y=x^2$, we can reach $y=2x^2+3$ in two algebraic "moves": First, we multiply the RHS by 2 (yielding $y=2x^2$), then add 3 to the RHS. The corresponding geometric transformations: Stretch vertically by a factor of 2, then shift up by 3. Hence, the graph we seek is a slightly skinnier U-shape opening up with its vertex is at $(0,3)$. ♦

Never forget: The RHS shortcuts apply only to *functions* ($y=f(x)$), not to equations in general!

* RHS = **R**ight-**H**and **S**ide of the equation $y=f(x)$.

† If $k<0$, this includes a vertical reflection. (With this understood, we can compress the table in the spirit of Exercise 24.)

‡ Note well: This argument worked only because our equation had the special form of a *function*, $y=f(x)$. Try it with an equation that does *not* have that form, such as $x^2+y^2=1$, and you'll see that the argument fails.

The RHS shortcuts are gratifyingly direct: To stretch vertically by 5, we just multiply the RHS by 5; to shift up by 3 units, we just **add** 3 to the RHS. A pleasant contrast to the topsy-turvy world of substitutions, where we must remember to take reciprocals and reverse algebraic signs! The RHS shortcuts are easy to use, easy to remember, and easy to understand. But once again, with feeling: They apply only to *functions*.

Let's practice the shortcuts by applying them to some sheep in wolves' clothing – harmless functions dressed up in enough transformational garb to make them *look* fearsome to the uninitiated.

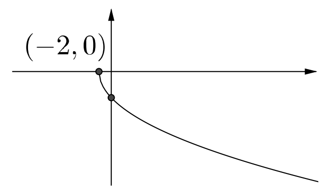
Example 2. Graph the function $y = -3\sqrt{x+2}$.

Solution. This is just a transformed version of a familiar function, $y = \sqrt{x}$, as this analysis shows:

$$y = \sqrt{x} \xrightarrow{\text{sub } (x+2) \text{ for } x} y = \sqrt{x+2} \xrightarrow{-3(\text{RHS})} y = -3\sqrt{x+2}$$

The corresponding *geometry*: Shift right by 2, then stretch vertically by a factor of 3 (accompanied by a vertical reflection).

Applying those successive geometric transformations to the graph of $y = \sqrt{x}$ will bring us to the graph at right, as you should verify. Hence, the figure at right must be the given function's graph. To provide more detail, we could find and label its intersection with the y-axis, which (as you should also verify) is $(0, -3\sqrt{2})$. ♦



In the previous example, we could have carried out the transformations in any order and obtained the same graph. (Try it and see.) In general, however, we must take care with the order of transformations.

Example 3. Graph the function $y = \frac{1}{2}x^2 - 3$.

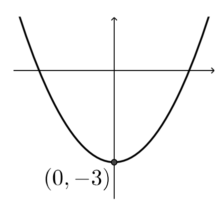
Solution. We recognize that this function is just a transformed version of $y = x^2$:

$$y = x^2 \xrightarrow{(1/2)\cdot\text{RHS}} y = \frac{1}{2}x^2 \xrightarrow{\text{RHS} - 3} y = \frac{1}{2}x^2 - 3$$

The corresponding *geometric* transformations: a vertical stretch by $1/2$, then a shift down by 3.

Applying these two geometric transformations (in the specified order!) to the familiar U-shaped graph of $y = x^2$ yields a graph whose basic shape is shown at right. Should we desire more detail, we could find and label its intersections with the x-axis in the usual way.

[These are $(-\sqrt{6}, 0)$ and $(\sqrt{6}, 0)$, as you should verify.] ♦



As you saw in exercise 12, it is essential that you do geometric transformations in the correct order. How can you tell whether a given order will yield the correct graph? This is easy: Just check whether the corresponding *algebraic* transformations, applied in the same order, yield the *equation* whose graph you want. If they do, then all's well. If they don't, then there's a problem.

In the previous example, suppose you had wondered if you could have *first* shifted down by 3 units, and *then* stretched vertically by a factor of $1/2$. Would that work? Well, consider the corresponding algebraic transformations: *First* subtract 3 from the original RHS, *then* multiply the resulting RHS by $1/2$. A little calculation shows that doing these to the original equation ($y = x^2$) yields $y = (1/2)x^2 - \mathbf{3/2}$, which, misses our algebraic target: $y = (1/2)x^2 - 3$. Hence, the corresponding geometric operations would have missed the geometric target, too; they would have led us to the graph of the wrong equation. Order is important. First you open the window, *then* you put your head through.

Here's a more challenging example – which we'll solve in two different ways.

Example 4. Graph the function $y = \sqrt{1-(3x+9)^2}$.

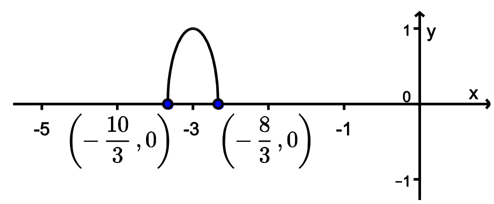
First Solution. This is a transformed version of $y = \sqrt{1-x^2}$, whose graph is the top half of the unit circle. Here's one sequence of algebraic steps that accomplishes this transformation:

$$y = \sqrt{1-x^2} \xrightarrow{\text{sub } (x+9) \text{ for } x} y = \sqrt{1-(x+9)^2} \xrightarrow{\text{sub } 3x \text{ for } x} y = \sqrt{1-(3x+9)^2}$$

The corresponding *geometric* transformations are:

Shift left by 9 units, then **stretch horizontally by $1/3$**.

Applying these geometric transformations *in the specified order* to the unit circle's top half gives us the graph of $y = \sqrt{1-(3x+9)^2}$, which, as you can verify by carrying the transformations out yourself, looks like the graph at right.[*] ♦



Second Solution. If we do some preliminary algebra (factoring out a 3 inside the parentheses), we can rewrite the given function in a new form,

$$y = \sqrt{1-(3(x+3))^2},$$

which suggests a different sequence of steps:

$$y = \sqrt{1-x^2} \xrightarrow{\text{sub } 3x \text{ for } x} y = \sqrt{1-(3x)^2} \xrightarrow{\text{sub } (x+3) \text{ for } x} y = \sqrt{1-(3(x+3))^2}$$

The corresponding *geometric* transformations are:

Stretch horizontally by $1/3$, then **shift left by 3**.

By drawing pictures, you should verify that the graph resulting from this alternate sequence of geometric transformations will be the same that we found in the first solution. Notice that the *shift* here was by only 3 units (as opposed to 9 units in the first solution); because we compressed the graph first, we didn't have to shift it as far.[†] ♦

[*] It's useful to follow a few individual points on their transformational journey. For example, let's follow $(1,0)$ on the unit circle. The shift left takes it to $(-8,0)$; the horizontal stretch then takes this to $(-8/3, 0)$.

[†] We can see this by following the journey of $(1,0)$ once again. First, the horizontal stretch takes it to $(1/3, 0)$. Then the shift left takes it over to $(-8/3, 0)$.

Exercises.

30. Give the *geometric* transformation corresponding to the following algebraic transformations of $y = f(x)$:

a) Multiply the RHS by 5 b) Substitute $(x + 2)$ for x c) Multiply the RHS by -1
d) Add 1 to the RHS e) Multiply the RHS by $7/9$ f) Substitute $6x$ for x
g) Substitute $(2/3)x$ for x h) Add 5 to the RHS i) Multiply the RHS by -3
j) Multiply the RHS by $(-3/5)$ k) Substitute $-4x$ for x l) Substitute $(y + 2)$ for y.

31. True or false (and explain why the false answers are false):

a) Multiplying the RHS of $x^2 + y^2 = 1$ by 4 stretches its graph vertically by a factor of 4.
b) Adding 3 to the RHS of $2x^2 + 2y^2 = 2$ shifts its graph up by 3 units.
c) Subtracting 7 from the RHS of $10xy = x^2 + 3y^3 - 5$ shifts its graph down by 7 units.
d) Multiplying the RHS of $y = 14x^8 + \sqrt[9]{x}$ by 3 stretches its graph vertically by a factor of 3.

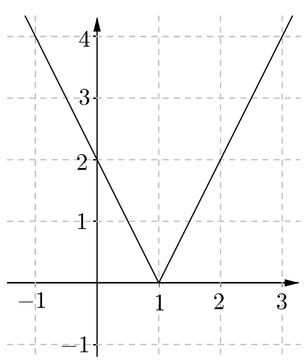
32. Give the *algebraic* transformation corresponding to the following geometric transformations of $y = f(x)$:

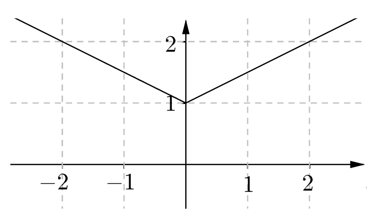
a) Shift right by 8 b) Shift left by 8 c) Shift down by 8 d) Shift up by 8 e) V-stretch by a factor of 6
f) V-stretch by a factor of $1/3$ g) H-stretch by a factor of 10 h) H-stretch by a factor of $1/10$
i) V-stretch by a factor of $3/16$ j) Reflection over the x-axis k) Reflection over the y-axis
l) V-stretch by a factor of 5 *and* a v-reflection m) V-stretch by a factor of $3/7$ *and* a vertical reflection

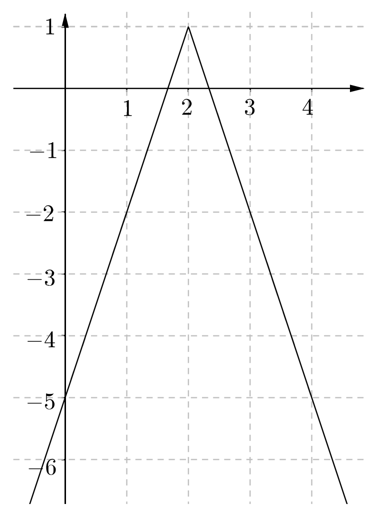
33. Graph the following functions, and label key points with their coordinates. (These will include intersections with the axes, and possibly endpoints or turning points when these exist).

a) $y = 5(x-1)^3$ b) $y = \sqrt{x+3} + 1$ c) $y = -2|x| + 3$ d) $y = \frac{2}{3}(x+2)^2 - 1$
e) $y = \frac{4}{x-2}$ f) $y = -3(x-1)^2 + 2$ g) $y = -\frac{3}{x} + 7$ h) $y = \sqrt[3]{8x - 8}$
i) $y = 2\sqrt{9 - x^2} - 3$ j) $y = 1 - \sqrt{4 - x^2}$ k) $y = -\frac{1}{2}(x+1)^3$

34. Find the functions corresponding to the following graphs, which are transformed versions of $y = |x|$.

a)


b)


c)


35. Using only shifts, stretches, and reflections, is it possible to transform line $y = x$ into the line through (x_0, y_0) with slope m? If not, why not? If so, give a sequence of transformations that will do it.

Graphs of Quadratic Functions

I'm very well acquainted too with matters mathematical,
I understand equations, both the simple and quadratical,
About binomial theorem I am teemin' with a lot o' news –
With many cheerful facts about the square of the hypotenuse!
- Major General Stanley (*Pirates of Penzance*, Act I)

We've already proven that the graphs of all linear functions (that is, functions of the form $y = ax + b$) are straight lines. In this section we'll prove that graphs of all quadratic functions (that is, functions of the form $y = ax^2 + bx + c$) are U-shapes. Moreover, we'll see later that they are not just any old U-shapes: They are *parabolas*. Before proving that all quadratic functions have U-shaped graphs, we'll consider one specific quadratic. This will contain, in a very tangible form, all the key ideas we'll need for the abstract universal proof. One idea we'll need is "completing the square", so you may wish to review that technique before reading on.

Example. Graph the quadratic function $y = 2x^2 + 12x + 13$.

Solution. We begin by rewriting the equation in an equivalent form by *completing the square*.

$$\begin{aligned} y &= 2x^2 + 12x + 13 \\ &= 2(x^2 + 6x) + 13 \\ &= 2(x^2 + 6x + \mathbf{9} - \mathbf{9}) + 13 \\ &= 2[(x+3)^2 - 9] + 13 \\ &= 2(x+3)^2 - 5. \end{aligned}$$

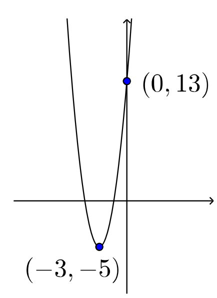
In this form, we see that our quadratic is just a transformed version of $y = x^2$, whose graph is the familiar U-shape. Specifically, it is related to $y = x^2$ by the following sequence of transformations:

$$y = x^2 \xrightarrow{2(\text{RHS})} y = 2x^2 \xrightarrow{\text{sub } (x+3) \text{ for } x} y = 2(x+3)^2 \xrightarrow{\text{RHS} - 5} y = 2(x+3)^2 - 5$$

The corresponding sequence of *geometric* transformations is:

Stretch vertically by 2, shift left by 3, shift down by 5.

Following these transformations of the graph of $y = x^2$ in your mind's eye, you'll see the rough shape of the graph we seek: a stretched out U, which bottoms out at $(-3, -5)$. For a bit more detail, we note that the equation tells us that when x is 0, we have $y = 13$. The graph we want is thus is a U whose lowest point is $(-3, -5)$ and which crosses the y-axis at $(0, 13)$. This much information is already enough to produce the graph at right.



To include more detail, if we desire it, we could find the points where the curve crosses the x-axis. [These, as you can verify, are $\left(-3 \pm \sqrt{5/2},\ 0\right)$.] ♦

The method we employed in the preceding example can be used to graph *any* quadratic function: Complete the square, then apply the appropriate transformations to the graph of $y = x^2$. Let us now tackle the problem in the abstract.

Claim. *Every* quadratic function has a U-shaped graph.

Proof. Consider the general quadratic function $y = ax^2 + bx + c$.
Rewriting the equation in an equivalent form by completing the square, we obtain

$$\begin{aligned} y &= ax^2 + bx + c \\ &= a\left(x^2 + \frac{b}{a}x\right) + c \\ &= a\left(x^2 + \frac{b}{a}x + \boldsymbol{\frac{b^2}{4a^2}} - \boldsymbol{\frac{b^2}{4a^2}}\right) + c \\ &= a\left[\left(x + \frac{b}{2a}\right)^2 - \frac{b^2}{4a^2}\right] + c \\ &= a\left(x + \frac{b}{2a}\right)^2 + \left(c - \frac{b^2}{4a}\right). \end{aligned}$$

This reveals that the general quadratic function is an algebraically transformed version of $y = x^2$. Consequently, its graph can be obtained from the U-shaped graph of $y = x^2$ by a sequence of geometric transformations: a vertical stretch by a factor of a (with a vertical *reflection* if $a < 0$), then a horizontal shift, and finally, a vertical shift. Since a stretched U is still a U (albeit a skinnier or a fatter one), and a shifted U obviously is still a U, we may therefore conclude that the graph of $y = ax^2 + bx + c$ must *always* be U-shaped, as claimed. ∎

One further observation: In the proof, we saw that if a, the quadratic's leading coefficient, is *negative*, then the U will be reflected vertically. If this happens, it will, of course, open downwards. Otherwise (if the leading coefficient is *positive*) the U will open upwards. To sum up what we've discovered,

The graph of every *quadratic* function (i.e. function of the form $y = ax^2 + bx + c$) **is U-shaped**.

- The U opens upwards if the leading coefficient is positive, and downwards if it is negative.
- To graph a quadratic, we complete the square and apply the transformations revealed thereby.

Finally, a little terminology: The turning point in a U-shape is called its *vertex*. In the graph of a quadratic function, the vertex lies either where the function attains its maximum output value (if the graph opens downwards) or its minimum value (if it opens upwards).

Exercises

36. Graph the following quadratics. Find the graphs' vertices and their intersections with the axes.

a) $y = x^2 + 8x + 1$ b) $y = 2x^2 + 4x + 2$ c) $y = 5x^2 - 3$ d) $y = -x^2 + 10x - 7$

e) $y = 3x^2 + 4x + 5$ f) $y = -5x^2 - 12x + 2$ g) $y = \frac{3}{2}x^2 + 6x$ h) $y = -\frac{3}{14}x^2 + \frac{2}{7}x + 1$

37. Can the graph of $y = x^2$ be contained between two vertical lines? If so, which ones? If not, why not?

38. Many textbooks state that in the graph of $y = ax^2 + bx + c$, the vertex's x-coordinate will be $-b/2a$.

a) Explain *why* this is so.

b) Use this result to find the coordinates of the vertex of $y = 5x^2 + 4x - 1$.

c) Although this result will allow you to solve certain homework problems (like those in exercise 36) quickly, memorizing it is counterproductive. It will *not* help you learn mathematics. In contrast, each time that you graph a quadratic by completing the square, you reinforce two important mathematical techniques: completing the square and transformations. (Besides, if you want a shortcut, why not just use a computer?) The second reason for eschewing this formula is that once you've learned a little calculus, you'll be able to find a quadratic's vertex in seconds without having to rummage in your memory for *anything*.

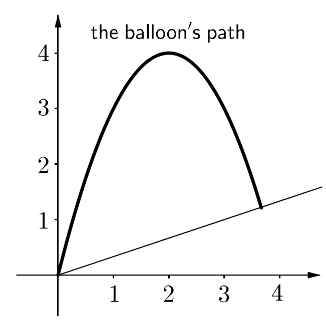
39. If two rectangles have the same *perimeter*, must they have the same *area*? If so, explain why. If not, provide a counterexample.

40. Mr. Square plans to fence off a rectangular area in the middle of field. He has 100 feet of fencing. He dimly remembers learning that rectangles with the same perimeter can have different areas, and he wants his rectangle's area as large as possible. Being Mr. Square, he's pretty sure that the way to maximize the area is to make a square, but then, he's been wrong before. Is he right this time? If not, why not? If so, prove it.

[**Hint:** *Consider a rectangle with perimeter 100. Let x be the length of one of its sides. Express the rectangle's area as a function of x. The function will be quadratic. Graph it, and think about its vertex in the context of Mr. Square's problem*.]

41. Mr. Square's neighbor, Lana Evitneter, also plans to fence off a rectangular area with 100 feet of fencing. However, since a wall of her house will serve as one side of the rectangle, she actually needs fencing for just three sides. She asks Mr. Square to help her maximize the area of her rectangle. Naturally, Mr. Square recommends making a square. Is he right this time? If so, prove it. If not, find the dimensions of the rectangle that will actually maximize the enclosed area.

42. Mortified by his mistake, Mr. Square flees the neighborhood in his hot-air balloon. He takes off from the base of a very long hill, which has a constant slope of $1/3$ (see the figure). The path that Mr. Square's balloon takes happens to be the graph of $y = -x^2 + 4x$. Alas, as you can see from the figure, his balloon doesn't make it over the hill. Assuming that each unit on the axes represents 1000 feet,



a) What are the coordinates of the point at which the balloon lands?

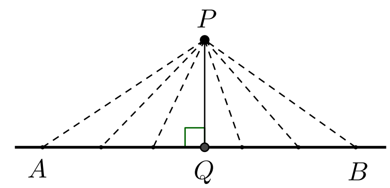
b) How far (in feet) is the launch point from the landing point?

c) If the launch point is at sea level, then what is the balloon's maximum altitude relative to sea level?

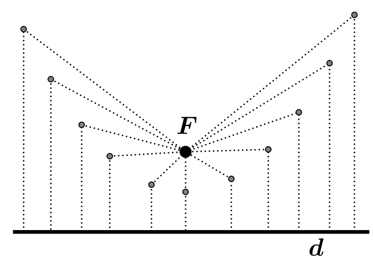
d) What is the balloon's maximum altitude relative to *the ground*?

Parabolas

> It has been observed that missiles and projectiles describe a curved path of some sort. However, no one has yet pointed out that this path is a *parabola*. This... I have succeeded in proving.
>
> - Galileo Galilei, *Dialogues Concerning Two New Sciences*, 3rd Day, Introduction



The distance from a point to a line is, by definition, the length of the shortest path joining them: a straight path meeting the line at right angles. Thus, in the figure at right, the distance from point P to line AB is the length of segment PQ.



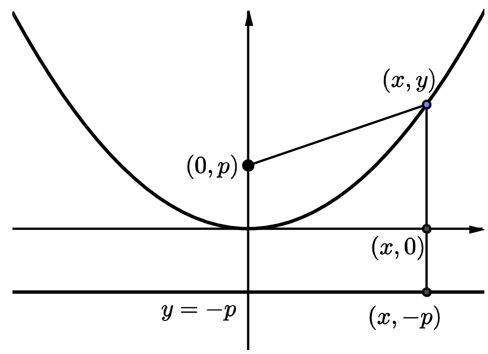
Draw a line and a point on a piece of paper. Call the point F and the line d. Locate a point equidistant from F and d, and mark it on the page. Then, find and mark as many other points as you can that are equidistant from F and d. After a while, you'll have a picture that looks something like the one at right. In your mind's eye, picture the curve that passes through each and every one of the infinitely many points equidistant from F and d. This curve is called a *parabola*.

> **Definition.** A *parabola* is the set of all points equidistant from some fixed point (called the parabola's *focus*) and some fixed line (called the parabola's *directrix*.)

While all parabolas are U-shapes, *very few U-shapes are parabolas*. If you draw a random U-shape on a page, it will almost certainly *not* be a parabola. Try it: First draw a random U-shape on a page, then try to guess where its focus and directrix would be if it were a parabola. Now start checking points on the U (preferably with the aid of a ruler). If you can find even one point on the U that is not exactly the same distance from your prospective focus and directrix, then your U-shape is not a parabola – at least not with those choices of focus and directrix. Difficult as it is to draw a reasonably accurate circle freehand (without the aid of a compass), it is still more difficult to draw a reasonably accurate parabola.

Mathematicians have studied parabolas for well over two thousand years on account of their remarkable geometric properties. Amazingly, in the early 17th century, Galileo proved that parabolas have physical significance as well: A projectile moving under the influence of gravity alone will always follow a parabolic path. When you toss a ball to your dog, the path that the ball follows when it leaves your hand is not merely U-shaped, but *parabolic*: As it moves through the air, the ball remains equidistant from an invisible focus and directrix. Take a physics class, and you'll learn why.

We can discover a parabola's equation by translating its definition into algebraic terms. This will be especially easy if we place the axes so that the parabola's vertex is at the origin and its focus lies on the positive y-axis. This setup ensures that the focus's coordinates will be $(0,\ p)$ for some positive number p. Moreover, since the vertex is p units from the focus, it must also (by the parabola's definition) lie p units up from the directrix; hence, the equation of the directrix must be $y = -p$.



Problem. Derive the equation of the parabola with vertex $(0,0)$ and focus $(0, p)$.

Solution. We seek an equation satisfied by the coordinates of *all* the parabola's points. Let (x, y) be a variable point on the parabola. Its distance to the *focus* $(0, p)$ is, by the distance formula,

$$\sqrt{(\Delta x)^2 + (\Delta y)^2} = \boldsymbol{\sqrt{x^2 + (y-p)^2}},$$

while its distance to the directrix is $\boldsymbol{y + p}$**.** (It lies y units above the x-axis, which itself lies p units above the directrix, as in the figure above.)

The two distances we've just computed are, by the parabola's definition, equal. That is,

$$\sqrt{x^2 + (y-p)^2} = y + p.$$

This equation is satisfied by our variable point (and thus by every point) on the parabola, so it is the parabola's equation. To polish it, we square both sides of the equation and then simplify. The resulting polished equation of the parabola is, as you should verify, $y = (1/4p)x^2$. ♦

The equation of the parabola with vertex $(0,0)$ and focus $(0, p)$ is

$$\boldsymbol{y = \frac{1}{4p}x^2}$$

We'll use this fact to establish two little preliminary results that we'll then parlay into a big theorem. Here's the first little result.

Claim 1. The graph of $y = x^2$ is not merely U-shaped, but *parabolic*.

Proof. The equation $y = x^2$ has the form in the box above, with $p = 1/4$. Hence, its graph is a parabola with vertex $(0, 0)$ and focus $(0, 1/4)$. (And its directrix is $y = -1/4$.) ■

The second of our two preliminary results concerns *stretched* parabolas. Suppose we stretch a parabola. The result will certainly be U-shaped, but will it still be parabolic? Remarkably, it turns out that if we stretch any parabola by any factor *in any direction* (not just vertically or horizontally), the result will still be a parabola! Although this full statement is too difficult for us to prove here, we can easily prove one very special – and very useful – case.

Claim 2. If we stretch the graph of $y = x^2$ vertically, the result is still a parabola.

Proof. If we stretch the graph of $y = x^2$ vertically by a factor of k, its new equation will be $y = kx^2$. This matches our boxed equation for a parabola with $k = 1/(4p)$; or equivalently, with $p = 1/4k$. Hence, the graph of $y = kx^2$ is indeed a *parabola* with vertex $(0, 0)$ and focus $(0, 1/(4k))$. ■

With our two preliminary results squared away, let's turn to our big theorem.

Theorem. The graphs of *all* quadratic function are *parabolas.*

Proof. In an earlier section ("Graphs of Quadratic Functions"), we proved that any quadratic's graph can be obtained from the graph of $y = x^2$ by following a certain sequence of transformations: first a *vertical stretch* (sometimes accompanied by a *reflection*), and then some *shifts*.

Since then, you've learned that the graph of $y = x^2$ is not just a U, but a parabola (Claim 1), and that a vertical stretch will preserve its parabolic nature (Claim 2). What about reflections and shifts? Might they disrupt a stretched parabola's parabolic nature, turning it into a mere U-shape? No, this won't happen. After all, reflections and shifts don't change a graph's shape at all (they only change the graph's location), so they must obviously preserve a parabola's parabolic nature, too.

To sum up, we've now seen that the graph of each and every quadratic function can be obtained from one specific *parabola* ($y = x^2$) by some sequence of "parabola-preserving" transformations. It follows that the graph of *every* quadratic function must be a parabola, as claimed. ■

Combining this last theorem with our earlier work on graphing quadratics, we may conclude that

The graph of every *quadratic* function (i.e. function of the form $y = ax^2 + bx + c$) **is a *parabola***.

- The parabola opens up if the leading coefficient is positive, and down if it is negative.
- To graph a quadratic, we complete the square and apply the transformations revealed thereby.

Exercises.

43. Pick a random number. Call it n. Square it. The point (n, n^2) lies on the graph of $y = x^2$. Compute its distances to the parabola's focus and directrix, and verify that these are indeed equal.

44. Is $y = 2x^2$ the graph of a parabola? If so, find its vertex, focus, and directrix.

45. Is $y = \left(\frac{1}{2}\right)x^2$ the graph of a parabola? If so, find its vertex, focus, and directrix.

46. Find the equation of the parabola with vertex $(0, 0)$ and focus $(0, 5)$.

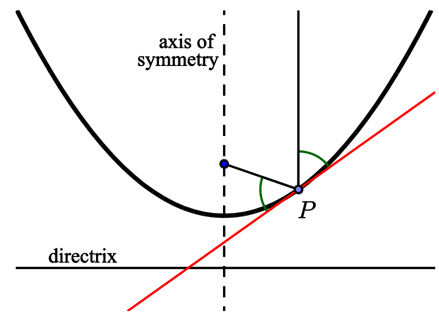
47. Find the equation of the parabola with vertex $(0, 0)$ and focus $(0, 1/5)$.

48. Prove that if we stretch the graph of $y = x^2$ *horizontally*, the result will still be a parabola.

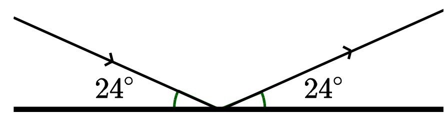
49. Find the focus and directrix of the parabola whose equation is $y = 3x^2 - 12x + 13$.
[**Hint:** C*omplete the square, think about how each geometric transformation affects the focus & directrix.*]

50. Find the focus and directrix of the parabola whose equation is $y = -3x^2 - 6x + 1$.

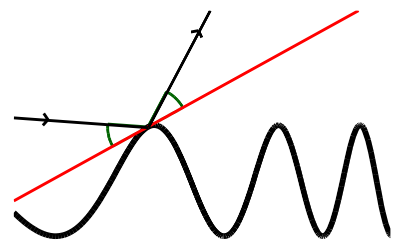
The Reflection Property of Parabolas



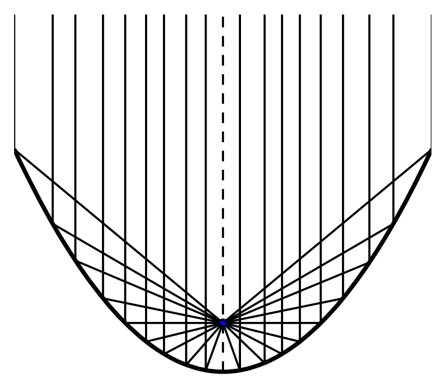
At any point P on a parabola, draw the parabola's tangent there. Next, draw the line segment joining P to the focus. Finally, draw the ray from P that is parallel to the parabola's axis of symmetry. We can prove that the line segment and ray make *equal angles* with the tangent. This so-called "**reflection property**" of parabolas does *not* hold for other U-shapes: only parabolas. Esoteric though it may seem, this parabolic reflection property has remarkable *physical* consequences, which you'll be able to appreciate after reading the following paragraphs about optics.



Light rays bounce off of flat surfaces in a simple manner: When a ray strikes a flat surface, it reflects off at the same angle at which it struck. The figure at right depicts a light ray striking (and then departing) a reflective surface at an angle of $24°$.



If a light ray strikes a *curved* surface, the same rule holds, but now the angles are measured between the light ray and the *tangent* to the surface where the ray hits it. You can see why this is so if you imagine zooming in on the point of tangency so closely that the curve and the tangent line become indistinguishable. This virtual identity of a curve and its tangent line – when viewed at a microscopic scale – is, incidentally, a major theme of calculus.

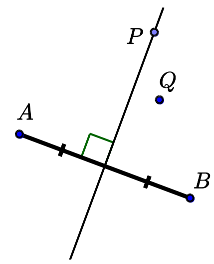


Now let us return to the reflection property of parabolas. If light leaves a parabola's *focus* and hits the parabola at point P, in what direction will it be reflected? If you've understood the three preceding paragraphs, you should be able to convince yourself of the following: Because of the reflection property, the ray that strikes the parabola at P will "bounce off" the parabola *parallel to the parabola's axis of symmetry*. This being so, suppose that we illuminate a bulb at the parabola's focus, so that light streams out of it in all directions. Amazingly, *all* the light that hits the parabola will bounce back in the same direction: parallel to the parabola's axis of symmetry! It is for this reason that parabolic mirrors are used in flashlights, headlights, and so forth.

Parabolic mirrors can also be used in reverse to concentrate parallel rays into a point. For example, solar rays that reach us on Earth are effectively parallel, since the Earth and Sun are so tiny compared to the vast distance between them. If we capture solar rays in a parabolic mirror, we can concentrate them into one very hot point at the mirror's focus. (Hence the name *focus*, which means "hearth" in Latin.) You can find images of such "burning mirrors" or "solar furnaces" online, including the world's largest, located in a French village in the Pyrenees. Another example: Satellite receiving dishes are parabolic in shape to concentrate the satellite's signals into the dish's focus, where the dish's transmitter is located.

Proof of the Parabola's Reflection Property

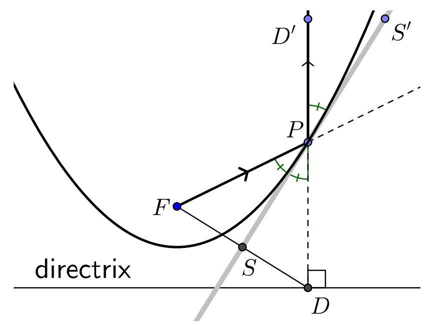
Our proof will make use of the following fairly obvious geometric fact.



*The **perpendicular bisector** of a line segment AB divides the plane into three sets of points: those on A's side of the bisector (which are closer to A than B); those on B's side of the bisector (which are closer to B than A); and those on the bisector itself (which are equidistant from A and B);*

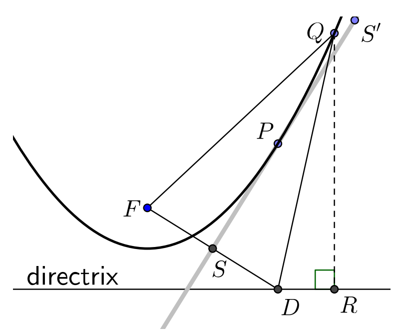
For example, in the figure at right, we must have $PA = PB$, but $QB < QA$.

Claim. Light emitted from a parabola's focus reflects off the parabola parallel to its axis of symmetry.



Proof. First we'll set the stage. Let P be any point on a parabola. Now draw segment FP joining it to the focus, and ray PD' parallel to the axis of symmetry. Among all the straight lines through P that do *not* enter angle $F\hat{P}D'$, its obvious that only one makes equal angles with FP and PD'. I've drawn this line (PS') in gray. Next, we extend ray PD' backwards until it hits the directrix at D. Since PD' is parallel to the axis of symmetry, the extension must be perpendicular to the directrix. Finally, let FD's intersection with the gray line be called S. The stage is now set.

In the proof's first act (of three), we will prove that the gray line is FD's *perpendicular bisector*. The key is to show that $\Delta FPS \cong \Delta DPS$, which we can do as follows: First, $FP = PD$ by the parabola's defining property. Next, $F\hat{P}S = D\hat{P}S$ (since both angles are equal to $D'\hat{P}S'$; one by the gray line's definition, the other by vertical angles). Finally, SP is common to both triangles. Therefore, by SAS-congruence, $\Delta FPS \cong \Delta DPS$ as claimed. It follows that $FS = DS$. That is, the gray line *bisects* FD. Moreover, $F\hat{S}P = D\hat{S}P$, and since they form a straight line, these equal angles must be *right* angles. Thus the gray line is indeed FD's *perpendicular* bisector.



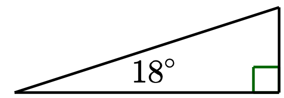
In act two, we'll prove the gray line is *tangent* to the parabola by showing that all the parabola's points (besides P, of course) lie on one side of the gray line. To this end, let Q be any other point on the parabola, and drop perpendicular QR to the directrix. By the parabola's definition, $QF = QR$, which is less than QD (in any right triangle, leg $<$ hypotenuse), so $QF < QD$. Thus, Q is on F's side of FD's perpendicular bisector, the gray line. Since Q was an arbitrary point, *all* the parabola's points (besides P) lie on F's side of the gray line, which is thus tangent to the parabola at P, as claimed.

The third and final act: Imagine that a ray of light emitted from the parabola's focus F strikes the parabola at P, as in the figure above. As explained earlier, the angle at which the ray is reflected equals the angle at which it strikes the parabola's *tangent* at P. But this tangent is the gray line. Since the ray strikes this line at angle $F\hat{P}S$, it reflects off at an angle of the same magnitude. But *by definition* of the gray line, that angle is $D'\hat{P}S'$. That is, the light reflects off the parabola along ray PD', which, by definition, is parallel to the parabola's axis of symmetry. ■

Chapter 10

The Unit Circle Definitions

Circling the Triangle: Generalizing Sine and Cosine

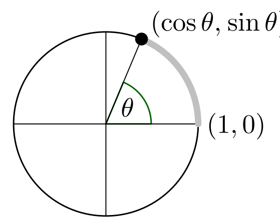


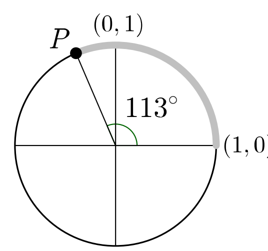
The triangle at right suggests that the numerical value of $\sin 18°$ is about $1/3$. Your calculator will confirm your intuition, thus reinforcing your faith in both. Surely SOH CAH TOA is in his heaven, and all's well in the world. But the Devil never sleeps! Should the evil one tempt you to enter $\sin 113°$ into your calculator, you'll be confronted with an enigma: $\sin 113° \approx 0.921$, an equation that hints at dark mysteries beyond SOH CAH TOA's realm, for what could $\sin 113°$ even mean? It clearly can't refer to the opposite-to-hypotenuse ratio in a right triangle containing a $113°$ angle, for *no such triangle exists*.

This mystery will dissolve once we've *redefined* the trigonometric functions. But can we do that? We used the original definitions to develop the subject, so wouldn't changing them now invalidate all our previous work? No it will not, because as you'll see, the new definitions have been carefully designed to preserve all the old results to which you've grown accustomed. The new definitions are, in computer terminology, backwards compatible. Here, then, are the most important definitions in all of trigonometry:

The Unit Circle Definitions of sine and cosine.

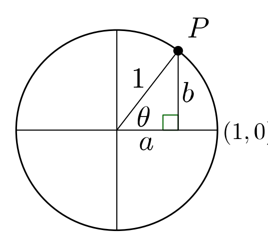
We define $\boldsymbol{\cos\theta}$ and $\boldsymbol{\sin\theta}$ to be, respectively, the x and y coordinates of the point that we reach when we rotate $(1, 0)$ counterclockwise about the origin through an angle of θ.



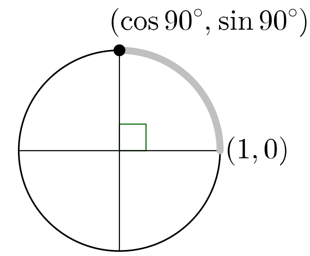


Sine's unit-circle definition (unlike its SOH CAH TOA definition) tells us what $\sin 113°$ means: It is simply the y-coordinate of point P in the figure at right. Clearly, its numerical value is slightly less than 1, so the calculator's assertion that $\sin 113° \approx 0.921$ now makes sense. From the same picture, we can also see that $\cos 113°$ (P's x-coordinate) is clearly *negative*, perhaps about -0.4. Your calculator will confirm this: $\cos 113° \approx -0.391$.

To verify that the unit circle definitions are compatible with the original SOH CAH TOA definitions, let θ be any *acute* angle, and consider the figure. By the old definition, $\cos\theta = a/1 = a$. By the new one, $\cos\theta$ is P's x-coordinate, which is also a. Thus, **the two definitions of cosine agree for all acute angles**. (So do the two definitions of sine, as you should verify.) Because the "new" sine and cosine extend the original functions' domains from acute angles to the more *general* context of all angles, we say that the new definitions *generalize* sine and cosine. We shall generalize the other trigonometric functions shortly.



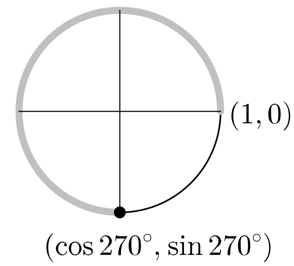
Example 1. Find the exact value of $\sin 90°$.



Solution. To use sine's unit-circle definition, we must rotate $(1, 0)$ counterclockwise about the origin through $90°$, as shown in the figure. The y-coordinate of its new location is, by definition, $\sin 90°$. Since, in this case, the point ends up at $(0, 1)$, we conclude that

$$\sin 90° = 1. \qquad ♦$$

Example 2. Find the exact value of $\cos 270°$.

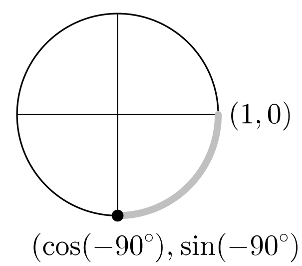


Solution. To use cosine's unit-circle definition, we must rotate $(1, 0)$ counterclockwise about the origin through $270°$, as shown in the figure. The x-coordinate of its new location is, by definition, $\cos 270°$. Since the point ends up at $(0, -1)$, we conclude that

$$\cos 270° = 0. \qquad ♦$$

Important Convention. A *negative* input for sine or cosine corresponds to a *clockwise* rotation of $(1, 0)$.

Example 3. Find the exact value of $\cos(-90°)$.



Solution. To use cosine's unit-circle definition, we must rotate $(1, 0)$ *clockwise* about the origin through $90°$, as shown in the figure at right. The x-coordinate of its new location is, by definition, $\cos(-90°)$. Since the point ends up at $(0, -1)$, we conclude that

$$\cos(-90°) = 0. \qquad ♦$$

To ensure that you understand what we've done so far, stop here and do a few simple exercises:

Exercises.

Find the exact values of….

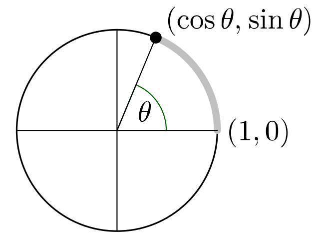
1. $\sin 270°$	**2.** $\sin 45°$	**3.** $\cos 90°$	**4.** $\sin 360°$
5. $\sin 0°$	**6.** $\cos 0°$	**7.** $\sin 60°$	**8.** $\cos 30°$
9. $\sin 720°$	**10.** $\cos(-180°)$	**11.** $-\cos(180°)$	**12.** $\sin(-270°)$
13. $\cos 60°$	**14.** $\sin(-720°)$	**15.** $\sin 360000°$	**16.** $\sin 450°$

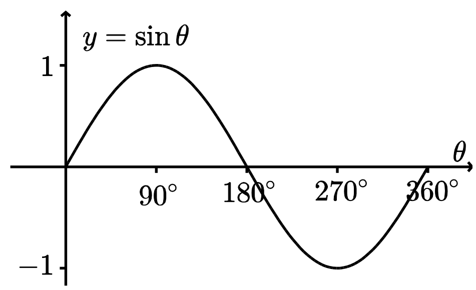
But Why Generalize?

Be wise: Generalize.

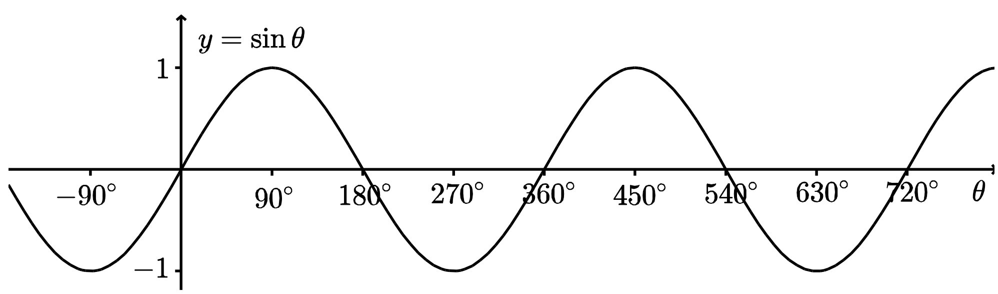
- An old mathematical saw.

The new definitions are not just exercises in generality for generality's sake. You can grasp their significance by thinking about the *graph* of $y = \sin\theta$, which we'll draw by considering the figure at right and noting how $\sin\theta$ varies as θ runs through a full rotation: As θ runs from $0°$ to $90°$, $\sin\theta$ increases from 0 to 1; as θ runs from $90°$ to $180°$, $\sin\theta$ decreases from 1 to 0; as θ runs from $180°$ to $270°$, $\sin\theta$ decreases from 0 to -1; and as θ runs from $270°$ to $360°$, $\sin\theta$ rises from -1 to 0. Putting this all together, we obtain the following graph.





But this is just the beginning. Taking a second lap around the unit circle, as θ runs from $360°$ to $720°$, $\sin\theta$ goes through the same cycle of values it went through during the first lap. The same, of course, holds true in the third lap, the fourth, and so on indefinitely. Throwing in negative values of θ for good measure, the graph of sine reveals its true character at last.



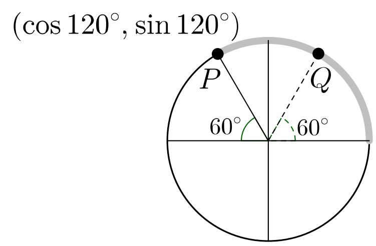
This is the famous *sine wave*, whose endlessly repeating pattern appears not only in mathematical models of waves (water, sound, light), but of periodic phenomena of all sorts, from the revolutions of planets to the vibrations of atoms. By generalizing the trigonometric functions, we thus open up new worlds of applications that would have been unthinkable in the restricted context of right triangles.

Exact Values of Sine and Cosine at Special Angles

God sends special angles into our lives.

- A roadside sign in southern Washington

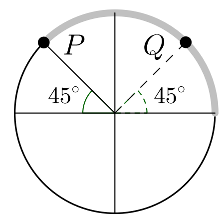
You've already learned the exact values of sine and cosine at the three special angles $30°, 60°$, and $45°$. We can now find the exact values of sine and cosine *at all integer multiples of these three special angles*. The method is simple: Draw a picture, and use it to relate the coordinate that you want to something in the *first* quadrant – the acute-angled land we know so well.



Example 1. Find the exact value of $\sin(120°)$.

Solution. Since $120°$ falls $60°$ short of a half-turn, we want the y-coordinate of point P. Clearly, points P and Q have the *same* y-coordinate, so $120°$ and $60°$ have the same sine. Thus,

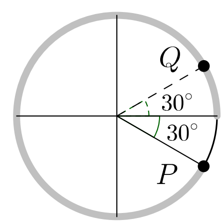
$$\sin(120°) = \sin(60°) = \frac{\sqrt{3}}{2}. \qquad \blacklozenge$$



Example 2. Find the exact value of $\cos(135°)$.

Solution. Since $135°$ falls $45°$ short of a half-turn, we must find point P's x-coordinate. Clearly, points P and Q have *equal but opposite* x-coordinates (same magnitude, opposite sign). Hence, $135°$ and $45°$ have equal but opposite cosines. Thus,

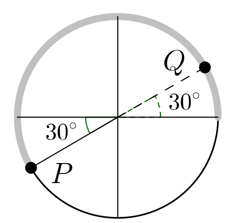
$$\cos(135°) = -\cos 45° = -\frac{\sqrt{2}}{2}. \qquad \blacklozenge$$



Example 3. Find the exact value of $\cos(330°)$.

Solution. Since $330°$ falls $30°$ short of a full rotation, we must find point P's x-coordinate. Points P and Q have the *same* x-coordinate, so $330°$ and $30°$ have the same sine. Thus,

$$\cos(330°) = \cos 30° = \frac{\sqrt{3}}{2}. \qquad \blacklozenge$$



Example 4. Find the exact value of $\sin(210°)$.

Solution. $210°$ falls $30°$ beyond a half-turn, so we want P's y-coordinate. Clearly, points P and Q have *equal but opposite* y-coordinates. Hence, $210°$ and $30°$ have equal but opposite sines. Thus,

$$\sin 210° = -\sin 30° = -\frac{1}{2}. \qquad \blacklozenge$$

Exercises.

Find the exact values of…

17. $\cos 120°$ **18.** $\sin 135°$ **19.** $\cos 150°$ **20.** $\sin 150°$ **21.** $\cos 300°$ **22.** $\sin 315°$

23. $\cos 270°$ **24.** $\cos 240°$ **25.** $\sin 225°$ **26.** $\cos 225°$ **27.** $\cos 1935°$ **28.** $\cos(-60°)$

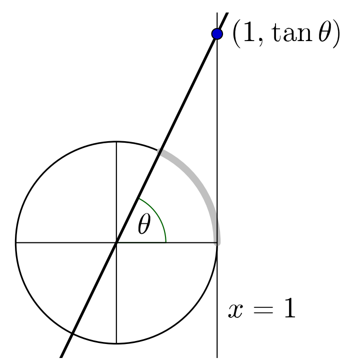
Circling the Triangle II: Generalizing the Tangent Function

> He experienced both bliss and horror in contemplating the way an inclined line, rotating spokelike, slid upwards along another, vertical one… The vertical one was infinite, like all lines, and the inclined one, also infinite, sliding along it and rising ever higher… was doomed to eternal motion, for it was impossible for it to slip off, and the point of their intersection, together with his soul, glided upwards along an endless path. But with the aid of a ruler he forced them to unlock: he simply redrew them, parallel to one another, and this gave him the feeling that out there, in infinity, where he had forced the inclined line to jump off, an unthinkable catastrophe had taken place, an inexplicable miracle, and he lingered long in those heavens where earthly lines go out of their mind.
>
> - Vladimir Nabokov, *The Defense* (Chapter 2).

Having generalized sine and cosine, we will now do as much for tangent, whose unit-circle definition explains this function's name, since it involves a line that is literally tangent to the unit circle.

The Unit Circle Definition of Tangent.

Imagine a line that initially coincides with the x-axis.
Rotate it about the origin through an angle of θ (counterclockwise).
It intersects the fixed line $x = 1$ at some point.
We define $\mathbf{tan}\,\boldsymbol{\theta}$ to be the y-coordinate of that point.

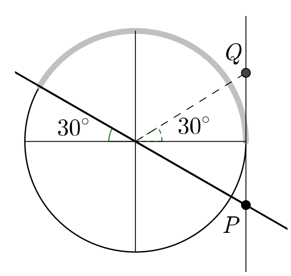


We can easily verify that this new definition of tangent agrees with the old one for all *acute* angles. Consider the figure above. By the old SOH CAH TOA definition, $\tan\theta$ is the ratio (in the right triangle) of the vertical to the horizontal leg. But the horizontal leg's length is 1, so $\tan\theta$ is simply the vertical leg's length. By the new unit-circle definition, $\tan\theta$ is the marked point's y-coordinate… which is obviously the length of the triangle's vertical leg. Hence, the two definitions of tangent agree for all acute angles. However, the unit circle definition also applies to *non*-acute angles, where SOH CAH TOA fears to tread.

Example. Find the exact value of $\tan(150°)$.

Solution. $150°$ falls $30°$ short of a half-turn, so we want P's y-coordinate, which is clearly equal but opposite to Q's. The tangents of $150°$ and $30°$ are therefore equal but opposite. Thus,

$$\tan 150° = -\tan 30° = -\frac{1}{\sqrt{3}}. \quad \blacklozenge$$



Exercises

29. Find the exact values of the tangents of: $210°$, $240°$, $330°$, $135°$, $120°$, $0°$, $180°$, $300°$.

30. Strictly speaking, tan is undefined at 90° and 270°. Informally, we say that tan is *infinite* at those angles. Why?

31. Explain why it makes some intuitive sense to say that $1/\infty = 0$. (We'll use this idea in the next section).

32. State the *ranges* of the sine, cosine, and tangent functions.

Circling the Triangle III: Last Generalizations

> In one word… one must *generalize*; this is a necessity
> that imposes itself on the most circumspect observer.
>
> - Henri Poincare, *The Value of Science*, Chapter 5

Generalizing the reciprocal trigonometric functions is trivial:

Definitions. For *all* values of θ, we define:

$$\boldsymbol{\sec\theta} = \frac{1}{\cos\theta}, \qquad \boldsymbol{\csc\theta} = \frac{1}{\sin\theta}, \qquad \boldsymbol{\cot\theta} = \frac{1}{\tan\theta}.$$

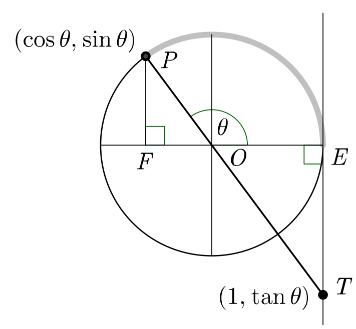
Note: When $\tan\theta$ is infinite, we define $\cot\theta$ to be 0. (See exercises 30 and 31 above.)

For example, since $\tan 150° = -1/\sqrt{3}$ (the previous page's example), it follows that $\cot 150° = -\sqrt{3}$.

Having generalized all of the trigonometric functions, it is now incumbent upon us to prove that the *identities* we proved in the previous chapter still hold in the trigonometric functions' enlarged domains. (Our original proofs of those identities used right triangles, so they are valid only for *acute* angles.)

Claim 1. For *all* values of θ at which the following expressions are defined, $\tan\theta = \frac{\sin\theta}{\cos\theta}$.

Proof. We've already proved that this holds for acute angles. Next, suppose θ lies in the *second* quadrant (as at right).



Drop a perpendicular PF from P to the x-axis, thus producing similar right triangles, ΔPOF and ΔTOE. Corresponding ratios of similar triangles are equal, so we have

$$\frac{ET}{OE} = \frac{FP}{FO}.$$

We'll now express each of these four lengths in terms of θ.

From the figure, observe that $\tan\theta$ is negative, while length ET is (like all lengths) positive. Consequently, $\boldsymbol{ET = -\tan\theta}$. Similarly, $\cos\theta$ is negative, but FO is positive, so $\boldsymbol{FO = -\cos\theta}$. The remaining legs are easy: $\boldsymbol{FP = \sin\theta}$, and $\boldsymbol{OE = 1}$. Substituting these boldface expressions into the proportion above and simplifying, we obtain

$$\tan\theta = \frac{\sin\theta}{\cos\theta},$$

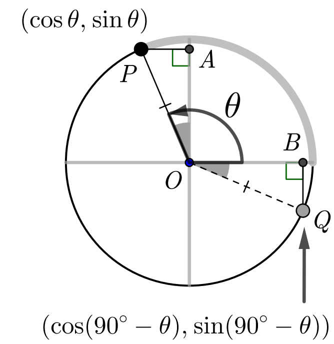
thus proving that the identity holds for all angles in the *second* quadrant. Similar arguments (with similar triangles) show that it holds in the third and fourth quadrants, too, as you should verify. ■

Next, we'll demonstrate that the cofunction identities hold for *all* angles. I'll provide some, but not all, of the details – enough for you to fill in the rest if you so desire.

Claim 2. Sine and cosine are cofunctions. That is, for *all* values of θ,

$$\boldsymbol{\sin(90° - \theta) = \cos\theta} \qquad \text{and} \qquad \boldsymbol{\cos(90° - \theta) = \sin\theta}.$$

Proof. We've already proved this for acute (i.e. quadrant 1) angles, so now let's prove that these cofunction identities still hold when θ lies in quadrant 2, as at right. We'll do so as follows:



First, let's locate $90° - \theta$ on the unit circle. Subtracting a 2nd-quadrant θ from 90° yields a *negative* angle (between $-90°$ and 0°), so rotating $(1, 0)$ through $90° - \theta$ puts it in quadrant 4 (at point Q).

Next, we'll get some congruent right triangles into the picture by dropping perpendiculars from P and Q to the axes. These congruent triangles tell us that $BQ = AP$. But carefully considering the figure shows that $BQ = -\sin(90° - \theta)$ and $AP = -\cos\theta$. Inserting these expressions into $BQ = AP$ and multiplying both sides by -1, we obtain $\boldsymbol{\sin(90° - \theta) = \cos\theta}$.

A nearly identical argument that begins by observing the equality $OB = OA$ leads, as you should verify, to the conclusion that $\boldsymbol{\cos(90° - \theta) = \sin\theta}$.

Arguments very similar to these will prove the identities when θ is in quadrants 3 or 4. ■

Thanks to the work we've done already, generalizing the other co-function identities is very easy.

Claim 3. Tangent and cotangent are cofunctions. That is, for *all* relevant values of θ,

$$\boldsymbol{\tan(90° - \theta) = \cot\theta} \qquad \text{and} \qquad \boldsymbol{\cot(90° - \theta) = \tan\theta}.$$

Proof. $\tan(90° - \theta) = \frac{\sin(90° - \theta)}{\cos(90° - \theta)} = \frac{\cos\theta}{\sin\theta} = \frac{1}{\frac{\sin\theta}{\cos\theta}} = \frac{1}{\tan\theta} = \cot\theta.$

Those equals are justified (in order) by: Claim 1, Claim 2, algebra, Claim 1, cotangent's definition. This establishes one identity. The other identity can be proved similarly, as you should verify. ■

Claim 4. Secant and cosecant are cofunctions. That is, for *all* relevant values of θ,

$$\boldsymbol{\sec(90° - \theta) = \csc\theta} \qquad \text{and} \qquad \boldsymbol{\csc(90° - \theta) = \sec\theta}.$$

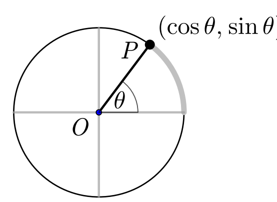
Proof. $\sec(90° - \theta) = \frac{1}{\cos(90° - \theta)} = \frac{1}{\sin\theta} = \csc\theta.$

The equals are justified, respectively, by secant's definition, Claim 2, and cosecant's definition. This establishes one identity. The other identity can be proved similarly, as you should verify. ■

Exercises

33. Find the exact values of the following:
$\cot 210°$, $\sec 240°$, $\csc 330°$, $\sec 135°$, $\csc 120°$, $\sec 0°$, $\csc 765°$, $\cot 300°$, $\cot 90°$.

34. Which of these numbers are in secant's range?: -5, $2/3$, $-e/\pi$, 10^{10}, π/e.



35. In the figure, we know P's coordinates in terms of θ, while O's are fixed at $(0, 0)$. Consequently, we should be able to express segment OP's slope in terms of θ. Do so. [A-ha! Now we have another nice new way to "see" the value of $\tan\theta$ on the unit circle: It's simply line OP's *slope*! Look back at exercise 30 with this in mind.]

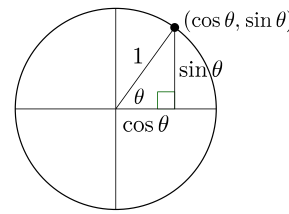
The Pythagorean Identity, Even and Odd Functions

Let's begin with the best-known of all trigonometric identities, often called "the Pythagorean Identity".

Claim 1. (The Pythagorean Trig Identity) For all values of θ, the following relationship holds:

$$\cos^2\theta + \sin^2\theta = 1.^*$$

Proof. For *all* values of θ, the definitions of sine and cosine ensure that the point $(\cos\theta, \sin\theta)$ lies on the unit circle. Its coordinates therefore satisfy the unit circle's equation, $x^2 + y^2 = 1$. That is, $\cos^2\theta + \sin^2\theta = 1$ for all values of θ, as claimed. ■

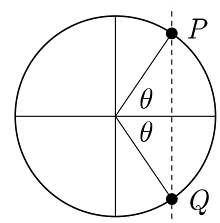


The "Pythagorean" nature of the identity is explained by the figure at right. You'll encounter the Pythagorean identity often in the future – and not just in this course. Because it relates $\cos\theta$ and $\sin\theta$, it lets us replace expressions involving $\cos\theta$ with expressions involving $\sin\theta$ (or vice versa) when doing so is convenient – which is surprisingly often.

An **even function** is one that always sends numbers and their negatives to precisely the same value. (In symbols, f is even if $f(-x) = f(x)$ for *all* x.) Examples: $f(x) = x^2$, $g(x) = x^4$, $h(x) = x^6$, and…

Claim 2. Cosine is an *even* function. That is, for all values of θ,

$$\boldsymbol{\cos(-\theta) = \cos\theta}.$$



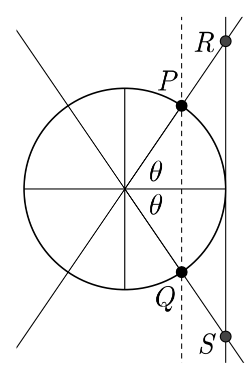
Proof. This should be obvious if you understand cosine's unit circle definition. The x-coordinates of P and Q will clearly be equal no matter what θ is (even if θ lies outside of the first quadrant). And these equal x-coordinates are, by the unit-circle definition of cosine, $\cos\theta$ and $\cos(-\theta)$ respectively. ■

Since cosine is even, we can quickly note, for example, that $\cos(-60°) = \cos 60° = 1/2$.

An **odd function** is one that always sends numbers and their negatives to *equal but opposite* values. (In symbols, f is odd if $f(-x) = -f(x)$ for *all* x.) Examples: $f(x) = x^3$, $g(x) = x^5$, $h(x) = x^7$, and…

Claim 3. Sine and tangent are *odd* functions. That is, for all values of θ,

$$\boldsymbol{\sin(-\theta) = -\sin\theta} \qquad \text{and} \qquad \boldsymbol{\tan(-\theta) = \tan\theta}.$$



Proof. The y-coordinates of P and Q will clearly be equal but opposite no matter what θ is (even if θ lies outside of the first quadrant). Thus, by the unit circle definition of sine, we have $\sin(-\theta) = -\sin\theta$ for all θ.

The same is true of R and S. Thus, by unit-circle definition of tangent, $\tan(-\theta) = \tan\theta$ for all θ. ■

Since sine is odd, we have, for example, $\sin(-60°) = -\sin 60° = -\sqrt{3}/2$.

[*] The peculiar symbol $\cos^2\theta$ is shorthand for $(\cos\theta)^2$.

The reciprocal trig functions inherit the evenness or oddity of their "parents". For example, secant is even for the following reason: For any number θ, we have $\sec(-\theta) = 1/\cos(-\theta) = 1/\cos\theta = \sec\theta$. (The equalities are justified by, in order, secant's definition, cosine's evenness, and secant's definition.) Trigonometry is unusual in that most of its main functions are even or odd. Most non-trigonometric functions are neither even nor odd.

Exercises.

36. Give an example of a function that is neither even nor odd, and prove that this is so.

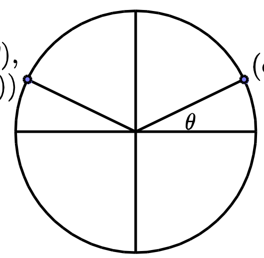
37. Is cosecant even, odd, or neither? Prove it.

38. Is cotangent even, odd, or neither? Prove it.

39. Dividing both sides of the Pythagorean identity by $\cos^2\theta$, we obtain a new identity. What is it?

40. Dividing both sides of the Pythagorean identity by $\sin^2\theta$, we obtain a new identity. What is it?

41. Discover an identity relating $\cos(180° - \theta)$ and $\cos\theta$ by thinking about the picture of the unit circle at right. This is a handy identity, but you need not memorize it, since any time you need it, you can recall it by drawing the appropriate picture and thinking for a moment.



42. Now come up with an identity for $\sin(180° - \theta)$. This identity, too, is often handy, but once again, you need not memorize it; whenever you need it, just draw the appropriate picture and recover it.

43. By drawing appropriate pictures of the unit circle and using identities you already know, find identities for the following expressions. As in the two previous exercises, you need not memorize the identities you discover; instead, you should be able to produce them when you need them by drawing a quick sketch and *thinking*.

a) $\sin(\theta + 90°)$ b) $\cos(\theta + 90°)$ c) $\tan(\theta + 90°)$ d) $\sin(\theta - 90°)$*
e) $\cos(\theta - 90°)$ f) $\tan(\theta - 90°)$ g) $\cos(\theta + 180°)$ h) $\cos(\theta - 180°)$
i) $\sin(\theta + 180°)$ j) $\sin(\theta - 180°)$ k) $\tan(180° + \theta)$ l) $\tan(\theta - 180°)$

44. Use identities to simplify the following expressions:

a) $\dfrac{(\sin\theta + \cos\theta)^2 - 1}{2\cos^2\theta}$ b) $\dfrac{\sin(-\beta)}{\tan(\beta)}$ c) $\left(\dfrac{\sin(180° - \alpha)}{\sin(\alpha + 90°)}\right)^2 + 1$ [**Hint:** *Exercise 46*.]

d) $\dfrac{\sin(180° - \gamma)}{\sin(\gamma - 180°)}$ e) $\dfrac{\cos(90° - x)}{\sin(x - 90°)}$ f) $\cos(180° + \theta)\cos(180° - \theta) - \sin(180° - \theta)\cos(\theta + 90°)$

45. Every even function's graph is symmetric about the y-axis. By thinking about the *definition* of an even function, explain why this is so.

46. The graphs of all odd functions display a peculiar sort of symmetry. Discover it and describe it.

47. I claim that if an odd function is defined at zero, then its graph must pass through the origin.
Either use the definition of an odd function to prove my claim or provide a counterexample that disproves it.

48. There's only one function whose domain is all real numbers that is both even *and* odd. What is it?
[**Hint:** *Think geometrically*.]

* Note well: the argument of sine is *not* $(90° - \theta)$, so this isn't just a cofunction identity. However, you can put it in that form by first using the algebraic identity $(a - b) = -(b - a)$, and then using the fact that sine is an odd function.