DISCOVER CALCULUS



Peter Keep
Moraine Valley Community College



Discover Calculus

Peter Keep
Moraine Valley Community College

This text is disseminated via the Open Education Resource (OER) LibreTexts Project (https://LibreTexts.org) and like the thousands of other texts available within this powerful platform, it is freely available for reading, printing, and "consuming."

The LibreTexts mission is to bring together students, faculty, and scholars in a collaborative effort to provide an accessible, and comprehensive platform that empowers our community to develop, curate, adapt, and adopt openly licensed resources and technologies; through these efforts we can reduce the financial burden borne from traditional educational resource costs, ensuring education is more accessible for students and communities worldwide.

Most, but not all, pages in the library have licenses that may allow individuals to make changes, save, and print this book. Carefully consult the applicable license(s) before pursuing such effects. Instructors can adopt existing LibreTexts texts or Remix them to quickly build course-specific resources to meet the needs of their students. Unlike traditional textbooks, LibreTexts' web based origins allow powerful integration of advanced features and new technologies to support learning.



LibreTexts is the adaptable, user-friendly non-profit open education resource platform that educators trust for creating, customizing, and sharing accessible, interactive textbooks, adaptive homework, and ancillary materials. We collaborate with individuals and organizations to champion open education initiatives, support institutional publishing programs, drive curriculum development projects, and more.

The LibreTexts libraries are Powered by NICE CXone Expert and were supported by the Department of Education Open Textbook Pilot Project, the California Education Learning Lab, the UC Davis Office of the Provost, the UC Davis Library, the California State University Affordable Learning Solutions Program, and Merlot. This material is based upon work supported by the National Science Foundation under Grant No. 1246120, 1525057, and 1413739.

Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation nor the US Department of Education.

Have questions or comments? For information about adoptions or adaptions contact info@LibreTexts.org or visit our main website at https://LibreTexts.org.

The LibreTexts name, logo, and book covers are trademarked by LibreTexts, Inc. (a 501(c)(3) not-for-profit organization) and protected against unauthorized use.

This text was compiled on 10/07/2026

TABLE OF CONTENTS

Licensing

1: Limits

1.1: The Definition of the Limit

1.2: Evaluating Limits

1.3: First Indeterminate Forms

1.4: Limits Involving Infinity

1.5: The Squeeze Theorem

1.6: Continuity and the Intermediate Value Theorem

1.7: Exploration- Limits

2: Derivatives

2.1: Introduction to Derivatives

2.2: Interpreting Derivatives

2.3: Some Early Derivative Rules

2.4: The Product and Quotient Rules

2.5: The Chain Rule

3: Implicit Differentiation

3.1: Implicit Differentiation

3.2: Derivatives of Inverse Functions

3.3: Logarithmic Differentiation

4: Applications of Derivatives

4.1: Mean Value Theorem

4.2: Increasing and Decreasing Functions

4.3: Concavity

4.4: Global Maximums and Minimums

4.5: Optimization

4.6: Linear Approximations

4.7: L'Hôpital's Rule

4.8: Exploration- Applications of Derivatives

5: Antiderivatives and Integrals

5.1: Antiderivatives and Indefinite Integrals

5.2: Riemann Sums and Area Approximations

5.3: The Definite Integral

5.4: The Fundamental Theorem of Calculus

5.5: More Results about Definite Integrals

5.6: Introduction to u Substitution

6: Applications of Integrals

6.1: Integrals as Net Change

6.2: Area Between Curves

6.3: Volumes of Solids of Revolution

6.4: More Volumes- Shifting the Axis of Revolution

6.5: Arc Length and Surface Area

6.6: Other Applications of Integrals

7: Techniques for Antidifferentiation

7.1: Improper Integrals

7.2: More on u Substitution

7.3: Manipulating Integrands

7.4: Integration by Parts

7.5: Integrating Powers of Trigonometric Functions

7.6: Trigonometric Substitution

7.7: Partial Fractions

8: Infinite Series

8.1: Introduction to Infinite Sequences

8.2: Introduction to Infinite Series

8.3: The Divergence Test and the Harmonic Series

8.4: The Integral Test

8.5: Alternating Series and Conditional Convergence

8.6: Common Series Types

8.7: Comparison Tests

8.8: The Ratio and Root Tests

8.9: Exploration- Infinite Series

9: Power Series

9.1: Polynomial Approximations of Functions

9.2: Power Series Convergence

9.3: How to Build Taylor Series

9.4: How to Use Taylor Series

9.5: Exploration- Power Series

10: Parametric and Polar Curves

10.1: Parametric Curves

10.2: Polar Curves

10.3: Calculus with Parametric Curves

10.4: Calculus with Polar Curves

Index

Glossary

Detailed Licensing

This page titled Discover Calculus is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Peter Keep.

Licensing

A detailed breakdown of this resource's licensing can be found in ***Back Matter/Detailed Licensing***.

This page titled Licensing is shared under a Public Domain license and was authored, remixed, and/or curated by LibreTexts.

1: LIMITS

1.1: The Definition of the Limit

1.2: Evaluating Limits

1.3: First Indeterminate Forms

1.4: Limits Involving Infinity

1.5: The Squeeze Theorem

1.6: Continuity and the Intermediate Value Theorem

1.7: Exploration- Limits

This page titled 1: Limits is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Peter Keep.

1.1: The Definition of the Limit

We're going to start this textbook by stating a definition. This is a common practice in math classes: we need to agree upon a common definition of the mathematical objects and adjectives we are thinking about. We will state a lot of definitions in this textbook.

What I hope we will do, though, is motivate these definitions. We want to arrive at a point where it makes sense to give a name to this phenomena or object that we're thinking of. Or maybe we arrive at a point where the specifics of the definition don't just come down to us out of nowhere, but feel like reasonable and obvious things to consider.

So for now, we're going to work on defining a very important and very key mathematical object that is used in calculus: the limit.

A limit is all about closeness, so let's first interact with the idea of closeness, and then work on a definition of a limit.

Defining a Limit

Activity 1.1.1: Close or Not?

We're going to try to think how we might define "close"-ness as a property, but, more importantly, we're going to try to realize the struggle of creating definitions in a mathematical context. We want our definition to be meaningful, precise, and useful, and those are hard goals to reach! Coming to some agreement on this is a particularly tricky task.

a. For each of the following pairs of things, decide on which pairs you would classify as "close" to each other.
   - You, right now, and the nearest city with a population of 1 million or higher
   - Your two nostrils
   - You and the door of the room you are in
   - You and the person nearest you
   - The floor of the room you are in and the ceiling of the room you are in

b. For your classification of "close," what does "close" mean? Finish the sentence: A pair of objects are close to each other if...

c. Let's think about how close two things would have to be in order to satisfy everyone's definition of "close." Pick two objects that you think everyone would agree are "close," if by "everyone" we meant:
   - All of the people in the building you are in right now
   - All of the people in the city that you are in right now.
   - All of the people in the country that you are in right now.
   - Everyone, everywhere, all at once.

d. Let's put ourselves into the context of functions and numbers. Consider the linear function $y = 4x - 1$. Our goal is to find some x-values that, when we put them into our function, give us y-value outputs that are "close" to the number 2. You get to define what close means.

First, evaluate $f(0)$ and $f(1)$. Are these -values "close" to 2, in your definition of "close?"

e. Pick five more, different, numbers that are "close" to 2 in your definition of "close." For each one, find the x-values that give you those y-values.

f. How far away from $x = \frac{3}{4}$ can you go and still have -value outputs that are "close" to 2?

To wrap this up, think about your points that you have: you have a list of x-coordinates that are clustered around $x = \frac{3}{4}$ where, when you evaluate $y = 4x - 1$ at those x-values, you get y-values that are "close" to 2. Great!

Do you think others will agree? Or do you think that other people might look at your list of y-values and decide that some of them *aren't* close to 2?

Do you think you would agree with other peoples' lists? Or you do think that you might look at other peoples' lists of y-values and decide that some of them *aren't* close to 2?

The balance that we need to find, as we discovered in Activity 1.1.1, is about being able to leave room for those with a very strict idea of what "close" might be. We will want to think of an idea kind of like "infinite closeness," but we're not going to frame it this way: we're going to think about a function's output being so close to some specific number that literally everyone can agree. It is so close that it is within every possible definition of closeness.

The general idea is that we want to think about the behavior of a function at inputs that are near some specific input. Is there a trend with the outputs? Are they all centered around a specific value or do they differ wildly?

Definition 1.1.1: Limit of a Function.

For the function $f(x)$ defined at all x-values around a (except maybe at $x = a$ itself), we say that the limit of $f(x)$ as x approaches a is L if $f(x)$ is arbitrarily close to the single, real number L whenever x is sufficiently close to, but not equal to, a. We write this as:

$$\lim_{x \to a} f(x) = L$$

or sometimes we write $f(x) \to L$ when $x \to a$.

We can clarify a couple of things here:

- There are two types of "close" in this definition: "arbitrarily close" and "sufficiently close." One of these is in references to x -values being close to a number and the other is in reference to function outputs being close to a specific number.
- We are concerned with the behavior of a function around, but not at, a specific x-value: $x = a$. We don't really care about what the function is doing at that input (if anything at all), and we already have words to describe that kind of behavior!
- When we talk about x-values that are near a, that might reference x-values that are a bit bigger than a or x-values that are a bit smaller than a. We can be more specific by simply changing this definition to focus on only one "side" individually.

We can go back to Activity 1.1.1 and think about how we chose x-values that were larger than $\frac{3}{4}$ and smaller than $\frac{3}{4}$. Let's define these ideas a bit more formally!

Definition 1.1.2: Left-Sided Limit

For the function $f(x)$ defined at all x-values around and less than a, we say that the left-sided limit of $f(x)$ as x approaches a is L if $f(x)$ is arbitrarily close to the single, real number L whenever x is sufficiently close to, but less than, a. We write this as:

$$\lim_{x \to a^-} f(x) = L$$

or sometimes we write $f(x) \to L$ when $x \to a^-$.

Definition 1.1.3: Right-Sided Limit

For the function $f(x)$ defined at all x-values around and greater than a, we say that the right-sided limit of $f(x)$ as x approaches a is L if $f(x)$ is arbitrarily close to the single, real number L whenever x is sufficiently close to, but greater than, a. We write this as:

$$\lim_{x \to a^+} f(x) = L$$

or sometimes we write $f(x) \to L$ when $x \to a^+$.

This should lead us to our first result in this textbook. This first result will do two things:

1. Introduce some language that we can use when we talk about limits as well as a classification that we can apply to them.

2. Introduce how we will build our results throughout the course of this text. We want to discover these results as things that are required for us to talk about (and do) calculus together, and hopefully we can motivate each one beforehand.

In lieu of a formal activity, let's just review Definition 1.1.1 Limit of a Function pose the following questions to think about:

- Why do we put emphasis on L being a number? What could happen if it wasn't?
- Why do we put the emphasis on the number L being a real number? What other type(s) of number could it be?
- Why do we put emphasis on L being a single number? How could we have the function be close to multiple real numbers?

We can look at one of the ways that we break the definition: by having two different values that the function gets close to.

Theorem 1.1.4: Mismatched Limits.

For a function $f(x)$, if both $\lim_{x\to a^-} f(x) \neq \lim_{x\to a^+} f(x)$, then we say that $\lim_{x\to a} f(x)$ does not exist.

Approximating Limits Using Our New Definition

We have defined a new term, and now we do the typical mathematical task: define a new thing and then investigate it.

Aside.

A common joke in mathematics is that we "make up a guy to get mad at." It's really only kind of a joke, because it really is a pretty good description of what we do! Here, we defined a new object and now we'll think about it and find ways that it frustrates us or some other weird behavior about it. That's mathematics!

We will eventually get really good at thinking about limits and using them, but for now we just want to get familiar with them. Let's approximate these values that our function is near by looking at some pictures of graphs and some tables of function outputs.

Later on, we'll formalize this more. For now, we just want to use these pictures and tables to get familiar with what a limit even is.

Activity 1.1.2: Approximating Limits.

For each of the following graphs of functions, approximate the limit in question. When you do so, approximate the values of the relevant one-sided limits as well.

a. Approximate $\lim_{x\to 1} f(x)$ using the graph of the function $f(x)$ below.

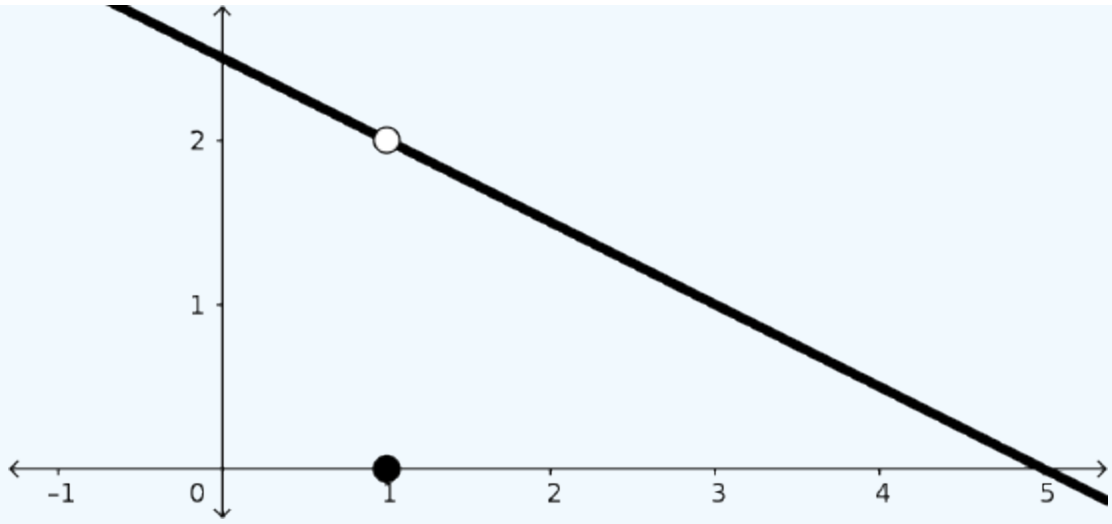


Figure 1.1.5

b. Approximate $\lim_{x\to 2} g(x)$ using the graph of the function $g(x)$ below.

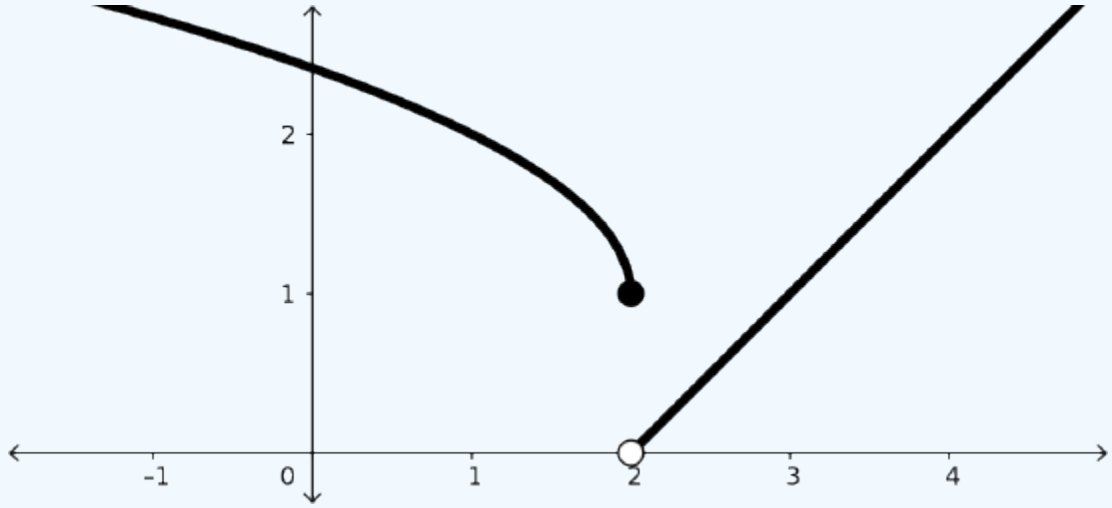


Figure 1.1.6.

c. Approximate the following three limits using the graph of the function $h(x)$ below.

- $\lim_{x\to -1} h(x)$
- $\lim_{x\to 0} h(x)$
- $\lim_{x\to 2} h(x)$

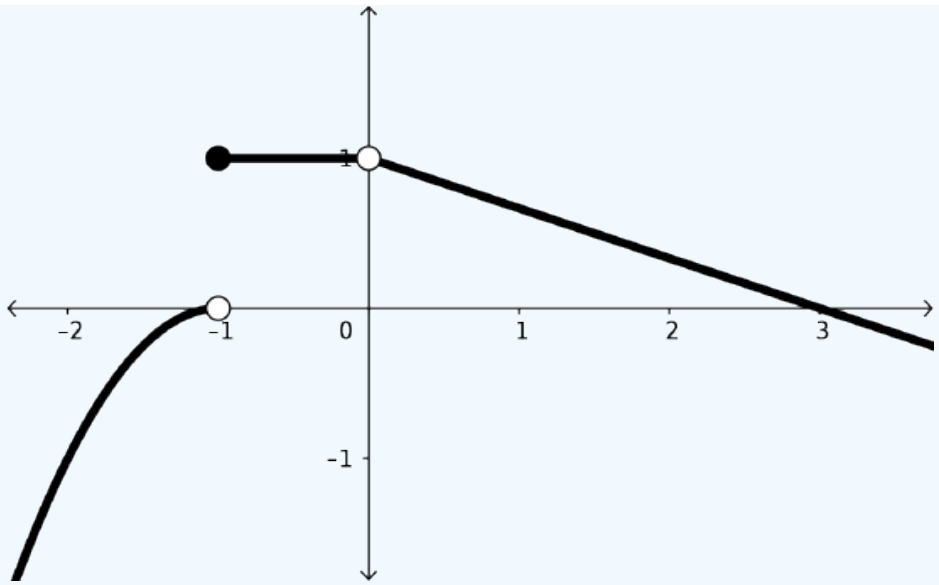


Figure 1.1.7.

d. Why do we say these are “approximations” or “estimations” of the limits we’re interested in?
e. Are there any limit statements that you made that you are 100% confident in? Which ones?
f. Which limit statements are you least confident in? What about them makes them ones you aren’t confident in?
g. What extra details would you like to see to increase the confidence in your estimations? Are there changes we could make to the way these functions are represented that would make these approximations better or easier to make?

It can be hard to focus on the aspects of a graph that we really care about for the purpose of a limit. Let’s build a small strategy to help us think about what we’re looking at. We’ll start by just considering some function, $f(x)$. Using our definition of the Limit of a Function as a guide, we’ll make sure that it’s defined around some x-value, $x = a$.

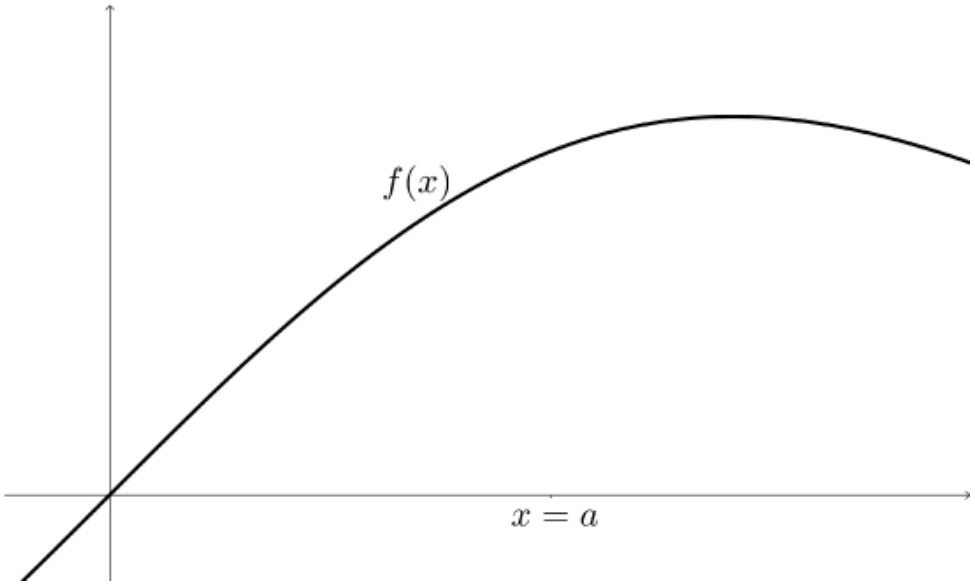


Figure 1.1.8 The function, $f(x)$.

Now we want to investigate more of our definition. We want to look at the x-values that are around, but not equal to, $x = a$.

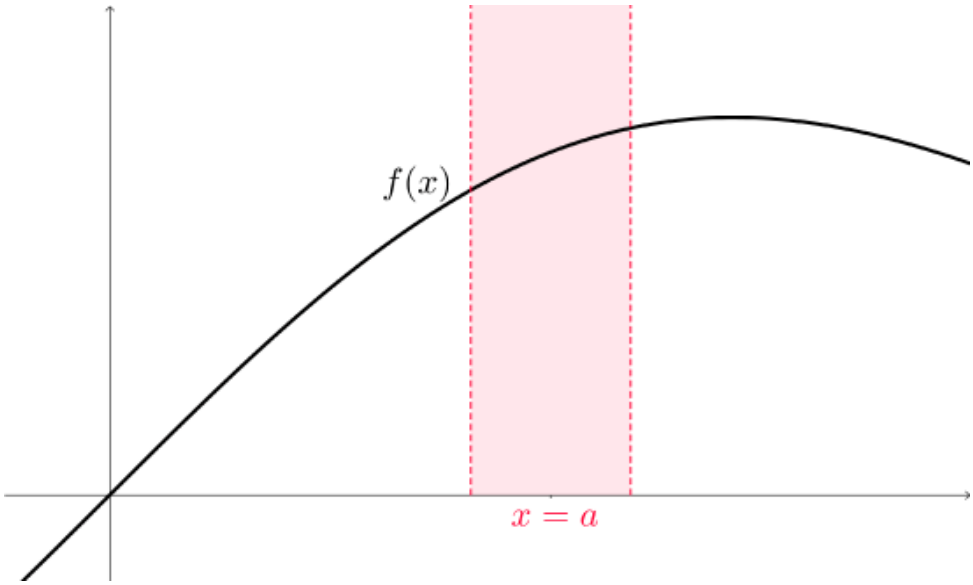


Figure 1.1.9 The x-values around $x = a$.

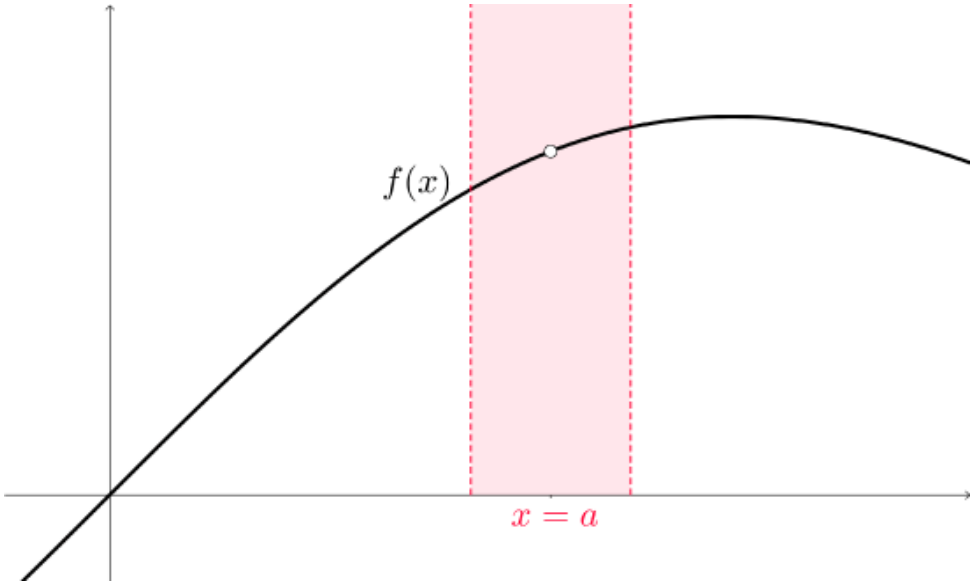


Figure 1.1.10 The x-values around, but not equal to, $x = a$.

We can see that we might as well remove any point at $x = a$ from our graph: we are only concerned with the behavior around that x-value instead of the function's behavior at it.

And now our focus can turn to the function outputs. For the x-values in this interval of inputs that we've constructed, is there some common real number that the corresponding function outputs are close to? We can visualize some interval of y-values. We'll think of this as a target: we want to build an interval of y-values that all of the function outputs from this interval of x-values land in.

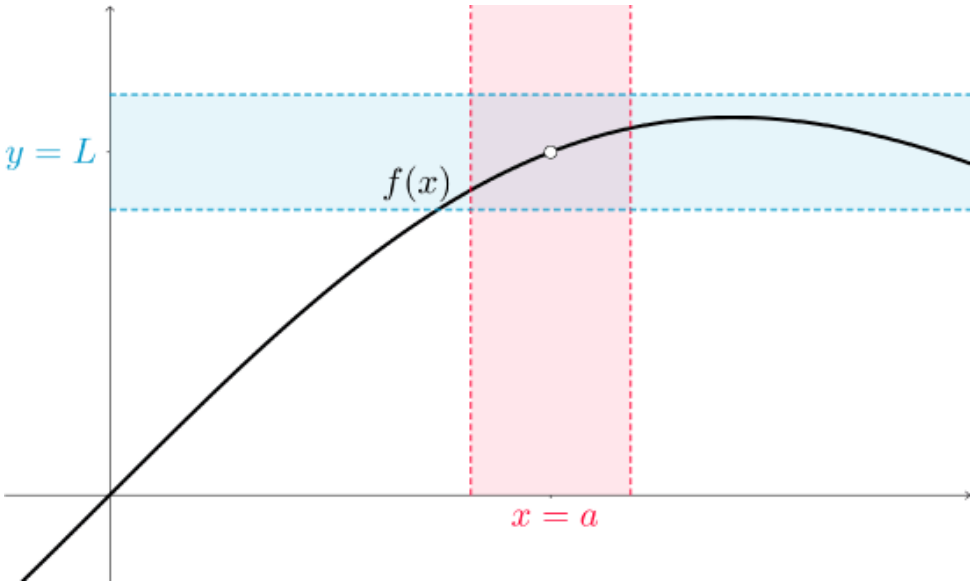


Figure 1.1.11 The corresponding function outputs $f(x)$ are all in the target interval of y-values.

This is a pretty wide range of y-values, but we can see that the graph of the function (when we limit to just the interval of x-values selected) produces function outputs that exist only in that interval. We don't fill the interval, but that's fine!

What we really care about, though, is if these function outputs are all close to the same, single, real number. What we can do is look at a more strict idea of "closeness" in the y-interval by shrinking it. In order for us to produce function outputs that are in this new, smaller, interval, we'll need to correspondingly shrink our interval of inputs to more closely surround $x = a$.

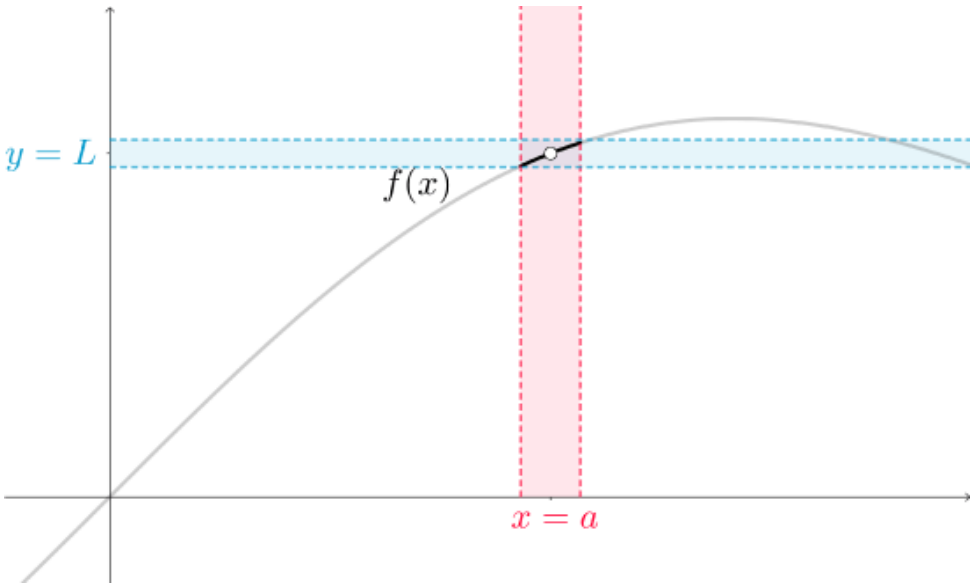


Figure 1.1.12 The x-values around, but not equal to, $x = a$.

In this visualization, we've also tried to focus on just the portion of our function that exists in this little intersection of intervals: we want to know what these functions values are close to, or more specifically if they are all close to the same thing. So we can de-emphasize the rest of our function!

All we're doing is working on a strategy to focus on the parts of this graph that matter: only the parts of the curve that are surrounding $x = a$ (but not that actual point specifically). From there, we just want to know what the function outputs are clustered around, if anything.

Let's look at this same kind of visualization for a limit that does not exist: we're going to think about the case where the one-sided limits don't match. We'll start a little further on in this visualization process: we have a function, and we can visualize an interval of x-values whose function outputs land inside a target interval of y-values.

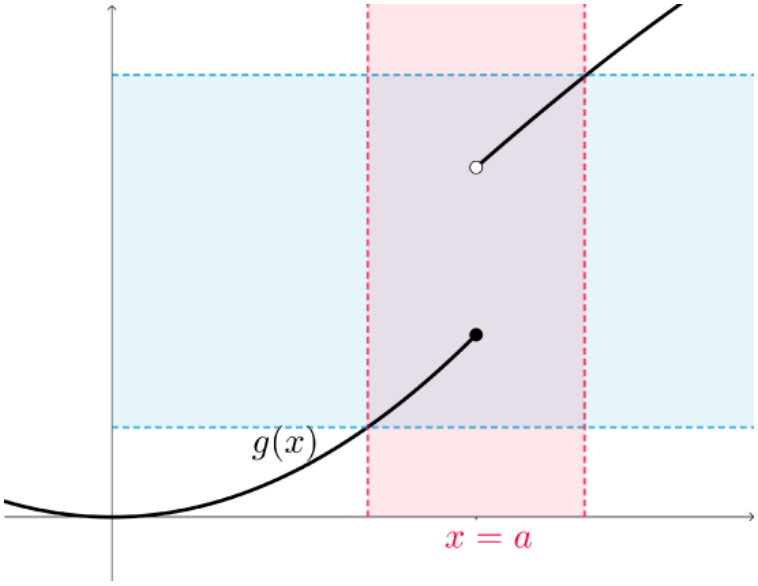


Figure 1.1.13

We can see the problem: that vertical space between the function on the left of $x = a$ and the where the function values are on the right of $x = a$ will make it so that horizontal bar cannot get much smaller. We can disregard the point at $x = a$ as well as the function outside of the interval, but once try to shrink the target interval of y-values, but we'll see the problem.

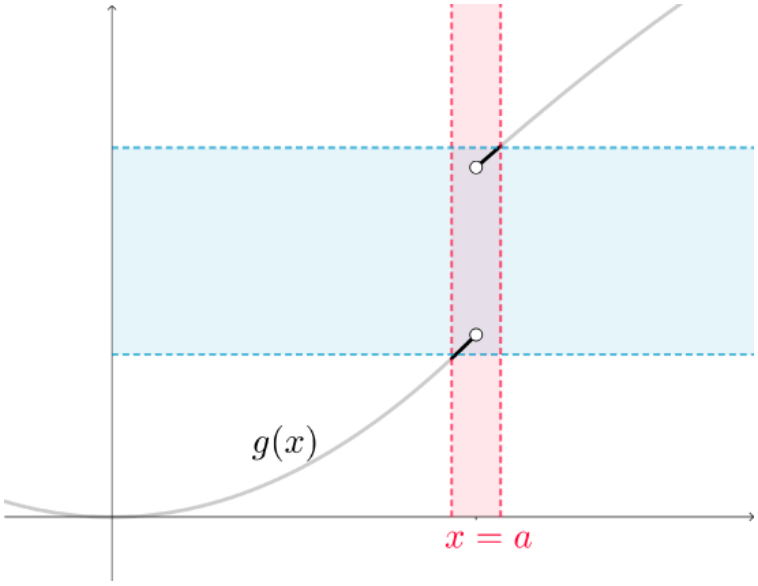


Figure 1.1.14

These function outputs are spread apart! They are not close to a single value. Instead, they're close to two! The function is close to a value on the left side, and then the function is close to a larger value on the right side.

$$\lim_{x \to a^-} g(x) \neq \lim_{x \to a^+} g(x)$$

and so $\lim_{x \to a} g(x)$ does not exist.

Now let's think about how we can approximate (and learn more about) limits using when we just think about the actual values of a function's inputs and corresponding outputs.

Activity 1.1.3: Approximating Limits Numerically.

For each of the following tables of function values, approximate the limit in question. When you do so, approximate the values of the relevant one-sided limits as well.

a. Approximate $\lim_{x\to 1} f(x)$ using the table of values of $f(x)$ below.

Table 1.1.15

x	0.5	0.9	0.99	1	1.01	1.1	1.5
$f(x)$	8.672	9.2	9.0001	-7	8.9998	9.5	7.59

b. Approximate $\lim_{x\to -3} g(x)$ using the table of values of $g(x)$ below.

Table 1.1.16

x	-3.5	-3.1	-3.01	-3	-2.99	-2.9	-2.5
$g(x)$	-4.41	-3.89	-4.003	-4	7.035	2.06	-4.65

c. Approximate $\lim_{x\to \pi} h(x)$ using the table of values of $h(x)$ below.

Table 1.1.17

x	3.1	3.14	3.141	π	3.142	3.15	3.2
$h(x)$	6	6	6	undefined	5.915	6.75	8.12

c. Are you 100% confident about the existence (or lack of existence) of any of these limits?
d. What extra details would you like to see to increase the confidence in your estimations? Are there changes we could make to the way these functions are represented that would make these approximations better or easier to make?

Overall, there's a common theme here: in either representation (graphically or numerically), we're making a best guess at the behavior of the function values around a point. We have limited information in these estimations, and so we're doing the best we can: in graphs, we're trying our best to make sense of the lack of precision in in the scales of our visual, and in the numerical tables we are only given a limited number of points to think about. In both cases, we are hoping to see more information to add more confidence to these estimations.

We want to make the jump from estimating these limits to evaluating them, and for that to happen, we'll need to add more information and more precision about the behavior of our function.

Practice Problems

Exercise 1.1.1

Explain in your own words the meaning of:

$$\lim_{x\to a^-} f(x) = L$$

Exercise 1.1.2

Explain in your own words the meaning of:

$$\lim_{x\to a^+} f(x) = L$$

Exercise 1.1.3

Explain in your own words the meaning of:

$$\lim_{x \to a} f(x) = L$$

Exercise 1.1.4

Say we know that $\lim_{x \to 3^-} f(x) = 2$ and $\lim_{x \to 3^+} f(x) = 2$. What do we know (specifically or in general, if anything) about each of the following?

a. f(3)
b. f(2.999)
c. f(3.001)
d. $\lim_{x \to 3} f(x)$

Answer available at the end of this section.

Exercise 1.1.5

Which of the following is possible? Explain why or why not, and any other conclusions that we can draw.

a. For some function $f(x)$, $\lim_{x \to 5} f(x) = 6$ and f(5) = –3
b. For some function $g(x)$,$\lim_{x \to 4^-} g(x) = -\frac{3}{2}$ and $\lim_{x \to 4^+} g(x) = \frac{4}{7}$.
c. For some function $\ell(t)$, $\lim_{t \to \alpha} \ell(t) = 2$ and $\lim_{t \to \alpha^+} \ell(t) = 1$.
d. For some function $r(\theta)$, $\lim_{\theta \to 0} r(\theta)$ does not exist, $\lim_{\theta \to 0^-} r(\theta) = \pi$, and $\lim_{\theta \to 0^+} r(\theta) = -\frac{\pi}{2}$.
e. For some function $j(w)$, $j(4) = \pi$ while $\lim_{w \to 4} j(w)$ does not exist.

Answer available at the end of this section.

Exercise 1.1.6

Fill in the following tables in order to satisfy the requirements listed. Afterwards, include a sentence or two justifying your choices.

a. Requirements: $\lim_{x \to 1} f(x) = 3$

x	–	0.93	–	1	–	–	1.04
$f(x)$	–	–	–	–	–	–	–

b. Requirements: $\lim_{x \to -5^-} f(x) = 2$, $f(-5) = 6$, and $\lim_{x \to -5} f(x)$ doesn't exist.

x	−5.2	–	–	−5	–	4.98	–
$f(x)$	–	–	–	–	–	–	–

c. Requirements: f(7) does not exist and $\lim_{x \to 7} f(x) = 3$

x	–	–	6.985	7	–	–	7.002
$f(x)$	–	–	–	–	–	–	–

d. Requirements: $\lim_{x \to 0^-} f(x) = \pi$ and $\lim_{x \to 0^+} f(x) = e$.

x	–	−0.14	–	0	–	–	0.5
$f(x)$	–	–	–	–	–	–	–

Exercise 1.1.7

From the following tables, estimate/report each of the requested values. Explain your choices.

a. Requested: $\lim_{x\to 1^-} f(x), \lim_{x\to 1^+} f(x), \lim_{x\to 1} f(x)$, and f(1)

x	0.9	0.99	0.999	1	1.001	1.01	1.1
$f(x)$	2.4	2.48	2.4998	9	2.5004	2.52	2.8

b. Requested: $\lim_{x\to 8^-} f(x), \lim_{x\to 8^+} f(x), \lim_{x\to 8} f(x)$, and f(8)

x	7.9	7.99	7.999	8	7.001	7.01	7.1
$f(x)$	-1.5	-1.9	-1.999	-2	7.0001	7.2	7.5

c. Requested: $\lim_{x\to \pi^-} f(x), \lim_{x\to \pi^+} f(x), \lim_{x\to \pi} f(x)$, and $f(\pi)$

x	3.1	3.14	3.141	π	3.142	3.15	3.2
$f(x)$	-3	-3	-3	does not exist	-3	-3	-3

Answer available at the end of this section.

Exercise 1.1.8

For each of the listed requirements, sketch a graph of a function that satisfies each. Afterwards, include a sentence or two justifying you sketch.

a. Requirements: $f(6) = 0, \lim_{x\to 6} f(x) = -2, \quad \lim_{x\to 2^-} f(x) = 1$, and $\lim_{x\to -2} f(x)$ does not exist.
b. Requirements: $\lim_{\omega\to 0} \rho(\omega) = 8, \lim_{\omega\to 2} \rho(\omega) = -2$, and $\rho(2)$ does not exist.
c. Requirements: $\lim_{t\to -3^-} q(t) = 0, \lim_{t\to -3^+} q(t) = 4, \lim_{t\to -1} q(t) = 9$, and $q(-1) = 9$.

Exercise 1.1.9

From the graph of $f(x)$ below, estimate of each of the requested values. Explain each of your choices.

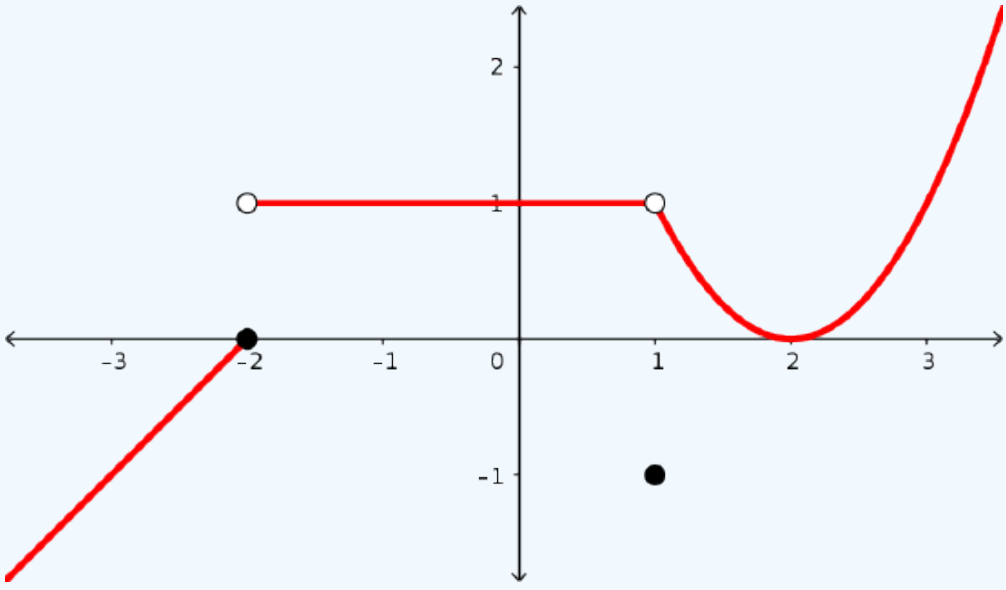


Figure 1.1.18 The function $f(x)$.

a. $\lim_{x\to -2^-} f(x)$

b. $\lim_{x\to -2^+} f(x)$
c. $\lim_{x\to -2} f(x)$
d. $\lim_{x\to 0^-} f(x)$
e. $\lim_{x\to 0^+} f(x)$
f. $\lim_{x\to 0} f(x)$
g. $\lim_{x\to 1^-} f(x)$
h. $\lim_{x\to 1^+} f(x)$
i. $\lim_{x\to 1} f(x)$
j. $\lim_{x\to 2^-} f(x)$
k. $\lim_{x\to 2^+} f(x)$
l. $\lim_{x\to 2} f(x)$

Answer available at the end of this section.

Answers

Exercise 1.1.4

a. We do not know anything about the value of f(3), including whether or not there is such a value.
b. The value of f(2.999) should be close to the number 2. We do not know exactly what the value is, though.
c. The value of f(3.0001) should be close to the number 2. We do not know exactly what the value is, though.
d. Since both the left and the right sided limits are the same single, real number (2), so is the limit: $\lim_{x\to 3} f(x) = 2$.

Exercise 1.1.5

a. This is possible.
b. This is possible, although it means that $\lim_{x\to 4} g(x)$ does not exist.
c. This is not possible, since for $\lim_{t\to\alpha} \ell(t) = 2$ we would need $\lim_{t\to\alpha^+} \ell(t) = 2$ and $\lim_{t\to\alpha^-} \ell(t) = 2$.
d. This is possible.
e. This is possible.

Exercise 1.1.7

a.
- *Estimated:*
  - $\lim_{x\to 1^-} f(x) = 2.5$
  - $\lim_{x\to 1^+} f(x) = 2.5$
  - $\lim_{x\to 1} f(x) = 2.5$
- *Reported:* f(1) = 9

b.
- *Estimated:*
  - $\lim_{x\to 8^-} f(x) = -2$
  - $\lim_{x\to 8^+} f(x) = 7$
  - $\lim_{x\to 8} f(x)$ doesn't exist
- *Reported:* f(8) = -2

c.
- *Estimated:*
  - $\lim_{x\to \pi^-} f(x) = -3$
  - $\lim_{x\to \pi^+} f(x) = -3$
  - $\lim_{x\to \pi} f(x) = -3$
- *Reported:* $f(\pi)$ does not exist

Exercise 1.1.9

a. $\lim_{x\to -2^-} f(x) = 0$
b. $\lim_{x\to -2^+} f(x) = 1$
c. $\lim_{x\to -2} f(x)$ doesn't exist, since the left and right sided limits don't match.
d. $\lim_{x\to 0^-} f(x) = 1$
e. $\lim_{x\to 0^+} f(x) = 1$
f. $\lim_{x\to 0} f(x) = 1$
g. $\lim_{x\to 1^-} f(x) = 1$
h. $\lim_{x\to 1^+} f(x) = 1$
i. $\lim_{x\to 1} f(x) = 1$
j. $\lim_{x\to 2^-} f(x) = 0$
k. $\lim_{x\to 2^+} f(x) = 0$
l. $\lim_{x\to 2} f(x) = 0$

This page titled 1.1: The Definition of the Limit is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Peter Keep.

1.2: Evaluating Limits

Adding Precision to Our Estimations

Activity 1.2.1: From Estimating to Evaluating Limits (Part 1).

Let's consider the following graphs of functions $f(x)$ and $g(x)$.

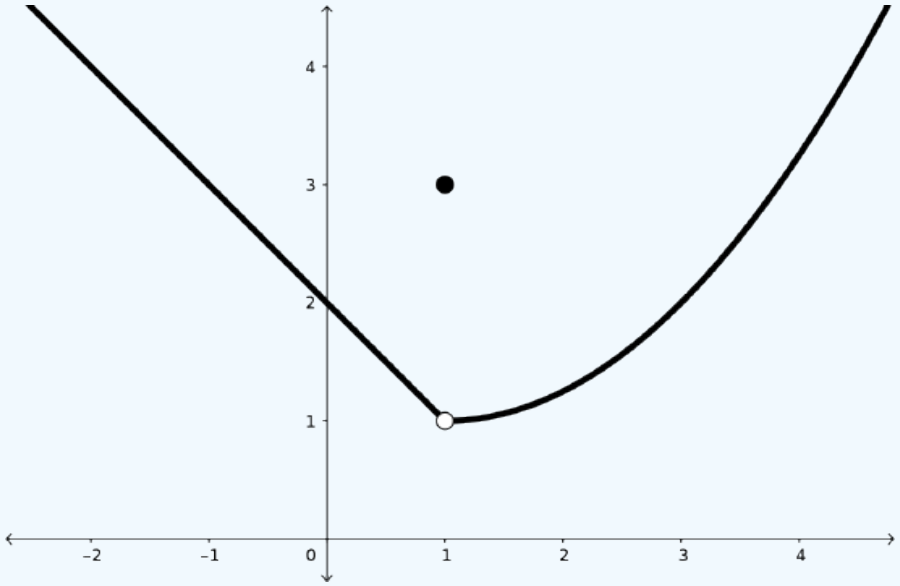


Figure 1.2.1 Graph of the function $f(x)$.

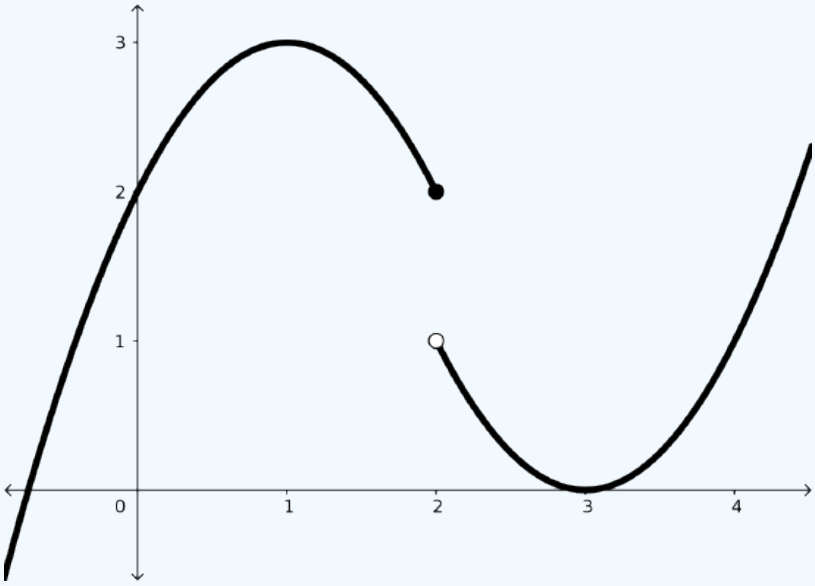


Figure 1.2.2 Graph of the function $g(x)$.

a. Estimate the values of the following limits. If you believe that the limit does not exist, say so and explain why.
   - $\lim_{x\to 1^-} f(x)$
   - $\lim_{x\to 1^+} f(x)$
   - $\lim_{x\to 1} f(x)$

b. Estimate the values of the following limits. If you believe that the limit does not exist, say so and explain why.
   - $\lim_{x \to 2^-} g(x)$
   - $\lim_{x \to 2^+} g(x)$
   - $\lim_{x \to 2} g(x)$

c. Find the values of $f(1)$ and $g(2)$.

d. For the limits and function values above, which of these are you most confident in? What about the limit, function value, or graph of the function makes you confident about your answer? Similarly, which of these are you the least confident in? What about the limit, function value, or graph of the function makes you not have confidence in your answer?

We're going to repeat this process, but with a slight change to the representation of each function. Hopefully this will be illuminating in our attempt to add more precision to our estimations.

Activity 1.2.2: From Estimating to Evaluating Limits (Part 2).

Let's consider the following graphs of functions $f(x)$ and $g(x)$, now with the added labels of the equations defining each part of these functions.

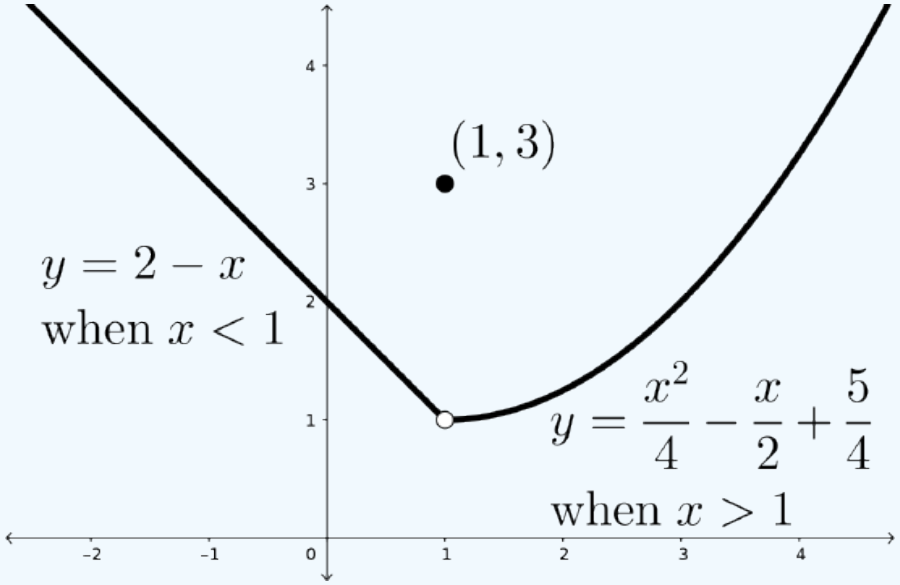


Figure 1.2.3 Graph of the function $f(x)$.

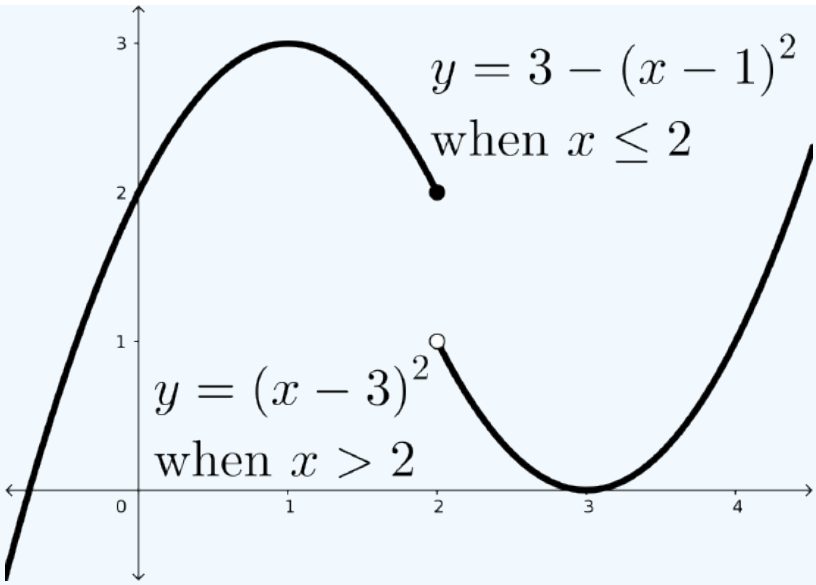


Figure 1.2.4 Graph of the function $g(x)$.

a. Estimate the values of the following limits. If you believe that the limit does not exist, say so and explain why.
   - $\lim_{x\to 1^-} f(x)$
   - $\lim_{x\to 1^+} f(x)$
   - $\lim_{x\to 1} f(x)$

b. Estimate the values of the following limits. If you believe that the limit does not exist, say so and explain why.
   - $\lim_{x\to 2^-} g(x)$
   - $\lim_{x\to 2^+} g(x)$
   - $\lim_{x\to 2} g(x)$

c. Does the addition of the function rules change the level of confidence you have in these answers? What limits are you more confident in with this added information?

d. Consider these functions without their graphs:

$$f(x) = \begin{cases} 2-x & \text{when } x < 1 \\ 3 & \text{when } x = 1 \\ \frac{x^2}{4} - \frac{x}{2} + \frac{5}{4} & \text{when } x > 1 \end{cases}$$

$$g(x) = \begin{cases} 3-(x-1)^2 & \text{when } x \leq 2 \\ (x-3)^2 & \text{when } x > 2 \end{cases}$$

Find the limits $\lim_{x\to 1} f(x)$ and $\lim_{x\to 2} g(x)$. Compare these values of $f(1)$ and $g(2)$: are they related at all?

These two examples are hopefully helpful for us to see that when we are given the actual rule for a function f that connects x to the corresponding output $f(x)$, we are able to move past estimation. We suddenly have whatever level of precision we'd like, since we can immediately see what is happening with every x input to produce the corresponding $f(x)$ output.

In order for us to formalize this evaluation of limits, we're going to think about some properties of this limit object.

Limit Properties

Activity 1.2.3: Combinations of Functions.

We want to remind ourselves how we can combine functions using different operations, and how we might find outputs based on the different combinations. Our goal is to then think about how this might work with limits: how can we summarize the behavior of combinations of functions around some point?

Let's consider some functions $f(x) = x^2 + 3$ and $g(x) = x - \frac{1}{x}$. We'll say that the domain of both functions is $(0, \infty)$ for our own convenience.

a. Let's consider the function $h(x) = f(x) + g(x)$. Describe at least two different ways of finding the value of $h(2)$.
b. If we instead define the function $h(x) = f(x) - g(x)$, how would you describe at least two different ways of finding the value of $h(2)$?
c. What about a scaled version of one of these functions? If we let $h(x) = 4f(x)$ and $j(x) = \frac{g(x)}{3}$, can you describe more than one way to find the value of $h(3)$ and $j(3)$?
d. You can probably guess where we're going: we're going to define a function that is the product of f and $g : h(x) = f(x) \cdot g(x)$. Describe more than one way of evaluating $h(4)$.
e. And finally, let's define $h(x) = \frac{f(x)}{g(x)}$. Now describe more than one way of finding $h(4)$.
f. If $h(x) = \frac{f(x)}{g(x)}$, then are there any x-values that are in the domain of f and g (the domain is $x > 0$) that $h(x)$ cannot be defined for? Why?

Ok, we can confront this big idea: when we combine functions, we can either evaluate the combination of the functions at some x-value or evaluate each function separately and just combine the answers! Of course, there are some limitations (like when the combination isn't nicely defined because of division by 0 or something else), but this is a good framework to move forward with!

Answer available at the end of this section.

Maybe this activity was obvious for you, but it might not have been! This isn't something that we always think about with functions, even if (deep down) we know it to be true.

A nice extension that we can make is that moving past functions evaluated at a specific x-value towards descriptions of the behavior of functions around that specific x-value.

We'll apply this same kind of thinking (combining things by looking at each piece individually first, and then combining the answers together) to limits of combinations of functions.

Theorem 1.2.5: Combinations of Limits.

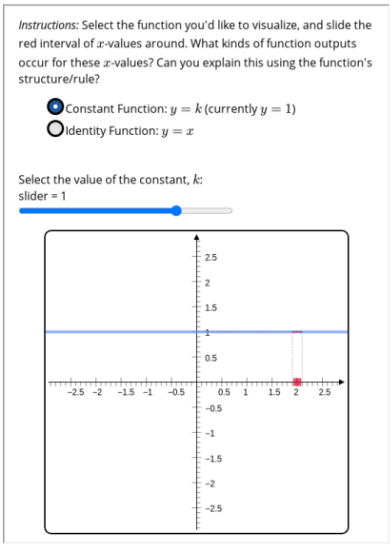
If $f(x)$ and $g(x)$ are two functions defined at x-values around, but maybe not at, $x = a$ and $\lim_{x\to a} f(x)$ and $\lim_{x\to a} g(x)$ both exist, then we can evaluate limits of combinations of these functions.

1. Sums: The limit of the sum of $f(x)$ and $g(x)$ is the sum of the limits of $f(x)$ and $g(x)$:
$$\lim_{x\to a}(f(x) + g(x)) = \lim_{x\to a} f(x) + \lim_{x\to a} g(x)$$
2. Differences: The limit of a difference of $f(x)$ and $g(x)$ is the difference of the limits of $f(x)$ and $g(x)$:
$$\lim_{x\to a}(f(x) - g(x)) = \lim_{x\to a} f(x) - \lim_{x\to a} g(x)$$
3. Coefficients: If k is some real number coefficient, then:
$$\lim_{x\to a} kf(x) = k \lim_{x\to a} f(x)$$
4. Products: The limit of a product of $f(x)$ and $g(x)$ is the product of the limits of $f(x)$ and $g(x)$:
$$\lim_{x\to a}(f(x) \cdot g(x)) = \left(\lim_{x\to a} f(x)\right) \cdot \left(\lim_{x\to a} g(x)\right)$$
5. Quotients: The limit of a quotient of $f(x)$ and $g(x)$ is the quotient of the limits of $f(x)$ and $g(x)$ (provided that you do not divide by 0):
$$\lim_{x\to a}\left(\frac{f(x)}{g(x)}\right) = \frac{(\lim_{x\to a} f(x))}{(\lim_{x\to a} g(x))} \quad \left(\text{ if } \lim_{x\to a} g(x) \neq 0\right)$$

We can summarize these properties: when we are thinking about our basic operations on functions, we can evaluate limits by just looking at the limits of each component function individually and then piecing those individual limit values back together.

This kind of structural "building-block" behavior is a really important one in mathematics. Whenever we define some new mathematical object, properties like this are typically good ideas for us to check in order to learn more about the object we've defined.

Ok, let's move on. We're going to turn our attention to something more concrete. We're going to think of two function types: constant functions and the identity function.



Standalone
Embed

Theorem 1.2.6: Limits of Two Basic Functions.

Let a be some real number.

1. Limit of a Constant Function: If k is some real number constant, then:
$$\lim_{x \to a} k = k$$
2. Limit of the Identity Function:
$$\lim_{x \to a} x = a$$

These two functions might seem pretty simplistic (most functions that we think of are more complicated than these), but we can use these to build more functions!

Activity 1.2.4: Limits of Polynomial Functions.

We're going to use a combination of properties from Theorem 1.2.5 and Theorem 1.2.6 to think a bit more deeply about polynomial functions. Let's consider a polynomial function:

$$f(x) = 2x^4 - 4x^3 + \frac{x}{2} - 5$$

a. We're going to evaluate the limit $\lim_{x \to 1} f(x)$. First, use the properties from Theorem 1.2.5 to rewrite this limit as 4 different limits added or subtracted together.
b. Now, for each of these limits, rewrite them as products of things until you have only limits of constants and identity functions, as in Theorem 1.2.6. Evaluate your limits.

c. Based on the definition of a limit (Definition 1.1.1), we normally say that $\lim_{x\to 1} f(x)$ is not dependent on the value of f(1). Why do we say this?
d. Compare the values of $\lim_{x\to 1} f(x)$ and f(1). Why do these values feel connected?
e. Come up with a new polynomial function: some combination of coefficients with x's raised to natural number exponents. Call your new polynomial function $g(x)$. Evaluate $\lim_{x\to -1} g(x)$ and compare the value to g(–1). Explain why these values are the same.
f. Explain why, for any polynomial function $p(x)$, the limit $\lim_{x\to a} p(x)$ is the same value as $p(a)$.

Answer available at the end of this section.

This leads us to an important result about a whole class of functions: polynomials! We can (finally) evaluate the limit of a polynomial without having to think too carefully about the distinction between the behavior of the function around $x = a$ and the behavior of the function at $x = a$.

Theorem 1.2.7: Limits of Polynomials.

If $p(x)$ is a polynomial function and a is some real number, then:

$$\lim_{x\to a} p(x) = p(a)$$

This result really just says that polynomials are friendly functions for limits: sure, a limit is really about the behavior of the function outputs around but not at x = a, but for polynomial functions, specifically, we can wave our hands and say "Ah, who cares, it's all the same anyways!"

Some questions that we might ask:

1. Are there other functions that have the same nice result about them that Theorem 1.2.7 says for polynomials?
2. Are there some typical functions that we'll work with where this result doesn't work (and we actually have to be aware of the behavior around a point instead of at it)?
3. What are we even going to use these limits for, anyways? Why do we care about these?

The answers to these questions will come slowly but surely, and we'll hopefully be able to start using these limits as a tool to think about more interesting and important topics soon: we just need to make sure we're familiar with them first.

Exercise 1.2.1

Given $\lim_{x\to 3} f(x) = 5$ and $\lim_{x\to 3} g(x) = -2$, evaluate the following limits. If the limit doesn't exist, explain why. Write out a few steps and explanations to justify your work.

a. $\lim_{x\to 3}\left(6f(x) - \frac{g(x)}{3}\right)$
b. $\lim_{x\to 3}(f(x))^2 g(x)$
c. $\lim_{x\to 3^-}\left(\frac{4g(x)}{f(x)} + 3f(x)\right)$

Answer available at the end of this section.

Exercise 1.2.2

Evaluate each limit. Justify your answers.

a. $\lim_{x\to 0}(4x^3 - 6x^2 + 7x - 10)$
b. $\lim_{x\to -2}(9 - 3x + x^2 + 3x^3 + x^4)$
c. $\lim_{t\to a}(9t^2 + 3at - 1)$ where a is some real number

d. $\lim_{s\to 1}\left(\dfrac{5s^2-6s+1}{s^2-4}\right)$

e. $\lim_{t\to 2}\left(\dfrac{4t-5}{6+t^2}\right)$

f. $\lim_{z\to 2}\left(\dfrac{z^2+z-4}{z+2}\right)$

Answer available at the end of this section.

Exercise 1.2.3

Evaluate each limit. If the limit does not exist, explain why not.

a. Let $f(x)=\begin{cases}3x-2 & \text{if } x<-1\\ x^2+x-4 & \text{if } x\geq -1\end{cases}$. Evaluate $\lim_{x\to -1^-} f(x), \lim_{x\to -1^+} f(x),$ and $\lim_{x\to -1} f(x)$.

b. Let $g(x)=\begin{cases}6+x & \text{if } x<-3\\ 6 & \text{if } x=-3\\ \dfrac{2x+15}{3} & \text{if } x>-3\end{cases}$. Evaluate $\lim_{x\to -3^-} g(x), \lim_{x\to -3^+} g(x),$ and $\lim_{x\to -3} g(x)$.

c. Let $s(t)=\begin{cases}t^2+1 & \text{if } t<1\\ 1-t^2 & \text{if } t\geq 1\end{cases}$. Evaluate $\lim_{t\to 1^-} s(t), \lim_{t\to 1^+} s(t),$ and $\lim_{t\to 1} s(t)$

d. Let $r(\theta)=\begin{cases}\theta-1+\theta^3 & \text{if } \theta<2\\ 3\theta+2 & \text{if } \theta\geq 2\end{cases}$. Evaluate $\lim_{\theta\to 2^-} r(\theta), \lim_{\theta\to 2^+} r(\theta),$ and $\lim_{\theta\to 2} r(\theta)$.

Answer available at the end of this section.

Answers

Activity 1.2.3: Combinations of Functions.

a.

Hint: You might think about writing out a function rule for $h(x)$. But can you also find h(2) without ever writing out a rule for $h(x)$?

Solution:

1. Since $h(x)=(x^2+3)+\left(x-\dfrac{1}{x}\right)$, we can evaluate $h(2)$ by:

$$\begin{aligned}h(2)&=2^2+3+2-\frac{1}{2}\\ &=\frac{17}{2}\end{aligned}$$

2. Since $h(x)=f(x)+g(x)$, we can evaluate $h(2)$ by:

$$\begin{aligned}f(2)&=7\\ g(2)&=\frac{3}{2}\\ h(2)&=f(2)+g(2)\\ &=7+\frac{3}{2}\\ &=\frac{17}{2}\end{aligned}$$

Activity 1.2.4: Limits of Polynomial Functions.

a. $\lim_{x\to 1}(2x^4)-\lim_{x\to 1}(4x^3)+\lim_{x\to 1}\left(\dfrac{x}{2}\right)-\lim_{x\to 1}5$

b. $2\left(\lim_{x\to 1}x\right)^4-4\left(\lim_{x\to 1}x\right)^3+\dfrac{1}{2}\left(\lim_{x\to 1}x\right)-\lim_{x\to 1}5$

Exercise 1.2.1

a.
$$\begin{aligned}\lim_{x\to 3}\left(6f(x)-\frac{g(x)}{3}\right)&=6\lim_{x\to 3}f(x)-\frac{\lim_{x\to 3}g(x)}{3}\\ &=6(5)-\frac{(-2)}{3}\\ &=\frac{92}{3}\end{aligned}$$

b.
$$\begin{aligned}\lim_{x\to 3}(f(x))^2g(x)&=\left(\lim_{x\to 3}f(x)\right)^2\left(\lim_{x\to 3}g(x)\right)\\ &=(5)^2(-2)\\ &=-50\end{aligned}$$

c.
$$\begin{aligned}\lim_{x\to 3^-}\left(\frac{4g(x)}{f(x)}+3f(x)\right) &= \frac{4\lim_{x\to 3^-} g(x)}{\lim_{x\to 3^-} f(x)} + 3\lim_{x\to 3^-} f(x)\\ &= \frac{4(-2)}{5}+3(5)\\ &= \frac{67}{5}\end{aligned}$$

Exercise 1.2.2

a. This is a polynomial, so we can evaluate this function at $x=0$: $4(0)^3-6(0)^2+7(0)-10=-10$

b. This is also a polynomial, and so we can evaluate the function at $x=-2$: $9-3(-2)+(-2)^2+3(-2)^3+(-2)^4=11$

c. $9(a)^2+3a(a)-1=12a^2-1$

d.
$$\begin{aligned}\lim_{s\to 1}\left(\frac{5s^2-6s+1}{s^2-4}\right) &= \frac{\lim_{s\to 1}(s^2-6s+1)}{\lim_{s\to 1}(s^2-4)}\\ &= \frac{0}{(-3)}\\ &= 0\end{aligned}$$

e.
$$\begin{aligned}\lim_{t\to 2}\left(\frac{4t-5}{6+t^2}\right) &= \frac{\lim_{t\to 2}(4t-5)}{\lim_{t\to 2}(6+t^2)}\\ &= \frac{3}{10}\end{aligned}$$

f.
$$\begin{aligned}\lim_{z\to 2}\left(\frac{z^2+z-4}{z+2}\right) &= \frac{\lim_{z\to 2}(z^2+z-4)}{\lim_{z\to 2}(z+2)}\\ &= \frac{2}{4}\\ &= \frac{1}{2}\end{aligned}$$

Exercise 1.2.3

a.
$$\lim_{x\to -1^-} f(x) = -5$$
$$\lim_{x\to -1^+} f(x) = -4$$

Since the left-sided limit and right-sided limit do not match, we say that

$$\lim_{x\to -1} f(x)$$

does not exist.

b.
$$\lim_{x\to -3^-} g(x) = 3$$
$$\lim_{x\to -3^+} g(x) = 3$$

Since the left-sided limit and right-sided limit match, we say that

$$\lim_{x\to -3} g(x) = 3$$

c.
$$\lim_{t\to 1^-} s(t) = 2$$
$$\lim_{t\to 1^+} s(t) = 0$$

Since the left-sided limit and right-sided limit do not match, we say that

$$\lim_{t\to 1} s(t)$$

does not exist.

d.
$$\lim_{\theta\to 2^-} r(\theta) = 9$$
$$\lim_{\theta\to 2^+} r(\theta) = 8$$

Since the left-sided limit and right-sided limit do not match, we say that

$$\lim_{\theta\to 2} r(\theta)$$

does not exist.

This page titled 1.2: Evaluating Limits is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Peter Keep.

1.3: First Indeterminate Forms

We're going to really focus on one of the main aspects of a limit in this next activity. The activity should serve two purposes:

1. We'll review a really important property or aspect of what a limit is!
2. We'll look at this thing that we already know from a slightly different perspective (or maybe just a specific perspective), and we'll discover a really important and helpful result from it!

Activity 1.3.1: Limits of (Slightly) Different Functions.

a. Using the graph of $f(x)$ below, approximate $\lim_{x\to 1} f(x)$.

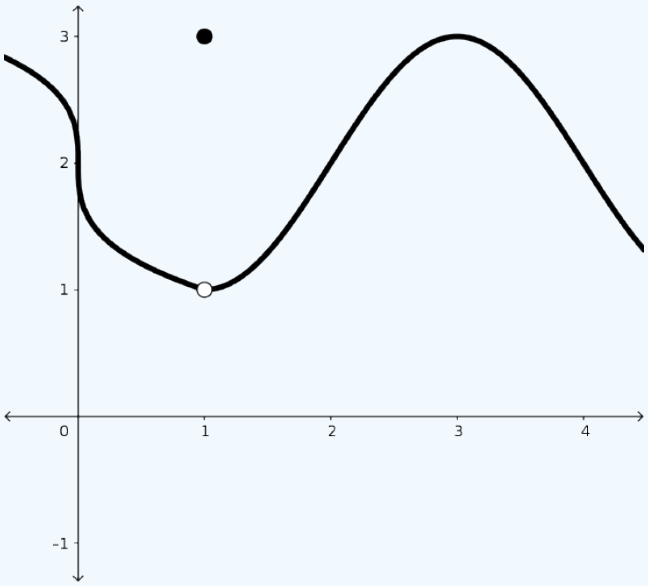


Figure 1.3.1

b. Using the graph of the slightly different function $g(x)$ below, approximate $\lim_{x\to 1} g(x)$.

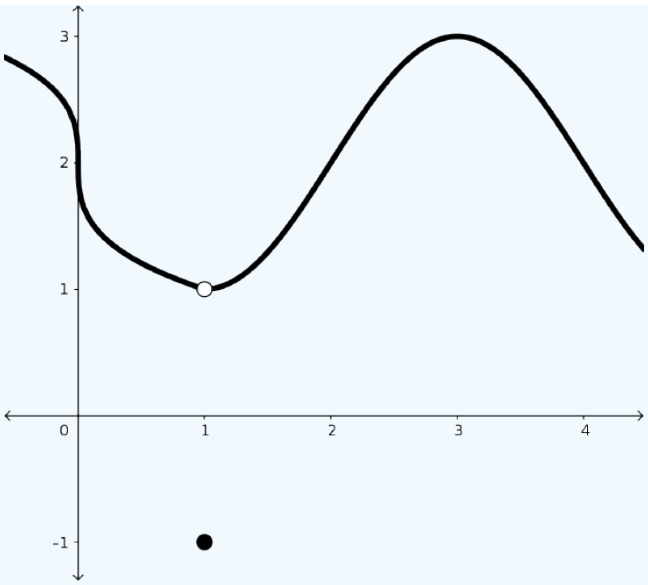


Figure 1.3.2

c. Compare the values of $f(1)$ and $g(1)$ and discuss the impact that this difference had on the values of the limits.
d. For the function $r(t)$ defined below, evaluate the limit $\lim_{t\to 4} r(t)$.

$$r(t) = \begin{cases} 2t - \dfrac{4}{t} & \text{when } t < 4 \\ 8 & \text{when } t = 4 \\ t^2 - t - 5 & \text{when } t > 4 \end{cases}$$

e. For the slightly different function $s(t)$ defined below, evaluate the limit $\lim_{t\to 4} s(t)$.

$$s(t) = \begin{cases} 2t - \dfrac{4}{t} & \text{when } t \leq 4 \\ t^2 - t - 5 & \text{when } t > 4 \end{cases}$$

f. Do the changes in the way that the function was defined impact the evaluation of the limit at all? Why not?

This is an important thing to notice: we can change our function by changing the value of the function at $x = a$ without changing the value of the limit as $x \to a$.

Theorem 1.3.3: Limits of (Slightly) Different Functions.

If $f(x)$ and $g(x)$ are two functions defined at x-values around a (but maybe not at $x = a$ itself) with $f(x) = g(x)$ for the x-values around a but with $f(a) \neq g(a)$ then $\lim_{x\to a} f(x) = \lim_{x\to a} g(x)$, if the limits exist.

Why will this be helpful? At the end of Section 1.2 we found that some functions (polynomials) are great: the limit of these functions is the same as the function value at the point (Theorem 1.2.7). This is a special case, and many functions won't be so nice to work with. But maybe we could use Theorem 1.3.3 to swap out an annoying-to-work-with function for a nice-to-work-with function!

A First Introduction to Indeterminate Forms

So before we begin applying this result, we will focus on a situation where we need it. We're going to do something strange: define a situation before we experience it.

Definition 1.3.4: Indeterminate Form.

We say that a limit has an indeterminate form if the general structure of the limit could take on any different value, or not exist, depending on the specific circumstances.

For instance, if $\lim_{x\to a} f(x) = 0$ and $\lim_{x\to a} g(x) = 0$, then we say that the limit $\lim_{x\to a}\left(\frac{f(x)}{g(x)}\right)$ has an indeterminate form. We typically denote this using the informal symbol $\frac{0}{0}$, as in:

$$\lim_{x\to a}\left(\frac{f(x)}{g(x)}\right) \overset{?}{\to} \frac{0}{0}$$

Ok, so why do we need this definition? What does the word "indeterminate" even mean, here?

We're going to see in this next activity that this kind of $\frac{0}{0}$ form of a limit can actually lead to very different behavior: we call it indeterminate because we cannot determine, based solely on the form $\frac{0}{0}$, what the limit is or even if it will exist.

Exercise 1.3.2

a. Were going to evaluate $\lim_{x\to 3}\left(\frac{x^2-7x+12}{x-3}\right)$.
   - First, check that we get the indeterminate form $\frac{0}{0}$ when $x \to 3$.
   - Now we want to find a new function that is equivalent to $f(x) = \frac{x^2-7x+12}{x-3}$ for all x-values other than $x = 3$. Try factoring the numerator, $x^2 - 7x + 12$. What do you notice?
   - "Cancel" out any factors that show up in the numerator and denominator. Make a special note about what that factor is.
   - This function is equivalent to $f(x) = \frac{x^2-7x+12}{x-3}$ except at $x = 3$. The difference is that this function has an actual function output at $x = 3$, while f(x) doesn't. Evaluate the limit as $x \to 3$ for your new function.

b. Now we'll evaluate a new limit: $\lim_{x\to 1}\left(\frac{\sqrt{x^2+3}-2}{x^2-5x+4}\right)$.
   - First, check that we get the indeterminate form $\frac{0}{0}$ when $x \to 1$.
   - Now we want a new function that is equivalent to $g(x) = \frac{\sqrt{x^2+3}-2}{x^2-5x+4}$ for all x-values other than $x = 1$. Try multiplying the numerator and the denominator by $(\sqrt{x^2+3}+2)$. We'll call this the "conjugate" of the numerator.
   - In your multiplication, confirm that $\left(\sqrt{x^2+3}-2\right)\left(\sqrt{x^2+3}+ +2\right) = (x^2+3) - 4$.
   - Try to factor the new numerator and denominator. Do you notice anything? Can you "cancel" anything? Make another note of what factor(s) you cancel.
   - This function is equivalent to $g(x) = \frac{\sqrt{x^2+3}-2}{x^2-5x+4}$ except at $x = 1$. The difference is that this function has an actual function output at $x = 1$, while $g(x)$ doesn't. Evaluate the limit as $x \to 1$ for your new function.

c. Our last limit in this activity is going to be $\lim_{x\to -2}\left(\frac{3-\frac{3}{x+3}}{x^2+2x}\right)$.
   - Again, check to see that we get the indeterminate form $\frac{0}{0}$ when $x \to -2$.
   - Again, we want a new function that is equivalent to $h(x) = \frac{3-\frac{3}{x+3}}{x^2+2x}$ for all x-values other than $x = -2$. Try completing the subtraction in the numerator, $3 - \frac{3}{x+3}$, using "common denominators."
   - Try to factor the new numerator and denominator(s). Do you notice anything? Can you "cancel" anything? Make another note of what factor(s) you cancel.

- For the final time, we've found a function that is equivalent to $h(x) = \dfrac{3 - \dfrac{3}{x+3}}{x^2+2x}$ except at $x = -2$. The difference is that this function has an actual function output at $x = -2$, while $h(x)$ doesn't. Evaluate the limit as $x \to -2$ for your new function.

d. In each of the previous limits, we ended up finding a factor that was shared in the numerator and denominator to cancel. Think back to each example and the factor you found. Why is it clear that these must have been the factors we found to cancel?

e. Let's say we have some new function $f(x)$ where $\lim_{x\to 5} f(x) \overset{?}{\to} \dfrac{0}{0}$. You know, based on these examples, that you're going to apply some algebra trick to rewrite your function, factor, and cancel. Can you predict what you will end up looking for to cancel in the numerator and denominator? Why?

This is great: we're applying Theorem 1.3.3 because in each algebraic manipulation, we change the domain of the function by removing some factor from the denominator!

These three algebra tricks are all we'll look at for now. In reality, there are plenty of little tricky manipulations we can use to slightly change functions, but if we focused on trying to build one for every situation we could run into, we'd spend the rest of this text just outlining different algebra tricks for different situations.

Algebra Tricks for Indeterminate Forms.

For limits with the $\frac{0}{0}$ indeterminate form, we can apply the following algebraic tricks:

1. Factor and cancel: This works well when we have polynomials divided by polynomials.
2. Conjugates: This works well when we have some difference of square roots in the numerator or denominator.
3. Combine fractions with common denominators: This works well when we have some subtraction with fractions inside of a numerator or denominator of another fraction.

What if There Is No Algebra Trick?

We've seen some nice examples above where we were able to use some algebra to manipulate functions in such was as to force some shared factor in the numerator and denominator into revealing itself. From there, we were able to apply Theorem 1.3.3 and swap out our problematic function with a new one, knowing that the limit would be the same.

But what if we can't do that? What if the specific structure of the function seems resistant somehow to our attempts at wielding algebra?

This happens a lot, and we'll investigate some more of those types of limits in Section 4.7. For now, though, let's look at a very famous limit and reason our way through the indeterminate form.

Exercise 1.3.3

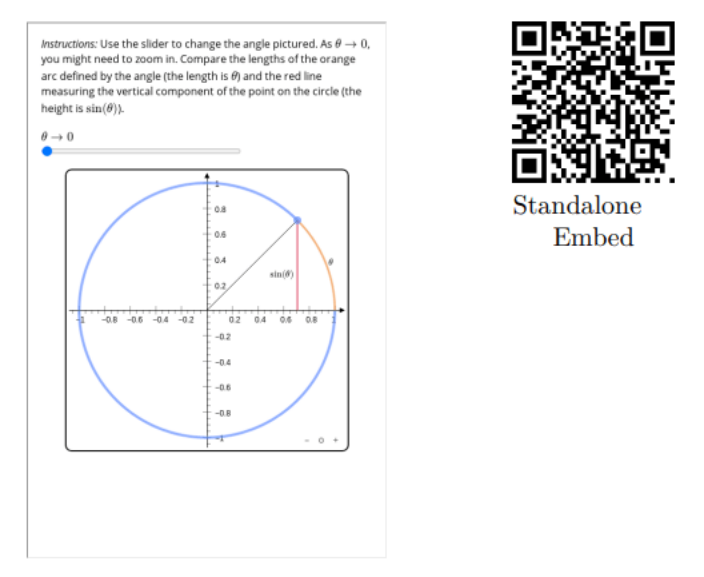
Let's consider a new limit:

$$\lim_{\theta\to 0} \frac{\sin(\theta)}{\theta}$$

This one is strange!

a. Notice that this function, $f(\theta) = \dfrac{\sin(\theta)}{\theta}$, is resistent to our algebra tricks:
   - There's nothing to "factor" here, since our trigonometric function is not a polynomial.
   - We can't use a trick like the "conjugate" to multiply and rewrite, since there's no square roots and also only one term in the numerator.
   - There aren't any fractions that we can combine by addition or subtraction.

b. Be frustrated at this new limit for resisting our algebra tricks.
c. Now let's think about the meaning of $\sin(\theta)$ and even θ in general. In this text, we will often use Greek letters, like θ, to represent angles. In general, these angles will be measured in radians (unless otherwise specified). So what does the sine function do or tell us? What is a radian?
   - Hint 1: On the unit circle, if we plot some point at an angle of θ, then the coordinates of that point can be represented with trig functions! Which ones?
   - Hint 2: The length of the curve defining a unit circle is 2π. This also corresponds to the angle we would use to represent moving all the way around the circle. What must the length of the portion of the circle be up to some point at an angle θ?
d. Let's visualize our limit, then, by comparing the length of the arc and the height of the point as $\theta \to 0$.



e. Explain to yourself, until you are absolutely certain, why the two lengths must be the same in the limit as $\theta \to 0$. What does this mean about $\lim_{\theta \to 0} \dfrac{\sin(\theta)}{\theta}$?

Practice Problems

Exercise 1.3.1

Explain, in your own words, why Theorem 1.3.3 is true.

Exercise 1.3.2

Consider the following limit:

$$\lim_{x \to 3} \left(\frac{2x^2 - 5x - 3}{x^2 + 5x - 24} \right)$$

a. Confirm that this limit has an indeterminate form.
b. Evaluate the limit.
c. When you were evaluating the limit, you likely "cancelled" a factor of $(x - 3)$ from the numerator and denominator. Why might you have known, before factoring anything, that $(x - 3)$ would be a factor shared in the numerator and denominator? For instance, how did you know it wasn't going to be $(x - 4)$ or $(x + 1)$ or something else?

Answer available at the end of this section.

Exercise 1.3.3

Consider the following limit:

$$\lim_{x\to 4}\left(\frac{\dfrac{2}{x+1}-\dfrac{x-2}{5}}{x-4}\right)$$

a. Confirm that this limit has an indeterminate form.
b. Evaluate the limit.
c. When you were evaluating the limit, you likely "cancelled" a factor of $(x-4)$ from the numerator and denominator. Why might you have known, before factoring anything, that $(x-4)$ would be a factor shared in the numerator and denominator? For instance, how did you know it wasn't going to be $(x-3)$ or $(x+1)$ or something else?

Answer available at the end of this section.

Exercise 1.3.4

Use the algebra tricks from Algebra Tricks for Indeterminate Forms to evaluate each limit.

a. $\lim_{x\to 2}\left(\dfrac{x^2-4}{x^2+3x-10}\right)$
b. $\lim_{x\to -1}\left(\dfrac{2x^2+x-1}{x+1}\right)$
c. $\lim_{x\to 4}\left(\dfrac{4-x}{x-4}\right)$
d. $\lim_{x\to 1}\left(\dfrac{\sqrt{x+3}-2}{x-1}\right)$
e. $\lim_{x\to 2}\left(\dfrac{\sqrt{x^2-1}-\sqrt{x+1}}{x-2}\right)$
f. $\lim_{h\to 0}\left(\dfrac{\sqrt{a+h}-\sqrt{a}}{h}\right)$ where a is some non-negative real number
g. $\lim_{t\to 3}\left(\dfrac{\dfrac{1}{t}-\dfrac{1}{3}}{t-3}\right)$
h. $\lim_{t\to 6}\left(\dfrac{\dfrac{t+1}{t-1}-\dfrac{7}{11-t}}{t-6}\right)$
i. $\lim_{h\to 0}\left(\dfrac{\dfrac{1}{a+h}-\dfrac{1}{a}}{h}\right)$ where a is some real number

Answer available at the end of this section.

Answers

Exercise 1.3.2

a. $\lim_{x\to 3}\left(\dfrac{2x^2-5x-3}{x^2+5x-24}\right)\overset{?}{\to}\dfrac{0}{0}$

b.
$$\lim_{x\to 3}\left(\frac{2x^2-5x-3}{x^2+5x-24}\right)=\lim_{x\to 3}\frac{(x-3)(2x+1)}{(x-3)(x+8)}$$
$$=\lim_{x\to 3}\frac{(2x+1)}{(x+8)}$$
$$=\frac{7}{11}$$

Exercise 1.3.3

a. $\lim_{x\to 4}\left(\dfrac{\dfrac{2}{x+1}-\dfrac{x-2}{5}}{x-4}\right)\overset{?}{\to}\dfrac{0}{0}$

b.
$$\lim_{x\to 4}\left(\frac{\frac{2}{x+1}-\frac{x-2}{5}}{x-4}\right)=\lim_{x\to 4}\left(\frac{\frac{10-(x^2-x-2)}{5(x+1)}}{x-4}\right)$$
$$=\lim_{x\to 4}\left(\frac{-(x-4)(x+3)}{5(x+1)}\right)\left(\frac{1}{x-4}\right)$$
$$=\lim_{x\to 4}\left(-\frac{x+3}{5(x+1)}\right)$$
$$=-\frac{7}{25}$$

Exercise 1.3.4

a. $lim_{x\to 2}\left(\dfrac{x^2-4}{x^2+3x-10}\right)=\dfrac{4}{7}$

b. $\lim_{x\to -1}\left(\dfrac{2x^2+x-1}{x+1}\right)=3$

c. $\lim_{x\to 4}\left(\dfrac{4-x}{x-4}\right)=-1$

d. $\lim_{x\to 1}\left(\dfrac{\sqrt{x+3}-2}{x-1}\right)=\dfrac{1}{4}$

e. $\lim_{x\to 2}\left(\dfrac{\sqrt{x^2-1}-\sqrt{x+1}}{x-2}\right)=\dfrac{3}{2\sqrt{3}}$

f. $\lim_{h\to 0}\left(\dfrac{\sqrt{a+h}-\sqrt{a}}{h}\right)=\dfrac{1}{2\sqrt{a}}$

g. $\lim_{t\to 3}\left(\dfrac{\dfrac{1}{t}-\dfrac{1}{3}}{t-3}\right)=-\dfrac{1}{9}$

h. $\lim_{t\to 6}\left(\dfrac{\dfrac{t+1}{t-1}-\dfrac{7}{11-t}}{t-6}\right)=-\dfrac{9}{25}$

i. $\lim_{h\to 0}\left(\dfrac{\dfrac{1}{a+h}-\dfrac{1}{a}}{h}\right)=-\dfrac{1}{a^2}$

This page titled 1.3: First Indeterminate Forms is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Peter Keep.

1.4: Limits Involving Infinity

Two types of limits involving infinity. In both cases, we'll mostly just consider what happens when we divide by small things and what happens when we divide by big things. We can summarize this here, though:

> Fractions with small denominators are big, and fractions with big denominators are small.

Infinite Limits

When we talked about indeterminate forms in Section 1.3, we noticed that the function value wasn't defined, since we divided by 0. Specifically, we were looking at the $\frac{0}{0}$ form in the limit. What happens in other cases when we divide by 0 with a nonzero numerator?

Activity 1.4.1: What Happens When We Divide by 0?

First, let's make sure we're clear on one thing: there is no real number than is represented as some other number divided by 0.

When we talk about "dividing by 0" here (and in Section 1.3), we're talking about the behavior of some function in a limit. We want to consider what it might look like to have a function that involves division where the denominator gets arbitrarily close to 0 (or, the limit of the denominator is 0).

a. Remember when, once upon a time, you learned that dividing one a number by a fraction is the same as multiplying the first number by the reciprocal of the fraction? Why is this true?
b. What is the relationship between a number and its reciprocal? How does the size of a number impact the size of the reciprocal? Why?
c. Consider $12 \div N$. What is the value of this division problem when:
   - $N = 6$?
   - $N = 4$?
   - $N = 3$?
   - $N = 2$?
   - $N = 1$?
d. Let's again consider $12 \div N$. What is the value of this division problem when:
   - $N = \frac{1}{2}$?
   - $N = \frac{1}{3}$?
   - $N = \frac{1}{4}$?
   - $N = \frac{1}{6}$?
   - $N = \frac{1}{1000}$?
e. Consider a function $f(x) = \frac{12}{x}$. What happens to the value of this function when $x \to 0^+$? Note that this means that the x-values we're considering most are very small and positive.
f. Consider a function $f(x) = \frac{12}{x}$. What happens to the value of this function when $x \to 0^-$? Note that this means that the x-values we're considering most are very small and negative.

So we want to formalize this kind of behavior. We know that the limits that we're looking at don't exist (since there isn't a single, real number that the function outputs are all close to), but there is definitely some sort of consistent behavior that we'd like to signify.

Definition 1.4.1: Infinite Limit.

We say that a function $f(x)$ has an infinite limit at a if $f(x)$ is arbitrarily large (positive or negative) when x is sufficiently close to, but not equal to, $x = a$. We would then say, depending on the sign of the values of $f(x)$, that:

$$\lim_{x \to a^-} f(x) = \pm\infty \qquad \lim_{x \to a^+} f(x) = \pm\infty$$

If the sign of both one-sided limits are the same, we can say that $\lim_{x \to a} f(x) = \pm\infty$ (depending on the sign), but it is helpful to note that, by the definition of the Limit of a Function, this limit does not exist, since $f(x)$ is not arbitrarily close to a single real number.

We know, from our adage "Fractions with small denominators are big, and fractions with big denominators are small," that this type of behavior happens when the denominator is tiny compared to the numerator. We can summarize this:

Theorem 1.4.2: Dividing by 0 in a Limit.

If $f(x) = \dfrac{g(x)}{h(x)}$ with $\lim_{x \to a} g(x) \neq 0$ and $\lim_{x \to a} h(x) = 0$, then $f(x)$ has an Infinite Limit at a. We will often denote this behavior as:

$$\lim_{x \to a} f(x) \overset{?}{\to} \frac{\#}{0}$$

where # is meant to be some shorthand representation of a non-zero limit in the numerator (often, but not necessarily, some real number).

These kinds of limits are great, in that they're consistent to identify. We just look for this tiny denominator compared to the numerator, and go from there. We also know a lot about these types of limits, and can summarize this below.

Evaluating Infinite Limits.

Once we know that $\lim_{x \to a} f(x) \overset{?}{\to} \dfrac{\#}{0}$, we know a bunch of information right away!

- This limit doesn't exist.
- The function $f(x)$ has a vertical asymptote at $x = a$, causing these unbounded y-values near $x = a$.
- The one sided limits must be either ∞ or $-\infty$.
- We only need to focus on the sign of the one sided limits! And signs of products and quotients are easy to follow.

So a pretty typical process is to factor as much as we can, and check the sign of each factor (in a numerator or denominator) as $x \to a^-$ and $x \to a^+$. From there, we can find the sign of $f(x)$ in both of those cases, which will tell us the one-sided limit.

Example 1.4.1

For each function, find the relevant one-sided limits at the input-value mentioned. If you can use a two-sided limit statement to discuss the behavior of the function around this input-value, then do so.

a. $\left(\dfrac{2x^2 - 5x + 1}{x^2 + 8x + 16}\right)$ and $x = -4$

Solution

Let's first factor the denominator: we want to see the factor $(x + 4)$, where we divide by 0.

$$\left(\frac{2x^2 - 5x + 1}{x^2 + 8x + 16}\right) = \left(\frac{2x^2 - 5x + 1}{(x + 4)^2}\right)$$

Now, when $x \to 4$, we have $\left(\frac{2x^2-5x+1}{(x+4)^2}\right) \overset{?}{\to} \frac{13}{0}$. This is our $\frac{\#}{0}$ form that tells us we have an infinite limit. We're going to look at the one-sided limits, but let's notice something:

- For x-values close to 4, we expect that the numerator will be close to 13. No matter what "side" of 13 it is, it's still going to be positive.
- For x-values close to 4, we expect the denominator to be close to 0. Depending on what side of 0 it is, it could be positive or negative.

So in our one-sided limits, we know that these both will be infinite limits (the function values will either approach $-\infty$ or ∞). The only difference in these is the sign. So let's check the signs!

First, we'll consider $x \to -4^-$. That's $x < -4$. We know the numerator is positive, since it is close to 13, and we also know that the factor $x + 4 < 0$.

$$\begin{aligned}\lim_{x\to -4^-}\left(\frac{2x^2-5x+1}{(x+4)^2}\right) &= \lim_{x\to -4^-}(\frac{\overbrace{2x^2-5x+1}^{\approx 13}}{\underbrace{(x+4)^2}_{\approx 0}}) \\ &= \frac{\oplus}{(\ominus)^2} \\ &= \oplus\end{aligned}$$

So we know that $\lim_{x\to -4^-}\left(\frac{2x^2-5x+1}{(x+4)^2}\right)$ is an infinite limit, and we know that it is positive.

$$\lim_{x\to -4^-}\left(\frac{2x^2-5x+1}{(x+4)^2}\right) = \infty$$

Now we can check the signs of these factors when $x \to -4^+$ (and so $x > -4$). The numerator is still close to 13, and so still negative. Now, $(x + 4)$ is positive, and is still being squared.

$$\begin{aligned}\lim_{x\to -4^+}\left(\frac{2x^2-5x+1}{(x+4)^2}\right) &= \lim_{x\to -4^+}(\frac{\overbrace{2x^2-5x+1}^{\approx 13}}{(x^{x+4})^2}) \\ &= \frac{\oplus}{(\oplus)^2} \\ &= \oplus\end{aligned}$$

Since this is an infinite limit and is also positive, we know that:

$$\lim_{x\to -4^+}\left(\frac{2x^2-5x+1}{(x+4)^2}\right) = \infty$$

Since the function approaches ∞ on both sides, we can say that

$$\lim_{x\to -4}\left(\frac{2x^2-5x+1}{(x+4)^2}\right) = \infty,$$

but we know that this limit does not exist (since the function values are not close to a single real number).

b. $\left(\frac{4x^2-x^5}{x^2-4x+3}\right)$ and $x = 1$

Solution

We can start this in a similar way: factor the denominator to see the places where we divide by 0:

$$\left(\frac{4x^2-x^5}{x^2-4x+3}\right) = \left(\frac{4x^2-x^5}{(x-3)(x-1)}\right)$$

Now, when we think about the limit as $x \to 1^-$, we're thinking about $x < 1$. We can check the signs, again! The numerator has $4x^2 - x^5 \to 3$, and so for x-values slightly smaller than 1, this numerator is close to 3, and so positive. Similarly, $(x-3) \to -2$, and so for x-values close to but less than 1, this is negative. Then, we can see that $(x-1) \to 0$. This is the part that gives us the $\frac{\#}{0}$ form.

For $x < 1$, we have $(x-1) < 0$, and so this is negative.

$$\lim_{x \to 1^-} \left(\frac{\overbrace{4x^2 - x^5}^{\oplus}}{\underbrace{(x-3)}_{\ominus}\underbrace{(x-1)}_{\ominus}} \right) \to \oplus$$

And so this limit is a positive infinite limit:

$$\lim_{x \to 1^-} \left(\frac{4x^2 - x^5}{(x-3)(x-1)} \right) = \infty$$

Now we can similarly check the signs when $x \to 1^+$, which is when $x > 1$. Note that, since x is still close to 1, the numerator and the factor $(x-3)$ will retain their sign. But, for the factor $(x-1)$, we get $(x-1) > 0$.

$$\lim_{x \to 1^-} \left(\frac{\overbrace{4x^2 - x^5}^{\oplus}}{\underbrace{(x-3)}_{\ominus}\underbrace{(x-1)}_{\oplus}} \right) \to \ominus$$

And so this limit is a negative infinite limit:

$$\lim_{x \to 1^+} \left(\frac{4x^2 - x^5}{(x-3)(x-1)} \right) = -\infty$$

c. $\sec(\theta)$ and $\theta = \frac{\pi}{2}$

Solution

We can think of $\sec(\theta)$ as a reciprocal: $\frac{1}{\cos(\theta)}$. Now, we can see that $\cos\left(\frac{\pi}{2}\right) = 0$, hence this is an infinite limit.

Let's visualize $\cos\left(\frac{\pi}{2}\right)$, so that we can tell the sign of this denominator when θ is on either side of $\frac{\pi}{2}$.

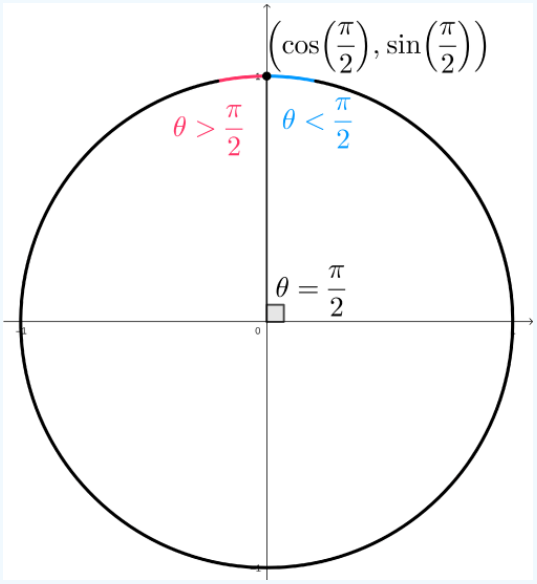


Figure 1.4.4

We can see that for $\theta < \frac{\pi}{2}$, we are looking at a point in the first quadrant with a positive horizontal component. So, in this case, $\cos(\theta) > 0$.

For the case when $\theta > \frac{\pi}{2}$, though, we are looking in the second quadrant with a negative horizontal component. So we see that $\cos(\theta) < 0$.

All of this to say:

$$\lim_{\theta \to \frac{\pi}{2}^-} \sec(\theta) = \infty$$
$$\lim_{\theta \to \frac{\pi}{2}^+} \sec(\theta) = -\infty$$

These are limits where $f(x) \to \pm\infty$ when $x \to a$ where a is a real number. What about the other way around? Is there a case where $x \to \pm\infty$, and what would we be looking at in the behavior of $f(x)$?

End Behavior Limits

Activity 1.4.2: What Happens When We Divide by Infinity?

Again, we need to start by making something clear: if we were really going to try divide some real number by infinity, then we would need to rebuild our definition of what it means to divide. In the context we're in right now, we only have division defined as an operation for real (and maybe complex) numbers. Since infinity is neither, then we will not literally divide by infinity.

When we talk about "dividing by infinity" here, we're again talking about the behavior of some function in a limit. We want to consider what it might look like to have a function that involves division where the denominator gets arbitrarily large (positive or negative) (or, the limit of the denominator is infinite).

a. Let's again consider $12 \div N$. What is the value of this division problem when:
   - $N = 1$?
   - $N = 6$?
   - $N = 12$?
   - $N = 24$?
   - $N = 1000$?
b. Let's again consider $12 \div N$. What is the value of this division problem when:
   - $N = -1$?
   - $N = -6$?
   - $N = -12$?
   - $N = -24$?
   - $N = -1000$?
c. Consider a function $f(x) = \frac{12}{x}$. What happens to the value of this function when $x \to \infty$? Note that this means that the x-values we're considering most are very large and positive.
d. Consider a function $f(x) = \frac{12}{x}$. What happens to the value of this function when $x \to -\infty$? Note that this means that the x-values we're considering most are very large and negative.
e. Why is there no difference in the behavior of $f(x)$ as $x \to \infty$ compared to $x \to -\infty$ when the sign of the function outputs are opposite ($f(x) > 0$ when $x \to \infty$ and $f(x) < 0$ when $x \to -\infty$)?

Definition 1.4.5: Limit at Infinity.

If $f(x)$ is defined for all large and positive x-values and $f(x)$ gets arbitrarily close to the single real number L when x gets sufficiently large, then we say:

$$\lim_{x\to\infty} f(x) = L$$

Similarly, if $f(x)$ is defined for all large and negative x-values and $f(x)$ gets arbitrarily close to the single real number L when x gets sufficiently negative, then we say:

$$\lim_{x\to-\infty} f(x) = L$$

In the case that $f(x)$ has a limit at infinity that exists, then we say $f(x)$ has a horizontal asymptote at $y = L$.

Lastly, if $f(x)$ is defined for all large and positive (or negative) x-values and $f(x)$ gets arbitrarily large and positive (or negative) when x gets sufficiently large (or negative), then we could say:

$$\lim_{x\to-\infty} f(x) = \pm\infty \text{ or } \lim_{x\to\infty} f(x) = \pm\infty$$

Because the primary focus for limits at infinity is the end behavior of a function, we will often refer to these limits as end behavior limits.

We're going to think about end-behavior of many types of functions, but we want to start with some small examples and build from there. So we're going to start by thinking about power functions. We're going to think about power functions in two ways:

1. Power functions in the form of x^p (with $p > 0$). It shouldn't be too much work to convince yourself that these functions always have limits where, as $x \to \pm\infty, x^p \to \pm\infty$ as well.
2. Reciprocal power functions, in the form $\dfrac{1}{x^p}$ (still with $p > 0$). We already should have an idea of what's going to happen in these, based on Activity 1.4.2.

Theorem 1.4.6: End Behavior of Reciprocal Power Functions.

If p is a positive real number, then:

$$\lim_{x\to\infty}\left(\frac{1}{x^p}\right) = 0 \text{ and } \lim_{x\to-\infty}\left(\frac{1}{x^p}\right) = 0$$

Our last result is one that you might already know, but we'll provide some more justification for this. We can use our knowledge of end-behavior limits for both type of power functions to think about the end-behavior limits of polynomials in general!

Theorem 1.4.7: Polynomial End Behavior Limits.

For some polynomial function:

$$p(x) = a_n x^n + a_{n-1}x^{n-1} + \ldots + a_2 x^2 + a_1 x + a_0$$

with n a positive integer (the degree) and all of the coefficients $a_0, a_1, \ldots, a_n$ real numbers (with $a_n \neq 0$), then

$$\lim_{x\to\infty} p(x) = \lim_{x\to\infty} a_n x^n$$

That is, the leading term (the term with the highest exponent) defines the end behavior for the whole polynomial function.

Proof. available at the end of this section.

Instead of spending our time thinking about these polynomial end-behavior limits too specifically (since we might already be

familiar with this result), and just focus on using these polynomials in the middle of larger problems.

Example 1.4.8

For each function, find the limits as $x \to \infty$ and $x \to -\infty$.

a. $\left(\dfrac{2x^2-5x+1}{x^2+8x+16}\right)$

Hint 1. You can think about the limit in the numerator and the limit in the denominator. Based on Theorem 1.4.7, which terms will be the ones to dictate the behavior of the numerator and denominator?

Hint 2. What happens when you think about just those dominant terms in the numerator and denominator and reduce the fraction? What is left?

Solution

We'll start with the same kind of factoring that is used in the proof of Theorem 1.4.7.

$$\begin{aligned}\left(\frac{2x^2-5x+1}{x^2+8x+16}\right) &= \left(\frac{2x^2}{x^2}\right)\left(\frac{1-\dfrac{5}{2x}+\dfrac{1}{2x^2}}{1+\dfrac{8}{x}+\dfrac{16}{x^2}}\right)\\ &= 2\left(\frac{1-\dfrac{5}{2x}+\dfrac{1}{2x^2}}{1+\dfrac{8}{x}+\dfrac{16}{x^2}}\right)\end{aligned}$$

Now we can apply the limits as $x \to \infty$ and $x \to -\infty$, since the reciprocal power functions will all $\to 0$.

$$\lim_{x\to\infty} 2\left(\frac{1-\dfrac{5}{2x}+\dfrac{1}{2x^2}}{1+\dfrac{8}{x}+\dfrac{16}{x^2}}\right) = 2(1) = 2$$

$$\lim_{x\to-\infty} 2\left(\frac{1-\dfrac{5}{2x}+\dfrac{1}{2x^2}}{1+\dfrac{8}{x}+\dfrac{16}{x^2}}\right) = 2(1) = 2$$

Alternatively, we could have done the following:

$$\begin{aligned}\lim_{x\to\infty}\left(\frac{2x^2-5x+1}{x^2+8x+16}\right) &= \frac{\lim_{x\to\infty}(2x^2-5x+1)}{\lim_{x\to\infty}(x^2+8x+16)}\\ &= \frac{\lim_{x\to\infty} 2x^2}{\lim_{x\to\infty} x^2}\\ &= \lim_{x\to\infty}\frac{2x^2}{x^2}\\ &= \lim_{x\to\infty} 2 = 2\end{aligned}$$

The same process could be used to show that:

$$\lim_{x\to-\infty}\left(\frac{2x^2-5x+1}{x^2+8x+16}\right) = 2$$

b. $\left(\dfrac{4x^2-x^5}{x^2-4x+3}\right)$

Hint. Be careful about which term (in the numerator) will persist!

c. $\dfrac{|x|}{3x}$

Hint. It might be helpful to remind yourself of the definition of the absolute value function:

$$|x| = \begin{cases} -x & \text{when } x < 0 \\ x & \text{when } x \geq 0 \end{cases}$$

This means you can replace $|x|$ with either x or $-x$ in the limits as $x \to \infty$ and $x \to -\infty$.

Practice Problems

Exercise 1.4.1

Use the graph of the function $f(x)$ below to determine each limit.

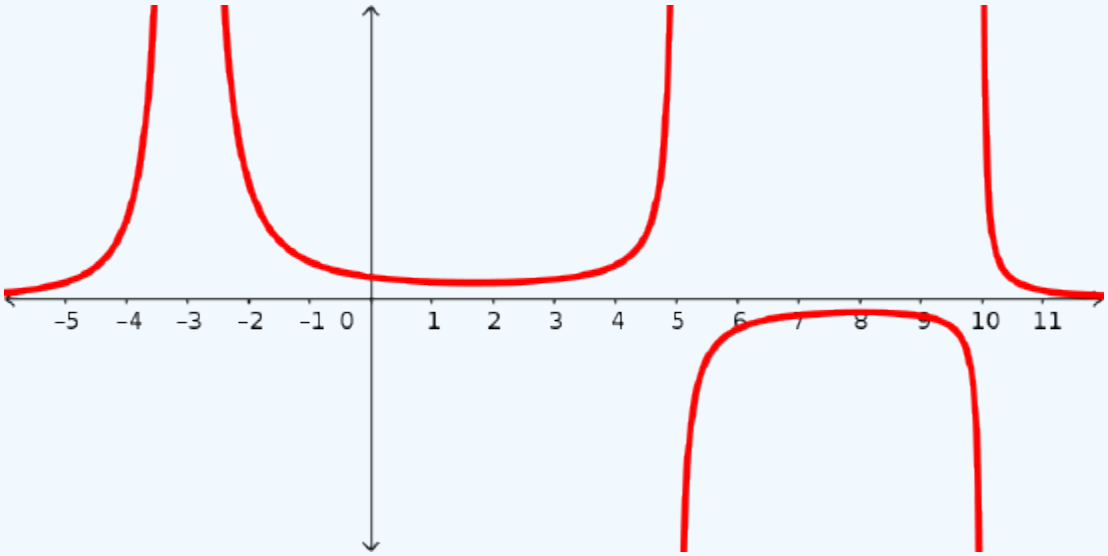


Figure 1.4.9 The function $f(x)$.

a. $\lim_{x \to -3^-} f(x)$
b. $\lim_{x \to -3^+} f(x)$
c. $\lim_{x \to -3} f(x)$
d. $\lim_{x \to 5^-} f(x)$
e. $\lim_{x \to 5^+} f(x)$
f. $\lim_{x \to 5} f(x)$
g. $\lim_{x \to 10^-} f(x)$
h. $\lim_{x \to 10^+} f(x)$
i. $\lim_{x \to 10} f(x)$

Answer available at the end of this section.

Exercise 1.4.2

Use the graph of the function $g(x)$ below to determine each limit.

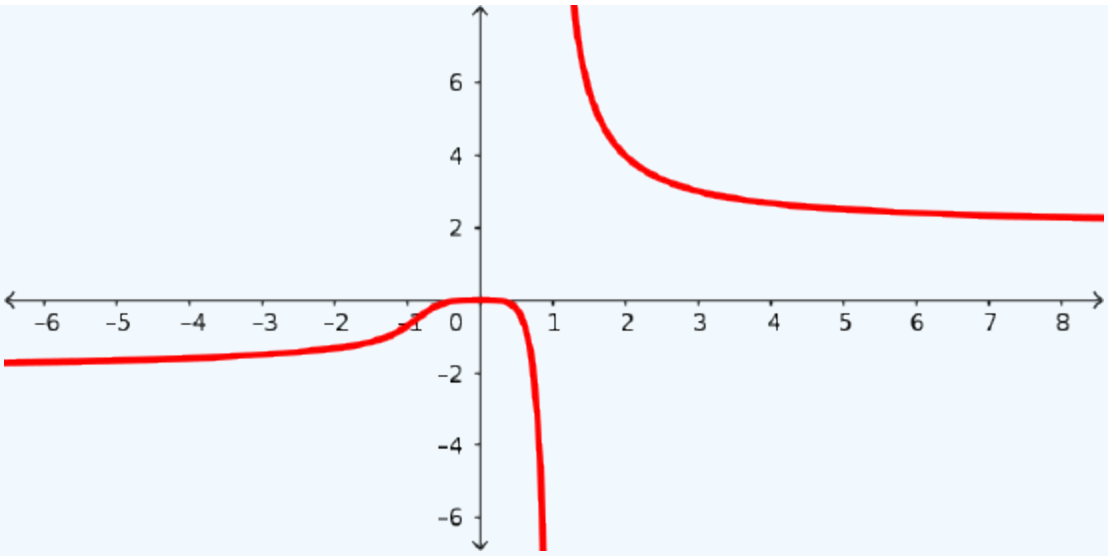


Figure 1.4.10 The function $g(x)$.

a. $\lim_{x\to-\infty} g(x)$
b. $\lim_{x\to\infty} g(x)$

Answer available at the end of this section.

Exercise 1.4.3

Use the graph of the function $h(x)$ below to determine each limit.

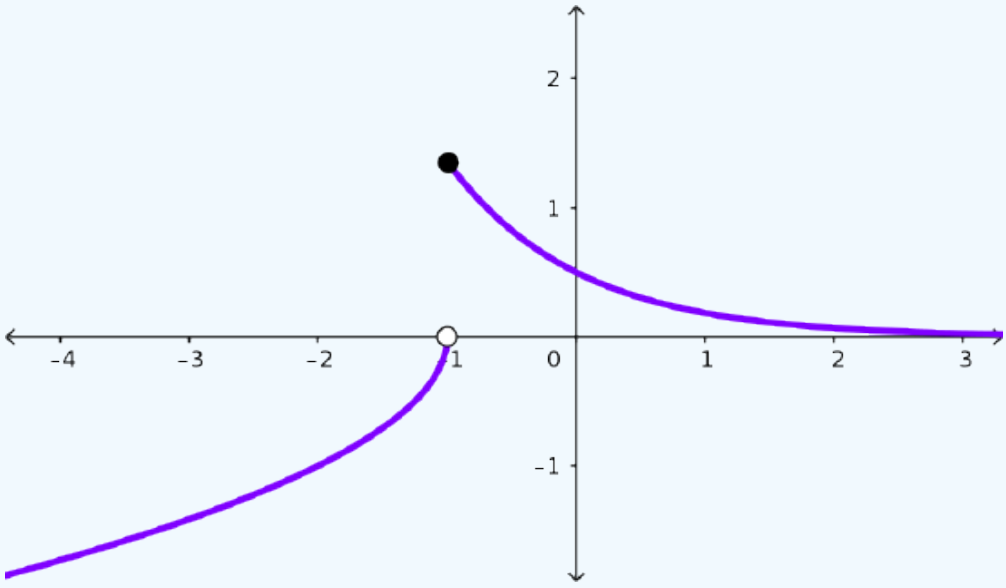


Figure 1.4.11 The function $h(x)$.

a. $\lim_{x\to-\infty} h(x)$
b. $\lim_{x\to\infty} h(x)$

Answer available at the end of this section.

Exercise 1.4.4

For each limit, fill in the table in a way that determines the limit, and report your answer.

a. $\lim_{x\to 4^-}\left(\dfrac{3x}{x-4}\right)$

x	$-$	$-$	$-$	4
$\dfrac{3x}{x-4}$	$-$	$-$	$-$	does not exist

b. $\lim_{x\to 4^+}\left(\dfrac{3x}{x-4}\right)$

x	4	$-$	$-$	$-$
$\dfrac{3x}{x^{-4}}$	does not exist	$-$	$-$	$-$

c. $\lim_{x\to 1}\left(\dfrac{x-3}{(2x-4)(x-1)^2}\right)$

x	$-$	$-$	$-$	1	$-$	$-$	$-$
$\dfrac{x-3}{(2x-4)(x-1)^2}$	$-$	$-$	$-$	does not exist	$-$	$-$	$-$

Answer available at the end of this section.

Exercise 1.4.5

Evaluate the following limits analytically. Also evaluate the relevant onesided limits when needed.

a. $\lim_{x\to-\infty}\left(2+4x-5x^2+\dfrac{x^3}{7}-\dfrac{8x^4}{9}\right)$

b. $\lim_{x\to\infty}\left(\dfrac{5x^3-4x+3}{2x+6}\right)$

c. $\lim_{x\to-3^-}\left(\dfrac{5x^3-4x+3}{2x+6}\right)$

d. $\lim_{x\to-\infty}\left(\dfrac{x-10}{3x+4}\right)$

e. $\lim_{x\to 2}\left(\dfrac{4x-1}{x^4-16}\right)$

f. $\lim_{x\to\infty}\left(\dfrac{6-5x^5}{2+3x^2}\right)$

g. $\lim_{x\to 1}\left(\dfrac{x^2+1}{x^3-2x^2+x}\right)$

h. $\lim_{\theta\to 3\pi/2}\sin(\theta)\tan(\theta)$

Answer available at the end of this section.

Exercise 1.4.6

For each of the following functions, find the locations of any vertical asymptotes and report the behavior of the functions around those asymptotes using limit statements.

a. $f(x)=\dfrac{|x-4|}{x+2}$

b. $g(x)=\dfrac{\sin(x)}{x^2-1}$

c. $h(t)=\dfrac{4}{t}+\dfrac{9}{t-1}$

d. $s(w) = \dfrac{w^3 - 3w^2 + w - 1}{w^2 - 5w + 6}$

e. $f(x) = \dfrac{e^x}{(x-3)^4}$

Answer available at the end of this section.

Exercise 1.4.7

For each of the following functions, find the locations of any horizontal asymptotes using limit statements (as $x \to \infty$ and $x \to -\infty$).

a. $f(x) = 2 + \dfrac{10}{x^2}$

b. $g(x) = \dfrac{4x^5 - 3x^3 + 5x - 1}{10x^2 + 5x^3 - 5x^4}$

c. $h(x) = \dfrac{3x^2 + 4x + \sqrt{x}}{1 - \sqrt{x} - 2x^2}$

d. $j(x) = \dfrac{3}{5 + 10^{1/x}}$

Answer available at the end of this section.

Answers

Theorem 1.4.7**: Polynomial End Behavior Limits.**

Consider the polynomial function:

$$p(x) = a_n x^n + a_{n-1}x^{n-1} + \ldots + a_2 x^2 + a_1 x + a_0$$

where n is some integer and ak is a real number for $k = 0, 1, 2, \ldots, n$. For simplicity, we will consider only the limit as $x \to \infty$, but we could easily repeat this exact proof for the case where $x \to -\infty$.

Before we consider this limit, we can factor out x n, the variable with the highest exponent:

$$\begin{aligned} p(x) &= a_n x^n + a_{n-1}x^{n-1} + \ldots + a_2 x^2 + a_1 x + a_0 \\ &= x^n \left(\frac{a^n x^n}{x^n} + \frac{a_{n-1}x^{n-1}}{x^n} + \ldots + \frac{a_2 x^2}{x^n} + \frac{a_1 x}{x^n} + \frac{a_0}{x^n} \right) \\ &= x^n \left(a_n + \frac{a_{n-1}}{x} + \ldots + \frac{a_2}{x^{n-2}} + \frac{a_1}{x^{n-1}} + \frac{a_0}{x^n} \right) \end{aligned}$$

Now consider the limit of this product:

$$\begin{aligned} \lim_{x\to\infty} p(x) &= \lim_{x\to\infty} x^n \left(a_n + \frac{a_{n-1}}{x} + \ldots + \frac{a_2}{x^{n-2}} + \frac{a_1}{x^{n-1}} + \frac{a_0}{x^n} \right) \\ &= \left(\lim_{x\to\infty} x^n \right) \left(\lim_{x\to\infty} a_n + \frac{a_{n-1}}{x} + \ldots + \frac{a_2}{x^{n-2}} + \frac{a_1}{x^{n-1}} + \frac{a_0}{x^n} \right) \end{aligned}$$

We can see that in the second limit, we have a single constant term, a_n, followed by reciprocal power functions. Then, due to Theorem 1.4.6, we know that the second limit will by a_n, since the reciprocal power functions will all approach 0.

$$\begin{aligned} \lim_{x\to\infty} p(x) &= \left(\lim_{x\to\infty} x^n \right) \left(\lim_{x\to\infty} a_n + \frac{a_{n-1}}{x} + \ldots + \frac{a_2}{x^{n-2}} + \frac{a_1}{x^{n-1}} + \frac{a_0}{x^n} \right) \\ &= \left(\lim_{x\to\infty} x^n \right) (a_n + 0 + \ldots + 0 + 0 + 0) \\ &= \left(\lim_{x\to\infty} x^n \right) (a_n) \\ &= \lim_{x\to\infty} a_n x^n \end{aligned}$$

And so $\lim_{x\to\infty} p(x) = \lim_{x\to\infty} a_n x^n$ as we claimed.

Exercise 1.4.1

a. $\lim_{x\to -3^-} f(x) = \infty$
b. $\lim_{x\to -3^+} f(x) = \infty$
c. $\lim_{x\to -3} f(x) = \infty$
d. $\lim_{x\to 5^-} f(x) = \infty$
e. $\lim_{x\to 5^+} f(x) = -\infty$
f. $\lim_{x\to 5} f(x)$ doesn't exist
g. $\lim_{x\to 10^-} f(x) = -\infty$
h. $\lim_{x\to 10^+} f(x) = \infty$
i. $\lim_{x\to 10} f(x)$ doesn't exist

Exercise 1.4.2

a. $\lim_{x\to -\infty} g(x) = -2$

b. $\lim_{x\to\infty} g(x) = 2$

Exercise 1.4.3

a. $\lim_{x\to-\infty} h(x) = -\infty$
b. $\lim_{x\to\infty} h(x) = 0$

Exercise 1.4.4

a. $\lim_{x\to4^-}\left(\dfrac{3x}{x-4}\right) = -\infty$
b. $\lim_{x\to4^+}\left(\dfrac{3x}{x-4}\right) = \infty$
c. $\lim_{x\to1}\left(\dfrac{x-3}{(2x-4)(x-1)^2}\right) = \infty$

Exercise 1.4.5

a. $\lim_{x\to-\infty}\left(2+4x-5x^2+\dfrac{x^3}{7}-\dfrac{8x^4}{9}\right) = -\infty$
b. $\lim_{x\to\infty}\left(\dfrac{5x^3-4x+3}{2x+6}\right) = \infty$
c. $\lim_{x\to-3^-}\left(\dfrac{5x^3-4x+3}{2x+6}\right) = \infty$
d. $\lim_{x\to-\infty}\left(\dfrac{x-10}{3x+4}\right) = \dfrac{1}{3}$
e. $\lim\limits_{x\to2^-}\left(\dfrac{4x-1}{x^4-16}\right) = -\infty$
$\lim\limits_{x\to2^+}\left(\dfrac{4x-1}{x^4-16}\right) = \infty$
f. $\lim_{x\to\infty}\left(\dfrac{6-5x^5}{2+3x^2}\right) = -\infty$
g. $\lim_{x\to1}\left(\dfrac{x^2+1}{x^3-2x^2+x}\right) = \infty$
h. $\lim\limits_{\theta\to3\pi/2^-}\sin(\theta)\tan(\theta) = -\infty$
$\lim\limits_{\theta\to3\pi/2^+}\sin(\theta)\tan(\theta) = \infty$

Exercise 1.4.6

a. $\lim\limits_{x\to-2^-} f(x) = -\infty$
$\lim\limits_{x\to-2^+} f(x) = \infty$
$\lim\limits_{x\to-1^-} g(x) = -\infty$
b. $\lim\limits_{x\to-1^+} g(x) = \infty$
$\lim\limits_{x\to1^-} g(x) = -\infty$
$\lim\limits_{x\to1^+} g(x) = \infty$
c. **Hint**: If you'd rather, you can add the terms together to get:

$$h(t) = \frac{13t-4}{t^2(t-1)}$$

$\lim\limits_{t\to0^-} \mathrm{h(t)} = \infty$
$\lim\limits_{t\to0^+} \mathrm{h(t)} = \infty$
$\lim\limits_{t\to1^-} \mathrm{h(t)} = -\infty$
$\lim\limits_{t\to1^+} \mathrm{h(t)} = \infty$

d. $\mathrm{s(w)} = \dfrac{\mathrm{w}^3-3\mathrm{w}^2+\mathrm{w}-1}{(\mathrm{w}-2)(\mathrm{w}-3)}$
$\lim\limits_{\mathrm{w}\to2^-} \mathrm{s(w)} = -\infty$
$\lim\limits_{\mathrm{w}\to2^+} \mathrm{s(w)} = \infty$
$\lim\limits_{\mathrm{w}\to3^-} \mathrm{s(w)} = -\infty$
$\lim\limits_{\mathrm{w}\to3^+} \mathrm{s(w)} = \infty$
e. $\lim\limits_{x\to3^-} f(x) = \infty$
$\lim\limits_{x\to3^+} f(x) = \infty$

Exercise 1.4.7

a. $\lim\limits_{x\to-\infty} f(x) = 0$
$\lim\limits_{x\to\infty} f(x) = 0$
b. $\lim\limits_{x\to-\infty} g(x) = \infty$
$\lim\limits_{x\to\infty} g(x) = -\infty$
c. $\lim\limits_{x\to-\infty} h(x) = -\dfrac{3}{2}$
$\lim\limits_{x\to\infty} h(x) = -\dfrac{3}{2}$
d. $\lim_{x\to-\infty} g(x) = \dfrac{1}{2}$
$\lim_{x\to\infty} g(x) = \dfrac{1}{2}$

This page titled 1.4: Limits Involving Infinity is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Peter Keep.

1.5: The Squeeze Theorem

We won't get enough time to spend thinking about all of the possible techniques that we could use to evaluate limits, but in this section we'll investigate one more.

Here, we'll introduce a new limit involving a type of function that we've not used in limits so far: trigonometric functions.

Weird Functions, Weird Behavior

Ok, trigonometric functions aren't actually weird. But we want to look at a function that is slightly more complicated than the ones we've looked at so far.

Activity 1.5.1: A Weird End Behavior Limit.

In this activity, we're going to find the following limit:

$$\lim_{x\to\infty}\left(\frac{\sin^2(x)}{x^2+1}\right)$$

This limit is a bit weird, in that we really haven't looked at trigonometric functions that much. We're going to start by looking at a different limit in the hopes that we can eventually build towards this one.

a. Consider, instead, the following limit:

$$\lim_{x\to\infty}\left(\frac{1}{x^2+1}\right)$$

Find the limit and connect the process or intuition behind it to at least one of the results from this text.
- **Hint 1**: Start with Theorem 1.2.5 to think about the numerator and denominator separately.
- **Hint 2**: Can you use Theorem 1.4.7 in the denominator?
- **Hint 3**: Is Theorem 1.4.6 relevant?

b. Let's put this limit aside and briefly talk about the sine function. What are some things to remember about this function? What should we know? How does it behave?

c. What kinds of values doe we expect $\sin(x)$ to take on for different values of x?

$$\_\_\_ \leq \sin(x) \leq \_\_\_$$

d. What happens when we square the sine function? What kinds of values can that take on?

$$\_\_\_ \leq \sin^2(x) \leq \_\_\_$$

e. Think back to our original goal: we wanted to know the end behavior of $\dfrac{\sin^2(x)}{x^2+1}$. Right now we have two bits of information:
- We know $\lim_{x\to\infty}\left(\dfrac{1}{x^2+1}\right)$.
- We know some information about the behavior of $\sin^2(x)$. Specifically, we have some bounds on its values.

Can we combine this information?

In your inequality above, multiply $\left(\dfrac{1}{x^2+1}\right)$ onto all three pieces of the inequality. Make sure you're convinced about the direction or order of the inequality and whether or not it changes with this multiplication.

$$\underbrace{\frac{}{x^2+1}}_{\text{call this } f(x)} \leq \frac{\sin^2(x)}{x^2+1} \leq \underbrace{\frac{}{x^2+1}}_{\text{call this } h(x)}$$

f. For your functions $f(x)$ and $h(x)$, evaluate $\lim_{x\to\infty} f(x)$ and $\lim_{x\to\infty} h(x)$.

g. What do you think this means about the limit we're interested in, $\lim_{x\to\infty}\left(\frac{\sin^2(x)}{x^2+1}\right)$?

Squeeze Theorem

This strategy is a really nice one to use when we know the behavior of some wellbehaved "bounding" functions. We can try to off-load the task of summarizing the behavior of a strangely behaved function to these bounding functions, and follow them! As long as they approach each other, than the strangely behaved function has to have the same behavior.

Let's formalize this result carefully.

Theorem 1.5.1: The Squeeze Theorem.

For some functions $f(x), g(x)$, and $h(x)$ which are all defined and ordered $f(x) \le g(x) \le h(x)$ for x-values near $x = a$ (but not necessarily at $x = a$ itself), and for some real number L, if we know that

$$\lim_{x\to a} f(x) = L = \lim_{x\to a} h(x)$$

then we also know that $\lim_{x\to a} g(x) = L$.

Aside.

Sometimes this theorem is called the Sandwich Theorem, since the upper bounding function and the lower bounding function act as the slices of bread, while the strangely behaved middle function acts like the toppings on the sandwich.

We choose not to use that naming convention in this text in order to preserve the flexibility of the definition of a sandwich. An open-faced sandwich is still a sandwich (in the opinion of the textbook author), but this result clearly doesn't hold when we only have one bounding function.

This theorem can be difficult to use, primarily because building the bounds for a function is difficult. In Activity 1.5.1, we were able to build the boundary functions by simply thinking about the bounds on the $\sin(x)$. This worked well, but we were only able to do this because of our familiarity with this function. With other functions, these bounds are harder to just come up with. This is especially true in that we need the bounds to accomplish multiple things at once:

- We need them to be ordered correctly with regard to the function we care about: one above it and one below it.
- We need the limits of these functions to be things we can actually evaluate! This is the whole point: we use these (hopefully easier) limits to evaluate the (probably hard) limit that we're interested in.
- We need the limits of these functions to be the same as $x \to a$, otherwise we're not certain about where our function actually is.

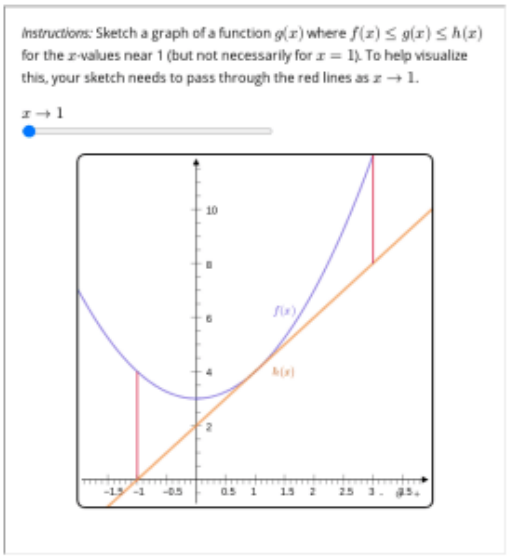
In practice, we'll try to build any bounding functions with some assistance, or start with bounding functions already stated.

Let's see another way of thinking about this result using our graphical intuition.

Activity 1.5.2: Sketch This Function Around This Point.

a. Sketch or visualize the functions $f(x) = x^2 + 3$ and $h(x) = 2x + 2$, especially around $x = 1$.

b. Now we want to add in a sketch of some function $g(x)$, all the while satisfying the requirements of the Squeeze Theorem.



Standalone
Embed

c. Use the Squeeze Theorem to evaluate and explain $\lim_{x\to 1} g(x)$ for your function $g(x)$.
d. Is this limit dependent on the specific version of $g(x)$ that you sketched? Would this limit be different for someone else's choice of $g(x)$ given the same parameters?
e. What information must be true (if anything) about $\lim_{x\to 3} g(x)$ and $\lim_{x\to 0} g(x)$? Do we know that these limits exist? If they do, do we have information about their values?

Practice Problems

Exercise 1.5.1

For each of the following statements, discuss the possible existence of the limit in question. Explain in detail how we might know whether the limit exists or not, or why it is impossible to tell.

a. We want to know about $\lim_{x\to 4} f(x)$, and we know that $a(x) \le f(x) \le b(x)$ for all x-values, with $a(4) = 3$ and $b(4) = 3$.
b. We want to know about $\lim_{x\to 0} f(x)$, and we know that $\lim_{x\to 0} g(x) = 5$ for the function $g(x) \ge f(x)$ for x-values around 0.
c. We want to know about $\lim_{x\to -1} f(x)$, and we know that $a(x) \le f(x) \le b(x)$ for x-values around –1, with $\lim_{x\to -1} a(x) = 9$ and $\lim_{x\to -1} b(x) = 9$.

Exercise 1.5.2

We want to use the Squeeze Theorem to evaluate $\lim_{x\to 0} (x^2 \sin(\pi/x))$.

a. Explain why we know that $-1 \le \sin(\pi/x) \le 1$ for any non-zero x-values. Why does this inequality not hold for $x = 0$?
b. Use the inequality from (a) to build an inequality of functions $f(x) \le x^2 \sin(\pi/x) \le g(x)$ for non-zero x-values.
c. With these functions $f(x)$ and $g(x)$ from (b), find $\lim_{x\to 0} f(x)$ and $\lim_{x\to 0} g(x)$.
d. Explain how the Squeeze Theorem tells us what the value of $\lim_{x\to 0} x^2 \sin(\pi/x)$ is.

This page titled 1.5: The Squeeze Theorem is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Peter Keep.

1.6: Continuity and the Intermediate Value Theorem

What does it mean for a function to be continuous?

This might be a familiar concept to you, and we all probably have some idea of what the word “continuous” means (or should mean), both in a colloquial context and in a math-specific context. Instead of focusing on some of the traditional ideas like “can you draw the graph without picking up your pencil”, we’re going to formalize the idea of continuity a bit.

And the first thing that we’ll have to do is shift our intuition from what continuity is describing. We won’t (at least at first) use continuity as a classification of functions, as a whole. Instead, we’re going to think of continuity as a description of some “local” behavior.

Continuity as Connectedness

We’re going to try to think about what we want to describe when we define continuity. The description of “drawing a graph without picking up a pencil” is a global property, and also a pretty vague and ambiguous one.

Activity 1.6.1: Classification and Continuity.

Let’s consider the following functions, graphed below.

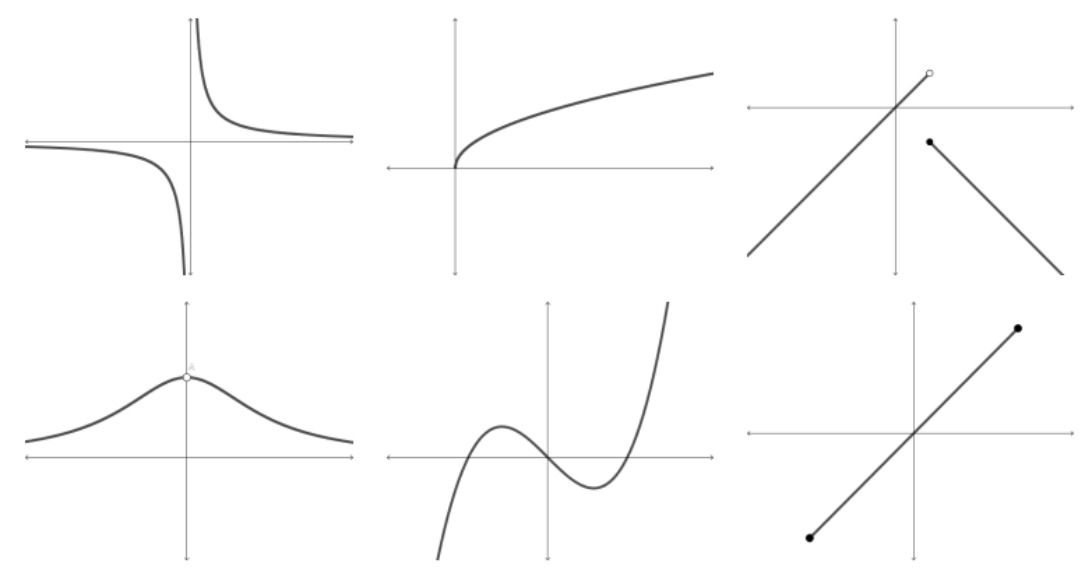


Figure 1.6.1 A variety of graphs for us to use to think about continuity.

a. Can you point out any points on the functions above that seem like the functions might not be continuous? Note that we’re not classifying each function as “continuous” or “not continuous”!
b. We want to build towards a definition of continuity using “connectedness” as the key: a function is continuous at a point if it is connected to itself. How does that work when we think about function values and limits? Use the graphs above and the points you looked at to help!

There are definitely some unknowns here. One of the questions that students ask a lot during an investigation like this is what we do about end-points. Is a function continuous at a closed ending-point in its domain? Is it connected to itself? To answer questions like this, we should try to write out what we mean by connectedness.

A pretty basic version of this is something like “A function is connected to itself at a point if, for the inputs near that point, the

function values are close to the function value at the point." Maybe even more concise is "Small deviations in the inputs produce small changes in the outputs."

Defining Continuity

We're going to take this idea of a function being connected to itself as "small deviations in the inputs produce small changes in the outputs" and write it down using limit notation.

Definition 1.6.2: Continuous at a Point.

The function $f(x)$ is continuous at an x-value in the domain of $f(x)$ if $x = a$ if $\lim_{x\to a} f(x) = f(a)$.

If $f(x)$ is not continuous at $x = a$, but one of the one-sided limits is equal to the function output, then we can define directional continuity at that point:

- We say $f(x)$ is continuous on the left at $x = a$ when $\lim_{x\to a^-} f(x) = f(a)$.
- We say $f(x)$ is continuous on the right at $x = a$ when $\lim_{x\to a^+} f(x) = f(a)$.

In using limits, we get some flexibility: we can talk about one-sided limits, specifically! This helps us take care of the pesky "Is a function connected to itself at a closed ending point," question! We can say, using this definition, "Yes, on the left side!" or "Yes, on the right side!"

Example 1.6.3

For the function $f(x)$ defined below, decide whether or not the function is continuous at the point listed. If it is not continuous, report whether it is continuous on one side.

$$f(x) = \begin{cases} 2x+1 & \text{when } x \le 2 \\ \dfrac{2x^2+7}{6^3} & \text{when } 2 < x < 4 \\ \dfrac{6^3-1}{x-3} & \text{when } x \ge 4 \end{cases}$$

a. Is $f(x)$ continuous at the point (2, 5)? **Hint**. Check the one-sided limits as $x \to 2$. Do they both equal 5? Does one of them?
b. Is $f(x)$ continuous at the point (4, 5)? **Hint**. Check the one-sided limits as $x \to 4$. Do they both equal 5? Does one of them?

Solution

a.

$$\begin{aligned} \lim_{x\to 2^-} f(x) &= \lim_{x\to 2^-} 2x+1 \\ &= 2(2)+1 \\ &= 5 \end{aligned}$$

$$\begin{aligned} \lim_{x\to 2^+} f(x) \lim_{x\to 2^+} \frac{2x^2+7}{3} \\ &= \frac{2(2)^2+7}{3} \\ &= 5 \end{aligned}$$

So $\lim_{x\to 2} f(x) = 5$. Since $f(2) = 5$ as well, we have $\lim_{x\to 2} f(x) = f(2)$ which means that $f(x)$ is continuous at $x = 2$.

b.

$$\begin{aligned}\lim_{x\to 4^-} f(x) &= \lim_{x\to 4^-} \frac{2x^2+7}{3}\\ &= \frac{2(4)^2+7}{3}\\ &= 13\end{aligned}$$

$$\begin{aligned}\lim_{x\to 4^+} f(x) &= \lim_{x\to 4^+} \frac{6}{x-1}+3\\ &= \frac{6}{4-1}+3\\ &= 5\end{aligned}$$

Since $\lim_{x\to 4^-} f(x) \neq \lim_{x\to 4^+} f(x)$ we know that $\lim_{x\to 4} f(x)$ doesn't exist, and so this function cannot be continuous at $x = 4$. But we can notice that $\lim_{x\to 4^+} f(x) = f(4)$, and so we can say that the function is still continuous on the right side at $x = 4$.

Another bit of flexibility that we have with this definition is that, once we know whether a function is continuous at a point or not, we can string together points in an interval and classify a function as continuous on an interval.

Definition 1.6.4: Continuous on an Interval.

We say that $f(x)$ is continuous on the interval (a, b) if $f(x)$ is continuous at every x-value with $a < x < b$.

If $f(x)$ is continuous on the right at $x = a$ and/or continuous on the left at $x = b$, then we will say that $f(x)$ is continuous on the interval $[a, b)$, $(a, b]$, or $[a, b]$, whichever is relevant.

Something to notice, here, is that we won't do this by checking each individual x-value in an interval. If we wanted to say that a function was continuous on (0, 1), we would then have to consider every real number between 0 and 1. But there are an infinite amount of these!

So when we talk about a function being continuous on an interval, we'll really think about this by pointing out the places where a function is not continuous (x-values where $f(x)$ violates some part of Definition 1.6.2).

Discontinuities

Since we'll want to label functions at places that they are not continuous (in order to write the intervals where it is continuous), we should summarize some things about discontinuities.

Where is a Function not Continuous?

Most of the functions that we consider in this text will be continuous everywhere that it makes sense: on their domain. That is, if there is a point defined at some x-value, it is likely that the function's limit matches the y-value of the point. More specifically, though:

- A function is discontinuous at any location that results in an infinite limit. These are locations where $f(x)$ is undefined and the limit is infinite (and so doesn't exist).
- A function is, in general, discontinuous at any x-value where $f(x)$ is not defined. This seems silly to say! We probably could have left this unsaid. Of course f isn't continuous at an x-value not in its domain.
- A function that is defined as a piecewise function could be discontinuous at locations where the pieces meet: maybe the limit doesn't exist, or maybe the function value is not defined, or maybe the limit exists and the function value is defined but they do not match.

Some quick notes:

- Every polynomial function is continuous everywhere on its domain. We knew this already, from Theorem 1.2.7 Limits of Polynomials: the limit is always the same as the function value at any input!

- Rational functions are only discontinuous when there is division by 0.
- Even-root functions (like square roots) are continuous whenever the radicand (the stuff under the root) is positive. We might need to investigate wherever the radicand is equal to 0, since the function might only be continuous on one side.
- The following functions are continuous on their domains:
  - $y = \sin(x)$ on $(-\infty, \infty)$
  - $y = \cos(x)$ on $(-\infty, \infty)$
  - $y = e^x$ on $(-\infty, \infty)$
  - $y = \ln(x)$ on $(0, \infty)$

Example 1.6.5

Let's revisit the same function we looked at earlier:

$$f(x) = \begin{cases} 2x+1 & \text{when } x \leq 2 \\ \dfrac{2x^2+7}{6^3} & \text{when } 2 < x < 4 \\ \dfrac{6^3-1}{x-3} & \text{when } x \geq 4 \end{cases}$$

We already looked at what was happening at $x = 2$ and $x = 4$. Are there any other x-values where this function might not be continuous? Why or why not? Can you report the intervals on which $f(x)$ is continuous?

Hint. Can you name the function type for $x < 2, 2 < x < 4$ and $x > 4$?

Solution

Since $f(x)$ is polynomial for $x < 2$ and $2 < x < 4$, we know that the function is continuous on those intervals as well. We can see that for $x > 4$, $f(x)$ is a rational function. Since the only x-value where we might divide by 0 is outside of the interval for which this function part is defined, we know that this function is continuous for $x > 4$. Since our function is continuous at $x = 2$, we can say that it is continuous on the interval $(-\infty, 4)$. We also say that the function was continuous on the right at $x = 4$, and so we can say that $f(x)$ is also continuous on the interval $[4, \infty)$.

We'll end this section with one last result that will be pretty important. Continuity, as a property, will be useful as we move forward, but only for a few specific reasons. One is the following result.

Theorem 1.6.6: Continuity with Composition.

For two functions $f(x)$ and $g(x)$, if $\lim_{x \to a} g(x) = L$ and $f(x)$ is continuous at $x = L$, then

$$\lim_{x \to a} f(g(x)) = f\left(\lim_{x \to a} g(x)\right) = f(L)$$

This is a generalization of the properties of limits that we explored earlier (Theorem 1.2.5 Combinations of Limits). We can use this new theorem to evaluate limits by just noticing if pieces of our function might behave nicely with limits.

Let's evaluate the following limit:

$$\lim_{x \to 1} \left(\sin\left(\sqrt{e^{x^2-2x+1}}\right)\right)$$

Notice that the sine function is continuous everywhere. This means that we can bring the limit statement "inside" by a layer:

$$\lim_{x \to 1} \left(\sin\left(\sqrt{e^{x^2-2x+1}}\right)\right) = \sin\left(\lim_{x \to 1} \sqrt{e^{x^2-2x+1}}\right)$$

Now we can think about the square root function: we know that the radicand function (the exponential) always produces positive outputs. So we can bring the limit statement "inside" by another layer of composition:

$$\sin\left(\lim_{x\to 1}\sqrt{e^{x^2-2x+1}}\right) = \sin\left(\sqrt{\lim_{x\to 1}e^{x^2-2x+1}}\right)$$

The exponential function is also continuous everywhere:

$$\sin\left(\sqrt{\lim_{x\to 1}e^{x^2-2x+1}}\right) = \sin\left(\sqrt{e^{\lim_{x\to 1}x^2-2x+1}}\right)$$

And we also know that polynomial functions are continuous everywhere!

$$\begin{aligned}\sin\left(\sqrt{e^{\lim_{x\to 1}x^2-2x+1}}\right) &= \sin\left(\sqrt{e^{(1)^2-2(2)+1}}\right)\\ &= \sin\left(\sqrt{e^0}\right)\\ &= \sin(1)\end{aligned}$$

This seemingly difficult limit is actually as easy as just evaluating the function at the input, since the function behaves nicely.

Intermediate Value Theorem

Let's introduce one of the biggest results coming from continuity of functions that exists. We're going to simply state the result and then discuss it, instead of building up towards it.

Theorem 1.6.7: Intermediate Value Theorem.

If $f(x)$ is a function that is continuous on $[a,b]$ with $f(a)\neq f(b)$ and L is any real number between $f(a)$ and $f(b)$ (either $f(a)<L<f(b)$ or $f(b)<L<f(a)$), then there exists some c between a and $b(a<c<b)$ such that $f(c)=L$.

Aside.

This theorem was stated as early as the 5th century BCE by Bryson of Heraclea. Back then, a really interesting problem was related to "squaring the circle." That is, given a circle with some measurable radius, can we construct a square with equal area? This is obviously true, in that we can just use a square with the side length $r\sqrt{\pi}$. What we typically mean by "construct," though, is to create this square using only a compass and straightedge (a ruler without length markings) and only a finite number of steps. This was finally proven to be impossible in 1882, approximately 2300 years later.

Bryson of Heraclea knew that the square itself existed (even if he couldn't construct it) because he was able to find a circle with area less than the square (by inscribing a circle inside of the square) and a circle with area greater than the square (where the square is inscribed in the circle). Since he posited that he could increase the size of the circle in a continuous manner (without using those words), he claimed that a square with area equal to that of the circle must exist, since the area of the circle passes through all values from the smaller area to the larger area.

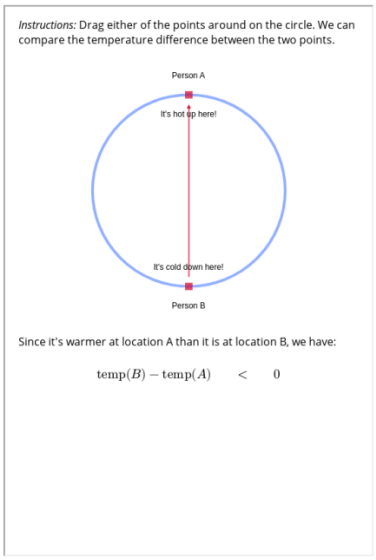
This is a seemingly simple result (even if it is stated in a seemingly complicated way): of course, in order to go from a point at one y-value to another point at a different y-value, the function has to pass through all of the y-values in between. This is pretty obvious, once we see through the phrasing.

We can use this result in some pretty interesting ways, though! Here's a favorite:

Let's consider two people, standing on polar opposite points on the earth. Let's say that these two people are carrying thermometers and are measuring the ambient temperature at their location. They coordinate their movement, and begin to move around the earth in a circle, remaining at polar opposite locations as they move.

How likely is it that the pair finds some location on their path where their temperature readings are exactly identical?

Let's pretend that the locations that they start in are different temperatures (or else yay, we found the spot where the temperatures are the same!), and that the starting position of person A is hotter than the starting position of person B.



Standalone
Embed

Can you see why we are guaranteed that there must be some pair of polar opposite locations where the ambient temperatures are identical?

If the ambient temperature differs continuously, then the Intermediate Value Theorem says that between the locations where the temperature differential is negative and where it is positive, the temperature differential function must pass through zero!

This argument can be used in a variety of contexts, and can be used to show that you can typically fix a wobbly table (where the legs don't all rest nicely on the ground) or chair by rotating it until all of the legs rest nicely on the ground.

The theorem itself might not seem that interesting, but it's important both historically and also as a means of making arguments about function values, especially when we claim that a function has to be 0 (because we show that there is a point where the function is positive, a point where it is negative, and the function is continuous in between the points).

Practice Problems

Exercise 1.6.1

Explain how the definition of continuity (Definition 1.6.2) describes a function being "connected to itself" at the point.

Exercise 1.6.2

What does it mean for a function to be continuous on the left at a point? What about on the right?

Exercise 1.6.3

Use the definition of continuity at a point to determine whether each function is continuous at the point listed. If it is not, is the function continuous on one side at the point? Explain.

a. $f(x) = \begin{cases} 4x+1 & \text{if } x < 1 \\ x^2+5 & \text{if } x \geq 1 \end{cases}$ Is $f(x)$ continuous at $x = 1$?

b. $g(x) = \begin{cases} 6x^2-2 & \text{if } x < 0 \\ 4 & \text{if } x = 0 \\ 1-3\cos(x) & \text{if } x > 0 \end{cases} Is\backslash(g(x)$ continuous at $x = 0$?

c. $h(x) = \begin{cases} \left(\dfrac{x^2-1}{x+1}\right) & \text{if } x \neq -1 \\ -2 & \text{if } x = -1 \end{cases}$ Is $h(x)$ continuous at $x = -1$?

Answer available at the end of this section.

Exercise 1.6.4

Find the values of a that make the function continuous at the point that is listed. If no such value of a exists, say so and explain why.

a. $f(x) = \begin{cases} x-6 & \text{if } x < 3 \\ ax^2+1 & \text{if } x \geq 3 \end{cases}$ at $x = 3$

b. $f(x) = \begin{cases} \dfrac{x^2-4}{x+2} & \text{if } x \neq -2 \\ a & \text{if } x = -2 \end{cases}$ at $x = -2$

Answer available at the end of this section.

Exercise 1.6.5

Use the properties of composition and continuity to evaluate the following limits.

a. $\lim_{x\to 0}\left(\sin\left(e^{\sqrt{x^2}}\right)\right)$

b. $\lim_{x\to 0} \ln\left(\dfrac{\cos\left(x^2\right)}{x^2+1}\right)$

c. $\lim_{x\to 0} \sqrt[3]{\dfrac{x}{\sqrt{16x+1}-1}}$

Answer available at the end of this section.

Exercise 1.6.6

"Even a broken clock is right twice a day," is a common phrase used to say that even an unreliable person can have a good point on accident. Explain how we can use the Intermediate Value Theorem to prove that a broken (analog) clock is guaranteed right twice per day. (Here, we're assuming that "broken" means stuck.)

Exercise 1.6.7

Explain why this argument doesn't work to show that an unreliable person will be guaranteed to have a good point on accident.

Exercise 1.6.8

Prove that at some point in your life you were exactly π feet tall.

Answers

Exercise 1.6.3

a.
- $\lim_{x\to 1} f(x)$ doesn't exist.
- $f(1) = 6$

- This function is not continuous at this point, since

$$\lim_{x\to 1} f(x) \neq f(1)$$

But, it is continuous on the right side, since

$$\lim_{x\to 1^+} f(x) = f(1)$$

b.
- $\lim_{x\to 0} g(x) = -2$
- $g(0) = 4$
- This function is not continuous at this point, since

$$\lim_{x\to -2} g(x) \neq g(-2)$$

c.
- $\lim_{x\to -1} h(x) = -2$
- $h(-1) = -2$
- This function is continuous at this point, since

$$\lim_{x\to -1} h(x) = h(-1)$$

Exercise 1.6.4

a. $a = -\frac{2}{9}$

b. $a = -4$

Exercise 1.6.5

a.
$$\begin{aligned}\lim_{x\to 0}\left(\sin\left(e^{\sqrt{x^2}}\right)\right) &= \sin\left(e^{\sqrt{\lim_{x\to 0} x^2}}\right)\\ &= \sin(1)\end{aligned}$$

b.
$$\begin{aligned}\lim_{x\to 0} \ln\left(\frac{\cos\left(x^2\right)}{x^2+1}\right) &= \ln\left(\frac{\cos\left(\lim_{x\to 0} x^2\right)}{\lim_{x\to 0}\left(x^2+1\right)}\right)\\ &= 0\end{aligned}$$

c.
$$\begin{aligned}\lim_{x\to 0}\sqrt[3]{\frac{x}{\sqrt{16x+1}-1}} &= \sqrt[3]{\lim_{x\to 0}\left(\frac{x}{\sqrt{16x+1}-1}\right)}\\ &= \frac{1}{2}\end{aligned}$$

This page titled 1.6: Continuity and the Intermediate Value Theorem is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Peter Keep.

1.7: Exploration- Limits

We're going to return to the definition of the Limit of a Function, and try to add some more precision to the way that we talk about these limits. We'll recreate the definition, without some of the ambiguous terminology, and then explore the new definition a bit, in order to get a good feel for what this means.

By the end of this exploration, we should have a pretty good understanding of what a limit is, it should match our intuition that we've built, and we should see a path that we could walk to investigate how these limit properties and results that we've talked about could be described.

Precise Definition of Limits

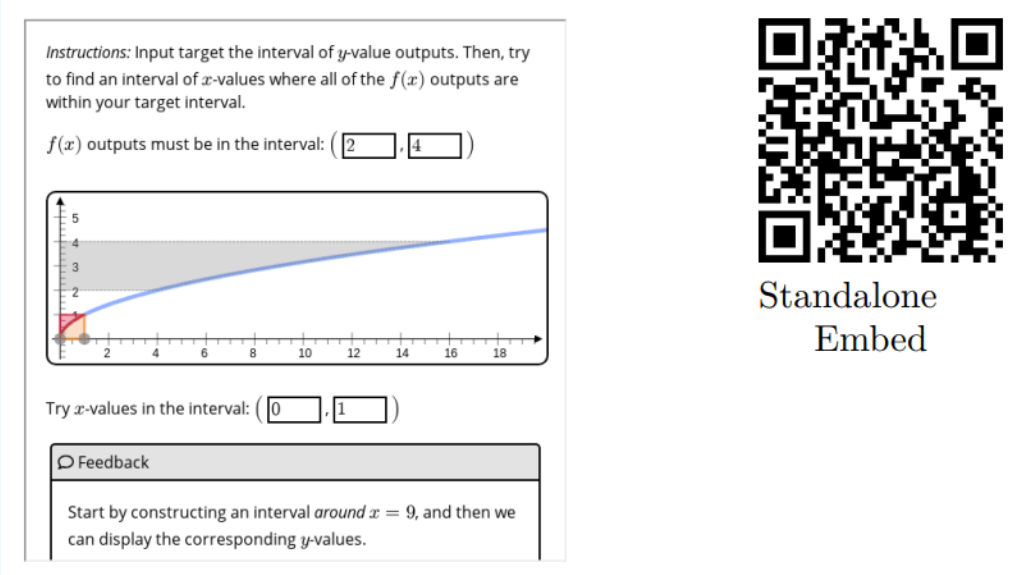
What are the areas in the definition of a limit where we need to add precision?

Exploration 1.7.1: A First Function to Think About.

Let's consider the function $f(x) = \sqrt{x}$. We're going to mostly be interested in the x-values around 9, since we'll be investigating $\lim_{x\to 9} \sqrt{x}$. We know that the y-value outputs of $f(x)$ are going to be around 3, giving us:

$$\lim_{x\to 9} \sqrt{x} = 3$$

Now we're going to perform some tasks that will hopefully help us add some precision to our definition of a limit. You can use the graph and inputs below while you work.



a. We want to come up with points on the graph of $f(x)$ whose y-values are within the interval (2, 4). Can you find an interval of x-values around $x = 9$ where the function outputs $f(x)$ are all in the interval (2, 4)?
b. Can you describe this in terms of distance? How far can x-values be from $x = 9$ where the corresponding $f(x)$ values are all in (2, 4)? Your answer should fill in the blank in the following sentence: "If I construct an interval of x-values that are within units of $x = 9$, then the corresponding function outputs will all be within (2, 4)." Note that another way of describing the interval, (2, 4) is to say that the function outputs are all within 1 unit of $y = 3$.
c. Repeat this process, but for the interval (2.9, 3.1). Can you find an interval of x-values around $x = 9$ where the function outputs $f(x)$ are all in the interval (2.9, 3.1)?

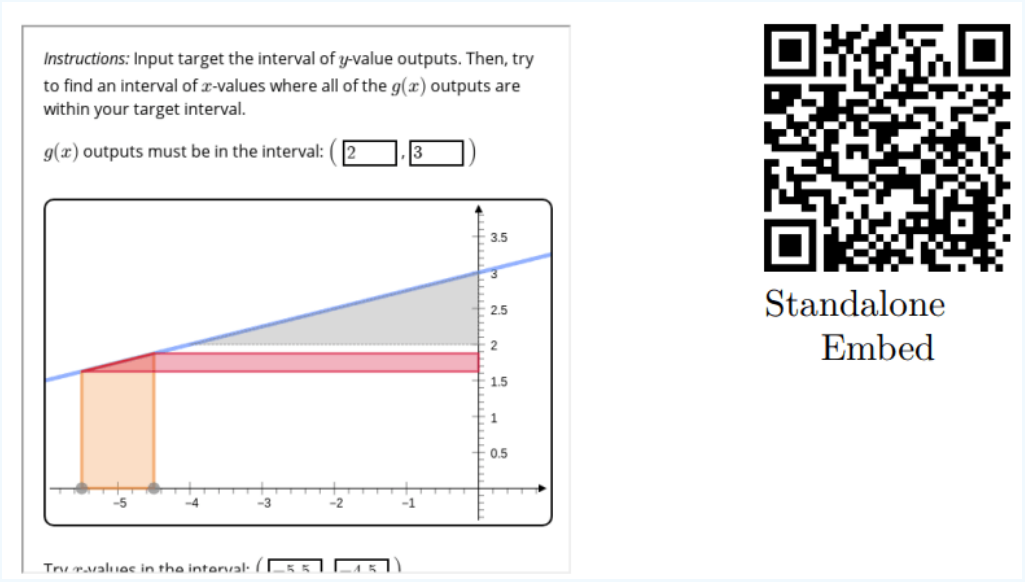
d. Can you describe this in terms of distance? How far can x-values be from $x = 9$ where the corresponding $f(x)$ values are all in (2.9, 3.1)? Your answer should fill in the blank in the following sentence: "If I construct an interval of x-values that are within units of $x = 9$, then the corresponding function outputs will all be within 0.1 units of $y = 3$."
e. Come up with an even smaller interval. Change the distance between the function outputs and $y = 3$ to something smaller than 0.1. Construct a new target interval of y-values, and see if you can find an interval of x-values such that all of the function outputs are in your target interval.
f. Have you convinced yourself that you can always do this? Do you think it really matters what the target interval of y -values is? Can you always find an interval of x-values around $x = 9$ (or x-values that are some non-zero distance from $x = 9$) where the outputs are in within that target distance from $y = 3$?
g. Let's formalize this: If ε is some real number with $\varepsilon > 0$, then can you find an interval of x-values around $x = 9$ where the function outputs $f(x)$ are all in the interval $(3 - \varepsilon, 3 + \varepsilon)$?
h. Can you describe this in terms of distance? How far can x-values be from $x = 9$ where the corresponding $f(x)$ values are all in $(3 - \varepsilon, 3 + \varepsilon)$? Your answer should fill in the blank in the following sentence: "If I construct an interval of x-values that are within _____ units of $x = 9$, then the corresponding function outputs will all be within ε units of $y = 3$."

Exploration 1.7.2: A Linear Function.

We'll try this again, but with a new function. This will change a bit of how we think about constructing our intervals. Consider the function $g(x) = \frac{x}{4} + 3$. We're going to mostly be interested in the x-values around -2, since we'll be investigating:

$$\lim_{x \to -2} \left(\frac{x}{4} + 3 \right) = \frac{5}{2}$$

You can, again, use the graph and inputs below while you work.



a. We want to come up with points on the graph of $g(x)$ whose y-values are within the interval (2, 3). Can you find an interval of x-values around $x = -2$ where the function outputs $g(x)$ are all in the interval (2, 3)?
b. Can you describe this in terms of distance? How far can x-values be from $x = -2$ where the corresponding $g(x)$ values are all in (2, 3)? Your answer should fill in the blank in the following sentence: "If I construct an interval of x-values that are within units of $x = -2$, then the corresponding function outputs will all be within (2, 3)." Note that another way of describing the interval, (2, 3) is to say that the function outputs are all within $\frac{1}{2}$ unit of $y = \frac{5}{2}$.
c. Repeat this process, but for the interval $\left(\frac{19}{8}, \frac{21}{8} \right)$. Can you find an interval of x-values around $x = -2$ where the function outputs $g(x)$ are all in the interval $\left(\frac{19}{8}, \frac{21}{8} \right)$?

d. Can you describe this in terms of distance? How far can x-values be from $x = -2$ where the corresponding $g(x)$ values are all in $\left(\frac{19}{8}, \frac{21}{8}\right)$? Your answer should fill in the blank in the following sentence: "If I construct an interval of x-values that are within units of $x = -2$, then the corresponding function outputs will all be within $\frac{1}{8}$ unit of $y = -2$."

e. If ε is any real number with $\varepsilon > 0$, then can you find an interval of x-values around $x = -2$ where the function outputs $g(x)$ are all in the interval $\left(\frac{5}{2} - \varepsilon, \frac{5}{2} + \varepsilon\right)$?

f. Can you describe this in terms of distance? How far can x-values be from $x = -2$ where the corresponding $f(x)$ values are all in $\left(\frac{5}{2} - \varepsilon, \frac{5}{2} + \varepsilon\right)$?

Our goal, now, is to summarize what we've done.

- We built a system to talk about y-values that are "arbitrarily close the single, real number, L." We constructed an interval, $(L - \varepsilon, L + \varepsilon)$ for any $\varepsilon > 0$. We also thought about this in terms of distance, by considering the y-values that were within ε units of $y = L$.
- We found a way to talk about x-values being "sufficiently close to, but not equal to, a." We constructed an interval of x-values around $x = a$, and then considered the function outputs from those x-values. We should clarify that we really are considering the function outputs from all of these x-values in the interval except $x = a$. We also talked about this in terms of distance! We want to construct x-values that are some non-zero distance away from $x = a$.
- We tried to come up with a way of describing this interval of x-values— whose function outputs (except $f(a)$) are all guaranteed to be in $(L - \varepsilon, L + \varepsilon)$—using a kind of function of ε. That way, we can know that we can repeat this process for any ε measure of "closeness." When we refer to that function of ε to give us the distance for the x-values, we often use the symbol, δ. Sometimes people call this δ_ε as a way of reminding ourselves that it depends on the value of ε.

This gives us two definitions that we can think of as a way of adding precision to the idea of a limit.

Definition 1.7.1: (Open Interval) Limit of a Function.

Suppose that $f(x)$ is a function that exists for all x-values in some open interval around and containing a, except possibly at $x = a$ itself. We say that

$$\lim_{x \to a} f(x) = L$$

if, for any open interval (p, q) of y-values containing the single, real number L, there is some corresponding open interval (c, d) around $x = a$ (that is, $c < a < d$) such that for all x-values in (c, d) with $x \neq a$, the function output $f(x)$ is in (p, q).

Definition 1.7.2: (Epsilon-Delta) Limit of a Function.

Suppose that $f(x)$ is a function that exists for all x-values in some open interval around and containing a, except possibly at $x = a$ itself. We say that

$$\lim_{x \to a} f(x) = L$$

if L is some single, real number, and for any real number $\varepsilon > 0$, there is a corresponding $\delta_\varepsilon > 0$ such that $|f(x) - L| < \varepsilon$ whenever $0 < |x - a| < \delta_\varepsilon$.

Both of these definitions are very important to the further study of calculus, but especially important it to use these definitions of limits to think about continuity.

We can say that a function is continuous at some point if every open interval in the range of the function around the y-value of the point comes from an open interval in the domain of the function around the x-value of the point. This is a very useful and important definition for the study of topology, and a nice way of generalizing the concept of continuity.

Practice Problems

Exercise 1.7.1

Consider the $\varepsilon - \delta$ definition of a limit. What does ε represent? Connect it with the standard definition of a limit that we've been using. Similarly, what does δ represent? Connect it with the standard definition of a limit we've been using.

Exercise 1.7.2

Draw a picture of the following situations:

a. A function $f(x)$ where if $0 < |x - 3| < 2$ then $|f(x) + 4| < 1$.
b. A function $g(x)$ where if $0 < |x + 1| < \frac{1}{2}$ then $|g(x) - 6| < 2$.

Exercise 1.7.3

For the following, find δ or ε. Include a sketch that illustrates the values you found with the function.

a. For $f(x) = 3x + 1$, if $0 < |x - 2| < \delta$, then $|f(x) - 7| < 1$.
b. For $g(x) = \sqrt{x + 3}$, if $0 < |x - 6| < 5$, then $|g(x) - 3| < \varepsilon$.
c. For the function $f(x) = -3x + 1$, explain how to find a value for δ that corresponds with any ε so that when $0 < |x - 2| < \delta$ we have $|f(x) + 5| < \varepsilon$.
d. Consider the open interval definition of a limit. What do (c, d) and (p, q) represent. Connect them to the standard definition of a limit that we've been using.

Exercise 1.7.4

Draw a picture of the following situations:

a. A function $f(x)$ with $a = 1$ and $L = 8.5$ where if x is in (0, 2) then $f(x)$ is in (8, 9).
b. A function $g(x)$ with $a = -4$ and $L = 10$ where if x is in (–4.5, –3.5) then $g(x)$ is in (9.8, 10.2).

Exercise 1.7.5

For the following, find values for (c, d) or (p, q). Include a sketch that illustrates the intervals you found with the function.

a. For $f(x) = 2x - 1$, if x is in (1, 2), then $f(x)$ is in (p, q).
b. For $g(x) = x^3 + 1$, if x is in (c, d), then $g(x)$ is in (8, 9).

Exercise 1.7.6

For the function $f(x) = x^2 + 4$, explain how to find an interval (c, d) around $x = 2$ if we are given some interval (p, q) around $L = 8$ such that if x is in (c, d) then $f(x)$ is in (p, q).

This page titled 1.7: Exploration- Limits is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Peter Keep.

2: DERIVATIVES

2.1: Introduction to Derivatives
2.2: Interpreting Derivatives
2.3: Some Early Derivative Rules
2.4: The Product and Quotient Rules
2.5: The Chain Rule

This page titled 2: Derivatives is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Peter Keep.

2.1: Introduction to Derivatives

We'll start this off by thinking about slopes. Before we begin, you should be able to answer the following questions:

- What is a slope? How could you describe it?
- How do you calculate the slope of a line between two points?
- If we have a function $f(x)$ and we pick two points on the curve of the function, what does the slope of a straight line connecting the two points tell us?

What kind of behavior about $f(x)$ does this slope describe?

Defining the Derivative

Activity 2.1.1: Thinking about Slopes.

We're going to calculate and make some conjectures about slopes of lines between points, where the points are on the graph of a function. Let's define the following function:

$$f(x) = \frac{1}{x+2}$$

a. We're going to calculate a lot of slopes! Calculate the slope of the line connecting each pair of points on the curve of f(x):
   - $(-1, f(-1))$ and $(0, f(0))$
   - $(-0.5, f(-0.5))$ and $(0, f(0))$
   - $(-0.1, f(-0.1))$ and $(0, f(0))$
   - $(-0.001, f(-.001))$ and $(0, f(0))$

b. Let's calculate another group of slopes. Find the slope of the lines connecting these pairs of points:
   - $(0, f(0))$ and $(1, f(1))$
   - $(0, f(0))$ and $(0.5, f(0.5))$
   - $(0, f(0))$ and $(0.1, f(0.1))$
   - $(0, f(0))$ and $(0.001, f(0.001))$

c. Just to make it clear what we've done, lay out your slopes in this table:

Between $(0, f(0))$ and...	Slope
$(1, f(1))$	
$(0.5, f(0.5))$	
$(0.1, f(0.1))$	
$(0.01, f(0.01))$	
$(-0.01, f(-0.01))$	
$(-0.1, f(-0.1))$	
$(-0.5, f(-0.5))$	
$(-1, f(-1))$	

d. Now imagine a line that is tangent to the graph of $f(x)$ at $x = 0$. We are thinking of a line that touches the graph at $x = 0$, but runs alongside of the curve at that point instead of through it. Make a conjecture about the slope of this line, using what we've seen above.

e. Can you represent the slope you're thinking of in the table above with a limit? What limit are we approximating in the slope calculations above? Set up the limit and evaluate it, confirming your conjecture.

Activity 2.1.2: Finding a Tangent Line.

Let's think about a new function, $g(x) = \sqrt{2-x}$. We're going to think about this function around the point at $x = 1$.

a. Ok, we are going to think about this function at this point, so let's find the coordinates of the point first. What's the y -value on our curve at $x = 11$?
b. Use a limit similar to the one you constructed in Activity 2.1.1 to find the slope of the line tangent to the graph of $g(x)$ at $x = 1$.
c. Now that you have a slope of this line, and the coordinates of a point that the line passes through, can you find the equation of the line?

Definition 2.1.1: Derivative at a Point.

For a function $f(x)$, we say that the derivative of $f(x)$ at $x = a$ is:

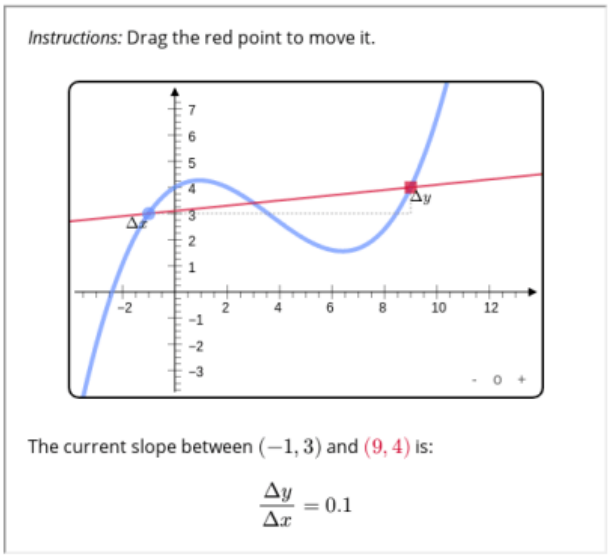
$$f'(a) = \lim_{x \to a} \left(\frac{f(x) - f(a)}{x - a} \right)$$

provided that the limit exists. If $[f'(a)$ exists, then we say $f(x)$ is differentiable at a.

We can investigate this definition visually. Consider the function $f(x)$ plotted below, where we will look at the point $(-1, f(-1))$. In the definition of the limit, we'll let $a = -1$, and so consider:

$$\lim_{x \to -1} \left(\frac{f(x) - f(-1)}{x - (-1)} \right)$$

Can you estimate the limit of the slope of the tangent line as $x \to -1$?



Standalone Embed

Does it look like the limit of the slope between $(-1, f(-1))$ and $(x, f(x))$ exists as $x \to -1$? What do you think it is?

Calculating a Bunch of Slopes at Once

Activity 2.1.3: Calculating a Bunch of Slopes.

Let's do this all again, but this time we'll calculate the slope at a bunch of different points on the same function.

Let's use $j(x) = x^2 - 4$.

a. Start calculating the following derivatives, using the definition of the Derivative at a Point:

- $j'(-2)$
- $j'(0)$
- $j'(1/3)$
- $j'(-1)$

b. Stop calculating the above derivatives when you get tired/bored of it. How many did you get through?

c. Notice how repetitive this is: on one hand, we have to set up a completely different limit each time (since we're looking at a different point on the function each time). On the other hand, you might have noticed that the work is all the same: you factor and cancel over and over. These limits are all ones that we covered in Section 1.3 First Indeterminate Forms, and so it's no surprise that we keep using the same algebra manipulations over and over again to evaluate these limits. Do you notice any patterns, any connections between the x-value you used for each point and the slope you calculated at that point? You might need to go back and do some more.

d. Try to evaluate this limit in general:

$$\begin{aligned} j'(a) &= \lim_{x\to a}\left(\frac{j(x)-j(a)}{x-a}\right) \\ &= \lim_{x\to a}\left(\frac{(x^2-4)-(a^2-4)}{x-a}\right) \end{aligned}$$

Remember, you know how this goes! You're going to do the same type of algebra that you did earlier! What is the formula, the pattern, the way of finding the slope on the $j(x)$ function at any x-value, $x=a$?

e. Confirm this by using your new formula to recalculate the following derivatives:

- $j'(-2)$
- $j'(0)$
- $j'(1/3)$
- $j'(-1)$

We're going to try to think about the derivative as something that can be calculated in general, as well as something that can be calculated at a point. We'll define a new way of calculating it, still a limit of slopes, that will be a bit more general.

Definition 2.1.2: The Derivative Function.

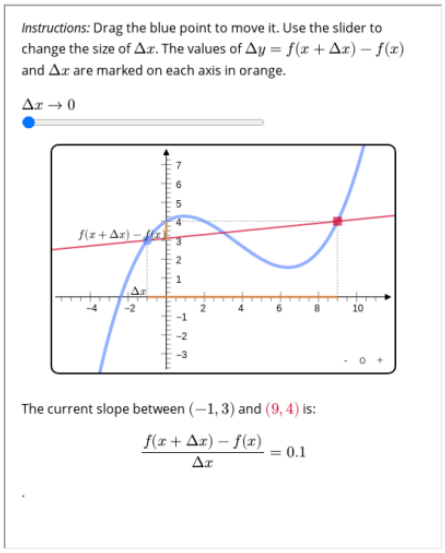
For a function $f(x)$, the derivative of $f(x)$, denoted $f'(x)$, is:

$$f'(x) = \lim_{\Delta x\to 0}\left(\frac{f(x+\Delta x)-f(x)}{\Delta x}\right)$$

for x-values in the domain of $f(x)$ where this limit exists.

This definition feels pretty different, but we hopefully can notice that this is really just calculating a slope. Notice in the following plot there is a significant difference. In the visualization of the Derivative at a Point, the first point was fixed into place and the second point was the one that we moved and changed. It was the one with the variable x-value.

Notice in the following visualization that the first point is the one that is moveable while the second point is defined based on the first one (and the horizontal difference between the points, Δx). This means that we don't need to define one specific point, and can find the slope of the line tangent to$f(x)$ at some changing x-value.



Standalone
Embed

Practice Problems

Exercise 2.1.1

We define a derivative of $f(x)$ at $x = a$ as $f'(x) = \lim_{x \to a} \left(\frac{f(x) - f(a)}{x - a} \right)$. What is this limit? What does it represent in terms of $f(x)$? Feel free to sketch some pictures in order to discuss it visually.

Exercise 2.1.2

For some function $f(x)$, we have that $f'(3) = -4$. What does this mean? Interpret this value in the context of $f(x)$.

Exercise 2.1.3

For each of the following functions, use the definition of the derivative at a point to find the slope of the line tangent to the function's curve at the specified point.

a. $f(x) = 6x + 1$ at (1, 7)
b. $g(x) = x^2 - 5$ at (–1, –4)
c. $j(x) = \sqrt{x}$ at (4, 2)
d. $f(x) = \frac{1}{4 - x}$ at (3, 1)
e. $g(x) = 4 + 3x - 2x^2$ at (2, 2)
f. $j(x) = \frac{1}{x^2}$ at (1, 1)

Answer available at the end of this section.

Exercise 2.1.4

We have talked about two limit-based definitions of derivatives:

$$f'(a) = \lim_{x \to a} \left(\frac{f(x) - f(a)}{x - a} \right)$$
$$f'(x) = \lim_{\Delta x \to 0} \left(\frac{f(x + \Delta x) - f(x)}{\Delta x} \right)$$

a. What are some similarities between these definitions? It might help to compare visual depictions of each.
b. What are some differences between these definitions? Focus on differences in how they are built as well as how we use them.
c. Pick a situation where it might make sense to use each definition. Explain your choice.

Exercise 2.1.5

For each of the functions below, find the derivative function $f'(x)$, and evaluate that function at $x = 1, x = 2, x = 3$, and $x = 4$. Interpret these values.

a. $f(x) = 5x + 2$
b. $f(x) = \sqrt{x + 4}$
c. $f(x) = x^3 - x$
d. $f(x) = \dfrac{x + 1}{x}$

Answer available at the end of this section.

Exercise 2.1.6

Find the derivative function for each of the following functions.

a. $f(x) = 4 - \sqrt{x}$
b. $f(x) = \dfrac{5}{x - 2}$
c. $f(x) = 3\sqrt{x - 1}$
d. $f(x) = x^2 + 3x$
e. $f(x) = \dfrac{1}{x^2 + 1}$
f. $f(x) = \sqrt{x^2 + 1}$

Answer available at the end of this section.

Answers

Exercise 2.1.3

a. $f'(1) = 6$
b. $g'(-1) = -2$
c. $j'(4) = \dfrac{1}{4}$
d. $f'(3) = 1$
e. $g'(2) = -5$
f. $j'(1) = -2$

Exercise 2.1.5

$f'(x) = 5$
$f'(1) = 5$
a. $f'(2) = 5$
$f'(3) = 5$
$f'(4) = 5$

$f'(x) = \dfrac{1}{2\sqrt{x+4}}$

$f'(1) = \dfrac{1}{2\sqrt{5}}$

b. $f'(2) = \dfrac{1}{2\sqrt{6}}$

$f'(3) = \dfrac{1}{2\sqrt{7}}$

$f'(4) = \dfrac{1}{2\sqrt{8}}$

$f'(x) = 3x^2 - 1$

$f'(1) = 2$

c. $f'(2) = 11$

$f'(3) = 26$

$f'(4) = 47$

$f'(x) = -\dfrac{1}{x^2}$

$f'(1) = -1$

d. $f'(2) = -\dfrac{1}{4}$

$f'(3) = -\dfrac{1}{9}$

$f'(4) = -\dfrac{1}{16}$

Exercise 2.1.6

a. $f'(x) = -\dfrac{1}{2\sqrt{x}}$

b. $f'(x) = -\dfrac{5}{(x-2)^2}$

c. $f'(x) = \dfrac{3}{2\sqrt{x-1}}$

d. $f'(x) = 2x + 3$

e. $f'(x) = -\dfrac{2x}{\left(x^2+1\right)^2}$

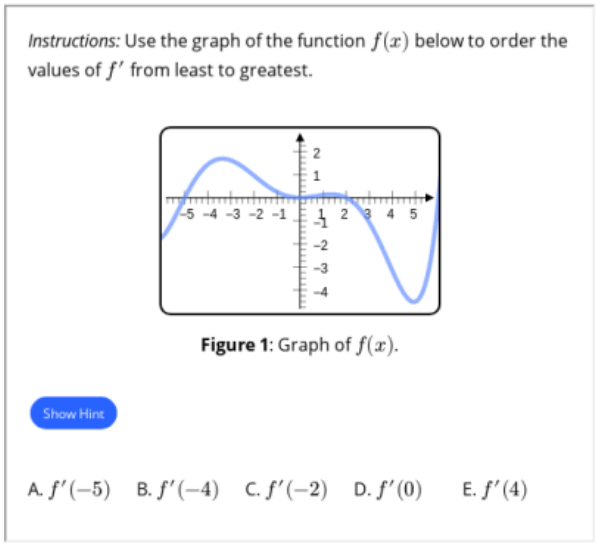
f. $f'(x) = \dfrac{x}{\sqrt{x^2+1}}$

This page titled 2.1: Introduction to Derivatives is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Peter Keep.

2.2: Interpreting Derivatives

What is a derivative?

This can feel like a silly question, since we're calculating derivatives and getting used to finding them. But what is a derivative?



Standalone Embed

In this section, we just want to remind ourselves of what this object—the derivative—is, how we should hold it in our minds as we move through the course, and then practice being flexible with this interpretation.

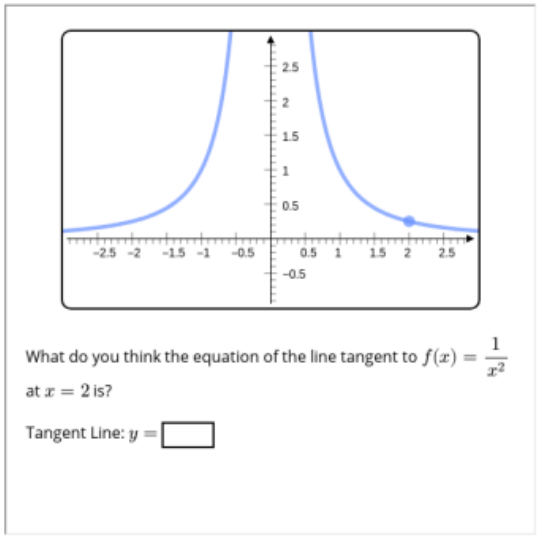
The Derivative is a Slope

Activity 2.2.1: Interpreting the Derivative as a Slope.

In Activity 2.1.1 Thinking about Slopes and Activity 2.1.2 Finding a Tangent Line, we built the idea of a derivative by calculating slopes and using them. Let's continue this by considering the function $f(x) = \frac{1}{x^2}$.

a. Use Definition 2.1.1 Derivative at a Point to find $f'(2)$. What does this value represent?
b. We want to plot the line that would be tangent to the graph of $f(x)$ at $x = 2$. Remember that we can write the equation of a line in two ways:
   - The equation of a line with slope m that passes through the point $(a, f(a))$ is:
   $$y = m(x - a) + f(a)$$
   - The equation of a line with slope m that passes the point $(0, b)$ (this is another way of saying that the y-intercept of the line is b) is:
   $$y = mx + b$$

   Find the equation of the line tangent to $f(x)$ at $x = 2$. Add it to the graph of $f(x) = \frac{1}{x^2}$ below to check your equation.



Standalone Embed

c. This tangent line is very similar to the actual curve of the function $f(x)$ near $x = 2$. Another way of saying this is that while the slope of $f(x)$ is not always the value you found for $f'(2)$, it is close to that for x-values near 2. Use this idea of slope to predict where the y-value of our function will be at 2.01.
d. Compare this value with $f(2.01) = \frac{1}{2.01^2}$. How close was it?

The Derivative is a Rate of Change

Activity 2.2.2: Interpreting the Derivative as a Rate of Change.

This is going to feel somewhat redundant, since we hopefully know that a slope is really just a rate of change. But we'll be able to explore this a bit more and see how we can use a derivative to tell us information about some specific context.

Let's say that we want to model the speed of a car as it races along a strip of the road. By the time we start measuring it (we'll call this time 0), the position of the car (along the straight strip of road) is:

$$s(t) = 73t + t^2$$

where t is time measured in seconds and $s(t)$ is the position measured in feet. Let's say that this function is only relevant on the domain $0 \leq t \leq 15$. That is, we only model the position of the car for a 15- second window as it speeds past us.

a. How far does the car travel in the 15 seconds that we model it? What was the car's average velocity on those 15 seconds?
b. Calculate $s'(t)$, the derivative of $s(t)$, using Definition 2.1.2 The Derivative Function. What information does this tell us about our vehicle?
c. Calculate $s'(0)$. Why is this smaller than the average velocity you found? What does that mean about the velocity of the car?
d. If we call $v(t) = s'(t)$, then calculate $v'(t)$. Note that this is a derivative of a derivative.
e. Find $v'(0)$. Why does this make sense when we think about the difference between the average velocity on the time interval and the value of $v(0)$ that we calculated?
f. What does it mean when we notice that $v'(t)$ is constant? Explain this by interpreting it in terms of both the velocity of the vehicle as well as the position.

The Derivative is a Limit

Look back at the definition of Derivative at a Point. The end of it is interesting: "provided that the limit exists." We need to keep

in mind that this is a limit, and so a derivative exists or fails to exist whenever that limit exists or fails to exist.

What are some ways that a limit fails to exist?

- A limit doesn't exist if the left-side limit and the right-side limit do not match: Theorem 1.1.4 Mismatched Limits.
- A limit doesn't exist if it is an Infinite Limit.

What do each of these situations look like when we're considering the limit of slopes?

When Does a Derivative Not Exist?

1. A derivative doesn't exist at points where the slopes on either side of the point don't match.
2. A derivative doesn't exist at points with vertical tangent lines.
3. A derivative doesn't exist at points where the function is not continuous.

The Derivative is a Function

Our last interpretation of the derivative will be thinking about it as a function. We will expand on this later, in Chapter 4.

Activity 2.2.3: Interpreting the Derivative as a Function.

In Activity 2.1.3 Calculating a Bunch of Slopes, we calculated the derivative function for $j(x) = x^2 - 4$. Using the definition of The Derivative Function, we can see that $j'(x) = 2x$. Let's explore that a bit more.

a. Sketch the graphs of $j(x) = x^2 - 4$ and $j'(x) = 2x$. Describe the shapes of these graphs.
b. Find the coordinates of the point at $x = \frac{1}{2}$ on both the graph of $j(x)$ and $j'(x)$. Plot the point on each graph.
c. Think back to our previous interpretations of the derivative: how do we interpret the y-value output you found for the \(j^{\prime \) function?
d. Find the coordinates of another point at some other x-value on both the graph of $j(x)$ and $j'(x)$. Plot the point on each graph, and explain what the output of j' tells us at this point.
e. Use your graph of $j'(x)$ to find the x-intercept of $j'(x)$. Locate the point on $j(x)$ with this same x-value. How do we know, visually, that this point is the x-intercept of $j'(x)$?
f. Use your graph of $j'(x)$ to find where $j'(x)$ is positive. Pick two x-values where $j'(x) > 0$. What do you expect this to look like on the graph of $j(x)$? Find the matching points (at the same x-values) on the graph of $j(x)$ to confirm.
g. Use your graph of $j'(x)$ to find where $j'(x)$ is negative. Pick two x-values where $j'(x) < 0$. What do you expect this to look like on the graph of $j(x)$? Find the matching points (at the same x-values) on the graph of $j(x)$ to confirm.
h. Let's wrap this up with one final pair of points. Let's think about the point (–3, 5) on the graph of $j(x)$ and the point (–3, –6) on the graph of $j'(x)$. First, explain what the value of –6 (the output of j' at $x = -3$) means about the point (–3, 5) on $j(x)$. Finally, why can we not use the value 5 (the output of j at $x = -3$) means about the point (–3, –6) on $j'(x)$?

Notation for Derivatives

So far we've been using the "prime" notation to represent derivatives: the derivative of $f(x)$ is $f'(x)$. We will continue to use this notation, but we'll introduce a bunch of other ways of writing notation to represent the derivative. Each new notation will emphasize some aspect of the derivative that will serve to be useful, even though they all represent essentially the same thing.

Function	Derivative	Derivative at $x = a$	Emphasis
$f(x)$	$f'(x)$	$f'(a)$	The derivative is a function. The function takes in -value inputs and returns the slope of f at that x-value.
y	y'	$y'\|_{x=a}$	We can find slopes on any curve, not just functions. This sometimes is also used as a way to simplify the notation, especially when we want to manipulate equations involving y'.

Function	Derivative	Derivative at $x = a$	Emphasis
y	$\frac{dy}{dx}$	$\left.\frac{dy}{dx}\right\|_{x=a}$	The derivative is a slope. It is $\frac{\Delta y}{\Delta x}$ as $\Delta x \to 0$, and we use dx and dy (called ***differentials***) to represent Δx and Δy as the limits as $\Delta x \to 0$. This notation also is useful to tell us what the rate of change is: what is changing (in this case y) and what is it changing based on (in this case x)?
$f(x)$	$\frac{d}{dx}(f(x))$	$\left.\frac{d}{dx}(f(x))\right\|_{x=a}$	The derivative is an action that we do to some function. We can call it an ***operator*** although we won't formally define that term in this text. We'll look at this idea more in **Section 3.1**. We can specify what we are expecting the input variable to be, based on the differential in the denominator.

Practice Problems

Exercise 2.2.1

For each of the following functions, use the definition of the derivative at a point to find $f'(a)$ and then find the equation of the line tangent to the graph of $f(x)$ at $x = a$.

Remember: The equation of a line given the slope m and (x_1, y_1), a point on the line is $y = m(x - x_1) + y_1$.

a. $f(x) = x^2 + x - 1$ with $a = 0$
b. $f(x) = x^3 + 1$ with $a = 2$
c. $f(x) = \frac{2}{\sqrt{x}}$ with $a = 6$
d. $f(x) = \sqrt{3x - 1}$ with $a = 2$
e. $f(x) = x - x^2$ with $a = 2$
f. $f(x) = \frac{x}{x+1}$ with $a = 1$

Answer available at the end of this section.

Exercise 2.2.2

We start watching a car driving down a straight road towards a stop sign. We can tell that their position function is $s(t) = -7.5t^2 + 81t$ with position measured in feet from their location when we began observing them and time measured in seconds.

a. How fast is the car driving at $t = 0$?
b. How fast is the car driving at $t = 2$?
c. At what value of t does the car stop (their velocity is 0)?
d. How far did they travel before they stopped?
e. At $t = 0$ what is the acceleration of the car?
f. At $t = 2$ what is the acceleration of the car?

Answer available at the end of this section.

Exercise 2.2.3

The monetary value of an ancient mathematical artifact (the pencil that Isaac Barrow used to prove the Fundamental Theorem of Calculus in 1667) follows the function $m(t) = \sqrt{1600t} + 10$ where t is the number of years since 1667 and $m(t)$ is the monetary value of the pencil in US Dollars.

a. What was the overall change in value from 1667 to 1967 (from $t = 0$ to $t = 300$)?
b. What was the average rate of change of the value from 1667 to 1967?
c. What was the instantaneous rate of change of the value in 1967?
d. Find $m(354)$ and $m'(354)$, and interpret both of these values.
e. Use $m(354)$ and $m'(354)$ to approximate the value of the pencil in 2022.
f. Use $m(t)$ to find the value in 2022. Was the approximation close or not? Explain.

Answer available at the end of this section.

Answers

Exercise 2.2.1

a. $f'(0) = 1$
$y = x - 1$
b. $f'(2) = 12$
$y = 12x - 15$
c. $f'(6) = \dfrac{1}{\sqrt{6}}$
$y = -\dfrac{x}{6\sqrt{6}} + \dfrac{3}{\sqrt{6}}$
d. $f'(2) = \dfrac{3}{2\sqrt{5}}$
$y = -\dfrac{3x}{2\sqrt{5}} + \dfrac{2}{\sqrt{5}}$
e. $f'(2) = -3$
$y = -3x + 4$
f. $f'(1) = \dfrac{1}{4}$
$y = \dfrac{x}{4} + \dfrac{1}{4}$

Exercise 2.2.2

a. $v(0) = s'(0) = 81$
b. $v(2) = s'(2) = 51$
c. $t = 5.4$
d. $s(5.4) - s(0) = 396.9$
e. $a(0) = v'(0) = -15$
f. $a(2) = v'(2) = -15$

Exercise 2.2.3

a. $m(300) - m(0) \approx \$292.82$
b. $\dfrac{m(300) - m(0)}{300 - 0} \approx \$0.98/\text{ year}$
c. $m'(300) \approx \$1.55/\text{ year}$
d. $m(354) \approx \$762.60$ is the value of the pencil 354 years after 1667 (in 2021). $m'(354) \approx \$1.06$ is the rate of change in the value of the pencil 354 years after 1667 (in 2021). This means that in 2021, we expect the price to increase at a rate of \$1.06/year.
e. $m(354) + m'(354) \approx \$763.66$
f. $m(355) \approx \$763.66$

This page titled 2.2: Interpreting Derivatives is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Peter Keep.

2.3: Some Early Derivative Rules

We are going to break this topic into two parts:

1. We will try to find some common patterns or connections between derivatives and specific functions. For instance, when we use Definition 2.1.2 The Derivative Function to build a derivative, are there patterns in the work of evaluating that limit that will allow us to get through the limit work quickly? Can we group some functions together based on how we might deal with the limit?
2. We will try to think about derivatives a bit more generally and show how we can build some basic properties to help us think about differentiating variations of the functions that we recognize.

Derivatives of Power Functions

Activity 2.3.1: Derivatives of Power Functions.

We're going to do a bit of pattern recognition here, which means that we will need to differentiate several different power functions. For our reference, a power function (in general) is a function in the form $f(x) = a\,(x^n)$ where n and a are real numbers, and $a \neq 0$. Let's begin our focus on the power functions x^2, x^3, and x^4. We're going to use Definition 2.1.2 The Derivative Function a lot, so feel free to review it before we begin.

a. Find $\frac{d}{dx}\,(x^2)$. As a brief follow up, compare this to the derivative $j'(x)$ that you found in Activity 2.1.3 Calculating a Bunch of Slopes. Why are they the same? What does the difference, the –4, in the $j(x)$ function do to the graph of it (compared to the graph of x^2) and why does this not impact the derivative?
b. Find $\frac{d}{dx}\,(x^3)$.
c. Find $\frac{d}{dx}\,(x^4)$.
d. Notice that in these derivative calculations, the main work is in multiplying $(x+\Delta x)^n$. Look back at the work done in all three of these derivative calculations and find some unifying steps to describe how you evaluate the limit/calculate the derivative after this tedious multiplication was finished. What steps did you do? Is it always the same thing? Another way of stating this is: if I told you that I knew what $(x+\Delta x)^5$ was, could you give me some details on how the derivative limit would be finished?
e. Finish the following derivative calculation:

$$\begin{aligned}\frac{d}{dx}\,(x^5) &= \lim_{\Delta x\to 0}\left(\frac{(x+\Delta x)^5 - x^5}{\Delta x}\right)\\ &= \lim_{\Delta x\to 0}\left(\frac{(x^5+5x^4\Delta x+10x^3\Delta x^2+10x^2\Delta x^3+5x\Delta x^4+\Delta x^5)-x^5}{\Delta x}\right)\\ &= \rightsquigarrow \ldots\end{aligned}$$

f. Make a conjecture about the derivative of a power function in general, $\frac{d}{dx}\,(x^n)$.

Something to notice here is that the calculation in this limit is really dependent on knowing what $(x+\Delta x)^n$ is. When n is an integer with $n \geq 2$, this really just translates to multiplication. If we can figure out how to multiply $(x+\Delta x)^n$ in general, then this limit calculation will be pretty easy to do. We noticed that:

1. The first term of that multiplication will combine with the subtraction of x^n in the numerator and subtract to 0.
2. The rest of the terms in the multiplication have at least one copy of Δx, and so we can factor out Δx and "cancel" it with the Δx in the denominator.
3. Once this is done, we've escaped the portion of the limit that was giving us the $\frac{0}{0}$ indeterminate form, and so we can evaluate the limit as $\Delta x \to 0$. The result is just whatever terms still have at least one remaining copy of Δx in them "go to" 0, and we're left with just the terms that do not have any copies of Δx in them.

There is a really nice theorem that helps to understand how the $(x+\Delta x)^n$ multiplication goes (for different positive integer exponents, n) called the Binomial Theorem, but we won't introduce it here, fully. What we'll do, instead, is summarize the results:

1. In each term, the exponents on x count down from n to 0.
2. Similarly, in each term, the exponents on Δx will count up from 0 to n.
3. The coefficients for each term are generated from a very cool pattern, displayed below.

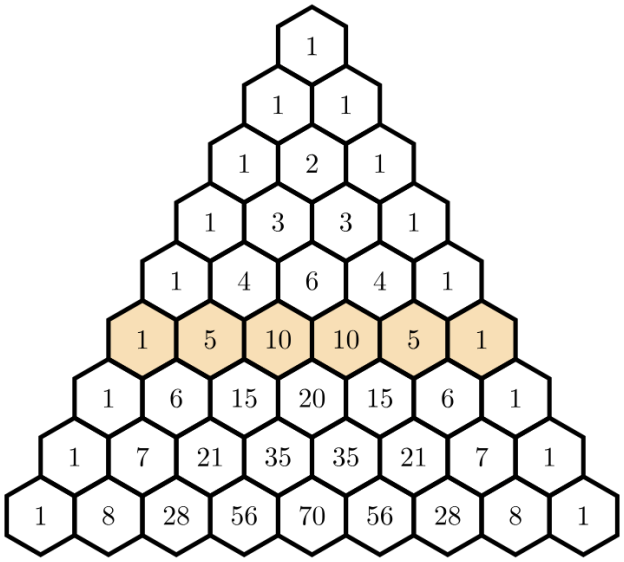


Figure 2.3.1

In each row, the coefficients are built by the sum of the two coefficients above it. We have highlighted the 5th row, in order to compare it to the work we did in Activity 2.3.1.

$$(x+\Delta x)^5 = x^5 + 5x^4\Delta x + 10x^3\Delta x^2 + 10x^2\Delta x^3 + 5x\Delta x^4 + \Delta x^5$$

We can see the coefficients on each term match the highlighted row!

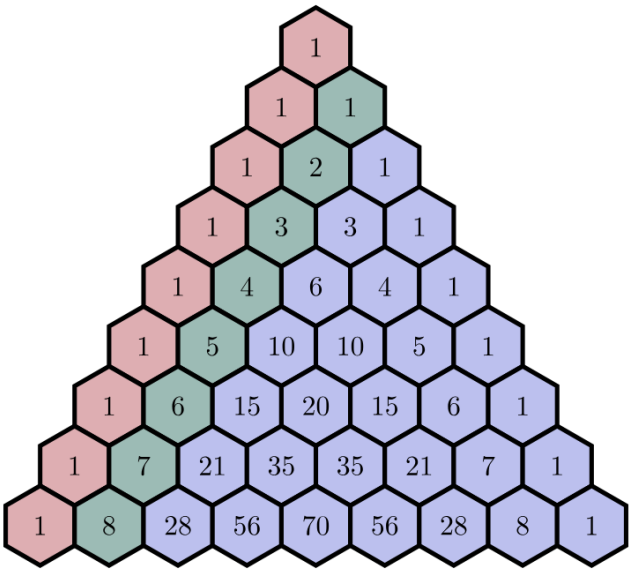


Figure 2.3.2

We can now think of each row, and figure out what was happening in the limit work we did.

1. The first term of that multiplication will combine with the subtraction of x^n in the numerator and subtract to 0. We can designate this in the triangle by shading in the first term in red.
2. The rest of the terms in the multiplication have at least one copy of Δx, and so we can factor out Δx and "cancel" it with the Δx in the denominator. We can designate the terms in the triangle without any remaining Δx's in them by shading the second term in green. The rest, that have at least one Δx in it, are shaded in blue.
3. Once this has done, we've escaped the portion of the limit that was giving us the $\frac{0}{0}$ indeterminate form, and so we can evaluate the limit as $\Delta x \to 0$. The result is just that whatever terms still have at least one remaining copy of Δx in it "go to" 0, and we're left with just the terms that do not have any copies of Δx in them. Note that all of the blue terms go to 0 in the limit, and we are left with the green terms, but without any Δx. Since these are the second terms, we have the coefficient and one fewer exponent on x.

Theorem 2.3.3: Power Rule for Derivatives.

$$\frac{d}{dx}\left(x^n\right) = nx^{n-1}$$

We have shown that this is true for $n = 2, 3, 4, \ldots$, but this is also true for any value of n (including $n = 1$, non-integers, and non-positives). We will prove this more formally later (in Section 3.3), and until then we will be free to use this result.

Example 2.3.4

Let's confirm this Power Rule for two examples that we are familiar with.

a. Find the derivative $\frac{d}{dx}(\sqrt{x})$ using the limit definition of the derivative function. Note that $\sqrt{x} = x^{1/2}$ and $\frac{1}{2\sqrt{x}} = \frac{1}{2}x^{-1/2}$.

b. Find the derivative $\frac{d}{dx}\left(\frac{1}{x}\right)$ using the limit definition of the derivative function. Note that $\frac{1}{x} = x^{-1}$ and $-\frac{1}{x^2} = -x^{-2}$.

In this activity, we also found one other result.

Theorem 2.3.5: Derivative of a Constant Function.

If $y = k$ where k is some real number constant, then $y' = 0$. Another way of saying this is:

$$\frac{d}{dx}(k) = 0$$

At this point, we have a pretty good handle on some really common function types: power functions of x. Let's look at some more functions that we'll use frequently.

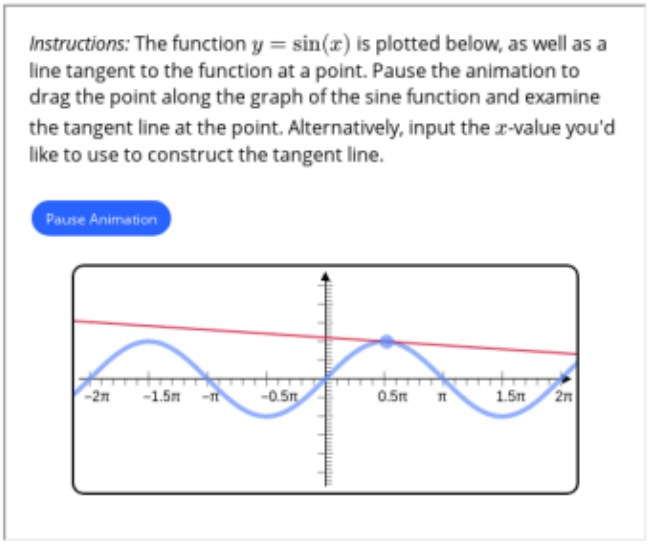
Derivatives of Trigonometric and Exponential Functions

We'll start this off by thinking of the two foundational trigonometric functions, $\sin(x)$ and $\cos(x)$.

Activity 2.3.2: Derivatives of Trigonometric Functions.

Let's try to think through the derivatives of $y = \sin(\theta)$ and $y = \cos(\theta)$. In this activity, we'll look at graphs and try to collect some information about the derivative functions. We'll be practicing our interpretations, so if you need to brush up on Section 2.2 before we start, that's fine!

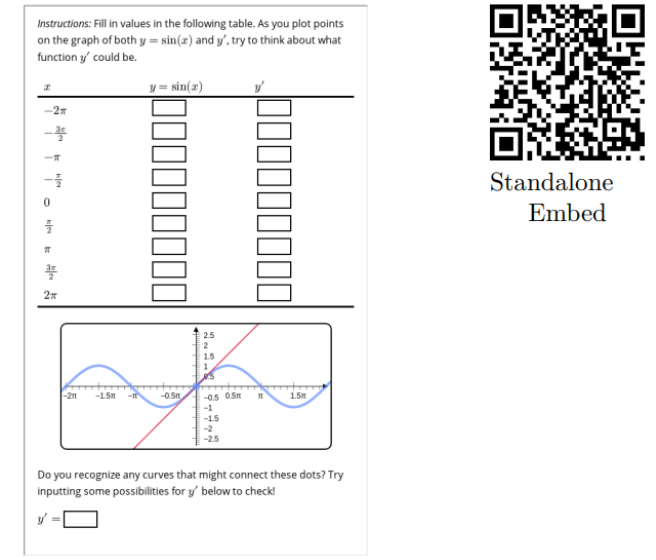
a. The following plot includes both the graph of $y = \sin(x)$, and the line tangent to $y = \sin(x)$. Watch as the point where we build the tangent line moves along the graph, between $x = -2\pi$ and $x = 2\pi$.



Standalone Embed

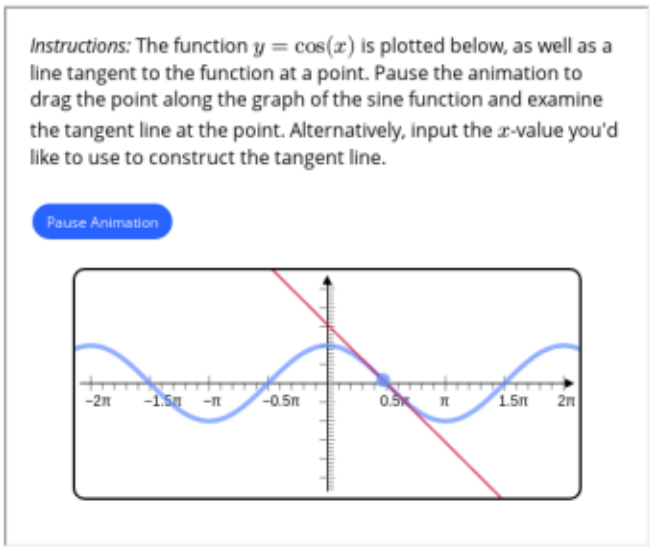
Collect as much information about the derivative, y', as you can. What do you know about it? What are some facts about the slopes of the tangent lines in this animation?

b. We're going to get more specific here: let's find the coordinates of points that are on both the graph of $y = \sin(x)$ and it's derivative y'. Remember, to get the values for y', we're really looking at the slope of the tangent line at that point.



Standalone Embed

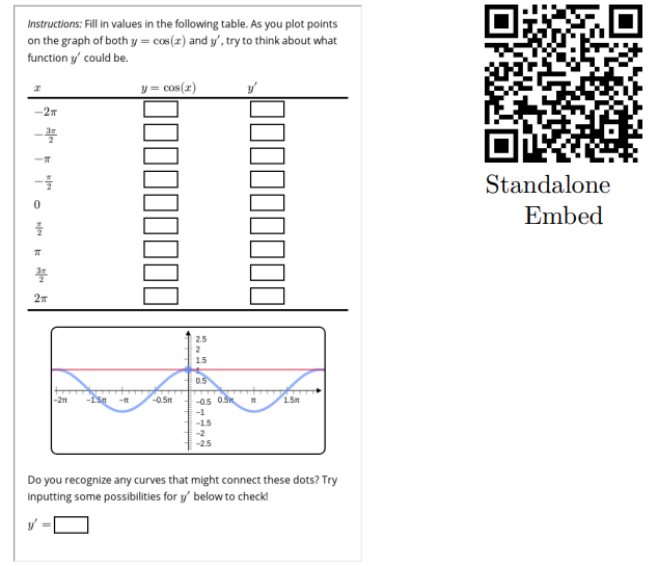
c. Let's repeat this process using the $y = \cos(x)$ function instead. The following plot includes both the graph of $y = \cos(x)$, and the line tangent to $y = \cos(x)$. Watch as the point where we build the tangent line moves along the graph, between $x = -2\pi$ and $x = 2\pi$.



Standalone Embed

Collect as much information about the derivative, y', as you can. What do you know about it? What are some facts about the slopes of the tangent lines in this animation?

d. We're going to get more specific here: let's find the coordinates of points that are on both the graph of $y = \cos(x)$ and it's derivative y'. Remember, to get the values for y', we're really looking at the slope of the tangent line at that point.



Standalone Embed

Theorem 2.3.6: Derivatives of the Sine and Cosine Functions.

$$\frac{d}{d\theta}(\sin(\theta)) = \cos(\theta)$$

$$\frac{d}{d\theta}(\cos(\theta)) = -\sin(\theta)$$

Proof available at the end of this section.

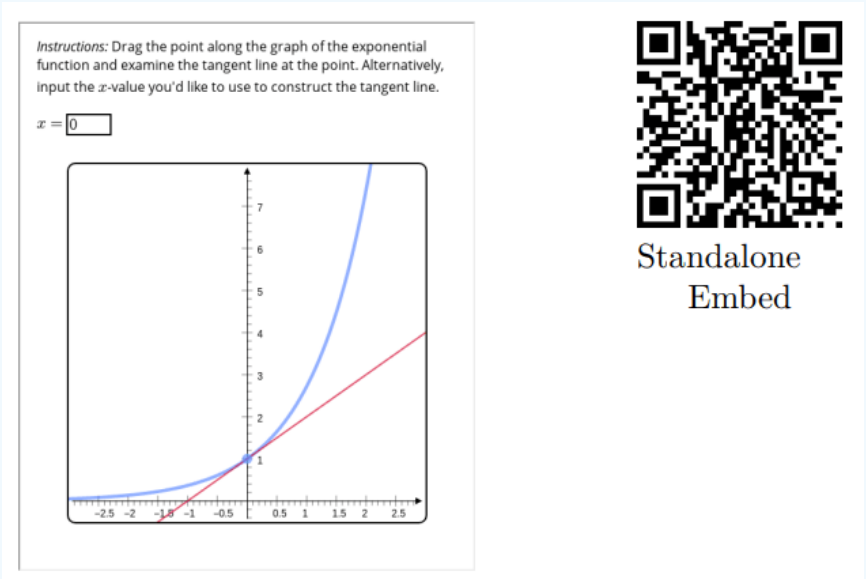
We'll think about the other trigonometric functions a bit later, in Section 2.4. Let's move on to another useful function, the

exponential function e^x.

Activity 2.3.3: Derivative of the Exponential Function.

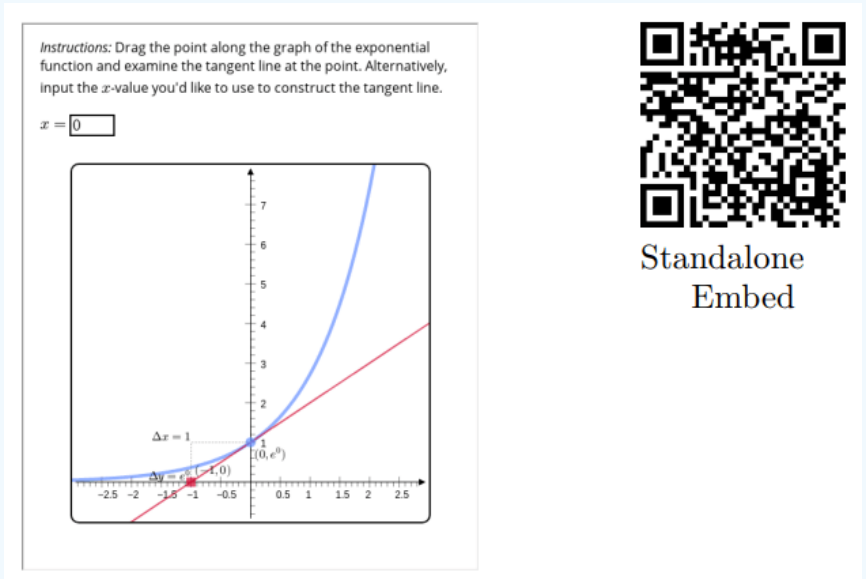
We're going to consider a maybe-unfamiliar function, $f(x) = e^x$. We'll explore this function similarly to how we thought of the derivatives of sine and cosine in Activity 2.3.2: we'll look at a tangent line at different points, and think about the slope.

a. The plot below includes both the graph of $y = e^x$ and the line tangent to $y = e^x$. Watch as the point moves along the curve.



Standalone Embed Collect as much information about the derivative, y', as you can. What do you know about it? What are some facts about the slopes of the tangent lines in this animation?

b. There is a slightly hidden fact about slopes and tangent lines in this animation. In the following animation, we'll add (and label) one more point. Let's look at this again, this time noting the point at which this tangent line crosses the x-axis. This will make it easier to think about slopes!



What information does this reveal about the slopes?

c. Make a conjecture about the slope of the line tangent to the exponential function $y = e^x$ at any x-value. What do you believe the formula/equation for y' is then?

Theorem 2.3.9: Derivative of the Exponential Function.

$$\frac{d}{dx}\left(e^x\right) = e^x$$

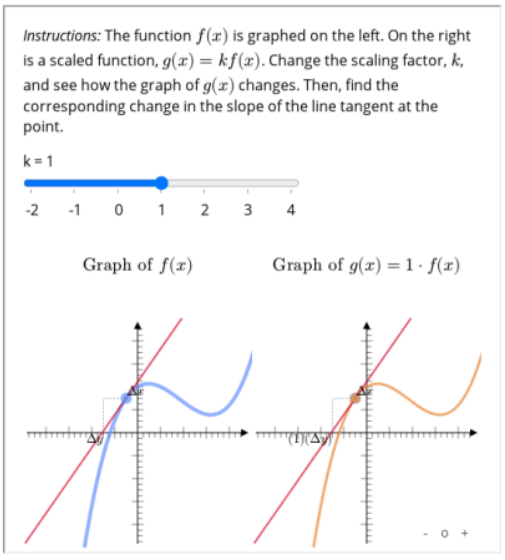
Some Properties of Derivatives in General

Theorem 2.3.10: Combinations of Derivatives.

If $f(x)$ and $g(x)$ are differentiable functions, then the following statements about their derivatives are true.

1. Sums: The derivative of the sum of $f(x)$ and $g(x)$ is the sum of the derivatives of $f(x)$ and $g(x)$:
$$\begin{aligned}\frac{d}{dx}(f(x)+g(x)) &= \left(\frac{d}{dx}f(x)\right)+\left(\frac{d}{dx}g(x)\right)\\ &= f'(x)+g'(x)\end{aligned}$$
2. Differences: The derivative of the difference of $f(x)$ and $g(x)$ is the difference of the derivatives of $f(x)$ and $g(x)$:
$$\begin{aligned}\frac{d}{dx}(f(x)-g(x)) &= \left(\frac{d}{dx}f(x)\right)-\left(\frac{d}{dx}g(x)\right)\\ &= f'(x)-g'(x)\end{aligned}$$
3. Coefficients: If k is some real number coefficient, then:
$$\begin{aligned}\frac{d}{dx}(kf(x)) &= k\left(\frac{d}{dx}f(x)\right)\\ &= kf'(x)\end{aligned}$$

We can think about each of these properties through the lense of how combining these functions impacts the slopes. For instance, if we wanted to visualize the property about coefficients (that $\frac{d}{dx}(kf(x)) = k\frac{d}{dx}(f(x))$), we can visualize this coefficient as a scaling factor.



Standalone Embed

Example 2.3.11: Putting These Together.

Find the following derivatives:

a. $\frac{d}{dx}\left(4x^5 - \frac{5x}{2} + 6\cos(x) - 1\right)$
b. $\frac{d}{dx}\left(\frac{1}{x} - \frac{\sqrt{x}}{4} + \sin(x) - 5e^x\right)$

Solution

a.
$$\begin{aligned}\frac{d}{dx}\left(4x^5 - \frac{5x}{2} + 6\cos(x) - 1\right) &= \frac{d}{dx}\left(4x^5\right) - \frac{d}{dx}\left(\frac{5x}{2}\right) + \frac{d}{dx}(6\cos(x)) - \frac{d}{dx}(1) \\ &= 4\frac{d}{dx}\left(x^5\right) - \frac{5}{2}\frac{d}{dx}(x) + 6\frac{d}{dx}(\cos(x)) - \frac{d}{dx}(1) \\ &= 4\left(5x^4\right) - \frac{5}{2}(1) + 6(-\sin(x)) - 0 \\ &= 20x^4 - \frac{5}{2} - 6\sin(x)\end{aligned}$$

b.
$$\begin{aligned}\frac{d}{dx}\left(\frac{1}{x} - \frac{\sqrt{x}}{4} + \sin(x) - 5e^x\right) &= \frac{d}{dx}\left(\frac{1}{x}\right) - \frac{d}{dx}\left(\frac{\sqrt{x}}{4}\right) + \frac{d}{dx}(\sin(x)) - \frac{d}{dx}\left(5e^x\right) \\ &= -\frac{1}{x^2} + \frac{1}{8\sqrt{x}} + \cos(x) - 5e^x\end{aligned}$$

Practice Problems

Exercise 2.3.1

Consider the following table of values of the derivative for f and g:

x	-1	0	1	2
$f'(x)$	4	$\frac{1}{3}$	$-\frac{5}{2}$	0
$g'(x)$	-3	-7	$\frac{1}{2}$	-2

For each version of $h(x)$, find the derivative indicated.

a. $h(x) = 3f(x)$; Find $h'(0)$
b. $h(x) = -2g(x)$; Find $h'(-1)$
c. $h(x) = \frac{1}{3}f(x) - \frac{2}{5}g(x)$; Find $h'(2)$
d. $h(x) = g(x) - \frac{f(x)}{6}$; Find $h'(1)$

Answer available at the end of this section.

Exercise 2.3.2

For each of the following functions, find $\frac{dy}{dx}$.

a. $y = 3x^2$
b. $y = 5x^3 + 1$
c. $y = 4x^3 - 5x^{-2}$
d. $y = \sqrt{x} + 4x^{15} + 400$
e. $y = e^3 + e^\pi - e^x$
f. $y = \frac{\sqrt{x^3}}{5} - \sqrt[3]{x^2}$

g. $y = \dfrac{5}{x^3} + \dfrac{x^3}{5}$
h. $y = \sqrt[3]{x^2} - 23$
i. $y = e^x - x^e$
j. $y = 4e^x + \dfrac{2}{\sqrt{x^5}}$

Answer available at the end of this section.

Exercise 2.3.3

We have not yet learned how to differentiate products and quotients of functions. For each of the following functions including products or quotients, rewrite the function and find the derivative.

a. $f(x) = \dfrac{x^2+1}{x}$
b. $g(x) = (3x+1)^2$
c. $p(x) = \dfrac{x^2-9}{x+3}$
d. $r(h) = \dfrac{h-2}{\sqrt{h}-\sqrt{2}}$
e. $\Gamma(x) = \sqrt{x}\left(x^{5/2}-5\right)$
f. $w(t) = \dfrac{te^t-1}{t}$
g. $g(w) = w(w+3)^2(2w-5)^2$
h. $s(t) = \dfrac{6t^{3/2}+4t^{5/3}}{5t^{1/3}}$

Answer available at the end of this section.

Exercise 2.3.4

For each of the following functions, find $f'(x)$, and then $\dfrac{d}{dx}\left(f'(x)\right) = f''(x)$.

a. $f(x) = 7x^5 - 2x^4 + x^3 - 7x^2 + 14x - 10$
b. $f(x) = 4e^x - \dfrac{3}{x^5} + x$

Answer available at the end of this section.

Exercise 2.3.5

For each of the following functions, use $f'(a)$ to find the equation of the line tangent to $f(x)$ at a (for the given value of a).

a. $f(x) = 4x^2 - x^3$ for $a = 3$
b. $f(x) = e^x - 4x$ for $a = 0$
c. $f(x) = x^2 - x + 1$ for $a = -2$
d. $f(x) = 4\sqrt{x} - x$ for $a = 9$

Answer available at the end of this section.

Exercise 2.3.6

Consider the function $f(x) = 4e^x - 5x$.

a. Find the derivative
b. Interpret $f'(x)$. What does it tell us about $f(x)$?
c. Find the value of x where the slope of the line tangent to $f(x)$ is 6.
d. Find the value for x where $f(x)$ has a horizontal tangent line.

Answer available at the end of this section.

Exercise 2.3.7

Consider the function $f(x) = \sqrt[3]{x} + x$.

a. Find the derivative $f'(x)$.
b. Find the value for x where the slope of the line tangent to $f(x)$ is 2.
c. Find the value for x where $f(x)$ has a vertical tangent line.

Answer available at the end of this section.

Exercise 2.3.8

Give an example of a function that meets the following requirements.

a. The equation of the line tangent to the curve at $x = 2$ is $y = -3x + 4$.
b. The tangent line is horizontal at $x = 5$, and the function also passes through the point (1, –2).
c. The derivative is positive when $x < -1$ and negative when $x > -1$.

Answer available at the end of this section.

Answers

Theorem 2.3.6: Derivatives of the Sine and Cosine Functions.

In order to show why $\frac{d}{d\theta}(\sin(\theta)) = \cos(\theta)$ and $\frac{d}{dx}(\cos(\theta)) = -\sin(\theta)$, we will work with the limit definitions of both. Consider both:

$$\frac{d}{d\theta}(\sin(\theta)) = \lim_{\Delta\theta \to 0}\left(\frac{\sin(\theta + \Delta\theta) - \sin(\theta)}{\Delta\theta}\right)$$
$$\frac{d}{d\theta}(\cos(\theta)) = \lim_{\Delta\theta \to 0}\left(\frac{\cos(\theta + \Delta\theta) - \cos(\theta)}{\Delta\theta}\right)$$

Our goal is to rewrite the numerators in both of these limits as something more usable. So far, we have been evaluating these types of limits (First Indeterminate Forms) using some algebraic manipulations. Instead of using algebra, we will use geometry. Consider the unit circle below. We have plotted the angle θ and are reminded that the point on the circle that corresponds with the terminal side of the angle θ has coordinates $(\cos(\theta), \sin(\theta))$. We can label the sides of the triangle pictured below.

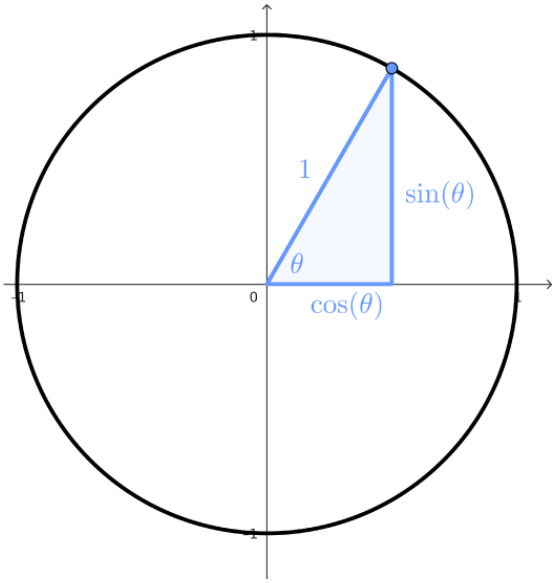


Figure 2.3.7 Unit circle with the angle θ.

Now we consider a second point on the circle, this one formed by the terminal side of the angle $(\theta + \Delta\theta)$. This point has coordinates $(\cos(\theta + \Delta\theta), \sin(\theta + \Delta\theta))$. Note below, we want to find expressions for $\sin(\theta + \Delta\theta) - \sin(\theta)$ and $\cos(\theta + \Delta\theta) - \cos(\theta)$. We can find these geometrically.

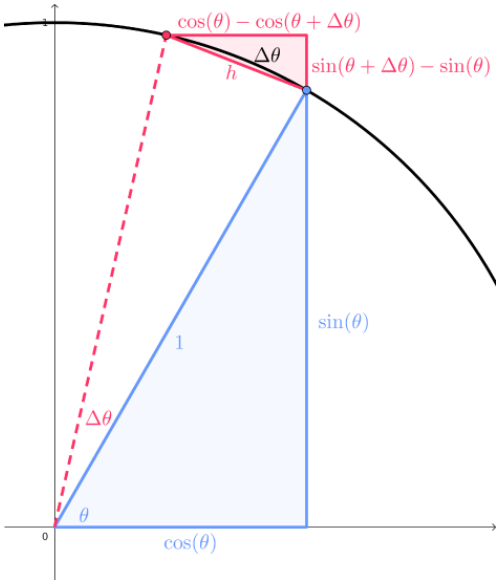


Figure 2.3.8 Angles θ and $(\theta + \Delta\theta)$.

Note, then, that the two triangles look to be similar triangles. We can say, then, that:

$$\sin(\theta + \Delta\theta) - \sin(\theta) \approx h\cos(\theta)$$
$$\cos(\theta) - \cos(\theta + \Delta\theta) \approx h\sin(\theta)$$

In fact, we will find that in the limit as $\Delta\theta \to 0$, the length h matches the arc length $\Delta\theta$ perfectly, and thus lays at a right angle to the terminal side of the angle $(\theta + \Delta\theta)$. Since as $\Delta\theta \to 0$ we have $h \to \Delta\theta$, we can say:

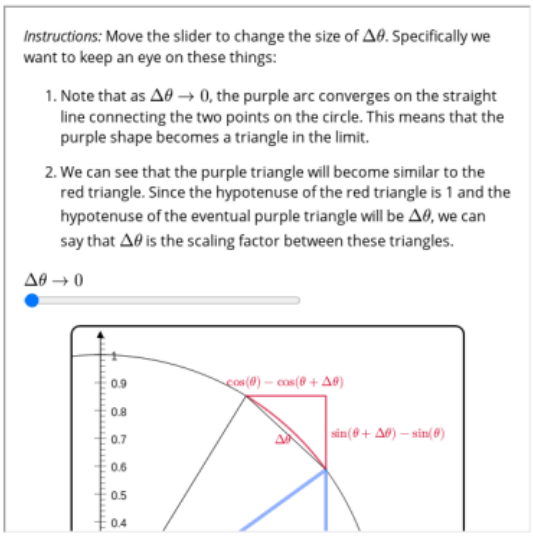
$$(\sin(\theta + \Delta\theta) - \sin(\theta)) \to \Delta\theta\cos(\theta)$$
$$(\cos(\theta) - \cos(\theta + \Delta\theta)) \to \Delta\theta\sin(\theta)$$

Consider, then, the limits involved in the derivative calculations that we built earlier.

$$\begin{aligned}
\frac{d}{d\theta}(\sin(\theta)) &= \lim_{\Delta\theta\to 0}\left(\frac{\sin(\theta+\Delta\theta)-\sin(\theta)}{\Delta\theta}\right)\\
&= \lim_{\Delta\theta\to 0}\left(\frac{\Delta\theta\cos(\theta)}{\Delta\theta}\right)\\
&= \lim_{\Delta\theta\to 0}(\cos(\theta))\\
&= \cos(\theta)\\
\frac{d}{d\theta}(\cos(\theta)) &= \lim_{\Delta\theta\to 0}\left(\frac{\cos(\theta+\Delta\theta)-\cos(\theta)}{\Delta\theta}\right)\\
&= \lim_{\Delta\theta\to 0}\left(\frac{-(\cos(\theta)-\cos(\theta+\Delta\theta))}{\Delta\theta}\right)\\
&= \lim_{\Delta\theta\to 0}\left(\frac{-\Delta\sin(\theta)}{\Delta\theta}\right)\\
&= \lim_{\Delta\theta\to 0}(-\sin(\theta))\\
&= -\sin(\theta)
\end{aligned}$$

So we have shown that $\frac{d}{d\theta}(\sin(\theta)) = \cos(\theta)$ and $\frac{d}{d\theta}(\cos(\theta)) = -\sin(\theta)$ as we claimed.

To see a clearer and more interactive version of this visual proof, see the figure below.





Exercise 2.3.1

a.
$$\begin{aligned}
h'(x) &= 3f'(x)\\
h'(0) &= 3f'(0)\\
&= 3\left(\frac{1}{3}\right)\\
&= 1
\end{aligned}$$

b.
$$\begin{aligned}
h'(x) &= -2g'(x)\\
h'(-1) &= -2g'(-1)\\
&= -2(-3)\\
&= 6
\end{aligned}$$

c.
$$\begin{aligned}
h'(x) &= \frac{1}{3}f'(x) - \frac{2}{5}g'(x)\\
h'(2) &= \frac{1}{3}f'(2) - \frac{2}{5}g'(2)\\
&= \left(\frac{1}{3}\right)(0) - \left(\frac{2}{5}\right)(-2)\\
&= -\frac{4}{5}
\end{aligned}$$

d.
$$\begin{aligned}
h'(x) &= g'(x) - \frac{f'(x)}{6}\\
h'(1) &= g'(1) - \frac{f'(1)}{6}\\
&= \frac{1}{2} - \left(-\frac{5}{12}\right)\\
&= \frac{11}{12}
\end{aligned}$$

Exercise 2.3.2

a. $\frac{dy}{dx} = 6x$

b. $\frac{dy}{dx} = 15x^2$

c. $\frac{dy}{dx} = 12x^2 + 10x^{-3}$

d. Since $y = x^{1/2} + 4x^{15} + 400$, we can write $\frac{dy}{dx} = \frac{1}{2}y^{-1/2} + 60x^{14}$, or:

$$\frac{dy}{dx} = \frac{1}{2\sqrt{x}} + 60x^{14}$$

e. Since e^3 and e^π are real numbers, we can write:

$$\frac{dy}{dx} = -e^x$$

f.
$$\begin{aligned} y &= \frac{1}{5}x^{1/3} - x^{2/3} \\ \frac{dy}{dx} &= \frac{1}{15}x^{-2/3} - \frac{2}{3}x^{-1/3} \\ &= \frac{1}{15\sqrt[3]{x^2}} - \frac{2}{3\sqrt[3]{x}} \end{aligned}$$

g.
$$\begin{aligned} y &= 5x^{-3} + \frac{1}{5}x^3 \\ \frac{dy}{dx} &= -15x^{-4} + \frac{3}{5}x^2 \\ &= -\frac{15}{x^4} + \frac{3x^2}{5} \end{aligned}$$

h.
$$\begin{aligned} y &= x^{2/3} - 23 \\ \frac{dy}{dx} &= \frac{2}{3}x^{-1/3} \\ &= \frac{2}{3\sqrt[3]{x}} \end{aligned}$$

i.
$$\begin{aligned} \frac{dy}{dx} &= e^x - ex^{e-1} \\ y &= 4e^x + 2x^{-5/2} \end{aligned}$$

j.
$$\begin{aligned} \frac{dy}{dx} &= 4e^x - 5x^{-7/2} \\ &= 4e^x - \frac{5}{\sqrt{x^7}} \end{aligned}$$

Exercise 2.3.3

a.
$$\begin{aligned} f(x) &= \frac{x^2}{x} + \frac{1}{x} \\ &= x + x^{-1} \\ f'(x) &= 1 - x^{-2} \\ &= 1 - \frac{1}{x^2} \end{aligned}$$

b.
$$\begin{aligned} g(x) &= (3x+1)(3x+1) \\ &= 9x^2 + 6x + 1 \\ g'(x) &= 18x + 6 \end{aligned}$$

c.
$$\begin{aligned} p(x) &= \frac{(x+3)(x-3)}{x+3} \\ &= x + 3 && \text{for } x \neq -3 \\ p'(x) &= 1 && \text{for } x \neq -3 \end{aligned}$$

d.
$$\begin{aligned} r(h) &= \sqrt{h} + \sqrt{2} && \text{for } h \neq 2 \\ &= h^{1/2} + \sqrt{2} && \text{for } h \neq 2 \\ r'(h) &= \frac{1}{2}h^{-1/2} && \text{for } h \neq 2 \\ &= \frac{1}{2\sqrt{h}} && \text{for } h \neq 2 \end{aligned}$$

e.
$$\begin{aligned} \Gamma(x) &= x^3 - 5x^{1/2} \\ \Gamma'(x) &= 3x^2 - \frac{5}{2}x^{-1/2} \\ &= 3x^2 - \frac{5}{2\sqrt{x}} \end{aligned}$$

f.
$$\begin{aligned} w(t) &= \frac{te^t}{t} - \frac{1}{t} \\ &= e^t - t^{-1} \\ w'(t) &= e^t + t^{-2} \\ &= e^t + \frac{1}{t^2} \end{aligned}$$

g. $g(w) = 4w^5 + 4w^4 - 59w^3 - 30w^2 + 225w$
$g'(w) = 20w^4 + 16w^3 - 177w^2 - 60w + 225$

h.
$$\begin{aligned} s(t) &= \frac{6}{5}t^{7/6} + \frac{4}{5}t \\ s'(t) &= \frac{7}{5}t^{1/6} + \frac{4}{5} \end{aligned}$$

Exercise 2.3.4

a.
$$\begin{aligned} f'(x) &= 35x^4 - 8x^3 + 3x^2 - 14x + 14 \\ f''(x) &= 140x^3 - 24x^2 + 6x - 14 \end{aligned}$$

b.
$$\begin{aligned} f(x) &= 4e^x - 3x^{-5} + x \\ f'(x) &= 4e^x + 15x^{-6} + 1 \\ f''(x) &= 4e^x - 90x^{-7} \end{aligned}$$

Exercise 2.3.5

a.
$$\begin{aligned} f'(x) &= 8x - 3x^2 \\ f'(3) &= 24 - 27 \\ &= -3 \end{aligned}$$

So the tangent line has a slope of -3 and passes through the point $(3, f(3)) = (3, -9)$:

$$\begin{aligned} y &= -3(x-3) - 9 \\ &= -3x \end{aligned}$$

b.
$$\begin{aligned} f'(x) &= e^x - 4 \\ f'(0) &= 1 - 4 \\ &= -3 \end{aligned}$$

So the tangent line has a slope of -3 and passes through the point $(0, f(0)) = (0, 0)$:

$$\begin{aligned} y &= -3(x-0) + 0 \\ &= -3x \end{aligned}$$

c.
$$\begin{aligned} f'(x) &= 2x - 1 \\ f'(-2) &= -4 - 1 \\ &= -5 \end{aligned}$$

So the tangent line has a slope of -5 and passes through the point $(-2, f(-2)) = (-2, 7)$:

$$\begin{aligned} y &= -5(x+2) + 7 \\ &= -5x - 3 \end{aligned}$$

d.
$$\begin{aligned} f'(x) &= \frac{2}{\sqrt{x}} - 1 \\ f'(9) &= \frac{2}{3} - 1 \\ &= -\frac{1}{3} \end{aligned}$$

So the tangent line has a slope of -3 and passes through the point $(0, f(0)) = (0, 0)$:

$$\begin{aligned} y &= -3(x-0) + 0 \\ &= -3x \end{aligned}$$

Exercise 2.3.6

a. $f'(x) = 4e^x - 5$

b.
$$\begin{aligned} f'(x) &= 4e^x - 5 \\ 6 &= 4e^x - 5 \end{aligned}$$

c.
$$\begin{aligned} \frac{11}{4} &= e^x \\ x &= \ln\left(\frac{11}{4}\right) \\ &= \ln(11) - \ln(4) \end{aligned}$$

d.
$$\begin{aligned} f'(x) &= 4e^x - 5 \\ 0 &= 4e^x - 5 \\ \frac{5}{4} &= e^x \\ x &= \ln\left(\frac{5}{4}\right) \\ &= \ln(5) - \ln(4) \end{aligned}$$

Exercise 2.3.7

a.
$$f'(x) = \frac{1}{3\sqrt[3]{x^2}} + 1$$

b.
$$\begin{aligned} f'(x) &= \frac{1}{3\sqrt[3]{x^2}} + 1 \\ 2 &= \frac{1}{3\sqrt[3]{x^2}} + 1 \\ 1 &= \sqrt[3]{x^2} \\ 1 &= x \end{aligned}$$

c. We are interested in finding any values of x that will make $f'(x)$ not exist, and specifically ones that will result in the derivative approaching $\pm\infty$.

$$f'(x) = \frac{1}{3\sqrt[3]{x^2}} + 1$$

Note that when $x = 0$, we end up with division by 0.

Exercise 2.3.8

This page titled 2.3: Some Early Derivative Rules is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Peter Keep.

2.4: The Product and Quotient Rules

We saw in Theorem 2.3.10 Combinations of Derivatives that when we want to find the derivative of a sum or difference of functions, we can just find the derivatives of each function separately, and then combine the derivatives back together (by adding or subtracting). This, hopefully, is pretty intuitive: of course a slope of a sum of things is just the slopes of each thing added together!

In this section, we want to think about derivatives of product and quotients of functions. What happens when we differentiate a function that is really just two functions multiplied together?

Activity 2.4.1: Thinking About Derivatives of Products.

Let's start with two quick facts:

$$\frac{d}{dx}\left(x^3\right) = 3x^2 \text{ and } \frac{d}{dx}(\sin(x)) = \cos(x)$$

a. What is $\frac{d}{dx}\left(x^3 + \sin(x)\right)$? What about $\frac{d}{dx}\left(x^3 - \sin(x)\right)$?

b. Based on what you just explained, what is a reasonable assumption about what $\frac{d}{dx}\left(x^3 \sin(x)\right)$ might be?

c. Let's just think about $\frac{d}{dx}\left(x^3\right)$ for a moment. What is x^3? Can you write this as a product? Call one of your functions $f(x)$ and the other $g(x)$. You should have that $x^3 = f(x)g(x)$ for whatever you used.

d. Use your example to explain why, in general, $\frac{d}{dx}(f(x)g(x)) \neq \frac{d}{dx}(f(x)) \cdot \frac{d}{dx}(g(x))$.

e. Let's show another way that we know this. Think about $\sin(x)$. We know two things:

$$\sin(x) = (1)(\sin(x)) \text{ and } \frac{d}{dx}(\sin(x)) = \cos(x)$$

What, though, is $\frac{d}{dx}(1) \cdot \frac{d}{dx}(\sin(x))$?

f. Use all of this to reassure yourself that even though the derivative of a sum of functions is the sum of the derivatives of the functions, we will need to develop a better understanding of how the derivatives of products of functions work.

I like to think about it this way: if $\frac{d}{dx}(f(x)g(x)) = f'(x)g'(x)$, then every function's derivative would be 0.

In the example with the $\sin(x)$ function, we noticed that we could write the function as $(1)(\sin(x))$. This is true of every function!

If $\frac{d}{dx}(f(x)g(x)) = f'(x)g'(x)$, then we could say that for any function $f(x)$ with a derivative $f'(x)$:

$$\begin{aligned}\frac{d}{dx}(f(x)) &= \frac{d}{dx}(1 \cdot f(x)) \\ &= \frac{d}{dx}(1)\frac{d}{dx}(f(x)) \\ &= 0 \cdot f'(x) \\ &= 0\end{aligned}$$

This, obviously, can't be true! We know of tons of functions that have non-zero slopes...most of them do!

So, we hopefully have some motivation for building a rule to that helps us think about derivatives of products of functions.

The Product Rule

Activity 2.4.2: Building a Product Rule.

Let's investigate how we might differentiate the product of two functions:

$$\frac{d}{dx}(f(x)g(x))$$

. We'll use an area model for multiplication here: we can consider a rectangle where the side lengths are functions $f(x)$ and $g(x)$ that change for different values of x. Maybe x is representative of some type of time component, and the side lengths change size based on time, but it could be anything.

If we want to think about $\frac{d}{dx}(f(x)g(x))$, then we're really considering the change in area of the rectangle.

a. Find the area of the two rectangles. The second rectangle is just one where the input variable for the side length has changed by some amount, leading to a different side length.

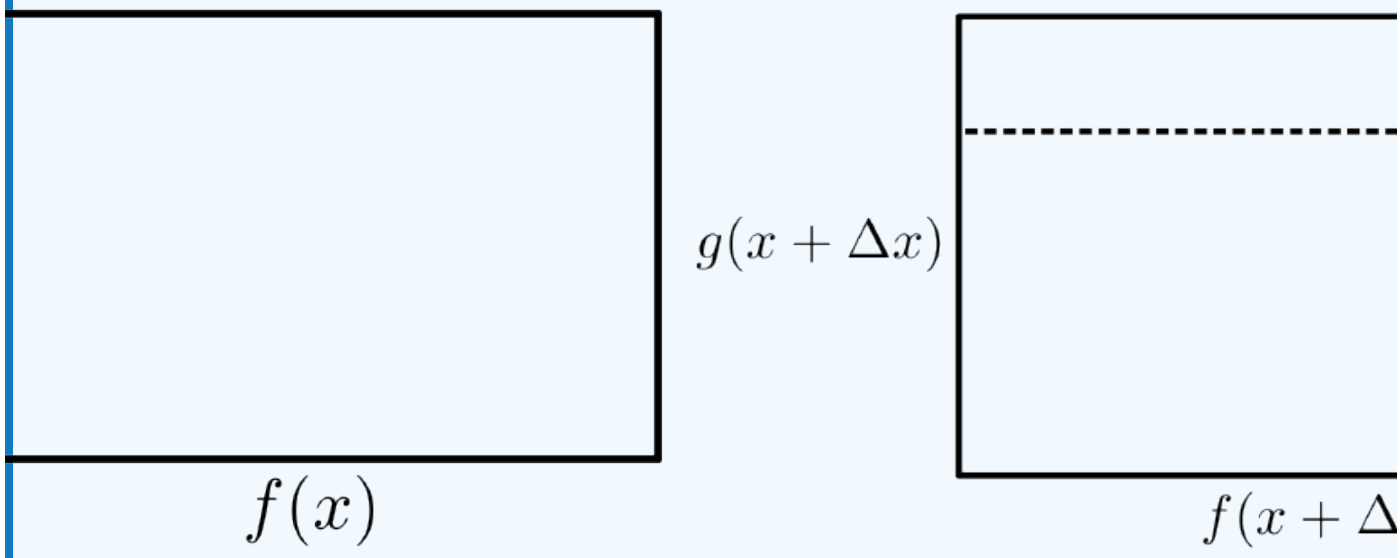


Figure 2.4.3

b. Write out a way of calculating the difference in these areas.
c. Let's try to calculate this difference in a second way: by finding the actual area of the region that is new in the second rectangle.

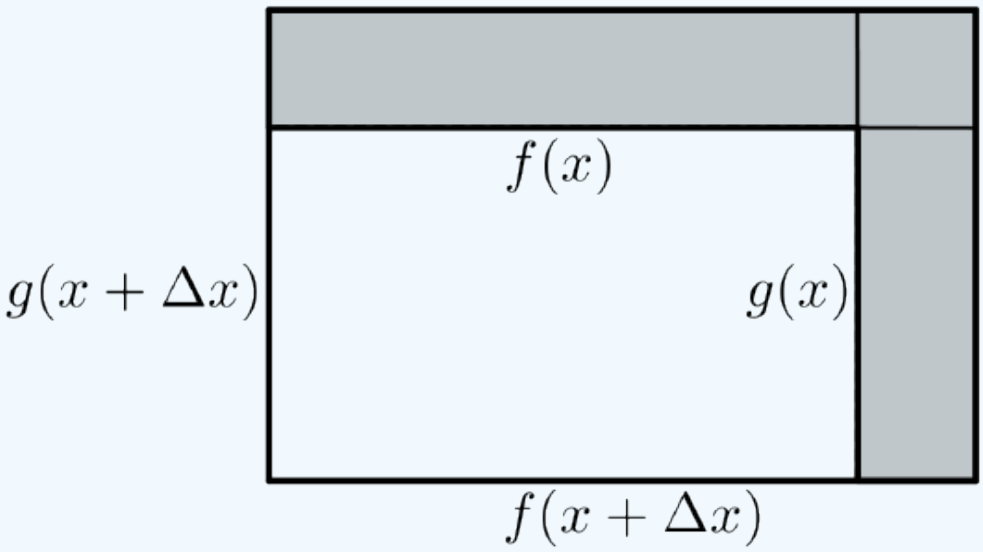


Figure 2.4.4

In order to do this, we've broken the region up into three pieces. Calculate the areas of the three pieces. Use this to fill in the following equation:

$$f(x+\Delta x)g(x+\Delta x) - f(x)g(x) =$$

d. We do not want to calculate the total change in area: a derivative is a rate of change, so in order to think about the derivative we need to divide by the change in input, Δx. We'll also want to think about this derivative as an instantaneous rate of change, meaning we will consider a limit as $\Delta x \to 0$. Fill in the following:

$$\frac{d}{dx}(f(x)g(x)) \lim_{\Delta x \to 0} \left(\frac{f(x+\Delta x)g(x+\Delta x) - f(x)g(x)}{\Delta x} \right)$$
$$= \lim_{\Delta x \to 0} \left(\frac{}{\Delta x} \right)$$

We can split this fraction up into three fractions:

$$\frac{d}{dx}(f(x)g(x)) = \lim_{\Delta x \to 0} \left(\frac{}{\Delta x} \right)$$
$$+ \lim_{\Delta x \to 0} \left(\frac{}{\Delta x} \right)$$
$$+ \lim_{\Delta x \to 0} \left(\frac{}{\Delta x} \right)$$

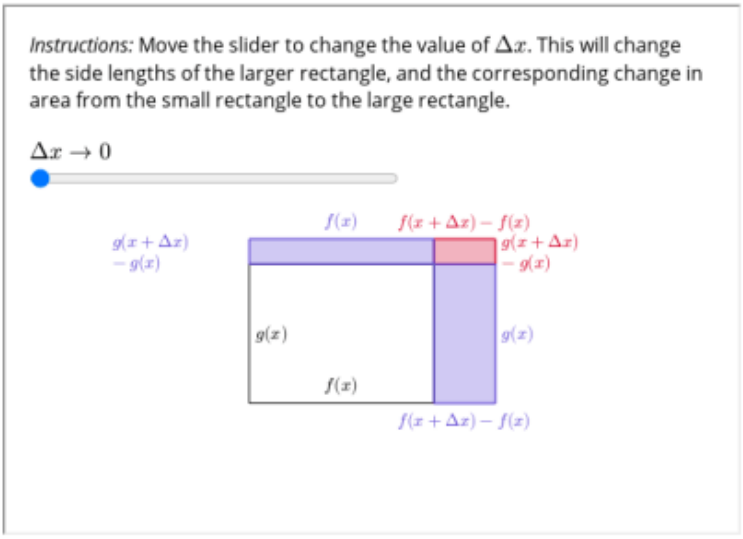
e. In any limit with $f(x)$ or $g(x)$ in it, notice that we can factor part out of the limit, since these functions do not rely on Δx, the part that changes in the limit. Factor these out. In the third limit, factor out either $\lim_{\Delta x \to 0}(f(x+\Delta x) - f(x))$ or $\lim_{\Delta x \to 0}(g(x+\Delta x) - g(x))$.

$$\frac{d}{dx}(f(x)g(x)) = f(x) \lim_{\Delta x \to 0} \left(\frac{}{\Delta x} \right)$$
$$+g(x) \lim_{\Delta x \to 0} \left(\frac{}{\Delta x} \right) \Big)$$
$$+ \lim_{\Delta x \to 0} () \left(\lim_{\Delta x \to 0} \left(\frac{}{\Delta x} \right) \right)$$

f. Use Definition 2.1.2 The Derivative Function to rewrite these limits. Show that when $\Delta x \to 0$, we get:

$$f(x)g'(x) + g(x)f'(x) + 0$$

We can investigate this visual a bit more dynamically: see the differences in area as $\Delta x \to 0$. What parts are left, when $\Delta x \to 0$? What areas aren't visible?



Standalone Embed

Theorem 2.4.5: Product Rule.

If $u(x)$ and $v(x)$ are functions that are differentiable at x and $f(x) = u(x) \cdot v(x)$, then:

$$\frac{d}{dx}(f(x)) = u'(x) \cdot v(x) + u(x) \cdot v'(x)$$

For convenience, this often is written as:

$$\frac{d}{dx}(u \cdot v) = u'v + uv' \quad \text{for} \quad \frac{d}{dx}(u \cdot v) = v\left(\frac{du}{dx}\right) + u\left(\frac{dv}{dx}\right)$$

Example 2.4.6

Use the Product Rule to find the following derivatives.

a. $\frac{d}{dx}\left(x^3 \sin(x)\right)$ **Hint.** Use $u = x^3$ and $v = \sin(x)$. Now find u' and v', and use:

$$\frac{d}{dx}(uv) = u'v + uv'$$

b. $\frac{d}{dx}\left(\left(x^3 + 4x\right)e^x\right)$

c. $\frac{d}{dx}(\sqrt{x}\cos(x))$

What about Dividing?

Since we can differentiate a product of functions, then the obvious next question should be about division: if we needed to build a reasonable way of differentiating a product, shouldn't we also need to build a new way of thinking about derivatives of quotients?

A nice thing that we can do is think about division as really just multiplication. For instance, if we want to differentiate $\frac{d}{dx}\left(\frac{\sin(x)}{x^2}\right)$, then we can just think about this quotient as really a product:

$$\frac{d}{dx}\left(\frac{\sin(x)}{x^2}\right) = \frac{d}{dx}\left(\frac{1}{x^2}(\sin(x))\right)$$

Now we can just apply the Product Rule!

$$\frac{d}{dx}\left(\frac{1}{x^2}(\sin(x))\right) = \frac{d}{dx}\left(x^{-2}\sin(x)\right)$$
$$u = \sin(x) \quad v = x^{-2}$$
$$u' = \cos(x) \quad v' = -2x^{-3}$$
$$\frac{d}{dx}\left(\sin(x)x^{-2}\right) = x^{-2}\cos(x) + \left(-2x^{-3}\sin(x)\right)$$
$$= \frac{\cos(x)}{x^2} - \frac{2\sin(x)}{x^3}$$

This works great! We can always think about quotients as just products of reciprocals! But the problem is that we can't always differentiate these reciprocals (for now). We were able to differentiate $\frac{1}{x^2}$ by rewriting this as just a power function (with a negative exponent).

What about this flipped example:

$$\frac{d}{dx}\left(\frac{x^2}{\sin(x)}\right)?$$

In order for us to do the same thing, we need to rewrite this as

$$\frac{d}{dx}\left(x^2(\sin(x))^{-1}\right)$$

but we don't know (or maybe don't know yet) how to differentiate $(\sin(x))^{-1}$.

So let's try to build a general way of differentiating quotients.

Activity 2.4.7: Constructing a Quotient Rule.

We're going to start with a function that is a quotient of two other functions:

$$f(x) = \frac{u(x)}{v(x)}$$

Our goal is that we want to find $f'(x)$, but we're going to shuffle this function around first. We won't calculate this derivative directly!

a. Start with $f(x) = \frac{u(x)}{v(x)}$. Multiply $v(x)$ on both sides to write a definition for $u(x)$.

$$u(x) =$$

b. Find $u'(x)$.

c. Wait: we don't care about $u'(x)$, right? We care about finding $f'(x)$! Use what you found for $u'(x)$ and solve for $f'(x)$.

$$f'(x) =$$

d. This is a strange formula: we have a formula for $f'(x)$ written in terms of $f(x)$! But we said earlier that $f(x) = \frac{u(x)}{v(x)}$. In your formula for f$f'(x)$, replace $f(x)$ with $\frac{u(x)}{v(x)}$.

$$f'(x) =$$

This formula is fine, but a little clunky. We'll try to rewrite it in a friendlier way, but it is a bit more complex than the Product Rule.

Theorem 2.4.8: Quotient Rule.

If $u(x)$ and $v(x)$ are differentiable at x and $f(x) = \frac{u(x)}{v(x)}$ then:

$$f'(x) = \frac{v(x)u'(x) - u(x)v'(x)}{(v(x))^2}$$

For convenience, this often is written as:

$$\frac{d}{dx}\left(\frac{u}{v}\right)=\frac{vu'-uv'}{v^2}$$

Example 2.4.9

Use the Quotient Rule to find the following derivatives.

a. $\frac{d}{dx}\left(\frac{\sin(x)}{x^2}\right)$ Once you have this derivative, confirm that it is the same as $\frac{\cos(x)}{x^2}-\frac{2\sin(x)}{x^3}$, the way that we found it using the Product Rule.

b. $\frac{d}{dx}\left(\frac{x^2}{\sin(x)}\right)$

c. $\frac{d}{dx}\left(\frac{x+4}{x^2+1}\right)$

Derivatives of (the Rest of the) Trigonometric Functions

You might remember that of the six main trigonometric functions, we can write four of them in terms of the other two: here are the different trigonometric functions written in terms of sine and cosine functions:

$$\tan(x)=\left(\frac{\sin(x)}{\cos(x)}\right)$$

$$\sec(x)=\left(\frac{1}{\cos(x)}\right)$$

$$\cot(x)=\left(\frac{\cos(x)}{\sin(x)}\right)$$

$$\csc(x)=\left(\frac{1}{\sin(x)}\right)$$

Example 2.4.10

Find the derivatives of the remaining trigonometric functions.

a. $\frac{d}{dx}(\tan(x))$ **Hint**. Write $\frac{d}{dx}(\tan(x))=\frac{d}{dx}\left(\frac{\sin(x)}{\cos(x)}\right)$ and use the Quotient Rule.

b. $\frac{d}{dx}(\sec(x))$ **Hint**. Write $\frac{d}{dx}(\sec(x))=\frac{d}{dx}\left(\frac{1}{\cos(x)}\right)$ and use the Quotient Rule.

c. $\frac{d}{dx}(\cot(x))$ **Hint**. Write $\frac{d}{dx}(\cot(x))=\frac{d}{dx}\left(\frac{\cos(x)}{\sin(x)}\right)$ and use the Quotient Rule.

d. $\frac{d}{dx}(\csc(x))$ **Hint**. Write $\frac{d}{dx}(\csc(x))=\frac{d}{dx}\left(\frac{1}{\cos(x)}\right)$ and use the Quotient Rule.

Practice Problems

Exercise 2.4.1

A student encounters the problem $\frac{d}{dx}\left(e^x\left(x^2+1\right)\right)$ and notes that since $\frac{d}{dx}\left(e^x\right)=e^x$ and $\frac{d}{dx}\left(x^2+1\right)=2x$, then $\frac{d}{dx}\left(e^x\left(x^2+1\right)\right)=2xe^x$. Explain not only why the student is incorrect, but also what kinds of pitfalls can occur if we use this "the derivative of a product is the product of the derivatives" method. Differentiate correctly, explaining your process for the student's benefit.

Exercise 2.4.2

Consider the following table of values of f, g, f', and g'.

Table 2.4.11.

x	$f(x)$	$f'(x)$	$g(x)$	$g'(x)$
0	-3	$\frac{1}{2}$	2	-4
2	$\frac{1}{7}$	-1	$-\frac{5}{2}$	$\frac{7}{3}$

Find the following derivatives, using the table.

a. $\left.\frac{d}{dx}\left(\frac{g(x)}{f(x)}\right)\right|_{x=2}$
b. $\left.\frac{d}{dx}(f(x)g(x))\right|_{x=0}$
c. $\left.\frac{d}{dx}\left(\frac{f(x)+1}{g(x)-2}\right)\right|_{x=0}$
d. $\left.\frac{d}{dx}(f(x)g(x))\right|_{x=2}$

Answer available at the end of this section.

Exercise 2.4.3

Use the product and quotient rules to differentiate each of the following.

a. $f(x)=(4x+4)\left(7x^2-6\right)$
b. $g(x)=\frac{x^2+1}{e^x}$
c. $j(x)=\frac{}{x-1}$
d. $s(t)=e^t\left(t^2+t-1\right)$
e. $\Gamma(\theta)=4e^\theta\sqrt{\theta}$
f. $r(t)=e^{2t}$
g. $w(x)=\left(5x^5-4x^2+5x-1\right)^2$
h. $p(\alpha)=\frac{\alpha^3+1}{\alpha^2-1}$
i. $f(x)=\frac{x+e}{x^3-4x}$
j. $g(x)=\frac{x+4}{e^x}$

Answer available at the end of this section.

Exercise 2.4.4

Differentiate the following.

a. $y = e^x(x+4)^2(4x-3)$
b. $y = \dfrac{xe^x+1}{x^2-1}$
c. $y = \dfrac{x^2+4x-1}{xe^x}$
d. $y = \left(\dfrac{x^2-4}{x+5}\right)(e^x(x+3))$

Answer available at the end of this section.

Exercise 2.4.5

Find the equation of the line tangent to the function f(x) at the given point.

a. $f(x) = \dfrac{x^2-5}{x^2+1}$ at (1, –2)
b. $f(x) = \dfrac{e^x}{x+1}$ at (0, 1)
c. $f(x) = e^x(x^2+6)$ at (1, 7e)
d. $f(x) = \dfrac{xe^x}{x+2}$ at (0, 0)

Answer available at the end of this section.

Answers

Exercise 2.4.2

a.
$$\frac{d}{dx}\left(\frac{g(x)}{f(x)}\right) = \frac{f(x)g'(x)-g(x)f'(x)}{(f(x))^2}$$
$$\left.\frac{d}{dx}\left(\frac{g(x)}{f(x)}\right)\right|_{x=2} = \frac{f(2)g'(2)-g(2)f'(2)}{(f(2))^2}$$
$$= \frac{\left(\frac{1}{7}\right)\left(\frac{7}{3}\right)-\left(-\frac{5}{2}\right)(-1)}{\left(\frac{1}{7}\right)^2}$$
$$= -\frac{637}{6}$$

b.
$$\frac{d}{dx}(f(x)g(x)) = f(x)g'(x)+f'(x)g(x)$$
$$\left.\frac{d}{dx}(f(x)g(x))\right|_{x=0} = f(0)g'(0)+f'(0)g(0)$$
$$= (-3)(-4)+\left(\frac{1}{2}\right)(2)$$
$$= 10$$

c.
$$\frac{d}{dx}\left(\frac{f(x)+1}{g(x)-2}\right) = \frac{(g(x)-2)(f'(x))-(f(x)+1)(g'(x))}{(g(x)-2)^2}$$
$$\left.\frac{d}{dx}\left(\frac{f(x)+1}{g(x)-2}\right)\right|_{x=0} = \frac{(g(0)-2)(f'(0))-(f(0)+1)(g'(0))}{(g(0)-2)^2}$$
$$= \frac{(0)\left(\frac{1}{2}\right)-(-2)(-4)}{0}$$

This derivative doesn't exist, since we have division by 0. We can notice that the actual quotient doesn't exist: there's a vertical asymptote on this function at $x = 0$.

d.
$$\frac{d}{dx}(f(x)g(x)) = f(x)g'(x)+f'(x)g(x)$$
$$\left.\frac{d}{dx}(f(x)g(x))\right|_{x=2} = f(2)g'(2)+f'(2)g(2)$$
$$= \left(\frac{1}{7}\right)\left(\frac{7}{3}\right)+(-1)\left(-\frac{5}{2}\right)$$
$$= \frac{17}{6}$$

Exercise 2.4.3

a. $f'(x) = (4)\left(7x^2 - 6\right) + (14x)(4x + 4)$
$= 84x^2 + 56x - 24$

b. $g'(x) = \dfrac{(2x)(x) - (1)\left(x^2 + 1\right)}{x^2}$
$= \dfrac{x^2 - 1}{x^2}$

c. $j'(x) = \dfrac{\left(e^x\right)(x - 1) - (1)\left(e^x\right)}{(x - 1)^2}$
$= \dfrac{xe^x - 2e^x}{(x - 1)^2} = \dfrac{e^x(x - 2)}{(x - 1)^2}$

d. $s'(t) = \left(e^t\right)(2t + 1) + \left(e^t\right)\left(t^2 + t - 1\right)$
$= t^2e^t + 3te^t = e^t\left(t^2 + 3t\right)$

e. $\Gamma'(\theta) = \left(4e^\theta\right)\left(\dfrac{1}{2\sqrt{\theta}}\right) + \left(4e^\theta\right)\left(\sqrt{\theta}\right)$
$= 4e^\theta\left(\dfrac{1}{2\sqrt{\theta}} + \sqrt{\theta}\right)$

f. Hint. This function can be written as $r(t) = \left(e^t\right)^2$

$$r(t) = \left(e^t\right)\left(e^t\right)$$
$$r'(t) = \left(e^t\right)\left(e^t\right) + \left(e^t\right)\left(e^t\right)$$
$$= 2\left(e^t\right)^2 = 2e^{2t}$$

g. $w(x) = \left(5x^5 - 4x^2 + 5x - 1\right)\left(5x^5 - 4x^2 + 5x - 1\right)$
$w'(x) = \left(25x^4 - 8x + 5\right)\left(5x^5 - 4x^2 + 5x - 1\right) + \left(25x^4 - 8x + 5\right)\left(5x^5 - 4x^2 + 5x - 1\right)$
$= 2\left(25x^4 - 8x + 5\right)\left(5x^5 - 4x^2 + 5x - 1\right)$
$= 250x^9 - 280x^6 + 300x^5 - 50x^4 + 64x^3 - 120x^2 + 66x - 10$

h. $p'(\alpha) = \dfrac{3\alpha^2\left(\alpha^2 - 1\right) - 2\alpha\left(\alpha^3 + 1\right)}{\left(\alpha^2 - 1\right)^2}$
$= \dfrac{\alpha^4 - 3\alpha^2 - 2\alpha}{\left(\alpha^2 - 1\right)^2}$

i. $f'(x) = \dfrac{\left(3x^2 - 4\right)(x + e) - \left(x^3 - 4x\right)}{\left(x^3 - 4x\right)^2}$
$= \dfrac{2x^3 + 3ex^2 - 4e}{\left(x^3 - 4x\right)^2}$

j. $g'(x) = \dfrac{e^x - e^x(x + 4)}{\left(e^x\right)^2}$
$= \dfrac{-xe^x - 5e^x}{\left(e^x\right)^2}$
$= -\dfrac{x + 5}{e^x}$

Exercise 2.4.4

a. $\dfrac{dy}{dx} = e^x\left(4x^3 + 41x^2 + 98x - 8\right)$

b. $\dfrac{dy}{dx} = \dfrac{e^x\left(x^3 - x^2 - x - 1\right) - 2x}{\left(x^2 - 1\right)^2}$

c. $\dfrac{dy}{dx} = \dfrac{-x^3 - 3x^2 + x + 1}{x^2e^x}$

d. $\dfrac{dy}{dx} = \dfrac{e^x\left(x^4 + 10x^3 + 29x^2 - 2x - 68\right)}{(x + 5)^2}$

Exercise 2.4.5

a. $f'(x) = \dfrac{2x\left(x^2 + 1\right) - 2x\left(x^2 - 5\right)}{\left(x^2 + 1\right)}$
$f'(1) = \dfrac{2(2) - 2(-4)}{(2)^2}$
$= 3$
So the equation of the tangent line is:
$y = 3(x - 1) - 2$
$= 3x - 5$

b. $f'(x) = \dfrac{e^x(x + 1) - e^x}{(x + 1)^2}$
$f'(0) = \dfrac{1 - 1}{(1)^2}$
$= 0$
So the equation of the tangent line is:
$y = 0(x - 0) + 1$
$= 1$

c. $f'(x) = e^x\left(x^2 + 6\right) + 2xe^x$
$f'(1) = 7e + 2e$
$= 9e$
So the equation of the tangent line is:
$y = 9e(x - 1) + 7e$
$= 9ex - 2e$

$$f'(x) = \frac{(xe^x + e^x)(x+2) - xe^x}{(x+2)^2}$$

$$f'(0) = \frac{(1)(2)}{(2)^2}$$

d. $$= \frac{1}{2}$$

So the equation of the tangent line is:

$$y = \frac{1}{2}(x-0) + 0$$

$$= \frac{x}{2}$$

This page titled 2.4: The Product and Quotient Rules is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Peter Keep.

2.5: The Chain Rule

We've been building up some intuition and rules to help us think about differentiating different functions and combinations of functions. We can find derivatives of scaled functions, sums of functions, differences of functions, products of functions, and also quotients of functions.

In this section, we'll look at our last operation between functions: composition.

Composition and Decomposition

An important part of finding derivatives of products and quotients is identifying the component functions that are being multiplied/divided (often labeled $u(x)$ or just u and $v(x)$ or just v). From there, we find the derivatives of each of the component functions, and then use the formula from the Product Rule or Quotient Rule to put the pieces together. Thinking about derivatives of composed functions will be the same: we'll need to identify what functions are being composed inside of other functions, and use those pieces in some formulaic way to represent the derivative. On that note, let's remind ourselves and practice working with composition (and decomposition) of functions.

Activity 2.5.1: Composition (and Decomposition) Pictionary.

This activity will involve a second group, or at least a partner. We'll go through the first part of this activity, and then connect with a second group/person to finish the second part.

a. Build two functions, calling them $f(x)$ and $g(x)$. Pick whatever kinds of functions you'd like, but this activity will work best if these functions are in a kind of sweet spot between "simple" and "complicated," but don't overthink this.
b. Compose $g(x)$ inside of $f(x)$ to create $(f \circ g)(x)$, which we can also write as $f(g(x))$.
c. Write your composed $f(g(x))$ function on a separate sheet of paper. Do not leave any indication of what your chosen $f(x)$ and $g(x)$ are. Just write your composed function by itself. Now, pass this composed $f(g(x))$ to your partner/a second group.
d. You should have received a new function from some other person/ group. It is different than yours, but also labeled $f(g(x))$ (with different choices of $f(x)$ and $g(x)$). Identify a possibility for $f(x)$, the outside function in this composition, as well as a possibility for $g(x)$, the inside function in this composition. You can check your answer by composing these!
e. Write a different pair of possibilities for $f(x)$ and $g(x)$ that still will give you the same composed function, $f(g(x))$.
f. Check with your partner/the second group: did you identify the pair of functions that they originally used? Did whoever you passed your composed function to correctly identify your functions?

A big thing to notice here is that when we pick the pieces of functions that we think were composed inside of each other, there's not a single obvious answer. This is pretty different compared to, say, using the Quotient Rule. In these quotients, we have a natural division (literally) between the pieces. Here, it's much more subjective for us when we decide to label an "inside" function and an "outside" function.

We will build up our intuition to find a good balance for how we pick these.

The Chain Rule, Intuitively

Before we build the Chain Rule for differentiating composed functions, we should talk about some notation. Earlier (in Notation for Derivatives) we talked about the derivative notation, $\frac{dy}{dx}$. One of the things we mentioned is that while we know that the derivative is an instantaneous rate of change, this notation is helpful to tell us what is changing with regard to what.

Aside.

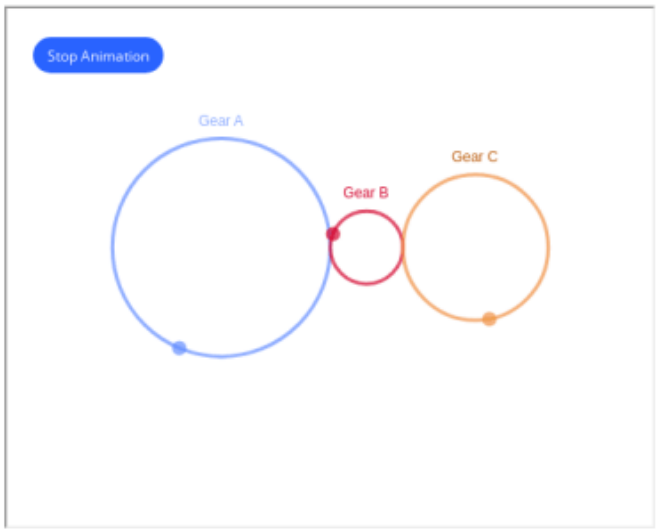
This notation, $\frac{dy}{dx}$, is called **Leibniz's Notation**, named after Gottfried Wilhelm Leibniz. Leibniz used dy and dx to represent infinitely small changes or increments in the y and x variables. This was meant to be a nod to Δy and Δx being used as finite changes in y and x.

While the general idea of calculus using these infinitely small changes hasn't persisted (we now use the concept of a limit to formalize the logic), we still use the notation where we can as a nod to Leibniz's contributions to the development of calculus and analysis as a field of mathematics.

In $\frac{dy}{dx}$, we are calculating how much the y-variable changes when x increases. If we talked about $\frac{df}{dt}$, then we are discussing how much f changes for an increase in t, whatever these variables represent.

Activity 2.5.2: Gears and Chains.

Let's think about some gears. We've got three gears, all different sizes. But the gears are linked together, and a small motor works to spin one of the gears. Since the gears are linked, when one gear spins, they all spin. But since they are different sizes, they complete a different number of revolutions: the smaller ones spin more times than the larger ones, since they have a smaller circumference.



Standalone Embed

For our purpose, let's say that Gear A is being driven by the motor.

a. Let's try to quantify how much "faster" Gear B is spinning compared to Gear A. How many revolutions does Gear B complete in the time it takes Gear A to complete one revolution?
b. Now quantify the speed of Gear C compared to its neighbor, Gear B. How many revolutions does Gear C complete in the time it takes Gear B to complete one revolution?
c. Use the above relative "speeds" to compare Gear C and Gear A: how many revolutions does Gear C complete in the time it takes Gear A to complete one revolution? More importantly, how do we find this?
d. Now, let's translate this into some derivative notation: we've really been finding rates at which one thing changes (the speed of the gear spinning) relative to another's. Call the speed of Gear B compared to Gear A: $\frac{dB}{dA}$. Now, call the speed of Gear C compared to Gear B: $\frac{dC}{dB}$. Come up with a formula to find $\frac{dC}{dA}$.

So what we need to do now is somehow translate this intuitive idea of multiplying rates of change to build a strategy for thinking about derivatives of composed functions.

We can think of these linked gears as functions: Gear C changes based on what is happening with Gear B, which changes based on Gear A. We can translate Gear A to be an input variable, like x. Then Gear B is a function based on that: we can call it $g(x)$. Then Gear C is a function that takes in the position of Gear B (the function $g(x)$), and so we can think of it as $f(g(x))$.

To build the derivative rule for composite functions, we need to find how the "outside" function changes as the "inside" function changes ($\frac{dC}{dB}$ in this case) and multiply that by how the "inside" function changes as the input variable changes ($\frac{dB}{dA}$ here).

Theorem 2.5.1: The Chain Rule.

For the composite function $y = f(g(x))$, if we define $u = g(x)$ and $y = f(u)$, then, as long as both f and g are differentiable at u and x respectively:

$$\frac{d}{dx}(f(g(x))) = \frac{d}{du}(f(u)) \cdot \frac{d}{dx}(g(x))$$

Alternatively, this can be written as:

$$\frac{dy}{dx} = \frac{dy}{du} \cdot \frac{du}{dx} \quad \text{or} \quad \frac{d}{dx}(f(g(x))) = f'(g(x)) \cdot g'(x)$$

Doing is Different than Knowing

It is lovely to know that the Chain Rule is really just linking the two rates of change together to connect a function with an input variable through a middle processing function. That's great!

But doing the Chain Rule is different than just knowing it, so let's walk through a first example. Let's find the following derivative:

$$\frac{d}{dx}\left(\sin\left(x^2\right)\right)$$

We'll call the "inside" function $u = x^2$, so we can really write the whole function (normally, we're calling this y) as $y = \sin(u)$.

$$\begin{aligned}\frac{d}{dx}(\underbrace{\sin(\overbrace{x^2}^{u})}_{y}) &= \frac{dy}{du} \cdot \frac{du}{dx} \\ &= \frac{d}{du}(\sin(u)) \cdot \frac{d}{dx}\left(x^2\right)\end{aligned}$$

What we can notice, here, is that $\sin(u)$ is just a function of some variable u, and we want to find $\frac{dy}{du}$, the rate at which $y = \sin(u)$ changes with regard to its input variable. This might feel a bit strange, since u isn't just an input variable: it means something, since we have that $u = x^2$. This is fine! The extra $\frac{du}{dx}$ that we multiply will take care of linking this derivative to the input variable x.

$$\begin{aligned}\frac{d}{dx}(\underbrace{\sin(\overbrace{x^2}^{u})}_{y}) &= \frac{d}{du}(\sin(u)) \cdot \frac{d}{dx}\left(x^2\right) \\ &= \cos(u) \cdot 2x \\ &= \cos\left(x^2\right) \cdot 2x \\ &= 2x\cos\left(x^2\right)\end{aligned}$$

After we finished differentiating $\frac{d}{du}(\sin(u))$, you'll notice that we used the fact that $u = x^2$ to write our combination of derivatives (the derivative of the "outside" function and the derivative of the "inside" function) in terms of the same input variable again.

The last line, rewriting $\cos(x^2) \cdot 2x$ as $2x\cos(x^2)$, is just for aesthetics.

Now you're ready to try some more examples! In each, focus on identifying a natural selection for the "inside" function, u.

Example 2.5.2

Use the Chain Rule to differentiate the following:

a. $\frac{d}{dx}\left(\sqrt{x^2+4}\right)$

Hint

Notice that x^2+4 is composed under the square root. Use $u = x^2+4$.

b. $\frac{d}{dx}\left(e^{\tan(x)}\right)$

Hint

Try letting $u = \tan(x)$, since it's composed inside the exponent of the exponential function.

c. $\frac{d}{dx}\left(\sin^5(x)\right)$

Hint

You could think about this as $\frac{d}{dx}(\sin(x)\sin(x)\sin(x)\sin(x)\sin(x))$ and try to use a very annoying product rule, but it might be easier to think about this as $\frac{d}{dx}\left((\sin(x))^5\right)$.

Generalizing the Derivative of the Exponential

Earlier, we looked at the specific derivative for $f(x) = e^x$ (Theorem 2.3.9), but we haven't talked about derivatives of other exponential functions. What about things like $y = 2^x$ or $y = \left(\frac{1}{2}\right)^x$? We can use a nice fact about exponentials and logarithms. We'll think more about log functions later (starting in Section 3.2), but we can think a bit about them now.

A big fact to recall is that a logarithm is a way of finding an exponent with a specific property. If we want to find the exponent that we would need to put on the number e to give us 9 as an answer, we could use $\ln(9)$.

$$e^{\ln(9)} = 9$$

This is just because logs are defined in this circular way: they are, by definition, the exponent you would need to output whatever number is inside the log.

This means that if we want to think about the number 2, written in a different way, we can think of $e^{\ln(2)}$. Ok, but why would we ever use this? This seems like a ridiculous way to write a number as basic as 2!

Consider the following:

$$2^x = (\underbrace{e^{\ln(2)}}_{=2})^x$$

But we also might notice that we can rewrite this using an exponent rule! We know that in general: $(a^b)^c = a^{b\cdot c}$. Let's rewrite this exponential function:

$$\begin{aligned} 2^x &= \left(e^{\ln(2)}\right)^x \\ &= e^{\ln(2)\cdot x} \end{aligned}$$

Remember, $\ln(2)$ is just a number: it's specifically the number you have to put in the exponent on e to get 2. So this is just a coefficient on x. We can differentiate and use the Chain Rule!

$$\begin{aligned}\frac{d}{dx}\left(2^x\right) &= \frac{d}{dx}\left(e^{\ln(2)\cdot x}\right)\\ &= e^{\ln(2)\cdot x}\cdot \ln(2)\end{aligned}$$

Now we can remember that $e^{\ln(2)\cdot x}$ is really $\left(e^{\ln(2)}\right)^x$ which is just 2^x.

So we get $\frac{d}{dx}\left(2^x\right) = 2^x \ln(2)$. We can notice that we can recreate this with any (reasonable) value for the base of this exponential function.

Theorem 2.5.3: Derivative of the Generalized Exponential Function.

If $b > 0$ and $b \neq 1$, then:

$$\frac{d}{dx}\left(b^x\right) = b^x \ln(b)$$

Exercise 2.5.1

For functions $f(x)$ and $g(x)$, explain how to use the Chain Rule to find $\frac{d}{dx}(f(g(x)))$.

Exercise 2.5.2

For functions $f(x)$ and $g(x)$, explain how to use the Chain Rule to find $\frac{d}{dx}(g(f(x)))$.

Exercise 2.5.3

Explain the differences in the two derivatives above.

Exercise 2.5.4

For functions $f(x)$ and $g(x)$, explain how to use the Chain Rule to find $\frac{d}{dx}(f(g(f(g(f(g(x)))))))$.

Exercise 2.5.5

Consider the following table of values of f, g, f' and g'.

Table 2.5.4

x	$f(x)$	$f'(x)$	$g(x)$	$g'(x)$
-1	2	3	0	-2
0	1	6	-1	-4
1	0	-4	1	$\frac{7}{3}$

Find the following derivatives, using the table.

a. $\left.\frac{d}{dx}(f(g(x)))\right|_{x=-1}$
b. $\left.\frac{d}{dx}(f(f(x)))\right|_{x=0}$
c. $\left.\frac{d}{dx}(g(f(x)))\right|_{x=1}$
d. $\left.\frac{d}{dx}(g(g(x)))\right|_{x=-1}$

Exercise 2.5.6

Find the following derivatives.

a. $\frac{d}{dx}\left(\tan\left(x^2\right)\right)$
b. $\frac{d}{dx}\left(e^{x^3+1}\right)$
c. $\frac{d}{dx}\left(\sin^4(x)\right)$
d. $\frac{d}{dx}(\sqrt{\sec(x)})$
e. $\frac{d}{dx}\left(\sqrt{e^{\cos(x)}}\right)$
f. $\frac{d}{dx}\left(\frac{4}{\sqrt[3]{x^2+x+1}}\right)$
g. $\frac{d}{dx}\left(\sec\left(e^x\right)\right)$

Exercise 2.5.7

Let's introduce a new kind of function: hyperbolic trigonometric functions. For now, let's just defined $y=\sinh(x)$ and $y=\cosh(x)$ (the hyperbolic sine and hyperbolic cosine functions) this way:

$$\sinh(x)=\frac{e^x-e^{-x}}{2}$$

$$\cosh(x)=\frac{e^x+e^{-x}}{2}$$

a. Show that $\frac{d}{dx}(\sinh(x))=\cosh(x)$.
b. Show that $\frac{d}{dx}(\cosh(x))=\sinh(x)$.
c. Let $\tanh(x)=\frac{\sinh(x)}{\cosh(x)}$. Find $\frac{d}{dx}(\tanh(x))$.
d. Find $\frac{d}{dx}\left(\sinh\left(x^2\right)\right)$.
e. Find $\frac{d}{dx}(\sqrt{\cosh(x)})$.

Exercise 2.5.8

Find the following derivatives.

a. $\frac{d}{dx}\left(\frac{\sin\left(e^x\right)}{\sqrt{x}}\right)$
b. $\frac{d}{dx}\left(e^{x\cos(x)}\right)$
c. $\frac{d}{dx}\left(\tan\left(\frac{x+1}{x^2+1}\right)\right)$

d. $\frac{d}{dx}\left((x+5)^2 e^{\tan(x)}\right)$

This page titled 2.5: The Chain Rule is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Peter Keep.

3: IMPLICIT DIFFERENTIATION

3.1: Implicit Differentiation
3.2: Derivatives of Inverse Functions
3.3: Logarithmic Differentiation

This page titled 3: Implicit Differentiation is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Peter Keep.

3.1: Implicit Differentiation

Let's quickly recap what we've done with this new calculus object, the derivative:

1. We defined the derivative at a point (Definition 2.1.1) to find the slope of a line touching a graph of a function $f(x)$ at a single point. We call this the "slope of the tangent line" at a point.
2. Once we calculated this slope, we quickly found a way to think about the derivative as a function (Definition 2.1.2) that connects x-values with the slope of the line tangent to $f(x)$ at that x-value.
3. We thought about how we could interpret the derivative as more than just a slope (Section 2.2). We can think about this as an instantaneous rate of change, and use it to add detail to how we think about mathematical models of different things.
4. We spent some time building up shortcuts, noticing patterns, and generalizing ways of finding these derivative functions for specific functions (Section 2.3) as well as combinations of those functions (Section 2.4 and Section 2.5).

Our goal, now, is to generalize this a bit. What happens when we push past the restriction of only dealing with functions? Can we think of some reasonable non-functions that might produce curves? Might we think about tangent lines and slopes in these contexts?

Explicit vs. Implicit Definitions

Definition 3.1.1: Explicitly and Implicitly Defined Curves.

A function or curve that is defined explicitly is one where the relationship between the variables is stated in with an equation like $y = f(x)$. Here, x is the input variable and we can find each corresponding value of the y-variable by applying some operations to x. As an example, we might consider the following function:

$$y = 3x + 1$$

A function or curve that is defined implicitly is one where the relationship between the variables is stated with an equation connecting the variables, but not necessarily one which is solved for a single variable. Here, the relationship between variables is not stated with the typical input and output variables. As an example, we might consider the same function as above, but defined as:

$$y - 3x - 1 = 0$$

Often, an implicitly defined curve is one where we cannot solve for a single variable by isolating it.

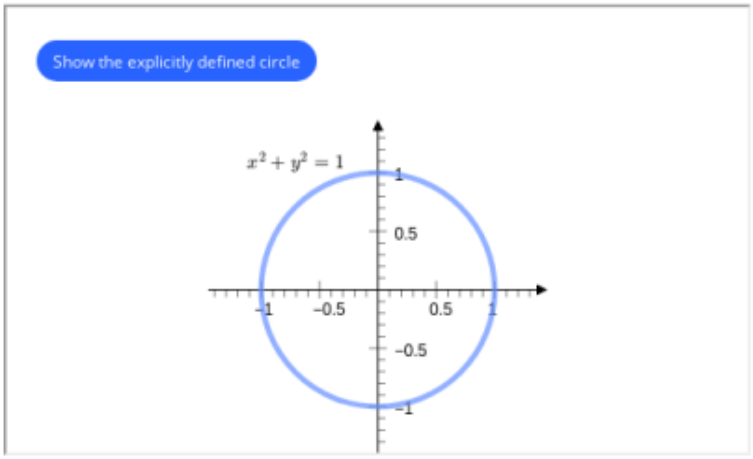
A classic and important implicitly defined curve is the unit circle:

$$x^2 + y^2 = 1$$

We can try to isolate for y and write this as an explicitly defined curve, and end up with:

$$y = \sqrt{1 - x^2}$$

Unfortunately, this only displays the upper half of the circle, since the square root will only output positive values. In this case, we can get around this by defining the circle with two functions.



Standalone Embed

As we move forward, let's work with this circle using the implicitly defined version ($x^2 + y^2 = 1$). How might we find a slope of a line tangent to this circle at some point?

Using a Derivative as an Operator

Let's recall back to Notation for Derivatives. We talked about how we can use the notation $\frac{d}{dx}(f(x))$ as a kind of action: the notation says "find the derivative of $f(x)$ with respect to x." When we say "with respect to x," we mean that we are treating x as an input variable, and trying to find out how f changes based on changes to that input. The notation says, "find the rate at which $f(x)$ changes as x increases."

Because this notation is a call to action, we can use it when we're dealing with an equation. We can call back to our early algebra days, where we learned that whatever we do to one side of an equation needs to be done to the other side, in order to maintain the equality.

We can apply this to our derivative actions: we can differentiate both sides of an equation!

Activity 3.1.1: Thinking about the Chain Rule.

a. Explain to someone how (and why) we use The Chain Rule to find the following derivative:

$$\frac{d}{dx}(\sqrt{\sin(x)})$$

b. Let's say that $f(x) = \sin(x)$. Explain how we find the following derivative:

$$\frac{d}{dx}(\sqrt{f(x)})$$

How is this different, or not different, than the previous derivative?

c. Let's say that we have some other function, $g(x)$. Explain how we find the following derivative:

$$\frac{d}{dx}(\sqrt{g(x)})$$

How is this different, or not different, than the previous derivatives?

d. What is the difference between the following derivatives:

$$\frac{d}{dx}(\sqrt{x}) \qquad \frac{d}{dx}(\sqrt{y}) \qquad \frac{d}{dy}(\sqrt{y})$$

Answer available at the end of this section.

Because we'll be applying our derivative notation to pieces of some equation, we need to be very aware of where we need to apply

the Chain Rule.

Now we can look at some examples of implicitly defined curves and think about questions about the derivative. Let's start with our circle.

Activity 3.1.2: Slopes on a Circle.

Visualize the unit circle. Feel free to draw one, or find the picture above. We're going to think about slopes on this circle.

a. Point out the locations on the unit circle where you would expect to see tangent lines that are perfectly horizontal. What do you think the value of the derivative, $\frac{dy}{dx}$, would be at these points?

b. Point out the locations on the unit circle where you would expect to see tangent lines that are perfectly vertical. What do you think the value of the derivative, $\frac{dy}{dx}$, would be at these points?

c. Find the point(s) where $x = \frac{1}{2}$. What do you think the value of the derivative, $\frac{dy}{dx}$, would be at these points? **Hint**: There are two points to consider here: $\left(\frac{1}{2}, \frac{\sqrt{3}}{2}\right)$ and $\left(\frac{1}{2}, \frac{-\sqrt{3}}{2}\right)$.

d. For the unit circle defined by the equation $x^2 + y^2 = 1$, apply the derivative to both sides of this equation to get the following:

$$\frac{d}{dx}\left(x^2 + y^2\right) = \frac{d}{dx}(1)$$
$$\frac{d}{dx}\left(x^2\right) + \frac{d}{dx}\left(y^2\right) = \frac{d}{dx}(1)$$

Carefully consider each of these derivatives (each of the terms). Which of these will you need to apply the Chain Rule for?

e. Differentiate. Solve for $\frac{dy}{dx}$ or y', whichever notation you decide to use. **Hint 1**: Make sure to use the Chain Rule when necessary! **Hint 2**: $\frac{d}{dx}\left(y^2\right) = \frac{d}{dy}\left(y^2\right)\frac{dy}{dx} = 2y \cdot \frac{dy}{dx}$ or $2yy'$

f. Go back to the first few questions, and try to answer them again:

   i. Find the locations of any horizontal tangent lines (where $\frac{dy}{dx} = 0$).

   ii. Find the locations of any vertical tangent lines (where $\frac{dy}{dx}$ doesn't exist, or where you would divide by 0).

   iii. Find the values of $\frac{dy}{dx}$ for the points on the circle where $x = \frac{1}{2}$.

Example 3.1.2

Let's repeat some of this process, but using a new curve. Consider the curve defined by the equation:

$$y^2 = x^3 - x + 1$$

This curve is a special curve with some interesting mathematical properties, and is actually a part of a family of curves called elliptic curves. For now, let's just consider it as a fun curve to look at, and use implicit differentiation to think about it.

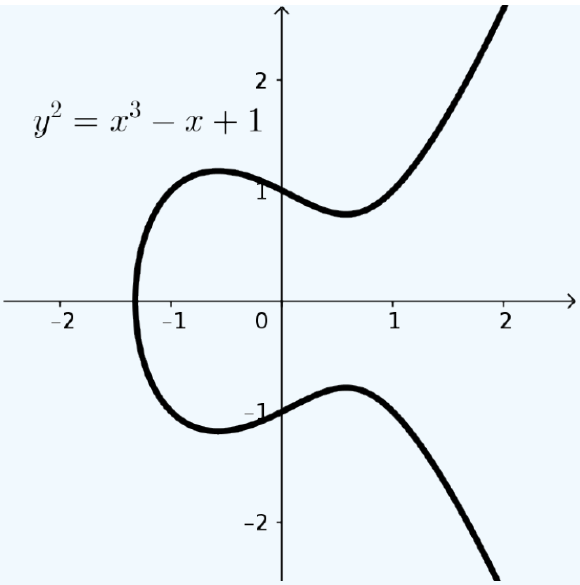


Figure 3.1.3

a. Mark the locations on the curve where it looks like the curve will have horizontal tangent lines. How many did you find?
b. Mark the locations on the curve where it looks like the curve will have vertical tangent lines. How many did you find?
c. Find the point(s) where $x = 0$. What do you think the value of the derivative, $\frac{dy}{dx}$, would be at these points?
d. For the elliptic curve defined by the equation $y^2 = x^3 - x + 1$, apply the derivative to both sides of this equation:

$$\frac{d}{dx}\left(y^2\right) = \frac{d}{dx}\left(x^3 - x + 1\right)$$
$$\frac{d}{dx}\left(y^2\right) = \frac{d}{dx}\left(x^3\right) - \frac{d}{dx}(x) + \frac{d}{dx}(1)$$

Carefully consider each of these derivatives (each of the terms). Which of these will you need to apply the Chain Rule for?

e. Differentiate. Solve for $\frac{dy}{dx}$ or y', whichever notation you decide to use. **Hint 1:** Make sure to use the Chain Rule when necessary! **Hint 2:** $\frac{d}{dx}\left(y^2\right) = \frac{d}{dy}\left(y^2\right)\frac{dy}{dx} = 2y \cdot \frac{dy}{dx}$ or $2yy'$
f. Go back to the first few questions, and try to answer them again:
   i. Find the locations of any horizontal tangent lines (where $\frac{dy}{dx} = 0$).
   ii. Find the locations of any vertical tangent lines (where $\frac{dy}{dx}$ doesn't exist, or where you would divide by 0).
   iii. Find the values of $\frac{dy}{dx}$ for the points on the curve where $x = 0$.

This example was pretty similar to the first activity: in both of these, we use the Chain Rule to differentiate $\frac{d}{dx}\left(y^2\right)$. Let's look at another example.

Activity 3.1.3: A New Curve.

Let's consider a new curve:

$$\sin(x) + \sin(y) = x^2y^2$$

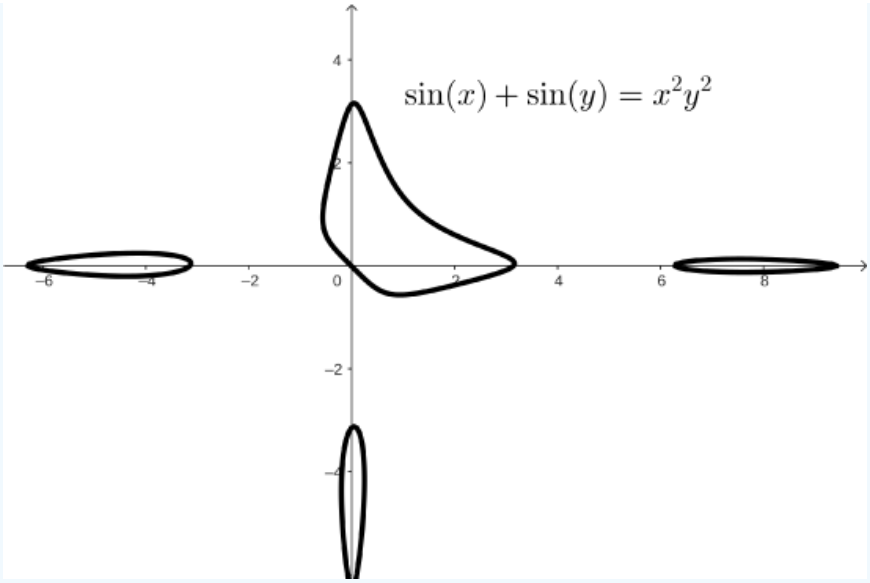


Figure 3.1.4

a. We are going to find $\frac{dy}{dx}$ or y'. Let's dive into differentiation:

$$\frac{d}{dx}(\sin(x) + \sin(y)) = \frac{d}{dx}\left(x^2y^2\right)$$
$$\frac{d}{dx}(\sin(x)) + \frac{d}{dx}(\sin(y)) = \frac{d}{dx}\left(x^2y^2\right)$$

Think carefully about these derivatives. For each of the three, how will you approach it? What kinds of nuances or rules or strategies will you need to think about? Why? **Hint**: Are any of these derivatives involving a variable other than x, the input variable (based on our $\frac{dy}{dx}$ notation, since we are thinking about how y changes with regard to x). Are any of these derivatives involving any other combination of functions? Are there products and/or quotients that we need to think about?

b. Implement your ideas or strategies from above to differentiate each term. **Hint**: We need to apply the Chain Rule to $\frac{d}{dx}(\sin(y))$ and then we need to apply the Product Rule $\frac{d}{dx}\left(x^2y^2\right)$. Notice that when we find the derivative of y^2 for the Product Rule, we need to use the Chain Rule!

c. Now we need to solve for $\frac{dy}{dx}$ or y', whichever you are using. While this equation can look complicated, we can notice something about the "location" of $\frac{dy}{dx}$ or y' in our equation. Why do we always know that $\frac{dy}{dx}$ or y' will be multiplied on a term whenever it shows up?

d. Now that we are confident that we will always know that we are multiplying this derivative, we can employ a consistent strategy:
   i. Rearrange our equation so that every term with a $\frac{dy}{dx}$ or y' is on one side, and everything without is on the other.
   ii. Now we are guaranteed that $\frac{dy}{dx}$ or y' is a common factor: factor it out.
   iii. Now we can solve for $\frac{dy}{dx}$ or y' by dividing! Solve for $\frac{dy}{dx}$ or y' in your equation!

e. Build the equation of a line that lays tangent to the curve at the origin. Does the value of $\frac{dy}{dx}$ at (0, 0) look reasonable to you?

Answer available at the end of this section.

Some Summary and Strategy

Let's wrap this up with some general strategy and summary of what we've seen.

The first thing we can notice is that we have talked through how to employ two of the three big derivative rules: we used the Chain Rule throughout these examples, and in Activity 3.1.3 we needed to use the Product Rule in order to differentiate $\frac{d}{dx}(x^2y^2)$. We have a glaring omission from our examples so far, though. Where is the Quotient Rule?

In these implicitly defined curves, we can manipulate the equations to never have to think about division! Consider the curve:

$$\frac{\sin(x)}{x} + \frac{\sin(y)}{x} = xy^2$$

Graph this curve in a graphing utility like Desmos. Does it look any different than the curve in Activity 3.1.3?

The only difference, really, is that the curve with the division is not defined at $x = 0$. As long as we keep those domain issues in mind, we can multiply everything by x to get our familiar equation:

$$\sin(x) + \sin(y) = x^2y^2$$

And there we go, we never have to think about the Quotient Rule in these kinds of contexts!

So we really only need a strategy for using the Chain Rule and the Product Rule.

Strategy for Implicit Differentiation.

- We use the Chain Rule whenever we differentiate something like $\frac{d}{dy}(f(y))$. We differentiate whatever the function is, and multiply by the derivative of y : $f'(y)y'$. This generalizes more: any time the variable in our derivative notation does not match the variable in the function/term, we need to use the Chain Rule:

$$\frac{d}{dy}(e^x) \qquad \frac{d}{dt}(\sin(x)) \qquad \frac{d}{dx}(y^4)$$

- We use the Product Rule whenever we differentiate a term with some combination of x and y variables. More generally, we can use the Product Rule any time we have a combination of at least two variables. We have to treat these as different kinds of functions getting multiplied!

$$\frac{d}{dy}(xe^y) \qquad \frac{d}{dt}(\cos(t)\sin(x)) \qquad \frac{d}{dx}\left(y^4\sqrt{x+1}\right)$$

Related Rates

What is the difference between the three derivatives below?

$$\frac{d}{dx}(x\sin(y)) \qquad \frac{d}{dy}(x\sin(y)) \qquad \frac{d}{dt}(x\sin(y))$$

We can notice that variable in the differential notation tells us when we apply the Chain Rule and when we don't need to.

This small idea allows us to think through some bigger application problems using implicit differentiation. We will call these Related Rates problems, since we'll use implicit differentiation to build an equation that relates at least two different rates of change to each other.

Activity 3.1.4: Dripping Water.

Two people are sitting at a table. Person A reaches across the table, grabs a water bottle, and begins dripping water onto the table right in front of Person B.

A puddle starts forming, in the shape of a circle.

a. If the puddle has a thickness of 0.5cm, what is the formula for the volume of the puddle?
b. Find a formula for how quickly the volume of the puddle is increasing per second.
c. Convince yourself that of course $\frac{dV}{dt}$ is connected to the change in the radius of the puddle, $\frac{dr}{dt}$.
d. If Person A is dripping the water at a rate of 0.2cm^3/s and the puddle is currently 13cm wide (in diameter), then how quickly is the edge of the puddle moving along the table?
e. How does this change when the puddle is larger? For instance, how does this change when the puddle is 20cm wide (in diameter)?

These related rates problems are a nice way for us to practice thinking about implicit differentiation, but also a nice way for us to think about how changes in one quantity might impact changes in another.

From here on out, we will use the ideas of implicit differentiation to accomplish two things:

1. We have a bit more flexibility with how we think of derivatives! We do not need to be restricted to only thinking about functions in order to think about rates of change or slopes at a point. We can think about any curve, any relationship between variables, and think about the relationship between them based on how one changes with regard to the other.
2. Implicit differentiation will be a very useful tool. Even when we have functions that can be written explicitly, they might be hard to deal with -- overly complex or maybe involving functions that we aren't used to. It is absolutely possible, and a really useful strategy, to rewrite the relationship between variables implicitly! We can maybe create a version of these equations that we can differentiate!

We're going to use this second idea first: in the next section we'll be thinking about inverse functions. We do not have any idea of how to think about derivatives of logarithmic functions, like $y = \ln(x)$.

We can rewrite this:

$$y = \ln(x) \longleftrightarrow x = e^y$$

This second representation is something we can differentiate!

Similarly, if we wanted to think about the derivative of $y = \tan^{-1}(x)$, we might instead think about rewriting this:

$$y = \tan^{-1}(x) \longleftrightarrow x = \tan(y)$$

There are some weird issues to think about with the domains and ranges of these functions, but this is how we'll approach these derivatives next.

Practice Problems

Exercise 3.1.1

Explain the difference between $\frac{d}{dx}(x^2)$ and $\frac{d}{dx}(y^2)$. How is the differentiation process different in each? Why?

Exercise 3.1.2

Explain why we need to use the product rule to differentiate $\frac{d}{dx}(ye^x)$.

Exercise 3.1.3

Use implicit differentiation to find $\frac{dy}{dx}$ or y'. Make a note of when you are using the product or chain rule, and make sure you know why you are using them.

a. $x^2 - y^2 = 0$
b. $4x + \sin(y) = 0$
c. $ye^x + 1 = 0$
d. $x^3 \tan(y) + e^y - 10x = 1$
e. $y^3 - xy = x^2 - 2x$
f. $4xy - y\sec(x) + x^5 = 0$
g. $\cos(xy) - x^3y^2 = 0$
h. $4x^3 = y^2(4 - x)$

Answer available at the end of this section.

Exercise 3.1.4

Pick 2 problems from Practice Problem 3, above, that you can solve for y explicitly. Find the derivative $\frac{dy}{dx}$ or y' this way, and confirm that your answers are the same.

Exercise 3.1.5

Pick 2 problems from Practice Problem 3, above, that you cannot solve for y explicitly. Explain why implicit differentiation is a useful method here.

Exercise 3.1.6

For the following curves, find the equation of the line tangent at the point given.

a. $3x^3 + 7y^3 = 10y$ at the point (1, 1)
b. $x^3 + y^3 = 9xy$ at the point (4, 2)
c. $y\left(x^2 + 4\right) = 8$ at the point (0, 2)
d. $\left(x^2 + y^2\right)^2 = \frac{25}{4}xy^2$ at (1, 2)

Answer available at the end of this section.

Exercise 3.1.7

Find the locations of the horizontal and vertical tangent lines for the curve:

$$x^2 + y^2 + 2xy = 12$$

Hint: Solve for $\frac{dy}{dx}$, and then think about where $\frac{dy}{dx}$ is either 0 (for the horizontal tangent lines) or doesn't exist (for the vertical ones).

Exercise 3.1.8

Consider the following curve:

$$(xt + 1)^3 = t - x^2 + 7$$

a. Find $\frac{dx}{dt}$. Explain where you need to use the chain rule and why.
b. Find $\frac{dt}{dx}$. Explain where you need to use the chain rule and why.

Answer available at the end of this section.

Related Rates. The rest of the exercises are related rates problems.

Exercise 3.1.9

Consider the equation:

$$3r + 5y^2 - 4 = 3\sin(x)$$

Differentiate both sides with regard to the variable t, and solve for $\frac{dr}{dt}$.

Answer available at the end of this section.

Exercise 3.1.10

Consider the equation:

$$3r + 5y^2 - 4 = 3\sin(x)$$

Differentiate both sides with regard to the variable t, and solve for $\frac{dy}{dt}$.

Answer available at the end of this section.

Exercise 3.1.11

A crab is walking along a path defined by the parabola $y = x^2 + 1$. It moves horizontally forward at a rate of 2 units/second, $\frac{dx}{dt} = 2$. If it is currently at the point (2, 5), what is the crab's vertical speed, $\frac{dy}{dt}$?

Answer available at the end of this section.

Exercise 3.1.12

Person A is holding a balloon 30m away from another Person B. If the balloon is released and floats into the air, rising straight up at a speed of 0.5m/s. Person B watches the balloon, their eye-line forming an angle θ with the ground. When the balloon is 10m up in the air, what is the speed at which θ changes, $\frac{d\theta}{dt}$?

Hint: You might use the fact that $\tan(\theta) = \frac{y}{x}$. Also, note that the horizontal distance between the people (x) is not changing.

Exercise 3.1.13

A paper cup shaped like a cone is leaking water out of the bottom. The cup is 4 inches tall, and the radius is 1/3 of the height. If the water is leaking at a rate of 1.5 in^3/s while the the cup is full, how quickly is the height of the water in the cup decreasing?

Hint: Remember, the volume of a cone is $V = \frac{\pi r^2 h}{3}$.

Exercise 3.1.14

A 12 foot ladder is leaning against a vertical wall. Someone is pushing the base of the ladder closer to the wall at a speed of 0.5 ft/s. If the ladder is currently 6 feet away from the wall, how quickly does the top of the ladder move up the wall?

Hint: The ladder and the wall form a right triangle, so maybe the Pythagorean Theorem will connect the side lengths! Also, note that the ladder is not changing length.

Exercise 3.1.15

A spherical water balloon is leaking water. The radius of the balloon is currently 2 inches, and we notice that it is decreasing at a rate of 1 in/min. How fast is the water leaking, in in^3/min?

Hint: The volume of a sphere is $V = \frac{4}{3}\pi r^3$

Answers

Activity 3.1.1**: Thinking about the Chain Rule.**

d. **Hint**: When do we need to use the Chain Rule? When do we need to use some linking derivative to connect the function we're looking at with the variable we care about?

$$\begin{aligned}
&\frac{d}{dx}(\sqrt{x}) = \frac{d}{dx}\left(x^{1/2}\right)\\
&= \frac{1}{2}\left(x^{-1/2}\right)\\
&= \frac{1}{2\sqrt{x}}\\
&\frac{d}{dx}(\sqrt{y}) = \frac{d}{dy}\left(y^{1/2}\right)\cdot\frac{dy}{dx}\\
&= \frac{1}{2}\left(y^{-1/2}\right)\cdot\frac{dy}{dx}\\
&= \frac{1}{2\sqrt{y}}\cdot\frac{dy}{dx}\\
&\text{or } \frac{y'}{2\sqrt{y}}\\
&\frac{d}{dy}(\sqrt{y}) = \frac{d}{dx}\left(x^{1/2}\right)\\
&= \frac{1}{2}\left(y^{-1/2}\right)\\
&= \frac{1}{2\sqrt{y}}
\end{aligned}$$

Activity 3.1.3**: A New Curve.**

b.

$$\begin{aligned}
\frac{d}{dx}(\sin(x)+\sin(y)) &= \frac{d}{dx}\left(x^2y^2\right)\\
\frac{d}{dx}(\sin(x)) + \frac{d}{dx}(\sin(y)) &= \frac{d}{dx}\left(x^2y^2\right)\\
\cos(x)+\cos(y)\cdot\frac{dy}{dx} &= 2xy^2 + 2x^2y\cdot\frac{dy}{dx}\\
\text{or } \cos(x)+y'\cos(y) &= 2xy^2 + 2x^2yy'
\end{aligned}$$

Exercise 3.1.3

$$\frac{d}{dx}\left(x^2\right) - \frac{d}{dx}\left(y^2\right) = \frac{d}{dx}(0)$$

a.
$$2x - 2y\frac{dy}{dx} = 0$$
$$\frac{dy}{dx}\ \frac{x}{y}$$

$$\frac{d}{dx}(4x) + \frac{d}{dx}(\sin(y)) = \frac{d}{dx}(0)$$

b.
$$4 + \cos(y)\frac{dy}{dx} = 0$$
$$\frac{dy}{dx} = -\frac{4}{\cos(y)}$$

$$\frac{d}{dx}\left(ye^x\right)+\frac{d}{dx}(1)=\frac{d}{dx}(0)$$

c.
$$e^x\frac{dy}{dx}+ye^x=0$$
$$\frac{dy}{dx}=-y$$

$$\frac{d}{dx}\left(x^3\tan(y)\right)+\frac{d}{dx}\left(e^y\right)-\frac{d}{dx}(10x)=\frac{d}{dx}(1)$$

d.
$$3x^2\tan(y)+x^3\sec^2(y)\frac{dy}{dx}+e^y\frac{dy}{dx}-10=0$$
$$\frac{dy}{dx}=\frac{10-3x^2\tan(y)}{x^3\sec^2(y)}$$

$$\frac{d}{dx}\left(y^3\right)-\frac{d}{dx}(xy)=\frac{d}{dx}\left(x^2\right)-\frac{d}{dx}(2x)$$

e.
$$3y^2\frac{dy}{dx}-\left(x\frac{dy}{dx}+y\right)=2x-2$$
$$\frac{dy}{dx}=\frac{2x-2+y}{3y^2-x}$$

$$\frac{d}{dx}(4xy)-\frac{d}{dx}(y\sec(x))+\frac{d}{dx}\left(x^5\right)=\frac{d}{dx}(0)$$

f.
$$4x\frac{dy}{dx}+4y-\left(\sec(x)\frac{dy}{dx}+y\sec(x)\tan(x)\right)+5x^4=0$$
$$\frac{dy}{dx}=\frac{y\sec(x)\tan(x)-4y-5x^4}{4x-\sec(y)}$$

$$\frac{d}{dx}(\cos(xy))-\frac{d}{dx}\left(x^3y^2\right)=\frac{d}{dx}(0)$$

$$-\sin(xy)\frac{d}{dx}(xy)-\left(3x^2y^2+2x^3y\frac{dy}{dx}\right)=0$$

g.
$$-\sin(xy)\left(y+x\frac{dy}{dx}\right)-3x^2y^2-2x^3y\frac{dy}{dx}=0$$
$$-y\sin(xy)-x\sin(y)\frac{dy}{dx}-3x^2y^2-2x^3y\frac{dy}{dx}=0$$
$$\frac{dy}{dx}=-\frac{y\sin(xy)+3x^2y^2}{x\sin(y)+2x^3y}$$

$$\frac{d}{dx}\left(4x^3\right)=\frac{d}{dx}\left(y^2(4-x)\right)$$

h.
$$12x^2=2y(4-x)\frac{dy}{dx}-y^2$$
$$\frac{dy}{dx}=\frac{12x^2+y^2}{2y(4-x)}$$

Exercise 3.1.6

$$\frac{d}{dx}\left(3x^3\right)+\frac{d}{dx}\left(7y^3\right)=\frac{d}{dx}(10y)$$
$$9x^2+21y^2\frac{dy}{dx}=10\frac{dy}{dx}$$
$$\frac{dy}{dx}=\frac{9x^2}{10-21y^2}$$

a.
$$\left.\frac{dy}{dx}\right|_{(1,1)}=\frac{9(1)^2}{10-21(1)^2}$$
$$=-\frac{9}{11}$$

The line tangent to the curve at $(1,1)$ is:

$$y=-\frac{9}{11}(x-1)+1$$
$$=-\frac{9}{11}x+\frac{20}{11}$$

$$\frac{d}{dx}\left(x^3\right)+\frac{d}{dx}\left(y^3\right)=\frac{d}{dx}(9xy)$$
$$3x^2+3y^2\frac{dy}{dx}=9y+9x\frac{dy}{dx}$$
$$\frac{dy}{dx}=\frac{9y-3x^2}{3y^2-9x}$$

b.
$$\left.\frac{dy}{dx}\right|_{(4,2)}=\frac{9(2)-3(4)^2}{3(2)^2-9(4)}$$
$$=\frac{5}{3}$$

The line tangent to the curve at $(4,2)$ is:

$$y=\frac{5}{3}(x-4)+2$$
$$=\frac{5}{3}x-\frac{14}{3}$$

$$\frac{d}{dx}\left(y\left(x^2+4\right)\right)=\frac{d}{dx}(8)$$
$$\frac{dy}{dx}\left(x^2+4\right)+2xy=0$$

c.
$$\frac{dy}{dx}=-\frac{2xy}{x^2+4}$$
$$\left.\frac{dy}{dx}\right|_{(0,2)}=-\frac{2(0)(2)}{(0)^2+4}$$
$$=0$$

The line tangent to the curve at $(0,2)$ is:
$y=0(x-0)+2$
$=2$

d.
$$\frac{d}{dx}\left(\left(x^2+y^2\right)^2\right)=\frac{d}{dx}\left(\frac{25}{4}xy^2\right)$$
$$2\left(x^2+y^2\right)\frac{d}{dx}\left(x^2+y^2\right)=\frac{25}{4}y^2+\frac{25}{2}xy\frac{dy}{dx}$$
$$2\left(x^2+y^2\right)\left(2x+2y\frac{dy}{dx}\right)=\frac{25}{4}y^2+\frac{25}{2}xy\frac{dy}{dx}$$
$$4x^3+4xy^2+4x^2y\frac{dy}{dx}+4y^3\frac{dy}{dx}=\frac{25}{4}y^2+\frac{25}{2}xy\frac{dy}{dx}$$
$$16x^3+16xy^2+16x^2y\frac{dy}{dx}+16y^3\frac{dy}{dx}=25y^2+50xy\frac{dy}{dx}$$
$$\frac{dy}{dx}=\frac{16x^3+16xy^2-25y^2}{50xy-16x^2y-16y^3}$$
$$\left.\frac{dy}{dx}\right|_{(1,2)}=\frac{16(1)^3+16(1)(2)^2-25(2)^2}{50(1)(2)-16(1)^2(2)-16(2)^3}$$
$$=\frac{1}{3}$$

Exercise 3.1.8

$$\frac{d}{dt}\left((xt+1)^3\right)=\frac{d}{dt}(t)-\frac{d}{dt}\left(x^2\right)+\frac{d}{dt}(7)$$
$$3(xt+1)^2\frac{d}{dt}(xt+1)=1-2x\frac{dx}{dt}+0$$

a.
$$3(xt+1)^2\left(x+t\frac{dx}{dt}\right)=1-2x\frac{dx}{dt}$$
$$\frac{dx}{dt}=\frac{1-3x(xt+1)^2}{3t(xt+1)^2-2x}$$

$$\frac{d}{dx}\left((xt+1)^3\right)=\frac{d}{dx}(t)-\frac{d}{dx}\left(x^2\right)+\frac{d}{dx}(7)$$
$$3(xt+1)^2\frac{d}{dx}(xt+1)=\frac{dt}{dx}-2x$$

b.
$$3(xt+1)^2\left(t+x\frac{dt}{dx}\right)=\frac{dt}{dx}-2x$$
$$\frac{dt}{dx}=\frac{2x+3t(xt+1)^2}{1-3x(xt+1)^2}$$

Exercise 3.1.9

$$\frac{d}{dt}(3r)+\frac{d}{dt}\left(5y^2\right)-\frac{d}{dt}(4)=\frac{d}{dt}(3\sin(x))$$
$$3\frac{dr}{dt}+10y\frac{dy}{dt}=3\cos(x)\frac{dx}{dt}$$
$$\frac{dr}{dt}=\frac{3\cos(x)\frac{dx}{dt}-10y\frac{dy}{dt}}{3}$$

Exercise 3.1.10

$$\frac{d}{dt}(3r)+\frac{d}{dt}\left(5y^2\right)-\frac{d}{dt}(4)=\frac{d}{dt}(3\sin(x))$$
$$3\frac{dr}{dt}+10y\frac{dy}{dt}=3\cos(x)\frac{dx}{dt}$$
$$\frac{dy}{dt}=\frac{3\cos(x)\frac{dx}{dt}-3\frac{dr}{dt}}{10y}$$

Exercise 3.1.11

$$\frac{d}{dt}(y)=\frac{d}{dt}\left(x^2\right)+\frac{d}{dt}(1)$$
$$\frac{dy}{dt}=2x\frac{dx}{dt}$$

When we evaluate with $x=2, y=5$, and $\frac{dx}{dt}=2$ we get

$$\frac{dy}{dt}=2(2)(2)$$
$$=8$$

The crab's vertical speed is 8 units/second.

This page titled 3.1: Implicit Differentiation is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Peter Keep.

3.2: Derivatives of Inverse Functions

We should start here by saying: we're going to be thinking about inverse functions, and so maybe it will be helpful to recap some facts about inverse functions.

- If $y = f(x)$ is some function, then we can use the inverse function to represent this relationship between variables: $x = f^{-1}(y)$. Some examples:
  - $y = e^x \longleftrightarrow x = \ln(y)$. That is, the logarithm function "solves" for the exponent (sometimes this is easier to just say that the logarithm is the exponent).
  - $y = \sin(\theta) \longleftrightarrow \theta = \sin^{-1}(y)$. That is, this inverse sine function (or sometimes $\arcsin(y)$) finds the angle at which sine of that angle is y. With these trigonometric functions, we need to make some restrictions: because there are an infinite number of angles that will produce the same output of the sine function (reflecting the angle across the y-axis will do it, as will adding any multiple of 2π), we restrict the angles that the inverse sine function can output to being in $\left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$.
- Based on this re-representation above, we can always compose a function and its inverse to get the identity function, $y = x$. In general, if $y = f(x)$ has an inverse function f^{-1}, then $\left(f \circ f^{-1}\right)(x) = f\left(f^{-1}(x)\right) = x$. Similarly, we can compose in the opposite order: $\left(f^{-1} \circ f\right)(x) = f^{-1}(f(x)) = x$. This can be a bit trickier to think about for the inverse trigonometric functions, since this only works on intervals of x where that inverse is defined. So we end up with strange things like:

  $$\sin^{-1}\left(\sin\left(\frac{3\pi}{2}\right)\right) = \sin^{-1}(-1) = -\frac{\pi}{2}$$

  This is because the inverse sine function finds only angles in the interval $\left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$, and the angles $\frac{3\pi}{2}$ and $-\frac{\pi}{2}$ are coterminal (and so have the same output from the sine function).

With these small facts in mind, we can think about derivatives of inverse functions.

Wielding Implicit Differentiation

We're going to do a very cool thing: in order to find derivatives of inverse functions, we can invert the relationship between x and y, and then use Implicit Differentiation to find $\frac{dy}{dx}$.

Activity 3.2.1: Building the Derivative of the Logarithm.

We're going to accomplish two things here:

1. By the end of this activity, we'll have a nice way of thinking about $\frac{d}{dx}(\ln(x))$, and we will now be able to differentiate functions involving logarithms!
2. Throughout this activity, we're going to develop a way of approaching derivatives of inverse functions more generally. Then we can apply this framework to other functions

Let's think about this logarithmic function!

a. We have stated (a couple of times now) how we define the log function:

$$y = e^x \longleftrightarrow x = \ln(y)$$

This arrow goes both directions: the log function is the inverse of the exponential, but the exponential is the inverse of the log function! Can you rewrite the relationship $y = \ln(x)$ using its inverse (the exponential)?

b. For your inverted $y = \ln(x)$ from above (it should be $x =$ ___), apply a derivative operator to both sides, and use implicit differentiation to find $\frac{dy}{dx}$ or y'.

c. A weird thing that we can notice is that when we use implicit differentiation, it is common to end up with our derivative written in terms of both x and y variables. This makes sense for our earlier examples: we needed specific coordinates of the point on the circle, for instance, to find the slope there. But if $y = \ln(x)$, we want $\frac{dy}{dx}$ or y' to be a function of

$$x : f(x) = \ln(x) \longrightarrow f'(x) = \_\_\_$$

Your derivative is written with only y values. Since $y = \ln(x)$, replace any instance of y with the log function. What do you have left?

d. Remember that $y = \ln(x)$. Substitute this into your equation for $\frac{dy}{dx}$. Can you write this in a pretty simplistic way?

e. Before we go much further, we should be a bit careful. What is the domain of this derivative? What are the values of x where $\frac{d}{dx}(\ln(x))$ makes sense to think about?

Theorem 3.2.1: Derivative of the Logarithmic Function.

$$\frac{d}{dx}(\ln(x)) = \frac{1}{x} \text{ for } x > 0$$

Further, since $\log_b(x) = \frac{\ln(x)}{\ln(b)}$ (for $b \geq 0$ and $b \neq 1$), we can say that:

$$\frac{d}{dx}(\log_b(x)) = \frac{1}{x\ln(b)} \text{ for } x > 0$$

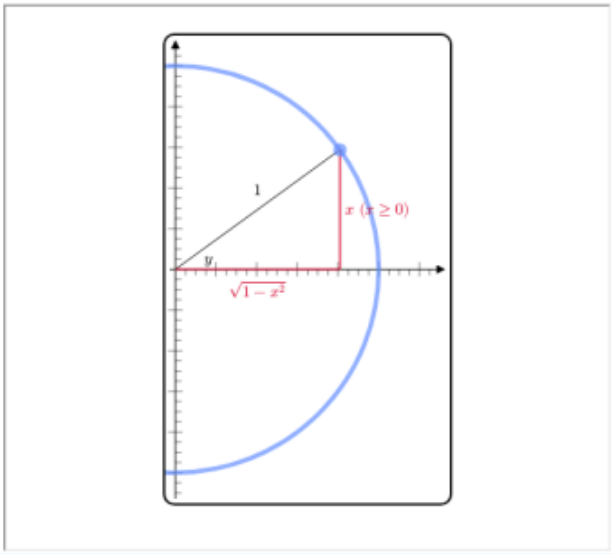
Derivatives of the Inverse Trigonometric Functions

Let's try a similar thing with inverse trigonometric functions. We'll start with the inverse sine function, $y = \sin^{-1}(x)$.

Activity 3.2.2: Finding the Derivative of the Inverse Sine Function.

We're going to do the same trick, except that there will be a couple of small differences due to thinking specifically about trigonometric functions. Let's think about the function $y = \sin^{-1}(x)$. We know that this is equivalent to $x = \sin(y)$ (for y-values in $\left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$).

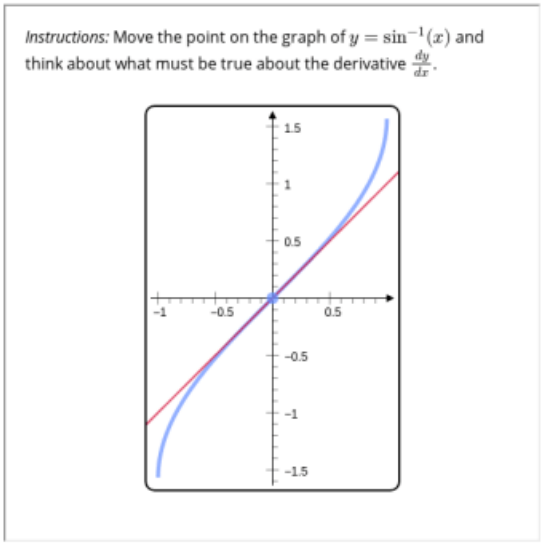
a. Move the point around the portion of the unit circle in the graph below. Convince yourself that:
- sin(y) = x
- sin(y) ≥ 0 when 0 ≤ y ≤ π 2
- sin(y) < 0 when − π 2 ≤ y < 0



Standalone Embed

What is $\cos(y)$ in this figure? Does the sign change depending on the value of y?

b. Use implicit differentiation and the equation $x = \sin(y)$ to find $\frac{dy}{dx}$ or y'
c. If you still have your derivative written in terms of y, make sure to write $\cos(y)$ in terms of x!
d. Let's think about the domain of this derivative: what x-values make sense to think about? Think about this both in terms of what x-values reasonably fit into your formula of $\frac{d}{dx}\left(\sin^{-1}(x)\right)$ as well as the domain of the inverse sine function in general.
e. Notice that in the denominator of $\frac{d}{dx}\left(\sin^{-1}(x)\right)$, you have a square root. Based on that information (and the visual above), what do you expect to be true about the sign of the derivative of the inverse sine function? Confirm this by playing with the graph of $y = \sin^{-1}(x)$ below.



Standalone Embed

f. Investigate the behavior of $\frac{dy}{dx}$ at the end-points of the function: at $x = -1$ and $x = 1$. What do the slopes look like they're doing, graphically? How does this work when you look at the function you built above? What happens when you try to find $\left.\frac{dy}{dx}\right|_{x=-1}$ or $\left.\frac{dy}{dx}\right|_{x=1}$?

Let's repeat the process to find the derivatives of $y = \tan^{-1}(x)$ and $y = \sec^{-1}(x)$.

Activity 3.2.3: Building the Derivatives for Inverse Tangent and Secant.

a. Consider the triangle representing the case when $y = \tan^{-1}(x)$.

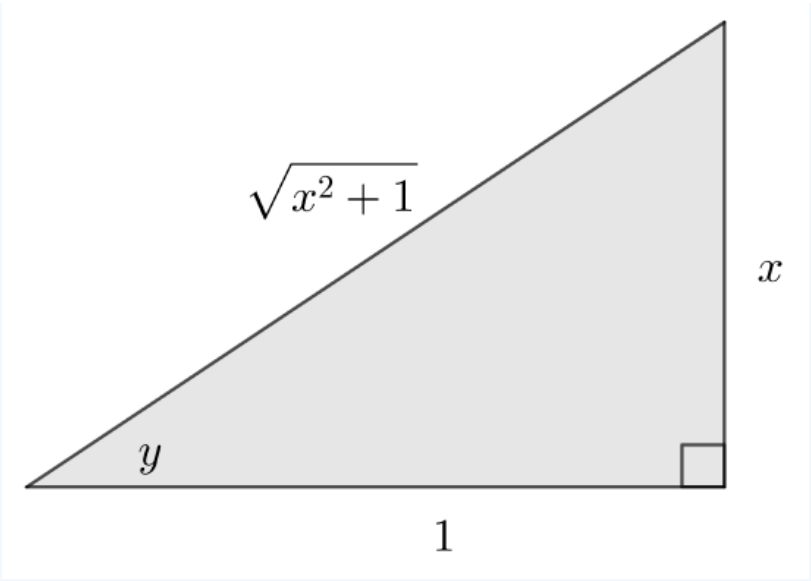


Figure 3.2.2

For $x = \tan(y)$, find $\frac{dy}{dx}$ using implicit differentiation. Find an appropriate expression for $\sec(y)$ based on the triangle above, but we will refer back to the version with the $\sec(y)$ in it later.

b. Consider the triangle representing the case when $y = \sec^{-1}(x)$.

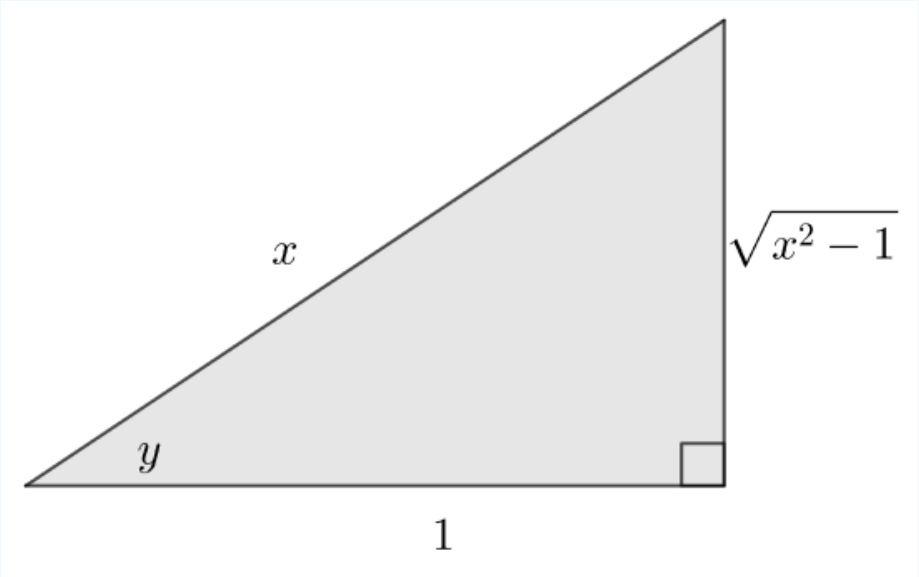
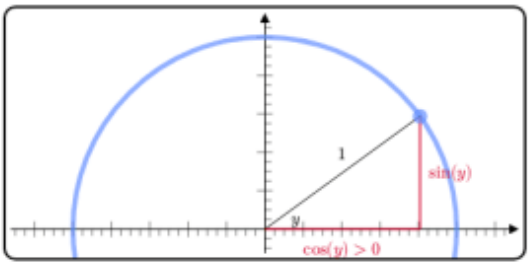


Figure 3.2.3

For $x = \sec(y)$, find $\frac{dy}{dx}$ using implicit differentiation. Find an appropriate expression for $\sec(y)$ and $\tan(y)$ based on the triangle above, but we will refer back to the version with the functions of y in it later.

c. Here's a graph of just the unit circle for angles $[0, \pi]$. We are choosing to focus on this region, since these are the angles that the inverse tangent and inverse secant functions will return to us. We want to investigate the signs of $\tan(y)$ and $\sec(y)$.

Instructions: For angles y in $[0, \pi]$, find the signs of $\tan(y)$ and $\sec(y)$. When are each of them positive or negative?



▸ Hint (click to open)

Standalone Embed

d. Go back to our derivative expressions for both the inverse tangent and inverse secant functions. What do we know about the signs of these derivatives? **Hint**:

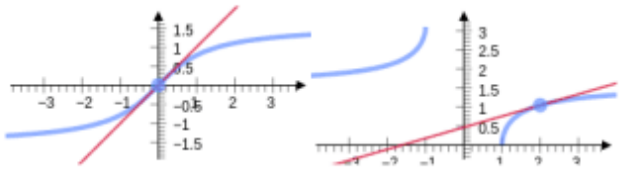
Notice that in $\frac{d}{dx}\left(\tan^{-1}(x)\right) = \frac{1}{\sec^2(y)}$, we know that the derivative must be positive. Even when $\sec(y) < 0$, we are squaring it.

In $\frac{d}{dx}\left(\sec^{-1}(x)\right) = \frac{1}{\sec(y)\tan(y)}$, we know that the derivative must also always be positive. Whenever $\sec(y) < 0$, we have $\tan(y) < 0$, and so the product must be positive.

e. Confirm your idea about the sign of the derivatives by investigating the graphs of each function.

Instructions: Move the points on the graphs of $y = \tan^{-1}(x)$ and $y = \sec^{-1}(x)$ and think about what must be true about the signs of the derivatives.

Graph of $y = \tan^{-1}(x)$ Graph of $y = \sec^{-1}(x)$



Standalone Embed

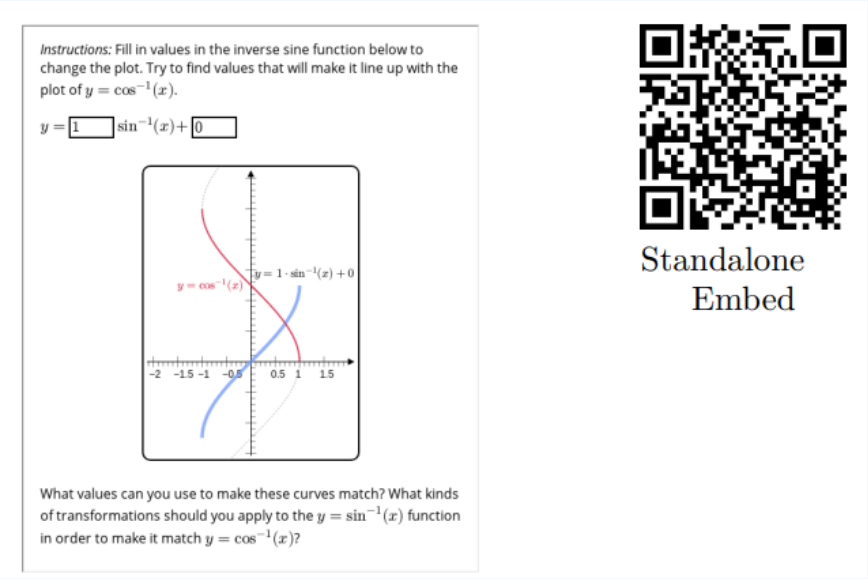
f. How do we need to write these derivatives, when we write them in terms of x to account for the sign of the derivative? **Hint**: Use an absolute value in the formula for $\frac{d}{dx}\left(\sec^{-1}(x)\right)$!

It is important to think carefully about how things might change when we start thinking about other trigonometric functions. For instance, what happens when we think about $y = \cos^{-1}(x)$ instead? We could repeat the process from Activity 3.2.2 with $y = \cos^{-1}(x)$ instead (and we'll do that for $y = \tan^{-1}(x)$), but for now let's think about the graph of $y = \cos^{-1}(x)$.

Activity 3.2.4: Connecting These Inverse Functions.

We're going to look at a graph of $y = \cos^{-1}(x)$, but we're specifically going to try to compare it to the graph of $y = \sin^{-1}(x)$. We'll use some graphical transformations to make these functions match up, and then later we'll think about derivatives.

a. Ok, consider the graph of $y = \cos^{-1}(x)$ and a transformed version of the inverse sine function. Apply some graphical transformations to make these match!



b. It might be fun to think about another reason that this connection between $\sin^{-1}(x)$ and $\cos^{-1}(x)$ exists. Consider this triangle:

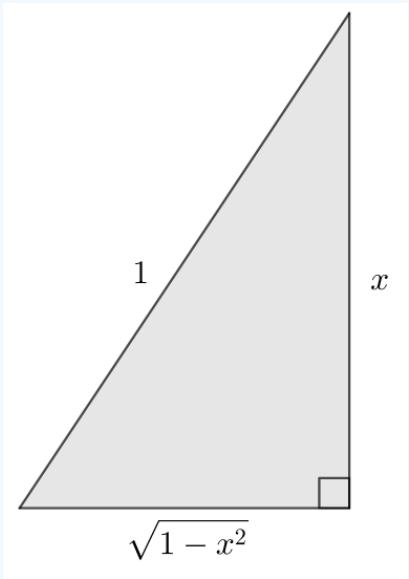


Figure 3.2.4

We're going to think about these inverse trigonometric functions as angles: let $\alpha = \cos^{-1}(x)$ and $\beta = \sin^{-1}(x)$. We can rewrite these as:

$$\cos(\alpha) = x$$
$$\sin(\beta) = x$$

Can you fill in your triangle using this information? Why does $\alpha + \beta = \frac{\pi}{2}$? Convince yourself that this is what we did with the graphical transformations above, as well.

c. Use this equation above, rewriting $\cos^{-1}(x)$ as some expression involving the inverse sine function, and then find $\frac{d}{dx}\left(\cos^{-1}(x)\right)$. **Hint**: $\frac{d}{dx}\left(\cos^{-1}(x)\right) = \frac{d}{dx}\left(-\sin^{-1}(x) + \frac{\pi}{2}\right)$

We could repeat this task to try to connect the graph of $y = \tan^{-1}(x)$ with $y = \cot^{-1}(x)$ as well as the graph of $y = \sec^{-1}(x)$ with $y = \csc^{-1}(x)$, but we can hopefully see what will happen. In each case, we have the same kind of connection that we saw in the triangle, since these are cofunctions of each other! We can summarize by believing that:

$$\frac{d}{dx}\left(\cos^{-1}(x)\right) = -\frac{d}{dx}\left(\sin^{-1}(x)\right) = -\frac{1}{\sqrt{1-x^2}}$$
$$\frac{d}{dx}\left(\cot^{-1}(x)\right) = -\frac{d}{dx}\left(\tan^{-1}(x)\right) = -\frac{1}{x^2+1}$$
$$\frac{d}{dx}\left(\csc^{-1}(x)\right) = -\frac{d}{dx}\left(\sec^{-1}(x)\right) = -\frac{1}{|x|\sqrt{x^2-1}}$$

Theorem 3.2.5: Derivatives of the Inverse Trigonometric Functions.

$$\frac{d}{dx}\left(\sin^{-1}(x)\right) = \frac{1}{\sqrt{1-x^2}} \qquad \text{Domain: } -1 < x < 1$$

$$\frac{d}{dx}\left(\cos^{-1}(x)\right) = -\frac{1}{\sqrt{1-x^2}} \qquad \text{Domain: } -1 < x < 1$$

$$\frac{d}{dx}\left(\tan^{-1}(x)\right) = \frac{1}{x^2+1} \qquad \text{Domain: all Real numbers}$$

$$\frac{d}{dx}\left(\cot^{-1}(x)\right) = -\frac{1}{x^2+1} \qquad \text{Domain: all Real numbers}$$

$$\frac{d}{dx}\left(\sec^{-1}(x)\right) = \frac{1}{|x|\sqrt{x^2-1}} \qquad \text{Domain: } x < -1 \text{ and } x > 1$$

$$\frac{d}{dx}\left(\csc^{-1}(x)\right) = -\frac{1}{|x|\sqrt{x^2-1}} \qquad \text{Domain: } x < -1 \text{ and } x > 1$$

Now that we have these new derivative rules for these functions, we can practice applying them! We can use these along with the Product, Quotient, and Chain Rules to differentiate many different functions.

Practice Problems

Exercise 3.2.1

Find the derivatives of the following functions.

a. $y = 7^x$
b. $y = \log_3(x)$
c. $A(t) = 400(1.001)^{12t}$
d. $f(x) = \frac{3^x}{3^x+1}$
e. $g(x) = \ln\left(x^2+1\right)$
f. $s(x) = \left(e^x\right)\left(x^e\right)$
g. $y = \ln\left(4x^3\right)$
h. $y = \ln\left(\sqrt{x^4+x^2}\right)$
i. $y = x\ln(x) - x$
j. $y = \sin\left(4^x\right)$
k. $y = e^x\ln(x)$

l. $r(\theta) = (\theta^2)\left(5^{\theta}\right)$
m. $y = \ln(\ln(x))$
n. $y = -\ln(\cos(x))$

Answer available at the end of this section.

Exercise 3.2.2

Find the derivatives of the following functions

a. $f(x) = \sin^{-1}(5x)$
b. $f(x) = x^2 \tan^{-1}(x)$
c. $f(x) = \left(\sec^{-1}(x)\right)^3$
d. $f(x) = \cos^{-1}\left(\dfrac{1}{x}\right)$
e. $f(x) = \ln\left(\sin^{-1}(x)\right)$
f. $f(x) = \cot^{-1}\left(e^{2x}\right)$

Answer available at the end of this section.

Exercise 3.2.3

Find the equation of the line tangent to the graph of the following functions at the point listed.

a. $f(x) = \tan^{-1}(x)$ at $\left(1, \dfrac{\pi}{4}\right)$
b. $f(x) = \sin^{-1}\left(\dfrac{x}{4}\right)$ at $\left(2, \dfrac{\pi}{6}\right)$
c. $f(x) = \cos^{-1}\left(x^2\right)$ at $\left(\dfrac{1}{\sqrt{2}}, \dfrac{\pi}{3}\right)$
d. $f(x) = \sec^{-1}\left(e^x\right)$ at $\left(\ln(2), \dfrac{\pi}{3}\right)$

Answer available at the end of this section.

Exercise 3.2.4

Graph the function $f(x) = \tan^{-1}(x) - x^3 - x$ using a graphing utility like Desmos.

a. Find $f'(x)$.
b. Graph $f'(x)$ using a graphing utility like Desmos.
c. Where is $f'(x) = 0$?
d. Where does $f(x)$ have a horizontal tangent line?

Answer available at the end of this section.

Answers

Exercise 3.2.1

a. $y' = 7^x \ln(7)$
b. $y' = \dfrac{1}{x \ln(3)}$
c. $A'(t) = 400(1.001)^{12t} \ln(1.001) \dfrac{d}{dt}(12t)$
$= 4800(1.001)^{12t} \ln(1.001)$

d. $$f'(x) = \frac{(3^x+1)(3^x)\ln(3) - (3^x)(3^x)\ln(3)}{(3^x+1)^2}$$
$$f'(x) = \frac{3^x \ln(3)}{(3^x+1)^2}$$

e. $$g'(x) = \frac{1}{x^2+1}\frac{d}{dx}(x^2+1)$$
$$= \frac{2x}{x^2+1}$$

f. $$s'(x) = (e^x)(x^e) + (e^x)(ex^{e-1})$$

g. $$y' = \frac{1}{4x^3}\frac{d}{dx}(4x^3)$$
$$= \frac{12x^2}{4x^3}$$
$$= \frac{3}{x}$$

h. $$y' = \frac{1}{\sqrt{x^4+x^2}}\frac{d}{dx}\left(\sqrt{x^4+x^2}\right)$$
$$= \frac{1}{\sqrt{x^4+x^2}}\left(\frac{1}{2\sqrt{x^4+x^2}}\right)\frac{d}{dx}(x^4+x^2)$$
$$= \frac{4x^3+2x}{2(x^4+x^2)}$$

i. $$y' = x\left(\frac{1}{x}\right) + \ln(x) - 1$$
$$= \ln(x)$$

j. $$y' = \cos(4^x)\,4^x \ln(4)$$

k. $$y' = e^x(\ln(x)) + e^x\left(\frac{1}{x}\right)$$
$$= e^x\left(\ln(x) + \frac{1}{x}\right)$$

l. $$r'(\theta) = (2\theta)(5^\theta) + (\theta^2)(5^\theta \ln(5))$$
$$= \theta 5^\theta(2 + \theta\ln(5))$$

m. $$y' = \frac{1}{\ln(x)}\frac{d}{dx}(\ln(x))$$
$$= \frac{1}{\ln(x)}\left(\frac{1}{x}\right)$$
$$= \frac{1}{x\ln(x)}$$

n. $$y' = -\frac{1}{\cos(x)}(-\sin(x))$$
$$= \frac{\sin(x)}{\cos(x)}$$
$$= \tan(x)$$

Exercise 3.2.2

a. $$f'(x) = \frac{1}{\sqrt{1-(5x)^2}}\frac{d}{dx}(5x)$$
$$= \frac{5}{\sqrt{1-25x^2}}$$

b. $$f'(x) = 2x\tan^{-1}(x) + x^2\left(\frac{1}{x^2+1}\right)$$
$$= 2x\tan^{-1}(x) + \frac{x^2}{x^2+1}$$

c. $$f'(x) = 3\left(\sec^{-1}(x)\right)^2\left(\frac{1}{|x|\sqrt{x^2-1}}\right)$$
$$= \frac{3\left(\sec^{-1}(x)\right)^2}{|x|\sqrt{x^2-1}}$$

d. $$f'(x) = -\frac{1}{\sqrt{1-\left(\frac{1}{x}\right)^2}}\frac{d}{dx}\left(\frac{1}{x}\right)$$
$$= -\frac{1}{\sqrt{1-\frac{1}{x^2}}}\left(-\frac{1}{x^2}\right)$$
$$= \frac{1}{x^2\sqrt{1-\frac{1}{x^2}}}$$
$$= \frac{1}{\sqrt{x^4-x^2}}$$

e. $$f'(x) = \frac{1}{\sin^{-1}(x)}\frac{d}{dx}\left(\sin^{-1}(x)\right)$$
$$= \frac{1}{\sin^{-1}(x)}\left(\frac{1}{\sqrt{1-x^2}}\right)$$
$$= \frac{1}{\sin^{-1}(x)\sqrt{1-x^2}}$$

f. $$f'(x) = -\frac{1}{1+(e^{2x})^2}\frac{d}{dx}\left(e^2 x\right)$$
$$= -\frac{2e^{2x}}{1+e^{4x}}$$

Exercise 3.2.3

$$f'(x) = \frac{1}{x^2+1}$$

a. $$f'(1) = \frac{1}{2}$$
$$y = \frac{1}{2}(x-1) + \frac{\pi}{4}$$

$$f'(x) = \frac{1}{4\sqrt{1-\frac{x^2}{16}}}$$

b. $$f'(2) = \frac{1}{4\sqrt{\frac{3}{4}}}$$
$$= \frac{1}{2\sqrt{3}}$$
$$y = \frac{1}{2\sqrt{3}}(x-2) + \frac{\pi}{6}$$

$$f'(x) = -\frac{2x}{\sqrt{1-x^4}}$$

c. $$f'(1/\sqrt{2}) = -\frac{4}{\sqrt{6}}$$
$$y = -\frac{4}{\sqrt{6}}\left(x - \frac{1}{\sqrt{2}}\right) + \frac{\pi}{3}$$

$$f'(x) = \frac{1}{\sqrt{e^{2x}-1}}$$

d. $$f'(\ln(2)) = \frac{1}{\sqrt{3}}$$
$$y = \frac{1}{\sqrt{3}}(x - \ln(2)) + \frac{\pi}{3}$$

Exercise 3.2.4

a. $$f'(x) = \frac{1}{x^2+1} - 2x - 1$$

This page titled 3.2: Derivatives of Inverse Functions is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Peter Keep.

3.3: Logarithmic Differentiation

We're going to start with a quick fact about logs and their derivatives. The derivative rule for $\frac{d}{dx}(\ln(x))$ is still relatively new for us, so it is ok to still be getting comfortable with it, but we should notice this nice fact.

Fact 3.3.1 Derivative of the Log of a Function.

$$\frac{d}{dx}(\ln(f(x))) = \frac{f'(x)}{f(x)}(\text{ when } f(x) > 0)$$

Note that there's nothing really special going on here: this is just an application of The Chain Rule to the Derivative of the Logarithmic Function.

Throughout this section, the goal is to convince any open-minded readers of one thing:

Logs are friends.

Let us be informal and technically not quite correct but hopefully clear in this. Logs really are friendly mathematical objects. They were created to be friendly objects! In a time when doing arithmetic with large numbers was difficult due to a lack of computing technology, logs were introduced to make arithmetic easier.

The general idea is that, if there is some kind of hierarchy of operations, then logs transform operations between things into different operations that are lower on the hierarchy of operations. So logs turn things like products (repeated addition) and quotients (repeated subtraction) into addition and subtraction. Logs turn exponents (repeated multiplication) into coefficients.

Using math notation, we can write the following log properties.

Fact 3.3.2 Properties of Logarithms.

We will make use of the following properties of logarithms.

- $\ln(xy) = \ln(x) + \ln(y)$
- $\ln\left(\frac{x}{y}\right) = \ln(x) - \ln(y)$
- $\ln\left(x^y\right) = y\ln(x)$

Because of the domain of the log function, we need $x, y > 0$ for these properties to make sense. We will use them relatively loosely, with functions that take on negative and positive values, and not worry too much about the domain issues when we don't need to.

Logs Are Friends!

Ok, so how will we use these new-found friends? We're going to think about some functions (and combinations of functions) that are new to us and that aren't so clear for us to use things like the Product, Quotient, or the Chain Rule. We'll try to use logs to rewrite our functions into some easier to approach implicitly defined relationships in order for us to differentiate.

But first, let's build an explanation for the Power Rule for Derivatives.

Activity 3.3.1: Returning to the Power Rule.

Back in Section 2.3 we built an explanation for why $\frac{d}{dx}(x^n) = nx^{n-1}$ that relied on thinking about exponents as repeated multiplication: it relied on n being some positive integer. We said, at the time, that the Power Rule generalizes and works for any integer, but did so without explanation.

Let's consider $y = x^n$ where n is just some real number without any other restrictions

a. Apply a logarithm to both sides of this equation:

$$\ln(y) = \ln(x^n)$$

Now use one of the Properties of Logarithms to rewrite this equation.

b. Use implicit differentiation to find $\frac{dy}{dx}$ or y'.

c. Explain to yourself why this is equivalent to the Power Rule that we built so long ago: $\frac{d}{dx}(x^n) = nx^{n-1}$.

d. Let's get weird. What if we have a not-quite-power function? Where the thing in the exponent isn't simply a number, but another variable? Let's use the same technique to think about $y = x^x$ and its derivative. First, though, confirm that this is not a power function (and so we cannot use the Power Rule to find the derivative) and is also not an exponential function (and so the derivative isn't itself or itself scaled by a log).

e. Now apply a log to both sides:

$$\ln(y) = \ln(x^x)$$

Rewrite this using the same log property as before, and then use implicit differentiation to find $\frac{dy}{dx}$ or y'.

f. Explain to yourself why logs are friends, especially when trying to differentiate functions in the form of $y = (f(x))^{g(x)}$.

This idea that we've just implemented (applying a logarithm to make some function more friendly and then using implicit differentiation to differentiate) is often called logarithmic differentiation. It works so well because logs are friends.

Wow, So Friendly!

This is incredible! We can now differentiate a whole new class of functions! Functions raised to functions, what could we think of next!

How about products and quotients of functions? I know, I know, we have The Product and Quotient Rules...what about big products and quotients? Annoying ones. Complicated ones.

Activity 3.3.2: Logarithmic Differentiation with Products and Quotients.

Let's say we've got some function that has products and quotients in it. Like, a lot. Consider the function:

$$y = \frac{(x-4)^2\sqrt{3x+1}}{(x+1)^7(x+5)^3}$$

a. Work out a general strategy for how you would find y' directly. Where would you have to use Quotient Rule? What are the pieces? Where would you have to use Product Rule? What are the pieces? Where would you have to use the Chain Rule? What are the pieces? To be clear: do not actually differentiate this. Just look at it in horror and try to outline a plan that some other fool would use.

b. Let's instead use logarithmic differentiation. First, apply the log to both sides to get:

$$\ln(y) = \ln\left(\frac{(x-4)^2\sqrt{3x+1}}{(x+1)^7(x+5)^3}\right)$$

Since this function is just a bunch of products of things with exponents all put into some big quotient, we can use our log properties to rewrite this!

c. We should have:

$$\ln(y) = 2\ln(x-4) + \frac{1}{2}\ln(3x+1) - 7\ln(x+1) - 3\ln(x+5)$$

Confirm this.

d. Now differentiate both sides! You'll have to use some Chain Rule (but not a lot)! Refer back to Fact 3.3.1 for help here.

e. Solve for $\frac{dy}{dx}$ or y'.

f. While this is not a nice looking expression for the derivative, spend some time confirming that this was a nicer process than differentiating directly. This is because logs are friends.

Answer available at the end of this section.

So how do we wrap this up? I hope we can see that logs are a useful and powerful tool: we can use logarithmic differentiation to transform our functions to become "easier to work with" versions of themselves: we put everything on a log-scale and allow our properties of logarithms to make the operations become a bit more accessible.

This is a commonly used trick in many applications of calculus. Specifically, this is used often enough in statistics that there is a whole paradigm of estimation (called Maximum Likelihood Estimation) that uses a log-transformed version of a likelihood function and then applies some basic calculus ideas (that we'll cover in Section 4.5) to perform some powerful estimations.

While I hope that we end up leaving this section knowing that logs are friends, we can probably add a second (and related result) that we're using over and over.

Using the Chain Rule is probably easier than any other option.

We apply logs in order to rewrite these functions in a friendly way because we would rather use the Chain Rule than any combination of other derivative strategies.

Practice Problems

Exercise 3.3.1

For each of the following functions, find the derivatives in two ways: directly, and after using properties of logs or exponentials to rewrite the function and differentiate.

a. $f(x) = \ln\left(\sqrt{5x^2e^x}\right)$

b. $g(x) = \left(e^{5x^3+\ln(x)}\right)^7$

Answer available at the end of this section.

Exercise 3.3.2

Use logarithmic differentiation to find the derivative of each of the following.

a. $y = (4x)^{10x}$

b. $y = \sin(x)^{\cos(x)}$

c. $y = (x+5)^6(x-5)^9(3x+1)^3$

d. $y = \dfrac{(x+2)^{2/3}(x+5)^{7/3}}{(x-4)^{5/3}}$

Answer available at the end of this section.

Exercise 3.3.3

For the function $f(x) = x^2 e^{-x^2}$, find the points where the graph of f has horizontal tangent lines.

Answer available at the end of this section.

Answers

Activity 3.3.2: Logarithmic Differentiation with Products and Quotients.

a.
$$\begin{aligned}\frac{dy}{dx} &= \frac{d}{dx}\left(\frac{(x-4)^2\sqrt{3x+1}}{(x+1)^7(x+5)^3}\right)\\ &= \frac{(x+1)^7(x+5)^3\frac{d}{dx}(\overbrace{(x-4)^2\sqrt{3x+1}}^{\text{Product Rule}}) - (x-4)^2\sqrt{3x+1}\frac{d}{dx}(\overbrace{(x+1)^7(x+5)^3}^{\text{Product Rule}})}{((x+1)^7(x+5)^3)^2}\\ &= \frac{(x+1)^7(x+5)^3\left(2(x-4)\sqrt{3x+1}+\frac{3(x-4)^2}{2\sqrt{3x+1}}\right) - (x-4)^2\sqrt{3x+1}\,(7(x+1)^6(x+5)^3+3(x+1)^7(x+5)^2)}{(x+1)^{14}(x+5)^6}\end{aligned}$$

Exercise 3.3.1

a.
- Solution 1:
$$\begin{aligned}f'(x) &= \frac{1}{\sqrt{5x^2e^x}}\left(\frac{1}{2\sqrt{5x^2e^x}}\right)(10xe^x+5x^2e^x)\\ &= \frac{10xe^x+5x^2e^x}{2\left(5x^2e^x\right)}\\ &= \frac{2+x}{2x}\\ &= \frac{1}{x}+\frac{1}{2}\end{aligned}$$
- Solution 2:
$$\begin{aligned}f(x) &= \ln\left(\left(5x^2e^x\right)^{1/2}\right)\\ &= \frac{1}{2}\ln\left(5x^2e^x\right)\\ &= \frac{1}{2}\ln\left(5x^2\right)+\frac{1}{2}\ln\left(e^x\right)\\ &= \ln(5x)+\frac{1}{2}x\\ &= \ln(5)+\ln(x)+\frac{x}{2}\\ f'(x) &= \frac{1}{x}+\frac{1}{2}\end{aligned}$$

b.
- Solution 1:
$$\begin{aligned}g'(x) &= 7\left(e^{5x^3+\ln(x)}\right)^6\left(e^{5x^3+\ln(x)}\right)\left(15x^2+\frac{1}{x}\right)\\ &= 7\left(e^{5x^3+\ln(x)}\right)^7\left(15x^2+\frac{1}{x}\right)\end{aligned}$$
- Solution 2:
$$\begin{aligned}g(x) &= e^{7(5x^3+\ln(x))}\\ &= e^{35x^3}e^{7\ln(x)}\\ &= e^{35x^3}x^7\\ g'(x) &= e^{35x^3}\left(7x^6\right)+x^7\left(e^{35x^3}\right)\left(105x^2\right)\\ &= 7x^6e^{35x^3}\left(1+15x^3\right)\end{aligned}$$

Exercise 3.3.2

a.
$$\begin{aligned}\ln(y) &= \ln\left((4x)^{(10x)}\right)\\ \ln(y) &= 10x\ln(4x)\\ \frac{d}{dx}(\ln(y)) &= \frac{d}{dx}(10x\ln(4x))\\ \frac{y'}{y} &= 10\ln(4x)+10\\ y' &= (4x)^{10x}(10\ln(4x)+10)\end{aligned}$$

$$\ln(y) = \ln\left((\sin(x))^{\cos(x)}\right)$$
$$\ln(y) = \cos(x)\ln(\sin(x))$$

b. $\frac{d}{dx}(\ln(y))\frac{d}{dx}(\cos(x)\ln(\sin(x)))$

$$\frac{y'}{y} = \frac{\cos(x)}{\sin(x)} - \sin(x)\ln(\sin(x))$$
$$y' = (\sin(x))^{\cos(x)}(\cot(x) - \sin(x)\ln(\sin(x)))$$

$$\ln(y) = \ln\left((x+5)^6(x-5)^9(3x+1)^3\right)$$
$$\ln(y) = 6\ln(x+5) + 9\ln(x-5) + 3\ln(3x+1)$$

c. $\frac{d}{dx}(\ln(y)) = 6\frac{d}{dx}(\ln(x+5)) + 9\frac{d}{dx}(\ln(x-5)) + 3\frac{d}{dx}(\ln(3x+1))$

$$\frac{y'}{y} = \frac{6}{x+5} + \frac{9}{x-5} + \frac{9}{3x+1}$$
$$y' = (x+5)^6(x-5)^9(3x+1)^3\left(\frac{6}{x+5} + \frac{9}{x-5} + \frac{9}{3x+1}\right)$$

$$\ln(y) = \ln\left(\frac{(x+2)^{2/3}(x+5)^{7/3}}{(x-4)^{5/3}}\right)$$
$$\ln(y) = \frac{2}{3}\ln(x+2) + \frac{7}{3}\ln(x+5) - \frac{5}{3}\ln(x-4)$$

d. $\frac{d}{dx}(\ln(y)) = \frac{2}{3}\frac{d}{dx}(\ln(x+2)) + \frac{7}{3}\frac{d}{dx}(\ln(x+5)) - \frac{5}{3}\frac{d}{dx}(\ln(x-4))$

$$\frac{y'}{y} = \frac{2}{3(x+2)} + \frac{7}{3(x+5)} - \frac{5}{3(x-4)}$$
$$y' = \frac{(x+2)^{2/3}(x+5)^{7/3}}{(x-4)^{5/3}}\left(\frac{2}{3(x+2)} + \frac{7}{3(x+5)} - \frac{5}{3(x-4)}\right)$$

Exercise 3.3.3

$$\ln(f(x)) = \ln\left(x^2e^{-x^2}\right)$$
$$\ln(f(x)) = 2\ln(x) - x^2$$
$$\frac{d}{dx}(\ln(f(x))) = 2\frac{d}{dx}(\ln(x)) - \frac{d}{dx}\left(x^2\right)$$
$$\frac{f'(x)}{f(x)} = \frac{2}{x} - 2x$$
$$f'(x) = x^2e^{-x^2}\left(\frac{2}{x} - 2x\right)$$

We can see that $f'(x) = 0$ when x^2, e^{-x^2} and $\left(\frac{2}{x} - 2x\right)$ are 0. Note that e^{-x^2} is only ever positive, and so is never 0. while $x^2 = 0$ when $x = 0$.

$$\frac{2}{x} - 2x = 0$$
$$2 - 2x^2 = 0$$
$$2(1-x)(1+x) = 0$$
$$x = \pm 1$$

Note that $f'(-1) = 0, f'(0) = 0$, and $f'(1) = 0$, and so these are the x-values where $f(x)$ has horizontal tangent lines.

This page titled 3.3: Logarithmic Differentiation is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Peter Keep.

4: APPLICATIONS OF DERIVATIVES

4.1: Mean Value Theorem

4.2: Increasing and Decreasing Functions

4.3: Concavity

4.4: Global Maximums and Minimums

4.5: Optimization

4.6: Linear Approximations

4.7: L'Hôpital's Rule

4.8: Exploration- Applications of Derivatives

This page titled 4: Applications of Derivatives is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Peter Keep.

4.1: Mean Value Theorem

Let's begin here with some weird questions. The questions aren't weird because of what they're asking. Instead, they're weird because of the logic of how we interpret them compared to how we want to interpret them.

1. Why is the derivative of a constant function 0?
2. Why do $y = x^2 + 7$ and $y = x^2 - 3$ have the same derivative?
3. If a function is only increasing on the interval (0, 1), what do we know about the derivative at any of these x-values in (0, 1)?

These questions are ones we can think through and answer.

Aside.

Here are some answers for these first three questions:

1. A constant function has all of the same y-values for any x-value in the domain: of course the slope everywhere is 0! There isn't any change in the y-values!
2. We can differentiate these functions term by term: we know that the x^2 term has a derivative of $2x$, and then for each function's constant, the derivative has to be 0. So it doesn't matter what each constant was, the derivatives wouldn't rely on that value!
3. If a function is increasing on an interval, then it means that for any pair of x-values, $x_1 < x_2$, we get y-values in the same order: $f(x_1) < f(x_2)$. If we find the limit of slopes as $x_1 \to x_2$, we'll always get a positive slope, so the derivative has to be positive.

What's tricky is that these don't always say what we want to say. For instance, I might sometimes want to say the following:

1. If you know a function's derivative is 0, then you know the function is constant. Another way of saying this is that the only functions whose derivative is 0 are constant functions.
2. Similarly, we might want to say that if you know two functions that have the same derivative, then the only difference between the functions is a constant. There aren't any other ways for functions to be different with the same derivative.
3. We might want to say that if you know the derivative is positive on an interval, that means that the function has to be increasing.

Do you see the (slight) difference?

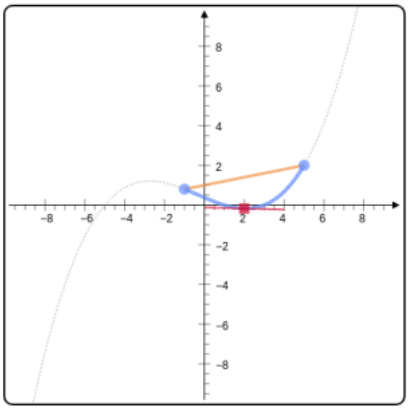
What we'll secretly see is that all of these statements are actually correct, but require a result for us to say them. The Mean Value Theorem will be the result that we use to build important and helpful results throughout the rest of this course.

Slopes

We have two different kinds of slopes that we think of right now: a slope between two points, and a slope at a single point.

We can translate this into a rate of change! We think of two rates of change: an average rate of change on some interval and the instantaneous rate of change at some point.

We will try to connect these two different slopes/rates of change for "wellbehaved" functions. Let's take a look at an example. In the graph below, we have a curve where we are considering some closed interval, as well as a point within that interval. Both slopes are visualized and calculated: the slope between the ending points of the interval is the average slope, while the slope of the line tangent to the curve at a point is the instantaneous slope. Move the point/interval around and get a feel for how these two different slopes relate (or don't relate!) to each other.

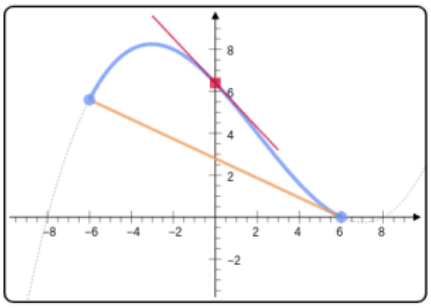


The average slope on the interval is: 0.2. The instantaneous slope at the point is: -0.025.

Standalone
Embed

If you move the interval/point around enough, you'll see that sometimes these two slopes are really similar and sometimes they're very different. But what if the point in the middle of the interval wasn't so "set" at being stuck in the exact middle of the interval? What if we stuck the interval in place, and then had the freedom to move the point anywhere in the interval?

Instructions: Move the point in the interval and try to make the slope of the line tangent to the curve at the point match the slope of the line connecting the endpoints of the interval.



You should notice when you find the point where the tangent line is parallel to the line connecting the endpoints.

Standalone
Embed

Do you think we will always be able to do this? What kinds of functions will have these points where the slope at the point matches the average slope on the interval?

The Mean Value Theorem

Theorem 4.1.1: Mean Value Theorem.

For a function $f(x)$ that is continuous on the closed interval $[a, b]$ and differentiable on the open interval (a, b), then there is some value $x = c$ with $a < c < b$ where:

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

This theorem is really just guaranteeing the existence of at least one of the points we found above: a point where the slope of the line tangent to the curve is the same as the slope between the endpoints of the interval. We can (and, very briefly, we will) use this theorem to find the point that is guaranteed to exist, but we will, more generally, use this theorem as a tool for connection.

We will try to use it as a way of connecting derivatives to the behavior of a function. The Mean Value Theorem gives us this equation where on one side we have a derivative, and on the other side we have function outputs. This is really similar to the definition of the Derivative at a Point, except that in this case we have no limit: we just get to use the behavior of the function on an interval around the point.

Secretly, the Mean Value Theorem is the driving force behind most of the results we will look at from here on out, at least when the requirements include continuity and differentiability on an interval. You can almost guarantee that if a theorem or result has these two requirements, then the Mean Value Theorem is likely at work.

Let's look at one example, at least, before we move on.

Example 4.1.2

Let's say that a person is planning on biking to their college campus, 8 miles away. At 2:39 p.m., they send a text to their friend with a selfie of them on their bike about to start their trip, captioned "Beautiful day for a ride!" At 3:27 p.m., they post a picture on their social media of them in front of the bike rack on campus.

a. What was the average velocity of the student between 2:39 p.m. and 3:27 p.m.?
b. The Mean Value Theorem connects average slopes with slopes of tangent lines. What does that mean, in general, in this context?
c. Make a claim about the instantaneous velocity of the student on their bike at some point in time.

Solution

a. The total trip took 48 minutes (or 0.8 hours), and the student traveled a total distance of 8 miles. The student's average velocity is $\frac{8}{0.8} = 10$ miles per hour.
b. In this case, the average rate of change of the function on the interval is the average velocity of the biker. So then the instantaneous rate of change must correspond with an instantaneous velocity, of their velocity at some specific point in time.
c. We know that at some point between 2:39 p.m. and 3:27 p.m., the cyclist had to be traveling at exactly 10 miles per hour.

Example 4.1.3

For a function $f(x) = \sqrt{x} + 1$ on the interval [1, 4], find the point that the Mean Value Theorem guarantees the existence of, and explain what it is.

Solution

Let's calculate the average slope on the interval:

$$\begin{aligned} \frac{f(4) - f(1)}{4 - 1} &= \frac{(\sqrt{4} + 1) - (\sqrt{1} + 1)}{3} \\ &= \frac{1}{3} \end{aligned}$$

So we know that there is some x-value between 1 and 4 where $f'(x) = \frac{1}{3}$.

The derivative is $f'(x) = \frac{1}{2\sqrt{x}}$, so we can solve for x:

$$f'(x) = \frac{1}{3}$$
$$\frac{1}{2\sqrt{x}} = \frac{1}{3}$$
$$2\sqrt{x} = 3$$
$$\sqrt{x} = \frac{3}{2}$$
$$x = \frac{9}{4}$$

So the point guaranteed to exist by the Mean Value Theorem is located at $\left(\frac{9}{4}, \frac{5}{2}\right)$, where the slope of the tangent line is $f'\left(\frac{9}{4}\right) = \frac{1}{3}$.

These examples are fine, but the real power of the Mean Value Theorem comes in how we can use it to generate more interesting results. Let's check those out!

More Results due to the Mean Value Theorem

First, let's remind ourselves of what it means for a function to be increasing or decreasing on an interval.

Definition 4.1.4: Increasing and Decreasing on an Interval.

A function $f(x)$ is increasing on the interval (a, b) if, for every pair of x-values x_1 and x_2 (with $x_1 < x_2$), $f(x_1) < f(x_2)$.

If $f(x_1) > f(x_2)$, then we say that f is **decreasing** on the interval (a, b).

Note that we classify a function as increasing or decreasing based on comparing two function outputs. This is a perfect time to use the Mean Value Theorem, since it can connect these pairs of function outputs with a derivative!

Theorem 4.1.5: Sign of the Derivative and Direction of a Function.

If f is a function that is differentiable on the interval (a, b) and $f'(x) > 0$ for all x in the interval (a, b), then f must be increasing on (a, b). Similarly, if $f'(x) < 0$ for all x in the interval (a, b), then f must be decreasing on (a, b).

Proof. available at the end of this section.

We'll use this result pretty extensively in the next couple of topics. It is helpful for us to use this to connect the signs of a derivative with our intuition about what that must mean for the direction of a function.

Theorem 4.1.6: If a Function's Derivative is 0, it's Constant.

If $f(x)$ is a function defined on (a, b) with $f'(x) = 0$ for all x-values in the interval (a, b), then $f(x)$ is a constant function.

Proof. available at the end of this section.

Theorem 4.1.7: Equal Derivatives Correspond with Related Functions.

For two functions f and g, both differentiable on (a, b), if $f'(x) = g'(x)$ for all x-values on (a, b), then we know that $f(x) = g(x) + C$ for some real number constant C. That is, the only differences in f and g are due to a difference in the constant term.

Proof. available at the end of this section.

We'll use this result a bit later on, but it's a great example of how we can use the Mean Value Theorem to connect information about the derivative to information about how a function might work.

Let me interject my own opinion here: I think the Mean Value Theorem is extremely important. But I also don't think that it is that important for you to use.

I think you should know the general idea of connecting the slopes of the line tangent to the curve and the average slope on an interval. I think you should remember the picture we looked at. I think you should be content to see some results later on in the course and smile fondly about how the Mean Value Theorem is under the surface, churning away and getting us cool results to think about. I think we should be happy that the Mean Value Theorem is here and we can recognize it as, maybe, the most important result in this textbook.

But we don't need to pretend that we need to actually use it directly for computations or anything like that. We'll just know it.

Practice Problems

Exercise 4.1.1

For each of the following scenarios, explain whether the Mean Value Theorem applies. If it does, explain what it tells us.

a. A student in a class is distracted by someone else tapping their fingers on their desk. The student decides to count the taps. In the 3 minutes that they counted, the other student tapped on their desk 245 times. **Hint:** Is there a continuous and differentiable function somewhere in this description?
b. On a sunny, Chicago day in May, the temperature was 58° at 7:15am. Later on, at 3:45pm, the temperature was 79°. **Hint:** What is changing here? Can you calculate an average rate of change? What does the Mean Value Theorem say should be happening?
c. In a small oceanside town, residents notice that the tide goes out and in every day. High tide is at 10am, and then again at 10pm, and, wow this is pretty incredible, the water level is exactly the same at both times. **Hint:** What's the average rate of change if the water level is the same at both times? What does this mean about the rate of change of the water level somewhere between those times?

Exercise 4.1.2

For each function:

- Sketch (or look up) a graph of the curve on the interval.
- Calculate the slope of the line connecting the end points.
- Find the location of any points that exist from the Mean Value Theorem.

a. $f(x) = x^2 - 4x + 1$ on the interval [0,3].
b. $g(x) = \dfrac{1}{\sqrt{x-3}}$ on the interval [4,7].
c. $r(x) = \sin(x)\cos(x)$ on the interval $\left[0, \dfrac{\pi}{2}\right]$.

Answers

Theorem 4.1.5**: Sign of the Derivative and Direction of a Function.**

Before we begin, let's just agree to think about only the case where $f'(x) > 0$ on the interval (a, b). The other case (where $f'(x) < 0$) ends up being the exact same argument, with some changes in signs and directions of inequalities. So we'll say $f'(x) > 0$ for all x-values in the interval (a, b).

Ok, let's begin!

Let's pick two arbitrary x-values from the interval (a, b). Call them x_1 and x_2, and we'll make sure that we name them in a way where $x_1 < x_2$. Now, f must be continuous on $[x_1, x_2]$ and differentiable on (x_1, x_2). We also know that $f'(x) > 0$ for every x in the interval $(x_1 < x < x_2)$.

The Mean Value Theorem says that there is some $x = c$ in (x_1, x_2) with

$$f'(c) = \frac{f(x_2) - f(x_1)}{x_2 - x_1}$$

Equivalently, this means that

$$f'(c)(x_2 - x_1) = f(x_2) - f(x_1)$$

Notice that $f'(c) > 0$ (since the derivative is positive everywhere in the interval) and $(x_2 - x_1) > 0$ (by the way we named these x-values). This means that $f'(c)(x_2 - x_1) > 0$, and so $(f(x_2) - f(x_1)) > 0$ as well.

So notice what just happened: we arbitrarily chose x-values x_1 and x_2 and noticed that for any of these pairs where $x_1 < x_2$, we found that $f(x_1) < f(x_2)$. This is exactly what it means for f to be increasing on the interval (a, b) (based on Definition 4.1.4).

Theorem 4.1.6**: If a Function's Derivative is 0, it's Constant.**

We can almost exactly replicate the proof from Theorem 4.1.5 here!

Let's pick two arbitrary x-values from the interval (a, b). Call them x_1 and x_2, and we'll again make sure that we name them in a way where $x_1 < x_2$. Now, f must be continuous on $[x_1, x_2]$ and differentiable on (x_1, x_2). This time, we know that $f'(x) = 0$ for every x in the interval $(x_1 < x < x_2)$.

The Mean Value Theorem says that there is some $x = c$ in (x_1, x_2) with

$$f'(c) = \frac{f(x_2) - f(x_1)}{x_2 - x_1}$$

Equivalently, this means that

$$f'(c)(x_2 - x_1) = f(x_2) - f(x_1)$$

Notice that $f'(c) = 0$ (since the derivative is zero everywhere in the interval). This means that $f'(c)(x_2 - x_1) = 0$, and so $(f(x_2) - f(x_1)) = 0$ as well.

This means that every y-value from the function is the same as every other one, since we picked these points arbitrarily. So f must be a constant function!

Theorem 4.1.7**: Equal Derivatives Correspond with Related Functions.**

This one is pretty easy to prove: we're going to use a fun little trick where we connect this theorem to Theorem 4.1.6.

Let's think about these two functions f and g with $f'(x) = g'(x)$, and let's think about a function $h(x) = f(x) - g(x)$. Now we can notice that $h'(x) = f'(x) - g'(x)$ which has to be 0.

Oh wow, $h(x)$ is a function with $h'(x) = 0$ on the interval (a, b). It must be a constant function (based on Theorem 4.1.6)! Let's call it $h(x) = C$, where C is some real number constant.

This means that $f(x) - g(x) = C$, and we can see that the only difference between these two functions is a constant.

This page titled 4.1: Mean Value Theorem is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Peter Keep.

4.2: Increasing and Decreasing Functions

Activity 4.2.1: How Should We Think About Direction?

Our goal in this activity is to motivate some new terminology and results that will help us talk about the "direction" of a function and some interesting points on a function (related to the direction of a function). For us to do this, we'll look at some different examples of functions and try to think about some unifying ideas

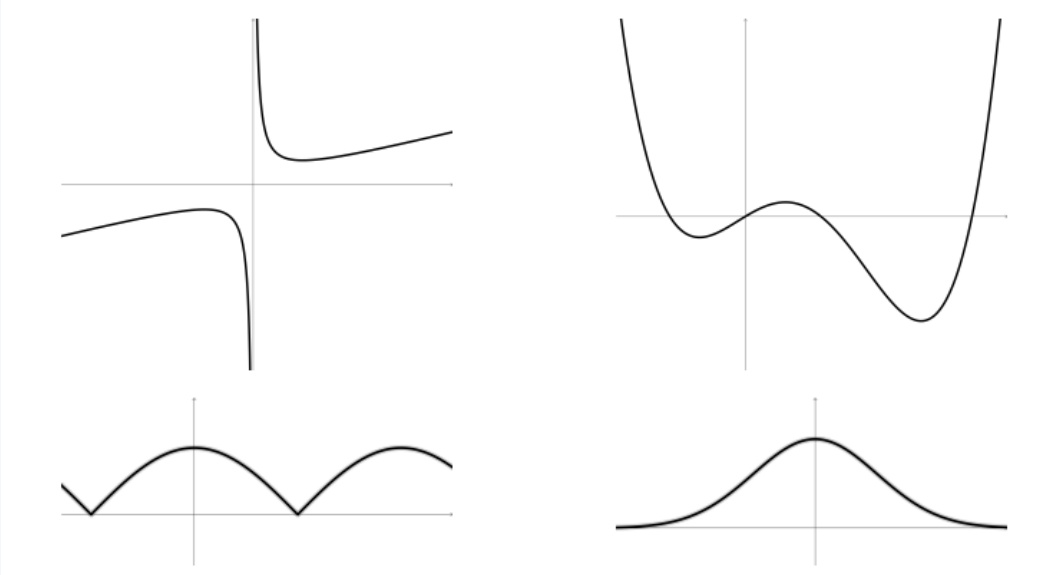


Figure 4.2.1

These examples do not cover all of the possibilities of how a function can act, but will hopefully provide us enough fertile ground to think about some different situations.

a. In each graph, find and identify:
   - The intervals where the function is increasing.
   - The intervals where the function is decreasing.
   - The points (or locations) around and between these intervals, the points where the function changes direction or the direction terminates.
b. Make a conjecture about the behavior of a function at any point where the function changes direction. **Hint:** What do you think has to be true about the derivative at these points?
c. Look at the highest and lowest points on each function. You can even include the points that are highest and lowest just compared to the points around it. Make a conjecture about the behavior of the function at these points. **Hint:** What do you think has to be true about the derivative at these points?

We want to turn this little bit of thinking and exploring into some useful definitions for us. To craft these definitions, we need to start with thinking about what we care about and why we might care about it.

Critical Points, Local Maximums, and Local Minimums

Let's start by saying what we're really looking for is the highest and lowest points on a function. These points are interesting, have useful applications, and are difficult to find in general without calculus. We hopefully noticed, though, that these points always end up showing up at the same kinds of locations! They're points where the direction of a function changes (or terminates).

We also noticed that there are some common characteristics of those points. They're points where the derivative was either 0 or

didn't exist. So we'll start by defining these points, and then we'll define the idea of "highest" and "lowest" points. Then we'll put together a result that we hopefully noticed here: that the highest and lowest points show up at these points where the derivative is 0 or doesn't exist.

Definition 4.2.2: Critical Point of a Function.

We say that a point $(c, f(c))$ on the graph of $y = f(x)$ is a **critical point** of the function f if $f'(c) = 0$ or $f'(c)$ doesn't exist. If $(c, f(c))$ is a critical point of f, then we sometimes will call $x = c$ a **critical number** and $y = f(c)$ a **critical value**.

So these are the points we will look for to find the "highest" and "lowest" points.

Now we need to define this idea so that we can build the result that tells us how to find these highest and lowest points.

Definition 4.2.3: Local Maximum/Minimum.

A point $(c, f(c))$ is a local maximum of $f(x)$ if there is some open interval of real numbers (a, b) around $x = c$ (that is, $a < c < b$) and $f(c) \geq f(x)$ for all x-values in the intersection of (a, b) and the domain of f.

Similarly, a point $(c, f(c))$ is a local minimum of $f(x)$ if there is some open interval of real numbers (a, b) around $x = c$ (that is, $a < c < b$) and $f(c) \leq f(x)$ for all x-values in the intersection of (a, b) and the domain of f.

These are really just slightly technical ways of saying that $f(c)$ is either the highest or lowest y-value produced by the function f for the x-values near $x = c$.

Aside.

If you look around online, or in other textbooks, you'll find different definitions of these same kinds of points. Some of those definitions have silly stipulations, like saying that an ending point of a function cannot be a local maximum/local minimum.

This seems to come from some sense that the derivative must be defined on each "side" of a local max/min. In this book, we'll not worry about this restriction, and instead just look at the highest and lowest points relative to the other points near it.

Now we want to build the connection between these points. In Activity 4.2.1, we pointed out that the highest and lowest points on a function all had the common theme of showing up at places where the derivative was 0 or didn't exist.

Theorem 4.2.4: Every Local Maximum/Minimum Occurs at a Critical Point.

Every local maximum or local minimum of f occurs at a critical point of f.

Another way of saying this is that if $(c, f(c))$ is a local maximum or local minimum of f, then it must also be a critical point of f.

WAIT! STOP! Before you move on, let's make sure we understand what this theorem is saying. Or maybe what this theorem is not saying.

Notice that we are not saying that every critical point is a local maximum or local minimum! This is a classic "every square is a rectangle but not every rectangle is a square" situation.

Every local maximum/minimum occurs at a critical point, but not every critical point is a local maximum/minimum.

Direction of a Function (and Where it Changes)

Let's build up a way of classifying critical points as local maximums, local minimums, or neither.

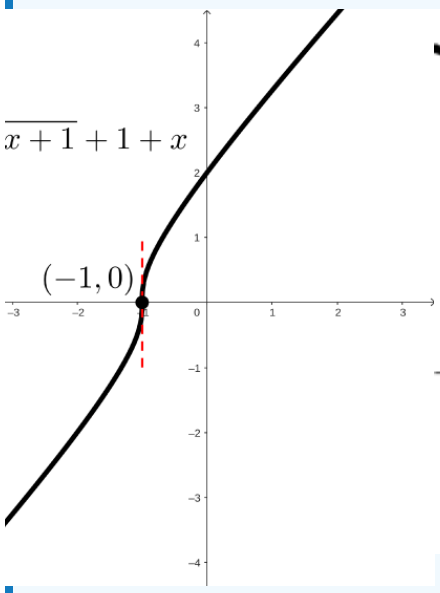
Activity 4.2.2: Comparing Critical Points.

Let's think about four different functions:

- $f(x) = 4 + 3x - x^2$
- $g(x) = \sqrt[3]{x+1} + 1 + x$
- $h(x) = (x-4)^{2/3}$
- $j(x) = 1 - x^3 - x^5$

Our goal is to find the critical points on the interval $(-\infty, \infty)$ and then to try to figure out if these critical points are local maximums or local minimums or just points that the function increases or decreases through.

a. To start, we're going to be finding critical points. Without looking at a picture of the graph of the function, find the derivative. Are there any x-values (in the domain of the function) where the derivative doesn't exist? We are normally looking for things like division by 0 here, but we could be finding more than that. Check out When Does a Derivative Not Exist? to remind yourself if needed. Are there any x-values (in the domain of the function) where the derivative is 0?
b. Now that we have the critical points for the function, let's think about where the derivative might be positive and negative. These will correspond to the direction of a function, based on Theorem 4.1.5 Sign of the Derivative and Direction of a Function. Write out the intervals of x-values around and between your list of critical points. For each interval, what is the sign of the derivative? What do these signs mean about the direction of your function?
c. Without looking at the graph of your function:
    - What changes about how your function increases up to or decreases down to a critical point based on whether the derivative was 0 or the derivative didn't exist?
    - Does your function change direction at a critical point? What will that look like, whether it does or does not change direction?
d. Give a basic sketch of your graph. It might be helpful to find the y-values for any critical points you've got. Then you can sketch your function increasing/decreasing in the intervals between these points. In your sketch, include enough detail to tell whether the derivative is 0 or doesn't exist at each critical point.
e. Compare your sketch to the actual graph of the function (you can find all of the graphs in the hint). **Hint**:



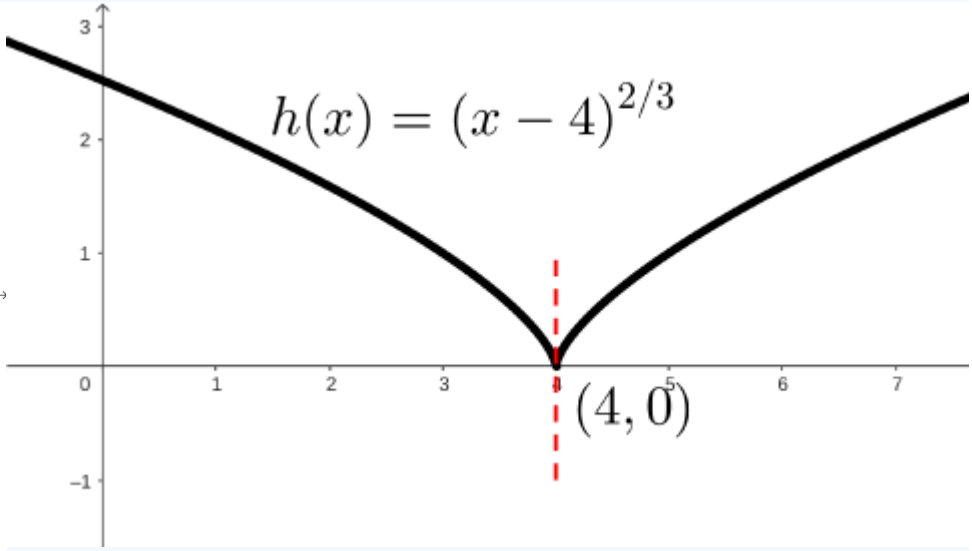


Figure **4.2.5.**

This is great, we have a nice strategy for thinking about critical points!

Something we can notice in finding these critical points (as well as thinking about the domain of the function), we found all of the locations where the derivative is both not positive and not negative. This is a weird way of saying that all of the intervals in between the critical points we found and any breaks in the domain of the function (like if there were vertical asymptotes or holes or something) are places where the derivative is positive or negative.

Even more exciting: if the derivative function we found is continuous, then the Intermediate Value Theorem says that it will only change signs at these critical points (or places like vertical asymptotes or holes). So this means that we can always construct a little chart or something, think about the x-values around and at critical points or other breaks in the domain, and then look at what the derivative does as we move through those intervals and x-values.

This will serve as a nice way of thinking about what's going on with our functions!

Theorem 4.2.5: First Derivative Test.

If $(c, f(c))$ is a critical point of f and we can evaluate the derivative f' on either side of this point, then we can use the signs of the first derivative to classify the critical point:

- If the sign of f' changes from positive to negative as x passes through $x = c$, then $(c, f(c))$ is a local maximum.
- If the sign of f' changes from negative to positive as x passes through $x = c$, then $(c, f(c))$ is a local minimum.
- If the sign of f' does not change as x passes through $x = c$, then the function f increases or decreases (depending on whether $f' > 0$ or $f' < 0$) through $x = c$. We will often lay these results out in a chart or table, like the following:

x		c	
f'	$\oplus$		$\ominus$
f	$\nearrow$		$\searrow$
		local max	

x		c	
f'	$\ominus$		$\oplus$
f	$\searrow$		$\nearrow$
		local min	

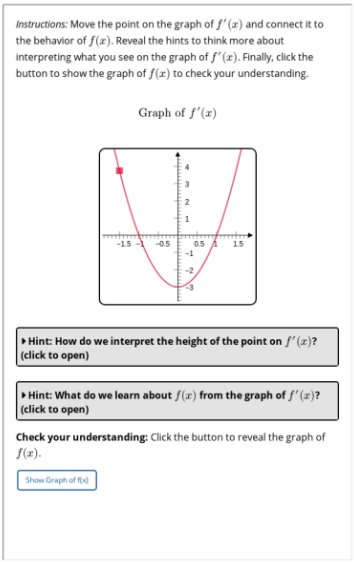
x		c	
f'	$\oplus$		$\oplus$
f	$\nearrow$		$\nearrow$
		increasing through	

x		c	
f'	$\ominus$		$\ominus$
f	$\searrow$		$\searrow$
		decreasing through	

Using the Graph of the First Derivative

Activity 4.2.3: First Derivative Test Graphically.

Let's focus on looking at a picture of a derivative, $f'(x)$, and trying to collect information about the function $f(x)$. This is what we've done already, except that we've done it by thinking about the representation of $f'(x)$ as a function rule written out with algebraic symbols. Here we'll focus on connecting all of that to the picture of the graphs. For all of the following questions, refer to the plot below. You can add information with the hints whenever you need to. Don't reveal the picture of $f(x)$ until you're really ready to check what you know.



Standalone
Embed

a. Based on the graph of $f'(x)$, estimate the interval(s) of x-values where $f(x)$ is increasing.
b. Based on the graph of $f'(x)$, estimate the interval(s) of x-values where $f(x)$ is decreasing.
c. Find the x-values of the critical points of $f(x)$. Once you've estimated these, classify them as local maximums, local minimums, or neither. Explain your reasoning.
d. What do you think the graph of $f(x)$ looks like? Do your best to sketch it or explain it before revealing it!
e. Why could we estimate the x-values of the critical numbers of $f(x)$, but not find the actual coordinates? How come we can't find the y-value based on looking at the graph of $f'(x)$?

Reading the graphs of functions is, in general, an important skill. But it's an especially important idea to be able to read and understand the graph of a function like a derivative and then interpret what we are seeing into some other context.

So for us to really excel here, we'll want to focus on the fact that a first derivative tells about the slope or direction of a function. Whatever y-values we find on the graph of a $f'(x)$ needs to be interpreted as a slope or rate of change of $f(x)$. Then we can string these slopes or rates of changes together to try to think about the behavior of $f(x)$ by knowing how the y-values are changing as we move along the curve of $y = f(x)$!

Strange Domains

We'll look at two more examples, both of them using functions whose domain is not $(-\infty, \infty)$.

Example 4.2.6

a. Consider the function

$$f(x) = \frac{x^2}{(x-3)^2}$$

Find the domain of f, the critical points of f, and then the intervals where f is increasing/decreasing. Then, classify any critical points local maximums/minimums if necessary.

Hint 1. $f'(x) = -\dfrac{6x}{(x-3)^3}$

Hint 2. The function $f(x)$ has one critical point at (0, 0). Why isn't there a critical point at $x = 3$? What is happening there instead?

Hint 3.

x	$(-\infty, 0)$	0	(0,3)	3	$(3, \infty)$
f'					
f					

b. Consider the function

$$g(x) = \sqrt{x} - x + 1$$

Find the domain of g, the critical points of g, and then the intervals where g is increasing/decreasing. Then, classify any critical points local maximums/minimums if necessary.

Hint 1. $g'(x) = \dfrac{1}{2\sqrt{x}} - 1$

Hint 2. Notice that, by our definition of critical points, both (0, 1) and $\left(\frac{1}{4}, \frac{3}{4}\right)$ are critical points.

Hint 3.

x	0	$\left(0, \frac{1}{4}\right)$	$\frac{1}{4}$	$\left(\frac{1}{4}, \infty\right)$
g'				
g				

Solution

a. The domain of $f(x)$ is $(-\infty, 3) \cup (3, \infty)$ due to the vertical asymptote at $x = 3$. The only critical point is at (0, 0). The table below shows where f is increasing and decreasing, as well as any local maximums or minimums.

x	$(-\infty, 0)$	0	(0,3)	3	$(3, \infty)$
f'	$\ominus$	0	$\oplus$		$\ominus$
f	$\searrow$ decreasing	(0,0) local min	$\nearrow$ increasing	asymptote	$\searrow$ decreasing

b. The domain of $g(x)$ is $[0, \infty)$. There are two critical points: one at (0, 1) and another at $\left(\frac{1}{4}, \frac{3}{4}\right)$. The table below shows where g is increasing and decreasing, as well as any local maximums or minimums.

x	0	$\left(0, \frac{1}{4}\right)$	$\frac{1}{4}$	$\left(\frac{1}{4}, \infty\right)$
g'	DNE	$\ominus$	0	$\oplus$
g	(0,1) local max	\(\searrow\) decreasing	$\left(\frac{1}{4}, \frac{3}{4}\right)$ local min	$\nearrow$ increasing

So we have two things to notice:

1. When we have some gap or missing spot in the domain of a function, that can still divide up the intervals where our function is increasing or decreasing! We should notice, though, that since this isn't actually a point on the curve of our function, it won't be a critical point and so we have to interpret it differently: we can't use the First Derivative Test!

2. An ending point of an interval is a location where the derivative cannot exist! We could define a one-sided derivative (similar to how we defined one-sided continuity in Definition 1.6.2), but for now we'll just say that the derivative doesn't exist, and call those ending points critical points. That means that depending on the direction that a function goes away (or leading up to) that ending point, we can classify it as a local maximum or minimum.

Lastly, just a couple of notes: in these little tables or charts (sometimes called sign charts, since they are showing the signs of the derivative), we'll use some shorthand notation. In Example 4.2.7, we used "DNE" to mean that a derivative "does not exist" at a point. Similarly, we used ff to represent the vertical asymptote at that x-value (so that we didn't accidentally think it was a local maximum or minimum based on the signs of the derivative around it).

Moving forward, we'll use this same kind of analysis to think about how the derivative might be changing on these intervals. This rate of change of the slopes, the second derivative, will be a useful tool for gathering more information about how a function might be acting.

Practice Problems

Exercise 4.2.1

Explain what a critical number of $f(x)$ is. What are the things we're looking for

Exercise 4.2.2

For the following functions, find the critical numbers and explain each one. Is it one where the derivative is 0, or one where the derivative doesn't exist? Feel free to confirm these by graphing the function!

a. $f(x) = x^2 + 5x - 4$ on [−4, 2]
b. $f(x) = \sqrt[3]{x+4}$ on [−6, −2]
c. $f(x) = \dfrac{x^2+1}{x-3}$
d. $f(x) = x^2 e^{-x}$
e. $f(x) = x^2 e^{-x}$ on [1, 10]
f. $f(x) = \ln\left(x^4 + x + 1\right)$

Exercise 4.2.3

Find the critical numbers of the function $f(x)$, graphed below. Explain each one.

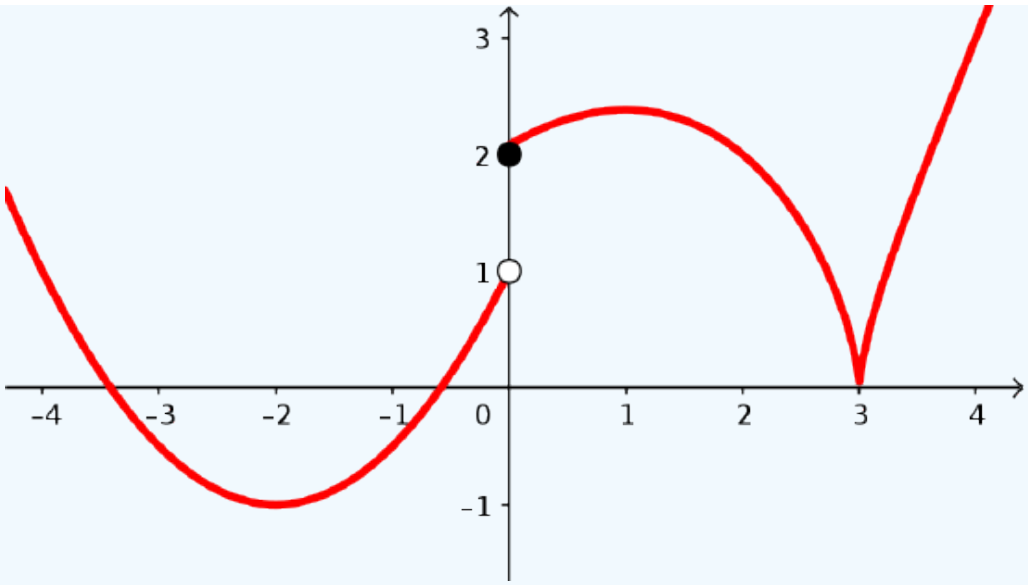


Figure 4.2.7 Graph of $f(x)$.

Exercise 4.2.4

Give an example of a function that has $f'(x) > 0$ on $(-\infty, 4)$ and $f'(x) < 0$ on $(4, \infty)$. Show graphically where the function is increasing and decreasing. Explain how to find out the location of a local maximum or minimum for your example.

Exercise 4.2.5

Sketch a graph of a function that increases everywhere, but has $f'(4) = 0$. Explain what sign the derivative is when $x \neq 4$.

Exercise 4.2.6

For each of the following functions, use the first derivative to find the intervals where the function is increasing or decreasing.

a. $f(x) = x^2 e^x$
b. $f(x) = x \ln(x) - x$
c. $f(x) = \dfrac{1}{x^2}$
d. $f(x) = \ln\left(x^4 + 1\right)$
e. $f(x) = \dfrac{x-4}{x^2+1}$
f. $f(x) = e^{-x^2-2x}$
g. $f(x) = \dfrac{\sqrt{x}}{e^x}$
h. $f(x) = \sin(x)$ on $\left[0, \dfrac{\sqrt{\pi}}{2}\right]$
i. $f(x) = \cos^3(x)$ on $(-2\pi, 2\pi)$

Exercise 4.2.7

Explain how the First Derivative Test helps us classify critical numbers as local maximums or minimums.

Exercise 4.2.8

Find any local maximums or local minimums for each of the following functions. Note that these are the same functions that you've already worked with in Practice Problem 6.

a. $f(x) = x^2 e^x$
b. $f(x) = x\ln(x) - x$
c. $f(x) = \dfrac{1}{x^2}$
d. $f(x) = \ln\left(x^4 + 1\right)$
e. $f(x) = \dfrac{x-4}{x^2+1}$
f. $f(x) = e^{-x^2-2x}$
g. $f(x) = \dfrac{\sqrt{x}}{e^x}$
h. $f(x) = \sin(x)$ on $\left[0, \dfrac{\sqrt{\pi}}{2}\right]$
i. $f(x) = \cos^3(x)$ on $(-2\pi, 2\pi)$

Exercise 4.2.9

Sketch a graph of a function $f(x)$ that is continuous on $(-\infty, \infty)$ and has the following characteristics. Label the intervals where f is increasing/ decreasing, and find the locations of any local maximums/minimums.

- $f'(x) > 0$ on $(-\infty, 3)$
- $f'(3)$ does not exist
- $f'(x) < 0$ on (3, 6)
- $f'(6) = 0$
- $f'(x) < 0$ on $(6, \infty)$

Exercise 4.2.10

Sketch a graph of a function $f(x)$ and has the following characteristics. Label the intervals where f is increasing/decreasing, and find the locations of any local maximums/minimums.

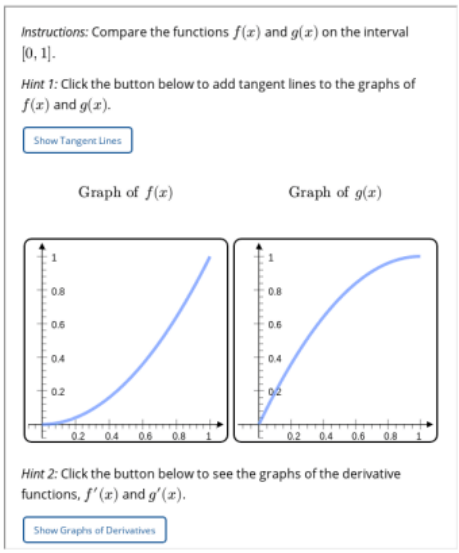
- $f'(x) < 0$ on $(-\infty, 3)$
- $f'(3) = 0$
- $f'(x) > 0$ on (3, 6)
- $f(x)$ has a vertical asymptote at $x = 6$
- $f'(x) < 0$ on $(6, \infty)$

This page titled 4.2: Increasing and Decreasing Functions is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Peter Keep.

4.3: Concavity

Activity 4.3.1: Compare These Curves.

a. Consider the two curves pictured below. Compare and contrast them. What characteristics of these functions are the same? What characteristics of these functions are different?



Standalone Embed

b. Explain the similarities you found by only talking about the slopes of each function (the values of $f'(x)$ and $g'(x)$).
c. Explain the differences you found by only talking about the slopes of each function (the values of $f'(x)$ and $g'(x)$).
d. Make a conjecture about the rate of change of both f' and g'. We'll call these things second derivative functions, $f''(x)$ and $g''(x)$.

Defining the Curvature of a Curve

Definition 4.3.1: Concavity and Inflection Points.

We say that a function is concave up on an interval (a, b) if $f'(x)$ is increasing on the interval. If $f'(x)$ is decreasing on the interval, then we say that $f(x)$ is concave down.

We say that a point $(c, f(c))$ is an inflection point if it is a point at which f changes concavity (from concave up to concave down or vice versa).

Note that when we think about a function being concave up or down on some interval, we can think about this in different ways. Curvature can sometimes be hard to recognize visually, but one of the things we can see from the visual above is the interaction between the tangent line and the curve:

- If the function is concave up on some interval, then the tangent line sits below the function at every point on that interval. The function curves upward away from the tangent line. Sometimes people will say that the curvature is concave "up, like a cup."
- If the function is concave down on some interval, then the tangent line sits above the function at every point on that interval. The function curves downward away from the tangent line. Sometimes people will say that the curvature is concave "down, like a frown."

So we have some visual ways of thinking about these different types of curvature. Annoyingly, though, it is still relatively difficult

to pinpoint the exact (or even close) location of an inflection point. We might be able to pretty easily point at the locations of local maximums and local minimums on a graph of a function, but it can be hard to see the exact point at which a graph of a function changes from concave up to down or vice versa. We'll focus on finding them algebraically first, but then we'll think a bit more about the graphs of functions later.

Activity 4.3.2: Finding a Function's Concavity.

We're going to consider a pretty complicated function to work with, and employ a strategy similar to what we did with the first derivative:

- Our goal is to find the sign of $f''(x)$ on different intervals and where that sign changes.
- To do' this, we'll find the places that $f''(x) = 0$ or where $f''(x)$ doesn't exist. These are the critical points of $f'(x)$.
- From there, we can build a little sign chart, where we split up the x-values based on the domain of f and the critical numbers of f'. Then we can attach each interval of x-values with a sign of the second derivative, f'', on that interval.
- We'll interpret these signs. For any intervals where $f''(x) > 0$, we know that $f'(x)$ must be increasing and so $f(x)$ is concave up. Similarly, for any intervals where $f''(x) < 0$, we know that $f'(x)$ must be decreasing and so $f(x)$ is concave down.

Let's consider the function

$$f(x) = \ln\left(x^2 + 1\right) - \frac{x^2}{2}$$

a. Find the first derivative, $f'(x)$, and use some algebra to confirm that it is really:

$$f'(x) = -\frac{x(x+1)(x-1)}{x^2+1}$$

While we have this first derivative, we could find the critical points of $f(x)$ and then classify those critical points using the First Derivative Test. For our goal of finding the intervals where $f(x)$ is concave up and concave down, and then finding the inflection points of f, let's move on.

b. Find the second derivative, $f''(x)$, and confirm that this is really:

$$f''(x) = -\frac{x^4 + 4x^2 - 1}{\left(x^2+1\right)^2}$$

c. Our goal is to find the x-values where $f''(x) = 0$ or where $f''(x)$ doesn't exist. Note that in order to find where $f''(x) = 0$, we are really looking at the x-values that make the numerator 0:

$$x^4 + 4x^2 - 1 = 0$$

Then, to find where $f''(x)$ doesn't exist, we are finding the x-values that make the denominator 0:

$$\left(x^2+1\right)^2 = 0$$

Solve these equations. **Hint**: Here are two strategies for solving $x^4 + 4x^2 - 1 = 0$:

1. Write some sort of substitution where $u = x^2$

$$\left(x^2\right)^2 + 4\left(x^2\right) - 1 = 0$$
$$u^2 + 4u + 1 = 0$$

Now solve this for u using the quadratic formula. Note that in the end, the two values you get will be possibilities for $u = x^2$. Make sure you get your answer to be in terms of x!

2. A similar technique (really the same thing) is to "complete the square" and write your equation this way:

$$\left(x^2+2\right)^2 - 5 = 0$$

Now solve for x in a natural way. Something to note here is that $x^2 + 2 > 0$, so $x^2 + 2 = \sqrt{5}$ (and notably not $-\sqrt{5}$).

d. You have two critical points of $f'(x)$ (let's just call them x_1 and x_2). These are possible inflection points of $f(x)$, but we need to check to see if the concavity changes at these points. Fill in the following sign chart and interpret.

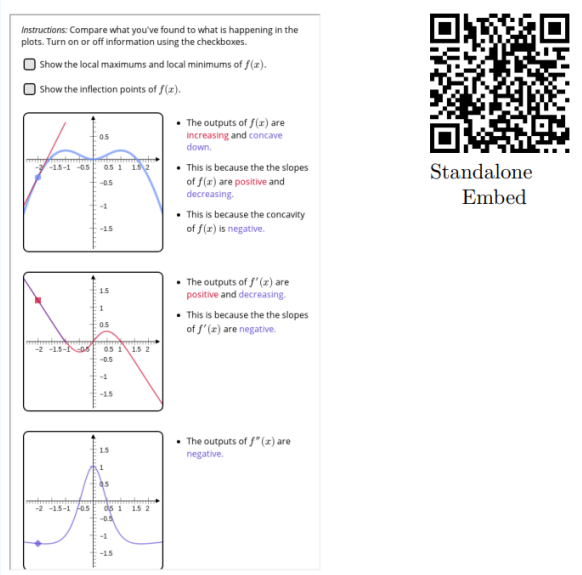
x	$(-\infty, x_1)$	x_1	(x_1, x_2)	x_2	(x_2, ∞)
f''					

f					

Let's confirm what we've just calculated graphically. Below, we have three graphs:

1. $f(x) = \ln(x^2+1) - \dfrac{x^2}{2}$
2. $f'(x) = -\dfrac{x(x+1)(x-1)}{x^2+1}$
3. $f''(x) = -\dfrac{x^4+4x^2-1}{(x^2+1)^2}$

Move the point on any graph and make sure the statements about signs, directions, and concavity match what you found! You should notice that signs of the first and second derivative change at the local maximums/ minimums and the inflection points that we found.



Answer available at the end of this section.

Let's circle back to the definition of Concavity and Inflection Points and think about these from the perspective of $f'(x)$.

We can notice that, by the definition, any inflection point is a point at which $f(x)$ changes concavity, and so is a point where $f'(x)$ changes direction. That means we are looking at local maximums or local minimums of $f'(x)$ (as long as they're not at the end points of some domain)! Similarly, these are points at which $f''(x)$ changes sign, and so we are thinking about the x-intercepts of these second derivatives (or other kinds of locations where the second derivative could change signs).

Example 4.3.2

Let's look at a few more examples, but this time they can be ones with derivatives that are a bit easier to work with.

a. Consider the function $f(x) = \ln(x^2+1)$. Find the intervals where f is concave up, the intervals where it is concave down, and then find the locations of any inflection points.
b. Consider the function $g(x) = e^{-x^2}$. Find the intervals where g is concave up, the intervals where it is concave down, and then find the locations of any inflection points.

At this point, we have three different functions that we are juggling: a function $f(x)$ (or whatever name we give it), the first derivative $f'(x)$, and the second derivative $f''(x)$. We'll want to keep in mind the role of each of these.

- $f(x)$ tells us the height, the y-value, of the function at some point.
- $f'(x)$\) tells us the direction, the slope, of the function at some point.
- $f''(x)$ tells us the curvature, the concavity, of the function at some point. We can try to summarize this in a small table:

$f(x)$	$\nearrow$	$\searrow$	$\smile$	$\frown$		
$f'(x)$	$\oplus$	$\ominus$	$\nearrow$	$\searrow$	$\smile$	$\frown$
$f''(x)$			$\oplus$	$\ominus$	$\nearrow$	$\searrow$
$f'''(x)$					$\oplus$	$\ominus$

Notice that we could extend this table and think about any triplet of functions/ derivatives in a row: we just need to think about what is positive/negative, what is increasing/decreasing, and what is concave up/down. If we wanted, we could try to define some adjective to describe what is happening to a function when $f'(x)$ is concave up/down, but let's not. It's hard enough to visualize concavity, and it will be difficult to visualize rates of change of the concavity.

Interpreting the Concavity at Critical Points

Theorem 4.3.3: Second Derivative Test for Local Maximums or Local Minimums.

If $(c, f(c))$ is a critical point of f with $f'(c) = 0$, then we can use the value of the second derivative at $x = c$ to classify the critical point:

1. If $f''(c) > 0$, then f is concave up at and around $x = c$, and so the function has a local minimum at $(c, f(c))$.
2. If $f''(c) < 0$, then f is concave down at and around $x = c$, and so the function has a local maximum at $(c, f(c))$.
3. If $f''(c) = 0$, then the Second Derivative Test is inconclusive.

Example 4.3.4

Find any critical points of the following functions. For each, use the Second Derivative Test to classify the critical point. If the Second Derivative Test fails (we get that the second derivative evaluated at the critical point is 0, and so is inconclusive), then classify the critical point in some other way.

a. $f(x) = x\ln(x) - x$ where the domain of f is $(0, \infty)$.
b. $g(x) = 5x + \dfrac{1}{x}$ where the domain of g is $(0, \infty)$.

Let's leave this here, with a few questions for you to think about:

- When, for you, do you think it would be reasonable to use the Second Derivative Test instead of the First Derivative Test? What goes into making this decision?
- When, for you, do you think it would be reasonable to use the First Derivative Test instead of the Second Derivative Test? What goes into making this decision?

Practice Problems

Exercise 4.3.1

Consider a function $f(x)$ that is continuous on $(-\infty, \infty)$ with $f''(x) > 0$ on $(-\infty, 5)$ and $f''(x) < 0$ on $(5, \infty)$. What do we know about $f'(x)$ on these intervals? What do we know about $f(x)$ on these intervals?

Exercise 4.3.2

Consider a function $g(x)$ that is continuous on $(-\infty, \infty)$ that is concave down on $(-\infty, -2)$, concave up on (–2, 4), and concave down on $(4, \infty)$. Where are the inflection points of $g(x)$? What is happening to $g'(x)$ at these points?

Exercise 4.3.3

Explain (with words and/or sketches) why the Second Derivative Test requires $f'(a) = 0$ for the second derivative to classify the critical number a as a local maximum or minimum. Why can't the second derivative classify a critical point where $f'(a)$ does not exist?

Exercise 4.3.4

For each of the following functions, use the second derivative to find the intervals where the function is concave up or concave down, and find the location of any inflection points.

a. $f(x) = e^x - x^2$
b. $f(x) = \ln(x^2 + 1)$
c. $f(x) = e^{-x^2}$
d. $f(x) = 1 - 2x + \dfrac{x^5}{20} - \dfrac{x^4}{12}$
e. $f(x) = \dfrac{4x^{5/2}}{15} - x^2$
f. $f(x) = 5\cos(x) - 2x + 1$ on $(0, 2\pi)$
g. $f(x) = x^2 e^x$
h. $f(x) = \tan^{-1}(x)$
i. $f(x) = \dfrac{1}{(x-1)(x-2)}$
j. $f(x) = 2x^4 + x$

Exercise 4.3.5

Find critical points of the following functions, and apply the Second Derivative Test to classify those critical points as local maximums/minimums. If the test fails, explain why.

a. $f(x) = 2x^4 + x$
b. $f(x) = \dfrac{4}{x} + 3x$ on $(0, \infty)$
c. $f(x) = (x-1)e^x$
d. $f(x) = \sqrt{x^3} - x$

Exercise 4.3.6

Sketch a graph of a function $f(x)$ and has the following characteristics. Label the intervals where f is concave up/down, and find the locations of any local inflection points.

- $f''(x) < 0$ on $(-\infty, -1)$
- $f''(x) > 0$ on $(-1, 4)$
- $f'(x) < 0$ on $(-\infty, 2)$
- $f'(x) > 0$ on $(2, 6)$
- $f'(6) = 0$
- $f''(x) < 0$ on $(4, 6)$
- $f'(x) > 0$ on $(6, \infty)$

- $f''(x) > 0$ on $(6, \infty)$

Exercise 4.3.7

Consider a function $f(x)$ that is differentiable on $(-\infty, \infty)$. We graph the first derivative, $f'(x)$, below.

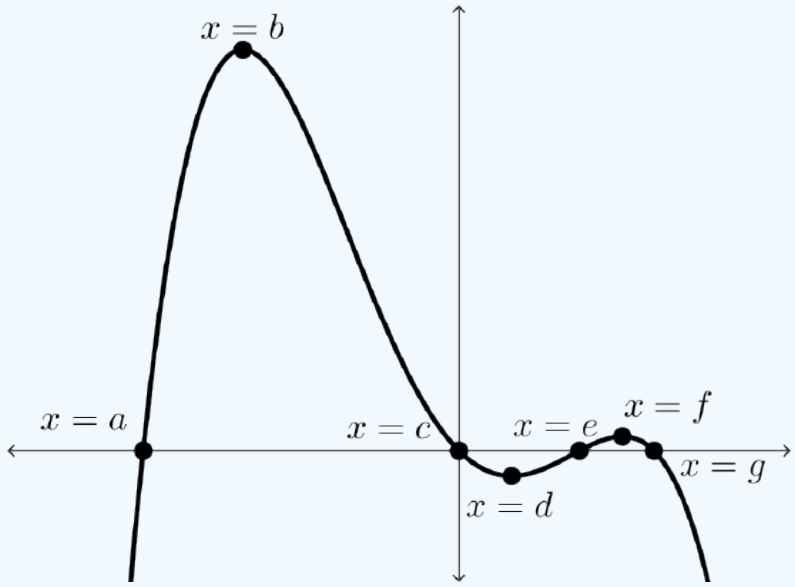


Figure 4.3.5 Graph of $f'(x)$.

a. Where is the function $f(x)$ concave up?
b. Where is the function $f(x)$ concave down?
c. Where are the inflection points of $f(x)$? What do they correspond with on the graph of $f'(x)$? Explain.
d. Explain whether $x = a$ is the location of a local max of $f(x)$, a local min of $f(x)$, or neither. Reference both the First and Second Derivative Test.
e. Where are the other local maximums/minimums of $f(x)$?

Exercise 4.3.8

Consider a function $g(x)$ that is differentiable on $(-\infty, \infty)$. We graph the second derivative, $g''(x)$, below.

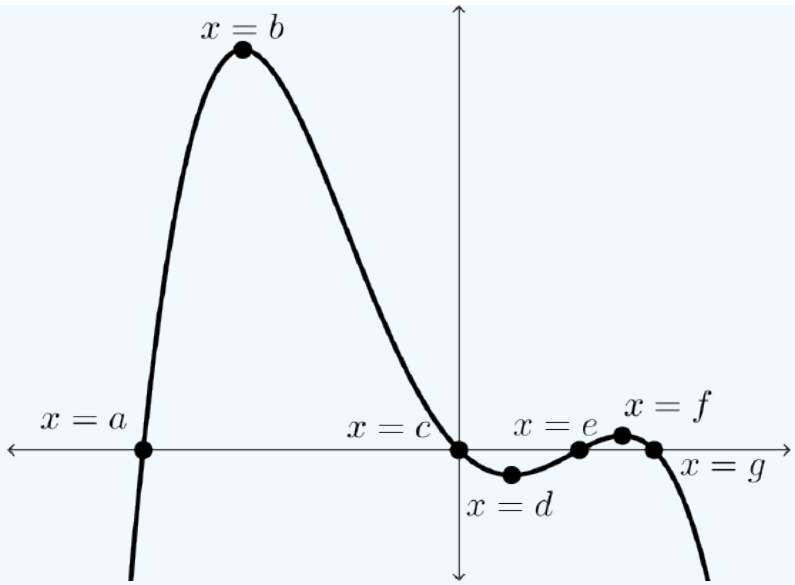


Figure 4.3.6 Graph of g ″(x).

a. Where is the function $g(x)$ concave up?
b. Where is the function $g(x)$ concave down?
c. Where are the inflection points of $g(x)$? What do they correspond with on the graph of $g''(x)$? Explain.
d. If there is some x-value, (call it x_0) with $a < x_0 < b$ where we know that $g'(x_0) = 0$, then what else can we say about this point on the graph of $g(x)$?

Answers

Activity 4.3.2**: Finding a Function's Concavity.**

c. You should get that the critical points of $f'(x)$ are at $x = -\sqrt{\sqrt{5}-2}$ and $x = \sqrt{\sqrt{5}-2}$.

d.

x	$(-\infty, x_1)$	x_1	(x_1, x_2)	x_2	(x_2, ∞)
f''	$\ominus$	0	$\oplus$	0	$\ominus$
f	$\frown$ concave down	inflection point	$\smile$ concave up	inflection point	$\frown$ concave down

This page titled 4.3: Concavity is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Peter Keep.

4.4: Global Maximums and Minimums

We need to start with a definition, and we can start with contrasting what we want the difference between a local maximum/minimum and a global maximum/minimum. Sometimes these are also called absolute maximum/minimums.

What do you want the difference to be?

we focus on the terms or the names we're giving, then the difference should be based on the distinction between the words "local" and "global." In one, we're considering some confined and relatively arbitrary geographic area, just the things around or in the neighborhood. In the other, our context grows until we're considering the whole picture, the whole space that we're interested in!

Definition 4.4.1: Global Maximum and Global Minimum.

A function has a **global maximum** value of $f(c)$ if $f(c) \geq f(x)$ for all x-values in the domain of f.

A function has a **global minimum** value of $f(c)$ if $f(c) \leq f(x)$ for all x-values in the domain of f.

Note that the difference between this definition and Definition 4.2.3 Local Maximum/Minimum is the types of y-values we're comparing $f(c)$ to: in this new definition, we just use all of the x-values in the domain. In the definition for a local max/min, we had to construct some interval to intersect with the domain in order to just consider the "local" picture.

Activity 4.4.1: When Would We Not Have Maximums or Minimums?

In this section, we're going to define these global maximums and then, most importantly, try to predict when these global maximums or global minimums might actually exist for a function. To start, maybe we should come up with some examples of functions that do not have them!

a. Come up with some situations where a function does not have some combination of global maximum/minimums. Sketch some graphs!
b. Come up with some examples of graphs of functions that are bounded (do not ever have y-values that tend towards infinity in a limit) that do not have some combination of global maximum/ minimums.
c. For the examples of graphs that you have built or collected, what are the features of the functions that allow for the examples you picked? If you could impose some requirements that would "fix" the examples you found (so that they had both a global maximum and a global minimum), what requirements could you use?

When Do We Guarantee Both a Global Maximum and a Global Minimum?

Theorem 4.4.2: Extreme Value Theorem.

If $f(x)$ is a continuous function on a closed interval $[a, b]$, then f must have both a global maximum and a global minimum on the interval.

The Extreme Value Theorem guarantees the existence of both the global maximums and minimums, but we actually get more than just this out of the Extreme Value Theorem. Once we know that both of the global maximums and minimums exist, we can find them pretty easily.

The global maximum of a function, if it exists for the function on the domain/interval, is just the local maximum with the largest y-value. Similarly, the global minimum, if it exists, is the local minimum with the lowest y-value.

So once we know they both exist for a function on an interval, we can simply collect the critical points of the function (including the ending points of the domain) and compare the y-value function outputs!

Example 4.4.3

Check to see if each function (on the stated domain) satisfies the conditions of the Extreme Value Theorem, and then find any global maximums/minimums of the function on the interval.

a. $f(x) = \ln(x^2 + 4x + 7)$ on [-1,3]
b. $g(x) = 3x^4 - 5x$ on [–3, 4]
c. $j(x) = \sqrt[3]{x+4}$ on [–6, –1]

What about Domains of Functions that Aren't Closed?

Without the conditions that imply the Extreme Value Theorem, things become trickier. For instance, if the function is not continuous, then the function might have some unbounded behavior at a vertical asymptote. In this case, we might need to look at the one-sided limits around that asymptote, in order to see if our function tends towards positive or negative infinity on either side of the asymptote. This could tell us that the function doesn't have a global maximum, a global minimum, or that it doesn't have either.

Similarly, if the function is not defined on a closed interval, then we need to investigate what happens to the function's behavior as the function moves towards the "ends" of the interval. This could mean that the function approaches some real number, but it could also mean that the function approaches positive or negative infinity, or something else. These end behavior limits could exist, in which case we need to compare these heights of horizontal asymptotes or open ends of an interval to the heights of any critical numbers.

But we might also find that the function tends towards infinity or negative infinity in the end behavior.

And there are other things to consider about discontinuity of a function (other than vertical asymptotes)!

All in all, it should be evident that if we remove one or both of the conditions on our function that guarantees the existence of a global maximum and a global minimum, it becomes much harder to find them, since we have so many different options to consider.

To simplify things, we will look at one case where things align in our favor: a continuous function that only has a single local maximum/minimum on an interval.

Theorem 4.4.4

If $f(x)$ is a continuous function on some interval, and f has only a single critical point (call it $(c, f(c))$) where the direction changes, then if that point is a local maximum of f, then $f(c)$ is the global maximum. Similarly, if $(c, f(c))$ is a local minimum, then $f(c)$ is the global minimum of f.

This is a great result to give us a path forward without having to check all of the edge cases and possibilities mentioned above. There are many functions that might have only a single critical point, or if it does have more than one critical point, only a single one of them acting as a local maximum/minimum.

Note here that we do need to classify the critical point as a local maximum or minimum! We'll use the First Derivative Test or the Second Derivative Test for Local Maximums or Local Minimums for this classification.

Example 4.4.5

For each function, find any global maximums/minimums that may exist.

a. $f(x) = x\ln(x) - x$ and note that the domain of this function is $(0, \infty)$
b. $g(x) = xe^{-x}$ and note that the domain of this function is $(-\infty, \infty)$.

Practice Problems

Exercise 4.4.1

What does the Extreme Value Theorem say? Why is it helpful?

Exercise 4.4.2

Where, in general, do we find global/absolute extrema?

Exercise 4.4.3

For each function, describe whether or not the Extreme Value Theorem applies (and explain why or why not). If it does apply, find the absolute maximum and absolute minimum.

a. $f(x) = x^2 + 5x - 4$ on [−4, 2]
b. $f(x) = \sqrt[3]{x+4}$ on [−6, −2]
c. $f(x) = |x - 1|$ on [−4, 0]
d. $f(x) = \dfrac{1}{x^2}$ on [−3, 1]
e. $f(x) = x^{2/3}(2-x)^2$ on [−1, 1]
f. $f(x) = \begin{cases} 4x & \text{if } 0 \le x < 1 \\ \sqrt{x} & \text{if } 1 \le x \le 4 \end{cases}$

This page titled 4.4: Global Maximums and Minimums is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Peter Keep.

4.5: Optimization

How do we use calculus to make decisions? If we're trying to find the best allocation of time between different tasks, or if we're trying to construct some object with limited resources, or if we're trying to find some other solution to the question, "How do we make the best/most/maximum/minimum of..." then we can think about translating these to a calculus context.

Optimization problems are some of the clearest application of calculus concepts to applied problems from any industry or field. While there certainly aren't teams of calculus experts in every office waiting for these kinds of problems to arrive, these problems are routinely solved using calculus, either done by computer software or coming from calculus experts who make sweeping recommendations across a whole field or some other source.

Optimization Framework

We'll start this discussion with a small example. This is not an example that shows up in "real life," but will be helpful to build a strategy for approaching these problems.

Here's our problem:

Example 4.5.1

Find two numbers that add to get 14 but give the largest possible product.

That's it. That's the problem. It's not super interesting, right? Let's think about how we'll solve it.

Solution

Solution 1.

First, we want to think about our two numbers. We'll need variable names for these. You can use whatever letters you'd like here: I'm going to be boring and pick a and b as the names of my two numbers. What do we know about a and b?

The first thing that we know is that $a+b=14$. This isn't a huge amount of information, but it does provide us with a nice connection between our numbers. If we know what one number is, then we automatically know the second one: if one number is 4, then the other number is whatever we need to add to 4 to get 14.

We also know that we're interested in the product: $P=ab$. We don't, currently, know anything about this product other than the fact that we want it to be large. Ok so automatically, we know that both a and b need to be positive: if we had one number that was positive (like 15) and one number that was negative (like –1), then the product will be negative, and even though they add to 14, we are not going to get a big product.

So we know that for both of these numbers $a,b\geq 0$. This doesn't seem like a lot of information, but we can put it all together really nicely. For instance, I can manipulate the fact that these two numbers add to get 14 into an equation that tells me what the value of one number is based on the other:

$$a+b=14\longrightarrow b=14-a$$

I can also write my product as an actual function of a single variable:

$$P=ab\longrightarrow P(a)=a(\underbrace{14-a}_{b})$$

Lastly, I actually know a domain for this function: we know that $a\geq 0$, but since $b\geq 0$, then $a\leq 14$. We can't have a number larger than 14, since the other number would be negative.

Aside.

I just said that we can't have a negative product, but nothing in this problem says that. We just know that a negative product will be small. But we don't have to limit each number to be between 0 and 14. We could just allow them to be bigger than 14 (and thus also allowed to be smaller than 0) and we'll just get a bunch of negative products that aren't the maximum one. Oh well!

We are making a similar choice by including the pair 0 and 14 as a possibility. Do we really think that this could give us the biggest product? NO! The product is 0! But it's maybe convenient to have a closed interval as the domain of this function, so why not?

All along the way, we're making choices that guide how we think about this problem. In real life, we'll do the same thing. The only difference is that the choices that we make about what are reasonable values for our inputs to take on are more meaningful, since those inputs represent real things. We have to take these choices seriously.

Ok, so we have a function representing the product, $P(a) = a(14 - a)$ and a domain for that function, [0, 14], and we know we want to find the maximum for it. This is a calculus problem that we can actually do! Take a moment to do it.

Solution 2.

$$P(a) = 14a - a^2$$
$$P'(a) = 14 - 2a$$

Now we can find the critical points.

$$P'(a) = 0$$
$$14 - 2a = 0$$
$$-2a = -14$$
$$a = 7$$

We also know that the ending points at $a = 0$ and $a = 14$ are critical, since the derivative cannot exist.

There is a point on the product function when $a = 7$ where we have a horizontal tangent line. Does this represent a maximum or a minimum? We have some options for how to do this:

1. FDT:

a	[0, 7)	7	(7, 14]
P'	$\oplus$	0	$\ominus$
P	$\nearrow$ increasing	(7, 49) local max	$\searrow$ decreasing

   Since this is the only turning point, it must be the global maximum!

2. SDT:

$$P''(a) = -2$$
$$P''(7) < 0$$

   Since the function is concave down at this point, then we know that the function reaches a local maximum when $a = 7$. Since this is the only turning point, it must be the global maximum!
3. EVT:

a	P	
0	0	global minimum
7	49	global maximum
14	0	global minimum

However we do this, we find that there is a maximum product when $a = 7$. What is the second number, b? Well we know that they add to 14, so:

$$\begin{aligned} b = 14 - a \longrightarrow b &= 14 - 7 \\ &= 7 \end{aligned}$$

So our two numbers are 7 and 7, and they multiply to get 49, the biggest possible product between two numbers that add to 14.

We have accomplished something, even if it's not much. Hurray, we solved a pretty unimportant problem about numbers!

More importantly, though, we set up a process for how we're going to approach optimization problems.

Optimization Process.

1. Label variables. What are the quantities that we care about? What are the kinds of measurements that we're given or need to find information about? Label them!
2. Find a formula to optimize. What are we trying to find the maximum or minimum of? Is it a formula that we might already know, or is there some other way of constructing that formula? This might involve drawing some geometric shape to get a clue!
3. Find a constraint. The constraint is really just some other connection between variables that guides their relationship. If we know some of the variables, there might be a way to find another one, since it would then depend on that variable.
4. Find some domains. For your variables, what are the smallest and largest possible values that they can reasonably take on? Are there any?
5. Set up a calculus statement. A calculus statement is a sentence that includes:
   a. A function you are optimizing. This should be a function with only one input, so we might need to use the constraint to rewrite our formula from earlier in order for it to only have a single input variable.
   b. A domain for that function. This should be the interval you found earlier that is relevant for your choice of input variables.
   c. Some indication of whether you're finding a maximum or a minimum. An example of a calculus statement might be something like: "We want to find the maximum/minimum of function on domain ."
6. Do calculus. We want to find a global maximum or a minimum of a function...we know how to do that! We'll find critical points, and then use the First Derivative Test, the Second Derivative Test for Local Maximums or Local Minimums, or the Extreme Value Theorem (and it's follow-up strategy) to find the global or absolute maximum/minimum!

In our example, we did the following:

1. Label variables. Our two numbers were a and b.
2. Find a formula to optimize. We said that $P = ab$ was the product.
3. Find a constraint. We knew that $a + b = 14$
4. Find some domains. We reasoned that $0 \leq a \leq 14$ and $0 \leq b \leq 14$.
5. Set up a calculus statement. We wanted to find the maximum of $P(a) = a(\underbrace{14 - a}_{b})$ on [0, 14].
6. Do calculus. We showed how we could use three different techniques to solve this.

Balancing Volume and Surface Area

Let's use this new Optimization Process to solve a real problem. This one is a classic problem that (in my opinion) every student should try. I hope you'll see why.

Activity 4.5.1: Constructing a Can.

A typical can of pop is 355 ml, and has around 15 ml of headspace (air). This results in a can that can hold approximately 23 cubic inches of volume.

Let's say we want to construct this can in the most efficient way, where we use the least amount of material. How could we do that? What are the dimensions of the can?

a. First, let's think of our can and try to translate this to some geometric shape with variable names. Collect as much information as we can about this setup! What is the shape? What are the names of the dimensions?
b. What is the actual measurement that we are trying to optimize? Are we finding a maximum or a minimum?
c. What other information about the can do we know? How do we translate this into a constraint, or a connection between our variables?
d. What kinds of values can our variables take on? Is there a smallest value for either? A largest? Are there other limitations to these?
e. Now we need to set up a calculus statement. This part mostly relies on us finding a way to build a single-variable function defining the surface area. Build that function, and then write down the calculus statement. **Hint**: Can you solve the constraint, $23 = \pi r^2 h$ for one of the variables? Can you then use that to replace or substitute out one of the variables in our surface area formula to create a function, $A(h)$ or $A(r)$?
f. Do some calculus to find the global maximum or minimum, and solve the optimization problem. **Hint**: Find a derivative, then find critical points, and test them using with the First Derivative Test or the Second Derivative Test for Local Maximums or Local Minimums. Make sure you find the other dimension after you've confirmed that you found the minimum of the function!
g. What is the relationship between r and h, here? How do they compare? What about the height and diameter of our can? Is this relationship noticeable in a typical can of pop? **Hint**: The height is twice the radius, or the height is equal to the diameter! We might call this a "square-cylinder", since the "width" and the height of the cylinder are equal. Do you ever see "square-cylinder" cans?

Answer available at the end of this section.

So why, then, do we never see cans that look like this? It is certainly worth thinking about how the setup and assumptions we made in this activity might not be the way things work in real life.

What are some reasons that we might not see these dimensions in a can of pop?

Note 4.5.2

This question (why do the dimensions of cans not match what we found as the optimal solution?) is interesting, and sometime in the in the 1980's, a math professor felt strongly enough about it that they wrote a letter to Carnation, a brand of food products that produces canned goods, asking why reality doesn't match mathematics.

The professor received a response from the Assistant Product Manager of Friskies Buffet with 5 reasons for the non-square dimensions. The full text of the letter can be found here: Appendix A Carnation Letter.

What Other Examples Can We Do?

There's a really important point to make from the Carnation Letter: in real life, our optimization problems are multi-variable problems. We're balancing many different aspects of a process or a problem to find an optimal solution. That's hard to do in a calculus course that focuses on single-variable functions!

So what kinds of problems can we actually do?

There are a bunch, but they don't really stand up to intense scrutiny: if we looked carefully at most of the "classic" optimization problems that we see in calculus texts, they'd fall apart just like the optimal can problem from Activity 4.5.1.

That said, it's good to practice thinking about using constraints and translating physical situations into formulas and functions, so try the practice problems to get used to this part of the optimization process!

Practice Problems

Exercise 4.5.1

When we set up an optimization problem, we normally look for some formula connecting information that we have with the thing we want to optimize, some sort of constraint or link between variables, and an clue to whether we want to find a maximum or minimum. From this, we work to create a "Calculus Statement" which includes a function (of one variable) that we want to find the absolute maximum or minimum of on a domain.

a. Why is it important to create a function of one variable? How do we normally do this?
b. Why is it important to find a domain for the function? How do we normally do this?
c. Why is it important to note whether we're looking for an absolute maximum or an absolute minimum?

Exercise 4.5.2

Let's say we're looking at two positive numbers that add up to 7. What is the maximum product that we can get when we multiply these two numbers together?

Exercise 4.5.3

Out of all of the rectangles that have an area of 24 square units, which one has the smallest perimeter? Report the dimensions and the perimeter.

Exercise 4.5.4

Consider a box created from an 18in square piece of material. We form the box by cutting an h-inch square out of each corner, and folding the sides up along the dotted lines.

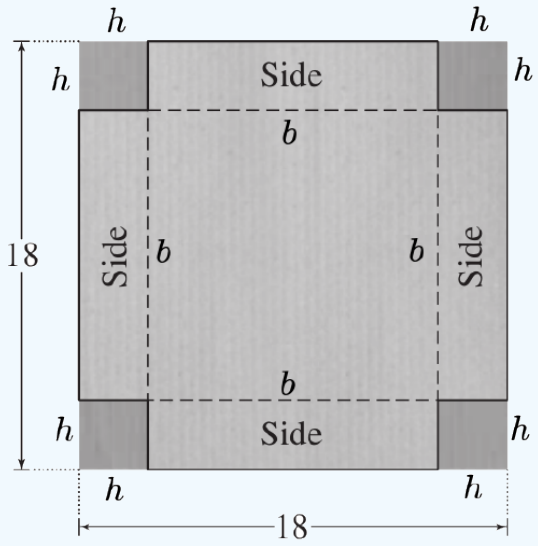


Figure 4.5.3 Material for the box.

a. What are the dimensions of the resulting open-topped box?
b. How are the values b and h related to each other?
c. What is the largest possible value for h? What is the corresponding value of b?
d. What is the smallest possible value for h? What is the corresponding value of b?
e. Find the value for h that maximizes the volume of the box. What is the maximum volume? What are the dimensions of the box?

Exercise 4.5.5

A trucking service is looking to add a route to their plan, but the road that they are considering passes through a tunnel. The top of the opening above the driving lane is formed by the curve $y = 4 - \frac{4}{27}x^3$. The width from the center line to the edge of the tunnel is 3 meters. What are the dimensions for the largest rectangular trailer the trucking service could pull through the tunnel?

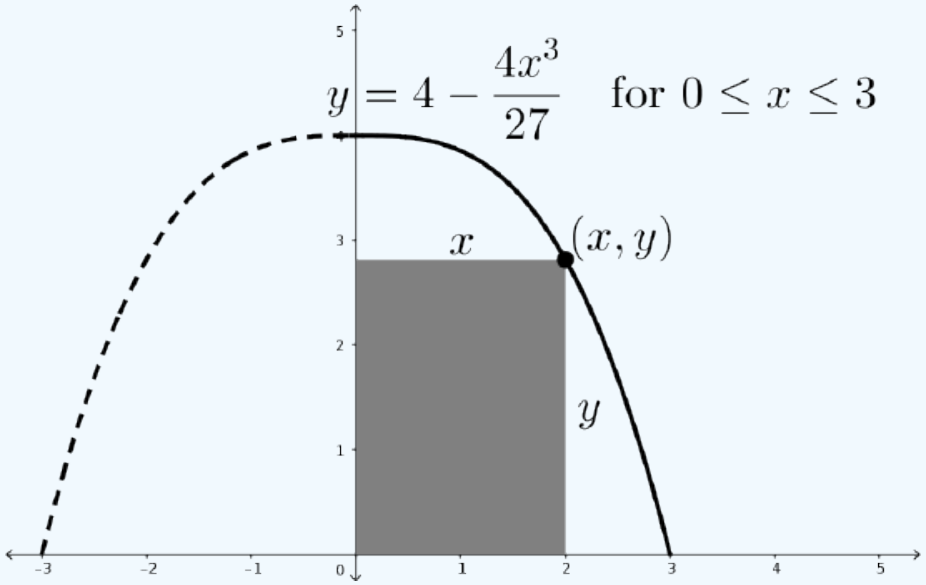


Figure 4.5.4 The truck in the tunnel.

Exercise 4.5.6

A circular area of radius 20ft. is surrounded by a paved pathway. A lamp post is placed in the middle of the circular area. What height should the light be to illuminate the walk most strongly? The intensity of illumination is $I = \frac{\sin\theta}{s}$ where s is the distance form the light source and θ is the angle at which the light hits the surface.

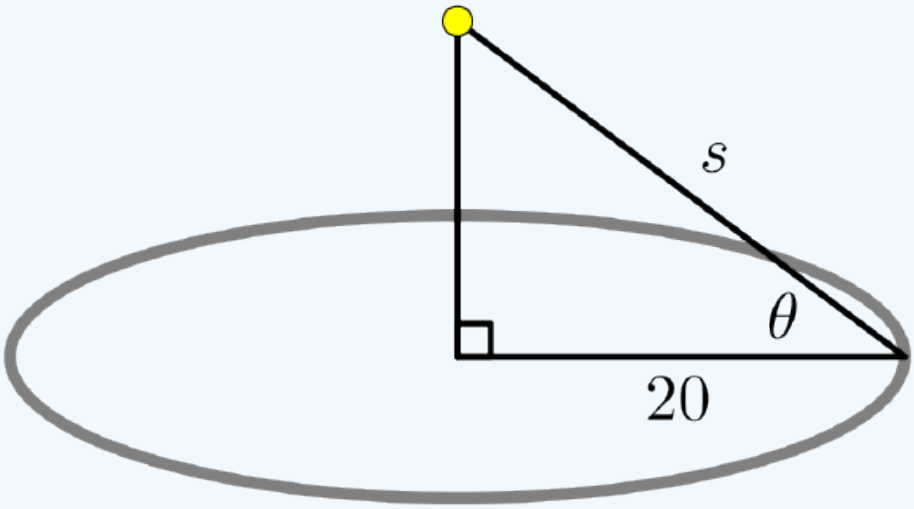


Figure 4.5.5 A solitary lampost.

Answers

Activity $4.5.1$**: Constructing a Can.**

a. Our can is probably a cylinder! For a cylinder, we can define it based on a radius and height, and we normally call them r and h. Both of these will be measured in inches.

b. We're trying to minimize the amount of material used to create the can: this should match up with surface area!
The formula for the surface area of a cylinder is $A = 2\pi r^2 + 2\pi rh$

c. We're told that the volume is 23 cubic inches! The volume formula for a cylinder is $V = \pi r^2 h$, so we have a constraint:

$$23 = \pi r^2 h$$

d. It's hopefully obvious that we need positive dimensions to make sense: we can't have a can with a negative radius or a negative height.
Notice, also, that neither dimension can be equal to 0: if one of them were , then the volume would also be , which doesn't work with our constraint!
We can also notice that as one dimension gets close to 0, think about the radius for example, then the other dimension, the height in this case, would need to get *really* large to compensate and keep the volume fixed.
All of this to say that our intervals are:

$$r > 0$$
$$h > 0$$

e. We have two options here, depending on which variable you isolated in the constraint and then substituted. These are the two ways of isolating for a variable in the constraint:
1. $23 = \pi r^2 h \longrightarrow h = \dfrac{23}{\pi r^2}$
2. $23 = \pi r^2 h \longrightarrow r = \sqrt{\dfrac{23}{\pi h}}$
   - We want to find the minimum of $A(r) = 2\pi r + \dfrac{46}{r}$ on $(0, \infty)$.
   - We want to find the minimum of $A(h) = 2\pi\sqrt{\dfrac{23}{\pi h}} + \dfrac{46}{h}$ on $(0, \infty)$
   - Personally, I think the first option is probably easier to work with.

f. Let's look at the version with

$$A(r) = 2\pi r^2 + \frac{46}{r} \text{ on } (0, \infty).$$
$$A'(r) = 4\pi r - \frac{46}{r^2}$$
$$\text{Solve } A'(r) = 0$$
$$0 = 4\pi r - \frac{46}{r^2}$$
$$\frac{46}{r^2} = 4\pi r$$
$$\frac{1}{r^2} = \frac{2\pi r}{23}$$
$$r^2 = \frac{23}{2\pi r}$$
$$r^3 = \frac{23}{2\pi}$$
$$r = \sqrt[3]{\frac{23}{2\pi}}$$

So we have a critical point on $A(r)$ at $r = \sqrt[3]{\frac{23}{2\pi}}$. Now we need to test it to ensure that it is a minimum (since that's what we're trying to find). We have two options to do this: the First Derivative Test, or the Second Derivative Test.

1. FDT:

r	$\left(0, \sqrt[3]{\frac{23}{2\pi}}\right)$	$\sqrt[3]{\frac{23}{2\pi}}$	$\left(\sqrt[3]{\frac{23}{2\pi}}, \infty\right)$
A'	$\ominus$	$\not\exists$	$\oplus$
A	$\searrow$ decreasing	local min	$\nearrow$ increasing

   Since this is the only turning point, then we know that there is a global minimum of A at $r = \sqrt[3]{\frac{23}{2\pi}}$.

2. SDT:

$$A''(r) = 4\pi + \frac{92}{r^3}$$

   Hang on one second: notice that no matter what, since $r > 0$, then we can see that $A''(r) > 0$. We don't even need to evaluate this second derivative at the critical point to know that the function is concave up here! So then we have to have a local minimum, and since it's the only turning point, a global minimum of A at $r = \sqrt[3]{\frac{23}{2\pi}}$.So we found our minimum! Now we just need to find the other dimension, h.

$$h = \frac{23}{\pi r^2} \qquad \text{This is the constraint.}$$
$$h = \frac{23}{\pi\left(\sqrt[3]{\frac{23}{2\pi}}\right)^2}$$
$$h = \left(\frac{23}{\pi}\right)\left(\frac{2\pi}{23}\right)^{2/3}$$
$$h = 2^{2/3}\left(\frac{23}{\pi}\right)^{1/3}$$
$$h2\sqrt[3]{\frac{23}{2\pi}}$$

   This is a lot of algebra, when we could have just plopped these values into a calculator and noticed the same, interesting thing.

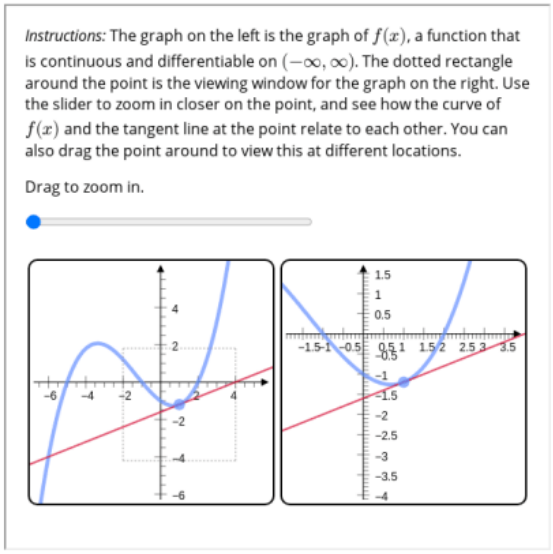
This page titled 4.5: Optimization is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Peter Keep.

4.6: Linear Approximations

We're going to return to a pretty central idea here, one that we've been using and developing and really exploring. But let's think about the very basic version of what we've been looking at over this whole chapter (and more):

The derivative of a function tells us the slope of the line tangent to the function at a point.

But what we'll do is explore how this tangent line and the graph of our function interact and relate to each other. Let's start with just playing with a graph and seeing if we can discover some things to say about the relationship between a tangent line and the function it is lying tangent to!



Standalone
Embed

The effect that we are seeing when we zoom in on a function is sometimes described as our function being locally linear. What do you think this means? Why is this a good description of what we're looking at, and how these differentiable functions are constructed?

Would this effect be noticeable for every function, even ones that are not differentiable at some points?

Convince yourself that a function will not look locally linear at a point where the function is not differentiable. You might want to remind yourself what it looks like, graphically, when a derivative doesn't exist: When Does a Derivative Not Exist?

Linearly Approximating a Function

The visual above should provide us with a nice framework to think about how we might approximate a function linearly, but we can recap some basic ideas:

- When we say "linear approximation," we're really just referring to the tangent line at some point.
- Our functions only look "locally linear" when we zoom in around some single point. Another way of saying this is that our tangent line only matches the behavior of our function really close to the point where we built the tangent line.
- We have a kind of vague or ambiguous idea of accuracy in approximation. While a tangent line follows the behavior of the function "around" that point where it was built, the actual rate at which it deviates from our function is different. If we move the point in the visual above, we'll see that at some locations, our function is pretty linear and doesn't move away from the tangent line very quickly. In other locations, the function turns quickly away from the tangent line!

Definition 4.6.1: Linear Approximation of a Function.

If $f(x)$ is differentiable at $x = a$, then we say that a l**inear approximation** of $f(x)$ centered at $x = a$ is:

$$L(x) = f'(a)(x - a) + f(a)$$

We know, then, that $L(x) \approx f(x)$ for x-values "close" to the center, $x = a$.

Note that the **center** is just the point at which we are building this linear approximation: the point at which we build the tangent line.

Let's see this in action!

Exercise 4.6.1: Approximating an Exponential Function.

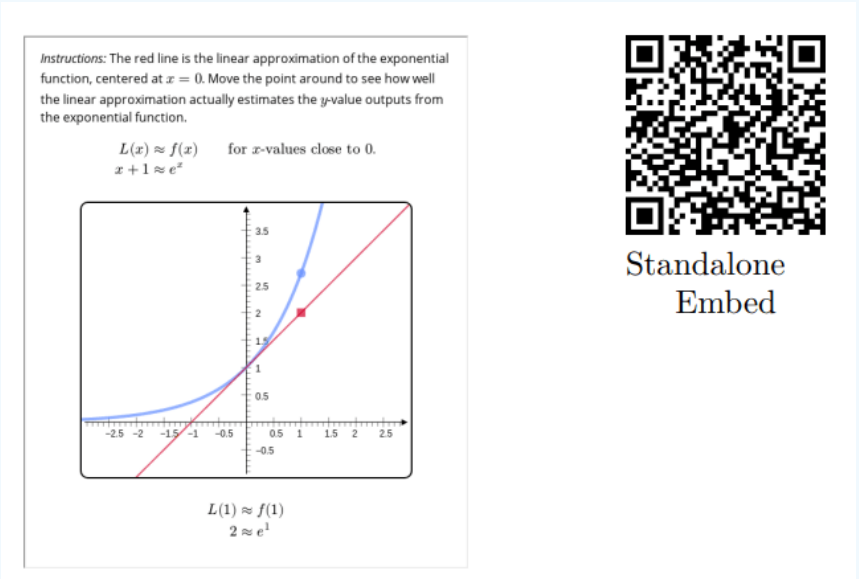
Let's consider the function $f(x) = e^x$. We're going to build the linear approximation, $L(x)$, but we also want to focus on understanding what the "center" is, and how we think about accuracy of our estimations.

a. We first need to find a "good" center for our linear approximation. We have two real requirements here:
   i. We need the center to be some x-value that will be "close" to the inputs we're most interested in. We know that $L(x) \approx f(x)$ for x-values "near" the center, so we should keep this in mind. We don't have a specific input that we're interested in (we are not specifically focused on estimating $f(7.35)$ for instance), so we don't need to worry about this for now.
   ii. We are going to need to evaluate the function and its derivative at the center: we use $f'(a)$ to find the slope and $f(a)$ to find a y-value for a point on the line. We'd like to choose a center that will make evaluating these functions reasonable, if we can! We are going to choose a center of 0: why?

b. Build a linear approximation of $f(x) = e^x$ centered at $x = 0$. **Hint**: Build a line with a slope of $f'(0)$ that goes through the point $(0, f(0))$

c. Use your linear approximation to estimate the value of $\sqrt[10]{e} = e^{\frac{1}{10}}$. **Hint**: Since $L(x) \approx f(x)$ for x-values near 0, we can say that $L(1/10) \approx f(1/10)$. So you can evaluate your linear approximation function $x = \frac{1}{10}$.

d. Let's visualize this approximation a bit:



Are you confident in your approximation of $\sqrt[10]{e}$? Would you be more or less confident in an approximation of $\frac{1}{e}$? Why?

e. Is your estimate of $\sqrt[10]{e}$ too big or too small? How can you tell, without even calculating the actual value of $\sqrt[10]{e}$? How can you tell that every estimate that you get out of any linear approximation of e^x (no matter what the center is) is going to be too small? **Hint**: It might be helpful to think about how the function moves away from the tangent line: how do the slopes of e^x change? Can you link this to ideas of concavity?

In this activity, we did not have to think much about an appropriate choice of center. We tried to justify our choice, but that's different than having to make a choice. Let's approach this a bit differently in our next examples.

Exercise 4.6.2: Approximating some Values.

Pick one of the following values to approximate:

- $\sin(-0.023)$
- $\ln(2)$
- $\sqrt{8}$
- $\sqrt[3]{10}$

Throughout the rest of this activity, use your value to build a linear approximation of some relevant function and estimate the value you chose.

a. To build a linear approximation of some function at some center, we need two things:
   i. A function.
   ii. A center. What function will you be using for $f(x)$? Why that one? **Hint**: Your value should look like the output of some function after you plug in some x-value. What function?
b. What center are you choosing? Why that one? **Hint**: Remember that we want some input for your function that is both close to the input you'd like to estimate your function at and also a reasonable one to know the value of your function and its derivative.
c. Build your linear approximation at your center! You should end up with an actual linear function. It might be helpful to plot this linear function and your actual function to confirm that you have actually built a tangent line.
d. Use your linear approximation function to estimate your value! Report the estimate, and comment on the accuracy of your estimate. Without calculating the actual value, can you tell if this is close or not? Do you have an overestimate or underestimate? **Hint**: Think about issues relating to the distance from the center, the concavity of the function, and even the rate at which the slopes change away as we move away from the center.

So far, we have been pretty limited in what we can actually do with these linear approximations. A function is only locally linear when we look at a very small interval of x-values. Once we move away from the center far enough (and it's often not that far), then our function curves away from the tangent line and our linear approximation is not at all accurate.

Aside.

If you want to see how we can make these linear approximations more accurate, then we can try to think about using a quadratic or cubic function instead: something with some curves built into it that can try to follow the function's behavior a bit! We'll cover that in Section 9.1 Polynomial Approximations of Functions.

Approximating Zeros of a Function

Let's look at one really cool application of linear approximations before we finish things up in this section.

In approximately 60 AD, Heron of Alexandria presented a method for approximating square roots (probably...historians know very little about exactly when Heron was born and died, but they think he saw an eclipse that matched one from 62 AD, so it's a good guess). This algorithm was presented along with different formulas for volumes and surface areas of a mixture of objects.

Aside.

You might know of Heron from Heron's formula for the area of a triangle!

More than 1000 years later, in the late 1660's, Isaac Newton was one of a long list of mathematicians to recreate this formula in a more general way, where we can use it to approximate roots of polynomials. This method was later extended by several different mathematicians, and is now known as the Newton-Raphson method, or sometimes more simply Newton's method.

Exercise 4.6.3: Walking in the Footsteps of Ancient Mathematicians.

Let's travel all the way back to the first (or maybe second) century AD and re-create Heron's method to approximate the value of $\sqrt{2}$. We'll develop this using modern calculus, and simple linear approximation. We're going to reframe the problem, and instead we're going to try to use a linear approximation of $f(x) = x^2 - 2$ to approximate the x-value where $f(x) = 0$. We know enough about quadratic functions to know that there are two values: $x = -\sqrt{2}$ and $x = \sqrt{2}$.

a. We're going to build a linear approximation of $f(x) = x^2 - 2$, and we need a reasonable center. Honestly, any integer will work, since we can evaluate f and f' really easily, but we want to find one that is close to $\sqrt{2}$. Let's center our approximation at $x = 2$. Find $f'(x)$, and then construct the linear approximation:

$$L(x) = f'(2)(x - 2) + f(2)$$

b. Now we know that $L(x) \approx f(x)$ for x-values near our center, $x = 2$. What if we estimate the x-value where $f(x) = 0$ by solving $L(x) = 0$ instead? Since $L(x) \approx f(x)$, the x-value where $L(x) = 0$ should make $f(x)$ pretty close to 0 at least. Solve $L(x) = 0$.

c. Ok, this might be kind of close to the value of $\sqrt{2}$, right? Let's visualize this. Hm...so this isn't that good of an approximation yet. We can check this by looking at the actual value of our function at $x = \frac{3}{2}$ and seeing if it's close to 0.

$$\begin{aligned} f\left(\frac{3}{2}\right) &= \left(\frac{3}{2}\right)^2 - 2 \\ &= \frac{9}{4} - 2 \\ &\frac{1}{4} \end{aligned}$$

This...isn't that close to 0. So let's try this again. This time, though, let's center our new linear approximation at $x = \frac{3}{2}$.

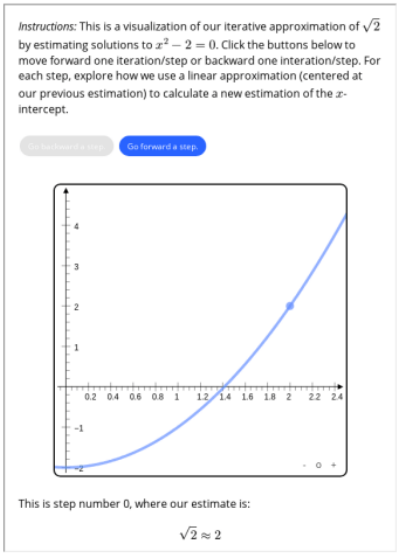
d. Now set this new linear approximation equal to 0 and solve $L(x) = 0$ to estimate the solution to $f(x) = 0$.

e. We can keep repeating this process, and that's exactly what the mathematicians we talked about discovered. Say we've built a linear approximation at some x-value (we'll call it x_{old}).

$$L(x) = f'(x_{\text{old}})(x - x_{\text{old}}) + f(x_{\text{old}})$$

Set this equal to 0 and solve.

f. Let's visualize these calculations.



Standalone
Embed

Something kind of strange happens in the last two steps. Why does the value of our estimation not change? What happens to our estimate?

Answer available at the end of this section.

Definition 4.6.2: Newton's Method for Approximating Zeros of Functions.

If x_0 is some initial estimation of a solution to $f(x) = 0$, then we can iteratively generate more estimations using the following formula

$$x_n = x_{n-1} - \frac{f(x_{n-1})}{f'(x_{n-1})}$$

provided that $f'(x_{n-1})$ exists and is non-zero.

A good question to ask is about when this process stops. If we want to estimate some x-intercept of a function, like $\sqrt{2}$ in Activity 4.6.3, then how many steps is enough? There are a couple of ways we can approach this:

1. We can just state at the beginning how many iterations we're going to do. This is what happened in Activity 4.6.3, since this activity was written to only make you calculate a specific number of these estimations. We could have started by saying that we'll calculate this 3 times, or maybe 100 times.
2. We can test to see what $f(x_n)$ is, and then stop when it is within some pre-determined distance from 0. We also did this when we noticed that $f\left(\frac{3}{2}\right) = \frac{1}{4}$ was not very close to 0 (after our first estimation), and so we should calculate this again. We could start by saying that we'll continue until we see a y-value that is within 0.0001 of 0, or some other small distance.
3. We can test to see how close our approximations are to each other, and stop when they're close enough. We saw this happen in the visualization: the last two estimates were the same! They actually weren't, but since the applet only displayed 4 decimal places, the numbers appeared the same after rounding. Maybe we set some criteria there, or we look at the distance between x_n and x_{n-1} (two successive estimates) and stop when they are within some distance from each other.

In reality, we often choose a combination of these. Maybe we set distance threshold for stopping, but use a maximum of 10 iterations as a backup plan. This happens often when we code this algorithm and have a computer run it. It is possible for this code to never give us two successive estimates that are close enough to stop, and so the code would run forever unless we cut it off at 100 iterations or some other value.

A wonderful thing about this small process is that, while it is ancient (dating back to Heron in the first or second century), it is still used today. This is a powerful estimation method that can be used in a variety of areas including statistics and data science.

Practice Problems

Exercise 4.6.1

What is the formula for the Linear Approximation of a function $f(x)$ centered at some x-value $x = a$? Explain what this is.

Exercise 4.6.2

or some function $f(x)$ that is differentiable everywhere, how could we use the concepts of linear approximation to find an approximate value $f(0.0001)$?

Exercise 4.6.3

Let's say that we have a function $f(x)$ that is differentiable at $x = 5$ and a Linear Approximation $L(x) = f'(5)(x - 5) + f(5)$. Give examples of two different x-values where $L(x)$ does a good job of approximating $f(x)$ accurately, and two examples of x-values where $L(x)$ does not do a good job of approximating $f(x)$ accurately.

Exercise 4.6.4

For each of the following functions and centers, build a Linear Approximation function.

a. $f(x) = \sqrt{x}$ centered at $x = 4$
b. $f(x) = \sin x$ centered at $x = 0$
c. $f(x) = \cos x$ centered at $x = 0$
d. $f(x) = \ln x$ centered at $x = 1$
e. $f(x) = \sqrt[3]{x}$ centered at $x = 8$
f. $f(x) = \dfrac{1}{x}$ centered at $x = 4$

Exercise 4.6.5

Use the linear approximating functions you created above to estimate the following values. Comment on the accuracy of these approximations.

a. $\sqrt{3.9}$
b. $\sin(-0.0002)$
c. $\cos(0.001)$
d. $\ln(2)$
e. $\sqrt[3]{9}$
f. $\dfrac{1}{4.0001}$

Exercise 4.6.6

Approximate each of the following values using a Linear Approximation. For each one:

- Report the function $f(x)$ and center.
- Write out the Linear Approximation function $L(x)$.
- Approximate the value.
- Justify whether your approximation is an over-approximation or under-approximation without referring to the actual value.

a. $(5.985)^2$
b. $(2.013)^3$
c. $\cos(3)$
d. $e^{0.01}$

Exercise 4.6.7

Explain how Newton's method uses Linear Approximation to approximate the x-intercept of a function.

Exercise 4.6.8

Explain how we cam use Newton's method to approximate:

a. the intersection point of two functions.
b. local maximums or minimums of a function.
c. inflection points of a function.

Exercise 4.6.9

Use Newton's method to find a zero for each function using the initial value given. Stop approximating when you get successive approximations that are within 0.001 of each other.

a. $f(x) = x^3 - 5$ with $x_0 = 2$
b. $f(x) = x\cos(x)$ with $x_0 = 2$

Exercise 4.6.10

Use Newton's method to find an intersection point between the two curves. Stop approximating when you get successive approximations that are within 0.001 of each other. Feel free to use a graphing utility like desmos.com to graph the curves and pick a reasonable stating value.

a. $y = x^2$ and $y = \sin(x)$ (Note: these curves intersect at $x = 0$. Find the second intersection point.)
b. $y = e^x$ and $y = \sqrt{x} + 2$

Exercise 4.6.11

Find/approximate the absolute maximum and absolute minimum of $y = x\sin(x)$ on [0, 6] using Newton's method (and any other techniques or concepts we've used).

Exercise 4.6.12

Find/approximate the inflection point of $y = \dfrac{x^3}{10} + e^x$ using Newton's method (and any other techniques or concepts we've used).

Answers

Exercise 4.6.3: Walking in the Footsteps of Ancient Mathematicians.

a. Since $f'(x) = 2x$, we have $f'(2) = 4$ and $f(2) = 2$. So then we end up with the following for our linear approximation:

$$L(x) = 4(x-2) + 2$$

b.
$$0 = 4(x-2)+2$$
$$-24(x-2)$$
$$-\frac{2}{4}x-2$$
$$2-\frac{1}{2}x$$
$$x=\frac{3}{2}$$

c.
$$L(x)f'\left(\frac{3}{2}\right)\left(x-\frac{3}{2}\right)+f\left(\frac{3}{2}\right)$$
$$=3\left(x-\frac{3}{2}\right)+\frac{1}{4}$$

d.
$$0=3\left(x-\frac{3}{2}\right)+\frac{1}{4}$$
$$-\frac{1}{4}=3\left(x-\frac{3}{2}\right)$$
$$-\frac{1}{12}=x-\frac{3}{2}$$
$$\frac{3}{2}-\frac{1}{12}$$
$$x=\frac{17}{12}$$

e.
$$0=f'\left(x_{\text{old}}\right)\left(x-x_{\text{old}}\right)+f\left(x_{\text{old}}\right)$$
$$-f\left(x_{\text{old}}\right)=f'\left(x_{\text{old}}\right)\left(x-x_{\text{old}}\right)$$
$$-\frac{f\left(x_{\text{old}}\right)}{f'\left(x_{\text{old}}\right)}x-x_{\text{old}}$$
$$x=x_{\text{old}}-\frac{f\left(x_{\text{old}}\right)}{f'\left(x_{\text{old}}\right)}$$

This page titled 4.6: Linear Approximations is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Peter Keep.

4.7: L'Hôpital's Rule

We're going to revisit limits, but with a slightly new problem-solving tool. Specifically, we'll be thinking about Indeterminate Forms. We noticed, back in Section 1.3, that we could evaluate limits for indeterminate forms by swapping out the function with another function that was mostly equivalent, only differing at the x-value we were approaching in the limit (Theorem 1.3.3 Limits of (Slightly) Different Functions).

We ended that section by thinking about a limit where this was difficult, in Activity 1.3.3.

We're now going to build a more systematic (and helpful) way of thinking about these limits using the ideas of Linear Approximation!

Indeterminate Forms

We have given a preliminary definition of Indeterminate Forms already (Definition 1.3.4), but let's remember how these work.

We said that $\frac{0}{0}$ is an indeterminate form, since a limit whose numerator and denominator approach 0 can end up taking on different values or even not exist. For instance, we can notice that the definition of the Derivative at a Point is a limit with this indeterminate form. As long as $f(x)$ is continuous (a necessity of it being differentiable) at $x = a$, then:

$$\lim_{x\to a}\left(\frac{f(x)-f(a)}{x-a}\right) \overset{?}{\to} \frac{f(a)-f(a)}{a-a} = \frac{0}{0}$$

But we have seen so many different values that this limit can end up being! We have spent most of the past two chapters in this text playing with derivatives and evaluating them: all of those values come from this limit! We have also seen that, even for continuous functions, this limit may not exist. A function can be non-differentiable at $x = a$.

We can show the same thing for a second indeterminate form: $\frac{\pm\infty}{\pm\infty}$, which we will simplify by just using the symbol $\frac{\infty}{\infty}$. For us to see that limits with this form can take on different values (or not exist), we just need to think about end behavior limits for rational functions (End Behavior Limits).

Let's think about the following limit:

$$\lim_{x\to\infty}\left(\frac{2x^m+1}{1-3x^n}\right)$$

As long as $m, n > 0$, then this limit looks like it's in the form of $\frac{\infty}{\infty}$. Sure, the denominator is really approaching $-\infty$, but we really just mean that there is an infinite numerator and an infinite denominator, regardless of sign.

We also know that the actual limit depends on the degrees m and n! Try to spend a couple of minutes confirming the next few claims:

- If $m < n$, then this limit is 0.
- If $m = n$, then this limit is $-\frac{2}{3}$.
- If $m > n$, then this limit doesn't exist.

All of this to show us that we have some forms of limits where we can't immediately tell what the actual value of the limit is (or if it even exists). L'Hôpital's Rule will be a way for us to navigate these limits a little easier than before, in some cases.

L'Hôpital's Rule

Activity 4.7.1: Building L'Hôpital's Rule.

We're going to take a closer look at the indeterminate form, $\frac{0}{0}$, and use our new ideas of linear approximation to think about how these types of things work.

We're going to be working with the following limit:

$$\lim_{x\to a}\frac{f(x)}{g(x)}$$

where $f(x)$ and $g(x)$ are differentiable at $x=a$ (since we're going to want to build linear approximations of them).

a. Write out the linear approximations for both $f(x)$ and $g(x)$, both centered at $x=a$. We'll call them $L_f(x)$ and $L_g(x)$. **Hint**: We're just using the formula for Linear Approximation of a Function, but with $f(x)$ for one of them and $g(x)$ in the other.
b. Describe how well or how poorly these linear approximations estimate the values from our functions $f(x)$ and $g(x)$? What happens to these approximations as we get close to the center $x=a$? What happens in the limit as $x\to a$? **Hint**: You can revisit the local linearity visualization from Section 4.6 to see what happens, in general, with a linear approximation of a function as we zoom in on the center.
c. Let's rewrite our limit. We can replace $f(x)$ with our formula for its linear approximation, $L_f(x)$ and replace $g(x)$ with its linear approximation, $L_g(x)$:

$$\lim_{x\to a}\frac{f(x)}{g(x)}=\lim_{x\to a}\left(\text{——}\right)$$

d. Up until now, we have not thought about indeterminate forms at all. Let's start now. If this limit is a $\frac{0}{0}$ indeterminate form, then that means that $\lim_{x\to a}f(x)=0$ and $\lim_{x\to a}g(x)=0$. Since our functions are, by definition, differentiable at $x=a$, then they also have to be continuous at $x=a$. What does this mean about the values of $f(a)$ and $g(a)$? **Hint**: Take a look back at our definition of function being Continuous at a Point. How does the function value relate to the limit? What does that mean in our case?
e. Use this new information about the values of $f(a)$ and $g(a)$ to revisit the limit. We rewrote $\lim_{x\to a}\frac{f(x)}{g(x)}$ by replacing each function with its linear approximation. What happens with the algebra when we know this information about $f(a)$ and $g(a)$? So we have a really nice result here! In the $\frac{0}{0}$ indeterminate form, we can replace the ratio of the y-values from our functions with the ratio of slopes (coming from the first derivatives) of our functions. In general, we'll put a step in between, where we find $f'(x)$ and $g'(x)$ first before trying to evaluate these derivatives at $x=a$. **Hint**: We rewrote $\lim_{x\to a}\frac{f(x)}{g(x)}=\lim_{x\to a}\left(\frac{f'(a)(x-a)+f(a)}{g'(a)(x-a)+g(a)}\right)$ and also know that in this case (due to the $\frac{0}{0}$ indeterminate form) that $f(a)=0$ and $g(a)=0$

Theorem 4.7.1: L'Hôpital's Rule.

If $f(x)$ and $g(x)$ are functions and a is some real number with f and g both being differentiable at a and $\lim_{x\to a}f(x)=0$ and $\lim_{x\to a}g(x)=0$, then

$$\lim_{x\to a}\frac{f(x)}{g(x)}=\lim_{x\to a}\frac{f'(x)}{g'(x)}$$

Similarly, this holds if $\lim_{x\to a}f(x)=\pm\infty$ and $\lim_{x\to a}(gx)=\pm\infty$. If f and g are both differentiable as $x\to\infty$ and either:

- $\lim_{x\to\infty}f(x)=0$ and $\lim_{x\to\infty}g(x)=0$
- $\lim_{x\to\infty}f(x)=\pm\infty$ and $\lim_{x\to\infty}g(x)=\pm\infty$

then

$$\lim_{x\to\infty}\frac{f(x)}{g(x)}=\lim_{x\to\infty}\frac{f'(x)}{g'(x)}$$

This is also true as $x\to-\infty$.

Example 4.7.2: Some First Limits.

Evaluate the following limits. You should first confirm that they are, actually, indeterminate forms!

a. $\lim_{x\to0}\left(\frac{\sin(x)}{x}\right)$
b. $\lim_{x\to5}\left(\frac{x^2-6x+5}{x-5}\right)$
c. $\lim_{x\to1}\left(\frac{\ln(x)}{x-1}\right)$
d. $\lim_{x\to\infty}\left(\frac{x^2-6x+5}{x-5}\right)$
e. $\lim_{x\to\infty}\left(\frac{\ln(x)}{x-1}\right)$

There are more indeterminate forms than these two! In each of the following cases, we mean that a limit with this form can take on different values (or not exist). Other indeterminate forms that we can consider include:

- $f(x)\cdot g(x)\overset{?}{\to}0\cdot\infty$
- $(f(x)-g(x))\overset{?}{\to}\infty-\infty$
- $f(x)^{g(x)}\overset{?}{\to}0^0$
- $f(x)^{g(x)}\overset{?}{\to}1^\infty$
- $f(x)^{g(x)}\overset{?}{\to}\infty^0$

The issue with these, though, is that L'Hôpital's Rule only applies to quotients! We needed that quotient for the algebra to work out when we canceled things out to end up with the ratio of slopes.

So our strategies for these other indeterminate forms will all require us to manipulate the product, difference, or exponential in order to force some division to show up somehow.

Forcing Division

Let's look at each new indeterminate form classified into groups based on the operation between the functions.

Products!

We can rewrite $f(x)\cdot g(x)$ as a quotient by dividing by a reciprocal. So either

$$f(x)\cdot g(x)=\frac{f(x)}{1/g(x)}$$

or

$$f(x)\cdot g(x)=\frac{g(x)}{1/f(x)}$$

Our choice ends up being based on what is most helpful.

Example 4.7.3

Evaluate the limit:

$$\lim_{x\to 0^+}(x\ln(x))$$

Note that since $x\to 0$ and $\ln(x)\to -\infty$, this is a $0\cdot\infty$ indeterminate form.

Hint: Rewrite this limit as:

$$\lim_{x\to 0^+}\left(\frac{\ln(x)}{1/x}\right)$$

Note that this is not an $\frac{\infty}{\infty}$ indeterminate form, and we can use L'Hôpital's Rule.

Solution

$$\begin{aligned}\lim_{x\to 0^+}(x\ln(x)) &= \lim_{x\to 0^+}\left(\frac{\ln(x)}{1/x}\right)\\ &= \lim_{x\to 0^+}\left(\frac{\frac{d}{dx}(\ln(x))}{\frac{d}{dx}(1/x)}\right)\\ &= \lim_{x\to 0^+}\left(\frac{1/x}{-1/x^2}\right)\\ &= \lim_{x\to 0^+}\left(\frac{1}{x}\right)\left(-\frac{x^2}{1}\right)\\ &= \lim_{x\to 0^+}(-x)\\ &= 0\end{aligned}$$

So $\lim_{x\to 0^+} x\ln(x)=0$.

Differences!

We can rewrite $f(x)-g(x)$ as a product by factoring something out of the difference. Then, if the product is a $0\cdot\infty$ indeterminate form, we can divide by a reciprocal to turn it into a quotient.

Choosing what to factor out is sometimes very difficult. But we should note that this is the strategy we used to evaluate Polynomial End Behavior Limits.

Example 4.7.4

Evaluate the following limit:

$$\lim_{x\to\infty}\left(2^x-x^2\right)$$

Note that since $2^x\to\infty$ and $x^2\to\infty$, this is an $\infty-\infty$ indeterminate form.

Hint: Try to factor out 2^x. You won't be able to actual factor it nicely, but you'll end up with a fraction term $\frac{x^2}{2^x}$ that is an $\frac{\infty}{\infty}$ indeterminate form!

Solution

$$\lim_{x\to\infty}\left(2^x-x^2\right)=\lim_{x\to\infty}2^x\left(1-\frac{x^2}{2^x}\right)$$

Let's focus on the limit $\lim_{x\to\infty}\left(\frac{x^2}{2^x}\right)$, since it is in an $\frac{\infty}{\infty}$ indeterminate form.

$$\begin{aligned}\lim_{x\to\infty}\left(\frac{x^2}{2^x}\right) &\overset{?}{\to} \frac{\infty}{\infty}\\ &= \lim_{x\to\infty}\frac{\frac{d}{dx}\left(x^2\right)}{\frac{d}{dx}\left(2^x\right)}\\ &= \lim_{x\to\infty}\frac{2x}{2^x\ln(2)}\\ \lim_{x\to\infty}\frac{2x}{2^x\ln(2)} &\overset{?}{\to} \frac{\infty}{\infty}\\ &= \lim_{x\to\infty}\frac{\frac{d}{dx}(2x)}{\frac{d}{dx}\left(2^x\ln(2)\right)}\\ &= \lim_{x\to\infty}\frac{2}{2^x\ln(2)\ln(2)}\\ \lim_{x\to\infty}\frac{2}{2^x\ln(2)\ln(2)} &\overset{?}{\to} \frac{2}{\infty} = 0\end{aligned}$$

So then we can go back to our original limit:

$$\begin{aligned}\lim_{x\to\infty}\left(2^x - x^2\right) &= \lim_{x\to\infty} 2^x\left(1-\frac{x^2}{2^x}\right)\\ &= \infty(1-0) = \infty\end{aligned}$$

Exponentials!

We can think about how we approached these types of functions raised to other functions when we learned about Logarithmic Differentiation.

We were able to use logarithms to rewrite these types of exponentials as products. So we can say that:

$$\begin{aligned}f(x)^{g(x)} &= e^{\ln\left(f(x)^{g(x)}\right)}\\ &= e^{g(x)\ln(f(x))}\end{aligned}$$

When we think about limits, the continuity of the exponential function allows us to just focus on the limit of the exponent, $g(x)\ln(f(x))$, which is likely an indeterminate form that we've seen!

Example 4.7.5

a. Evaluate the following limit:

$$\lim_{x\to 0^+} x^x$$

Note that this is the 0^0 indeterminate form. **Hint**: We can rewrite x^x as $e^{\ln(x^x)}$ which is the same as $e^{x\ln(x)}$. Now we can evaluate the limit $\lim_{x\to 0^+} x\ln(x)$, and make sure to return the value into the exponent.

b. Evaluate the following limit:

$$\lim_{x\to\infty}\sqrt[x]{x}$$

Note that this is the 0^∞ indeterminate form. **Hint**: Rewrite this as $\lim_{x\to\infty} x^{\frac{1}{x}}$. Now you can use the exponential/log composition to rewrite this again.

Solution

a. We know from Example 4.7.3 that $\lim_{x\to 0^+} x\ln(x) = 0$. So then:

$$\begin{aligned}\lim_{x\to 0^+} x^x &= \lim_{x\to 0^+} e^{x\ln(x)} \\ &= e^{\lim_{x\to 0^+} x\ln(x)} \\ &= e^0 = 1\end{aligned}$$

So $\lim_{x\to 0^+} x^x = 1$.

b.

$$\begin{aligned}\lim_{x\to\infty} \sqrt[x]{x} &= \lim_{x\to\infty} x^{\frac{1}{x}} \\ &\lim_{x\to\infty} e^{\frac{1}{x}\ln(x)} \\ &= e^{\lim_{x\to\infty} \frac{1}{x}\ln(x)}\end{aligned}$$

Now we can notice that the limit in the exponent is in a $\frac{\infty}{\infty}$ indeterminate form.

$$\begin{aligned}\lim_{x\to\infty} \frac{\ln(x)}{x} &= \lim_{x\to\infty} \frac{1/x}{1} \\ &= 0\end{aligned}$$

So then we have:

$$\begin{aligned}\lim_{x\to\infty} \sqrt[x]{x} &= e^{\lim_{x\to\infty} \frac{\ln(x)}{x}} \\ &e^0 \\ &= 1\end{aligned}$$

Practice Problems

Exercise 4.7.1

Explain how to use L'Hôpital's Rule to evaluate a limit with the following indeterminate forms:

a. $0/0$
b. ∞/∞
c. $0\cdot\infty$
d. $0^0, 1^\infty$, or ∞^0

Exercise 4.7.2

Consider the following limit: $\lim_{x\to -2}\left(\frac{x+2}{x^2-4}\right)$

a. Evaluate this limit using L'Hôpital's Rule.
b. Evaluate this limit without using L'Hôpital's Rule. Instead, use some algebraic manipulation of the function.

Exercise 4.7.3

Consider the following limit: $\lim_{x\to 3}\left(\frac{2+\frac{3}{x}-x}{x-3}\right)$

a. Evaluate this limit using L'Hôpital's Rule.

b. Evaluate this limit without using L'Hôpital's Rule. Instead, use some algebraic manipulation of the function.

Exercise 4.7.4

Consider the following limit: $\lim_{x\to\infty}\left(\dfrac{2x^3-x^2+9x-1}{10-5x+3x^2-8x^4}\right)$

a. Evaluate this limit using L'Hôpital's Rule.
b. Evaluate this limit without using L'Hôpital's Rule. Instead, use some algebraic manipulation of the function.

Exercise 4.7.5

Give an example of:

a. A limit with a 0/0 indeterminate form that we wouldn't be able to evaluate using the basic algebraic manipulations we have. Use L'Hôpital's Rule to evaluate that limit.
b. A limit with an ∞/∞ indeterminate form, and evaluate it using L'Hôpital's Rule.
c. A limit with a 0/0 indeterminate form that equals 6.
d. A limit with an ∞/∞ indeterminate form that equals –2.

Exercise 4.7.6

Evaluate the following limits using L'Hôpital's Rule.

a. $\lim_{x\to\infty}\left(\dfrac{e^{2x}+4x-2}{x^2+1}\right)$
b. $\lim_{x\to 3}\left(\dfrac{x^3-x^2-4x-6}{x^2-5x+6}\right)$
c. $\lim_{x\to\infty}\left(\dfrac{\ln x}{x^2}\right)$
d. $\lim_{x\to 0}\left(\dfrac{\sin(x)+x}{\cos(x)-2x}\right)$
e. $\lim_{x\to 0}\left(\dfrac{e^x-\dfrac{x^3}{6}-\dfrac{x^2}{2}-x-1}{\cos x-\sin x-\dfrac{x^3}{6}+\dfrac{x^2}{2}+x-1}\right)$
f. $\lim_{x\to\infty}\left(\dfrac{x^2+1}{e^{4x+1}}\right)$

Exercise 4.7.7

Evaluate the following limits using L'Hôpital's Rule.

a. $\lim_{x\to 0^+}(x\ln x)$
b. $\lim_{x\to 0^+}(\sin(x)\ln(x))$
c. $\lim_{x\to\infty}(x^2e^{-x})$
d. $\lim_{x\to\infty}(e^{-x}\sqrt{x})$
e. $\lim_{x\to\infty}x^{1/x}$
f. $\lim_{x\to 0^+}x^x$
g. $\lim_{x\to 0}(\cos(x))^{(1/x)}$
h. $\lim_{x\to\infty}\left(1+\dfrac{1}{x}\right)^x$
i. $\lim_{x\to\infty}\left(1+\dfrac{a}{x}\right)^x$ with a a real number.

j. $\lim_{x\to\infty}(\ln(x))^{1/x}$

This page titled 4.7: L'Hôpital's Rule is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Peter Keep.

4.8: Exploration- Applications of Derivatives

We have a lot of practice with calculating derivatives, and we hopefully have a good idea of how we might interpret what this derivative is: a function that tells us slopes of a line tangent to a curve at a specific x-value.

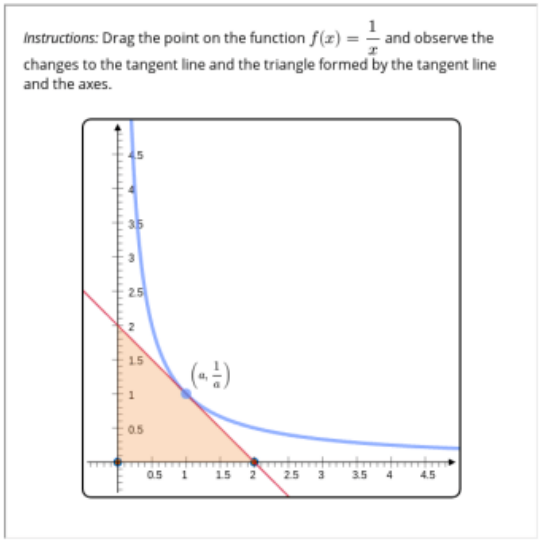
Let's think about this idea.

Exploration 4.8.1: A Tangent Triangle.

Let's consider the function $f(x) = \frac{1}{x}$.

a. Consider a point on this curve at some x-value $x = a$ where $a > 0$. Find the equation of the line tangent to $f(x)$ at this point.

b. We can graph this line, as well as the function $f(x)$.



Standalone Embed

Find the area of the triangle formed by the tangent line and the axes.

c. Hopefully you found something surprising about this area. Show that this surprising fact also occurs for the function $g(x) = \frac{k}{x}$, where $k > 0$.

d. Are there other functions that have this interesting property about the area of a triangle formed by the tangent line and the axes? What are some requirements that these functions would need to have?

This page titled 4.8: Exploration- Applications of Derivatives is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Peter Keep.

5: ANTIDERIVATIVES AND INTEGRALS

5.1: Antiderivatives and Indefinite Integrals
5.2: Riemann Sums and Area Approximations
5.3: The Definite Integral
5.4: The Fundamental Theorem of Calculus
5.5: More Results about Definite Integrals
5.6: Introduction to u Substitution

This page titled 5: Antiderivatives and Integrals is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Peter Keep.

5.1: Antiderivatives and Indefinite Integrals

We've been spending a lot of time thinking about derivatives! We've done this in a couple of different ways:

1. We have thought carefully about what derivatives are, what they measure, and how to interpret them.
2. We have built up a whole list of tools that we can use to actually find or calculate these derivatives. We know how to differentiate most functions (and combinations of functions) that we can think of!
3. We've been able to apply these derivatives to some specific contexts to solve problems or analyze functions and mathematical models.

Let's think about derivatives in a slightly different way!

Activity 5.1.1: Find a Function Where....

For each of the following derivatives, find a function $f(x)$ whose first (or second) derivative matches the listed derivative.

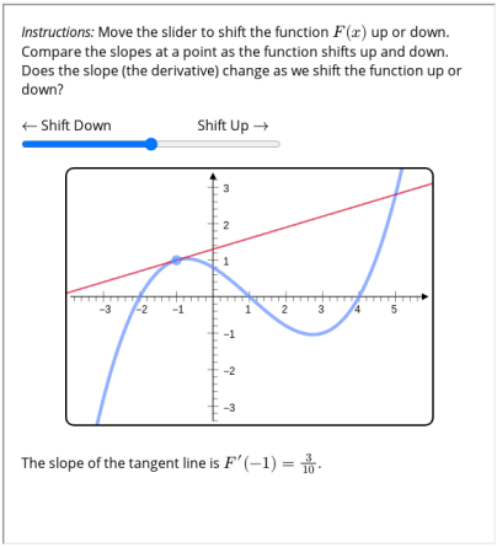
a. $f'(x) = 4x^3 + 6$
b. $f'(x) = 8x^7 - x^4 + x$
c. $f'(x) = \sqrt{x} - \dfrac{1}{x^2}$
d. $f''(x) = x^3 + \cos(x)$
e. $f''(x) = e^x - \dfrac{1}{\sqrt[3]{x}}$
f. Go back through each of the above derivatives, and find a different option for $f(x)$ that still works. Make sure that it is something completely unique, and not just an equivalent function that is written differently.

Why are you able to find multiple answers in these questions, but not when we are given a function and need to find a derivative?

We've done two things here: thought about how we might "undo" differentiation, and discovered a nice property about constants.

Note that we've already discovered this rule! We proved it, back when we were playing with the Mean Value Theorem. We built a related theorem that showed that two functions can have the same derivative, and if they do then they are off by, at most, a constant: Theorem 4.1.7 Equal Derivatives Correspond with Related Functions.

Let's visualize this phenomena!



Standalone
Embed

Antiderivatives

We want to try to define and name these "backwards derivatives." Instead of calling them the "negative first" derivative, we will name them as **antiderivatives**.

Definition 5.1.1: Antiderivative.

For a function $f(x)$, we say that $F(x)$ is an **antiderivative** of $f(x)$ on an interval if $F'(x) = f(x)$ on the interval.

We call $F(x) + C$ the **family of antiderivatives** of $f(x)$, where C represents any real number constant.

Example 5.1.2

For each of the following functions, find the family of antiderivatives.

a. $f(x) = 7x + \sec^2(x)$ **Hint**. Do we know a function whose derivative is $\sec^2(x)$?

b. $g(x) = \dfrac{5}{x} - \dfrac{3}{1+x^2}$ **Hint**. We won't be undoing the Power Rule with either of these! We might try to think about functions whose derivatives are $\dfrac{1}{x}$ and $\dfrac{1}{1+x^2}$.

c. $j(x) = x^5 - 4x + 1 - \dfrac{4}{3x^5}$ **Hint**. It might help to write the function as $j(x) = x^5 - 4x + 1 - \dfrac{4}{3}x^{-5}$.

Solution

a. $F(x) = \dfrac{7x^2}{2} + \tan(x) + C$

b. $G(x) = 5\ln|x| - 3\tan^{-1}(x) + C$ We use absolute values in the logarithm because we want to find a function whose derivative is $g(x)$ on the whole interval that $g(x)$ is defined. The log function is only defined for positive inputs, but we would like to be able to put any non-zero input into our function (since that's the domain of g).

c. $J(x) = \dfrac{x^6}{6} - 2x^2 + x + \dfrac{1}{3x^4} + C$

Initial Value Problems

Activity 5.1.2: A File Sorting Speed Test.

A computer program is trying to sort a group of computer files based on their size. The program isn't very efficient, and the time that it takes to sort the files increases when it tries to sort more files. The time that it takes, measured in seconds, based on the total, cumulative size of the files g, measured in gigabytes, is modeled by a function $T(g)$. We don't know the function, but we do know that the time increases at an instantaneous rate of $0.0001g$ seconds when the total size, g increases slightly. That is, the rate of change is a function of the file size, g, itself.

a. We can build a function for $T'(g)$. What is it?
b. Find all of the possibilities for the function modeling the time, T, that it takes the computer program to sort files with a total size of g. **Hint**. We are looking for the family of antiderivatives of $T'(g)$.
c. What does your constant C represent, here? You can interpret it graphically, interpret it by thinking about derivatives, but you should also interpret it in terms of the time that it takes this program to sort these files by size.
d. Let's say that we feed some number of files totaling up to 1.4GB in size into this program. It takes 0.24 seconds to sort the files by size. Find the function, $T(g)$, that models how quickly this program sorts these files.

Answer available at the end of this section.

We call this type of problem an "initial value problem." Here, we ended up solving for a family of antiderivatives, but then using some more information about that antiderivative (in this case, information about file size and time) to find the specific antiderivative function that was relevant.

Solving Initial Value Problems.

For some function $f(x)$, if we want to find an antiderivative function $F(x)$ and we know some "initial value," $F(a)$, then we can find the exact antiderivative by:

1. Finding the family of antiderivatives: $F(x) + C$.
2. Using the initial value to solve for the constant C, by evaluating $F(x) + C$ at $x = a$ and solving the resulting equation.

Example 5.1.3

a. For $f(x) = \frac{x^5}{2} + \sin(x)$, find $F(x)$ where $F(0) = 3$.
b. For $g'(x) = e^x$, find $G(x)$ where $G(0) = 4$ and $g(0) = 2$.

Indefinite Integrals

To finish this out, we'll just build some notation that represents these families of antiderivatives. We can use words to describe them, but it will be helpful to introduce some quick way of writing this using notation.

Definition 5.1.4: Indefinite Integral.

An indefinite integral represents a family of antiderivatives:

$$\int f(x)dx = F(x) + C$$

where

- Z is a symbol directing us to find a family of antiderivatives (or integrate)
- $f(x)$ is called the integrand
- dx is a differential, and represents both the "end" of the integral as well as an indicator of what the input variable of the integrand should be (or what variable we antidifferentiate "with regard to").
- $F(x)$ is an antiderivative of $f(x)$ (where $F'(x) = f(x)$).
- C is called the "constant of integration" and represents any real number

Example 5.1.5

Find families of antiderivatives according to each of the following indefinite integrals.

a. $\int \left(\frac{4}{x} - \sqrt{x}\right) dx$

b. $\int (x+4)\left(x^2-7\right) dx$ **Hint**. While we do not know how to antidifferentiate products of functions yet, we can just multiply the integrand function!

$$(x+4)\left(x^2-7\right) = x^3 + 4x^2 - 7x - 28$$

Antidifferentiate this.

c. $\int \left(\frac{xe^x - 1}{x}\right) dx$ **Hint**. Similar to the previous problem, we do not know how to antidifferentiate quotients, but we can rewrite this function before we antidifferentiate!

$$\begin{aligned} \frac{xe^x - 1}{x} &= \frac{xe^x}{x} - \frac{1}{x} \\ &= e^x - \frac{1}{x} \end{aligned}$$

Antidifferentiate this!

All we have left to do now is to just formalize the antiderivative rules we've been intuitively building and using.

Theorem 5.1.6: Power Rule for Antiderivatives.

$$\int x^n dx = \frac{x^{n+1}}{n+1} + C$$

for $n \neq -1$

Theorem 5.1.7: Antiderivatives Related to the Exponential and Log Functions.

$$\int e^x dx = e^x + C$$

$$\int b^x dx = \frac{b^x}{\ln(b)} + C \qquad \text{for } b > 0 \text{ and } b \neq 1$$

$$\int \frac{1}{x} dx = \ln|x| + C$$

Theorem 5.1.8: Antiderivatives of Trigonometric Functions.

$$\int \sin(x) dx = -\cos(x) + C$$

$$\int \cos(x) dx = \sin(x) + C$$

Theorem 5.1.1: Combinations of Indefinite Integrals.

1. Sums: $\int (f(x) + g(x))dx = \int f(x)dx + \int g(x)dx$
2. Differences: $\int (f(x) - g(x))dx = \int f(x)dx - \int g(x)dx$
3. Coefficients: $\int (kf(x))dx = k\int f(x)dx$

These should all be very familiar, since they are really just restatements of the results from Section 2.3 Some Early Derivative Rules.

We should also be comfortable recognizing derivatives of functions that we know, in order to find more functions that we can antidifferentiate.

If we are following the path set out by us already when we learned about derivatives, then at some point we will need to think about how to interpret these antiderivatives. What does $F(x)$ tell us about $f(x)$?

What does $f(x)$ tell us about $f'(x)$? We're probably so used to thinking about what $f'(x)$ tells us about $f(x)$ that it might be hard to reverse the interpretation. And that's ok!

Instead of worrying about uncovering this connection naturally, let's just take a different approach, by stating the relationship first: Over the next few sections, we'll discover that antiderivatives of $f(x)$ are deeply connected to areas carved out by the graph of $f(x)$.

Practice Problems

Exercise 5.1.1

Our definition of an antiderivative is that $F(x)$ is an antiderivative of $f(x)$ if $F'(x) = f(x)$. Why does this not mean that $F(x)$ is the antiderivative of $f(x)$?

Exercise 5.1.2

Let's consider the connection between derivatives, functions, antiderivatives, and indefinite integrals.

a. For a first derivative $f'(x) = 3x^2 + \cos(x) + 2$, find a suitable function $f(x)$
b. For a second derivative $f''(x) = 3x^2 + \cos(x) + 2$, find a suitable first derivative $f'(x)$
c. For a function $f(x) = 3x^2 + \cos(x) + 2$, find a suitable antiderivative function $F(x)$
d. For a function $f(x) = 3x^2 + \cos(x) + 2$, find the family of antiderivatives.
e. Find $\int (3x^2 + \cos(x) + 2)\, dx$

Exercise 5.1.3

Explain the difference between a single antiderivative of a function $g(x)$ and the family of antiderivatives of $g(x)$.

Exercise 5.1.4

What does a family of antiderivatives look like, graphically?

Exercise 5.1.5

For each of the following indefinite integrals, find the family of antiderivatives.

a. $\int \left(4x^3 - \frac{1}{x}\right) dx$
b. $\int \left(8x - \sec^2(x) - \frac{1}{x^2}\right) dx$
c. $\int \left(x^5 - 3\sin(x) + 2e^x\right) dx$
d. $\int \left(\sqrt[3]{t} + \frac{4}{t^3} + 2\right) dt$
e. $\int \left(\frac{4y}{3} - \frac{3}{4y}\right) dy$
f. $\int \left(5\sec(x)\tan(x) + 2\sqrt[5]{x^3}\right) dx$
g. $\int \left(\frac{1}{\sqrt{1-x^2}} + \frac{4x^2}{3}\right) dx$
h. $\int \left(\frac{1}{1+x^2} - \cos(x)\right) dx$

Exercise 5.1.6

For each of the following indefinite integrals, rewrite the integrand (the function inside the integral) before finding a family of antiderivatives.

a. $\int \left(\frac{x^2-1}{x-1}\right) dx$
b. $\int \left(x\left(x^2+1\right)\right) dx$
c. $\int \left(\frac{\sec(x)+\tan(x)}{\cos(x)}\right) dx$
d. $\int \left(\frac{xe^x+1}{x}\right) dx$
e. $\int \left(\ln\left(e^{2x}\right)\right) dx$
f. $\int \left(\frac{x^2+2}{x^2+1}\right) dx$ **Hint**. It might be helpful to write x^2+2 as $(x^2+1)+1$. Then, try to split up that fraction.

Exercise 5.1.7

For each of the functions below, find the antiderivative that satisfies the condition given.

a. $f(x) = x^2 - 4\sin(x)$ with $F(0) = 3$
b. $g(x) = x^3 - 4x^2 + 5x$ with $G(1) = 0$
c. $m(x) = \frac{9}{x} - 3e^x$ with $M(1) = 4$

Answers

Activity 5.1.2: A File Sorting Speed Test.

a. $T'(g) = 0.0001g$

This page titled 5.1: Antiderivatives and Indefinite Integrals is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Peter Keep.

5.2: Riemann Sums and Area Approximations

One of the last things we said in Section 5.1 was that antiderivatives will be connected to areas. We're going to eventually show this! For now, though, we want to focus on areas defined by curves.

Activity 5.2.1: Approximating Areas.

We're going to consider two different functions, and some areas based on them. Let's think about two functions: $f(x) = 2x + 1$ and $g(x) = x^2 + 1$. For both of these functions, we'll focus on the interval [0, 2]. Instead of thinking about the function only, we'll be considering the two-dimensional region bounded between the graph of our function and the x-axis between $x = 0$ and $x = 2$.

a. Find the area of the region bounded between the graph of $f(x) = 2x + 1$ and the x-axis between $x = 0$ and $x = 2$.

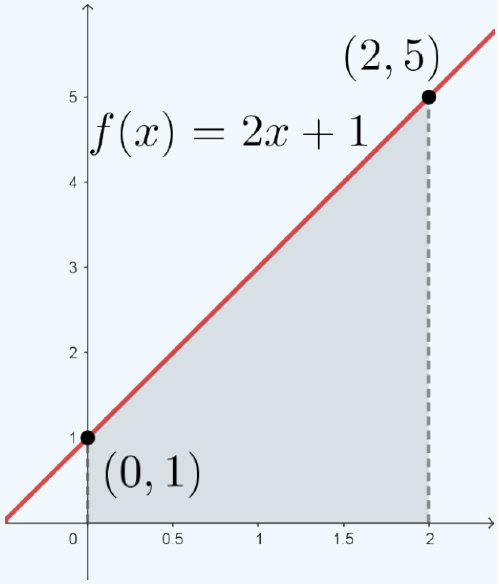


Figure 5.2.1

How did you evaluate this area? What kind(s) of shape(s) did you think about?

b. Estimate the area of the region bounded between the graph of $g(x) = x^2 + 1$ and the x-axis between $x = 0$ and $x = 2$.

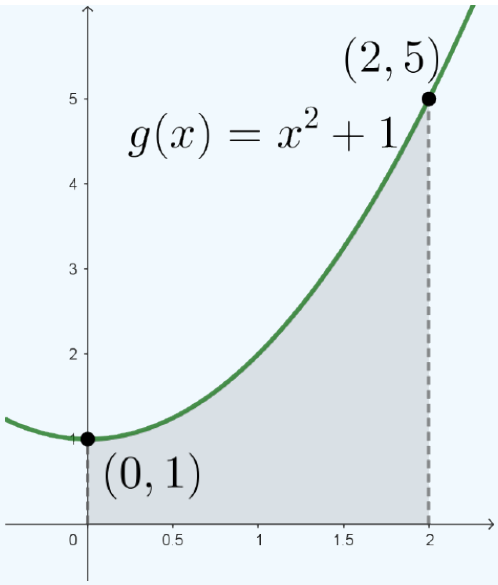


Figure 5.2.2

How did you estimate this area? What kind(s) of shape(s) did you think about?

c. Come up with an upper and lower bound for this area. In other words, give an underestimate and overestimate for the actual area we would like to know. How did you come up with these estimates? How "good" do you think your estimates are? Can you come up with "better" (or closer) ones?

Hopefully we've had a chance to think about and compare a couple of different strategies for estimating this area. What we want to do, though, is build a systematic way of estimating this area. We'd like it to have a couple of features:

- Easy area calculations. We don't want to have to spend a lot of time thinking about tricky area formulas, so simple shapes will be nicer to use.
- Flexibility. We want to be able to apply our approach to an area defined by any curve.
- Precision. We want to be able to make our estimates as precise as we'd like. It's fine to come up with rough estimates, but we would like a method that allows us to increase the accuracy in our estimations.

Aside.

Archimedes of Syracuse discovered how to calculate this area exactly, without estimation, around 300 BC, writing his results in the now-famous "Quadrature of the Parabola." This is, notably, before the formalization of Calculus (during the 1600's). It might be unfair to say that Archimedes proved this "without using calculus," though, since his technique, the "Method of Exhaustion," is really a version of what we do in calculus, but without a formal framework of limits.

This is the problem with Archimedes' method: without a framework that could be flexible while maintaining precision, it only worked for areas defined by parabolas. Once we change our function to something else, Archimedes would have to come up with a completely new area formula for calculation. The techniques we're looking at now have the advantage of flexibility!

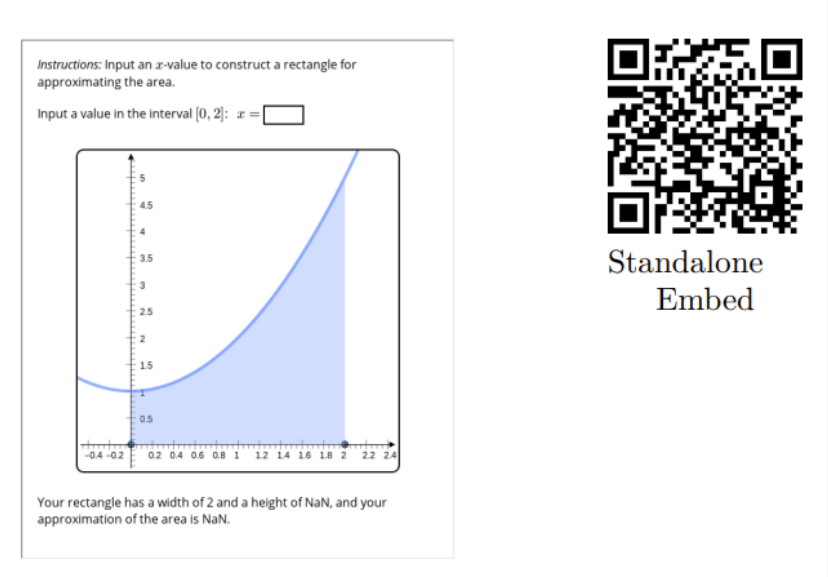
Rectangular Approximations

We're going to revisit the same region as before, but this time we'll outline a process that should help us approximate the area with as much precision as we'd like.

Activity 5.2.2: Approximating the Area using Rectangles.

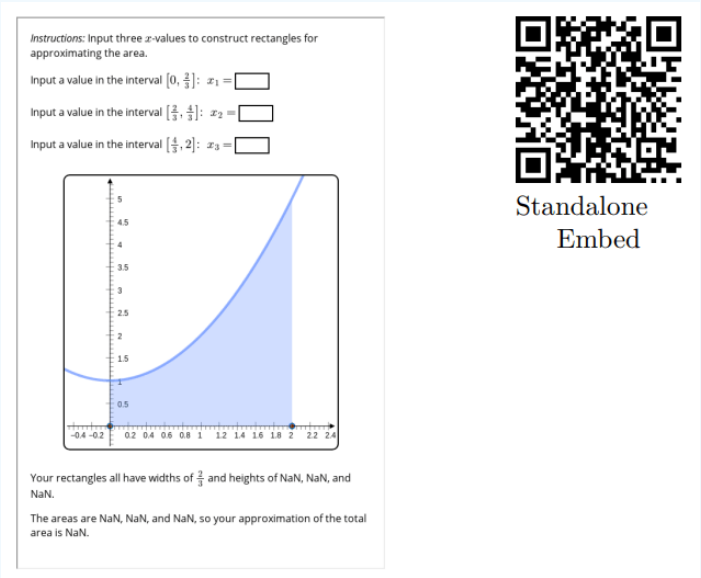
We're going to stick with the function $g(x) = x^2 + 1$ on the interval [0, 2], and keep thinking about the area bounded by the curve and the x-axis on this interval. We're going to approximate the area in a couple of different tries, each one more accurate than the one before. By the end of this activity, we'll have a pretty good process built!

a. Let's start with approximating this region with a single rectangle. We're going to define the rectangle by picking some x -value in the interval [0, 2]. Then, we'll use the point at that x-value to define the height of our rectangle. Essentially, we are picking a single point on our function on the interval, and our approximation is pretending that the single point we picked is representative of the whole function on the interval.



b. Can you try repicking an x-value, and try to find one that gets you an area approximation that is pretty good?

c. We're going to use more rectangles. Let's jump up to 3 rectangles. If we split up the interval between $x = 0$ and $x = 2$ into 3 rectangles, we can make them all the same width, and pick an x-value that we can use to get a representative point for each of the 3 rectangles. We'll need to pick 3 x-values this time.



d. Can you try repicking your x-values, and trying to find one that gets you an area approximation that is pretty good?

e. Let's scale this up a bit. Pick a good number for your number of rectangles. We'll call this value n. (If you're working in a classroom, maybe it would be good to pick the number of groups or the number of students, or some other number between 10 and 20 or something like that.) For your value n, we're going to divide up the interval between $x = 0$ and $x = 2$ into n pieces. These will be the intervals that we pick from to get our rectangles. What are the subintervals? What are the widths of each subinterval (and then the widths of the rectangles)? Call this with Δx.
f. For each subinterval, pick an$x$-value in the subinterval to represent it.
g. Evaluate the function $f(x) = x^2 + 1$ at each of the x-values you picked. These are the heights of your rectangles!
h. Find the areas of each rectangle by multiplying the height of each rectangle by Δx, the width of each rectangle.
i. Add these areas up to get a total approximation of the actual area! What do you think: is it worth fiddling with what x -value to pick from each subinterval to try to get a better approximation? If n is large, do you think it matters how we pick the x-values from each subinterval?

This is our process! We'll refer to it often as the **slice-and-sum process**, since we are slicing our region into a bunch of pieces, approximating the area on each piece (by using one point to represent the whole slice), and then summing the areas back up.

More formally, we can call this the Riemann Sum process, because the sum of the areas is a special form of summation.

Definition 5.2.3: Riemann Sum.

For a closed interval $[a, b]$ with a partition $\{x_0, x_1, \ldots, x_n\}$ with $a = x_0 < x_1 < \ldots < x_n = b$, consider some x_k^*, any x-value in the interval $[x_{k-1}, x_k]$ and Δx_k, the length of the interval $[x_{k-1}, x_k]$. If f is a function that is defined on the interval $[a, b]$, then we call the sum

$$\sum_{k=1}^{n} f(x_k^*)\,\Delta x_k = f(x_1^*)\,\Delta x_1 + f(x_2^*)\,\Delta x_2 + \ldots + f(x_n^*)\,\Delta x_n$$

a **Riemann Sum** for f on $[a, b]$.

Note 5.2.4

In practice, we typically choose a Regular Partition, where each subinterval $[x_{k-1}, x_k]$ is equally-wide, and so $\Delta x_k = \dfrac{b-a}{n}$ for every $k = 1, 2, \ldots, n$. We then normally write our Riemann sum as

$$\sum_{k=1}^{n} f(x_k^*)\,\Delta x = f(x_1^*)\,\Delta x + f(x_2^*)\,\Delta x + \ldots + f(x_n^*)\,\Delta x$$

where Δx is the value of the widths of all of the equally-sized subintervals.

Selection Strategies

This is great! We have a scalable way to approximate areas, and it seems like we can pretty easily increase the precision of our approximations by increasing n, the number of slices/rectangles that we use. And the great thing about this is that when we do increase n, we don't increase the complexity of our calculations!

Sure, it would be tedious to calculate and add 100 areas of rectangles by hand, but those area calculations don't get more difficult: there are just more of them.

The only real downside is that when we increase the number of slices/ rectangles, we are really increasing the number of decisions that we have to make: we have to choose an x_k^* for each subinterval, and so while it isn't hard to just calculate a bunch of areas and add them up, it is difficult, on a human level, to make a bunch of decisions about which x-value to choose from each subinterval. But this decision isn't even that important!

We use the "star" notation on the x_k^* to represent the fact that it really doesn't matter which x-value gets chosen from the

subinterval: as long as we pick one, we get an approximation! And when n increases, it matters less and less what the actual x-value is: as long as our function $f(x)$ is continuous, then there will be not much variation among the y-value outputs for any x-values in each (small) interval!

All of this to say: let's make a single decision about picking n x-values from n subintervals instead of having to make n decisions (one for each x-value).

Left, Right, and Midpoint Riemann Sums.

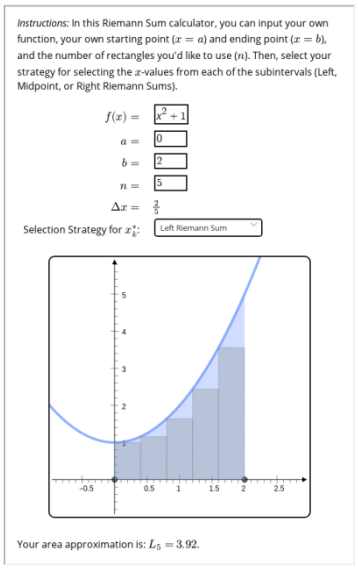
When we build a Riemann sum, we can make a choice to systematically choose the values for x_k^* (for $k = 1, 2, ..., n$). There are many ways of doing this, but here are three:

- Left Riemann Sum: We pick the left-most x-value from each subinterval. That is, if the partition is $\{a = x_0, x_1, x_2, \dots, b = x_n\}$, then we choose $\{a, x_1, x_2, \dots, x_{n-1}\}$ as our x-values to evaluate f at for the rectangle heights. We refer to these as L_n, a Left Riemann sum with n rectangles.
- Right Riemann Sum: We pick the right-most x-value from each subinterval. That is, if the partition is $\{a = x_0, x_1, x_2, \dots, b = x_n\}$, then we choose $\{x_1, x_2, \dots, b\}$ as our x-values to evaluate f at for the rectangle heights. We refer to these as R_n, a Right Riemann sum with n rectangles.
- Midpoint Riemann Sum: We pick the x-value that is in the middle of each subinterval. That is, if the partition is $\{a = x_0, x_1, x_2, \dots, b = x_n\}$, then we choose $\left\{\frac{a+x_1}{2}, \frac{x_1+x_2}{2}, \dots, \frac{x_{n-1}+b}{2}\right\}$ as our x-values to evaluate f at for the rectangle heights. We refer to these as M_n, a Midpoint Riemann sum with n rectangles.

None of this is a requirement for a Riemann sum, but we will consistently find that when we limit the number of decisions that we have to make, the complexity of the calculation decreases.

Notice that we've already made a similar choice with how we calculate Δx: it is not required that each rectangle have the same width, but it is very nice to not have to think about n different widths!

Lastly, we'll finish with a nice interactive Riemann sum calculator. Feel free to explore some different graphs and see how the Riemann sums work when we change how we select the values for x_k^* as well as when we change the number of rectangles, n.



Standalone
Embed

Practice Problems

Exercise 5.2.1

Explain in words how we will approximate the area bounded between a function and the x-axis between two x-values, $x = a$ and $x = b$.

Exercise 5.2.2

Consider the function $f(x) = x^2 + 1$ on the interval [0, 2].

a. Divide the interval [0, 2] into 5 equally sized subintervals.
b. What is the width of each subinterval, Δx?
c. Pick an x-value from each of the subintervals. Label them $x_1^*, x_2^*, \ldots, x_5^*$.
d. Find the function outputs at each of the x_k^* values for $k = 1, 2, \ldots, 5$, $f(x_1^*), f(x_2^*), \ldots, f(x_5^*)$.
e. Multiply each of the function values above by Δx.
f. Add each of the values above together.
g. Interpret this sum. What does this represent? What does it approximate?
h. Draw the Riemann sum you built above. Your graph should include the graph of $f(x) = x^2 + 1$ as well as 5 rectangles, with heights defined by the points you found above.

Exercise 5.2.3

Let's consider a function $f(x)$ that is continuous on some interval $[a, b]$.

a. In the construction of a Riemann sum, we use x-values notated x_k^*. What are these? Explain them.
b. What does it mean to construct a Left Riemann sum?
c. What does it mean to construct a Right Riemann sum?
d. What does it mean to construct a Midpoint Riemann sum?

Exercise 5.2.4

Consider the function g(x) with the following function values.

Table 5.2.5 Function values for $g(x)$

x	0	0.5	1	1.5	2	2.5	3	3.5	4
$g(x)$	0	2	-4	7	0.5	-1	3.5	-3	1

a. Approximate the area bounded between the curve $y = g(x)$ and the x-axis between $x = 0$ and $x = 4$ using a Right Riemann sum with $n = 4$ rectangles (R_4).
b. Approximate the area bounded between the curve $y = g(x)$ and the x-axis between $x = 0$ and $x = 4$ using a Left Riemann sum with $n = 4$ rectangles (L_4).
c. Approximate the area bounded between the curve $y = g(x)$ and the x-axis between $x = 0$ and $x = 4$ using a Midpoint Riemann sum with $n = 4$ rectangles (M_4).
d. Approximate the area bounded between the curve $y = g(x)$ and the x-axis between $x = 0$ and $x = 4$ using a Right Riemann sum with $n = 8$ rectangles (R_8).
e. Approximate the area bounded between the curve $y = g(x)$ and the x-axis between $x = 0$ and $x = 4$ using a Left Riemann sum with $n = 8$ rectangles (L_8).

Exercise 5.2.5

Evaluate the following Riemann sums to approximate areas described. Sketch a picture of the region and Riemann sum for each.

a. The area bounded by the graph of $f(x) = \frac{x^3}{2} - 3x^2 + 4x + 2$ and the x-axis between $x = 0$ and $x = 2$ using an L_5 Riemann sum.
b. The area bounded by the graph of $g(x) = \frac{x^3}{2} - 3x^2 + 4x + 2$ and the x-axis between $x = 2$ and $x = 4$ using an R_3 Riemann sum.
c. The area bounded by the graph of $j(x) = \sqrt{x} - x$ and the x-axis between $x = 0$ and $x = 1$ using an L_4 Riemann sum.
d. The area bounded by the graph of $j(x) = \sqrt{x} - x$ and the x-axis between $x = 0$ and $x = 1$ using an M_4 Riemann sum.
e. The area bounded by the graph of $j(x) = \sqrt{x} - x$ and the x-axis between $x = 0$ and $x = 2$ using an R_5 Riemann sum.

Exercise 5.2.6

Consider a function that is increasing on the interval $[a, b]$. Will a Left Riemann sum overapproximate or underapproximate the true area bounded between the curve of $f(x)$ and the x-axis on the interval $[a, b]$? What about a Right Riemann sum?

Exercise 5.2.7

Consider a function that is decreasing on the interval $[a, b]$. Will a Left Riemann sum overapproximate or underapproximate the true area bounded between the curve of $f(x)$ and the x-axis on the interval $[a, b]$? What about a Right Riemann sum?

Exercise 5.2.8

For each pair of Riemann sums, explain which one you think most accurately estimates the area under the curve $f(x)$ from $[a, b]$, or explain why it is impossible or difficult to tell.

a. L_5 or R_5
b. L_8 or M_8
c. M_6 or R_{10000}

This page titled 5.2: Riemann Sums and Area Approximations is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Peter Keep.

5.3: The Definite Integral

The big result from our last section on Riemann sums is not just that we can approximate areas by thinking about a bunch of small (thin) rectangles. The big result is that this strategy is scalable: we can increase n, the number of slices/rectangles, and essentially guarantee that, eventually, our approximations will be very accurate.

Now, we move from a concrete process for building rectangles to calculate areas to a more conceptual framework: what happens when $n \to \infty$?

Evaluating Areas (Instead of Approximating Them)

Our goal is to move from approximating area to evaluating areas: calculating the real value of the area of these regions bounded between curves and the x-axis. We have already decided (when we built the framework for Riemann sums and made scalability and precision in our estimates a focus) that the area we're interested in is the result of some limiting process: we increase the number of slices, n, and in turn decrease the width of each slice, Δx.

Definition 5.3.1: Definite Integral.

If $f(x)$ is some function defined on the interval $[a, b]$ and $\sum_{k=1}^{n} f\left(x_k^*\right) \Delta x$ is a Riemann sum with n slices and $\Delta x = \dfrac{b-a}{n}$, then we say that the definite integral of $f(x)$ from $x = a$ to $x = b$ is:

$$\int_{x=a}^{x=b} f(x)dx = \lim_{n\to\infty} \sum_{k=1}^{n} f\left(x_k^*\right) \Delta x$$

if this limit exists. When this limit exists, we say that $f(x)$ is integrable on the interval $[a, b]$. We call $x = a$ and $x = b$ the **limits of integration** for this definite integral, and we read $\int_{x=a}^{x=b} f(x)dx$ as "the integral from $x = a$ to $x = b$ of $f(x)$ with regard to x ," or sometimes we might just say "of $f(x)dx$" for short.

Note 5.3.2

This is assuming we're using a Regular Partition (Note 5.2.4). If we are not, and each slice has its own width called Δx_k, then the definition of a definite integral requires that as $n \to \infty$ we see $\Delta x_k \to 0$ for all $k = 1, 2, ..., n$. Essentially, we need all of the widths to eventually get tiny: we can't let one slice take up half of the width and then let all of the other slices get tiny, since that would still be an approximation of the area we want.

We don't need to worry about this, though, since we'll always just choose to make all of the Δx_k's the same size: $\Delta x = \dfrac{b-a}{n}$.

Let us make something very clear: we will absolutely not calculate these areas this way.

Let's see why not. Let's say we want to calculate $\int_{x=0}^{x=2} \left(x^2+1\right) dx$. This is the area we were estimating in Activity 5.2.2 Approximating the Area using Rectangles. How many slices did you pick at the end of this activity? How annoying was it to add up those areas?

Whatever you did, it's not enough: even if we decided to divide this region up into $n = 1000$ pieces, this is merely an approximation of the limit we want:

$$\int_{x=0}^{x=2} \left(x^2+1\right) dx = \lim_{n\to\infty} \sum_{k=1}^{n} \underbrace{\left(\left(x_k^*\right)^2+1\right)}_{f\left(x_k^*\right)} \underbrace{\left(\frac{2}{n}\right)}_{\Delta x}$$

There are some ways of evaluating this specific limit using some known formulas for sums of squares and end behavior limits of

rational functions. But these techniques are extremely limited: we might get lucky being able to fiddle with this limit of this sum for this function, but we won't be so lucky in general.

Instead, let's just think about these areas, focus on what types of functions are integrable, and then build towards our end goal of connecting these areas to antiderivatives.

Signed Area

We're going to now deal with the consequences of our decisions. A truth about mathematics, sometimes not an obvious truth, is that every time we state a definition, what we are actually doing is making a decision. We are deciding on some common way of classifying and describing an object. These classifications and descriptions are choices that we are making: choices to prioritize some property or aspect over a different one, choices to include or exclude a type of object into the group of things we're interested in, and choices that come with downstream effects.

We chose to define the area bounded between a curve defined by the function $f(x)$ and the x-axis between $x = a$ and $x = b$ as:

$$\int_{x=a}^{x=b} f(x)dx = \lim_{n\to\infty} \sum_{k=1}^{n} f\left(x_k^*\right) \Delta x$$

We are going to stand by this definition. It's a good one, for the reasons we described at the beginning of Section 5.2 Riemann Sums and Area Approximations.

But there are some weird things to notice. Let's notice them!

Activity 5.3.1: Weird Areas.

Let's think about a simple linear function, $f(x) = 4 - 3x$. We'll both approximate and evaluate the area bounded between $f(x)$ and the x-axis from $x = 0$ to $x = 3$:

$$\int_{x=0}^{x=3} (4 - 3x)dx$$

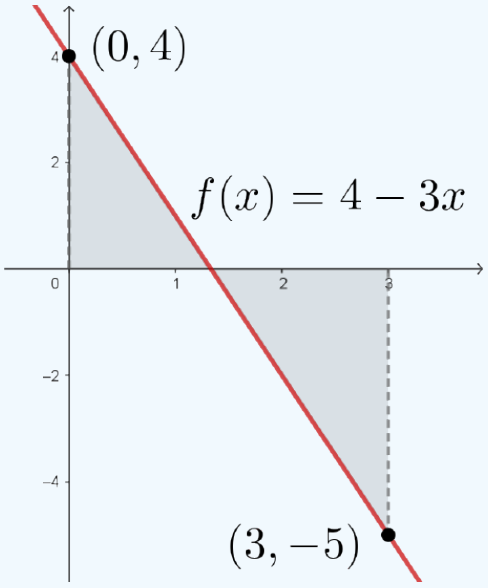


Figure 5.3.3

a. Explain why we do not need to think about Riemann sums in order for us to calculate the shaded area. How would you calculate this without using calculus? Calculate the area! **Hint**. Are there some shapes that you recognize? What are the dimensions of these shapes?

b. Let's approximate this area using a Riemann sum. Calculate L_3, the Left Riemann sum with $n = 3$ rectangles. **Hint**. You're going to divide up the interval from $x = 0$ to $x = 3$ into 3 subintervals: [0,1], [1,2], and [2,3]. Note that $\Delta x = 1$.Then you're picking the left-most x-value from each subinterval ($x_1^* = 0$, $x_2^* = 1$, and $x_3^* = 2$) to plug into $f(x)$ in order to find the heights of your rectangles.

c. Let's approximate this area a second time, but with a different selection strategy for our x-values. Calculate R_3, the Right Riemann sum with $n = 3$ rectangles. **Hint**. You're still going to divide up the interval from $x = 0$ to $x = 3$ into 3 subintervals: [0,1], [1,2], and [2,3]. Note that $\Delta x = 1$.Then you're picking the right-most x-value from each subinterval ($x_1^* = 1$, $x_2^* = 2$, and $x_3^* = 3$) to plug into $f(x)$ in order to find the heights of your rectangles.

d. Compare your answers for L_3 and R_3. They should be very different. Why? What is happening that makes R_3 specifically such a weird value?

e. Do you need to go back and recalculate the area geometrically (from the first part of this activity)? Explain why your answer for $\int_{x=0}^{x=3}(4-3x)dx$ should be negative, based on the Riemann sums we calculated. **Hint**. Did you account for "negative" area in the second triangular region in this integral?

f. Find a value $x = b$ such that: $\int_{x=0}^{x=b}(4-3x)dx = 0$.

g. Find a different value $x = b$ such that:

$$\int_{x=0}^{x=b}(4-3x)dx = 0$$

Is there a second way of making this area 0? **Hint**. Depending on what you did earlier, you might have to find some ending x-value that "balances" the area above and below the x-axis. If you already did this, then you might have to find an ending x-value that collapses this shape down to a 1-dimensional shape with no area.

Weird areas, right? Negative? That's not how we normally think about areas. So we have to be slightly careful with how we describe this new object, the definite integral, that we've built. We don't need to go back and change anything about the object itself: we just need to change how we talk about it.

It's common to think about $\int_{x=a}^{x=b} f(x)dx$ as the "area under the curve $f(x)$ from $x = a$ to $x = b$," but we know that's not really true. Instead, we'll think about it as a **signed area** of the region bounded between the curve $f(x)$ and the x-axis from $x = a$ to $x = b$. When we say "signed area," we're just referring to the consequence of using y-values to define "heights" of the rectangles: when the curve is under the x-axis, we end up with negative values for heights, and so those rectangles have negative area.

Activity 5.3.2: Weird Areas - Part 2.

We're going to think about the same region, kind of.

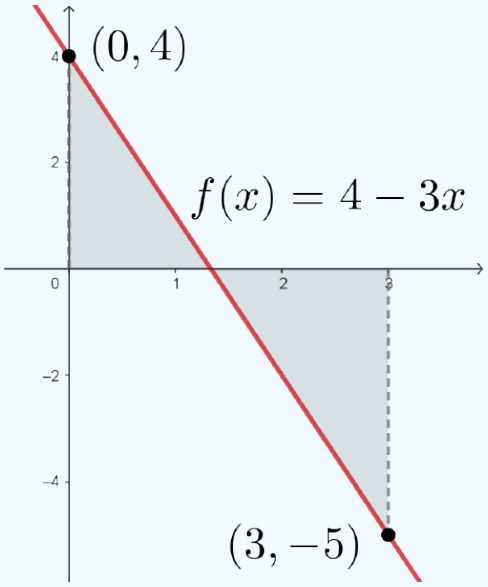


Figure 5.3.4

Let's think about the same linear function, $f(x) = 4 - 3x$, but this time we'll approximate and evaluate the area bounded between $f(x)$ and the x-axis from $x = 3$ to $x = 0$:

$$\int_{x=3}^{x=0} (4 - 3x)dx$$

a. Use geometry to calculate the area. Compare this to the result from Activity 5.3.1.
b. Let's approximate this using a Riemann sum. Calculate M_3, the Midpoint Riemann sum with $n = 3$ rectangles.
c. Do you need to go back and recalculate the area geometrically (from the first part of this activity)? Explain why your answer for $\int_{x=3}^{x=0}(4 - 3x)dx$ should be positive, based on the Riemann sums we calculated.

Ok, so we have some intuition about how the signs of these areas work, and we've also built up some nice properties that we can talk through. Let's finish this section by just summarizing some of the things we've done and thinking about what kinds of functions this works for!

Properties of Definite Integrals

First, this result should be reasonable: we can always calculate these areas for continuous functions!

Theorem 5.3.5: Continuous Functions are Integrable.

If $f(x)$ is continuous on the interval $[a, b]$, then $f(x)$ is integrable on $[a, b]$. That is, the limit $\lim_{n\to\infty} \sum_{k=1}^{n} f(x_k^*)\,\Delta x$ exists and so we can evaluate the definite integral:

$$\int_{x=a}^{x=b} f(x)dx$$

We'll come back to this, but first, let's summarize some properties that we've discovered.

Theorem 5.3.6: Properties of Definite Integrals.

If a, b, and c are real numbers and $f(x)$ is a function that is continuous on the intervals $[a, b]$ and $[b, c]$, then:

- The signed area under a single point is 0:
$$\int_{x=a}^{x=a} f(x)dx = 0$$
- We can cut a region into pieces and evaluate the areas separately:
$$\int_{x=a}^{x=c} f(x)dx = \int_{x=a}^{x=b} f(x)dx + \int_{x=b}^{x=c} f(x)dx$$
- When we integrate a function "backwards" through an interval, we get an area with an opposite sign:
$$\int_{x=a}^{x=b} f(x)dx = -\int_{x=b}^{x=a} f(x)dx$$

Ok, that's enough of this: let's get to the point and try to figure out how to actually calculate these areas without relying on our functions being "nice" enough that we can use geometry!

Exercise 5.3.1

Explain what $\int_{x=a}^{x=b} f(x)dx$ means geometrically/graphically.

Exercise 5.3.2

What does “signed” area or “net” area mean?

Exercise 5.3.3

Sketch the integral $\int_{x=0}^{x=2\pi}(\sin(x))dx$. Estimate this value from the sketch you drew. Explain your answer.

Exercise 5.3.4

Using the idea of Riemann sums, explain why area “under” the x-axis is considered negative.

Exercise 5.3.5

Using the idea of Riemann sums, explain why the $\int_{x=a}^{x=b} f(x)dx = -\int_{x=b}^{x=a} f(x)dx$

Exercise 5.3.6

For each of the following definite integrals, sketch the integral and use geometry to evaluate the definite integral. Feel free to use desmos to visualize the graphs of functions.

a. $\int_{x=0}^{x=4}(5)dx$
b. $\int_{x=0}^{x=3}(2x)dx$
c. $\int_{x=1}^{x=2}(x+6)dx$
d. $\int_{x=-3}^{x=3}\left(\sqrt{9-x^2}\right)dx$
e. $\int_{x=0}^{x=2}\left(-\sqrt{4-x^2}\right)dx$

Exercise 5.3.7

Consider the function $f(x)$, graphed below, with areas of different regions noted (without noting sign of the area).

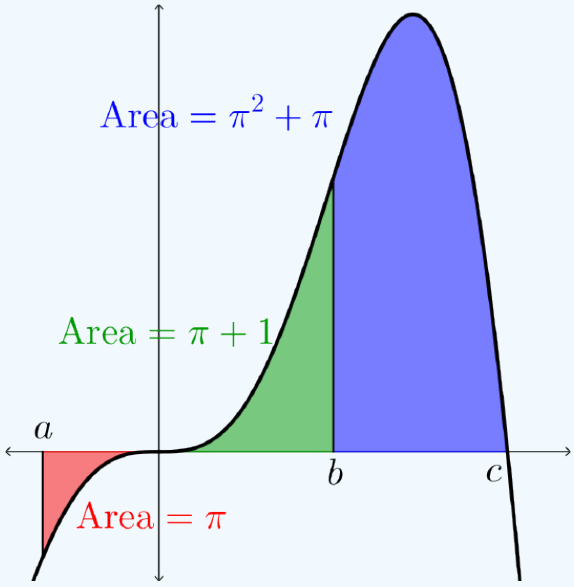


Figure 5.3.7 Graph of $f(x)$ with labeled areas.

Evaluate the following definite integrals.

a. $\int_{x=a}^{x=0} f(x)dx$
b. $\int_{x=a}^{x=b} f(x)dx$
c. $\int_{x=0}^{x=c} f(x)dx$
d. $\int_{x=a}^{x=c} f(x)dx$
e. $\int_{x=a}^{x=c} |f(x)|dx$
f. $\int_{x=a}^{x=c} (f(x)+1)dx$

Exercise 5.3.8

Say we are given that $\int_{x=0}^{x=5} f(x)dx = 10$, $\int_{x=5}^{x=9} f(x)dx = -2$ and $\int_{x=0}^{x=9} g(x)dx = -6$. Evaluate each of the following.

a. $\int_{x=0}^{x=9} (f(x)-g(x))dx$
b. $\int_{x=0}^{x=5} \left(\frac{f(x)}{3}\right) dx$
c. $\int_{x=0}^{x=9} (3g(x)+4)dx$
d. $\int_{x=0}^{x=9} (2g(x)+f(x))dx$

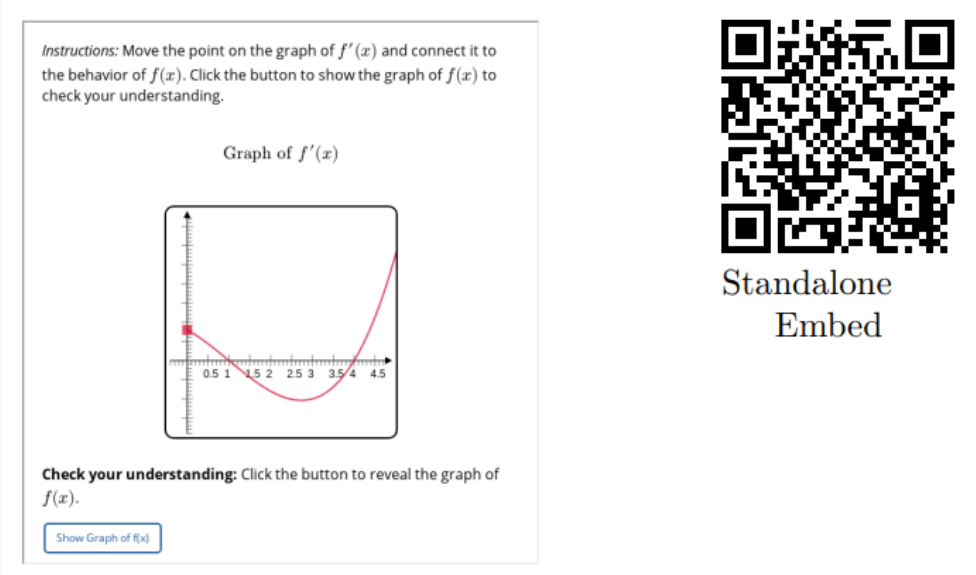
This page titled 5.3: The Definite Integral is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Peter Keep.

5.4: The Fundamental Theorem of Calculus

Let's remind ourselves of how we interpret derivatives. We are going to repeat a task that we did in Activity 4.2.3 First Derivative Test Graphically. It should feel familiar, which is good! We're going to use the intuition to make the big connection we've been forecasting so far.

Activity 5.4.1: Interpreting the Graph of a Derivative.

Let's look at a picture of a graph of the first derivative, $f'(x)$, and try to get some information about $f(x)$ from it. Use the following graph of $f'(x)$, the first derivative, to answer the questions about $f(x)$.



Since we don't have a huge amount of detail, you'll likely have to estimate the x-values for intervals and points in the following questions, but that's ok! Estimate away! Just make sure you know what you're looking for in the graph of $f'(x)$ to answer these questions.

a. List the intervals on which $f(x)$ is increasing. What about decreasing?
b. Find the x-values of any local maximums and/or local minimums of $f(x)$.
c. List the intervals on which $f(x)$ is concave up. What about concave down?
d. Find the x-values of any inflection points of $f(x)$.

Areas and Antiderivatives

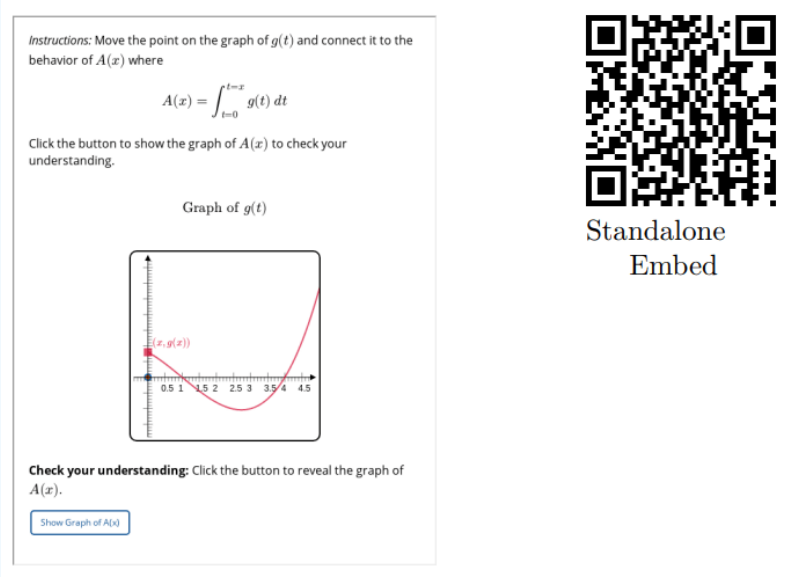
Activity 5.4.2: Interpreting Area.

First, we're going to define a bit of a weird function. Sometimes it's called the **Area function**:

$$A(x) = \int_{t=0}^{t=x} g(t)dt$$

This is a strange function, because we're defining the function as an integral of another function. Specifically, note that the input for our area function $A(x)$ is the ending limit of integration: we're calculating the signed area "under" the curve of $g(t)$ from $t = 0$ up to some variable ending point $t = x$.

We can visualize this function by looking at the areas we create as we change x. For now, get used to just seeing the area "under" g when we move the point around. The areas themselves are the outputs of the function $A(x)$.



Now we can think about this area function, and try to connect it to the graph of $g(t)$.

a. List the intervals on which $A(x)$ is increasing. What about decreasing?
b. Find the x-values of any local maximums and/or local minimums of $A(x)$.
c. List the intervals on which $A(x)$ is concave up. What about concave down?
d. Find the x-values of any inflection points of $A(x)$.
e. Compare your answers here to your answers about the behavior of $f(x)$ based on the (same) graph of $f'(x)$ in Activity 5.4.1. What does this mean about the connection between areas and derivatives, or areas and antiderivatives?

There it is! The way that we can interpret antiderivatives of functions! We found that the derivative of the function that tells us the signed area trapped between a curve and the x-axis between a fixed starting point and a variable ending point is the curve itself.

Another way of saying this, though, is that the function that tells us the signed area trapped between a curve and the x-axis between a fixed starting point and a variable ending point is an antiderivative of the curve itself! This is the Fundamental Theorem of Calculus, or at least half of it.

Theorem 5.4.1: Fundamental Theorem of Calculus (Part 1).

For a function f that is continuous on an interval $[a, b]$, and a function $A(x) = \int_{t=a}^{t=x} f(t)dt$ defined for x-values in $[a, b]$, then $A'(x) = f(x)$. That is:

$$\frac{d}{dx}\left(\int_{t=a}^{t=x} f(t)dt\right) = f(x)$$

Proof. available at the end of this section.

This theorem is going to be the big result that we use to show how to actually evaluate an area, and so it is easy to think of it as purely support for a "more important" result coming next. But we should pause and think about what this result tells us. What we've done here is come up with a way of:

1. Guaranteeing that every continuous function has an antiderivative family. We have found a function whose derivative is whatever continuous function we want!
2. Generating antiderivatives. Until now, we have had to rely on being able to recognize functions as derivatives of other things, or be able to "undo" derivative rules. And this will continue to be an important way for us to antidifferentiate functions. But

now we have a way of constructing antiderivatives, albeit weird looking ones. We are not yet used to thinking about a function that is defined as a definite integral with a variable ending point.

We will play with this idea more later (in Section 6.1), and so for now we will push forward towards our goal of evaluating a definite integral without directly calculating a limit of Riemann sums.

Evaluating Definite Integrals

Activity 5.4.3: Evaluating Areas and Antiderivatives.

In this short activity, we'll just collect information about antiderivatives and this new area function,

$$A(x) = \int_{t=a}^{t=x} f(t)dt$$

for a function $f(t)$ that is continuous on the interval $a \leq t \leq x$. For our purposes in this activity, let's say that $f(x) = x + \cos(x)$.

a. From the Fundamental Theorem of Calculus (Part 1), we know that $A(x)$ is an antiderivative of $f(x)$, since $A'(x) = f(x)$. Write out the function $A(x)$, and then name/write out one other antiderivative of $f(x)$, some $F(x)$. **Hint**. $A(x) = \int_{t=a}^{t=x} t + \cos(t)dt$ and for $F(x)$ you can use your antidifferentiation rules from Section 5.1 Antiderivatives and Indefinite Integrals.

b. We know that all of the antiderivatives of a function are connected to each other. Describe the connection between $A(x)$ and your $F(x)$. **Hint**. This is the result that we proved in Theorem 4.1.7 and used to define a **family of antiderivatives** in Definition 5.1.1.

c. What is the value of $A(a)$? What is the value of $F(a)$? How are they different from each other? **Hint**.

$$A(a) = \int_{t=a}^{t=a} f(t)dt = 0$$

For $F(a)$ you can evaluate your antiderivative at $x = a$. The important part is thinking about how these two values are different from each other.

d. What is the value of $A(b)$? What is the value of $F(b)$? How are they different from each other? **Hint**.

$$A(b) = \int_{t=a}^{t=b} f(t)dt$$

For $F(b)$ you can evaluate your antiderivative at $x = b$. The important part is thinking about how these two values are different from each other. Is the difference between these values the same, or different from the difference between $A(a)$ and $F(a)$?

e. What about the differences: $A(b) - A(a)$ compared to $F(b) - F(a)$? **Hint**. It is worth noting that:

$$\begin{aligned} A(b) - A(a) &= A(b) - 0 \\ &= \int_{t=a}^{t=b} f(t)dt \end{aligned}$$

Theorem 5.4.4: Fundamental Theorem of Calculus (Part 2).

For a function $f(x)$ continuous on the closed interval $[a, b]$ and some $F(x)$, an antiderivative of $f(x)$, then

$$\int_{x=a}^{x=b} f(x)dx = F(x)|_{x=a}^{x=b} = F(b) - F(a)$$

The vertical bar means "evaluated," and $F(x)|_{x=a}^{x=b}$ is typically read as "$F(x)$ evaluated from $x = a$ to $x = b$."

Phew, this was a lot! Let's sit back a bit and enjoy the fruits of all of this deep, mathematical thinking: we have a relatively straightforward way of evaluating definite integrals!

1. Find an antiderivative of the integrand. (Any antiderivative will do, so we can just choose the one with 0 as the constant term!)
2. Evaluate that antiderivative at the end points of the interval we're integrating over, and subtract.

Example 5.4.5

Evaluate the following definite integrals. Interpret the answers.

a. $\int_{x=0}^{x=2} \left(x^2+1\right) dx$
b. $\int_{x=0}^{x=2\pi} (\sin(x) - \cos(x)) dx$
c. $\int_{x=1}^{x=4} \left(\sqrt{x} - e^x\right) dx$

Solution

a.
$$\begin{aligned}\int_{x=0}^{x=2} \left(x^2+1\right) dx &= \underbrace{\left(\frac{x^3}{3}+x\right)}_{F(x)}\Bigg|_{x=0}^{x=2} \\ &= \underbrace{\left(\frac{2^3}{3}+2\right)}_{F(2)} - \underbrace{\left(\frac{0^3}{3}+0\right)}_{F(0)} \\ &= \left(\frac{8}{3}+2\right) - (0) \\ &= \frac{14}{3}\end{aligned}$$

This is the area we were approximating in Section 5.2!

b.
$$\begin{aligned}\int_{x=0}^{x=2\pi} (\sin(x) - \cos(x)) dx &= (-\cos(x) - \sin(x))|_{x=0}^{x=2\pi} \\ &= (-\cos(2\pi) - \sin(2\pi)) - (-\cos(0) - \sin(0)) \\ &= (-1-0) - (-1-0) \\ &= 0\end{aligned}$$

Why is this area 0? What does that mean about the region trapped between $y = \sin(x) - \cos(x)$ and the x-axis between $x = 0$ and $x = 2\pi$?

c.
$$\begin{aligned}\int_{x=1}^{x=4} \left(\sqrt{x} - e^x\right) dx &= \int_{x=1}^{x=4} \left(x^{1/2} - e^x\right) dx \\ &= \left(\frac{x^{3/2}}{3/2} - e^x\right)\Bigg|_{x=1}^{x=4} \\ &= \left(\frac{2(4)^{3/2}}{3} - e^4\right) - \left(\frac{2(1)^{3/2}}{3} - e^1\right) \\ &= \frac{14}{3} - e^4 + e\end{aligned}$$

This value is $\frac{14}{3} - e^4 + e \approx -47.21$. Why is this value negative? What does that mean about the region we're looking at, and the function we're looking at?

Practice Problems

Exercise 5.4.1

The Fundamental Theorem of Calculus comes in two parts. Explain why the Fundamental Theorem of Calculus (Part 1) guarantees that any function that is continuous on an interval has an antiderivative on that interval.

Exercise 5.4.2

The Fundamental Theorem of Calculus comes in two parts. Explain what the Fundamental Theorem of Calculus (Part 2) says about evaluating definite integrals.

Exercise 5.4.3

Consider the function $f(x) = x$.

a. Sketch the graph of $f(x)$.
b. Use geometry to find the area under the curve $f(x) = x$ from $x = 0$ to $x = 1$.
c. Use geometry to find the area under the curve $f(x) = x$ from $x = 0$ to $x = 4$.
d. Use geometry to find the area under the curve $f(x) = x$ from $x = 0$ to some unknown x-value with $x > 0$.
e. Write your area formula from above as a function $A(x)$.
f. Find $A'(x)$. How does this connect with the function $f(x) = x$?

Exercise 5.4.4

Consider the function $f(x) = 4x + 1$.

a. Sketch the graph of $f(x)$.
b. Use geometry to find the area under the curve $f(x) = x$ from $x = 2$ to $x = 3$.
c. Use geometry to find the area under the curve $f(x) = x$ from $x = 2$ to $x = 5$.
d. Use geometry to find the area under the curve $f(x) = x$ from $x = 2$ to some unknown x-value with $x > 2$.
e. Write your area formula from above as a function $A(x)$.
f. Find $A'(x)$. How does this connect with the function $f(x) = 4x + 1$?

Exercise 5.4.5

Consider the function $f(x) = 3x^2 + 6$ on the interval $x = -1$ to $x = 3$.

a. Find an antiderivative of $f(x)$. Call it $F(x)$.
b. Evaluate $F(3)$ and $F(-1)$.
c. Find a different antiderivative of $f(x)$. Call it $\tilde{F}(x)$.
d. Evaluate $\tilde{F}(3)$ and $\tilde{F}(-1)$.
e. Compare your answers in (b) and (d). What of these is similar? What of these are different?
f. Evaluate the definite integral $\int_{x=-1}^{x=3} (3x^2 + 6)\, dx$ using the Fundamental Theorem of Calculus.
g. When evaluating a definite integral using the Fundamental Theorem of Calculus, why does it not matter which antiderivative of $f(x)$ you use?

Exercise 5.4.6

Consider the graph of the first derivative $f'(x)$ below.

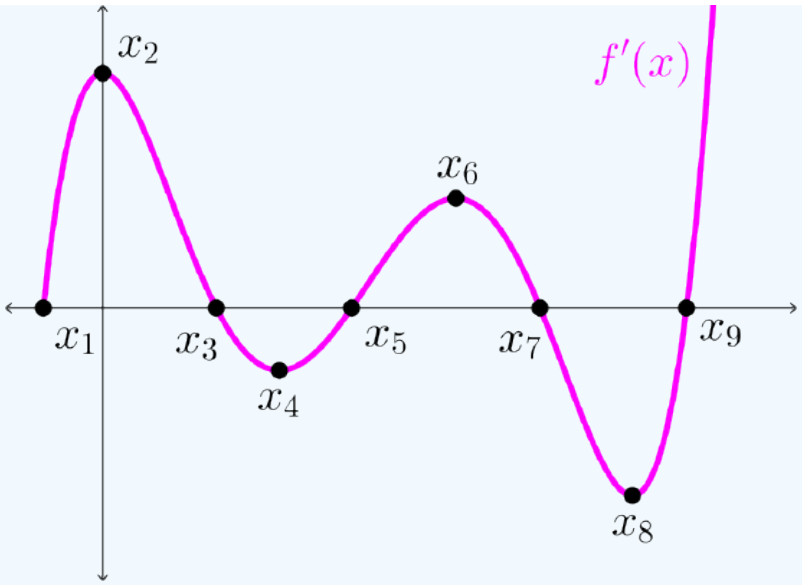


Figure 5.4.6 Graph of a first derivative, $f'(x)$.

a. Where is the function $f(x)$ increasing? Where is it decreasing?
b. Where are the local maximums and/or local minimums of $f(x)$?
c. Where is the function $f(x)$ concave up? Where is it concave down?
d. Where are the inflection points of $f(x)$?
e. Sketch a graph of $f(x)$.

Exercise 5.4.7

Consider the graph of the function $g(t)$ below with the Area function $A(x) = \int_{t=x_1}^{t=x} g(t)dt$ shaded in.

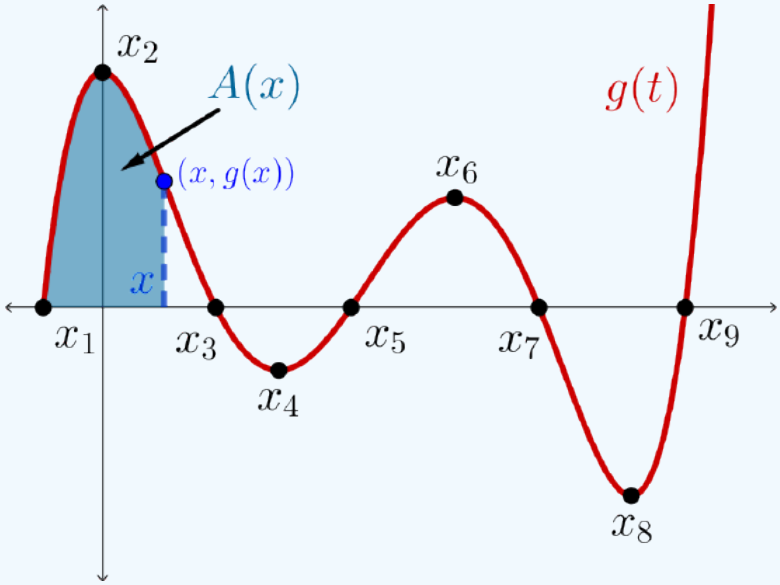


Figure 5.4.7 Graph of $g(t)$ and the area function, $A(x)$.

a. Where is the function $A(x)$ increasing? Where is it decreasing?
b. Where are the local maximums and/or local minimums of $A(x)$?

c. Where is the function $A(x)$ concave up? Where is it concave down?
d. Where are the inflection points of $A(x)$?
e. Sketch a graph of $A(x)$.
f. Compare your answers to the answers above. What can you conclude about the relationship between $A(x)$ and $g(x)$?

Exercise 5.4.8

For each of the definite integrals below, evaluate using the Fundamental Theorem of Calculus.

a. $\int_{x=1}^{x=5} (3x^2 + 6)\, dx$
b. $\int_{x=1}^{x=4} \sqrt{x} dx$
c. $\int_{x=0}^{x=\pi/4} (\cos(x) + 4\sin(x)) dx$
d. $\int_{x=-1}^{x=1} \left(\frac{1}{1+x^2}\right) dx$
e. $\int_{x=0}^{x=\pi} (\sin(x)) dx$
f. $\int_{t=-2}^{t=1} (t^5 + 2t^4 - 2t^3 + t^2 - 4t - 2)\, dt$
g. $\int_{x=1}^{x=3} \left(\frac{1}{x^2} + \frac{1}{x}\right) dx$
h. $\int_{x=0}^{x=5} (e^x)\, dx$
i. $\int_{x=\pi/4}^{x=\pi/3} (-5\sec^2(x))\, dx$
j. $\int_{x=1}^{x=2} \left(\frac{xe^x - 1}{x}\right) dx$
k. $\int_{x=1}^{x=6} \left(\frac{3}{\sqrt{x}} + \frac{1}{\sqrt[3]{x^2}}\right) dx$
l. $\int_{x=-1}^{x=2} (x(x-1)(x-4)) dx$
m. $\int_{x=2}^{x=4} \left(\frac{4}{3x} - \frac{3x}{4}\right) dx$
n. $\int_{\theta=\pi/3}^{\theta=\pi/2} \left(\frac{\cos\theta}{\sin^2\theta}\right) d\theta$

Answers

Theorem 5.4.1: Fundamental Theorem of Calculus (Part 1).

The proof of this theorem is one of the most delightful proofs we'll see. This is a "connector" theorem: a theorem that brings together several big ideas or objects from one common area of math and links them together. Let's enjoy the proof together.

Let $f(t)$ be a function that is continuous on the interval $a \le t \le b$. Then, we'll define the area function as $A(x) = \int_{t=a}^{t=x} f(t)dt$ for $a \le t \le b$. We are interested in $A'(x)$. From Definition 2.1.2, we know:

$$A'(x) = \lim_{\Delta x \to 0} \left(\frac{A(x+\Delta x) - A(x)}{\Delta x}\right)$$

If we just focus on the numerator, $A(x+\Delta x) - A(x)$, we have:

$$\begin{aligned} A(x+\Delta x) - A(x) &= \left(\int_{t=a}^{t=x+\Delta x} f(t)dt\right) - \left(\int_{t=a}^{t=x} f(t)dt\right) \\ &= \int_{t=x}^{t=x+\Delta x} f(t)dt \end{aligned}$$

Let's approximate this integral with a Riemann sum with $n = 1$ rectangle.

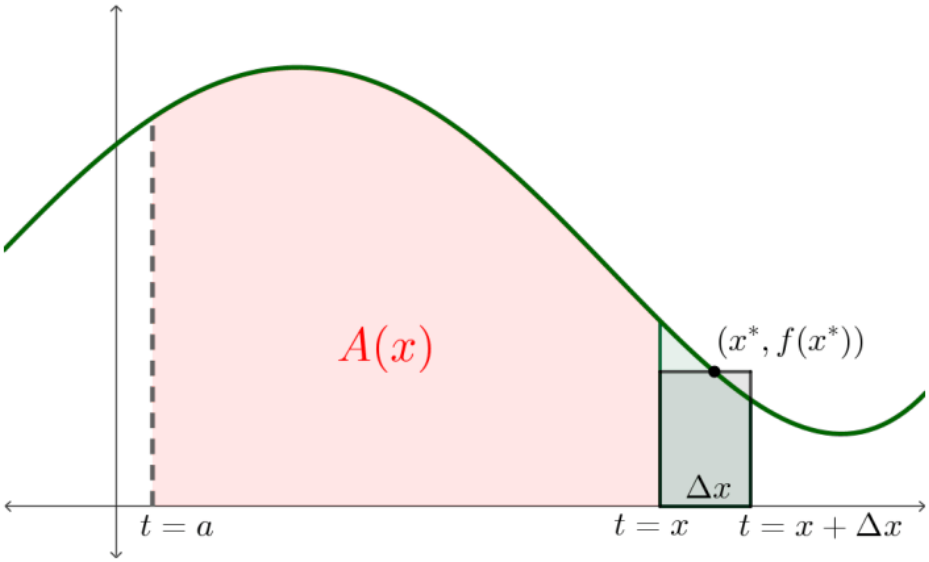


Figure 5.4.2

The total width of our interval is Δx, so we have that

$$\int_{t=x}^{t=x+\Delta x} f(t)dt \approx f(x^*)\,\Delta x$$

where x^* is some x-value in $[x, x + \Delta x]$. Note that we don't have a sum, as we normally would, since we are only "adding" a single area of a single rectangle. This is only an approximation of the difference $A(x + \Delta x) - A(x)$, and so we can say, for small values of Δx,

$$A'(x) \approx \left(\frac{A(x + \Delta x) - A(x)}{\Delta x}\right)$$

$$A'(x) \approx \left(\frac{f(x^*)\,\Delta x}{\Delta x}\right)$$

$$A'(x) \approx f(x^*)$$

All that is left to do is to convince ourselves of two facts:

1. This approximation gets better as Δx gets smaller, and as $\Delta x \to 0$ we have $A'(x) \to f(x^*)$.

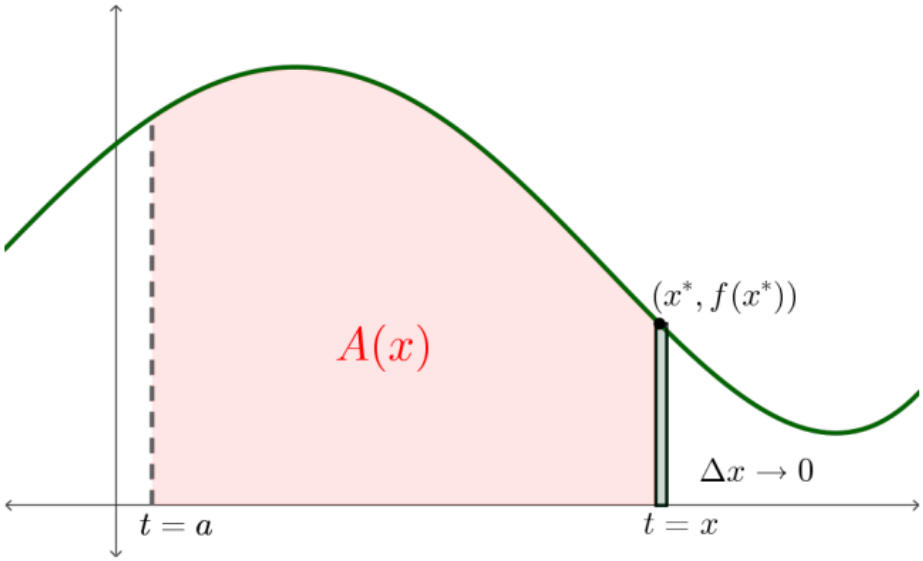


Figure 5.4.3

2. As $\Delta x \to 0$, the options for x^* in $[x, x + \Delta x]$ reduce to just x, since the interval collapses towards the single value. So as $\Delta x \to 0$, we have $x^* \to x$.

To be convinced that $A'(x) \to f(x^*)$, we just have to rely on the fact that, while our Riemann sum only has $n = 1$ rectangle, as $\Delta x \to 0$ the width(s) of "all" of our rectangles (our only one) approach 0, and so we end up with the definition of a definite integral in the limit:

$$\begin{aligned}\lim_{\Delta x \to 0} f(x^*)\,\Delta x &= \lim_{\Delta x \to 0} \sum_{k=1}^{1} f(x_k^*)\,\Delta x \\ &= \lim_{\Delta x \to 0} \int_{t=x}^{t=x+\Delta x} f(t)dt\end{aligned}$$

Hopefully it is easy to see that $x^* \to x$, since $[x, x + \Delta x]$ collapses on x.

Once we are convinced of these two facts, then it is clear that $A'(x) = f(x)$, since:

$$\begin{aligned} A'(x) &= \lim_{\Delta x \to 0} \left(\frac{A(x + \Delta x) - A(x)}{\Delta x} \right) \\ &\lim_{\Delta x \to 0} \left(\frac{f(x^*)\,\Delta x}{\Delta x} \right) \\ &= \lim_{\Delta x \to 0} f(x^*) \\ &= f(x) \end{aligned}$$

This completes the proof! Most of the proofs that you might see for this theorem use the Mean Value Theorem to help, since we can see a connection between the derivative $A'(x)$ and the average rate of change of the area function:

$$\left(\frac{A(x + \Delta x) - A(x)}{\Delta x} \right)$$

The Mean Value Theorem really is behind many of the most important results in calculus!

This page titled 5.4: The Fundamental Theorem of Calculus is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Peter Keep.

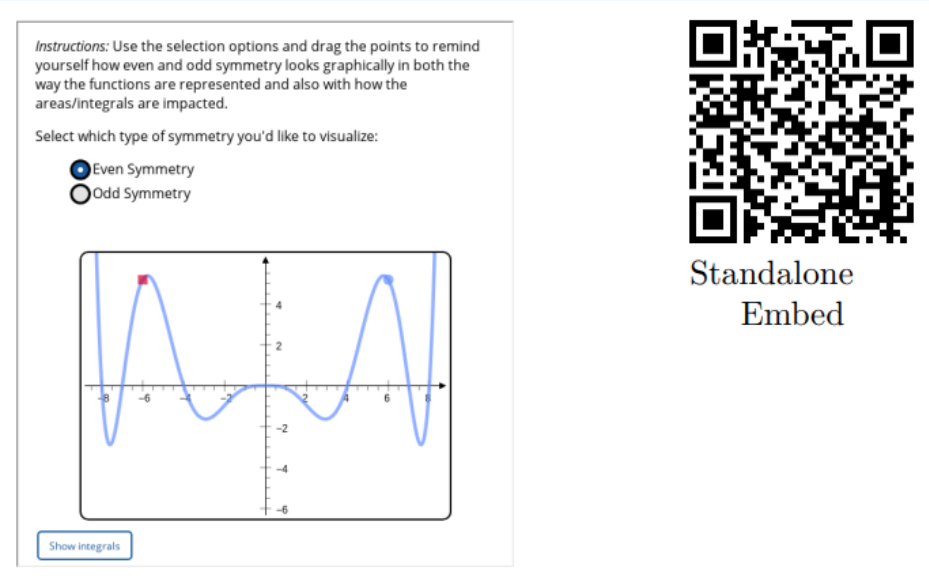
5.5: More Results about Definite Integrals

We'll end this chapter by looking a bit more closely at definite integrals and pulling a couple of small results out of our understanding of them, as well as some prior knowledge.

Symmetry

Activity 5.5.1: Symmetry in Functions and Integrals.

First, let's take a moment to remind ourselves (or see for the first time) what two types of "symmetry" we'll be considering. We call them "even" and "odd" symmetry, but sometimes we think of them as a "reflective" symmetry and a "rotational" symmetry in the graphs of our functions.



a. Convince yourself that you know what we mean when we say that a function is even symmetric on an interval if $f(-x) = f(x)$ on the interval. Similarly, convince yourself that you know what we mean when we say that a function is odd symmetric on an interval if $f(-x) = -f(x)$ on the interval. **Hint**. Look at the relationship between the points on the graphs when you select the different symmetries: How do their x-values relate to each other? How do their y-values relate to each other?

b. Now let's think about areas. Before we visualize too much, let's start with a small question: How does the height of a function impact the area defined by a definite integral? It should be helpful to think about Riemann sums and areas of rectangles here. The important question then, is how does a function being even or odd symmetric tell us information about areas defined by definite integrals of that function? **Hint**.

   If we know that for an even symmetric function, there are some heights/y-values that are the same, then we know that there are some areas/integrals that should also be the same. Which ones?

   If we know that for an odd symmetric function, there are some heights/y-values that are opposite, then we know that there are some areas/integrals that should also be opposite. Which ones?

Theorem 5.5.1: Definite Integrals of Symmetric Functions.

If $f(x)$ is a continuous function on $[-a, a]$ for some real number $a > 0$, then:

- If $f(x)$ is even symmetric on $[-a, a]$, then:

$$\int_{x=-a}^{x=0} f(x)dx = \int_{x=0}^{x=a} f(x)dx$$

- If $f(x)$ is odd symmetric on $[-a, a]$, then:

$$\int_{x=-a}^{x=0} f(x)dx = -\int_{x=0}^{x=a} f(x)dx$$

Activity 5.5.2: Connecting Symmetric Integrals.

We're going to do some sketching here, and I want you to be clear about something: your sketches can be absolutely terrible. It's ok! They just need to embody the kind of symmetry we're talking about. You will probably sketch something and notice that your areas aren't to scale (or maybe even the wrong sign!), and that's fine.

It might be helpful to practice sketching graphs accurately, but don't worry if that part is a struggle.

a. Sketch a function $f(x)$ with the following properties:
   - $f(x)$ is even symmetric on the interval [–6, 6]
   - $\int_{x=0}^{x=6} f(x)dx = 4$
   - $\int_{x=-6}^{x=-2} f(x)dx = -1$

b. Find the values of the following integrals:
   - $\int_{x=0}^{x=2} f(x)dx$
   - $\int_{x=-6}^{x=6} f(x)dx$
   - **Hint**. Since $f(x)$ is even symmetric, what are the two other integrals that we know about? How can we use those to help us find these two?

c. Sketch a function g(x) with the following properties:
   - $g(x)$ is odd symmetric on the interval [–9, 9]
   - $\int_{x=0}^{x=4} g(x)dx = 5$
   - $\int_{x=-9}^{x=0} g(x)dx = 2$

d. Find the values of the following integrals:
   - $\int_{x=-9}^{x=-4} g(x)dx$
   - $\int_{x=-4}^{x=9} g(x)dx$
   - **Hint**. Since $g(x)$ is even symmetric, what are the two other integrals that we know about? How can we use those to help us find these two?

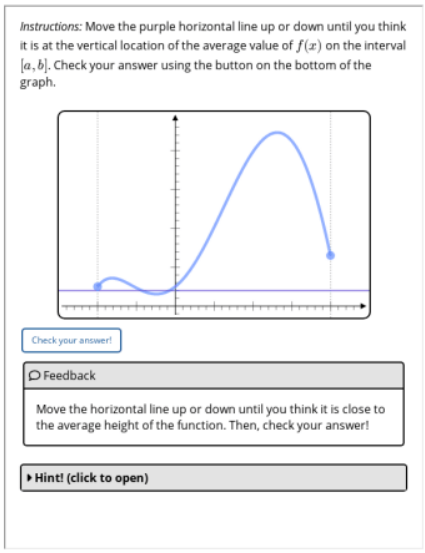
Answer available at the end of this section.

Average Value of a Function

Activity 5.5.3: Visualizing the Average Height of a Function.

We are going to build a formula to find the"average height" or "average value" of a function $f(x)$ on the interval $[a, b]$. We're going to look at a function and try to find the average height. Along the way, we'll think a bit about areas!

a. Consider the following function. Find the average height of the function on the interval pictured!



Standalone Embed

b. How does the area "under" the curve $f(x)$ on the interval compare to the area of the rectangle formed by the average height line?

c. How do you define the two areas? **Hint**.

One of these is the area under $f(x)$ from $x = a$ to $x = b$, which we can use calculus for!

The other is the area of a rectangle with a height (the average height of $f(x)$) and a width (the width of the interval).

d. Set up an equation connecting the two areas, and solve for the average height of $f(x)$. **Hint**. If $\text{area} = \text{height} \times \text{width}$, then doesn't it make sense that $\text{height} = \dfrac{\text{area}}{\text{width}}$? How, then, do we find average height by dividing an area and a width?

Theorem 5.5.2: Average Value of a Function.

If a function $f(x)$ is continuous on the interval $[a, b]$, then the average value of $f(x)$ on $[a, b]$ is:

$$\frac{\int_{x=a}^{x=b} f(x)dx}{b-a}$$

Example 5.5.3

A small model glider airplane is thrown and travels for 10 seconds before it hits the ground. The height of the glider is modeled by the function $h(t) = 6 + \frac{7t}{5} - \frac{t^2}{5}$ on the interval [0, 10]. Find the average height of the glider on the time interval.

Solution

$$\begin{aligned}\frac{1}{10}\int_{t=0}^{t=10} 6 + \frac{7t}{5} - \frac{t^2}{5}dt &= \frac{1}{10}\left(6t + \frac{7t^2}{10} - \frac{t^3}{15}\right)\Big|_{t=0}^{t=10} \\ &= \frac{1}{10}\left(\frac{190}{3}\right) \\ &= \frac{190}{30}\end{aligned}$$

The glider has an average height of 6 feet and 4 inches.

Practice Problems

Exercise 5.5.1

Explain how we can use the different types of symmetry can help us as we evaluate the definite integral $\int_{x=-a}^{x=a} f(x)dx$.

Exercise 5.5.2

Use the definition of even symmetric (where $f(-x) = f(x)$ for all x-values) to explain why $\int_{x=-4}^{x=4} f(x)dx = 2\int_{x=0}^{x=4} f(x)dx$ for an even symmetric function $f(x)$.

Exercise 5.5.3

Use the definition of odd symmetric (where $f(-x) = -f(x)$ for all x-values) to explain why $\int_{x=-4}^{x=4} f(x)dx = 0$ for an odd symmetric function $f(x)$.

Exercise 5.5.4

Assume that $f(x)$ is some even symmetric function with $\int_{x=0}^{x=6} f(x)dx = 4$ and $\int_{x=-3}^{x=0} f(x)dx = 9$.

a. Sketch a possible graph of the function $f(x)$, and shade in the areas from the integrals listed.
b. Evaluate $\int_{x=0}^{x=3} f(x)dx$.
c. Evaluate $\int_{x=3}^{x=6} f(x)dx$.
d. Evaluate $\int_{x=-6}^{x=-3} f(x)dx$.
e. Evaluate $\int_{x=-6}^{x=6} f(x)dx$.

Exercise 5.5.5

Assume that $g(x)$ is some odd symmetric function with $\int_{x=0}^{x=6} g(x)dx = 4$ and $\int_{x=-3}^{x=0} g(x)dx = 9$.

a. Sketch a possible graph of the function $g(x)$, and shade in the areas from the integrals listed.
b. Evaluate $\int_{x=0}^{x=3} g(x)dx$.
c. Evaluate $\int_{x=3}^{x=6} g(x)dx$.
d. Evaluate $\int_{x=-6}^{x=-3} g(x)dx$.
e. Evaluate $\int_{x=-6}^{x=6} g(x)dx$.

Exercise 5.5.6

Explain how to find the average height of a function on an interval. Why does this work?

Exercise 5.5.7

Note that $y = \sin(x)$ is odd symmetric and $y = x^2$ is even symmetric.

a. Show that $f(x) = x^2\sin(x)$ is odd symmetric by showing that $f(-x) = -f(x)$.
b. Evaluate $\int_{x=-\pi}^{x=\pi} x^2\sin(x)dx$.
c. Find the average height of $f(x)$ on the interval $[-\pi, \pi]$.

Exercise 5.5.8

Note that $y = \cos(x)$ is even symmetric and $y = x^2$ is even symmetric.

a. Show that $f(x) = x^2 \cos(x)$ is even symmetric by showing that $f(-x) = f(x)$.
b. Given that $\int_{x=0}^{x=\pi} x^2 \cos(x) dx = -2\pi$, evaluate $\int_{x=-\pi}^{x=\pi} x^2 \cos(x) dx$.
c. Find the average height of $f(x)$ on the interval $[-\pi, \pi]$.

Exercise 5.5.9

Find the average height of the following functions on the intervals listed.

a. $f(x) = e^x - x$ on [0, 3]
b. $g(x) = \cos(x) + x$ on $\left[0, \frac{\pi}{2}\right]$
c. $j(x) = \frac{1}{1+x^2}$ on [–1, 1]
d. $\ell(x) = \frac{1}{x^2}$ on [1, 4]

Answers

Activity 5.5.2: Connecting Symmetric Integrals.

b.
- $\int_{x=0}^{x=2} f(x) dx = 5$
- $\int_{x=-6}^{x=6} f(x) dx = 8$

d.
- $\int_{x=-9}^{x=-4} g(x) dx = 7$
- $\int_{x=-4}^{x=9} g(x) dx = -2$

This page titled 5.5: More Results about Definite Integrals is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Peter Keep.

5.6: Introduction to u Substitution

We have spent some time thinking about integrals, both indefinite and definite integrals, and what they represent. Let's end this discussion of integration with some classifications of indefinite integrals.

This classification is completely unserious, but maybe helpful. Let's say that we had to classify each of the following integrals as **Easy**, **Medium**, or **Hard**. Again, these are completely ambiguous and not at all defined, but I hope that we can see the spirit of what we're thinking. Here are the three integrals that we want to classify:

$$\int x^5 dx \qquad \int \sec^2(x) dx \qquad \int \frac{x^3}{1+x^2} dx$$

Which one could have the **Easy**? What about a **Medium** difficulty integral? What makes this one a bit harder, but not so hard to be classified as **Hard**? Can you even solve the **Hard** right now?

Let's think about a legitimately difficult integral, but one that we can actually think about.

Activity 5.6.1: A Hard Integral.

We're going to be thinking about two integrals here, but before we do, we should remind ourselves about how we can "rephrase" an integration question.

If we are asked to find $\int f(x)dx$, then we are really being asked to find some function $F(x)$ whose derivative is $f(x)$. Of course, we're actually being asked to find all of the possible functions that fit this requirement, but we know that the constant of integration covers all of the differences. This means that we can (and maybe should) check our answers pretty consistently: just find a derivative and check that it matches the integrand!

1. Find $\int 3x^2 e^{x^3} dx$.
2. Explain the role of $3x^2$ in the integral. Explain why the positioning of x^3 matters.
3. Let's try another one.
$$\int 7x^6 \cos\left(x^7\right) dx$$
4. How can you tell, in general, that the derivative you're looking at is one that was produced through the Chain Rule and not the Product Rule?

Undoing the Chain Rule

Let's try to formalize this process of "undoing the Chain Rule" that we noticed in Activity 5.6.1. It might be helpful to, first, think about the Chain Rule and how the differentiation process works. Let's look at the differentiating $\sin\left(x^7\right)$ using the Chain Rule.

The first thing we should do when finding $\frac{d}{dx}\left(\sin\left(x^7\right)\right)$ is recognize and identify the composition. We might call x^7 the "inside" function, or relabel it as u. Then we know that the Chain Rule will tell us to differentiate the "outside" function with regard to the "inside" function (or u if we relabel things) and then multiply by the derivative of the inside function $\left(\text{or} \frac{du}{dx}\right)$.

$$\begin{aligned}
\frac{d}{dx}(\sin(\underbrace{x^7}_{u})) &= \frac{d}{dx}(\sin(u)) \\
&= \frac{d}{du}(\sin(u))\left(\frac{du}{dx}\right) \\
&= \cos(u)\frac{d}{dx}\left(x^7\right) \\
&= \cos\left(x^7\right)7x^6
\end{aligned}$$

From here, we might rewrite things to make it look nicer (coefficients look weird when they're not in the front, in my opinion) and write:

$$\frac{d}{dx}\left(\sin\left(x^7\right)\right) = 7x^6\cos\left(x^7\right)$$

When we work through this process backwards, we'll need to identify the "inside" function, but also find the derivative of that "inside" function, $\frac{du}{dx}$. This derivative gets introduced in the Chain Rule, and so it will have to be picked out when we undo the Chain Rule.

Let's build this process and go through the integral $\int 7x^6\cos\left(x^7\right)dx$ from Activity 5.6.1.

Process for u-Substitution.

1. Identify an "inside" function and/or a "function-derivative pair." We'll label the "inside" function, or the "function" part of the "function-derivative pair," as u. **Example**: In our integral $\int 7x^6\cos\left(x^7\right)dx$ we can see the "inside" function is x^7, and also we have x^7 and $7x^6$ as the function-derivative pair. Let $u = x^7$.
2. Define the substitution for the differential, $du = u'dx$. We can think about this as a result from knowing that $\frac{du}{dx} = u'$, or by thinking that a small change in u corresponds with u' multiplied by a small change in dx. **Example**: In our integral $\int 7x^6\cos\left(x^7\right)dx$, we labeled $u = x^7$. This means that $du = 7x^6dx$, since this is the derivative of x^7.
3. Substitute! Rewrite the integral, replacing the parts that you've labeled as u and du. **Example**: We can rewrite our integral to make this a bit easier to see:
$$\begin{aligned}\int 7x^6\cos(x)dx &= \int \cos(\underbrace{x^7}_{u})\,\underbrace{7x^6dx}_{du}\\ &= \int \cos(u)du\end{aligned}$$
4. Antidifferentiate! We should have an integral that is "written in terms of u," and so we can antidifferentiate the function as if u was our input variable. Notice that what we've done with our substitution is to undo the "multiply by the derivative of the inside function" step of the Chain Rule. Now we can antidifferentiate the "outside" function! **Example**: $\int \cos(u)du = \sin(u) + C$
5. Substitute "back" to have our antiderivative family written in terms of our original input variable, x in this case. We'll replace u with whatever we had defined our substitution to be in step 1. **Example**: We defined $u = x^7$, so:
$$\sin(u) + C = \sin\left(x^7\right) + C$$

Let's practice this with some more examples.

Activity 5.6.2: Picking the Pieces of a Substitution.

We're going to look at three integrals. Instead of working through them one-at-a-time, we'll look at all three simultaneously, where we can practice identifying, substituting, and antidifferentiating all at the same time.

a. Let's consider these three integrals:
   - $\int \frac{3x^2+1}{\left(x^3+x-2\right)^2}dx$
   - $\int \cos(x)\sqrt{\sin(x)}dx$
   - $\int \frac{(\ln(x))^3}{x}dx$

   For each of these integrals, identify the substitution: define u as some function of x. **Hint**. Can you find the composition in each one? If you notice the function-derivative pairing instead, then you can think about making the "function" portion!

b. For each substitution, define $du = u'dx$.
c. For each integral, use your substitution (for both u and the differential du) to rewrite the integral.

d. Antidifferentiate each integral, and then use your substitution to write each integral back in terms of x.

Answer available at the end of this section.

Let's try to explain a little bit of what is happening. This style of problem solving is really useful in mathematics, and shows up in many places.

Aside.

The first time I saw Figure 5.6.1 was in a Differential Equations class. We were learning about Laplace transformations, a technique that is very useful for solving a variety of problems in the field of differential equations. My professor was explaining why and how Laplace transformations were so powerful, and drew a version of the figure I've included to explain u -substitution. It was so helpful for me to understand what was happening, but the most helpful thing was when the professor said, offhand, "But that's exactly the same type of thing that u-substitution does, too." So many things fell into place for me because of that comment! You are guaranteed to see different versions of this picture throughout this textbook, but you can also keep an eye out for this in different problem-solving techniques in mathematics.

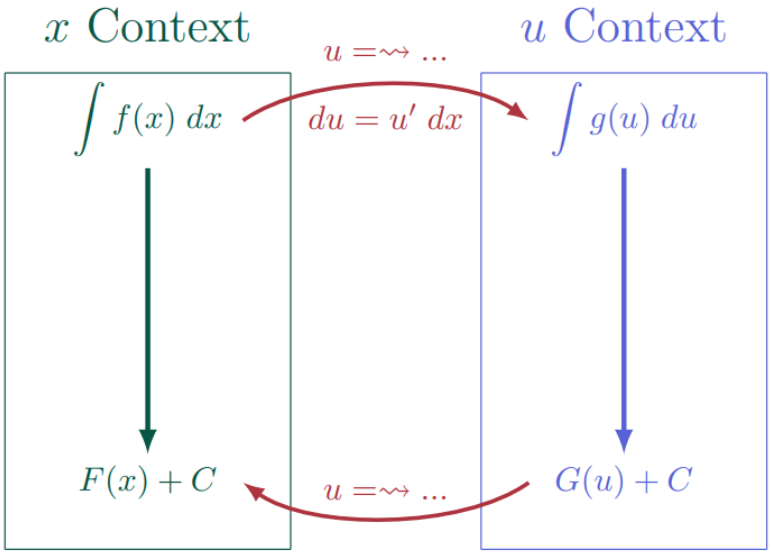


Figure 5.6.1 General idea of how a variable substitution in an integral works.

Let's explain what is happening in this picture. First, we typically are presented with integrals in some context. Our goal is to antidifferentiate. So for $\int f(x)dx$, we hope to antidifferentiate and end up with $F(x) + C$, the family of antiderivative of $f(x)$. These two things (the indefinite integral and the family of antiderivatives) exist in the same context (since they're defined with the same variable). We have spent some time moving directly from $f(x)dx$ to $F(x) + C$. But now we're seeing that this isn't always as direct of a path as we might wish: some integrals are hard.

In this case, we can try to identify what the problem is (in this case, composition) and find some transformation to apply to our integral. We choose a variable substitution, and we translate our integral to a different context (in this case, by writing it in terms of a different variable). In order for us to do this, we ned to define some substitution u, and then translate dx into du using the relationship $du = u'dx$.

Once we have this new integral, of a different function $g(u)$ in a different context with u acting as the input, we try to antidifferentiate again. If we pick our substitution carefully and we know what kind of problem we're trying to fix (in this case, getting around the composition), then this new integral in the u context could be "easier" to antidifferentiate. So our goal is to antidifferentiate the integral, but we can antidifferentiate it after translating it to a different context.

Once we have this antiderivative, we can translate that antiderivative family back to the original context (in this case, we write it

in terms of x). We do this by utilizing the same translation or substitution that we defined earlier: we have something defined to link u and x. We can notice that, since the object we're translating is not an integral anymore, we do not have a differential to translate.

And there we have it: u-substitution works by identifying a problem that makes our task hard, translating our object to a friendlier context based on what we know about our problem, solving the problem in this friendlier context, and then finally translating the solution back to our original context.

Aside.

Doesn't this feel like what we do with Logarithmic Differentiation? Use logs to translate our function to a friendlier form, differentiate in the log-context, and then try to translate back by solving for y'?

Let's practice this new strategy!

Activity 5.6.3: Compare Two Integrals.

Let's compare two integrals, and use this to build a more general strategy for performing u-substitution.

a. Consider the following integral:

$$\int -4x^3 \sec^2\left(1-x^4\right) dx$$

Select and justify a choice for u. **Hint**. Where is the composition? Do you see a function-derivative pair?

b. Perform the u-substitution and antidifferentiate, and then substitute back to write your antiderivative in terms of x.

c. Compare that integral to this one:

$$\int x^3 \sec^2\left(1-x^4\right) dx$$

What is different about this new integral? What has remained the same? How does that impact your choice for u, or does it impact your choice for u? **Hint**. Notice that nothing has changed about the "inside" function!

d. Has that changed what du should be? **Hint**. Remember that $du = u'dx$. If we didn't change our selection of u, then shouldn't du remain the same as well?

e. Ok, so we've noticed an issue here. There are plenty of good ways of solving this problem, where du doesn't "show up" perfectly in our integral. In this case, we have that we're missing a necessary coefficient. We have the x^3 part, but we are missing the –4.Try to rewrite our integral with a –4 coefficient in there. We'll do that by multiplying the integrand function by 1, disguised as $\frac{-4}{-4}$ or $\left(-\frac{1}{4}\right)(-4)$.

f. Now we can use the same u-substitution as before, and integrate in a similar way! Notice, though, that we will retain the coefficient of $-\frac{1}{4}$. (This should be reasonable: our integral is $-\frac{1}{4}$ of the original one, since our coefficient was 1 to the original's –4.) Go ahead and integrate!

Answer available at the end of this section.

This gives us a general strategy that we can use: if we pick a u-substitution, but we cannot find in our integral, we can try to manipulate the integrand function to find it! We normally do this by applying some operation and its inverse (like multiplying by and dividing by it as well, as in Activity 5.6.3).

Fixing Coefficients for du.

If we choose to perform a u-substitution in the integral $\int f(x)dx$, but we require a coefficient k in our definition of du, we can "fix" the coefficient in our integral:

$$\frac{1}{k}\int kf(x)dx$$

This strategy works well for coefficients, since we can factor out the $\frac{1}{k}$ from the integral.

Substitution for Definite Integrals

How would we evaluate the following definite integral?

$$\int_{x=0}^{x=2}\left(\frac{x+1}{x^2+2x+1}\right)dx$$

We can think back to what the Fundamental Theorem of Calculus (Part 2) says about evaluating a definite integral. We need to do two things:

1. Find an antiderivative of our function, $F(x)$. Any antiderivative will do, and we often pick the one where the constant term is 0.
2. Evaluate the antiderivative at the end points of the interval and subtract: $F(b)-F(a)$.

So for us to evaluate this definite integral, we can split the work into two parts.

Antidifferentiate, then Evaluate

Part 1: Antidifferentiation.

We can think about the function $f(x)=\frac{x+1}{x^2+2x+1}$ and find the family of antiderivatives. Then, we can disregard the constant term (by selecting the antiderivative where the constant is 0 for convenience).

So we'll use u-substitution on the integral $\int\left(\frac{x+1}{x^2+2x+1}\right)dx$

We can use $u=x^2+2x+1$, which gives us $du=(2x+2)dx$ or $du=2(x+1)dx$.

$$\begin{aligned}\int\left(\frac{x+1}{x^2+2x+1}\right)dx &= \frac{1}{2}\int\left(\frac{2(x+1)}{x^2+2x+1}\right)dx\\ &= \frac{1}{2}\int\frac{1}{u}du\\ &= \frac{1}{2}\ln|u|+C\\ &= \frac{1}{2}\ln\left|x^2+2x+1\right|+C\end{aligned}$$

So let's choose $F(x)=\frac{1}{2}\ln\left|x^2+2x+1\right|$ as the antiderivative we'll use.

Part 2: Evaluate at the End Points.

For the integral $\int_{x=0}^{x=2}\left(\frac{x+1}{x^2+2x+1}\right)dx$, our ending points are $x=2$ and $x=0$ so let's evaluate!

$$\begin{aligned} F(2) &= \frac{1}{2}\ln\left|(2)^2+2(2)+1\right| \\ &= \frac{1}{2}\ln(9) \\ F(0) &= \frac{1}{2}\ln\left|(0)^2+2(0)+1\right| \\ &= \frac{1}{2}\ln(1) = 0 \\ F(2)-F(1) &= \frac{1}{2}\ln(9) - 0 \\ &= \frac{1}{2}\ln(9) \\ &= \ln(3) \end{aligned}$$

Great, so we have a way of evaluating this integral!

$$\int_{x=0}^{x=2}\left(\frac{x+1}{x^2+2x+1}\right)dx = \ln(3)$$

A More Wholistic Substitution

Hang on, wait.

When we substituted our integral, we were substituting the indefinite integral. What if we applied our substitution to the definite integral.

The only difference is the limits of integration (other than the interpretation of area vs. family of antiderivatives, of course). So let's substitute the limits of integration.

Consider the definite integral. Really think about it.

$$\int_{x=0}^{x=2}\left(\frac{x+1}{x^2+2x+1}\right)dx$$

Here, we label the limits of integration as x-values: $x=0$ and $x=2$.

Can't we use our substitution rule to find corresponding u-values? What happens then? Let's approach this definite integral using the same substitution: we will think of $u=x^2+2x+1$ again. But now we can find corresponding values of u when $x=0$ and $x=2$. All we need to do is evaluate the formula for u at those x-values!

$$\begin{aligned} \int_{x=0}^{x=2}\left(\frac{x+1}{x+2x+1}\right) &= \frac{1}{2}\int_{x=0}^{x=2}\left(\frac{2(x+1)}{x^2+2x+1}\right)dx \\ &= \frac{1}{2}\int_{u=(0)^2+2(0)+1}^{u=(2)^2+2(2)+1}\frac{1}{u}du \\ &= \frac{1}{2}\int_{u=1}^{u=9}\frac{1}{u}du \\ &\left(\frac{1}{2}\ln|u|\right)\Big|_{u=1}^{u=9} \\ &= \frac{1}{2}\ln(9) - \frac{1}{2}\ln(1) \\ &= \frac{1}{2}\ln(9) \\ &= \ln(3) \end{aligned}$$

So notice that we end up with the same thing here...we can substitute the limits of integration, and this matches the same value that we would get when we evaluate our antiderivatives at the endpoints of the x-interval.

We can amend our picture from Figure 5.6.1 to include definite integrals: in this case, we can evaluate the definite integral in either context, as long as we translate the limits of integration as well.

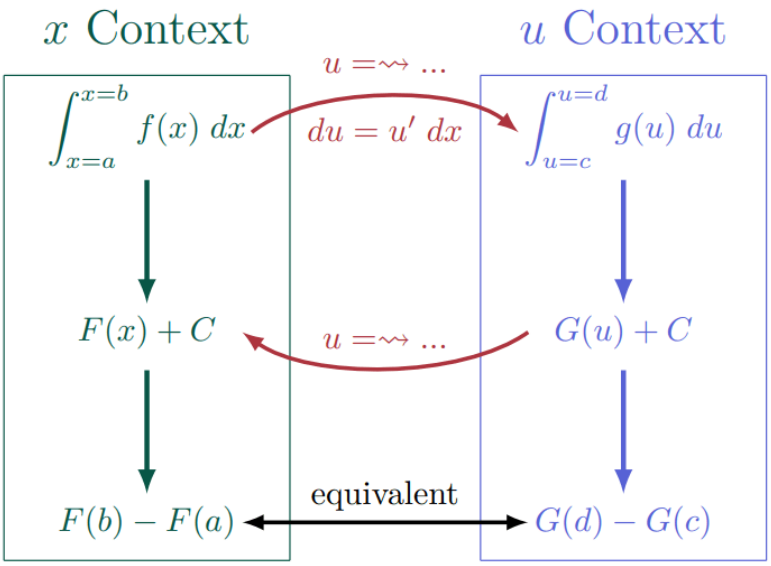
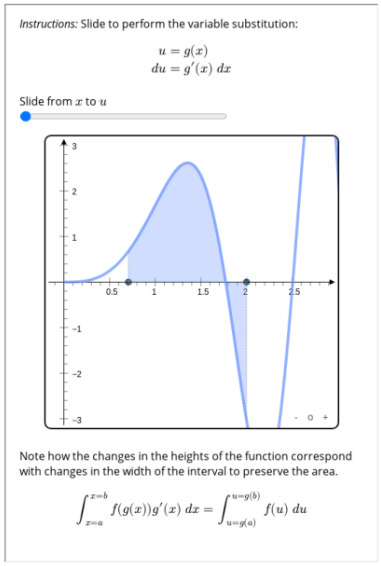


Figure 5.6.2 General idea of how a variable substitution in a definite integral works.

To see a visualization of what is happening, we can look below: move the slider to see the continuous deformation of the integral as we apply the variable substitution:



Standalone
Embed

Something we can say is that, since the area doesn't change when we do a variable substitution, then area is invariant under the transformation we're applying with that variable substitution.

More to Translate

There are more little tricks and nuances that we can think about with u-substitution: in general, this is an extremely flexible integration technique that we'll use in a variety of ways. For now, let's leave things off with one more interesting example.

In this example, we'll see a similar kind of issue to the one we saw in Activity 5.6.3: when we pick our substitution, there will be some issues "finding" du.

Example 5.6.3

Integrate the following, making sure to translate the whole integrand function to be written in terms of u.

$$\int \left(\frac{x^3}{\sqrt{x^2+1}} \right) dx$$

Hint 1. First, notice that $u = x^2 + 1$ is a great choice: we really want to focus on that composition. If this is the case, though, then $du = 2xdx$.

Hint 2. We can write x^3 as $x^2 \cdot x$, or if you really want to, we can write it as $\frac{1}{2}x^2 \cdot (2x)$

Hint 3. Our u-substitution formula can be written in a whole bunch of different ways!

$$\begin{aligned} u &= x^2 + 1 \\ x^2 &= u - 1 \\ x &= \pm\sqrt{u-1} \end{aligned}$$

These are all equivalent, but the first two might be the most helpful:

- Anywhere in our integral that we can see an $x^2 + 1$, we can replace that with u.
- We can also replace any extra x^2 pieces with $u - 1$!

Solution

$$\begin{aligned} \int \left(\frac{x^3}{\sqrt{x^2+1}} \right) dx u &= x^2 + 1 \\ du &= 2xdx \\ \int \left(\frac{x^3}{\sqrt{x^2+1}} \right) dx &= \frac{1}{2} \int \left(\frac{x^2 \cdot (2x)}{\sqrt{x^2+1}} \right) dx \\ &= \frac{1}{2} \int \frac{(u-1)}{\sqrt{u}} du \qquad u = x^2 + 1 \leftrightarrow x^2 = u - 1 \\ &= \frac{1}{2} \int \frac{u}{\sqrt{u}} - \frac{1}{\sqrt{u}} du \\ &= \frac{1}{2} \int u^{1/2} - u^{-1/2} du \\ &= \frac{1}{2} \left(\frac{2u^{3/2}}{3} - 2u^{1/2} \right) + C \\ &= \frac{(x^2+1)^{3/2}}{3} - \sqrt{x^2+1} + C \end{aligned}$$

We'll spend more time thinking about u-substitution (as well as other variable substitutions) later on in Chapter 7. For now, this is a good stopping point, and should give us enough of a handle on u-substitution to integrate some difficult integrals!

Practice Problems

Exercise 5.6.1

The u-substitution integration method was designed to undo a specific derivative rule. Explain.

Exercise 5.6.2

The u-substitution integration method deals with composition of functions. If you're looking at a composite function $f(g(x))$, how do you identify the "inside" function, $g(x)$?

Exercise 5.6.3

For each composite function $f(g(x))$ below, identify two different possibilities for both the "inside" $g(x)$ function and the "outside" $f(x)$ function.

a. $f(g(x)) = \sqrt{x^2+1}$
b. $f(g(x)) = \sin\left(e^x+1\right)$
c. $f(g(x)) = e^{\cos(x)}$
d. $f(g(x)) = (\tan(x)+x)^6$

Exercise 5.6.4

Use the chain rule to differentiate each of the following. Carefully note for each problem which part of the composite function acts as the "inside"piece.

a. $\frac{d}{dx}\left(e^{x^2}\right)$
b. $\frac{d}{dx}(\ln(\sin(x)))$
c. $\frac{d}{dx}(-\cos(\sqrt{x}))$
d. $\frac{d}{dx}\left((5x+1)^{10}\right)$
e. $\frac{d}{dx}\left(\tan\left(e^x+\pi\right)\right)$
f. $\frac{d}{dx}\left(\sqrt{x^3+1}\right)$

Exercise 5.6.5

Integrate the following indefinite integrals using the suggested u-substitution.

a. $\int 2xe^{x^2}dx; u = x^2$
b. $\int\left(\frac{\cos(x)}{\sin(x)}\right)dx; u = \sin(x)$
c. $\int\left(\frac{\sin(\sqrt{x})}{2\sqrt{x}}\right)dx; u = \sqrt{x}$
d. $\int 50(5x+1)^9 dx; u = 5x+1$
e. $\int e^x \sec^2\left(e^x+\pi\right)dx; u = e^x+\pi$
f. $\int\left(\frac{3x^2}{2\sqrt{x^3+1}}\right)dx; u = x^3+1$

Exercise 5.6.6

Use u-substitution to integrate the following indefinite integrals.

a. $\int\left(\frac{2x+1}{\left(x^2+x+1\right)^3}\right)dx$

b. $\int \left(\dfrac{x}{x^2+1}\right) dx$
c. $\int xe^{-x^2} dx$
d. $\int \tan(x) dx$
e. $\int \dfrac{1}{x\ln(x)} dx$
f. $\int \sin^6(x)\cos(x) dx$
g. $\int x\sec\left(x^2\right)\tan\left(x^2\right) dx$
h. $\int \left(\dfrac{x}{1+x^4}\right) dx$

Exercise 5.6.7

For each of the following definite integrals, identify a u-substitution and write the definite integral with regard to u, including new (translated) limits of integration.

a. $\int_{x=0}^{x=3} \left(\dfrac{1}{x+1}\right) dx$
b. $\int_{\theta=\sqrt{\pi/2}}^{\theta=\sqrt{\pi}} \left(\theta\cos\left(\theta^2\right)\right) d\theta$
c. $\int_{x=0}^{x=1} \left(x\sqrt{1-x^2}\right) dx$
d. $\int_{t=-1}^{t=1} \left(\dfrac{2t+1}{\sqrt[3]{t^2+t+2}}\right) dt$

Exercise 5.6.8

Evaluate all of the definite integrals above.

Answers

Activity 5.6.2: Picking the Pieces of a Substitution.

a.
- For $\int \dfrac{3x^2+1}{(x^3+x-2)^2} dx$, let $u = x^3+x-2$
- For $\int \cos(x)\sqrt{\sin(x)} dx$, let $u = \sin(x)$
- For $\int \dfrac{(\ln(x))^3}{x} dx$, let $u = \ln(x)$

b.
- For $u = x^3+x-1, du = (3x^2+1)\,dx$
- For $u = \sin(x), du = \cos(x)dx$
- For $u = \ln(x), du = \dfrac{1}{x}dx$

c.
- $\int \dfrac{3x^2+1}{(x^3+x-2)^2} dx = \int \dfrac{1}{u^2} du$
- $\int \cos(x)\sqrt{\sin(x)} dx = \int \sqrt{u} du$
- $\int \dfrac{(\ln(x))^3}{x} dx = \int u^3 du$

d.
- $\int \dfrac{1}{u^2} du = -\dfrac{1}{u} + C$
  $= -\dfrac{1}{x^3+x-2} + C$
- $\int \sqrt{u} du = \dfrac{2u^{3/2}}{3} + C$
  $= \dfrac{2(\sin(x))^{3/2}}{3} + C$
- $\int u^3 du = \dfrac{u^4}{4} + C$
  $= \dfrac{(\ln(x))^4}{4} + C$

Activity 5.6.3: Compare Two Integrals.

b. We'll let you $u = 1-x^4$, and so $du = -4x^3 dx$.

$$\begin{aligned}\int -4x^3\sec^2\left(1-x^4\right)dx &= \int \sec^2(\underbrace{1-x^4}_{u})\underbrace{\left(-4x^3\right)dx}_{du}\\ &= \int \sec^2(u)du\\ &= \tan(u)+C\\ &= \tan\left(1-x^4\right)+C\end{aligned}$$

c.

d.

e. $$\begin{aligned}\int x^3\sec^2\left(1-x^4\right)dx &= \int\left(-\frac{1}{4}\right)(-4)x^3\sec^2\left(1-x^4\right)dx\\ &= -\frac{1}{4}\int -4x^3\sec^2\left(1-x^4\right)dx\end{aligned}$$

f. $\int x^3\sec^2\left(1-x^4\right)dx = -\frac{1}{4}\tan\left(1-x^4\right)+C$

This page titled 5.6: Introduction to \(u\) Substitution is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Peter Keep.

6: APPLICATIONS OF INTEGRALS

6.1: Integrals as Net Change
6.2: Area Between Curves
6.3: Volumes of Solids of Revolution
6.4: More Volumes- Shifting the Axis of Revolution
6.5: Arc Length and Surface Area
6.6: Other Applications of Integrals

This page titled 6: Applications of Integrals is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Peter Keep.

6.1: Integrals as Net Change

We have some rudimentary ideas of what an integral is, but we want to challenge and expand those ideas by examining the object at the root of the definition of the Definite Integral: a Riemann Sum. We want to generalize a bit more our notion of what a Riemann sum is. So for now, let's think about how we can use a Riemann sum to think about a measurement that will not be an area. That's been our only real context so far, so let's try to stretch that thinking.

Estimating Movement

Activity 6.1.1: Estimating Movement.

We're observing an object traveling back and forth in a straight line. Throughout a 5 minute interval, we get the following information about the velocity (in feet/second) of the object

Table 6.1.1 Velocity of an Object

t	$v(t)$
0	0
30	2
60	4.25
90	5.75
120	3.5
150	0.75
180	-1.25
210	-3.5
240	-2.75
270	-0.5
300	-0.25

a. Describe the motion of the object in general. **Hint**. How do we interpret the different values of velocity? How do we interpret the sign of velocity? What about how velocity changes from one of the 30-second time points to the next?
b. When was the acceleration of the object the greatest? When was it the least? **Hint**. You can decide how to interpret the "least" acceleration: it is either where the acceleration is closest to 0, or it is the most negative value of the acceleration. These are interpreted differently, but it's a bit ambiguous what we might mean when we say "least acceleration."
c. Estimate the total displacement of the object over the 5 minute interval. What is the overall change in position from the start to the end? **Hint**. How do we use velocity and some time interval to estimate the distance traveled? How do we estimate/assume the velocity on each 30-second time interval?
d. Is this different than the total distance that the object traveled over the 5 minute interval? Why or why not? **Hint**. How do we think about (or ignore) the direction of the object? Why is this important here?
e. If we know the initial position of the object, how could we find the position of the object at some time, t, where t is a multiple of 30 between 0 and 300? **Hint**. Can we limit the time intervals that we use to calculate the object's displacement? How do we use displacement and a starting point to find an ending point?

So what are the big ideas in this short activity? There are a lot, and many of them already are things we know, at least to some level. So we are really focusing on adding depth to our understanding of these big ideas. Let's list them in the order that they showed up in this activity:

1. We interpret the velocity as the derivative of the position of the object. So when we interpret the value of the velocity of the

object (large vs small, positive vs negative, etc.) we are interpreting these through the lens of a rate of change.
2. Acceleration is the derivative of the velocity function. While we don't have the full picture of the velocity function at any value of t, we still were interested in the rates at which velocity changes with regard to time.
3. We can estimate the total displacement of the object by predicting how far it traveled in each 30-second time interval. We might pick the starting velocity for each 30-second interval and multiply that by 30 seconds. We could alternatively pick the ending velocity of each 30-second interval. Then, we can add all of these products of velocity and time together to approximate a total change in position! Doesn't this feel like a Riemann sum?
4. When we calculate displacement, the negative velocities get multiplied out to get negative changes in position for the object -- that's because a negative velocity means that the object is moving backwards. If we wanted to calculate the distance traveled, then we need to not account for negative velocities. We can just disregard the sign of the velocity on each time interval and repeat the process above. So, another Riemann sum then?
5. In order to forecast some position at time t, we just need to start with the initial position, and then calculate (or approximate) the displacement from $t = 0$ to whatever time $t \leq 300$ we care about, and then add the displacement to the initial position.

Ok, now let's formalize those results!

Position, Velocity, and Acceleration

We know that the velocity of an object is really a rate of change of the position of that object with regard to time. Similarly, the acceleration of an object is the rate of change of the velocity of the object with regard to time. So we're really thinking about derivatives!

Definition 6.1.2: Position, Velocity, and Acceleration Functions.

For an object moving along a straight line, if $s(t)$ represents the position of that object at time t, then the **velocity** of the object at time t is $v(t) = s'(t)$ and the **acceleration** of the object at time t is $a(t) = v'(t) = s''(t)$.

Once we establish this relationship, we can answer questions about movement of an object using the same interpretations of derivatives that we practiced in Chapter 3 of this text.

Exercise 6.1.2: A Friendly Jogger.

Consider a jogger running along a straight-line path, where their velocity at t hours is $v(t) = 2t^2 - 8t + 6$, and velocity is measured in miles per hour. We begin observing this jogger at $t = 0$ and observe them over a course of 3 hours.

a. When is the jogger's acceleration equal to 0 mi/hr^2 ? **Hint**. Solve $a(t) = v'(t) = 0$
b. Does this time represent a maximum or minimum velocity for the jogger? **Hint**. You can use the First Derivative Test or the Second Derivative Test here!
c. When is the jogger's velocity equal to 0 mi/hr?
d. Describe the motion of the jogger, including information about the direction that they travel and their top speeds.

Displacement, Distance, and Speed

Let's revisit Activity 6.1.1. When we approximated the displacement of the object, we built a Riemann sum:

$$\sum_{k=1}^{10} v\left(t_k^*\right) \Delta t$$

We chose our t_k^* as either the time at the beginning of each 30-second interval or the time at the end of the 30-second interval, but that was only because of the limited information that we had about different values of $v(t)$. If we had information about the $v(t)$ function at any values of t ($0 \leq t \leq 300$), then we could pick any time in each 30-second time interval for our Riemann sum! We might note, though, that if we did have this kind of information about the velocity at any time in the 5-minute interval, then we also would build a more precise approximation by subdividing the time interval into smaller/shorter pieces. So maybe the

Riemann $\sum_{k=1}^{100} v\left(t_k^*\right) \Delta t$ (where we are dividing up the 5 minute interval into 100, 3-second intervals) would do a better job! But why stop there? If we have the definition of the velocity function, and so we can truly obtain the velocity of the object at any time in the 5 minute interval, then we can use the definition of the definite integral as the limit of a Riemann sum:

$$\lim_{n\to\infty} \sum_{k=1}^{n} v\left(t_k^*\right) \Delta t = \int_{t=0}^{t=300} v(t)dt$$

This should work out well with our first understanding of displacement: the displacement of an object is just the difference in position from the starting time to the ending time. So we could say that if $s(t)$ is the position function, then we might expect to represent displacement from $t = a$ to $t = b$ as $s(b) - s(a)$. But isn't this just the Fundamental Theorem of Calculus, since $s'(t) = v(t)$?

Definition 6.1.3: Displacement of an Object.

If an object is moving along a straight line with velocity $v(t)$ and position $s(t)$, then the displacement of the object from time $t = a$ to $t = b$ is

$$\int_{t=a}^{t=b} v(t)dt = s(b) - s(a)$$

Let's keep revisiting the same activity. We also noticed that when we looked at the distance compared to the displacement, the only difference was that we were integrating the absolute value of the velocity function, since we didn't care about the sign of the velocity (the direction that the object was traveling) on each interval.

Definition 6.1.4: Distance Traveled.

If an object is moving along a straight line with velocity $v(t)$, then the distance traveled by the object from time $t = a$ to $t = b$ is:

$$\int_{t=a}^{t=b} |v(t)|dt$$

Here, we call $|v(t)|$ the speed of the object (instead of the velocity).

We should note that we don't have any quick and easy ways of dealing with the integral of the absolute value of a function.

$$|v(t)| = \begin{cases} -v(t) \text{ when } v(t) < 0 \\ v(t) \text{ when } v(t) \geq 0 \end{cases}$$

So, in order for us to integrate $|v(t)|$, we need to think about where the velocity passes through 0, so that we can see where it might change from positive to negative.

Exercise 6.1.3: Tracking our Jogger.

Let's revisit our jogger from Activity 6.1.2.

a. Calculate the total displacement of the jogger from $t = 0$ to $t = 3$. **Hint**. Set up and evaluate a definite integral here, using the velocity function.
b. Think back to our description of the jogger's movement: when is this jogger moving backwards? Split up the time interval from $t = 0$ (the start of their run) to $t = c$ (where c is the time that the jogger changed direction) to $t = 3$. Calculate the displacements on each of these two intervals.
c. Calculate the total distance that the jogger traveled in their 3 hour run. **Hint**. Remember that we're really calculating:

$$\left|\int_{t=0}^{t=c} v(t)dt\right| + \left|\int_{t=c}^{t=3} v(t)dt\right|$$

Finding the Future Value of a Function

We can again think back to Activity 6.1.1 and build our last result of this section. Remember when we were looking to predict the location of our object at different times: we said it was reasonable to start at our initial position, and then add the displacement of the object from that initial time up to the time that we were interested in. So, to estimate the object's position after 150 seconds, we would calculate:

$$s(0) + \int_{t=0}^{t=150} v(t)dt$$

But we said we could do this to estimate the object's position at any value for time, t.

Theorem 6.1.5: Future Position of an Object.

For some object moving along a straight line with velocity $v(t)$ and an initial position of $s(a)$, the **future position of the object** at some time t (with $t \geq a$) is:

$$\underbrace{s(t)}_{\substack{\text{future}\\ \text{position}}} = \underbrace{s(a)}_{\substack{\text{initial}\\ \text{position}}} + \underbrace{\int_{x=a}^{x=t} v(x)dx}_{\substack{\text{displacement from}\\ \text{a to } t}}$$

Note that we change the variable in the velocity function while we integrate: since we want our position function to be in terms of t, the ending time that we calculate the displacement up to, we need to choose to write the velocity function in terms of a different variable. Mechanically, there is no difference, since we're just swapping out the variables and naming them x.

We can note that this relationship between velocity and position can exist in many other contexts: any pair of functions that are derivatives/antiderivatives of each other can have this relationship!

Theorem 6.1.6: Net Change and Future Value.

Suppose the value $F(t)$ changes over time at a known rate $F'(t)$. Then the **net change** in F between $t = a$ and $t = b$ is:

$$F(b) - F(a) = \int_{t=a}^{t=b} F'(t)dt$$

Similarly, given the initial value $F(a)$, the **future value** of F at time $t \geq a$ is:

$$F(t) = F(a) + \int_{x=a}^{x=t} F'(x)dx$$

Practice Problems

Exercise 6.1.1

Explain the following terms in reference to an object moving along a straight path from time $t = a$ to time $t = b$.

a. Position of the object at time t.
b. Displacement of the object.
c. Distance traveled by the object.
d. Velocity of the object at time t.
e. Speed of the object at time t.

Exercise 6.1.2

Consider the graph of a velocity function, v(t), of some object moving along a line on the time interval [0, 7].

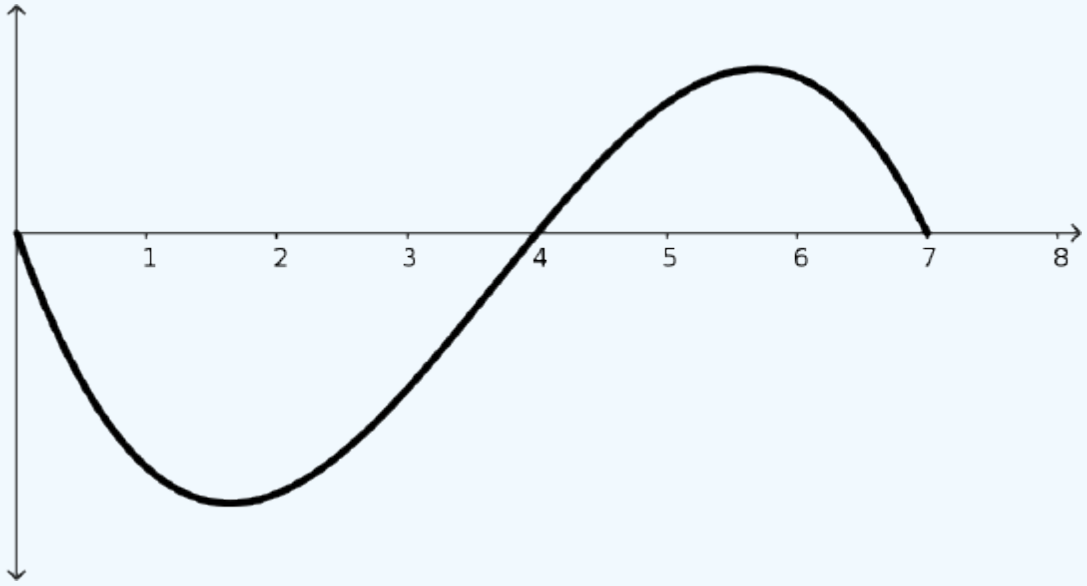


Figure 6.1.7

a. Do you expect the displacement of the object from $t = 0$ to $t = 7$ to be positive, negative, or 0?
b. Write two different expressions that represent the total displacement of the object from $t = 0$ to $t = 7$.
c. Do you expect the distance traveled by the object from $t = 0$ to $t = 7$ to be positive, negative, or 0?
d. Write two different expressions that represent the total distance traveled by the object from $t = 0$ to $t = 7$.

Answer available at the end of this section.

Exercise 6.1.3

Let's consider an animal running along a straight path with the velocity function:

$$\begin{aligned} v(t) &= \frac{t^4}{10} - t^3 + \frac{27t^2}{10} - \frac{9t}{5} \\ &= \frac{t}{10}(t-1)(t-3)(t-6) \end{aligned}$$

on the time interval [0, 6].

a. What is the total displacement of the animal on the time interval [0, 1]?
b. What is the total displacement of the animal on the time interval [1, 3]?
c. What is the total displacement of the animal on the time interval [3, 6]?
d. What is the total displacement of the animal on the time interval [0, 6]?
e. What is the total distance traveled by the animal on the time interval [0, 6]?
f. Write a short summary of the animal's movement, including notes about direction, speed, and where the animal travels.

Answer available at the end of this section.

Exercise 6.1.4

Consider an object with velocity function $v(t) = t^2 - 4t + 2$ on the interval [0, 100] with the initial position $s(0) = 3$.

a. Determine the position function, $s(t)$, for $0 \le t \le 100$ using the Future Position of an Object..
b. Determine the position function, $s(t)$, for $0 \le t \le 100$ using the Solving Initial Value Problems strategy.
c. Compare the results from both methods. Explain why these are equivalent.

Answer available at the end of this section.

Exercise 6.1.5

Consider an object with an acceleration function $a(t) = t + \sin(2\pi t)$ for $t \ge 0$ with $v(0) = 5$.

a. Determine the velocity function, $v(t)$, for $t \ge 0$ using the Future Position of an Object.
b. Determine the velocity function, $v(t)$, for $t \ge 0$ using the Solving Initial Value Problems strategy.
c. Can you obtain the position function, $s(t)$? Explain why or why not, based on the information given.

Answer available at the end of this section.

Exercise 6.1.6

During a brake test for a heavy truck, the truck decelerates from an initial velocity of 88 ft/s with the acceleration function $a(t)$ = –17 ft/s^2. Assume that the initial position of the truck is $s(0) = 0$.

a. Find the velocity function for the truck.
b. When does the truck stop? In this situation, the truck won't have a negative velocity (since it's just braking and not eventually going in reverse). What time interval is the velocity function relevant on?
c. What is the total displacement of the truck on this time interval?
d. Safety standards say that for a truck like this, it needs to be able to stop (from a speed of 88ft/s) in, at most, 200 feet. Do we need to make changes to the braking mechanism, in order to have the acceleration function change? If so, what does the acceleration need to be (assuming it is constant and we are just replacing it with a new negative number)?

Answer available at the end of this section.

Answers

Exercise 6.1.2

a. Displacement will likely be negative, since it looks like the size of the integral between $t = 0$ and $t = 4$ is larger than the size of the integral from $t = 4$ up to $t = 7$. Another way to say this is that there is more negative velocity accumulated than positive velocity.
b. There are plenty of ways to do this! Here are two:
   i. $\int_{t=0}^{t=7} v(t)dt$
   ii. $\int_{t=0}^{t=4} v(t)dt + \int_{t=4}^{t=7} v(t)dt$
c. Distance is always non-negative, and since the velocity was not just constant at 0 (with no movement), then we expect the distance to be positive. It is the integral of speed, which is $|v(t)|$, velocity without taking into account the direction.
d. There are plenty of ways to do this! Here are two:
   i. $\int_{t=0}^{t=7} |v(t)|dt$
   ii. $-\int_{t=0}^{t=4} v(t)dt \int_{t=4}^{t=7} v(t)dt$

Exercise 6.1.3

a.
$$\int_{t=0}^{t=1} v(t)dt = \int_{t=0}^{t=1} \frac{t^4}{10} - t^3 + \frac{27t^2}{10} - \frac{9t}{5}dt$$
$$\left(\frac{t^5}{50} - \frac{t^4}{4} + \frac{27t^3}{30} - \frac{9t^2}{10}\right)\Big|_{t=0}^{t=1}$$
$$= -\frac{23}{100}$$

b.
$$\int_{t=1}^{t=3} v(t)dt = \int_{t=1}^{t=3} \frac{t^4}{10} - t^3 + \frac{27t^2}{10} - \frac{9t}{5}dt$$
$$\left(\frac{t^5}{50} - \frac{t^4}{4} + \frac{27t^3}{30} - \frac{9t^2}{10}\right)\Big|_{t=1}^{t=3}$$
$$= \frac{26}{25}$$

c. $$\begin{aligned}\int_{t=3}^{t=6} v(t)dt &= \int_{t=3}^{t=6} \frac{t^4}{10} - t^3 + \frac{27t^2}{10} - \frac{9t}{5} dt \\ &\left(\frac{t^5}{50} - \frac{t^4}{4} + \frac{27t^3}{30} - \frac{9t^2}{10}\right)\Big|_{t=3}^{t=6} \\ &= -\frac{729}{100}\end{aligned}$$

d. We can integrate from $t = 0$ to $t = 6$, or just add up the previous displacements.

$$\begin{aligned}\int_{t=0}^{t=6} v(t)dt &= \int_{t=0}^{t=6} \frac{t^4}{10} - t^3 + \frac{27t^2}{10} - \frac{9t}{5} dt \\ &\left(\frac{t^5}{50} - \frac{t^4}{4} + \frac{27t^3}{30} - \frac{9t^2}{10}\right)\Big|_{t=0}^{t=6} \\ &= -\frac{162}{25}\end{aligned}$$

e. $$\begin{aligned}\int_{t=0}^{t=6} |v(t)|dt &= \left|\int_{t=0}^{t=1} v(t)dt\right| \\ &+ \left|\int_{t=1}^{t=3} v(t)dt\right| \\ &+ \left|\int_{t=3}^{t=6} v(t)dt\right| \\ &= \frac{214}{25}\end{aligned}$$

f. From $t = 0$ to $t = 1$, the animal is moving backwards, slowly. It speeds up until around $t \approx 0.4$ and then slows down until it stops and turns around at $t - 1$. From then until $t = 3$ it moves forwards. It speeds up until around $t \approx 2.1$, and then slows down until it stops to turn back around at $t = 3$. Then, it speeds up quickly in the negative direction until $t = \approx 5$ when it begins to slow down to turn around at $t = 6$. It accelerates very quickly in the positive direction until $t = 7$.

Exercise 6.1.4

a. $$\begin{aligned}s(t) &= s(0) + \textstyle\int_{x=0}^{x=t} v(x)dx \\ &= 3 + \textstyle\int_{x=0}^{x=t} x^2 - 4x + 2dx \\ &= 3 + \left(\frac{x^3}{3} - 2x^2 + 2x\right)\Big|_{x=0}^{x=t} \\ &= \frac{t^3}{3} - 2t^2 + 2t + 3 \qquad \text{on the interval } [0, 100]\end{aligned}$$

b. We'll find the family of antiderivatives, of which $s(t)$ is one.

$$\begin{aligned}\int v(t)dt &= \int t^2 - 4t + 2dt \\ &= \frac{t^3}{3} - 2t^2 + 2t + C\end{aligned}$$

Now, $s(t)$ is the antiderivative where $s(0) = 3$.

$$\begin{aligned}3\frac{0^3}{3} - 2(0)^2 + 2(0) + C \\ 3 = C\end{aligned}$$

So $s(t) = \frac{t^3}{3} - 2t^2 + 2t + 3$.

Exercise 6.1.5

a. $$\begin{aligned}a(t) &= v(0) + \int_{x=0}^{x=t} a(x)dx \\ &= 5 + \int_{x=0}^{x=t} (x + \sin(2\pi x))dx \\ &= 5 + \left(\frac{x^2}{2} - \frac{\cos(2\pi x)}{2\pi}\right)\Big|_{x=0}^{x=t} \\ &= \frac{t^2}{2} - \frac{\cos(2\pi t)}{2\pi} + 5 + \frac{1}{2\pi} \quad \text{for } t \geq 0\end{aligned}$$

b. $$\begin{aligned}\int a(t)dt &= \int t + \sin(2\pi t)dt \\ &= \frac{t^2}{2} - \frac{\cos(2\pi t)}{2\pi} + C\end{aligned}$$

We know that $v(t)$ is one of these antiderivatives, so we can use $v(0) = 5$ to solve for C.

$$\begin{aligned}5 = \frac{0^2}{2} - \frac{\cos(0)}{2\pi} + C \\ 5 + \frac{1}{2\pi} = C\end{aligned}$$

So $v(t) = \frac{t^2}{2} - \frac{\cos(2\pi t)}{2\pi} + 5 + \frac{1}{2\pi}$ for $t \geq 0$.

Exercise 6.1.6

a. $$\begin{aligned}v(t) &= v(0) + \int_{x=0}^{x=t} a(x)dx \\ &88 + \int_{x=0}^{x=t} -17dx \\ &88 - 17t\end{aligned}$$

b. The truck stops when $v(t) = 0$:

$$0 = 88 - 17t$$
$$t\frac{88}{17}$$

This means that the velocity function is only relevant on the interval $[0, \frac{88}{17}]$.

c.
$$\int_{t=0}^{t=88/17} v(t)dt = \int_{t=0}^{t=88/17} 88 - 17tdt$$
$$= \left(88t - \frac{17t^2}{2}\right)\Big|_{t=0}^{t=88/17}$$
$$\frac{3872}{17}$$

So the truck moved approximately 227.76 feet.

d. Note that for an acceleration function $a(t) = -k$ for some positive number k, the velocity function is $v(t) = 88 - kt$ and the truck stops at $t = \frac{88}{k}$. Now we can set the displacement integral equal to 200 in order to solve for k.

$$200 = \int_{t=0}^{t=88/k} 88 - ktdt$$
$$= \left(88t - \frac{kt^2}{2}\right)\Big|_{t=0}^{t=88/k}$$
$$= \frac{88^2}{k} - \frac{88^2}{2k}$$
$$200 = \frac{88^2}{2k}$$
$$400k88^2$$
$$k = \frac{484}{25} = 19.36$$

So the truck needs to have an acceleration function of $a(t) = -19.36$ ft/s^2 or more negative for it to stop in time.

This page titled 6.1: Integrals as Net Change is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Peter Keep.

6.2: Area Between Curves

We're going to stick with our theme of thinking about a Riemann Sum, but this time we'll get back to thinking about area. First, we'll try to remind ourselves not just of what a Riemann sum is, but how we actually constructed it.

Remembering Riemann Sums

Activity 6.2.1: Remembering Riemann Sums.

Let's start with the function $f(x)$ on the interval $[a, b]$ with $f(x) > 0$ on the interval. We will construct a Riemman sum to approximate the area under the curve on this interval, and then build that into the integral formula.

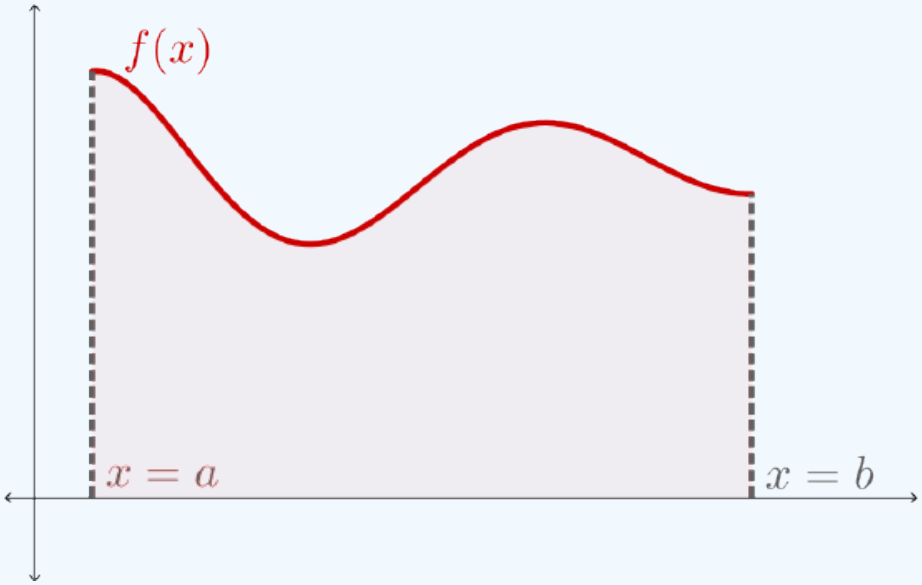


Figure 6.2.1

a. Divide the interval $[a, b]$ into 4 equally-sized subintervals.
b. Pick an x_{x^*} for k = 1, 2, 3, 4, one for each subinterval. Then, plot the points $(x_1^*, f(x_1^*)), (x_2^*, f(x_2^*)), (x_3^*, f(x_3^*))$, and $(x_4^*, f(x_4^*))$. **Hint**. These points are just general ones, and you don't have to come up with actual numbers for the x-values or the corresponding y-values. Instead, just draw them in on the curve, somewhere in each of the subintervals.
c. Use these 4 points to draw 4 rectangles. What are the dimensions of these rectangles (the height and width)? **Hint**. You won't have any numbers to calculate here, really: instead, see if you can calculate the widths by thinking about the total width of your interval. Then calculate the heights by thinking about the points you created.
d. Find the area of each rectangle by multiplying the heights and widths for each rectangle.
e. Add up the areas to construct a Riemman sum. Is this sum very accurate? Why or why not? **Hint**. Try to think about the accuracy of your area approximation by looking at it visually. Are there any places where your approximation looks far away from the actual area we're thinking about?
f. Now we will generalize a little more. Let's say we divide this up into n equally-sized pieces (instead of 4). Instead of trying to pick an x_k^* for the unknown number of subintervals (since we don't have a value for n yet), let's just focus on one of these: the arbitrary kth subinterval.

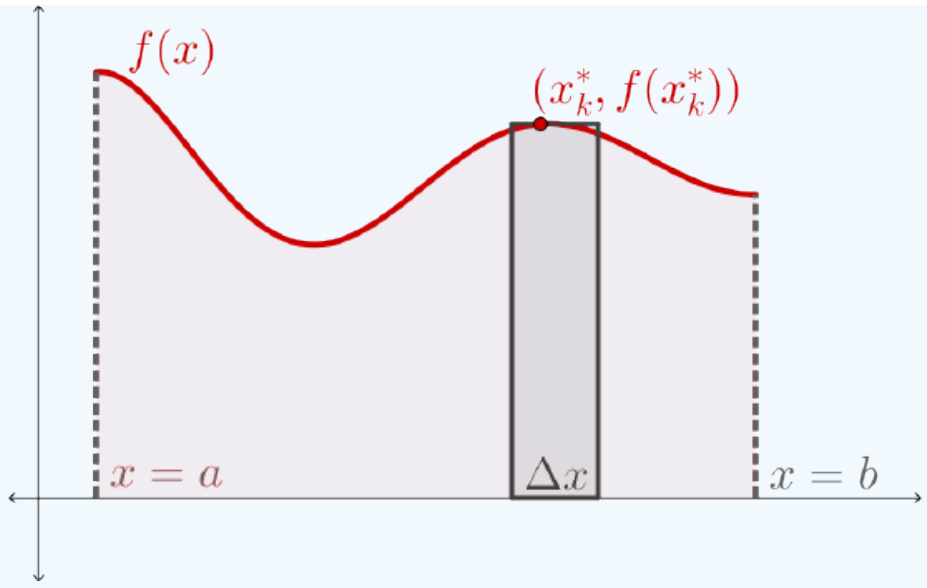


Figure 6.2.2

What are the dimensions of this kth rectangle?

g. Find A_k, the area of this kth rectangle.
h. Add up the areas of A_k for $k = 1, 2, 3, \ldots, n$ to approximate the total area, A. **Hint**. You might want to use summation notation, starting with $\sum_{k=1}^{n}$
i. Apply a limit as $n \to \infty$ to this Riemann sum in order to construct the integral formula for the area under the curve $f(x)$ from $x = a$ to $x = b$.

Answer available at the end of this section.

Hopefully this is helpful. If you'd like more reminders on this, you can always revisit Section 5.2 Riemann Sums and Area Approximations. For now, though, we mainly want to think about the general process we're using:

1. We slice the region from $x = a$ to $x = b$ into n pieces, and, for convenience, we choose an equal width: $\Delta x = \dfrac{b-a}{n}$.
2. From each of the slices, we select some x-value (called $x_k^* x_k^*$ from the kth slice). We use that to evaluate the function on each slice: $f\left(x_k^*\right)$.
3. We multiply the function value, $f\left(x_k^*\right)$, by the width of the slice, Δx, to get the measured area of each slice, $A_k = f\left(x_k^*\right)\Delta x$.
4. We can estimate the total measured area of the region by adding all of the areas of the slices together:

$$A \approx \sum_{k=1}^{n} f\left(x_k^*\right)\Delta x$$

5. If we keep adding more and more slices (that keep getting thinner and thinner), then we eventually (in the limit) evaluate the area exactly:

$$A = \lim_{n\to\infty} \sum_{k=1}^{n} f\left(x_k^*\right)\Delta x$$

We're going to use this process (we'll call it the **slice-and-sum** process) for other measurements! Let's see how we can change this slightly to measure a different area.

Building an Integral Formula for the Area Between Curves

Activity 6.2.2: Area Between Curves.

Let's start with our same function $f(x)$ on the same interval $[a, b]m$ but also add the function $g(x)$ on the same interval, with $f(x) > g(x) > 0$ on the interval. We will construct a Riemann sum to approximate the area between these two curves on this interval, and then build that into the integral formula.

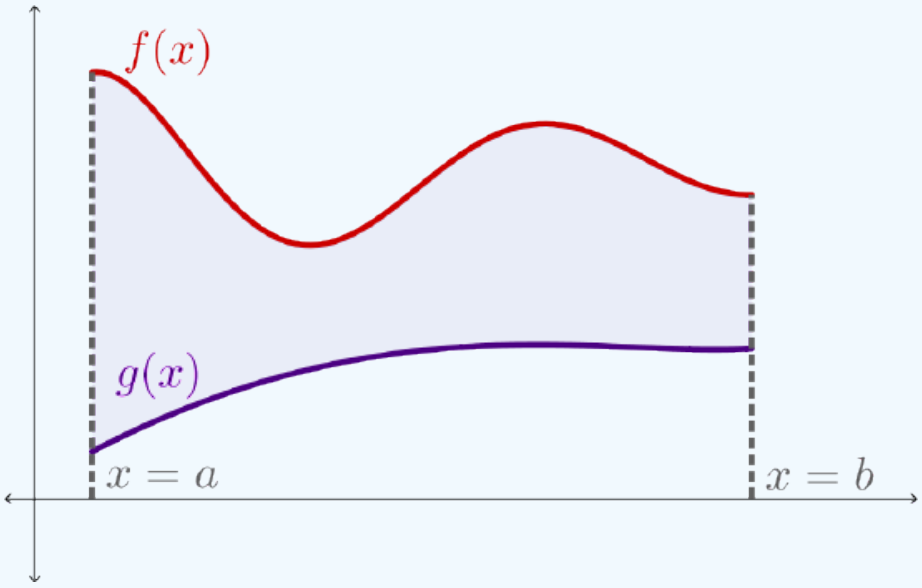


Figure 6.2.3

a. Divide the interval $[a, b]$ into 4 equally-sized subintervals.
b. Pick an x_k^* for k = 1, 2, 3, 4, one for each subinterval. Plot the points $(x_1^*, f(x_1^*))$, $(x_2^*, f(x_2^*))$, $(x_3^*, f(x_3^*))$, and $(x_4^*, f(x_4^*))$. Then plot the corresponding points on the g function: $(x_1^*, g(x_1^*))$, $(x_2^*, g(x_2^*))$, $(x_3^*, g(x_3^*))$, and $(x_4^*, g(x_4^*))$.
c. Use these 8 points to draw 4 rectangles, with the points on the f function defining the tops of the rectangles and the points on the g function defining the bottoms of the rectangles. What are the dimensions of these rectangles (the height and width)?
d. Find the area of each rectangle by multiplying the heights and widths for each rectangle.
e. Add up the areas to construct a Riemman sum.
f. Now we will generalize a little more. Let's say we divide this up into n equally-sized pieces (instead of 4). Instead of trying to pick an x_k^* for the unknown number of subintervals (since we don't have a value for n yet), let's just focus on one of these: the arbitrary kth subinterval.

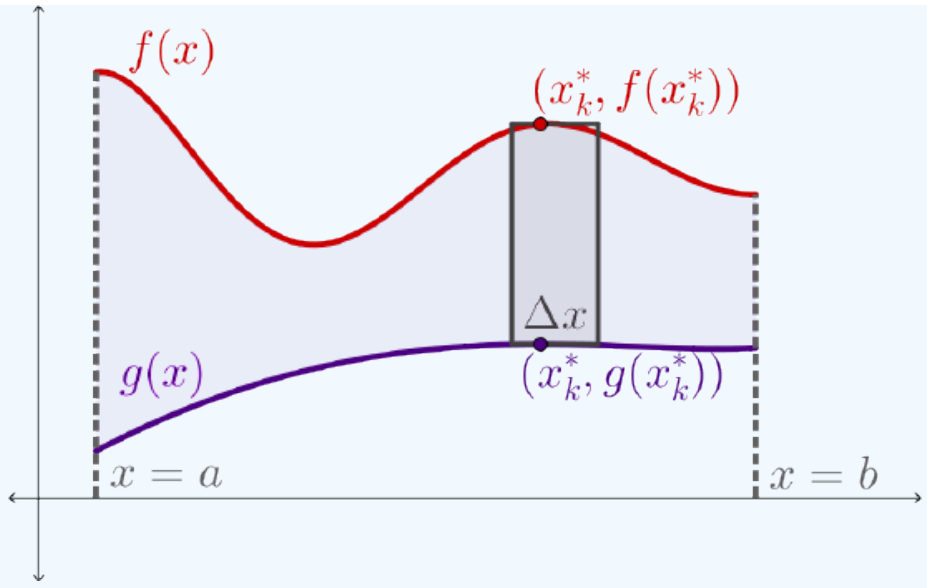


Figure 6.2.4

What are the dimensions of this kth rectangle?

g. Find A_k, the area of this kth rectangle.

h. Add up the areas of A_k for $k = 1, 2, 3, \ldots, n$ to approximate the total area, A. **Hint**. You might want to use summation notation, starting with $\sum_{k=1}^{n}$

i. Apply a limit as $n \to \infty$ to this Riemann sum in order to construct the integral formula for the area between the curves $f(x)$ and $g(x)$ from $x = a$ to $x = b$.

Answer available at the end of this section.

Definition 6.2.5: Area Between Curves.

If $f(x)$ and $g(x)$ are continuous functions with $f(x) \geq g(x)$ on the interval $[a, b]$, then the area bounded between the curves $y = f(x)$ and $y = g(x)$ between $x = a$ and $x = b$ is

$$A = \int_{x=a}^{x=b} (f(x) - g(x))dx$$

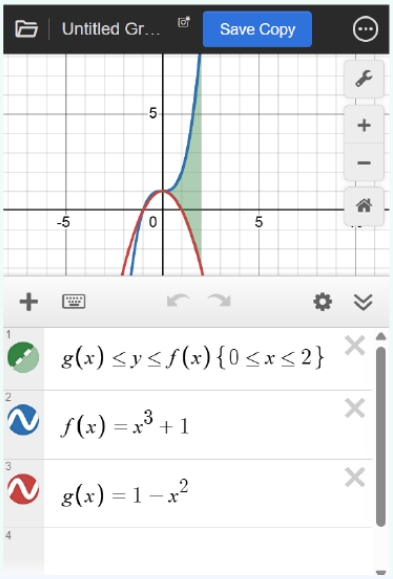
When we're applying this formula for the area between curves, we won't need to recreate the process from Activity 6.2.2 to create the integral formula. We simply can identify the functions bounding the region and the end points of the interval, and set up the integral. We'll use the slice-and-sum process for two reasons:

1. To justify these formulas that we continue to build! While this one isn't that difficult (you could have just built the formula by thinking about the area between curves as a difference in areas under each curves), some of the formulas we play with in this chapter will not be as intuitive.
2. To help us understand what a Riemann sum actually is. It's a product of a function value from a subinterval multiplied by the width of that subinterval, summed up across some larger interval.

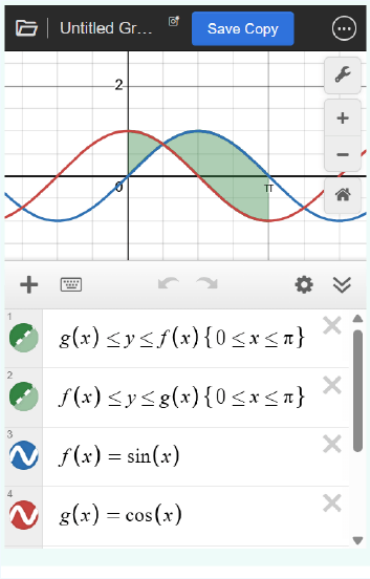
Example 6.2.1

For each of the following regions, set up an integral expression representing the area of the region. We also can practice evaluating these integrals to actually calculate the areas. For each of these described regions, the hint will reveal a visualization of the region (using Desmos). Feel free to use that to set up the integral expression!

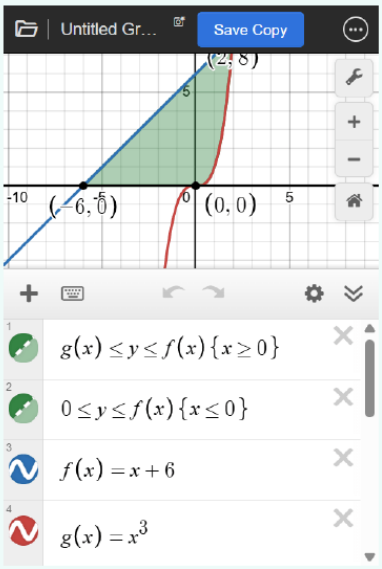
a. The region bounded between the graphs $y = x^3 + 1$ and $y = 1 - x^2$ from $x = 0$ to $x = 2$.



b. The region bounded between the graphs $y = \sin(x)$ and $y = \cos(x)$ from $x = 0$ to $x = \pi$.



c. The region bounded between the curves $y = x + 6$ and $y = x^3$ and the x-axis.



Solution

a.
$$\int_{x=0}^{x=2} (x^3+1)-(1-x^2)\,dx = \int_{x=0}^{x=2} (x^3+x^2)\,dx$$
$$= \frac{20}{3}$$

b. Notice that the boundary functions intersect at $x = \frac{\pi}{4}$, and they switch order. We'll need to split this region into two different regions in order to identify the "top" and "bottom" boundary functions.
$$\int_{x=0}^{x=\pi/4} \cos(x)-\sin(x)dx + \int_{x=\pi/4}^{x=\pi} \sin(x)-\cos(x)dx = (\sqrt{2}-1)+(1+\sqrt{2})$$
$$= 2\sqrt{2}$$

c. On the interval $-6 \le x \le 0$, the region is bounded above by $y = x+6$ and below by the x-axis ($y = 0$). On the interval $0 \le x \le 2$, the region is bounded above by $y = x+6$ and below by $y = x^3$.
$$\int_{x=-6}^{x=0} (x+6)-0dx + \int_{x=0}^{x=2} (x+6)-(x^3)\,dx = \int_{x=-6}^{x=0} x+6dx + \int_{x=0}^{x=2} 6+x-x^3dx$$
$$= 18+10$$
$$= 28$$

Changing Perspective

This last example had two interesting regions: we had to split them into two pieces because the boundary functions changed order or, in the case of the last example, changed completely to different boundary functions.

We're going to redo the last problem and work on trying to change our perspective a bit in order to get a single integral to evaluate the area.

Activity 6.2.3: Trying for a Single Integral.

Let's consider the same setup as earlier: the region bounded between two curves, $y = x+6$ and $y = x^3$, as well as the x-axis (the line $y = 0$). We'll need to name these functions, so let's call them $f(x) = x^3$ and $g(x) = x+6$. But this time, we'll approach the region a bit differently: we're going to try to find the area of the region using only a single integral.

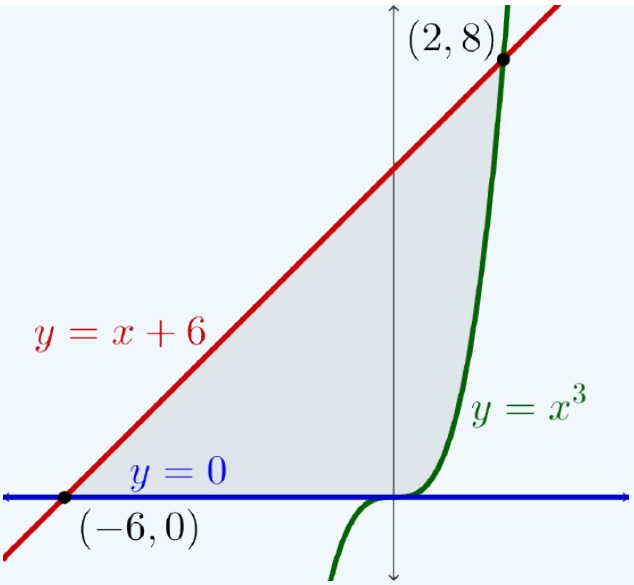


Figure 6.2.7

a. The range of y-values in this region span from $y = 0$ to $y = 8$. Divide this interval evenly into 4 equally sized-subintervals. What is the height of each subinterval? We'll call this Δy. **Hint**. $\Delta y = \frac{8-0}{4}$
b. Pick a y-value from each sub-interval. You can call these y_1^*, y_2^*, y_3^*, and y_4^*.
c. Find the corresponding x-values on the $f(x)$ function for each of the y-values you selected. These will be $f^{-1}(y_1^*), f^{-1}(y_2^*), f^{-1}(y_3^*)$, and $f^{-1}(y_4^*)$. **Hint**. You're really just putting your y-values into the equation $y = x + 6$ and solving for x. Or you can solve for $f^{-1}(y)$ in general, by solving for x while leaving y as a variable.
d. Do the same thing for the g function. Now you have 8 points that you can plot: $(f^{-1}(y_1^*), y_1^*), (f^{-1}(y_2^*), y_2^*), (f^{-1}(y_3^*), y_3^*)$, and $(f^{-1}(y_4^*), y_4^*)$ as well as $(g^{-1}(y_1^*), y_1^*), (g^{-1}(y_2^*), y_2^*), (g^{-1}(y_3^*), y_3^*)$, and $(g^{-1}(y_4^*), y_4^*)$. Plot them.
e. Use these points to draw 4 rectangles with points on f and g, determining the left and right ends of the rectangle. What are the dimensions of these rectangles (height and width)?
f. Find the area of each rectangle by multiplying the height and widths of each rectangle.
g. Add up the areas to construct a Riemann sum.
h. Again, we'll generalize this and think about the kth rectangle, pictured below.

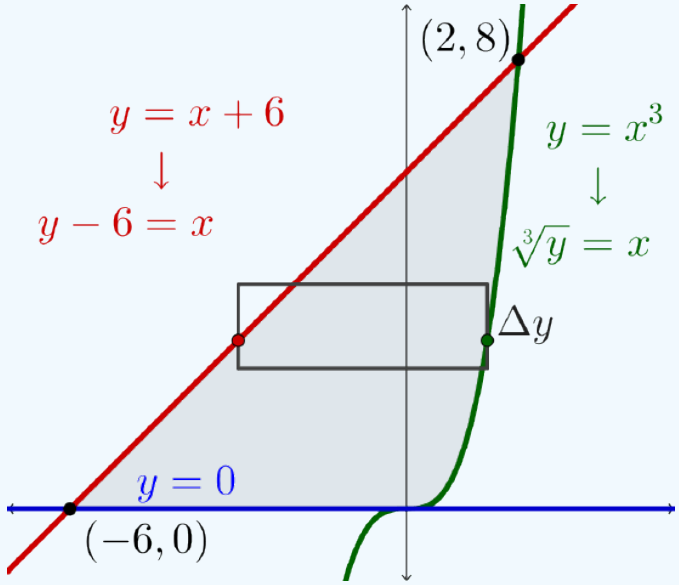


Figure 6.2.8

Which variable defines the location of the kth rectangle, here? That is, if you were to describe where in this graph the kth rectangle is laying, would you describe it with an x or y variable? This will act as our general input variable for the integral we're ending with.

i. What are the dimensions of the kth rectangle?
j. Find A_k, the area of this kth rectangle.
k. Add up the areas of A_k for $k = 1, 2, 3, \ldots, n$ to approximate the total area, A. **Hint**. You might want to use summation notation, starting with $\sum_{k=1}^{n}$
l. Apply a limit as $n \to \infty$ to this Riemann sum in order to construct the integral formula for the area between the curves $f(x)$ and $g(x)$ from $x = a$ to $x = b$.
m. Now that you have an integral, evaluate it! Find the area of this region to compare with the work we did previously, where we used multiple integrals to measure the size of this same region.

Answer available at the end of this section.

We can rewrite our definition of the area between curves (Definition 6.2.5) to account for this change in perspective, by thinking about these same functions in terms of y.

Definition 6.2.9: Area Between Curves (in terms of y).

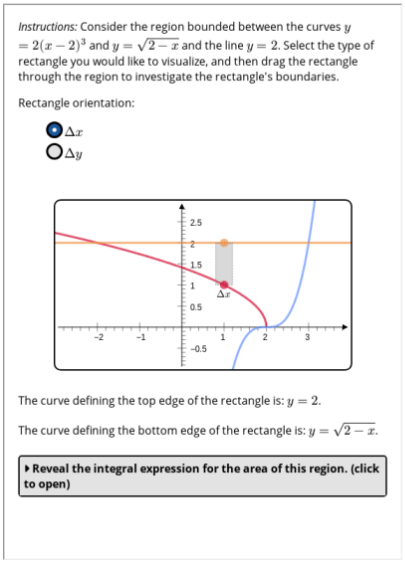
If $f(y)$ and $g(y)$ are continuous functions with $f(y) \geq g(y)$ on the interval of y-values $[c, d]$, then the area bounded between the curves $x = f(y)$ and $x = g(y)$ from $y = c$ to $y = d$ is

$$A = \int_{y=c}^{y=d} (f(y) - g(y))dy$$

This strategy of inverting our functions to change the variable that we integrate with regard to is useful, but a tricky part of this is deciding when to change variables.

Something that we can look for is intersection points in the region we're working with. If, in our plan for setting up an integral, we would stack rectangles that would pass through an intersection point, then this indicates that we would need to split our region up to set up the integrals (since the boundary functions are changing). If we change the orientation of the rectangles, would they still pass through an intersection point? Are the functions that we're working with relatively easy to invert? Can we antidifferentiate these functions, or their inverted versions?

These are some of the things we'll consider as we make these decisions. To finish things up, let's look at a nice little interactive graph that can help show the differences between finding area with regard to x (using Δx in our rectangles and dx in our integrals) and finding area with regard to y (using Δy in our rectangles and dy in our integrals).



Standalone
Embed

Practice Problems

Exercise 6.2.1

Explain how we use the ”slice and sum” method to build an integral formula for the area bounded between curves. Give some details, enough to make sure you understand how the Riemann sums are constructed and how they turn into our integral formula.

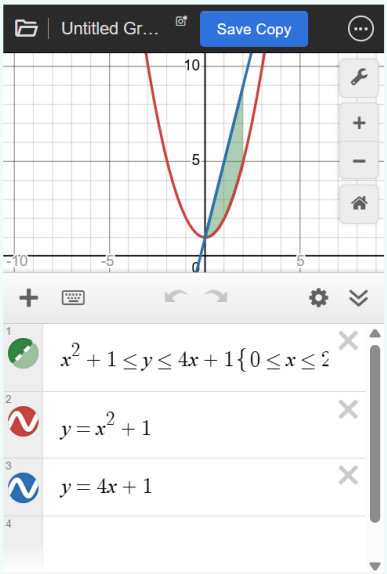
Exercise 6.2.2

What are some changes/considerations that we need to make when we decide to set up our integral in terms of y instead of x?

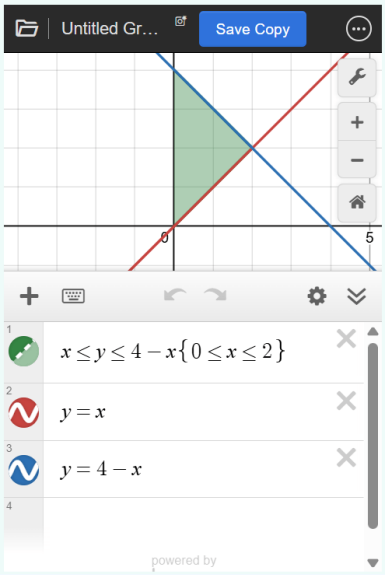
Exercise 6.2.3

Set up (and practice evaluating) an integral expression representing the area of each of the regions described below. The hint for each problem will open a graph of the region.

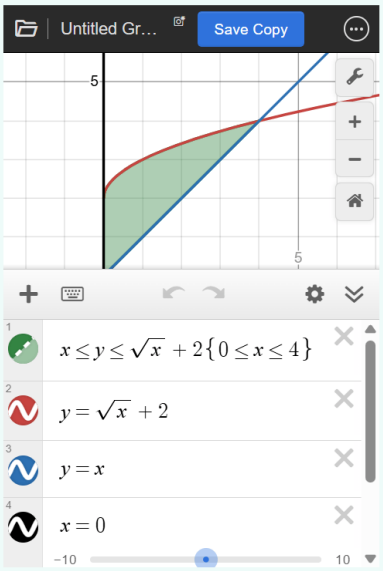
a. The region bounded by the curves $y = x^2 + 1$ and $y = 4x + 1$ between $x = 0$ and $x = 2$. **Hint**.



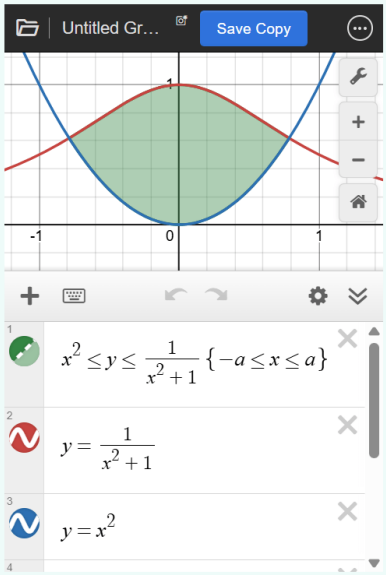
b. The region bounded by the curves $y=x$ and $y=4-x$ between $x=0$ and $x=2$. **Hint**.



c. The region bounded by the curves $y=\sqrt{x}+2$ and $y=x$ and the line $x=0$. **Hint**.



d. The region bounded by the curves $y = \dfrac{2}{x^2+1}$ and $y = x^2$. **Hint.**

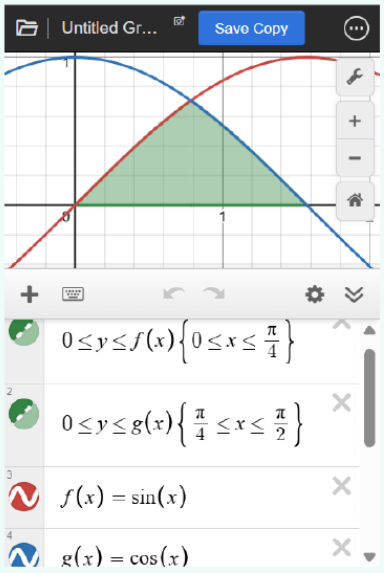


Answer available at the end of this section.

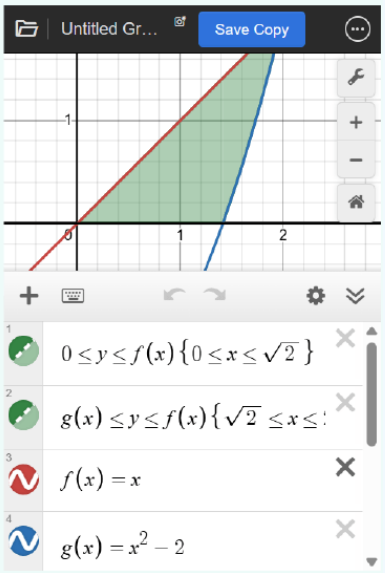
Exercise 6.2.4

Set up and evaluate an integral representing the area of each of the regions described below. Explain whether you chose to integrate with respect to x or y, and why you made that choice.

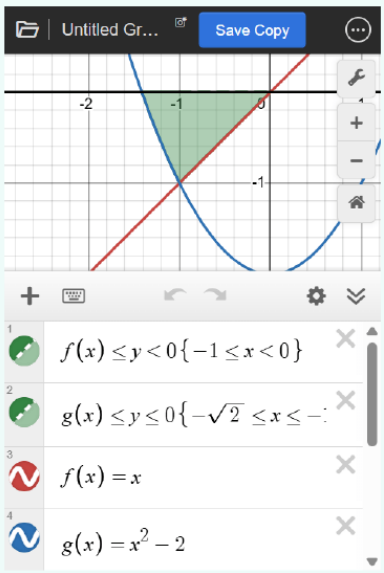
a. The region bounded by the curves $y = \sin(x)$ and $y = \cos(x)$ and the line $y = 0$ between $x = 0$ and $x = \dfrac{\pi}{2}$. **Hint.**



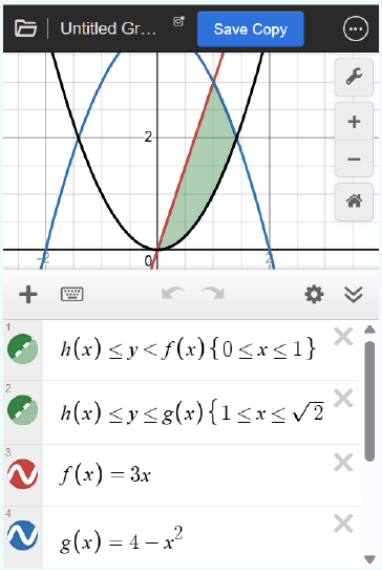
b. The region bounded by the curves $y = x$ and $y = x^2 - 2$ and the line $y = 0$ in the first quadrant. **Hint**.



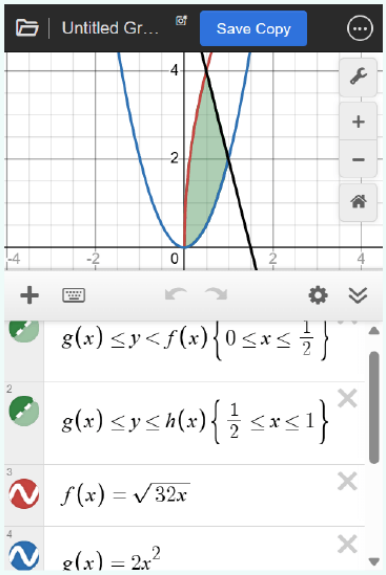
c. The region bounded by the curves $y = x$ and $y = x^2 - 2$ and the line $y = 0$ in the third quadrant. **Hint**.



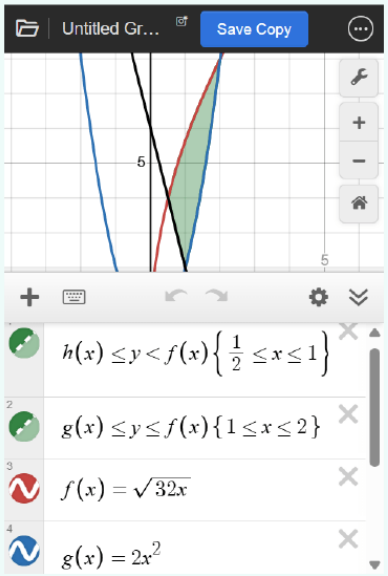
d. The region bounded by the curves $y = 3x, y = 4 - x^2$, and $y = x^2$ in the first quadrant with $0 \le x \le \sqrt{2}$. **Hint**.



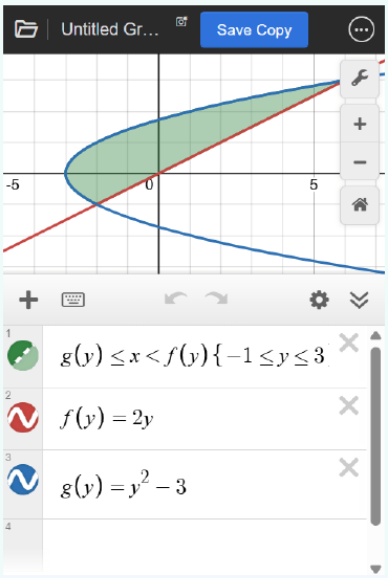
e. The region bounded by the curves $y = \sqrt{32x}, y = 2x^2$, and $y = -4x + 6$ in the first quadrant. **Hint**.



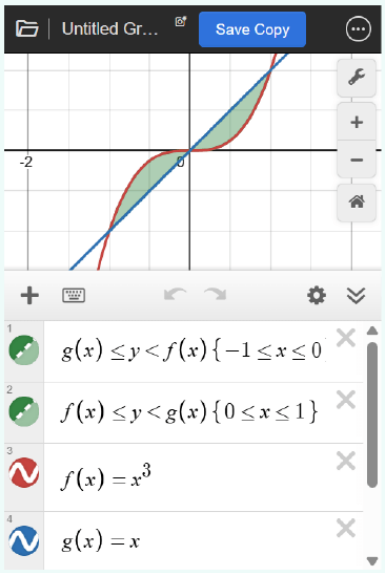
f. The other region bounded by the curves $y = \sqrt{32x}, y = 2x^2$, and $y = -4x + 6$ in the first quadrant. **Hint**.



g. The region bounded by the curves $x = 2y$ and $x = y^2 - 3$. **Hint**.



h. The region(s) bounded by the curves $y = x^3$ and $y = x$. **Hint**.



Answer available at the end of this section.

Answers

Activity 6.2.1: Remembering Riemann Sums.

f. Height: $f(x_k^*)$
Width: Δx

g. $A_k = f(x_k^*)\,\Delta x$

h. $A \approx \sum_{k=1}^{n} f(x_k^*)\,\Delta x$

i. $$A = \lim_{n\to\infty} \sum_{k=1}^{n} f(x_k^*)\,\Delta x = \int_{x=a}^{x=b} f(x)dx$$

Activity 6.2.2: Area Between Curves.

f. Height: $f(x_k^*) - g(x_k^*)$
Width: Δx

g. $A_k = (f(x_k^*) - g(x_k^*))\,\Delta x$

h. $A \approx \sum_{k=1}^{n} (f(x_k^*) - g(x_k^*))\,\Delta x$

i. $$A = \lim_{n\to\infty} \sum_{k=1}^{n} \left(f\left(x_k^*\right) - g\left(x_k^*\right)\right) \Delta x$$
$$= \int_{x=a}^{x=b} (f(x) - g(x))dx$$

Activity 6.2.3**: Trying for a Single Integral.**

i. Height: Δy
Width: $\sqrt[3]{y_k^*} - (y_k^* - 6)$

j. $A_k = \left(\sqrt[3]{y_k^*} - (y_k^* - 6)\right) \Delta y$

k. $A \approx \sum_{k=1}^{n} \left(\sqrt[3]{y_k^*} - (y_k^* - 6)\right) \Delta y$

l. $$A = \lim_{n\to\infty} \sum_{k=1}^{n} \left(\sqrt[3]{y_k^*} - (y_k^* - 6)\right) \Delta y$$
$$= \int_{y=0}^{y=8} (\sqrt[3]{y} - (y - 6))dy$$

m. $$\int_{y=0}^{y=8} (\sqrt[3]{y} - (y-6))dy = \int_{y=0}^{y=8} \left(y^{1/3} - y + 6\right) dy$$
$$= 28$$

Exercise 6.2.3

a. $\int_{x=0}^{x=2} (4x+1) - (x^2+1)\, dx = \frac{16}{3}$

b. $\int_{x=0}^{x=2} (4-x) - (x) dx = 4$

c. $\int_{x=0}^{x=4} (\sqrt{x}+2) - (x) dx = \frac{16}{3}$

d. This one has annoying limits of integration, and so evaluating the integral is more annoying than it is worth.
$$\int_{x=-\sqrt{\frac{\sqrt{5}-1}{2}}}^{x=\sqrt{\frac{\sqrt{5}-1}{2}}} \left(\frac{1}{x^2+1}\right) - \left(x^2\right) dx$$

Exercise 6.2.4

a. $$\int_{x=0}^{x=\frac{\pi}{4}} \sin(x)dx + \int_{x=\frac{\pi}{4}}^{x=\frac{\pi}{2}} \cos(x)dx$$

Or, alternatively:

$$\int_{y=0}^{y=\frac{\sqrt{2}}{2}} \cos^{-1}(y) - \sin^{-1}(y)dy$$

Admittedly, it is annoying to evaluate this in terms of y, even though it is satisfying to have only a single integral.
The area itself is $2 - \sqrt{2}$.

b. $$\int_{x=0}^{x=\sqrt{2}} xdx + \int_{x=\sqrt{2}}^{x=2} x - \left(x^2 - 2\right) dx$$

Or alternatively:

$$\int_{y=0}^{y=2} \sqrt{y+2} - ydy$$

In either case, the area is $\frac{10 - 4\sqrt{2}}{3}$

c. $$\int_{x=-\sqrt{2}}^{x=-1} 0 - \left(x^2 - 2\right) dx + \int_{x=-1}^{x=0} 0 - (x) dx$$

Or, alternatively:

$$\int_{y=-1}^{y=0} y - \sqrt{y+2} dy$$

In either case, the area is $\frac{4\sqrt{2}}{3} - \frac{7}{6}$

d. $$\int_{x=0}^{x=1} (3x) - \left(x^2\right) dx + \int_{x=1}^{x=\sqrt{2}} \left(4 - x^2\right) - \left(x^2\right) dx$$

Or, alternatively:

$$\int_{y=0}^{y=2} \sqrt{y} - \left(\frac{y}{3}\right) dy + \int_{y=2}^{y=3} \sqrt{4-y} - \left(\frac{y}{3}\right) dy$$

In either case, the area is $\frac{8\sqrt{2}}{3} - \frac{13}{6}$

This page titled 6.2: Area Between Curves is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Peter Keep.

6.3: Volumes of Solids of Revolution

Hopefully, by now we're feeling pretty comfortable with the use of a Riemann sum to create an integral formula. So far, these integral formulas have matched with our intuition somewhat. We probably can justify the integral formula for displacement of an object (Definition 6.1.3 Displacement of an Object) by thinking about the fact that position is an anti-derivative of velocity. We probably can convince ourselves about the integral formula for the area between curves (Definition 6.2.5 Area Between Curves) by thinking about subtracting areas, geometrically.

We're going to make a jump from a 2-dimensional measurement of size, area, to a 3-dimensional measurement of size, volume.

From Area To Volume

Here's the basic idea, in a broad overview: if we want to calculate a volume, then we are going to be working with a 3-dimensional solid. We'll use the slice-and-sum process:

1. Slice the object into uniformly thick slices along some axis.
2. For each slice, we'll approximate the volume. We can do this by thinking about the cross-sectional area. If we assume that the area is constant all the way through the slice (in the same way that we assumed earlier that the heights of our rectangles were constant), then we simply can multiply the cross-sectional area by the thickness to get the volume of each slice:
$$V_k = A\left(x_k^*\right)\Delta x$$
3. Approximate the total volume of the solid by adding the volumes of the slices together:
$$V \approx \sum_{k=1}^{n} A\left(x_k^*\right)\Delta x$$
4. Apply a limit, where the number of slices gets infinitely big (and the thickness of each slice gets infinitely small):
$$\begin{aligned} V &= \lim_{n\to\infty} \sum_{k=1}^{n} A\left(x_k^*\right)\Delta x \\ &= \int_{x=a}^{x=b} A(x)dx \end{aligned}$$
Here, a and b are the x-values that define the interval we're slicing along and $A(x)$ is a formula for the cross-sectional area of the object at x.

The biggest issue here is going to be thinking about that formula for area. In order for us to do that, we're going to think about a specific type of 3- dimensional solid, built in a systematic way so that we can find the cross-sectional areas easily.

Solids of Revolution

A solid of revolution is a strange type of solid: we're going to define it based on a 2-dimensional region (we'll use functions in a normal xy-plane) that we then imagine revolving around a straight line axis. Maybe we define some region in the upper half of the plane, but then revolve it around the x-axis. While we imagine this revolution, we want to think about the three dimensional solid that gets "traced" by the curve spinning around the axis.

Let's dive into an example to see. Let's visualize some function $f(x)$ defined (and continuous) on the interval $[a, b]$ and with $f(x) \geq 0$ on that interval. We'll see why this is useful, but for now, we're just thinking of some function.

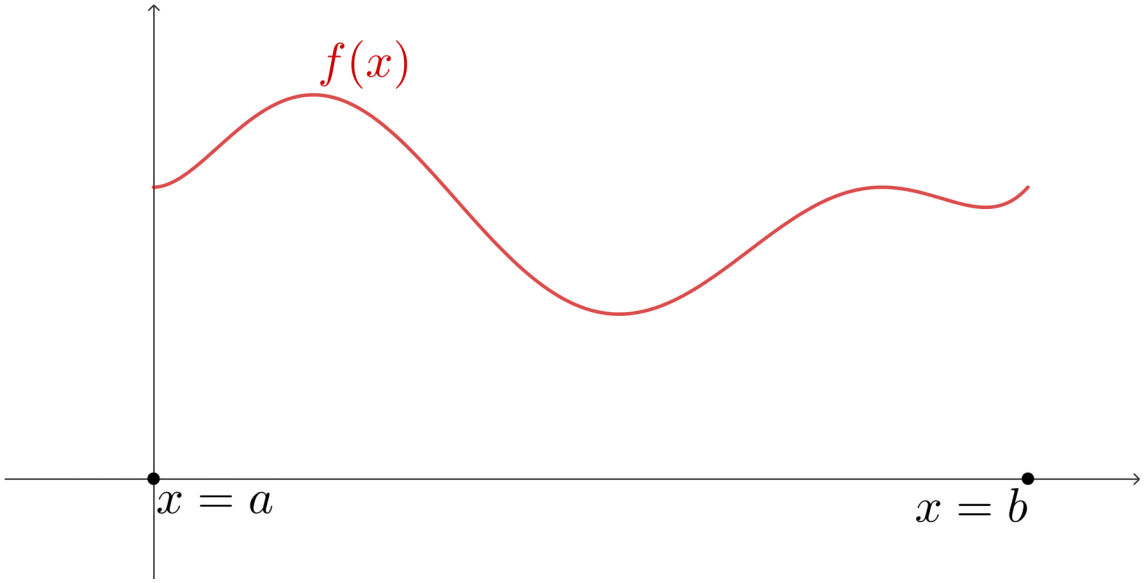


Figure 6.3.1

We're going to revolve this curve (and the region bounded between it and the x-axis) around the x-axis. This will create the following shape.

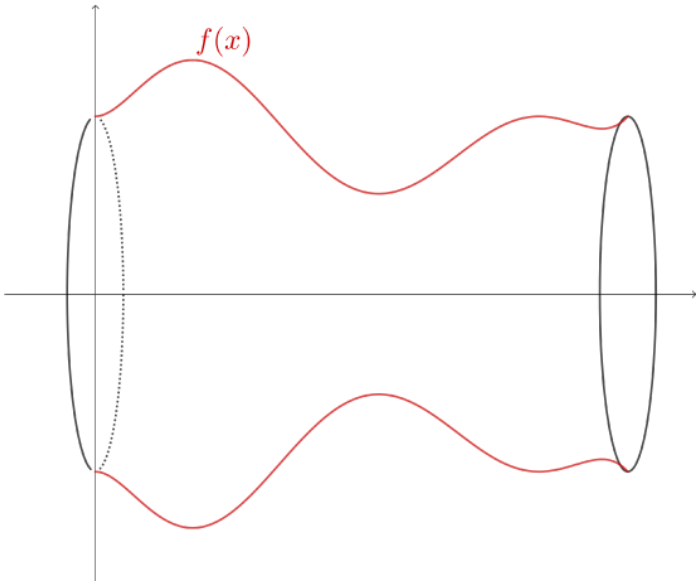


Figure 6.3.2

So our goal is to find the volume of this type of solid. The curve defining the edge of it can change, but the way that we create it will be systematic enough that we can build a formulaic integral expression for it.

As you would imagine, we'll start with a rectangle.

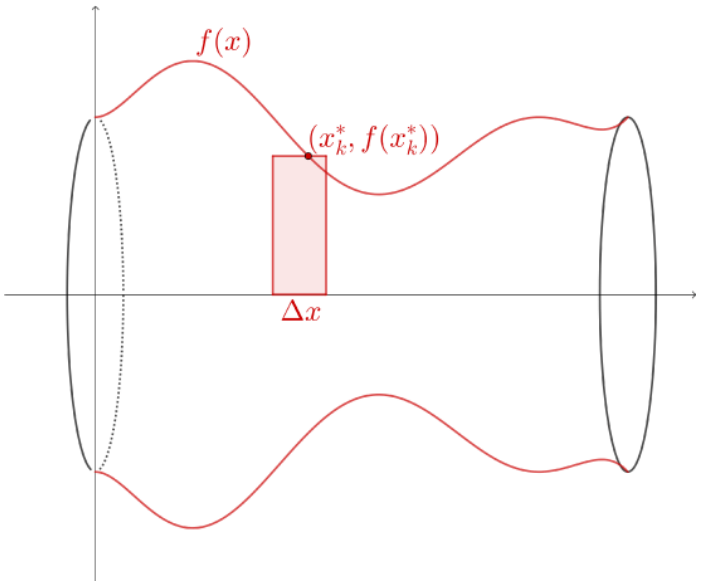


Figure 6.3.3

This rectangle will represent a single, generic slice. We really want to imagine a slice of the 3-dimensional solid, though, and so we will revolve this rectangle around the x-axis. This will create a slice of our solid of revolution. From there, we can think about the volume of this generic kth slice, and fall into the rhythm of our slice-and-sum process.

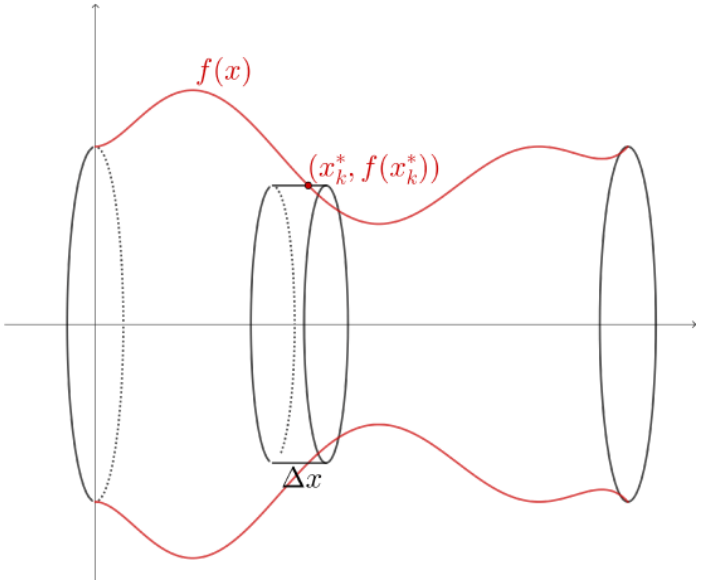


Figure 6.3.4

We want to find the volume of this specific slice. To do this, we can remove this stubby cylinder from the solid and think about it directly. We can see the thickness of the slice is represented by Δx, and so we need to think about the cross-sectional area of the "face" of this slice.

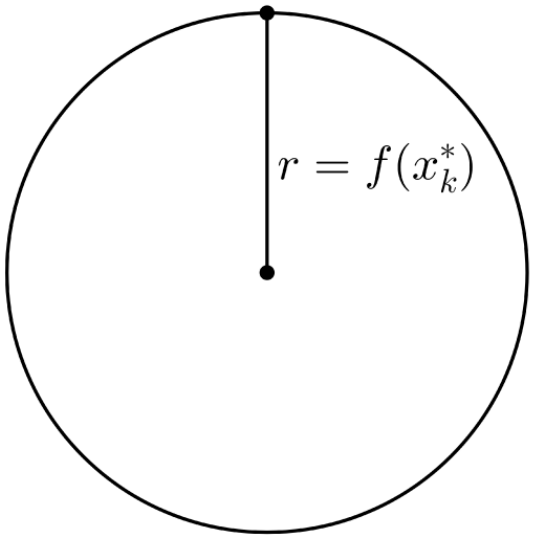


Figure 6.3.5

This is something we easily can find the area of! We know the formula for the area of a circle: $A = \pi r^2$. We'll notice that the radius of this circle is the distance from the center of our slice to the outer edge: this is the height of the rectangle in Figure 6.3.3. So we can use $r = f(x_k^*)$, giving us the cross-sectional area of the kth slice:

$$A(x_k^*) = \pi (f(x_k^*))^2$$

Now, we can drop into our slice and sum process:

1. The volume of the kth slice is:

$$\begin{aligned} V_k &= A(x_k^*)\,\Delta x \\ &= \pi (f(x_k^*))^2\,\Delta x \end{aligned}$$

2. We can approximate the volume by adding the slices:

$$V \approx \sum_{k=1}^{n} \pi (f(x_k^*))^2\,\Delta x$$

Sometimes this can be hard to visualize. We're approximating the solid in Figure 6.3.2 by thinking about a bunch of these circular disks stacked next to each other.

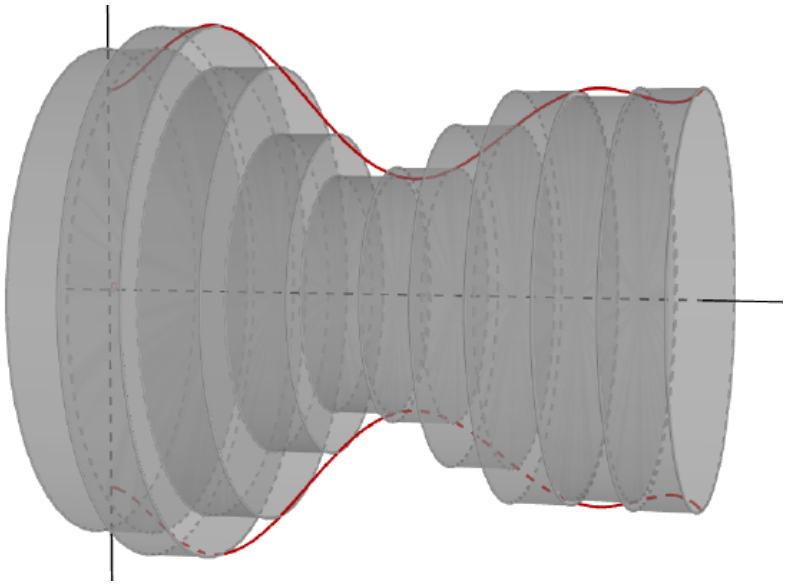


Figure 6.3.6

3. We can apply a limit to evaluate the actual volume of the solid and construct a definite integral.

$$V = \lim_{n\to\infty} \sum_{k=1}^{n} \pi \left(f\left(x_k^*\right)\right)^2 \Delta x$$
$$= \int_{x=a}^{x=b} \pi (f(x))^2 dx$$

This is great! We'll call this volume integral the Disk Method, since each cross section is a circular disk.

What happens if we add a second curve defining a lower boundary to the region, like we did in Section 6.2 Area Between Curves for areas?

Activity 6.3.1: Carving out a Hole in the Center.

We're going to look at the same solid as in Figure 6.3.2. But this time, when we define the 2-dimensional region that we're going to revolve around the x-axis, we're going to add a lower boundary function, $g(x)$.

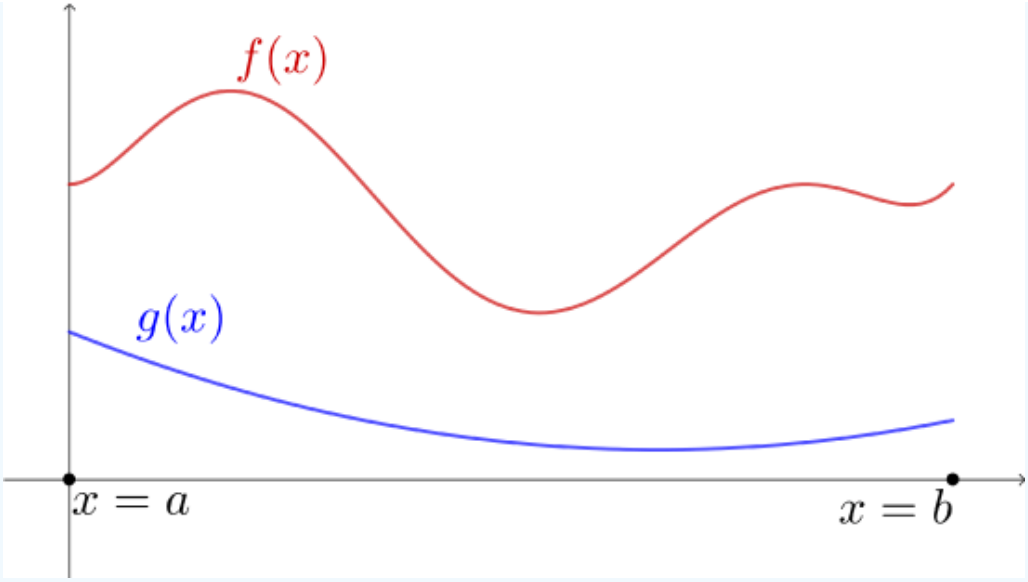


Figure 6.3.7

When we revolve this region around the x-axis, we get the following 3-dimensional solid.

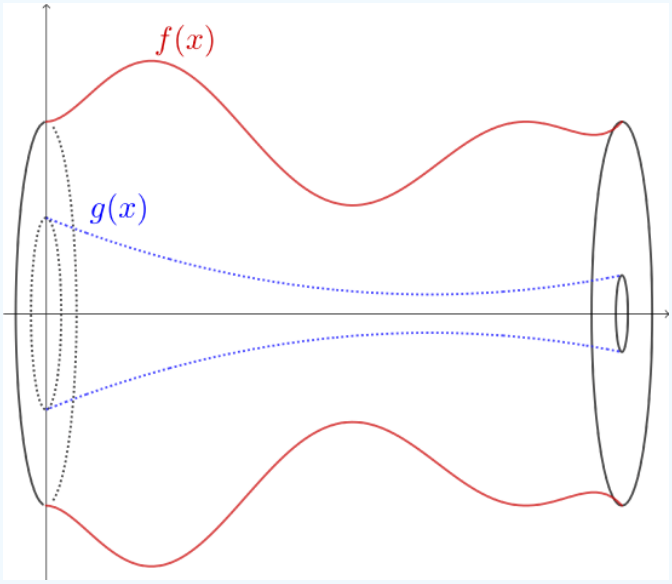


Figure 6.3.8

a. How is a single generic slice on this solid different than the one in Figure 6.3.2?
Hint 1.

Here is the rectangle that will define that slice!

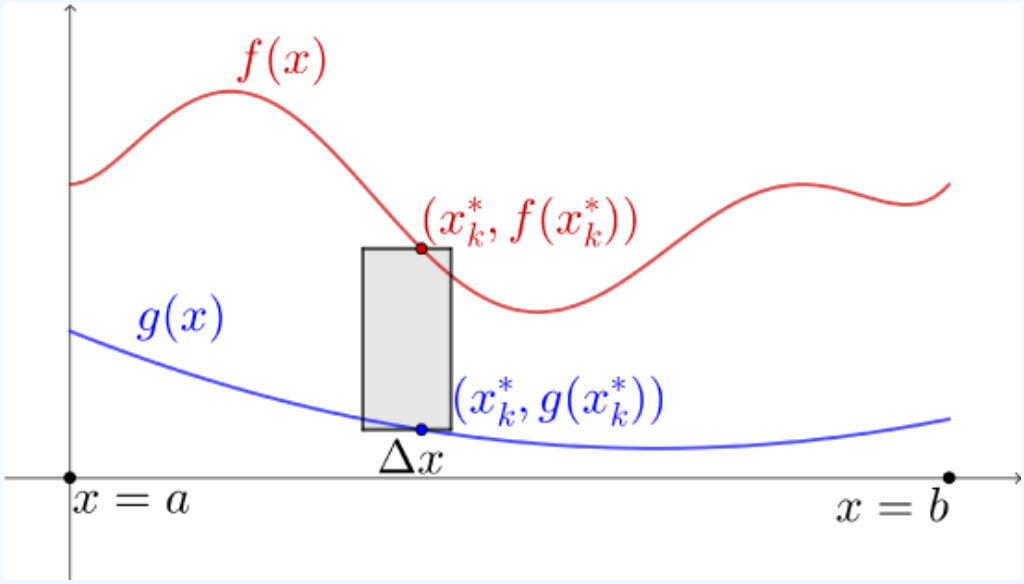


Figure 6.3.9

Hint 2.

Here is the slice formed when the rectangle revolves around the axis!

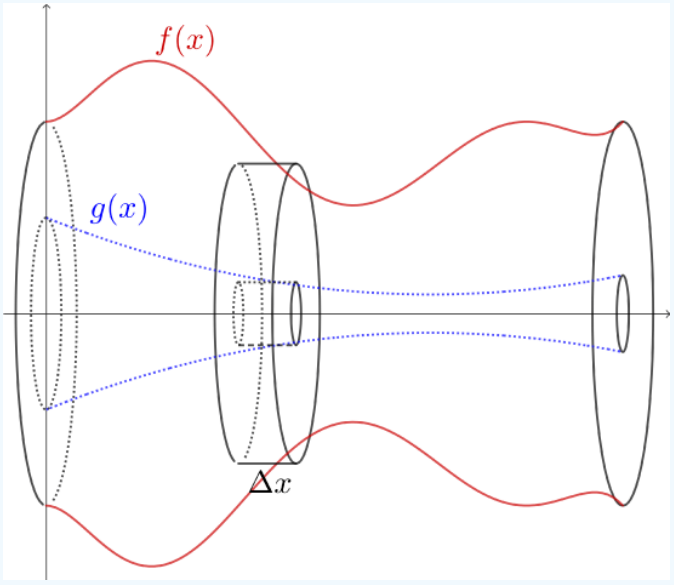


Figure 6.3.10

b. Find a formula for the area of the face of the cross-sectional slice. **Hint**. Here's a picture of the face of the slice!

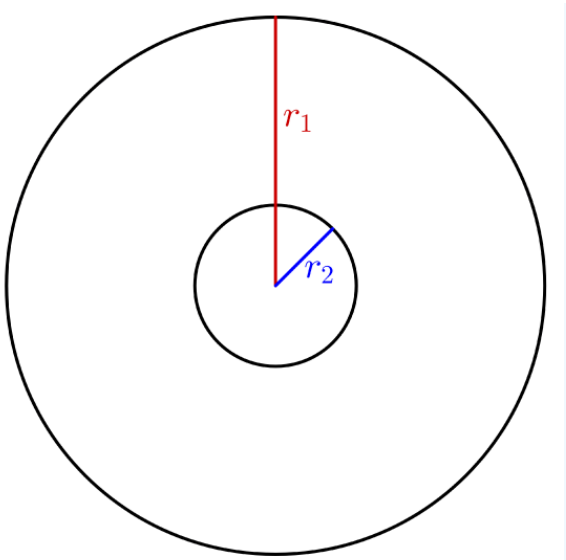


Figure 6.3.11

c. Use the slice-and-sum process to create an integral expression representing the volume of this solid.

Answer available at the end of this section.

Definition 6.3.12: Volume by Disks/Washers.

If f and g are continuous functions with $f(x) \geq g(x) \geq 0$ on the interval $[a, b]$, then the volume of the solid formed by revolving the region bounded between the curves $y = f(x)$ and $y = g(x)$ from $x = a$ to $x = b$ around the x-axis is:

$$V = \pi \int_{x=a}^{x=b} \left((f(x))^2 - (g(x))^2 \right) dx$$

This is called the **Washer Method**. Note that if $g(x) = 0$, then the resulting volume is:

$$V = \pi \int_{x=a}^{x=b} (f(x))^2 dx$$

This is called the **Disk Method**.

We'll walk through two examples where we construct these integral expressions before pretending to be too comfortable. Let's start with something similar to what we've just done.

Activity 6.3.2: Volumes by Disks/Washers.

Consider the region bounded between the curves $y = 4 + 2x - x^2$ and $y = \dfrac{4}{x+1}$. We will create a 3-dimensional solid by revolving this region around the x-axis. Figure 6.3.10

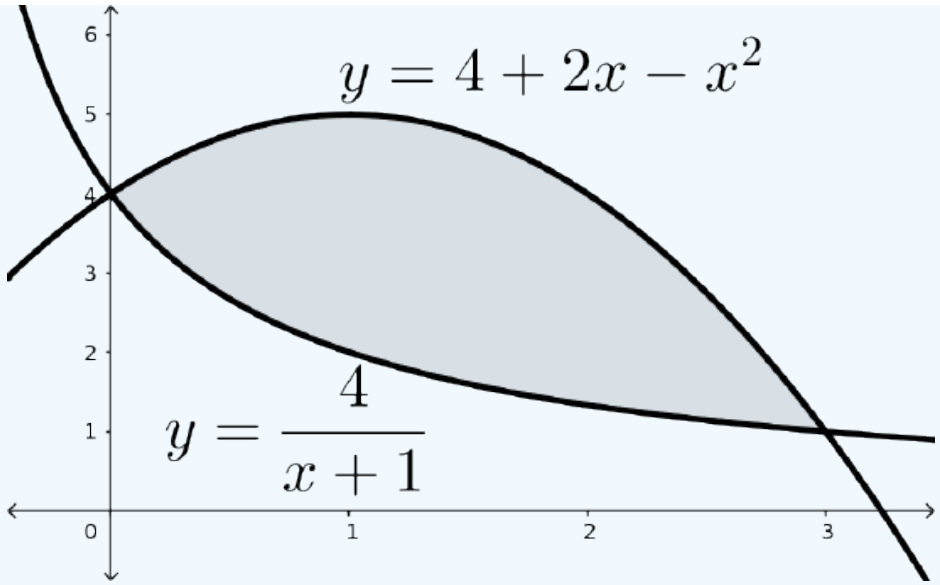


Figure 6.3.12

a. Visualize the solid you'll create when you revolve this region around the x-axis.
b. Draw a single rectangle in your region, standing perpendicular to the x-axis.
c. Let's use this rectangle to visualize the kth slice of this 3- dimensional solid. What does the "face" of it look like? **Hint**. It should be a circle within a circle like in Figure 6.3.11, but can you be more detailed with labeling the radii?
d. Find the area of the face of the kth slice. **Hint**. Note that this is a 2-dimensional shape, and we're just finding the area of it.
e. Set up the integral representing the volume of the solid. **Hint**. The sum is going to be adding up all of the areas multiplied by δx. What will that look like in the integral?
f. Can you describe how you would antidifferentiate and evaluate this integral? **Hint**. What happens when you square these functions? What kinds of strategies will you use for the types of functions you're left with?

Answer available at the end of this section.

Ok, so when we're creating these integrals, we really are focussing on using the rectangle we drew to show us which functions serve as the large radius compared to the small radius. In the next example, we'll see another key thing to notice from a single rectangle.

Activity 6.3.3: Another Volume.

Now let's consider another region: this time, the one bounded between the curves $y = x$ and $y = 3\sqrt{x}$. We will, again, create a 3-dimensional solid by revolving this region around the y-axis.

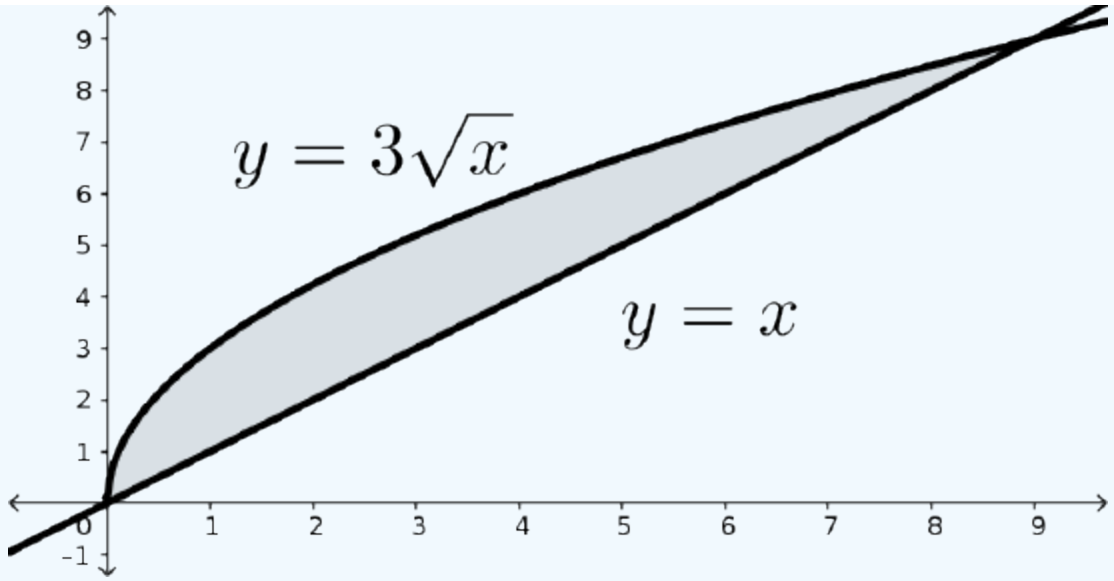


Figure 6.3.13

a. Visualize the solid you'll create when you revolve this region around the y-axis.
b. Draw a single rectangle in your region standing perpendicular to the y-axis. **Hint**. Notice that your rectangle is sitting on its side now! This will change some things for us in a familiar way!
c. Let's use this rectangle to visualize the kth slice of this 3- dimensional solid. What does the "face" of it look like? **Hint**. It should be a circle within a circle like in Figure 6.3.11, but can you be more detailed with labeling the radii?
d. Find the area of the face of the kth slice. **Hint**. Note that this is a 2-dimensional shape, and we're just finding the area of it. You'll also notice that the radii are measuring a horizontal distance in terms of a differing height, so you'll want to express these as functions of y.
e. Set up the integral representing the volume of this solid.

Answer available at the end of this section.

Notice that the rectangle was the clue that we were going to be using Δy when we calculated volumes. This ended up being the reason that we integrated with regard to y, since the $\Delta y \to dy$ in the integral.

A single rectangle, carefully drawn, can give us a large amount of information as we try to juggle these volumes!

Reorienting our Rectangles

We saw in Activity 6.3.3 that thinking about the single rectangle we drew was helpful. We'll see that again in this next formula that we build.

Notice that, in all of the previous work we've done, we've drawn our rectangle so that it is standing perpendicular to the axis of revolution. This is the kind of rectangle that, when we revolve it, traces out the "washer" shape!

So what happens when we change the orientation of our rectangle? What happens when we draw a rectangle that is parallel to the axis of revolution? Let's consider the same region as before (the one we visualized in Figure 6.3.7) with the same rectangle as before (the one we visualized in Figure ??), but we'll revolve around the y-axis.

When we revolve this region around the y-axis, we end up with the following solid.

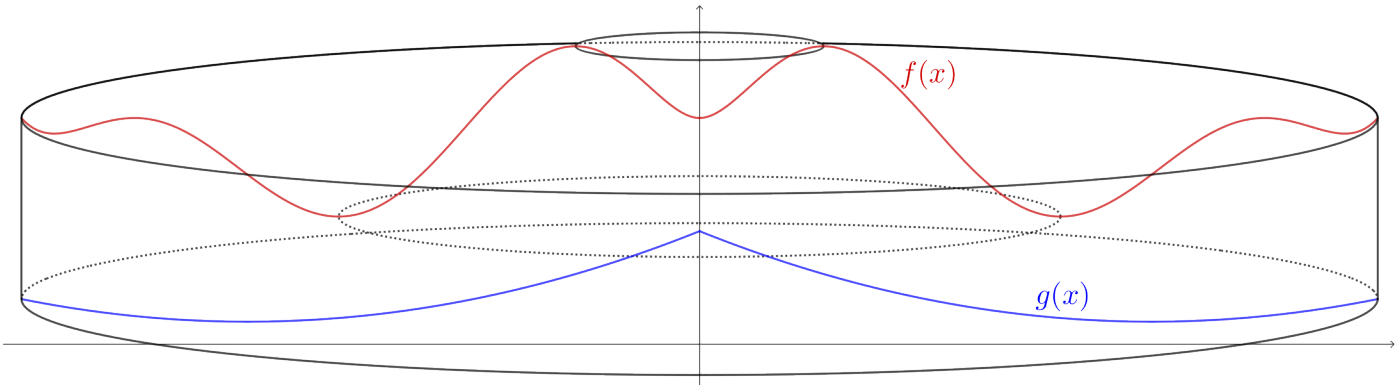


Figure 6.3.14

We want to focus on the single rectangle and the shape that it forms when we revolve it around the y-axis. From there, we can fall into our slice and sum process by thinking about how we might calculate the volume of this single sliced piece and then adding them up.

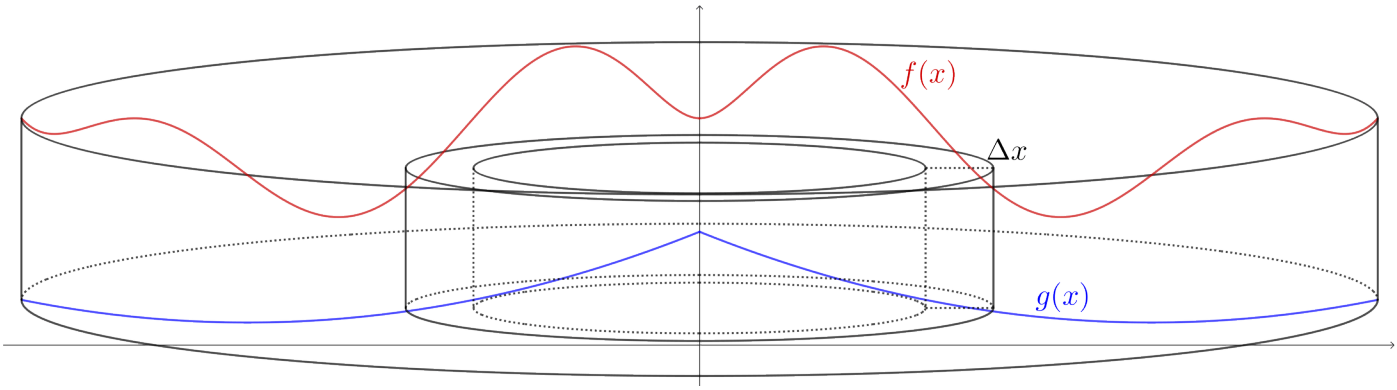


Figure 6.3.15

For this rectangle, we can notice that when we revolve it around the y-axis, we create a hollow cylinder. We'll focus more specifically on this cylinder.

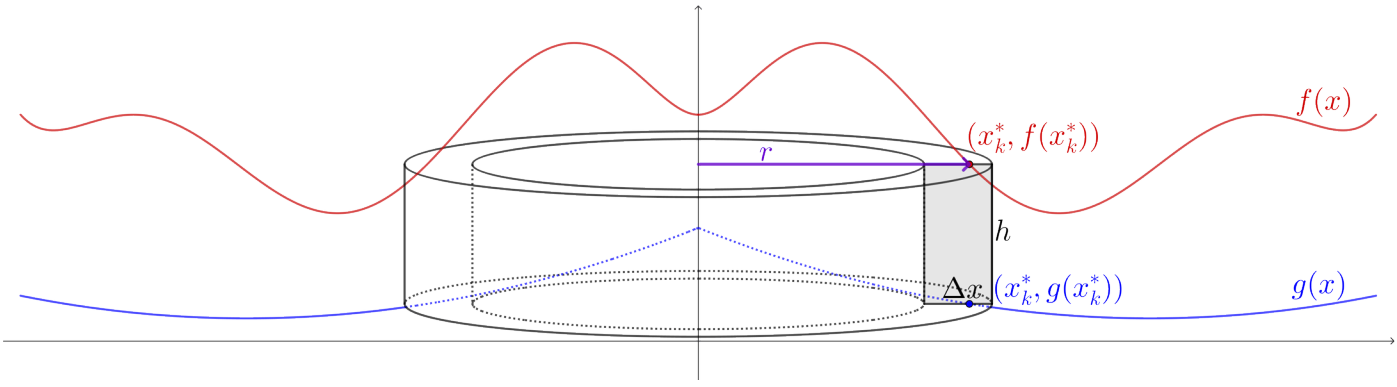


Figure 6.3.16

Let's focus more on the cylinder. We'll need to find the volume of this cylinder. We can think of this volume as really the surface area of the cylinder multiplied by the thickness. Another way to visualize it is to think about cutting the cylinder open, and unfurling it to create a rectangular solid with some thickness.

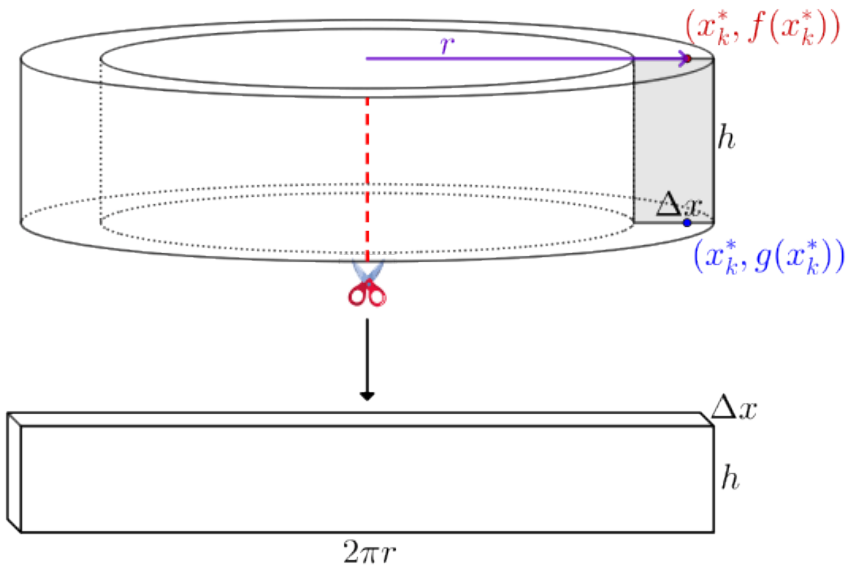


Figure 6.3.17

So we can see that to find V_k, we're going to multiply $A\left(x_k^*\right)$ and Δx again, where $A\left(x_k^*\right)$ is the area of the cross-sectional "face." In this case, we can see how we'll construct this from the unfurled cylinder.

$$\begin{aligned} V_k &= 2\pi r \Delta x \\ &= 2\pi\left(x_k^*\right)\left(f\left(x_k^*\right) - g\left(x_k^*\right)\right)\Delta x \\ V &\approx \sum_{k=1}^{n} 2\pi\left(x_k^*\right)\left(f\left(x_k^*\right) - g\left(x_k^*\right)\right)\Delta x \\ V &= \lim_{n\to\infty} \sum_{k=1}^{n} 2\pi\left(x_k^*\right)\left(f\left(x_k^*\right) - g\left(x_k^*\right)\right)\Delta x \\ &= \int_{x=a}^{x=b} 2\pi x(f(x) - g(x))dx \end{aligned}$$

Definition 6.3.16: Volume by Shells.

If $f(x)$ and $g(x)$ are continuous functions with $f(x) \geq g(x)$ on the interval $[a, b]$ (with $a \geq 0$), then the volume of the solid formed when the region bounded between the curves $y = f(x)$ and $y = g(x)$ from $x = a$ to $x = b$ is revolved around the y-axis is

$$V = 2\pi \int_{x=a}^{x=b} x(f(x) - g(x))dx$$

We can apply this formula in a familiar example, and also practice changing variables.

Activity 6.3.4: Volume by Shells.

Let's consider the region bounded by the curves $y = x^3$ and $y - x + 6$ as well as the line $y = 0$. You might remember this region from Activity 6.2.3. This time, we'll create a 3-dimensional solid by revolving the region around the x-axis

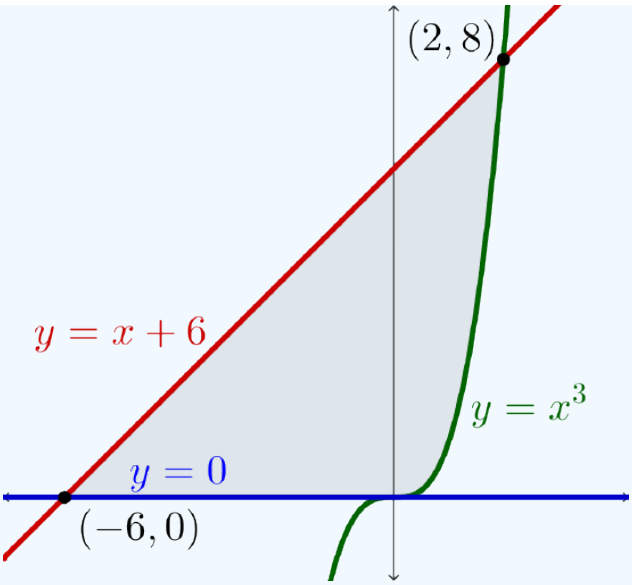
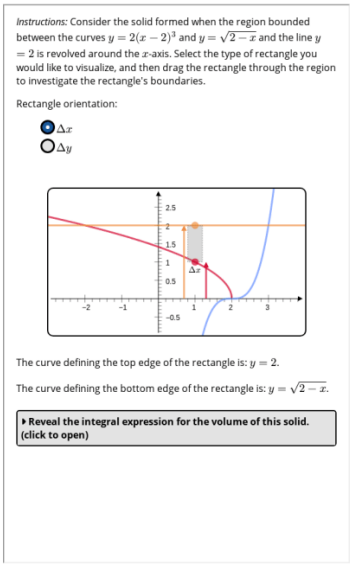


Figure 6.3.18

a. Sketch one or two rectangles that are perpendicular to the x-axis. Then set up an integral expression to find the volume of the solid using them. **Hint**. Note that in this context, we're actually using disks and washers. Also note that the bottom of the rectangles are bounded by $y = 0$ from $x = -6$ to $x = 0$ and then switches to being bounded by $y = x^3$ from $x = 0$ to $x = 2$.
b. Now, draw a single rectangle in the region that is parallel to the axis of revolution. Use this rectangle to visualize the k th slice of this 3-dimensional solid. What does that single rectangle create when it is revolved around the x-axis? **Hint**. This won't create a disk or washer! We'll have to change variables, and try to see how we can create a shell.
c. Set up the integral expression representing the volume of the solid.
d. Confirm that your volumes are the same, no matter your approach to setting it up.

Answer available at the end of this section.

To finish things up, let's look at another interactive graph (similar to how we ended Section 6.2 Area Between Curves) that can help show the differences between finding volume with regard to x (using Δx in our rectangles and dx in our integrals) and finding volume with regard to y (using Δy in our rectangles and dy in our integrals), and how this choice changes our method from washers to shells depending on the axis of revolution.




Standalone
Embed

Practice Problems

Exercise 6.3.1

We say that the volume of a solid can be thought of as $\int_{x=a}^{x=b} A(x)dx$ where $A(x)$ is a function describing the cross-sectional area of our solid at an x-value between $x = a$ and $x = b$. Explain how this integral formula gets built, referencing the slice-and-sum (Riemann sum) method.

Exercise 6.3.2

Explain the differences and similarities between the disk and washer methods for finding volumes of solids of revolution.

Exercise 6.3.3

When do we integrate with regard to x (using a dx in our integral and writing our functions with x-value inputs) and when do we integrate with regard to y (using a dy in our integral and writing our functions with y-value inputs) when we're finding volumes using disks and washers? How do we know?

Exercise 6.3.4

For each of the solids described below, set up an integral using the disk/ washer method that describes the volume of the solid. It will be helpful to visualize the region, a rectangle on that region, as well as the rectangle revolved around the axis of revolution.

a. The region bounded by the curve $y = 2x$ and the lines $y = 0$ and $x = 3$, revolved around the x-axis.
b. The region bounded by the curve $y = e^{-2x}$ and the x-axis between $x = 0$ and $x = \ln(2)$, revolved around the x-axis.
c. The region bounded by the curves $y = \ln(x)$ and $y = \sqrt{x}$ between $y = 0$ and $y = 1$, revolved around the y-axis.
d. The region bounded by the curves $y = 2x + 1$ and $y = x$ between $x = 0$ and $x = 3$, revolved around the x-axis.
e. The region bounded by the curve $y = x^3$, the x-axis, and the line $x = 2$, revolved around the y-axis.
f. The region bounded by the curve $y = x^3$ and the y-axis between $y = 0$ and $y = 2$, revolved around the y-axis.

Exercise 6.3.5

Explain where the pieces of the shell formula come from. How is this different than using disks/washers?

Exercise 6.3.6

Say we're revolving a region around the x-axis to create a solid. Using the disk/washer method, we will integrate with respect to x. Using the shell method, we integrate with respect to y. Explain the difference, and why this difference occurs.

Exercise 6.3.7

For each of the solids described below, set up an integral using the shell method that describes the volume of the solid. It will be helpful to visualize the region, a rectangle on that region, as well as the rectangle revolved around the axis of revolution.

a. The region bounded by the curve $y = 3x$ and the lines $x = 0$ and $y = 5$, revolved around the y-axis.
b. The region bounded by the curve $y = \sqrt{x}$ and the x-axis between $x = 0$ and $x = 9$, revolved around the x-axis.
c. The region bounded by the curves $y = 2 - x^2$ and $y = x$ and the line $x = 0$ revolved around the y-axis.
d. The region bounded by the curves $y = \sin(x^2) + 2$ and $y = x$ from $x = 0$ to $x = 1$, revolved around the y-axis.

e. The region bounded by the curves $y = x^2 - 6x + 10$ and $y = 2 + 4x - x^2$ revolved around the y-axis.
f. The region bounded by the curves $y = \sqrt{2x}$ and $y = 4 - x$ and the x-axis between $x = 0$ and $x = 4$, revolved around the x-axis.

Exercise 6.3.8

Pick at least 2 integrals from Exercise 4 to rewrite using shells instead. What about those regions did you look for to choose which ones to rewrite and which ones to not?

Exercise 6.3.9

Pick at least 2 integrals from Exercise 7 to rewrite using disks/washers instead. What about those regions did you look for to choose which ones to rewrite and which ones to not?

Exercise 6.3.10

For each of the following solids, set up an integral expression using either the disk/washer method or the shell method. You don't need to evaluate them, but you should do some careful thinking about how you set these up, especially as you choose between methods and what variable you are integrating with.

a. The region bounded by the curves $y = x^2 + 1$ and $y = x^3 + 1$ in the first quadrant, revolved around the x-axis.
b. The region bounded by the curves $y = x^2 + 1$ and $y = x^3 + 1$ in the first quadrant, revolved around the y-axis.
c. The region bounded by the curves $y = \dfrac{1}{x}$ and $y = 1 - (x-1)^2$ in the first quadrant, revolved around the x-axis.

Answers

Activity 6.3.1: Carving out a Hole in the Center.

2. $A = \pi r_1^2 - \pi r_2^2$
$A(x_k^*) = \pi (f(x_k^*))^2 - \pi (g(x_k^*))^2$
$V_k = \left(\pi (f(x_k^*))^2 - \pi (g(x_k^*))\right) \Delta x$
$= \pi \left((f(x_k^*))^2 - (g(x_k^*))^2\right) \Delta x$
3. $V \approx \sum_{k=1}^{n} \pi \left((f(x_k^*))^2 - (g(x_k^*))^2\right) \Delta x$
$V = \lim_{n\to\infty} \sum_{k=1}^{n} \pi \left((f(x_k^*))^2 - (g(x_k^*))^2\right) \Delta x$
$= \int_{x=a}^{x=b} \pi \left((f(x))^2 - (g(x))^2\right) dx$

Activity 6.3.2: Volumes by Disks/Washers.

d. $A(x_k^*) = \pi \left(\left(4 + 2x_k^* - (x_k^*)^2\right)^2 - \left(\dfrac{4}{x_k^* + 1}\right)^2\right)$
e. $V = \lim_{n\to\infty} \sum_{k=1}^{n} \pi \left(\left(4 + 2x_k^* - (x_k^*)^2\right)^2 - \left(\dfrac{4}{x_k^* + 1}\right)^2\right) \Delta x$
$= \pi \int_{x=0}^{x=3} \left(\left(4 + 2x - x^2\right)^2 - \left(\dfrac{4}{x+1}\right)^2\right) dx$
f. The first function, the quadratic, will be annoying to square. We'll end up with some big degree 4 polynomial, though, and antidifferentiating will be easy, since we can use the Power Rule.

The second function squared will give us $\dfrac{4}{(x+1)^2}$. We can use a u-substitution here with $u = x + 1$. Then, we have a negative exponent and we can use the Power Rule!

Activity 6.3.3: Another Volume.

d. The outer radius comes from the function $y = x$, but we'll invert it to be $x = y$.

The inner radius comes from the function $y = 3\sqrt{x}$, but we'll invert it be written as $x = \left(\dfrac{y}{3}\right)^2$ or $x = \dfrac{y^2}{9}$.

$$A(y_k^*) = \pi \left((y_k^*)^2 - \left(\frac{(y_k^*)^2}{9}\right)^2\right)$$

e. $$V = \lim_{n\to\infty} \sum_{k=1}^{n} \pi \left((y_k^*)^2 - \left(\frac{(y_k^*)^2}{9} \right)^2 \right) \Delta y$$
$$= \pi \int_{y=0}^{y=9} y^2 - \frac{y^4}{81} dy$$

Activity 6.3.4**: Volume by Shells.**

a. $V = \pi \int_{x=-6}^{x=0} (x+6)^2 dx + \pi \int_{x=0}^{x=2} (x+6)^2 - \left(x^3\right)^2 dx$

b.

c. $V = 2\pi \int_{y=0}^{y=8} y(\sqrt[3]{y} - (y-6)) dy$

This page titled 6.3: Volumes of Solids of Revolution is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Peter Keep.

6.4: More Volumes- Shifting the Axis of Revolution

We have introduced some methods for creating and calculating the volume of different 3-dimensional solids of revolution.

What Changes?

Let's first consider a volume created using disks or washers.

Activity 6.4.1: What Changes (in the Washer Method) with a New Axis?

Let's revisit Activity 6.3.2 Volumes by Disks/Washers, and ask some more follow-up questions. First, we'll tinker with the solid we created: instead of revolving around the x-axis, let's revolve the same solid around the horizontal line $y = -3$.

a. What changes, if any, do you have to make to the rectangle you drew in Activity 6.3.2?
b. What changes, if any, do you have to make to the area of the "face" kth washer?
c. What changes, if any, do you have to make to the eventual volume integral for this solid?

Now, let's consider a volume created using shells.

Activity 6.4.2: What Changes (in the Shell Method) with a New Axis?

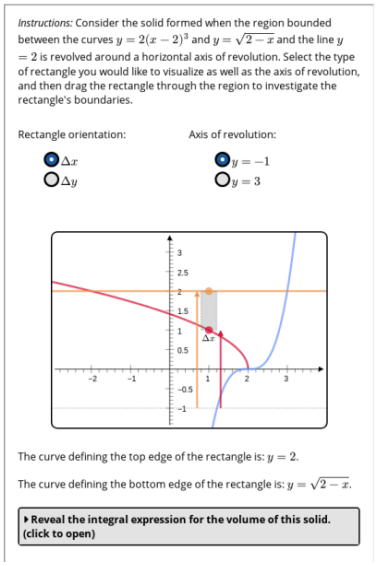
Let's revisit Activity 6.3.4 Volume by Shells, and ask some more followup questions about the shell method. Again, we'll tinker with the solid we created: instead of revolving around the x-axis, let's revolve the same solid around the horizontal line $y = 9$.

a. What changes, if any, do you have to make to the rectangle you drew in Activity 6.3.4?
b. What changes, if any, do you have to make to the area of the rectangle formed by "unrolling" up kth cylinder?
c. What changes, if any, do you have to make to the eventual volume integral for this solid?

In both of these cases, we can notice that the only changes we make are to the radii: we just need to remeasure the distance from axis of revolution to either the ends of the rectangle (in the washer method) or the side of the rectangle (in the shell method).

Formalizing These Changes in the Washers and Shells

We can look at yet another interactive graph (similar to how we ended Section 6.2 Area Between Curves and Section 6.3 Volumes of Solids of Revolution). This time, we'll think about how our axis of revolution as well as our choice of rectangle orientation impacts how we construct the washers or shells.



Standalone
Embed

Notice that in each case, we're remeasuring the radius! Whether we're measuring the radii of a washer by thinking about how far the function outputs are away from the axis of revolution, or if we're measuring the radius of a shell by thinking about how far the input variable is away from the axis of revolution, we need to rethink this and do some subtraction.

Activity 6.4.3: More Shifted Axes.

We're going to spend some time constructing several different volume integrals in this activity. We'll consider the same region each time, but make changes to the axis of revolution. For each, we'll want to think about what kind of method we're using (disks/washers or shells) and how the different axis of revolution gets implemented into our volume integral formulas.

Let's consider the region bounded by the curves $y = \cos(x) + 3$ and $y = \dfrac{x}{2}$ between $x = 0$ and $x = 2\pi$.

a. Let's start with revolving this around the x-axis and thinking about the solid formed. While you set up your volume integral, think carefully about which method you'll be using (disks/washers or shells) as well as which variable you are integrating with regard to (x or y). **Hint**. Note that in this region, we definitely want to use rectangles that stand up vertically. That means that they'll have a very small width, Δx, and sit perpendicular to the axis of revolution.
b. Now, let's create a different solid by revolving this region around the y-axis. Set up a volume integral, and continue to think carefully about which method you'll be using (disks/washers or shells) as well as which variable you are integrating with regard to (x or y). **Hint**. We still will use the same tall rectangle with a small Δx side length, but this time it will be parallel to our axis of revolution.
c. We'll start shifting our axis of revolution now. We'll revolve the same region around the horizontal line $y = -1$ to create a solid. Set up an integral expression to calculate the volume. **Hint 1**. Note that we're still using the same rectangle (perpendicular to this horizontal axis), and so still integrating with regard to x, and using the washer method. **Hint 2**. Since in the washer method our function outputs represent the radii, we need to remeasure the distance from our curves to the axis of revolution to find each circle's radius in the washer formula. For a y-value on each curve, how do we find the vertical distance down to the line $y = -1$?
d. Now, revolve the region around the line $y = 5$ to create a solid of revolution, and write down the integral representing the volume. **Hint**. Note, now, that the y-value of the axis of revolution is larger than all of the y-values on the curves, meaning that to measure the distance from the axis of revolution to the curves, we might measure them in the opposite direction. Also, which curve is further away from the axis of revolution, representing the larger/outer radius?
e. Let's change things up. Revolve the region around the vertical line $x = -1$ to create a new solid. Set up an integral representing the volume of that solid. **Hint 1**. Note that the same rectangle that we used before is standing parallel to our axis of revolution. We're going to change methodology, and use the shell method! **Hint 2**. Normally we use the input variable (x in this case) to measure the radius from the rectangles at different x-value to the axis of revolution, the y-axis. Now, though, we're not looking at the distance from x-values to $x = 0$. We're looking to find the radius, the distance from x-values in this region to $x = -1$.

f. We'll do one more solid. Let's revolve this region around the line $x = 7$. Set up an integral representing the volume. **Hint**. Note that this time, the axis of revolution's x-value is larger than all of the x-values in our region. So when we subtract to measure the radius, we need to subtract from $x = 7$ down to the varying x-values in the region.

Practice Problems

Exercise 6.4.1

Consider the integral formula for computing volumes of a solid of revolution using the disk/washer method. What part of this integral formula represents the radius/radii of any circle(s)? Why is the radius represented using the function output from the curve(s) defining the region?

Exercise 6.4.2

Consider the integral formula for computing volumes of a solid of revolution using the shell method. What part of this integral formula represents the radius/radii of any circle(s)? Why is the radius not represented using the function output from the curve(s) defining the region?

Exercise 6.4.3

For each of the solids described below, set up an integral expression using disks/washers representing the volume of the solid.

a. The region bounded by the curve $y = 1 - \sqrt{x}$ in the first quadrant, revolved around $x = 2$.
b. The region bounded by the curve $y = 1 - \sqrt{x}$ in the first quadrant, revolved around $x = -1$.
c. The region bounded by the curve $y = 1 - \sqrt{x}$ in the first quadrant, revolved around $y = -2$.
d. The region bounded by the curve $y = 1 - \sqrt{x}$ in the first quadrant, revolved around $y = 3$.

Exercise 6.4.4

For each of the solids described below, set up an integral expression using shells representing the volume of the solid.

a. The region bounded by the curves $y = \sqrt{x}$ and $y = x$ in the first quadrant, revolved around the line $x = 2$.
b. The region bounded by the curves $y = \sqrt{x}$ and $y = x$ in the first quadrant, revolved around the line $x = -1$.
c. The region bounded by the curves $y = \sqrt{x}$ and $y = x$ in the first quadrant, revolved around the line $y = -2$.
d. The region bounded by the curves $y = \sqrt{x}$ and $y = x$ in the first quadrant, revolved around the line $y = 3$.

This page titled 6.4: More Volumes- Shifting the Axis of Revolution is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Peter Keep.

6.5: Arc Length and Surface Area

We're going to continue to think about different applications of definite integrals: what they can measure and how we can construct these integral formulas. In this section, we're going to add two more formulas for two more measurements. Before we get far into this discussion, we want to center the important parts of our discussion.

Sure, it is worth noting that, in this section, we'll add a 1-dimensional measurement of size to our list of things an integral can measure. We have talked about a 2-dimensional measure of size (area) and a 3-dimensional measure of size (volume), but we'll add length to the list now! We'll also add a 2-dimensional extension of perimeter to the list when we talk about surface area. That's cool!

But, more importantly, we're going to see how we can construct an integral formula from a Riemann sum, and we're going to get some experience constructing a Riemann sum to measure the thing we care about. In our study of integrals, it might not actually be that important to know how to calculate the specific kinds of volumes or lengths that we're talking about. But we can get some experience with using some formulas from a pretty comfortable field (geometry) to get some experience with the slice-and-sum process. And this process is a useful one to know! We want to see that a definite integral is more than just an area under a curve, and we want to be able to look at an integral formula for some measurement or calculation and see some of the parts of that formula that could be familiar.

Anyways, let's calculate some arc lengths.

Integrals for Evaluating the Length of a Curve

When we talk about arc length, we might think of the length of some portion of a circle. Here, we'll use it to refer to the length of some more general curve. We'll graph a function and think about how long the curve of the graph is from some point to another point.

Activity 6.5.1: Measuring Distance.

a. Consider the following right-triangle with the normal names of side lengths.

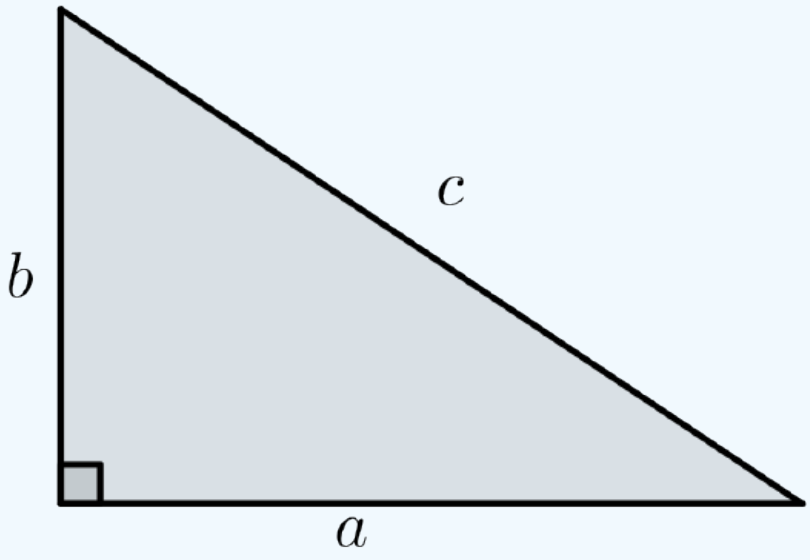


Figure 6.5.1

How do we use the Pythagorean Theorem to find the length of c?

b. Consider the two points (x_1, y_1) and (x_2, y_2) below.

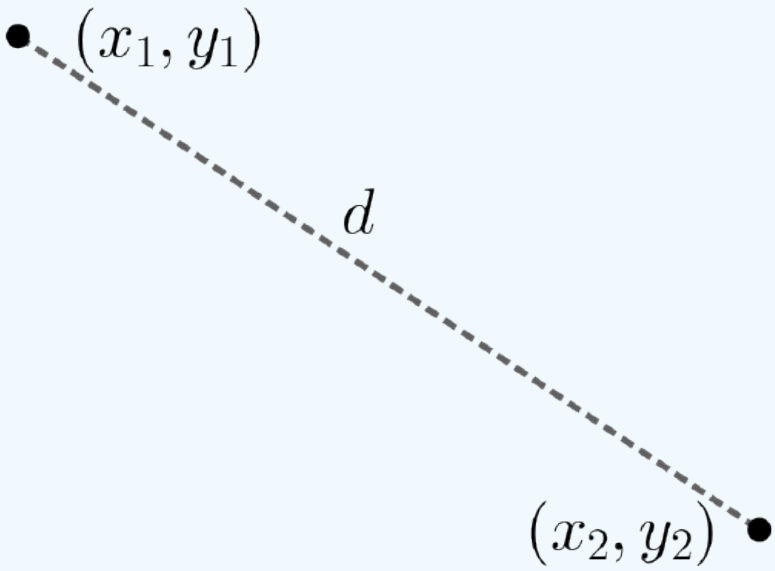


Figure 6.5.2

How do we use the distance formula to find the length of the line connecting the two points, d?

c. How are these two things the same? How are they different?

This might be a reminder of something we already knew, but let's make sure we are certain: when we calculate distances, we're really just using the Pythagorean Theorem! We can square the vertical distance between the points and the horizontal distance between the points, and then the length of the straight line connecting two points is:

$$\ell = \sqrt{(\Delta x)^2 + (\Delta y)^2}$$

This will be a useful formula for us to find an integral expression for the length of a curve.

If we think about the slice-and-sum technique, then we'll want to visualize the kth slice of whatever we're trying to measure. In this case, that means we'll divide the curve up into equally-wide slices and calculate the length of each subsection of the curve. We'll make a recognizable assumption: we'll assume that the curve is actually a straight line between the end points, and calculate that length.

Aside.

We have made similar assumptions along the way!

1. When we calculated area, we assumed that the curve(s) defining our region were constant on each subinterval. This is what gave us rectangular slices, with flat tops and bottoms!
2. When we calculated volumes by disks and washers, we again assumed that the curve(s) were constant on each subinterval. This is what gave us constant radii on each disk/washer!
3. When we calculated volumes by shells, this assumption of constant curves made the cylindrical shells have flat tops and bottoms!

Let's visualize the kth slice.

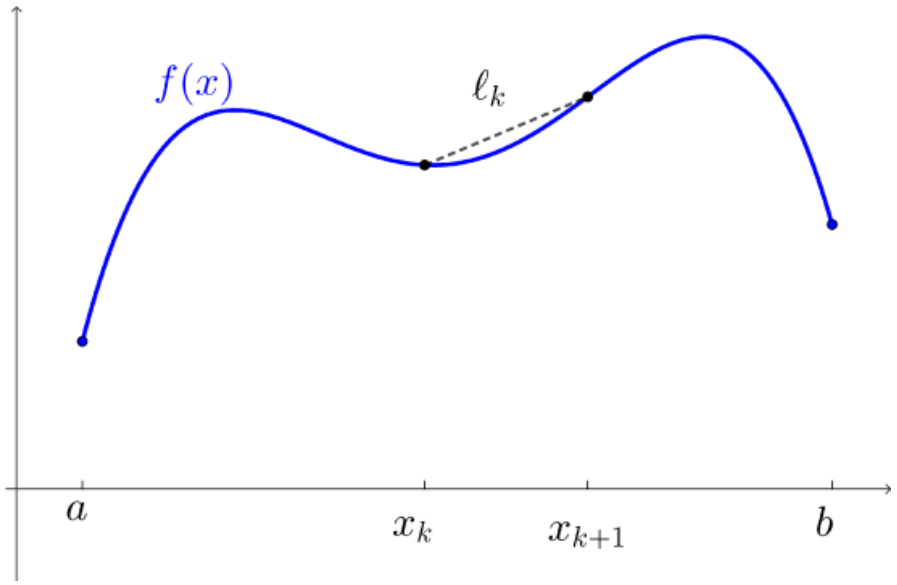


Figure 6.5.3

In order to calculate ℓ_k, the straight-line length connecting the end points of the kth subinterval, we can use the Pythagorean Theorem or distance formula (from Activity 6.5.1).

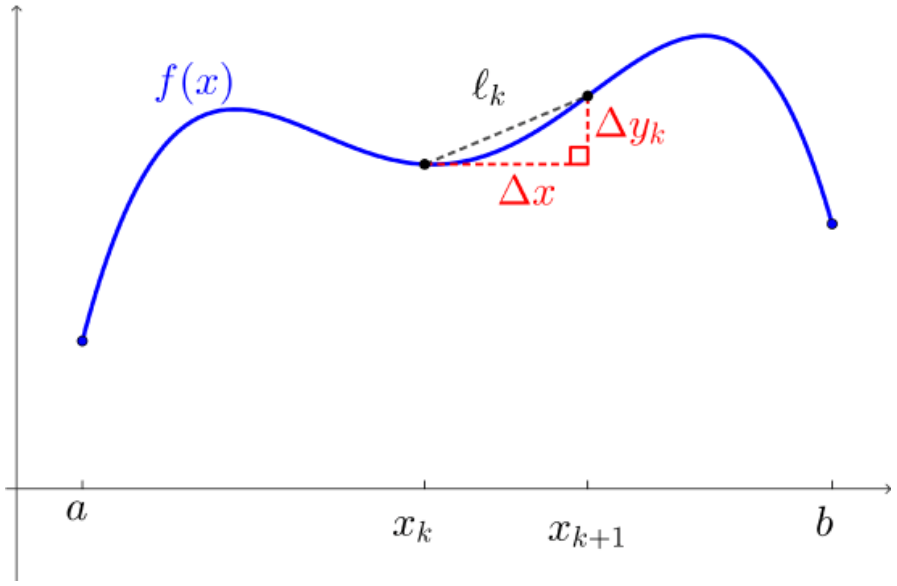


Figure 6.5.4

Let's start the slice-and-sum process.

$$\ell_k = \sqrt{(\Delta x)^2 + (\Delta y_k)^2}$$
$$\ell \approx \sum_{k=1}^{n} \sqrt{(\Delta x)^2 + (\Delta y_k)^2}$$

Two notes:

1. We're using Δy_k to denote the vertical distance between the end points of the kth subinterval because we expect these to differ for each subinterval. We don't need to do this for Δx, since we've been slicing things into equally-wide subintervals this whole time.

2. This isn't a Riemann sum. This is much more important, and much more pressing.

Before we can do anything, we need to try to manipulate this sum so that it is in the form of a Riemann sum. What does this mean? What are some of the things required for the Riemann sum structure that we don't have here? Feel free to look back at Definition 5.2.3 Riemann Sum to remind yourself what elements are needed for a Riemann sum.

Notice, first, that we need a function evaluated at any single input on the subinterval: $f(x_k^*)$. In our version, we have a function evaluated twice at very specific inputs:

$$\Delta y_k = f(x_{k+1}) - f(x_k)$$

We'll need to rethink about how we represent this part in order to get a single function output.

We also need to have this function multiplied by Δx. In our sum, we have Δx as a part of the function itself, under the square root.

We'll want to move this Δx outside of the root. Let's start there. We'll start by looking at the sum to approximate the length ℓ and factoring Δx outside of the root.

$$\begin{aligned}
\ell &\approx \sum_{k=1}^{n} \sqrt{(\Delta x)^2 + (\Delta y_k)^2} \\
&= \sum_{k=1}^{n} \sqrt{(\Delta x)^2 \left(1 + \frac{(\Delta y_k)^2}{(\Delta x)^2}\right)} \\
&\sum_{k=1}^{n} \Delta x \sqrt{1 + \left(\frac{\Delta y_k}{\Delta x}\right)^2} \\
\ell &\approx \sum_{k=1}^{n} \sqrt{1 + \left(\frac{\Delta y_k}{\Delta x}\right)^2} \Delta x
\end{aligned}$$

This looks better! We have Δx floating around at the end of our sum, ready to turn into dx once we apply the limit as $n \to \infty$.

The inside of our root, though, is still a bit messed up. We would like a single function of x_k^*, any x-value from the kth subinterval. Instead, we have a function involving the two x-values of the end points and we still have Δx involved in this part!

But we can notice something about $\frac{\Delta y_k}{\Delta x}$: it really is the slope of the straight line! Can we use a function to represent this? We absolutely can approximate this slope using a function: the derivative!

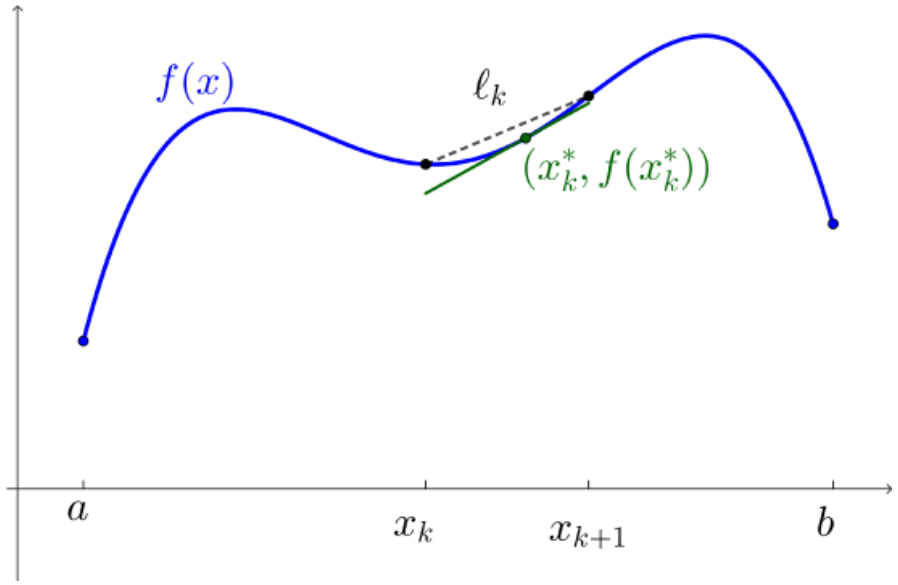


Figure 6.5.5

If we pick some point, $(x_k^*, f(x_k^*))$, on the kth subinterval, then we can approximate $\frac{\Delta y_k}{\Delta x}$ with $f'(x_k^*)$. This is a fine approximation of this slope (and the Mean Value Theorem guarantees that there is a point on the subinterval where $f'(x_k^*) = \frac{\Delta y_k}{\Delta x}$ exactly), but the real magic will happen when $n \to \infty$. The definition of the Derivative at a Point will make sure that these slopes are equal in the limit!

Let's return to our slice-and-sum process.

$$\begin{aligned} \ell &\approx \sum_{k=1}^{n} \sqrt{1 + \left(\frac{\Delta y_k}{\Delta x}\right)^2} \Delta x \\ &\approx \sum_{k=1}^{n} \sqrt{1 + \left(f'(x_k^*)\right)^2} \Delta x \end{aligned}$$

This is a Riemann sum! We can apply a limit and get an integral!

$$\begin{aligned} \ell &\approx \sum_{k=1}^{n} \sqrt{1 + \left(f'(x_k^*)\right)^2} \Delta x \\ \ell &= \lim_{n\to\infty} \sum_{k=1}^{n} \sqrt{1 + \left(f'(x_k^*)\right)^2} \Delta x \\ &= \int_{x=a}^{x=b} \sqrt{1 + \left(f'(x)\right)^2} dx \end{aligned}$$

Definition 6.5.6: Length of a Curve.

If $f(x)$ is continuous on the interval $[a, b]$ and differentiable on (a, b), then the length of the curve $y = f(x)$ from $x = a$ to $x = b$ is:

$$\int_{x=a}^{x=b} \sqrt{1 + \left(f'(x)\right)^2} dx$$

Example 6.5.7

Find an integral expression representing the length of the following curves.

a. The curve $y = \dfrac{1}{x}$ from $x = 1$ to $x = 2$.
b. The curve $y = \sin^{-1}(x)$ from $x = -1$ to $x = 1$.

Solution

a. Since $y' = -\dfrac{1}{x^2}$, then we can construct the following integral:

$$\begin{aligned} \ell &= \int_{x=1}^{x=2} \sqrt{1 + (y')^2} dx \\ &= \int_{x=1}^{x=2} \sqrt{1 + \left(-\frac{1}{x^2}\right)^2} dx \\ &= \int_{x=1}^{x=2} \sqrt{1 + \frac{1}{x^4}} dx \end{aligned}$$

Instead of worrying about actually evaluating this integral, we'll leave it like this. If you do want to fiddle with this integral, than it might be helpful to note that we can rewrite it:

$$\int_{x=1}^{x=2}\sqrt{1+\frac{1}{x^4}}dx=\int_{x=1}^{x=2}\sqrt{\frac{x^4+1}{x^4}}dx$$
$$=\int_{x=1}^{x=2}\frac{\sqrt{1+x^4}}{x^2}dx$$

b. We know that $y'=\frac{1}{\sqrt{1-x^2}}$, so we can construct the following integral:

$$\ell=\int_{x=-1}^{x=1}\sqrt{1+\left(\frac{1}{\sqrt{1-x^2}}\right)^2}dx$$
$$=\int_{x=-1}^{x=1}\sqrt{1+\frac{1}{1-x^2}}dx$$

We can leave this integral like this for now. Similar to the first example, though, we can rewrite this if you'd like to explore it more!

$$\int_{x=-1}^{x=1}\sqrt{1+\frac{1}{1-x^2}}dx=\int_{x=-1}^{x=1}\sqrt{\frac{2-x^2}{1-x^2}}dx$$

Integrals for Evaluating the Surface Area of a Solid

Moving from the length of some curve towards calculating the surface area of some solid of revolution won't be hard: we'll use the length formula in our procedure!

Let's build this surface area formula. Consider some function, $f(x)$, on the interval from $x=a$ to $x=b$.

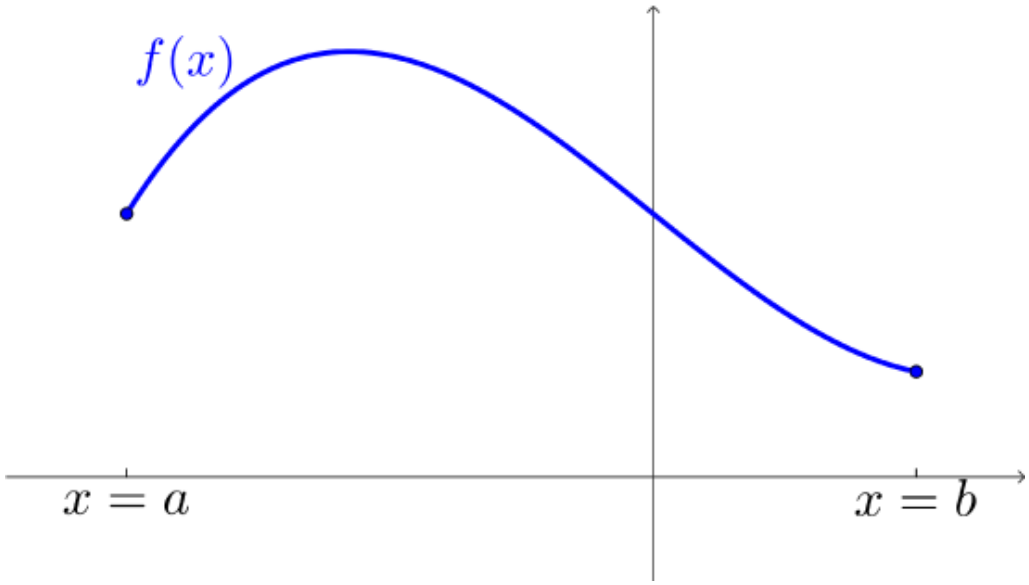


Figure 6.5.8

Instead of forming a rectangle for the kth slice, we'll do the same thing that we did for arc length: we'll connect the end points of the kth subinterval. This will create a trapezoid.

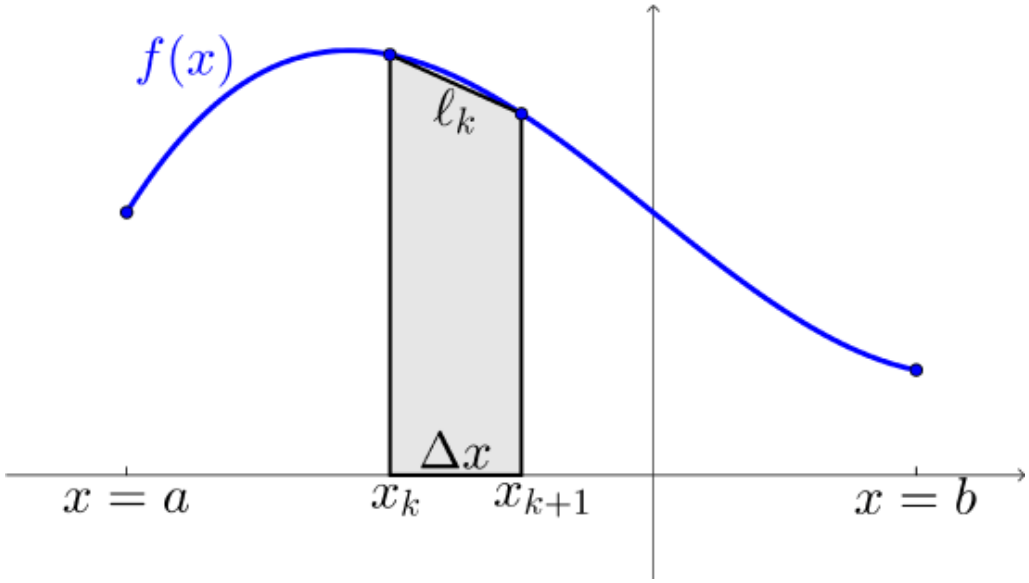


Figure 6.5.9

We'll use ℓ_k to represent the diagonal length of the line connecting the endpoints. Notice that this is going to become the arc length.

When we revolve the curve $y = f(x)$ around the x-axis, we can see not just the solid created by the curve, but the solid representing this kth slice.

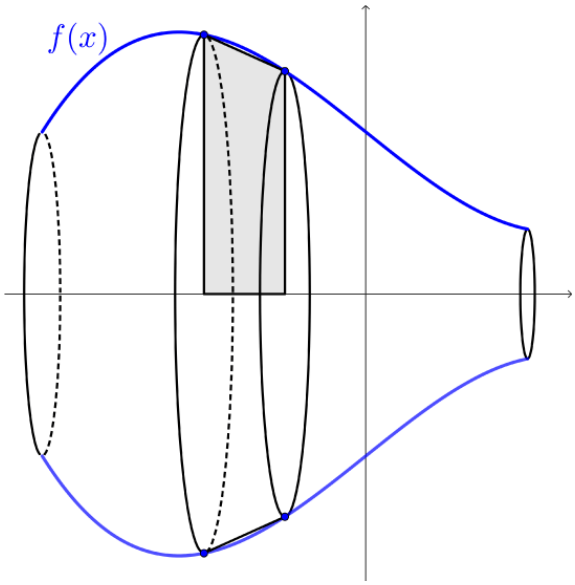


Figure 6.5.10

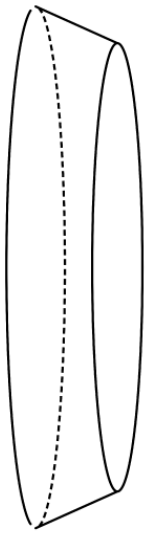


Figure 6.5.11 The kth frustum-shaped slice.

In order for us to find the surface area of this kth slice, we'll think about how "far" the diagonal line revolves. This is based on the circumference of each circular end of our slice, which means we have two radii to consider: the function outputs at both endpoints of the kth subinterval:

$$A_k = 2\pi \left(\frac{f(x_k) + f(x_{k+1})}{2} \right) \ell_k$$

This is going to become problematic, since we need only one function output evaluated at some x_k^* on the kth subinterval.

Instead, we can select some x-value on the interval and use the function output at that point to represent the radius of our kth slice.

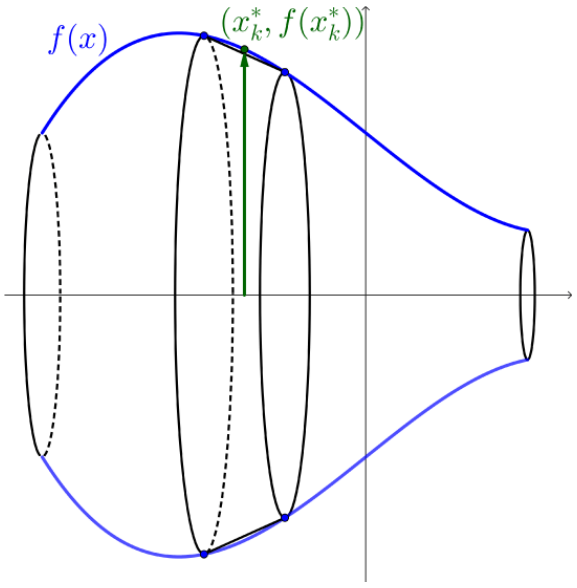


Figure 6.5.12

Instead of averaging the large and small radii from the end-points, we'll just select the one function output to represent this "average" radius. In the limit as $n \to \infty$, things will work out, since this randomly selected radius will become exactly equal to the average radius in the limit since $\Delta x \to 0$.

Now we can slice and sum!

$$\begin{aligned} A_k &= 2\pi f\left(x_k^*\right)\ell_k \\ &= 2\pi f\left(x_k^*\right)\sqrt{1+\left(f'\left(x_k^*\right)\right)^2}\Delta x \\ A &\approx \sum_{k=1}^{n} 2\pi f\left(x_k^*\right)\sqrt{1+\left(f'\left(x_k^*\right)\right)^2}\Delta x \\ A &= \lim_{n\to\infty}\sum_{k=1}^{n} 2\pi f\left(x_k^*\right)\sqrt{1+\left(f'\left(x_k^*\right)\right)^2}\Delta x \\ &= \int_{x=a}^{x=b} 2\pi f(x)\sqrt{1+\left(f'(x)\right)^2}dx \end{aligned}$$

Definition 6.5.13: Surface Area.

Let $f(x)$ be a continuous function with $f(x)\geq 0$ on the interval $[a,b]$ and differentiable on (a,b). If the region bounded by $f(x)$ and the x-axis from $x=a$ to $x=b$ is revolved around the x-axis, then the surface area of the resulting solid is:

$$A = 2\pi\int_{x=a}^{x=b} f(x)\sqrt{1+\left(f'(x)\right)^2}dx$$

Practice Problems

Exercise 6.5.1

The formula for arc length of the function $f(x)$ on the interval $[a,b]$ is $L=\int_{x=a}^{x=b}\sqrt{1+f'(x)^2}dx$. Explain how this definition is built, using the slice-and-sum method. Make sure to explain how we the Pythagorean Theorem is involved.

Exercise 6.5.2

Why do we use $f'(x)$ in the formula for arc length?

Exercise 6.5.3

For each of the following curves on intervals, evaluate the arc length.

a. $y=-6x+2$ on [1, 5]
b. $y=\dfrac{x^{3/2}}{3}$ on [0, 60]
c. $y=\dfrac{x^4}{4}+\dfrac{1}{8x^2}$ on [1, 3]

Exercise 6.5.4

For each of the following curves on intervals, set up an integral representing the arc length. Do not evaluate.

a. $y=\tan^2(x)$ on $\left[-\dfrac{\pi}{4},\dfrac{\pi}{4}\right]$
b. $y=\ln\left(x^2\right)$ on $[1,e]$
c. $y=\sqrt{x+1}$ on [–1, 8]

Exercise 6.5.5

Why is the formula for arc length seemingly involved in the integral formula for surface area of a solid of revolution?

Exercise 6.5.6

In the integral formula $A = \int_{x=a}^{x=b} 2\pi f(x)\sqrt{1+f'(x)^2}dx$, what does $f(x)$ represent? What about 2π?

Exercise 6.5.7

For each of the following curves and intervals, find the surface area of the solid formed when the curve is revolved around the x-axis.

a. $y = 2x + 3$ on [0, 3]
b. $y = 4\sqrt{x}$ on [4, 9]
c. $y = \dfrac{x^4}{4} + \dfrac{1}{8x^2}$ on [1, 3]

Exercise 6.5.8

For each of the following curves and intervals, set up the surface area of the solid formed when the curve is revolved around the x-axis. Do not evaluate the integral.

a. $y = \sin\left(e^{x^2}\right) + 1$ on [0, 2]
b. $y = \ln\left(x^2\right)$ on $[1, e]$

This page titled 6.5: Arc Length and Surface Area is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Peter Keep.

6.6: Other Applications of Integrals

We should pause and think (even briefly) about the whole point of this chapter.

Do you think you will need to know how to calculate the volume of a solid of revolution?

Is being able to think about the surface area of a solid of revolution important?

Are you going to need to calculate a bunch of arc lengths constantly in your other classes, jobs, or spare time?

No. Probably not. It's fun to think about this stuff, but the goal is not really the formulas: it's the construction.

What we really should be leaving this chapter with is a renewed view of what a definite integral is. After Chapter 5 Anti-derivatives and Integrals, we probably thought about a definite integral as a measurement of the (signed) area "under the curve" (bounded between the curve and the x-axis). Hopefully now, though, we have a deeper view of the definite integral through this slice-and-sum process.

A definite integral is an accumulation of some function outputs multiplied by the space between the inputs.

So when we do move past these topics into other classes, jobs, or your spare time, then you will run into formulas involving definite integrals. Hopefully, based on this chapter, we have the tools to deconstruct those formulas to find out what they are measuring and how they are measuring it.

To end the chapter, let's look at some applications of definite integrals into formulas from other academic fields.

Physics Application: Mass

Let's consider some object made out of a material with a constant density. To calculate the mass of the object, we need to know the size (the volume of the object) and the density, and then we can multiply:

$$\text{mass} \;=\; \text{density} \;\cdot\; \text{volume}$$

We've seen how we can use the slice-and-sum formula to calculate volumes by slicing the object into thin pieces and approximating the volume by multiplying the cross-sectional area by the thickness of the slice:

$$V_k = A\left(x_k^*\right)\Delta x$$

We can easily construct an integral formula for the mass of an object now! We just need to find some measurement of the mass of a single slice of an object, and then start summing until we end up with a definite integral! So, for an object spanning from $x = a$ to $x = b$ with density, ρ, we have:

$$\begin{aligned} M_k &= \rho V_k \\ &= \rho A\left(x_k^*\right)\Delta x \\ M &\approx \sum_{k=1}^{n} \rho A\left(x_k^*\right)\Delta x \\ M &= \lim_{n\to\infty} \sum_{k=1}^{n} \rho A\left(x_k^*\right)\Delta x \\ &= \int_{x=a}^{x=b} \rho A(x)dx \end{aligned}$$

This is great! We can easily extend this to other situations, as well.

Imagine if the density of our object isn't constant. Maybe we have some metal alloy, where the mixture of the metal changes as we traverse through the object. This variable density can be represented, then, as $\rho(x)$. We can replace the constant ρ with $\rho\left(x_k^*\right)$ and then eventually $\rho(x)$ to get:

$$M = \int_{x=a}^{x=b} \rho(x)A(x)dx$$

Theorem 6.6.1: Mass of an Object.

For some solid spanning from $x = a$ to $x = b$ where $A(x)$ is the cross-sectional area of the object at x and $\rho(x)$ is the density of the object at x, then the total mass of the object is:

$$M = \int_{x=a}^{x=b} \rho(x)A(x)dx$$

For very thin objects, like wires, we can not consider the cross-sectional area and simply use what we call the linear density: $\rho(x)$ measures the amount of mass per unit length (instead of mass per cubic unit of volume). In this case, our formula changes.

Theorem 6.6.2: Mass of a Thin Wire.

For some thin object (like a wire) spanning from $x = a$ to $x = b$, then if $\rho(x)$ measures the mass per unit of length ($\rho(x)$ is a linear density function), then the mass of the object is:

$$M = \int_{x=a}^{x=b} \rho(x)dx$$

Example 6.6.3

For each of the following objects, set up an integral expression to calculate the mass. Evaluate the integrals!

a. A thin wire that is 30 cm long with a density function $\rho(x) = 2 + \frac{x}{60}$, where $\rho(x)$ measures the density of the wire at x in g/cm.
b. A thin wire that is 100 cm long, with an unknown density. We assume that the density changes linearly from one end to the other, with the measured density at both ends being 5.1 g/cm and 4.7 g/cm. **Hint**. Construct the density function using two points: (0, 5.1) and (100, 4.7). You could switch the orders and use (0, 4.7) and (100, 5.1) as well. What is the linear function connecting these points?

Solution

a.
$$\begin{aligned} M &= \int_{x=0}^{x=30} \left(2 + \frac{x}{60}\right) dx \\ &= \left(2x + \frac{x^2}{120}\right)\Big|_{x=0}^{x=30} \\ &= 60 + \frac{900}{120} \\ &= 67.5 \end{aligned}$$

The mass of the wire is 67.5 g.

b. We'll use the points (0, 5.1) and (100, 4.7) to construct $\rho(x)$.

$$\begin{aligned} \rho(x) &= (\underbrace{\frac{4.7 - 5.1}{100 - 0}}_{\text{slope}})(x - 0) + 5.1 \\ &= -0.004x + 5.1 \end{aligned}$$

Now we can integrate:

$$\begin{aligned} M &= \int_{x=0}^{x=100} 5.1 - 0.004x dx \\ &= \left(5.1x - 0.002x^2\right)\Big|_{x=0}^{x=100} \\ &= 510 - 20 \\ &= 490 \end{aligned}$$

So the mass of the wire is 490 g.

Physics Application: Work

Another example of a measurement found in physics contexts is Work. Work is, generally, the amount energy transferred to (or from, depending on perspective) an object by some force across some distance or displacement.

$$\text{Work} = \text{Force} \cdot \text{distance}$$

In general, we can use the "slice-and-sum" process when the force applied to our object is some function of the position:

$$\begin{aligned} W_k &= F\left(x_k^*\right)\Delta x \\ W &\approx \sum_{k=1}^{n} F\left(x_k^*\right)\Delta x \\ W &= \lim_{n\to\infty} \sum_{k=1}^{n} F\left(x_k^*\right)\Delta x \\ &= \int_{x=a}^{x=b} F(x)dx \end{aligned}$$

Theorem 6.6.4: Work Required.

If $F(x)$ is a function measuring the force applied to an object at some x-value in the interval $[a, b]$, then the work done to move the object across the interval is:

$$W = \int_{x=a}^{x=b} F(x)dx$$

In many contexts, we can notice that the force being applied to the object is dependent on the position of the object along its path. For instance, in the 1600's, British physicist Robert Hooke claimed (and proved) that the force required to stretch or compress a spring is directly proportional to the distance that it is away from "equilibrium" (the position the spring naturally rests in).

Work: Springs

We'll first look at the springs, since the force functions are relatively simple. Robert Hooke's claim (Hooke's Law) says that the force function for a spring is always in the form $F(x) = kx$ where k is just some constant proportion connecting distance to the force required for that particular spring.

Example 6.6.5

Calculate the work required to stretch or compress the following springs.

a. A spring stretched to 0.1 m past its equilibrium position, where a Force of $F(x) = 180x$ is applied, measured in N.
b. A force of 81 N was applied to a spring to stretch it 0.8 m from its equilibrium position. Calculate the work required to stretch it 0.2 m further.

Solution

a.

$$\begin{aligned} W &= \int_{x=0}^{x=0.1} 180x\,dx \\ &= 90x^2\Big|_{x=0}^{x=0.1} \\ &= 0.9 \end{aligned}$$

So the work required to stretch the spring is 0.9 J.

b.

$$\begin{aligned} F(x) &= \frac{81}{0.8}x \\ &= \frac{405x}{4} \\ W &= \int_{x=0.8}^{x=1} \frac{405x}{4}\,dx \\ &= \left(\frac{405x^2}{8}\right)\Big|_{x=0.8}^{x=1} \\ &= \frac{405}{8} - \frac{162}{5} \\ &= \frac{729}{40} \end{aligned}$$

So the work required to stretch the spring is 18.225 J.

Work: Pumping Problems

We can explore another class of examples with some more complications in the setup. We still will be thinking about force and distance, but we can reframe them with a new example.

What if we pump liquid out of the top of a tank? How does that work? There are some fun things to think about. Let's visualize this tank below, and then talk through the construction of the formula.

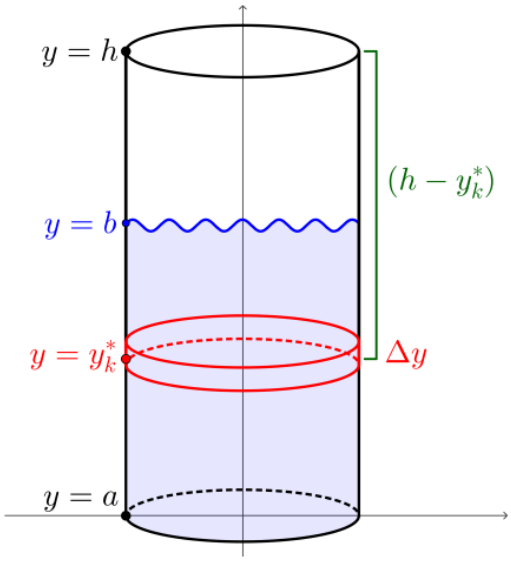


Figure 6.6.6 Diagram of a pumping problem.

We can see a couple of notable things here:

- We're going to slice up the liquid. So from $y = a$, the bottom of the tank, up to $y = b$, the height at the top of the volume of liquid, we will create k slices (each at some y_k^* y-value) with Δy representing the thickness of the slice.
- The distance portion of the formula is not represented by the width of a subinterval from the whole distance. Instead, we'll note that liquid at the top of the tank needs to be pumped a shorter distance than the liquid near the bottom of the tank. The

distance part, then, is a function based on the y-value. So, if $y = h$ is the height of the tank, then $(h - y_k^*)$ is the distance that the liquid in the kth slice will move.

- To calculate the force required to pump the liquid in the kth slice, we need to know the mass of the liquid and the acceleration needed to pump the liquid out of the tank. The acceleration will be whatever the acceleration needed to overcome gravity: we'll use g, the positive gravitational constant. For the mass, we'll think about the density of the liquid (a constant, ρ) and the volume of the kth slice.
- To find the volume of the kth slice, we'll think about the cross-sectional area of the slice at y_k^* (we'll call it $A(y_k^*)$) multiplied by the thickness, Δy.

This gives us the following in the slice-and-sum process:

$$\begin{aligned} W_k &= F(y_k^*)(h - y_k^*) \\ &= (\rho A(y_k^*)\Delta y)\, g\,(h - y_k^*) \\ &= \rho g A(y_k^*)(h - y_k^*)\Delta y \end{aligned}$$

$$W \approx \sum_{k=1}^{n} \rho g A(y_k^*)(h - y_k^*)\Delta y$$

$$\begin{aligned} W &= \lim_{n\to\infty} \sum_{k=1}^{n} \rho g A(y_k^*)(h - y_k^*)\Delta y \\ &= \rho g \int_{y=a}^{y=b} A(y)(h - y)dy \end{aligned}$$

So, when we consider these pumping problems, we really need to take note of only a few things:

- What are the limits of integration? Where is the liquid we're pumping (based on y-value heights)?
- What is the geometry of our tank? When we consider a single slice of the liquid, what kind of shape will we have, and how do we calculate the cross-sectional area?

Example 6.6.7

For each of the following tanks, calculate the work required to empty the liquid from the tank by pumping it through the top of the tank.

a. Consider a cylindrical tank, similar to the one in Figure 6.6.6. The radius of the tank is 0.5 m, and the height is 2 m. The tank is half full of some liquid with density ρ.
b. Now box-shaped tank, where the base of the tank is 1 meters by 3 meters and the height is 2 meter. The tank is filled all the way to the top with some liquid with density ρ.
c. Consider a conical tank, where the radius of the base is 3 meters and the cone is 2 meters tall. The liquid is filled up to the 1.5 meter mark.

Solution

a. The cross sectional area is:

$$\begin{aligned} A &= \pi r^2 \\ A(y) &= \pi(0.5)^2 \\ &= \frac{\pi}{4} \end{aligned}$$

Then the limits of integration will be from $y = 0$ to $y = 1$, since the height is 2 m and the liquid, then, reaches to 1 m.

$$\begin{aligned} W &= \rho g \int_{y=0}^{y=1} A(y)(2-y)dy \\ &= \frac{\rho g \pi}{4} \int_{y=0}^{y=1} 2-ydy \\ &= \frac{\rho g \pi}{4} \left(2y - \frac{y^2}{2}\right)\Big|_{y=0}^{y=1} \\ &= \frac{\rho g \pi}{4} \left(\frac{3}{2}\right) \\ &= \frac{3\rho g \pi}{8} \end{aligned}$$

b. The cross sectional area is:

$$\begin{aligned} A &= \ell w \\ A(y) &= 1(3) \\ &= 3 \end{aligned}$$

Then the limits of integration will be from $y = 0$ to $y = 2$, since the height is 2 m and the liquid, then, reaches to the top.

$$\begin{aligned} W &= \rho g \int_{y=0}^{y=2} A(y)(2-y)dy \\ &= 3\rho g \int_{y=0}^{y=2} 2-ydy \\ &= 3\rho g \left(2y - \frac{y^2}{2}\right)\Big|_{y=0}^{y=2} \\ &= 3\rho g(2) \\ &= 6\rho g \end{aligned}$$

c. Finding the cross-sectional area function will be a bit harder. It is the area of a circle, but the radius changes based on the y-value. At $y = 0$, we have $r = 3$. At $y = 2$, we have $r = 0$. So the rate of change of the radius is -1.5 meters per meter of height.

$$A(y) = \pi(3 - 1.5y)^2$$

The limits of integration are from $y = 0$ to $y = 1.5$:

$$\begin{aligned} W &= \rho g \pi \int_{y=0}^{y=1.5} (3-1.5y)^2(2-y)dy \\ &= \rho g \pi \int_{y=0}^{y=1.5} 18 - 27y + 13.5y^2 - 2.25y^3 dy \\ &= \rho g \pi \left(18y - \frac{27y^2}{2} + \frac{13.5y^3}{3} - \frac{2.25y^4}{4}\right)\Big|_{y=0}^{y=1.5} \\ &= \frac{2295\rho g \pi}{256} \end{aligned}$$

Practice Problems

Exercise 6.6.1

Consider a thin wire whose density is variable and is written as $\rho(x)$ on some interval. For each of the following density functions (and their intervals), build and evaluate an integral representing the mass of the wire.

a. $\rho(x) = \dfrac{2x+1}{10}$ on [0, 10]
b. $\rho(x) = x^2 - 2$ on [100, 200]

Exercise 6.6.2

Consider a spring that requires 100 N of force to stretch the spring 0.3 m from its equilibrium position.

a. Using Hooke's Law, find the spring constant k such that $F(x) = kx$.
b. How much work is required to stretch the spring 0.4 m from its equilibrium position?
c. How much work is required to stretch the spring another 0.1 m after this?

Exercise 6.6.3

Consider a spring that is stretched 0.1 m from its equilibrium position. It requires 10 N to stretch it an additional 0.2 m.

a. Using Hooke's Law, find the spring constant k such that $F(x) = kx$.
b. How much work is required to stretch the spring 1 m from its equilibrium position?
c. How much work is required to compress the spring 0.2 m from is equilibrium position?

Exercise 6.6.4

For each tank, assume the density of the liquid is ρ, and set up and evaluate the integral representing the work required to pump the liquid out of the top of the tank.

a. A square-based tank that is 2 meters tall with a 0.5 m × 0.5 m base, full of liquid.
b. A cylindrical tank where the base has a radius of 1 m and the height is 3 m, half-full of liquid.
c. A frustum-shaped tank where the bottom radius is 4 m, the top radius is 1 m, the height is 2 m, and the tank is filled to a height of $\frac{2}{3}$ m.

This page titled 6.6: Other Applications of Integrals is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Peter Keep.

7: TECHNIQUES FOR ANTIDIFFERENTIATION

7.1: Improper Integrals

7.2: More on u Substitution

7.3: Manipulating Integrands

7.4: Integration by Parts

7.5: Integrating Powers of Trigonometric Functions

7.6: Trigonometric Substitution

7.7: Partial Fractions

This page titled 7: Techniques for Antidifferentiation is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Peter Keep.

7.1: Improper Integrals

We're going to think a bit about integration with a twist: what happens when our "definite" integrals can't actually be evaluated? First, let's try to sink ourselves back into the context we've been in for a while now: what kinds of problems have we encountered so far, and how do we use our calculus intuition to get around those problems?

Exercise 7.1.1: Remembering a Theme so Far.

a. Let's say that we want to find what the y-values of some function $f(x)$ are when the x-values are "infinitely close to" some value, $x = a$. Since there is no single x-value that is "infinitely close to" a that we can evaluate $f(x)$ at, we need to do something else. How do we do this?
b. Let's say that we want to find the rate of change of some function instantaneously at a point with $x = a$. We can't find a rate of change unless we have two points, since we need to find some differences in the outputs and inputs. How do we do this?
c. Suppose you want to find the total area, covered by an infinite number of infinitely thin rectangles. You have a formula for finding the dimensions and areas for some finite number of rectangles, but how do we get an infinite number of them?
d. Can you find the common calculus theme in each of these scenarios?

So moving forward, we want to remember how we typically have solved these problems. Now, let's try to identify the types of problems with integrals that we need to figure our way around.

Activity 7.1.2: Remembering the Fundamental Theorem of Calculus.

We want to think about generalizing our notion of integrals a bit. So in this activity, section, we're going to think about some of the requirements for the Fundamental Theorem of Calculus and try to loosen them up a bit to see what happens. We'll try to construct meaningful approaches to these situations that fit our overall goals of calculating area under a curve. This practice, in general, is a really good and common mathematical process: taking some result and playing with the requirements or assumptions to see what else can happen. So it might feel like we're just fiddling with the "What if?" questions, but what we're actually doing is good mathematics!

a. What does the Fundamental Theorem of Calculus say about evaluating the definite integral $\int_{x=a}^{x=b} f(x)dx$?
b. What do we need to be true about our setup, our function, etc. for us to be able to apply this technique to evaluate $\int_{x=a}^{x=b} f(x)dx$?

We are going to introduce the idea of "Improper Integrals" as kind-of-but-not-quite definite integrals that we can evaluate. They are going to violate the requirements for the Fundamental Theorem of Calculus, but we'll work to salvage them in meaningful ways.

This should build a pretty good idea of a new "class" of integrals: ones that aren't quite definite integrals that we can evaluate with the Fundamental Theorem of Calculus, but ones that we can use limits to get at.

Improper Integrals

Definition 7.1.1: Improper Integral.

An integral is an **improper integral** if it is an extension of a definite integral whose integrand or limits of integration violate a requirement in one of two ways:

1. The interval that we integrate the function over is unbounded in width, or infinitely wide.
2. The integrand is unbounded in height, or infinitely tall, somewhere on the interval that we integrate over.

With this definition, we can think about the strategies that we got from Activity 7.1.1: we're going to identify the problems in our integral (infinite width of the interval or infinite height of the integrand function) and use a limit!

Before we formalize that, though, let's try to think about how this works by being really explicit about what this limit actually is doing.

Activity 7.1.3: Approximating Improper Integrals.

In this activity, we're going to look at two improper integrals:

1. $\int_{x=2}^{\infty} \frac{1}{(x+1)^2} dx$
2. $\int_{x=-1}^{x=2} \frac{1}{(x+1)^2} dx$

a. First, let's just clarify to ourselves what it means for an integral to be improper. Why are each of these integrals improper? Be specific!

b. Let's focus on $\int_{x=2}^{\infty} \frac{1}{(x+1)^2} dx$ first. We're going to look at the slightly different integral:

$$\int_{x=2}^{x=t} \frac{1}{(x+1)^2} dx$$

As long as t is some real number with $t > 2$, then our function is continuous and bounded on $[2, t]$, and so we can evaluate this integral:

$$\int_{x=2}^{x=t} \frac{1}{(x+1)^2} dx = F(t) - F$$

where $F(x)$ is an antiderivative of $f(x) = \frac{1}{(x+1)^2}$. Find and antiderivative, $F(x)$.

c. Now we're going to evaluate some areas for different values of t. Use your antiderivative $F(x)$ from above!

- Let's start with making $t = 99$. So we're going to evaluate:

$$\int_{x=2}^{x=99} \frac{1}{(x+1)^2} dx = F(99) - F(2)$$

- Now let $t = 999$. Evaluate:

$$\int_{x=2}^{x=999} \frac{1}{(x+1)^2} dx = F(999) - F(2)$$

- Now let $t = 9999$. Evaluate:

$$\int_{x=2}^{x=9999} \frac{1}{(x+1)^2} dx = F(9999) - F(2)$$

d. Based on what you found, what do you think is happening when $t \to \infty$ to the definite integral

$$\int_{x=2}^{x=t} \frac{1}{(x+1)^2} dx = F(t) - F(2)?$$

e. Ok, we're going to switch our focus to the other improper integral,

$$\int_{x=-1}^{x=2} \frac{1}{(x+1)^2} dx$$

again, we'll look at a slightly different integral:

$$\int_{x=t}^{x=2} \frac{1}{(x+1)^2} dx$$

As long as t is some real number with $-1 < t < 2$, then our function is continuous and bounded on $[t, 2]$, and so we can evaluate this integral:

$$\int_{x=t}^{x=2} \frac{1}{(x+1)^2} dx = F(2) - F(t)$$

where $F(x)$ is an antiderivative of $f(x) = \dfrac{1}{(x+1)^2}$. We can use the same antiderivative as before! We're going to evaluate this intergal for different values of t again, but this time we'll use values that are close to –1, but slightly bigger, since we want to be in the interval [–1, 2].

- Let's start with making $t = -\dfrac{9}{10}$. So we're going to evaluate:

$$\int_{x=-9/10}^{x=2} \frac{1}{(x+1)^2}dx = F(2) - F\left(-\frac{9}{10}\right)$$

- Now let $t = -\dfrac{99}{100}$. Evaluate:

$$\int_{x=-99/100}^{x=2} \frac{1}{(x+1)^2}dx = F(2) - F\left(-\frac{99}{100}\right)$$

- Now let $t = -\dfrac{999}{1000}$. Evaluate:

$$\int_{x=-999/1000}^{x=2} \frac{1}{(x+1)^2}dx = F(2) - F\left(-\frac{999}{1000}\right)$$

f. Based on what you found, what do you think is happening when $t \to -1^+$ to the definite integral

$$\int_{x=t}^{x=2} \frac{1}{(x+1)^2}dx = F(2) - F(t)?$$

We can think about putting this a bit more generally into limit notation, but we'll get to this later.

Ok, let's formalize these limits with some strategies for evaluating improper integrals!

Strategies for Evaluating Improper Integrals

Evaluating Improper Integrals (Infinite Width).

For a function $f(x)$ that is continuous on $[a, \infty)$, we can evaluate the improper integral $\int_{x=a}^{\infty} f(x)dx$:

$$\int_{x=a}^{\infty} f(x)dx = \lim_{t\to\infty} \int_{x=a}^{x=t} f(x)dx$$

If $f(x)$ is continuous on $(-\infty, b]$, we can evaluate the improper integral $\int_{-\infty}^{x=b} f(x)dx$:

$$\int_{-\infty}^{x=b} f(x)dx = \lim_{t\to-\infty} \int_{x=t}^{x=b} f(x)dx$$

Finally, if $f(x)$ is continuous on $(-\infty, \infty)$ and m is some real number, then we can evaluate the improper integral $\int_{-\infty}^{\infty} f(x)dx$:

$$\begin{aligned}\int_{-\infty}^{\infty} f(x)dx &= \int_{-\infty}^{x=m} f(x)dx + \int_{x=m}^{\infty} f(x)dx \\ &= \lim_{t\to-\infty} \int_{x=t}^{x=m} f(x)dx + \lim_{t\to\infty} \int_{x=m}^{x=t} f(x)dx\end{aligned}$$

Example 7.1.2

Evaluate the improper integral $\int_{x=2}^{\infty} \dfrac{1}{(x+1)^2}dx$ by evaluating the limit:

$$\lim_{t\to\infty} \int_{x=2}^{x=t} \frac{1}{(x+1)^2}dx = \lim_{t\to\infty} (F(t) - F(2))$$

Try to interpret this limit. What does it mean if this limit doesn't exist? What does it mean if the limit does exist? What does the actual number represent?

Evaluating Improper Integrals (Infinite Height).

For a function $f(x)$ that has an unbounded discontinuity (a vertical asymptote) at $x = m$ with $a < m < b$, but is otherwise continuous on $[a, b]$, then we can evaluate the improper integrals:

$$\int_{x=m}^{x=b} f(x)dx = \lim_{t\to m^+} \int_{x=t}^{x=b} f(x)dx$$

$$\int_{x=a}^{x=m} f(x)dx = \lim_{t\to m^-} \int_{x=a}^{x=t} f(x)dx$$

$$\begin{aligned}\int_{x=a}^{x=b} f(x)dx &= \int_{x=a}^{x=m} f(x)dx + \int_{x=m}^{x=b} f(x)dx \\ &= \lim_{t\to m^-} \int_{x=a}^{x=t} f(x)dx + \lim_{t\to m^+} \int_{x=t}^{x=b} f(x)dx\end{aligned}$$

Example 7.1.3

Evaluate the improper integral $\int_{x=-1}^{x=2} \frac{1}{(x+1)^2}dx$ by evaluating the limit:

$$\lim_{t\to -1^+} \int_{x=t}^{x=2} \frac{1}{(x+1)^2}dx = \lim_{t\to -1^+} (F(2) - F(t)).$$

Try to interpret this limit. What does it mean if this limit doesn't exist? What does it mean if the limit does exist? What does the actual number represent?

Ok, let's note that we can classify these improper integrals into two categories. We have already classified them based on the reason that they're improper. But now we also can classify them based on the outcome of the limit:

1. Improper integrals (of any type) whose limit exists.
2. Improper integrals (of any type) where the limit doesn't exist.

Let's define a term for this, so that our classification isn't so wordy.

Convergence and Divergence of an Improper Integral

Definition 7.1.4: Convergence of an Improper Integral.

We say that an improper integral **converges** if the limit of the appropriate definite integral exists. If the limit does not exist, then we say that the improper integral **diverges**.

All we've done here is added some language: we'll say that an improper integral diverges if the limit doesn't exist. And if the limit exists, we'll say that the improper integral "converges to _____."

Practice Problems

Exercise 7.1.1

Explain what it means for an integral to be improper. What kinds of issues are we looking at?

Exercise 7.1.2

Give an example of an integral that is improper due to an unbounded or infinite interval of integration (infinite width).

Exercise 7.1.3

Give an example of an integral that is improper due to an unbounded integrand (infinite height).

Exercise 7.1.4

What does it mean for an improper integral to "converge?" How does this connect with limits?

Exercise 7.1.5

What does it mean for an improper integral to "diverge?" How does this connect with limits?

Exercise 7.1.6

Why do we need to use limits to evaluate improper integrals?

Exercise 7.1.7

For each of the following improper integrals:

- Explain why the integral is improper. Be specific, and point out the issues in detail.
- Set up the integral using the correct limit notation.
- Antidifferentiate and evaluate the limit.
- Explain whether the integral converges or diverges.

a. $\int_{x=0}^{\infty} \frac{1}{\sqrt{x+1}} dx$
b. $\int_{x=0}^{\infty} e^{-2x} dx$
c. $\int_{x=-1}^{x=3} \frac{1}{x+1} dx$
d. $\int_{-\infty}^{x=0} \sqrt{e^x} dx$
e. $\int_{x=2}^{x=8} \frac{5}{(x-2)^3} dx$
f. $\int_{x=1}^{x=12} \frac{dx}{\sqrt[5]{12-x}}$

Exercise 7.1.8

One of the big ideas in probability is that for a curve that defines a probability density function, the area under the curve needs to be 1. What value of k makes the function $\frac{kx}{(x^2+3)^{5/4}}$ a valid probability distribution on the interval $[0, \infty)$?

Exercise 7.1.9

Let's consider the integral $\int_{x=1}^{\infty} \frac{\sqrt{x^2+1}}{x^2} dx$. This is a difficult integral to evaluate!

a. First, compare $\sqrt{x^2+1}$ to $\sqrt{x^2}$ using an inequality: which one is bigger?

b. Second, use this inequality to compare the function $\frac{\sqrt{x^2+1}}{x^2}$ to $\frac{1}{x}$ for $x > 0$: which one is bigger? Again, use your inequality from above to help!

c. Now compare $\int_{x=1}^{\infty} \frac{\sqrt{x^2+1}}{x^2} dx$ to $\int_{x=1}^{\infty} \frac{1}{x} dx$. Which one is bigger?

d. Explain how we can use this result to make a conclusion about whether our integral, $\int_{x=1}^{\infty} \frac{\sqrt{x^2+1}}{x^2} dx$ converges or diverges.

This page titled 7.1: Improper Integrals is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Peter Keep.

7.2: More on u Substitution

We're going to do some more thinking about an integration topic that we've already introduced in Section 5.6. We're going to do a bit more with it now, and try to build some more flexibility, so definitely review that introduction section if you'd like!

Before moving on, we should work through these few examples, just to make sure we remember what we're up to.

Activity 7.2.1: Recapping u-Substitution.

We're going to consider a few integrals, and work through each of the questions for all integrals.

1. $\int \left(\frac{1}{\sqrt{4-3x}} \right) dx$
2. $\int x^2 \left(5+x^3\right)^7 dx$
3. $\int \left(\frac{\sin(x)}{\cos(x)} \right) dx$
4. $\int \left(xe^{-x^2} \right) dx$

a. For each integral, explain why u-substitution is a good choice. How can you tell, just by looking at the integral, that this strategy will be a reasonable thing to try?
b. For each integral, explain your choice of u, and what that means for how we define du.
c. For each integral, is your definition of du present in the integrand function? How do you go about making this substitution when the integrand function isn't set up perfectly?
d. Finish the substitution and integration, and substitute back to the original variable.

Variable Substitution in Integrals

In u-substitution, we focus a lot on one specific kind of structure: composition in our integrand function and/or some function-derivative pairing present. We do this because we're undoing the The Chain Rule. Variable substitutions can be much more general in their goal, but this is a good one to focus on because it solves a specific problem that we might run into while integrating.

Strategy for u-Substitution.

Goal: undo the Chain Rule, or antidifferentiate functions with composition.

Process: We'll translate $\int f(x)dx$ to $\int g(u)du$ by labeling some "inside" function as u, and substituting its derivative, $du = u'dx$.

But we can use u-substitution more generally as a kind of grouping mechanism.

Activity 7.2.2: Turn Around Problems.

The two integrals that we're going to look at are "just" some u substitution problems, but I like to call integrals like these turnaround problems. We'll see why!

a. Consider the integral:

$$\int x\sqrt[3]{x+5}dx$$

. First, explain why u-substitution is reasonable here.

b. Identify du for your chosen u-substitution. When you substitute, you should notice that there are some extra bits in this integrand function that have not been assigned to be translated over to be written in terms of u. Which parts?
c. We need to think about how to write x in terms of u. Luckily, we already have everything we need! We have defined a link between the x variable and the u variable. We defined it as u being written as some function of u, but can we "turn around" that link to write x in terms of u?
d. Substitute the integral to be fully written in terms of u.
e. Before antidifferentiating, compare this integral with the original one. Specifically thinking about how we might multiply, describe the differences between the integrals with regard to composition and rewriting our integrand. Then, go ahead and use this nicely rewritten version to antidifferentiate and substitute back to x.
f. Apply this same strategy to the following integral:

$$\int \frac{x}{x+5}dx$$

This integral might be a bit trickier to find the composition in order to identify the u-substitution! Give some things a try!
g. Compare your integral in terms of x with the substituted version, in terms of u. Why was the second one so much easier to think about or rewrite?

In both of these examples, we got around not being able to multiply (using the distributive property) or divide (by splitting up our fraction into two with common denominators) by grouping some terms together with our substitution. Once we wrote $x+5$ as u in both of these, we were able to distribute $u^{1/3}$ across the two other terms, and we were able to divide $u-5$ by u through splitting the single fraction into two fractions. The term turn-around problem is a good one because we're turning around two things:

1. The substitution itself, by solving for u instead of x.
2. The structure of the integral, by grouping $x+5$ into one term, u, and expanding x into two terms, $u-5$. This allowed us to change how the algebra would work, making it much friendlier!

Example 7.2.1

Find the following indefinite integral:

$$\int \frac{x^2+3x-1}{x-1}dx$$

Hint. Try letting $u = x-1$ so that $du = dx$. Then we can say that $x = u+1$.

Solution

$$\begin{aligned} u = x-1 \longrightarrow x &= u+1 \\ du &= dx \\ \int \frac{x^2+3x-1}{x-1}dx &= \int \frac{(u+1)^2+3(u+1)-1}{u}du \\ &= \int \frac{u^2+5u+3}{u}du \\ &= \int u+5+\frac{3}{u}du \\ &= \frac{u^2}{2}+5u+3\ln|u|+C \\ &= \frac{(x-1)^2}{2}+5(x-1)+3\ln|x-1|+C \end{aligned}$$

There are some ways of rewriting this antiderivative family: we could try to group up all of the constant terms by multiplying everything out. Feel free to do this, but it is completely unnecessary.

This specific example is an interesting one, because we actually have a couple of different options as to how we approach it. This is true in a lot of cases: there is very rarely only a single approach to an integral that will eventually work out. Sometimes there are approaches or more techniques that are more obvious to some people, and sometimes there are approaches that seem more

easy/difficult for some people. But even still, we are often presented with many choices we could make in how we approach our integration.

Moving forward in this chapter, we'll present a whole host of strategies for how we might integrate different types of functions and how we might approach different structures that we see in the integrals we'll look at. We'll try to balance a difficult duality:

- There is rarely no single "right" way to do things! We can't summarize things with strongly worded rules like "if your function looks like this, then you have to do this to antidifferentiate."
- We would like to build some good intuition, and so having some tried-and-true strategies to fall back on will help! We can try to identify some intuitive strategies, even if they're not the only ones that will work.

All of this to simply say: we are going to present a lot of problems with a lot of solutions, and there simply isn't enough space to write out alternative approaches for each one. We will try to revisit some integrals to think about alternative strategies when we are able to, though!

Practice Problems

Exercise 7.2.1

Why do we say that u-substitution undoes the Chain Rule? Explain how the process of our variable substitution accomplishes that.

Exercise 7.2.2

For each of the following indefinite integrals, identify a useful substitution and antidifferentiate.

a. $\int \dfrac{x^2+2}{x^3+6x-5}dx$
b. $\int \sec(4x-1)\tan(4x-1)dx$
c. $\int \left(\dfrac{x+5}{x-1}\right)dx$
d. $\int (x+5)\sqrt{x-1}dx$
e. $\int (x+5)^{51}(x-1)^2dx$
f. $\int \dfrac{x+1}{\sqrt{x-3}}dx$
g. $\int \dfrac{1}{(x+2)\sqrt{x+1}}dx$

This page titled 7.2: More on \(u\) Substitution is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Peter Keep.

7.3: Manipulating Integrands

We've looked at how to use a variable substitution to antidifferentiate composite functions. We've already seen, though, that sometimes identifying and actually using a helpful substitution can be difficult to do. In this section, we want to introduce some different strategies for noticing and setting up useful substitutions in some specific instances.

Rewriting the Integrand

We're going to look at a few different examples or strategies that revolve around the same idea: we're going to reveal a reasonable function to antidifferentiate, whether its through finding a substitution or putting our function into some other recognizable form.

Example 7.3.1

For each of the following integrals, rewrite the integrand function using some algebraic manipulation, trigonometric identity, or some other strategy. Then, once the integrand function is in a friendlier form, antidifferentiate.

a. $\int \tan^2(\theta)d\theta$ **Hint**. Can you think of a trigonometric identity that can help translate the squared tangent function into some other squared trigonometric function that we recognize as the derivative of something?

b. $\int \left(\frac{x^2-9}{x+3}\right) dx$ **Hint**. Try some factoring! Can you factor and cancel?

c. $\int \left(\frac{\sqrt{x}-4}{x^2}\right) dx$ **Hint**. Split this fraction into $\int \left(\frac{\sqrt{x}x^2}{-}\frac{4}{x^2}\right) dx$. Then, can you write these two terms as power functions?

d. $\int \sec(x)dx$ **Hint**. This is a hard one an annoying one, and we'll revisit it later with a better strategy, but for now you can notice something nice happens when you multiply the numerator and denominator by $(\sec(x)+\tan(x))$:

$$\int \sec(x)\left(\frac{\sec(x)+\tan(x)}{\sec(x)+\tan(x)}\right)dx = \int \frac{\sec^2(x)+\sec(x)\tan(x)}{\sec(x)+\tan(x)}dx$$

This strategy is not intuitive, in my opinion: the nice thing to multiply seemingly comes out of nowhere!

Let's look at one more type of example, just to reiterate what we're thinking about with these rewritten functions.

Activity 7.3.1: A Negative Exponent.

Let's think about this integral:

$$\int \frac{1}{1+e^{-x}}dx$$

a. Is there any composition in this integral? Pick it out, and either explain or show that using this to guide your substitution will not be helpful.

b. What does e^{-x} mean? What does $\frac{1}{e^{-x}}$ mean?

c. Rewrite the integral, specifically focusing on the negative exponent. You should find that the function looks worse! How can you clean that up?

d. Why is this new integral set up so much better for the purpose of u-substitution? How could we tell this just by looking at the initial integral?

Example 7.3.2

Rewrite the integrand function for $\int \frac{1}{x+x^{-1}}dx$, and then integrate using an appropriate substitution. **Hint**. Try to rewrite this integral by noticing that $x^{-1} = \frac{1}{x}$. Then try to make the resulting fraction a bit nicer to look at, since it has a fraction inside of the denominator of another fraction.

Solution

$$\begin{aligned}
\int \frac{1}{x+x^{-1}}dx &= \int \frac{1}{x+\frac{1}{x}} \cdot \frac{x}{x}dx \\
&= \int \frac{x}{x^2+1}dx \\
u = x^2+1 \quad & du = 2xdx \\
\int \frac{x}{x^2+1}dx &= \frac{1}{2}\int \frac{2x}{x^2+1}dx \\
&= \frac{1}{2}\int \frac{1}{u}du \\
&= \frac{1}{2}\ln|u| + C \\
&\frac{1}{2}\ln(x^2+1) + C
\end{aligned}$$

This last example is a good one to help us transition into thinking about a whole class of functions: rational functions! As a reminder, these are just polynomials divided by polynomials.

Antidifferentiating Rational Functions

Strategies for antidifferentiating rational functions are just that: strategies. There isn't one consistent rule to use to antidifferentiate these (like there is with derivatives and the Quotient Rule), but we'll find some common tactics to apply and try to build our intuition for noticing the different kinds of structure we can have in these rational functions. All of these strategies are based around cleverly rewriting our rational functions (using some algebraic manipulations) to reveal some structure. We'll try to notice the structure, so that we know what we're trying to reveal.

Activity 7.3.2: Integrating a Rational Function Three Ways.

We're going to think about the integral:

$$\int \left(\frac{x^2+3x-1}{x-1}\right)dx$$

Let's find 3 different ways of integrating this. This is kind of misleading, since we're actually going to look at 2, since we've already used u-substitution to integrate this in Example 7.2.1.

a. Let's just notice some things about this rational function.
   - Are there any vertical asymptotes? How do you know where to find them?
   - Are there any horizontal asymptotes? How do you know that there aren't?
   - When you zoom really far out on the graph of this function, it looks like a different kind of function. What kind of function? Why is that?

b. Now we're going to rewrite the function itself: $\frac{x^2+3x-1}{x-1}$ means we're dividing (x^2+3x-1) by $(x-1)$. So let's do the division!

$$x-1\overline{)\,x^2+3x-1}$$

c. Rewrite your integral using this new version of the function. Notice that we haven't done any calculus or antidifferentiating yet. Explain why this new version of this integrand function is easier to antidifferentiate. What do you get?
d. Let's approach this integral differently. We said earlier that this function is really an "almost" linear function in disguise: when we divide the quadratic numerator by a linear denominator, we expect a linear function to be left over. In the long division, we saw this happen! We ended with a linear function and some remainder. Let's try to uncover this linear function. If we're looking to find what linear functions multiply together to get (x^2+3x-1), then we can try factoring!

$$\frac{x^2+3x-1}{x-1} = \frac{(\quad)(\quad)}{x-1}$$

In order for this factoring to be useful, we want to be able to "cancel" out the $x-1$ factor in the denominator. We're really only interested in what linear factor will multiply by $(x-1)$ to get (x^2+3x-1).

$$\frac{x^2+3x-1}{x-1} = \frac{(x-1)(\quad)}{x-1}$$

First, explain why there is no linear function factor that accomplishes this.
e. What if we're able to "almost" factor this? If there was a linear factor that multiplied by $(x-1)$ to get (x^2+3x-1), then the linear portions would multiply together to get x^2. What does this mean about the first linear term of our factor?
f. What does the constant term of our missing factor need to be? We are hoping that whatever it is can multiply by x (from $x-1$) and combine with the $-x$ (from the constant –1 multiplied by x in our missing factor) to match the $3x$ in x^2+3x-1. What is it?
g. Note that we have not factored (x^2+3x-1)! We almost did: we found two factors:

$$(x-1)(\quad) = x^2+3x+(\quad)$$

How far off is the actual polynomial that we are working with, x^2+3x-1? Write x^2+3x-1 as your two factors, plus or minus some remainder.
h. You should get the same thing that we got from using long division! Great! The rest of the integral will work the same. Before we end, though, compare this antiderivative to the one we got in Example 7.2.1. It's different. Why? Is this a problem?

This gives us a good approach for whenever division will help us rewrite our rational function as some polynomial and a remainder.

Let's look at two more rational functions: these ones won't be good candidates to use long division, but we'll try to build some intuition for why we will need to rewrite one of them to get a substitution that works.

Activity 7.3.3: Comparing Two Very Similar Integrals.

We're going to compare these two integrals:

$$\int \frac{x+2}{x^2+4x+5}dx \qquad \int \frac{2}{x^2+4x+5}dx$$

a. Describe why $u=x^2+4x+5$ is such a useful choice for the first integral, but not for the second. How do the differences in these two integrals influence this substitution, even though the denominators are the same? **Hint**. If $u=x^2+4x+5$, then $du=(2x+4)dx$ for both integrals. Why is this good for one integral but not the other?
b. Why would it be useful to have a linear substitution rule (instead of the quadratic one that we picked) for the second integral? Why would that match the structure of the numerator better?Go ahead and integrate the first integral.
c. We're going to write the denominator, x^2+4x+5 in a different way, in order to get a linear function composed into something familiar. **Complete the square** for this polynomial: that is, find some linear factor $(x+h)$ and a real number b such that $(x+h)^2+b=x^2+4x+5$. This should feel familiar, since we already haves tried to force polynomials to factor cleverly in Activity 7.3.2. **Hint**. We want to find a constant term so that $(x+\quad)^2$ gives us $x^2+4x+\quad$. Then, we can compare the quadratic to x^2+4x+5 to see how far off it is!
d. There is an intuitive substitution to pick, since we now have more obvious composition. Pick it. What kind of integral do we end up with and how do we antidifferentiate? Complete this problem!

Answer available at the end of this section.

Going forward, when you see a quadratic denominator in a rational function, what are some things you can think about and strategies you can use, based on what the rest of the function looks like? We want to summarize this a bit!

Integrating Rational Functions.

If $f(x) = \dfrac{p_n(x)}{p_m(x)}$ where $p_n(x)$ is a degree n polynomial and $p_m(x)$ is a degree m polynomial, then we can think about how we might integrate $\int f(x)dx$ based on degrees.

- If $n \geq m$ (the degree in the numerator is at least the degree in the denominator), then we can use long division to write $f(x)$ as some polynomial with degree $n-m$ and some remainder.
- If $n = m-1$ (the degree of the numerator is one less than the degree of the denominator), then we can try a u -substitution where $u = p_m(x)$, since the derivative of $p_m(x)$ is a polynomial of degree n. If this substitution works, we can antidifferentiate to get some sort of logarithm. This is not guaranteed to work, but for now (without other strategies), this is something we can think about.
- If we can reduce $f(x)$ (through some transformation or substitution) to a rational function that is a constant term divided by quadratic function (or if it already is), then we can complete the square in the denominator to get to a form that can be antidifferentiated to an inverse tangent function.

In the last point, we are referencing the strategy we found in Activity 7.3.3. We have a bit more of a general version of this strategy.

Theorem 7.3.3: Generalized Inverse Tangent Forms.

$$\int \frac{1}{u^2+a^2}du = \frac{1}{a}\tan^{-1}\left(\frac{u}{a}\right)+C$$

Proof. available at the end of this section.

We have much more to talk about with integration. From here, we can move on to more systematic strategies — ones that have some goals based on familiar things like operations that we might notice or specific variable substitutions that can be useful.

Practice Problems

Exercise 7.3.1

Use polynomial division or some clever factoring to rewrite and find the following indefinite integrals or evaluate the following definite integrals.

a. $\int \left(\dfrac{x+4}{x-3}\right) dx$
b. $\int \left(\dfrac{x^2+4}{x-4}\right) dx$
c. $\int \left(\dfrac{t^2+t+6}{t^2+1}\right) dt)$
d. $\int_{x=2}^{x=4} \left(\dfrac{x^3+1}{x-1}\right) dx$
e. $\int_{x=0}^{x=1} \left(\dfrac{x^4+1}{x^2+1}\right) dx$

Exercise 7.3.2

Complete the square in order to find the following indefinite integrals.

a. $\int \left(\dfrac{1}{x^2 - 2x + 10} \right) dx$
b. $\int \left(\dfrac{x}{x^2 + 4x + 8} \right) dx$
c. $\int \left(\dfrac{2x}{x^4 + 6x^2 + 10} \right) dx$

Exercise 7.3.3

Find the following indefinite integrals.

a. $\int \left(\dfrac{1}{x^{-1} + 1} \right) dx$
b. $\int \left(\dfrac{\sin(\theta) + \tan(\theta)}{\cos^2(\theta)} \right) d\theta$
c. $\int \left(\dfrac{1 - x}{1 - \sqrt{x}} \right) dx$
d. $\int \left(\dfrac{1}{1 - \sin^2(\theta)} \right) d\theta$
e. $\int \left(\dfrac{x^{2/3} - x^3}{x^{1/4}} \right) dx$
f. $\int \left(\dfrac{4 + x}{\sqrt{1 - x^2}} \right) dx$

Answers

Activity 7.3.3: Comparing Two Very Similar Integrals.

b.
$u = x^2 + 4x + 5 \quad du = (2x + 4)dx$

$$\begin{aligned} \int \frac{x+2}{x^2+4x+5} dx &= \frac{1}{2} \int \frac{2(x+2)}{x^2+4x+5} dx \\ &= \frac{1}{2} \int \frac{1}{u} du \\ &= \frac{1}{2} \ln|u| + C \\ &= \frac{1}{2} \ln|x^2+4x+5| + C \end{aligned}$$

c. We can use $(x+2)^2$ since $(x+2)^2 = x^2 + 4x + 4$. Then, we can write:

$$x^2 + 4x + 5 = \overbrace{(x+2)^2}^{x^2+4x+4} + 1$$

d.
$$\int \frac{4}{x^2+4x+5} dx = \int \frac{4}{(x+4)^2+1} dx$$
$$\int \frac{4}{(x+4)^2+1} dx = \int \frac{4}{u^2+1} du$$
$$= 4\tan^{-1}(u) + C$$
$$= 4\tan^{-1}(x+4) + C$$

Theorem 7.3.3: Generalized Inverse Tangent Forms.

This is really just based on a clever substitution. Once we see this specific constant over a sum of squares, we can factor out a convenient coefficient to force the denominator to look like a sum of something squared and 1.

$$\begin{aligned} \int \frac{1}{u^2+a^2} du &= \frac{1}{a^2} \int \frac{1}{\frac{u^2}{a^2} + 1} du \\ &= \frac{1}{a^2} \int \frac{1}{\left(\frac{u}{a}\right)^2 + 1} du \end{aligned}$$

Now we can let $w = \dfrac{u}{a}$ and $dw = \dfrac{1}{a} du$.

$$\begin{aligned}\frac{1}{a^2}\int \frac{1}{\left(\frac{u}{a}\right)^2+1}du &= \frac{1}{a}\int \frac{1}{\left(\frac{u}{a}\right)^2+1}\left(\frac{1}{a}\right)du \\ &= \frac{1}{a}\int \frac{1}{w^2+1}dw \\ &= \frac{1}{a}\tan^{-1}(w)+C \\ &= \frac{1}{a}\tan^{-1}\left(\frac{u}{a}\right)+C\end{aligned}$$

This strategy also can be used for other inverse trigonometric derivatives. But we will use the inverse tangent form most of all, and thus we want to outline it fully.

This page titled 7.3: Manipulating Integrands is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Peter Keep.

7.4: Integration by Parts

We've seen now that Introduction to u-Substitution is a useful technique for undoing The Chain Rule. We set up the variable substitution with the specific goal of going backwards through the Chain Rule and antidifferentiating some composition of functions.

A reasonable next step is to ask: What other derivative rules can we "undo?" What other operations between functions should we think about? This brings us to Integration by Parts, the integration technique specifically for undoing Product Rule.

Discovering the Integration by Parts Formula

Activity 7.4.1: Discovering the Integration by Parts Formula.

The product rule for derivatives says that:

$$\frac{d}{dx}(u(x)\cdot v(x)) = \square + \square$$

We know that we intend to "undo" the product rule, so let's try to reframe the product rule from a rule about derivatives to a rule about antiderivatives.

a. Antidifferentiate the product rule by antidifferentiating each side of the equation.

$$\begin{aligned}\int \left(\frac{d}{dx}(u\cdot v)\right) dx &= \int \square + \square \\ &= \int \square dx + \int \square dx\end{aligned}$$

Hint. Note that on the left side of this equation, you're antidifferentiating a derivative. What will that give you? Then, on the right side, we're just splitting up the terms of the product rule into two different integrals.

b. On the right side, we have two integrals. Since each of them has a product of functions (one function and a derivative of another), we can isolate one of them in this equation and create a formula for how to antidifferentiate a product of functions! Solve for $\int uv'dx$.

c. Look back at this formula for $\int uv'dx$. Explain how this is really the product rule for derivatives (without just undoing all of the steps we have just done).

We typically use the substitutions $du = u'dx$ and $dv = v'dx$ to rewrite the integrals.

Integration by Parts.

Suppose $u(x)$ and $v(x)$ are both differentiable functions. Then:

$$\int udv = uv - \int vdu$$

When we select the parts for our integral, we are selecting a function to be labeled u and a function to be labeled as dv. We begin with one of the pieces of the product rule, a function multiplied by some other function's derivative. It is important to recognize that we do different things to these functions: for one of them, u, we need to find the derivative, du. For the other, dv, we need to find an antiderivative, v. Because of these differences, it is important to build some good intuition for how to select the parts.

Intuition for Selecting the Parts

Activity 7.4.2: Picking the Parts for Integration by Parts.

Let's consider the integral:

$$\int x\sin(x)dx$$

We'll investigate how to set up the integration by parts formula with the different choices for the parts.

a. We'll start with selecting $u = x$ and $dv = \sin(x)dx$. Fill in the following with the rest of the pieces:

$$\begin{array}{ll} u = x & v = \square \\ du = \square & dv = \sin(x)dx \end{array}$$

b. Now, set up the integration by parts formula using your labeled pieces. Notice that the integration by parts formula gives us another integral. Don't worry about antidifferentiating this yet, let's just set up the pieces. **Hint**. $\int udv = uv - \int vdu$

c. Let's swap the pieces and try the setup with $u = \sin(x)$ and $dv = xdx$. Fill in the following with the rest of the pieces:

$$\begin{array}{ll} u = \sin(x) & v = \square \\ du = \square & dv = xdx \end{array}$$

d. Now, set up the integration by parts formula using this setup. **Hint**. $\int udv = uv - \int vdu$

e. Compare the two results we have. Which setup do you think will be easier to move forward with? Why? **Hint**. When we say we need to keep moving forward with our setup, what we mean is that we have another integral to antidifferentiate. Which one will be easier to work with: $\int(-\cos(x))dx$ or $\int\left(\frac{x^2}{2}\cos(x)\right)$?

f. Finalize your work with the setup you have chosen to find $\int x\sin(x)dx$.

What made things so much better when we chose $u = x$ compared to $dv = xdx$? We know that the new integral from our integration by parts formula will be built from the new pieces, the derivative we find from u and the antiderivative we pick from dv. So when we differentiate $u = x$, we get a constant, compared to antidifferentiating $dv = xdx$ and getting another power function, but with a larger exponent. We know this will be combined with a $\cos(x)$ function no matter what (since the derivative and antiderivatives of $\sin(x)$ only will differ in their sign). So picking the version that gets that second integral to be built from a trig function and a constant is going to be much nicer than a trig function and a power function. It was nice to pick x to be the piece that we found the derivative of!

Let's practice this comparison with another example in order to build our intuition for picking the parts in our integration by parts formula.

Activity 7.4.3: Picking the Parts for Integration by Parts.

This time we'll look at a very similar integral:

$$\int x\ln(x)dx$$

Again, we'll set this up two different ways and compare them.

a. We'll start with selecting $u = x$ and $dv = \ln(x)dx$. Fill in the following with the rest of the pieces:

$$\begin{array}{ll} u = x & v = \square \\ du = \square & dv = \ln(x)dx \end{array}$$

Hint. You're not forgetting how to antidifferentiate $\ln(x)$. This is just something we don't know yet!

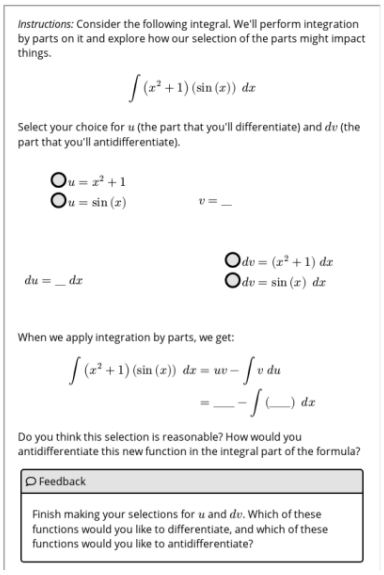
b. Ok, so here we have to swap the pieces and try the setup with $u = \ln(x)$ and $dv = xdx$, since we only know how to differentiate $\ln(x)$. Fill in the following with the rest of the pieces:

$$u = \ln(x) \quad v = \square$$
$$du = \square \quad dv = xdx$$

c. Now set up the integration by parts formula using this setup. **Hint**. $\int udv = uv - \int vdu$
d. Why was it fine for us to antidifferentiate x in this example, but not in Activity 7.4.2?
e. Finish this work to find $\int x\ln(x)dx$. **Hint**. Notice that $\left(\frac{x^2}{2}\right)\left(\frac{1}{x}\right) = \frac{x}{2}$

So here, we didn't actually get much choice. We couldn't pick $u = x$ in order to differentiate it (and get a constant to multiply into our second integral) since we don't know how to antidifferentiate $\ln(x)$ (yet: once we know how, it might be fun to come back to this problem and try it again with the parts flipped). But we can also notice that it ended up being fine to antidifferentiate x: the increased power from our power rule didn't really matter much when we combined it with the derivative of the logarithm, since the derivative of the log is also a power function! So we were able to combine those easily and actually integrate that second integral.

Let's try this again, but with a different set of functions. For this one, we'll not worry about finding derivatives and antiderivatives ourselves: we'll just play with the selection of the pieces to see how that changes things.



Standalone
Embed

Hopefully we can see a bit of strategy in the selection process! Let's try some examples.

Aside.

It is common for students to want to place functions into some sort of hierarchy or classification guidelines for choosing the parts. Some students have found that the acronym LIPET (logs, inverse trig, power functions, exponentials, and trig functions) can be a useful tool for selecting the parts. When you have two different types of functions, it might help to select u to be whichever function shows up first in that list.

This can be a helpful tool to start, but we shouldn't tie ourselves too tightly to a mnemonic like this. Instead, we should focus on building intuition beyond a memory tool like this.

Example 7.4.1

Integrate the following:

a. $\int x^2 e^x dx$ **Hint**. It doesn't matter whether we differentiate or antidifferentiate e^x, since we'll get the same thing. Let's pick $u = x^2$ so that we can differentiate it.
b. $\int 2x \tan^{-1}(x) dx$ **Hint**. We don't know how to antidifferentiate $\tan^{-1}(x)$, but we do know how to differentiate it!

Solution

a.
$$\begin{aligned} u &= x^2 \quad v = e^x \\ du &= 2x dx \quad dv = e^x dx \end{aligned}$$
$$\int x^2 e^x dx = x^2 e^x - \int 2x e^x dx$$

We need to do more integration by parts!

$$\begin{aligned} u &= 2x \quad v = e^x \\ du &= 2dx \quad dv = e^x dx \end{aligned}$$
$$\begin{aligned} \int x^2 e^x dx &= x^2 e^x - \left(2x e^x - \int 2e^x dx\right) \\ &= x^2 e^x - 2x e^x + 2e^x + C \end{aligned}$$

b.
$$\begin{aligned} u &= \tan^{-1}(x) \quad v = x^2 \\ du &= \frac{1}{x^2+1} dx \quad dv = 2x dx \end{aligned}$$
$$\begin{aligned} \int 2x \tan^{-1}(x) dx &= x^2 \tan^{-1}(x) - \int \frac{x^2}{x^2+1} dx \\ &= x^2 \tan^{-1}(x) - \int \frac{(x^2+1)-1}{x^2+1} dx \quad \text{Alternatively, use long division.} \\ &= x^2 \tan^{-1}(x) - \int 1 - \frac{1}{x^2+1} dx \\ &= x^2 \tan^{-1}(x) - x + \tan^{-1}(x) + C \end{aligned}$$

Some Flexible Choices for Parts

We're going to look at a couple of examples where we can showcase some of the flexibility we have with our choices of parts. First, we'll revisit Example 7.4.1. In this example, when we got to that second integral, we noticed that for the fraction $\frac{x^2}{x^2+1}$, we could either do some long division (since the degrees in the numerator and denominator are the same) or do some clever rewriting of the numerator. Either way, we know that this fraction is almost 1...It's really 1± some bit (in this case, the extra bit was a fraction $\frac{1}{x^2+1}$).

What if we chose our parts differently? Not the u and dv parts, though, since we still haven't figured out how to antidifferentiate $\tan^{-1}(x)$. But we get one more choice!

Once we choose u, we don't really get a separate choice for du: it's simply the derivative of u with regard to x multiplied by the differential dx. But consider our choice of dv, and the subsequent process of finding v. In a sense, yes, there's only one possible answer for the antiderivative, but we also know that we have a lot of flexibility in our choice of antiderivative. There are an infinite number of them, all in the same family of antiderivatives! We know, due to the Mean Value Theorem and then later due to Theorem 4.1.7, that there are an infinite number of antiderivatives, all differing by, at most, a constant term. So let's pick a more appropriate antiderivative!

Example 7.4.2

Integrate $\int 2x \tan^{-1}(x) dx$, this time making a more intentional choice for v. **Hint**. Note that if we pick $v = x^2 + 1$, then the second integral will be just delightful.

Solution

$$u = \tan^{-1}(x) \qquad v = x^2 + 1$$
$$du = \frac{1}{x^2+1}dx \quad dv = 2xdx$$

$$\begin{aligned}\int 2x\tan^{-1}(x)dx &= \left(x^2+1\right)\tan^{-1}(x) - \int \frac{x^2+1}{x^2+1}dx \\ &= x^2\tan^{-1}(x) + \tan^{-1}(x) - \int dx \\ &= x^2\tan^{-1}(x) + \tan^{-1}(x) - x + C\end{aligned}$$

So we get the same thing, but didn't have to think through the long division or the forced factoring. But the trade off here is that we almost have to see this coming to notice it. This flexibility doesn't always come into play for us. But we can look at a different kind of flexibility.

We've looked at integrals with both $\ln(x)$ and $\tan^{-1}(x)$. For these, and for other inverse functions specifically, we pick them to be the u part in our integration by parts problems because we don't know how do antidifferentiate them.

So let's look at $\int \ln(x)dx$, and we'll solve this integral by, specifically, differentiating $\ln(x)$ instead of antidifferentiating it.

Example 7.4.3: Antidifferentiating the Log Function.

Integrate $\int \ln(x)dx$. **Hint**. Pick $u = \ln(x)$, since we can differentiate it. What does that leave for dv?

Solution 1

$$u = \ln(x) \quad v = x$$
$$du = \frac{1}{x}dx \quad dv = dx$$
$$\begin{aligned}\int \ln(x)dx &= x\ln(x) - \int \frac{x}{x}dx \\ &= x\ln(x) - x + C\end{aligned}$$

Solution 1

An alternate approach is to use a substitution first. We're going to be using a lot of different variable names here, so let's use a t-substitution. Let $t = \ln(x)$ so that $dt = \frac{1}{x}dx$. In order to induce this derivative of the log, let's multiply by $\frac{x}{x}$ inside the integral:

$$\begin{aligned}\int \ln(x)dx &= \int \frac{x}{x}\ln(x)dx \\ &= \int x\underbrace{\ln(x)}_{t}\underbrace{\left(\frac{1}{x}\right)dx}_{dt}\end{aligned}$$
$$t = \ln(x) \longrightarrow x = e^t$$
$$\int x\ln(x)\left(\frac{1}{x}\right)dx = \int te^t dt$$

This integral can be done using the standard integration by parts!

$$u = tv = e^t$$
$$du = dtdv = e^t dt$$
$$\begin{aligned}\int te^t dt &= uv - \int vdu \\ &= te^t - \int e^t dt \\ &= te^t - e^t + C\end{aligned}$$

Now, we can substitute back to x:

$$te^t - e^t + C = x\ln(x) - x + C$$

We can use this same strategy to find antiderivatives of $\tan^{-1}(x)$, $\sin^{-1}(x)$, and eventually $\sec^{-1}(x)$.

Aside.

For $\int \sec^{-1}(x)dx$, we'll need to use this same tactic of setting $u = \sec^{-1}(x)$ and $dv = dx$, but then later on we'll need to use a technique called Trigonometric Substitution to finish the problem.

Now that we know the antiderivative family for $\ln(x)$, we can revisit the problem in Activity 7.4.3, $\int x\ln(x)dx$, and try to work through the integration by parts when $u = x$ and $dv = \ln(x)dx$.

Example 7.4.4

Integrate $\int x\ln(x)dx$.

Solution

$$\begin{aligned} u &= x & v &= x\ln(x) - x \\ du &= dx & dv &= \ln(x)dx \end{aligned}$$

$$\begin{aligned} \int x\ln(x)dx &= x(x\ln(x) - x) - \int x\ln(x) - xdx \\ &= x^2\ln(x) - x^2 - \int x\ln(x)dx + \int xdx \\ &= x^2\ln(x) - x^2 + \frac{x^2}{2} - \int x\ln(x)dx \end{aligned}$$

Note that this last integral is really recognizable: it's the one we started with! Let's "solve" this equation for that integral by adding it to both sides of our equation.

$$\begin{aligned} 2\int x\ln(x)dx &= x^2\ln(x) - \frac{x^2}{2} \\ \int x\ln(x)dx &= \frac{x^2\ln(x)}{2} - \frac{x^2}{4} + C \end{aligned}$$

Solving for the Integral

In this last example, we ended up seeing the original integral repeated when we did integration by parts. This is a useful technique, especially when we deal with functions that have a kind of "repeating" structure to their derivatives or antiderivatives. We'll look at a couple of classic integrals where we see this kind of technique employed. Let's have you explore this idea.

Activity 7.4.4: Squared Trig Functions.

Let's look at two integrals. We'll talk about both at the same time, since they're similar.

$$\int \sin^2(x)dx \qquad \int \cos^2(x)dx$$

a. What does it mean to "square" a trig function? Write these integrals in a different way, where the meaning of the "squared" exponent is more clear. What do you notice about the structure of these integrals, the operation in the integrand function? What does this mean about our choice of integration technique? **Hint**.

$$\int \sin^2(x)dx = \int \sin(x)\sin(x)dx$$
$$\int \cos^2(x)dx = \int \cos(x)\cos(x)dx$$

b. If you were to use integration by parts on these integrals, does your choice of u and dv even matter here? Why not? **Hint**. Is there a difference in the two functions being multiplied?
c. Apply the integration by parts formula to each. What do you notice?
d. Instead of applying another round of integration by parts to the resulting integral, use the Pythagorean identities to rewrite these integrals:

$$\sin^2(x) = 1 - \cos^2(x)$$
$$\cos^2(x) = 1 - \sin^2(x)$$

e. You should notice that, in your equation for the integration of $\sin^2(x)dx$, you have another copy of $\int \sin^2(x)dx$. Similarly, in your equation for the integration of $\cos^2(x)dx$, you have another copy of $\int \cos^2(x)dx$. Replace these integrals with a variable, like I (for "integral"). Can you "solve" for this variable (integral)?

Answer available at the end of this section.

This "solving for the integral" approach works well, but works best when we can see it coming. Notice that it happened here due to the repeating structure of the derivatives of the sine and cosine functions, as well as the Pythagorean identities. We can see some more examples of this in play with similar functions!

Example 7.4.5

For each of the following integrals, use integration by parts to solve.

a. $\int \sin(x)\cos(x)dx$ **Hint**. This one is pretty straightforward, since it doesn't really matter what we select as our parts. Notice, though, that this isn't the only way we can approach this! We can use u-substitution, or even rewrite this using a trigonometric identity.
b. $\int e^x \cos(x)dx$

Solution

a.
$$\begin{aligned} u &= \sin(x) & v &= \sin(x) \\ du &= \cos(x)dx & dv &= \cos(x)dx \end{aligned}$$
$$\int \sin(x)\cos(x)dx = \sin^2(x) - \int \sin(x)\cos(x)dx$$
$$2\int \sin(x)\cos(x)dx = \sin^2(x)$$
$$\int \sin(x)\cos(x)dx = \frac{\sin^2(x)}{2} + C$$

b.
$$\begin{aligned} u &= e^x & v &= \sin(x) \\ du &= e^x dx & dv &= \cos(x)dx \end{aligned}$$
$$\int e^x \cos(x)dx = e^x \sin(x) - \int e^x \sin(x)dx$$
$$\begin{aligned} u &= e^x & v &= -\cos(x) \\ du &= e^x dx & dv &= \sin(x)dx \end{aligned}$$

$$\int e^x \cos(x)dx = e^x \sin(x) - \int e^x \sin(x)dx$$
$$\int e^x \cos(x)dx = e^x \sin(x) + e^x \cos(x) - \int e^x \cos(x)dx$$
$$2\int e^x \cos(x)dx = e^x \sin(x) + e^x \cos(x)$$
$$\int e^x \cos(x)dx = \frac{e^x \sin(x) + e^x \cos(x)}{2} + C$$

Notice that we can come up with a bunch of different examples that are similar to Example 7.4.5. If we put trigonometric functions inside our integral, we'll have some options with how we approach them! We can use u-substitution, since the derivatives of trigonometric functions are other trigonometric functions. In Example 7.4.5, for instance, we could write $u = \sin(x)$ and $du = \cos(x)dx$, or even chose $u = \cos(x)$ and $du = -\sin(x)dx$.

The real issues will come when our integrand is not just a product of two trigonometric functions, but when they are products of trigonometric functions raised to exponents. We'll have some combinations of these products (which maybe makes us think about integration by parts) and composition (which points towards u-substitution). In the next section, we'll develop some strategies to deal with these kinds of integrals.

Practice Problems

Exercise 7.4.1

Explain how we build the Integration by Parts formula, as well as what the purpose of this integration strategy is.

Exercise 7.4.2

How do you choose options for u and dv? What are some good strategies to think about?

Exercise 7.4.3

Let's say that you make a choice for u and dv and begin working through the Integration by Parts strategy. How can you tell if you've made a poor choice for your parts? Can you always tell?

Exercise 7.4.4

Integrate the following.

a. $\int 3x \sin(x)dx)$
b. $\int 5xe^x dx$
c. $\int x^2 e^{-x} dx$
d. $\int x^2 \ln(x)dx$
e. $\int x^2 \cos(x)dx$
f. $\int x^3 e^{-x} dx$
g. $\int x \sin(x) \cos(x)dx$
h. $\int e^x \sin(x)$
i. $\int \sin^{-1}(x)dx$
j. $\int \tan^{-1}(x)dx$

Exercise 7.4.5

Evaluate the following definite integrals.

a. $\int_{x=1}^{x=e} x\ln(x)dx)$
b. $\int_{x=0}^{x=\pi/4} x\cos(2x)dx$
c. $\int_{x=0}^{x=\ln(5)} xe^x dx$

Exercise 7.4.6

In this problem, we'll consider the integral $\int \sin^2(x)dx$. We'll integrate this in two different ways!

a. We know that:

$$\int \sin^2(x)dx = \int \sin(x)\sin(x)dx$$

Use the Integration by Parts strategy, and especially note that you can solve for the integral (Solving for the Integral).

b. We can use a trigonometric identity to rewrite the integral:

$$\sin^2(x) = \frac{1-\cos(2x)}{2}$$

So we have:

$$\int \sin^2(x)dx = \int \frac{1}{2} - \frac{1}{2}\cos(2x)dx$$

Use u-substitution.

c. Were your answers the same or different? Should they be the same? Why or why not? Are they connected somehow?

Exercise 7.4.7

For these next problems, we'll use $x = u^2$ and $dx = 2udu$ to substitute into the integral as written. Then use Integration by Parts.

a. $\int \sin(\sqrt{x})dx$
b. $\int e^{\sqrt{x}}dx$

Answers

Activity 7.4.4: Squared Trig Functions.

c. For the integral $\int \sin^2(x)dx$:

$$\begin{aligned} u &= \sin(x) & v &= -\cos(x) \\ du &= \cos(x)dx & dv &= \sin(x)dx \\ \int \sin^2(x)dx & & &= -\sin(x)\cos(x) + \int \cos^2(x)dx \end{aligned}$$

For the integral $\int \cos^2(x)dx$:

$$\begin{aligned} u &= \cos(x) & v &= \sin(x) \\ du &= -\sin(x)dx & dv &= \cos(x)dx \\ \int \cos^2(x)dx &= \sin(x)\cos(x) + \int \sin^2(x)dx \end{aligned}$$

d.

e. For $\int \sin^2(x)dx$:

$$\int \sin^2(x)dx = -\sin(x)\cos(x) + \int 1 - \sin^2(x)dx$$
$$I = -\sin(x)\cos(x) + \int 1dx - \underbrace{\int \sin^2(x)dx}_{I}$$
$$I = -\sin(x)\cos(x) + x - I$$
$$2I = -\sin(x)\cos(x) + x$$
$$I = \frac{-\sin(x)\cos(x) + x}{2} + C$$

So we end up with:

$$\int \sin^2(x)dx = \frac{x - \sin(x)\cos(x)}{2} + C$$

For $\int \cos^2(x)dx$:

$$\int \cos^2(x)dx = \sin(x)\cos(x) + \int 1 - \cos^2(x)dx$$
$$I = \sin(x)\cos(x) + \int 1dx - \underbrace{\int \cos^2(x)dx}_{I}$$
$$I = \sin(x)\cos(x) + x - I$$
$$2I = \sin(x)\cos(x) + x$$
$$I = \frac{\sin(x)\cos(x) + x}{2} + C$$

So we end up with:

$$\int \cos^2(x)dx = \frac{x + \sin(x)\cos(x)}{2} + C$$

This page titled 7.4: Integration by Parts is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Peter Keep.

7.5: Integrating Powers of Trigonometric Functions

Let's remind ourselves of two example problems that we've done in the past.

In Example 5.6.3, we performed a u-substitution, but needed to work to rewrite our whole integrand in terms of u. Specifically, we found that in the numerator, there was an x^3, but $du = 2xdx$. We were substituting out a linear function of x in the numerator, but the actual function was cubic. This wasn't a problem: we rewrote $x^3 = x^2 \cdot x$, and noticed that the extra x^2 was able to be substituted, since we could rewrite our substitution rule: we noted that $u = x^2 + 1$ is equivalent to $x^2 = u - 1$. This meant that even though we had an extra factor of x^2 "in" the part that we were using for substituting in the differential du, we were still able to translate the whole function to be written in terms of u.

Then, more recently, in Example 7.4.5, we noted that we could use a mix of methods to integrate this:

$$\int \sin(x)\cos(x)dx$$

On one hand, we can look at the structure of the integrand and notice that we have a product of two functions! Integration by parts was a fine strategy to employ, and that's what we did in the example. On the other hand, we noticed that since we have this function-derivative pairing, a u-substitution also was appropriate.

In this section, we'll explore more combinations of trigonometric functions and build a strategy for anti-differentiating them that includes some ideas from both of these previous examples.

Building a Strategy for Powers of Sines and Cosines

Activity 7.5.1: Compare and Contrast.

Let's do a quick comparison of two integrals, keeping the above examples in mind. Consider these two integrals:

$$\int \sin^4(x)\cos(x)dx \qquad \int \sin^4(x)\cos^3(x)dx$$

a. Consider the first integral, $\int \sin^4(x)\cos(x)dx$. Think about and set up a good technique for anti-differentiating. Without actually solving the integral, explain why this technique will work.
b. Now consider the second integral, $\int \sin^4(x)\cos^3(x)dx$. Does the same integration strategy work here? What happens when you apply the same thing?
c. We know that $\sin(x)$ and $\cos(x)$ are related to each other through derivatives (each is the derivative of the other, up to a negative). Is there some other connection that we have between these functions? We might especially notice that we have a $\cos^2(x)$ left over in our integral. Can we write this in terms of $\sin(x)$, so that we can write it in terms of u?
d. Why would this strategy not have worked if we were looking at the integrals $\int \sin^4(x)\cos^2(x)dx$ or $\int \sin^4(x)\cos^4(x)dx$? What, specifically, did we need in order to use this combination of substitution and trigonometric identity to solve the integral?

Integrating Powers of Sine and Cosine.

For integrals in the form $\int \sin^p(x)\cos^q(x)dx$ where p and q are real number exponents:

- If q, the exponent on $\cos(x)$ is odd, we should use $u = \sin(x)$ and $du = \cos(x)dx$). Then, we can apply the Pythagorean Identity $\cos^2(x) = 1 - \sin^2(x)$.
- If p, the exponent on $\sin(x)$ is odd, we should use $u = \cos(x)$ and $du = -\sin(x)dx$. Then, we can apply the Pythagorean Identity $\sin^2(x) = 1 - \cos^2(x)$.
- If both p and q are even, we can either use Integration by Parts or use the following power-reducing trigonometric identities:

$$\sin^2(x) = \frac{1-\cos(2x)}{2} = \frac{1}{2} - \frac{\cos(2x)}{2}$$
$$\cos^2(x) = \frac{1+\cos(2x)}{2} = \frac{1}{2} + \frac{\cos(2x)}{2}$$

A strange note, here, is that we typically pick our u-substitution based on looking to see a suitable candidate for u: we look for functions that are composed "inside" of other functions or we look for a function whose derivative is in the integral (the "function-derivative pair" that we talk about in Section 5.6). Here, though, we're selecting our substitution based on du: we're looking to see which function we can set aside one copy of for the differential, and then have an even power left over so that we can apply the Pythagorean Identity to translate the rest.

Example 7.5.1

For each of the following, identify an appropriate substitution, make a note of which trigonometric identity you'll use, and then integrate.

a. $\int \sin^5(x)\cos^6(x)dx$ **Hint**. Notice that the exponent on $\sin(x)$ is odd: if you let $u = \cos(x)$, you'll end up with $\sin^4(x)$ left over in your integral, and you can write it as $(\underbrace{1-\cos^2(x)}_{\sin^2(x)})^2$.

b. $\int \sin^{4/3}(x)\cos^5(x)dx$ **Hint**. This one seems scary at first, because of the fraction exponent. Notice, though, that you have no hope of converting fractional exponents of sine functions into cosine functions easily. So, pick $u = \sin(x)$ and try to convert any remaining cosines using $\cos^2(x) = 1 - \sin^2(x)$.

c. $\int \sin^3(x)\cos^9(x)dx$ **Hint**. You get a choice here! Both exponents are odd, so picking either function as u will leave you with an even exponent on the other function on which to use the Pythagorean Identity. Is there a choice of u that will be easier than the other choice?

Building a Strategy for Powers of Secants and Tangents

Activity 7.5.2: Compare and Contrast (Again).

We're going to do another Compare and Contrast, but this time we're only going to consider one integral:

$$\int \sec^4(x)\tan^3(x)dx$$

We're going to employ another strategy, similar to the one for Integrating Powers of Sine and Cosine.

a. Before you start thinking about this integral, let's build the relevant version of the Pythagorean Identity that we'll use. Our standard version of this is:

$$\sin^2(x) + \cos^2(x) = 1 \tag{7.5.1}$$

Since we want a version that connects $\tan(x)$, which also is written as $\frac{\sin(x)}{\cos(x)}$, with $\sec(x)$, or $\frac{1}{\cos(x)}$, let's divide everything in the Pythagorean Identity by $\cos^2(x)$:

$$\frac{\sin^2(x)}{\square} + \frac{\cos^2(x)}{\square} = \frac{1}{\square}$$
$$\square + \square = \square$$

b. Now start with the integral. We're going to use two different processes here, two different u-substitutions. First, set $u = \tan(x)$. Complete the substitution and solve the integral. **Hint**. Here $du = \sec^2(x)dx$. We'll also use $\sec^2(x) = \tan^2(x) + 1$.

c. Now try the integral again, this time using $u = \sec(x)$ as your substitution. **Hint**. Now $du = \sec(x)\tan(x)dx$, and we'll use the same Pythagorean Identity, just rewritten as $\tan^2(x) = \sec^2(x) - 1$.

d. For each of these integrals, why were the exponents set up just right for u-substitution each time? How does the structure of the derivatives of each function play into this? **Hint**. Notice we had an even exponent on the $\sec(x)$ function, but an odd exponent on the $\tan(x)$ function.

e. Which substitution would be best for the $\int \sec^4(x)\tan^4(x)dx$. Why?
f. Which substitution would be best for the $\int \sec^3(x)\tan^3(x)dx$. Why?

Answer available at the end of this section.

Integrating Powers of Secant and Tangent.

For integrals in the form $\int \sec^p(x)\tan^q(x)dx$ where p and q are real number exponents:

- If q, the exponent on $\tan(x)$, is odd, we can use $u = \sec(x)$ and $du = \sec(x)\tan(x)dx$. Then we can apply the Pythagorean Identity $\tan^2(x) = \sec^2(x) - 1$.
- If p, the exponent on $\sec(x)$, is even, we can use $u = \tan(x)$ and $du = \sec^2(x)dx$. Then we can apply the Pythagorean Identity $\sec^2(x) = \tan^2(x) + 1$.
- If p is odd and q is even, we can use Integration by Parts.

Example 7.5.2

For each of the following, identify an appropriate substitution, make a note of which trigonometric identity you'll use, and then integrate.

a. $\int \tan^3(x)\sec^8(x)dx$ **Hint**. You have some choices here! If you use $u = \sec(x)$, then you'll end up with a remaining $\tan^2(x)$ to convert using a Pythagorean Identity. Alternatively, if you use $u = \tan(x)$, you'll end up needing to convert the remaining $\sec^6(x) = \left(\sec^2(x)\right)^3$.
b. $\int \tan^{-5/2}(x)\sec^6(x)dx$ **Hint**. Another fraction exponent that is pushing us towards using $u = \tan(x)$. There will still be a $\sec^4(x)$ to convert after substituting in du.
c. $\int \tan^5(x)\sec^3(x)dx$ **Hint**. The odd exponent on tangent is fine, since we can use $u = \sec(x)$ to leave us with a $\tan^4(x)$ to convert.

Practice Problems

Exercise 7.5.1

For an integral $\int \sin^a(x)\cos^b(x)dx$, how do you know whether to use $u = \sin(x)$ or $u = \cos(x)$ as the substitution?

Exercise 7.5.2

For an integral $\int \tan^a(x)\sec^b(x)dx$, how do you know whether to use $u = \tan(x)$ or $u = \sec(x)$ as the substitution?

Exercise 7.5.3

Integrate the following.

a. $\int \sin^3(x)\cos^2(x)dx$
b. $\int \sin^2(x)\cos^3(x)dx$
c. $\int \sin^3(x)\cos^3(x)dx$
d. $\int \tan^4(x)\sec^4(x)dx$
e. $\int \tan^3(x)\sec^3(x)dx)$
f. $\int \tan^3(x)\sec^4(x)dx)$

Exercise 7.5.4

Integrate the following.

a. $\int \sin^{3/4}(x)\cos^5(x)dx)$
b. $\int \tan^5(x)\sec^{-1/2}(x)dx$
c. $\int \sin^{3/4}(x)\cos^5(x)dx$
d. $\int \tan^5(x)\sec^{-1/2}(x)dx$

Exercise 7.5.5

Consider the integral $\int \sin^2(x)dx$.

a. Use the trigonometric identity:

$$\sin^2(x) = \frac{1-\cos(2x)}{2}$$

to integrate.
b. Use integration by parts to integrate.
c. Which of these techniques do you think was easier to implement and use? Why is that?

Exercise 7.5.6

Consider the integral $\int \cos^4(x)dx$.

a. Use the trigonometric identity:

$$\cos^2(x) = \frac{1+\cos(2x)}{2}$$

to integrate.
b. Use integration by parts to integrate.
c. Which of these techniques did you prefer? Why?

Exercise 7.5.7

Integrate the following integrals.

a. $\int \tan^2(x)dx$
b. $\int \sec^3(x)dx$
c. $\int \tan^5(x)dx$
d. $\int \sin^5(x)dx$

Answers

Activity 7.5.2: Compare and Contrast (Again).

a. $\tan^2(x)+1=\sec^2(x)$

$$\begin{aligned}\int \sec^4(x)\tan^3(x)dx &= \int \underbrace{\sec^2(x)}_{\tan^2(x)+1}\tan^3(x)\sec^2(x)dx \\ &= \int (u^2+1)\,u^3du \\ &= \int u^5+u^3du \\ &= \frac{u^6}{6}+\frac{u^4}{4}+C \\ &= \frac{\tan^6(x)}{6}+\frac{\tan^4(x)}{4}+C\end{aligned}$$

b.

c.

$$\begin{aligned}\int \sec^4(x)\tan^3(x)dx &= \int \sec^3(x)\underbrace{\tan^2(x)}_{\sec^2(x)-1}\sec(x)\tan(x)dx \\ &= \int u^3\left(u^2-1\right)du \\ &= \int u^5 - u^3 du \\ &= \frac{u^6}{6} - \frac{u^4}{4} + C \\ &= \frac{\sec^6(x)}{6} - \frac{\sec^4(x)}{4} + C\end{aligned}$$

This page titled 7.5: Integrating Powers of Trigonometric Functions is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Peter Keep.

7.6: Trigonometric Substitution

We're going to look at an integral that requires a variable substitution, but our goal for the substitution will be a bit different. We're going to focus on the structure of our integrand function, but we won't be focusing on composition. Instead, we're going to focus on some trigonometric identities that we've used already:

$$1-\sin^2(\theta)=\cos^2(\theta)$$
$$\sec^2(\theta)-1=\tan^2(\theta)$$
$$\tan^2(\theta)+1=\sec^2(\theta)$$

Activity 7.6.1: Difference of Squares.

Consider the integral:

$$\int \sqrt{1-x^2}dx$$

a. First, convince yourself that a normal u-substitution will not be an effective strategy for integration in this case. Why not? **Hint**. There is composition in this integrand function, but what part of our u-substitution is missing?
b. Second, convince yourself that $\sqrt{1-x^2}\neq\sqrt{1}-\sqrt{x^2}$. Why can we not distribute roots across sums and differences like this? When can we "distribute" roots across multiple things? **Hint**. A root is really an exponent: why can't we distribute exponents across sums and differences? Try this with $(1-x^2)^2$, and convince yourself that this isn't $1^2-(x^2)^2$. Notice, though, that something like $(4x^2)^2=4^2(x^2)^2$, and $\sqrt{4x^2}$ could act similarly.
c. Our goal, then, is to utilize a substitution (using trigonometric functions) to somehow transform this difference of squared terms under the square root into a single product of squared things under the square root. Which trigonometric identities from our list of them above utilize differences of thing squared, and equate them to a single term? Can you use the order of the subtraction to help guide which substitution we should use? **Hint**. We want to pick some trigonometric function $T(\theta)$ so that when we let $x=T(\theta)$, we end up with one of the Pythagorean Identities above. Is there something we can put in for x such that $1-x^2$ becomes $1-(T(\theta))^2$ in a convenient way?
d. When we do a variable substitution in an integral, we are not only finding a way of transforming x to be in terms of some other variable (in this case, θ). We also need to transform the differential, dx. Based on your substitution of $x=T(\theta)$, what is dx? **Hint**. We'll use $dx=T'(\theta)d\theta$. What is $T'(\theta)$ for you?
e. Perform your substitution! Use your substitution $x=T(\theta)$ and $dx=T'(\theta)d\theta$. Note that we have picked this substitution with a very specific goal: we are hoping to notice a Pythagorean Identity. After you have performed your substitution, apply the relevant Pythagorean Identity to the radicand: the bit of our function underneath the radical or root. What integral are we left with (in terms of θ)? **Hint**. Since we are using $x=\sin(\theta)$, we should notice $1-\sin^2(\theta)$ underneath the square root.

This new integral is something we can antidifferentiate now! We already have done this one in Activity 7.4.4 Squared Trig Functions. So we can end up with:

$$\begin{aligned}\int \sqrt{1-x^2}dx &= \int \cos^2(\theta)d\theta \\ &= \frac{\theta+\sin(\theta)\cos(\theta)}{2}+C\end{aligned}$$

It is up to us, now, to translate this antiderivative family to be written in terms of x. We can utilize our substitution to do this, but let's first think about how this variable substitution works a bit more.

Answer available at the end of this section.

Another Type of Variable Substitution

We're going to employ another variable substitution, in the same way that we use u-substitution. The main difference is the goal: we're going to select our substitution not based on uncovering the composition in our function (like in u-substitution).

Instead, we'll focus on selecting a trigonometric function in order to utilize the relevant Pythagorean identity to rewrite our sum or difference of squares.

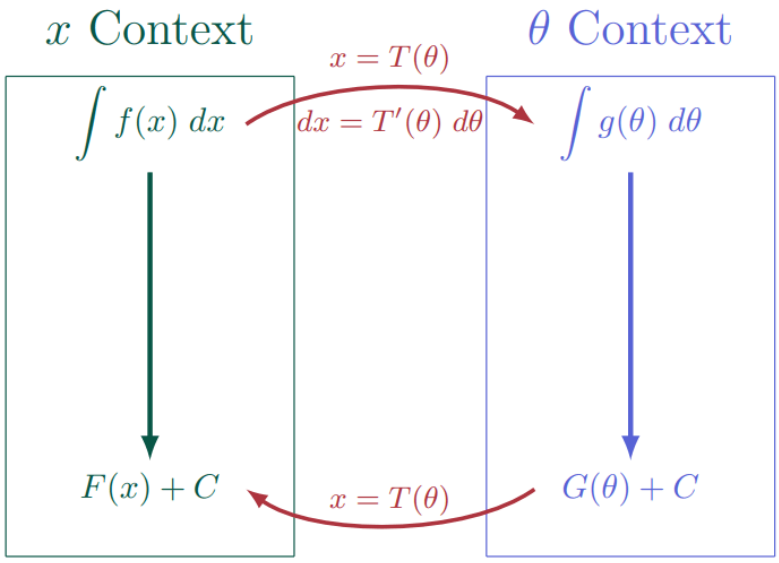


Figure 7.6.1 General idea of how this variable substitution works.

Ok, but how do we choose which trigonometric function to use in our substitution? Since we're focusing on sums or differences of squared terms, we can think of the different arrangements of terms, connect them with different Pythagorean Identities, and set up some strategies for picking a trigonometric substitution.

$$1 - x^2 \longleftrightarrow 1 - \sin^2(\theta) = \cos^2(\theta)$$
$$x^2 - 1 \longleftrightarrow \sec^2(\theta) - 1 = \tan^2(\theta)$$
$$x^2 + 1 \longleftrightarrow \tan^2(\theta) + 1 = \sec^2(\theta)$$

We can note that the sum is commutative, so we can treat $1 + x^2$ the same way that we treat $x^2 + 1$.

We'll also notice that the constant term can differ: we can scale our Pythagorean identities by some constant easily to make sure that they match.

$$a^2 - x^2 \longleftrightarrow a^2 - (a\sin(\theta))^2 = (a\cos(\theta))^2$$
$$x^2 - a^2 \longleftrightarrow (a\sec(\theta))^2 - a^2 = (a\tan(\theta))^2$$
$$x^2 + a^2 \longleftrightarrow (a\tan(\theta))^2 + a^2 = (a\sec(\theta))^2$$

This can be confusing, and we want to keep thinking about how we might recognize these structures to pick a substitution. Yes, we can recognize these Pythagorean identities. We can rely on the order of subtraction or noticing addition. But we also can think about this geometrically. The Pythagorean Identities come from the Pythagorean Theorem, relating the squared lengths of the sides of a right triangle together. Let's visualize our substitutions geometrically.

We'll consider three triangles, each with side lengths of x and a. The third side length will vary between $\sqrt{x^2 - a^2}$, $\sqrt{a^2 - x^2}$, and $\sqrt{x^2 + a^2}$ (or the equivalent $\sqrt{a^2 + x^2}$) based on which length is representing the hypotenuse.

Activity 7.6.2: Trig Substitution Geoemtry.

We're going to consider three triangles, and we're going to fill in side lengths. In each of these, we'll assume that the lengths x and a are real numbers and are positive.

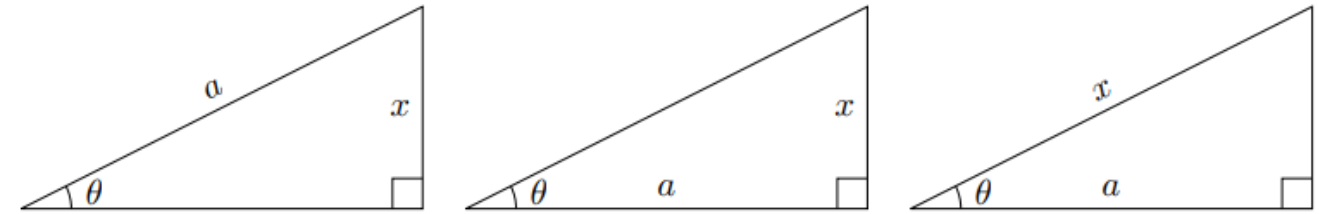


Figure 7.6.2 Three triangles to guide our trigonometric substitutions.

a. Use the Pythagorean theorem to label the missing side length in each of the three triangles.
b. For each triangle, explain how you can tell which side length represents the hypotenuse when you see the lengths x, a, and then the missing lengths you found above: $\sqrt{x^2-a^2}$, $\sqrt{a^2-x^2}$, or $\sqrt{x^2+a^2}$. **Hint**. We know that the hypotenuse is the longest side length in a triangle. Just based on the square root length, how can you tell which length is longest?
c. For each triangle, find a trigonometric function of θ that connects lengths x and a to each other. Solve each for x to reveal the relevant substitution.
d. For each substitution, find the corresponding substitution for the differential, dx.

Answer available at the end of this section.

This gives us a nice strategy for substitution!

Trigonometric Substitution.

We have three (typical) ways of using trigonometric substitution to transform a sum or difference of squared terms into a product of squares.

- For an integral containing (a^2-x^2), we can use the following triangle to build our substitution:

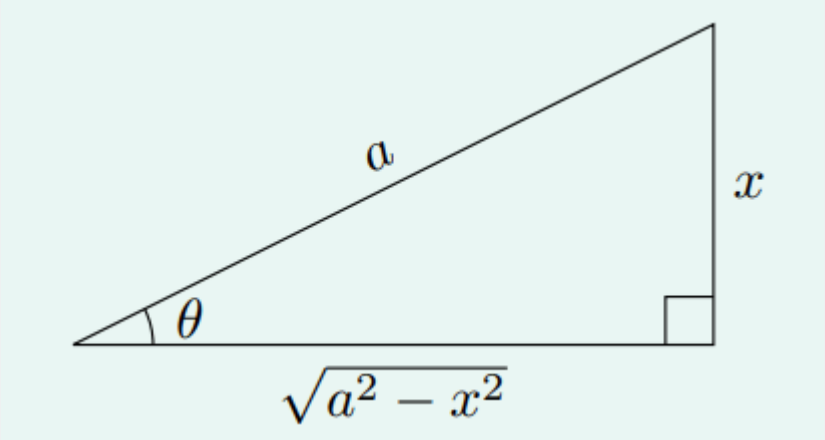


Figure 7.6.3

This results in using the following substitution and identity.

$$\begin{aligned} x &= a\sin(\theta) \\ dx &= a\cos(\theta)d\theta \\ a^2-(a\sin(\theta))^2 &= (a\cos(\theta))^2 \end{aligned}$$

- For an integral containing (x^2-a^2), we can use the following triangle to build our substitution:

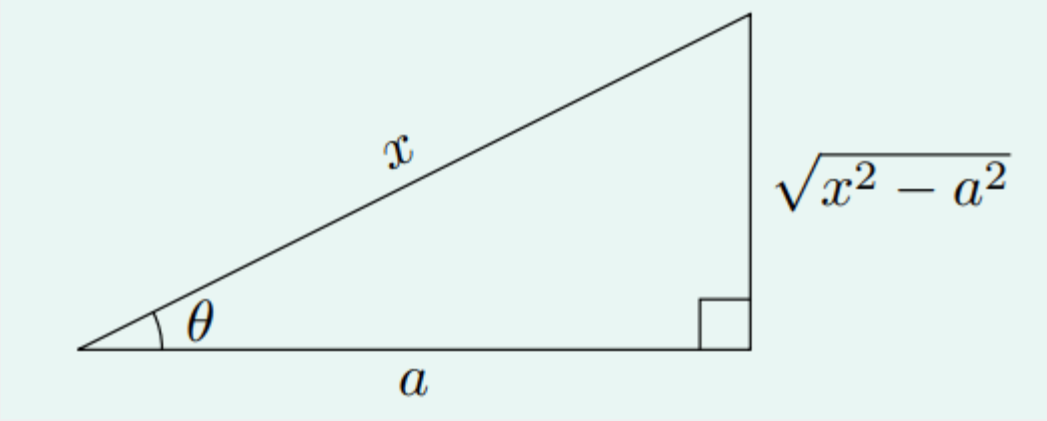


Figure 7.6.4

This results in using the following substitution and identity.

$$\begin{aligned} x &= a\sec(\theta) \\ dx &= a\sec(\theta)\tan(\theta)d\theta \\ (a\sec(\theta))^2 - a^2 &= (a\tan(\theta))^2 \end{aligned}$$

- For an integral containing $(x^2 + a^2)$, we can use the following triangle to build our substitution:

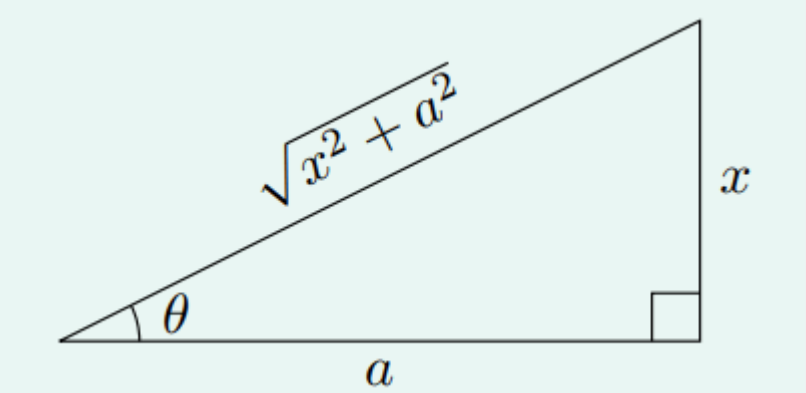


Figure 7.6.5

This results in using the following substitution and identity.

$$\begin{aligned} x &= a\tan(\theta) \\ dx &= a\sec^2(\theta)d\theta \\ (a\tan(\theta))^2 + a^2 &= (a\sec(\theta))^2 \end{aligned}$$

Two things to note before we move on:

1. There are really 6 main trigonometric substitutions. If you go back to Activity 7.6.2 and place the angle θ in the opposite corner of the triangle, the substitutions you build will all be using the "co-functions": cosine, cosecant, and cotangent. Each of these has a very similar structure with regard to derivatives (for the differential substitution) and Pythagorean Identities. Each is equivalent to the respective sine, secant, and tangent substitutions. We often choose to use sine, secant, and tangent just due to familiarity.
2. We can use the triangle as a kind of key for our substitution! After antidifferentiating, we have some antiderivative family written in terms of an angle θ: we can use the triangle to substitution trigonometric functions of θ to be written in terms of x.

Example 7.6.5

We can finish the substitution we started in Activity 7.6.1. We used the substitution $x = \sin(\theta)$, but we now can construct the relevant triangle. Since we were hoping so use a substitution to rewrite the difference of squares, $\sqrt{1-x^2}$, we had the following triangle:

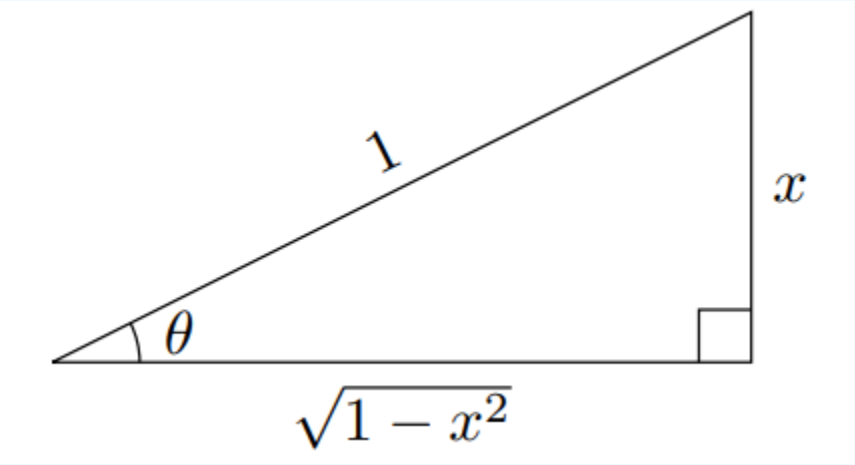


Figure 7.6.7 Substitution used in Activity 7.6.1.

We can see that $\sin(\theta) = \frac{x}{1}$ or $x = \sin(\theta)$, which was our substitution. But we also were left with the following antiderivative:

$$\begin{aligned}\int \sqrt{1-x^2}dx &= \int \cos^2(\theta)d\theta \\ &= \frac{\theta + \sin(\theta)\cos(\theta)}{2} + C\end{aligned}$$

Now we can substitute that antiderivative! We can see from our triangle that $\cos(\theta) = \sqrt{1-x^2}$, $\sin(\theta) = x$ (this also was our original substitution anyways), and we also can invert our substitution to get $\theta = \sin^{-1}(x)$.

$$\begin{aligned}\int \sqrt{1-x^2}dx &= \int \cos^2(\theta)d\theta \\ &= \frac{\theta + \sin(\theta)\cos(\theta)}{2} + C \\ &= \frac{1}{2}\left(\sin^{-1}(x) + x\sqrt{1-x^2}\right) + C\end{aligned}$$

Activity 7.6.6: Practicing Trigonometric Substitution.

Let's look at three integrals, and practice the kind of thinking we'll need to use to apply trigonometric substitution to them.

1. $\int \frac{\sqrt{x^2-9}}{x_2}dx$
2. $\int \frac{}{(4-x^2)^{3/2}}dx$
3. $\int \frac{1}{x^2\sqrt{x^2+1}}dx$

For each integral, do the following:

a. Identify the term (or terms) that signify that trigonometric substitution might be a reasonable strategy. **Hint**. In each case, you're looking for some sum or difference of squared terms, normally (but not always) nested inside of some square root. It also might be nice to rewrite the second integral in order to notice the root:

$$\int \frac{2}{(4-x^2)^{3/2}}dx = \int \frac{2}{\left(\sqrt{4-x^2}\right)^3}dx$$

b. Use that portion of the integral to compare three side lengths of a triangle. Which one is the largest (and so must represent the length of the hypotenuse)? **Hint**. 1>For the side lengths x, 3, and $\sqrt{x^2-9}$, which must be the largest? Think about $\sqrt{x^2-9}$ and what we can learn from it. 2> For the side lengths x, 2, and $\sqrt{4-x^2}$, which must be the largest? Think about $\sqrt{4-x^2}$ and what we can learn from it. 3> For the side lengths x, 1, and $\sqrt{x^2+1}$, which must be the largest? Think about $\sqrt{x^2+1}$ and what we can learn from it.
c. Construct the triangle, label an angle θ, and use a trigonometric function to connect the two single-term side lengths. (Feel free to change the angle you label in order to use the sine, secant, or tangent functions instead of their co-functions).
d. Define your substitution (for both x and the differential dx), and identify the Pythagorean Identity that will be relevant for the integral.
e. Substitute and antidifferentiate! **Hint**.

$$\int \frac{\sqrt{x^2-9}}{x}dx = \int \frac{\sqrt{9\sec^2(\theta)-9}}{3\sec(\theta)}(3\sec(\theta)\tan(\theta))d\theta$$

Also note that $\sqrt{9\sec^2(\theta)-9} = \sqrt{9\tan^2(\theta)}$.

$$\int \frac{2}{(4-x^2)^{3/2}}dx = \int \frac{2}{\left(\sqrt{4-4\sin^2(\theta)}\right)^3}(2\cos(\theta))d\theta$$

Also note that $\sqrt{4-4\sin^2(\theta)} = \sqrt{4\cos^2(\theta)}$.

$$\int \frac{1}{x^2\sqrt{x^2+1}}dx = \int \frac{1}{\tan^2(\theta)\sqrt{\tan^2(\theta)+1}}\left(\sec^2(\theta)\right)d\theta$$

Also note that $\sqrt{\tan^2(\theta)+1} = \sqrt{\sec^2(\theta)}$.

f. Use your triangle to substitute your antiderivative back in terms of x.

Answer available at the end of this section.

Trigonometric substitution is a pretty involved technique! Setting up the substitution is definitely not trivial. Because our substitution involves trigonometric functions, we end up with integrals of trigonometric functions that we then have to work to antidifferentiate. And substituting back to x relies on us having set up a robust substitution strategy from the beginning.

It can sometimes seem like this strategy is barely relevant: the goal of it is so focussed on the specific structure of the Pythagorean identities, and these might not feel very present.

A friend of mine, though, says that once we start recognizing sums and differences of squares as being connected to Pythagoras, it's hard to not see them.

For instance, we can go back to Theorem 7.3.3 and see the sum of squares in the denominator. Instead of doing any tricky factoring to get the u-substitution to work, we could try a trigonometric substitution and get the same thing!

Another friend of mine says that trigonometric substitution only exists so that we can evaluate arc length integrals (Integrals for Evaluating the Length of a Curve).

Whatever the case, this new substitution strategy should, at the very least, generalize the concept of a variable substitution in an integral to show that we can define these for a variety of purposes, all based on the kinds of structures that we're seeing in the integrand function itself.

Practice Problems

Exercise 7.6.1

Explain how trigonometric substitution helps to convert sums or differences of squares to products of squares. Why is this helpful? When is it helpful?

Exercise 7.6.2

Draw a right triangle with $\sqrt{x^2-4}$ as one of the non-hypotenuse side lengths. What is the length of the hypotenuse? What about the other side length? What would be an appropriate substitution for an integral containing $\sqrt{x^2-4}$?

Exercise 7.6.3

Draw a right triangle with $\sqrt{4-x^2}$ as one of the non-hypotenuse side lengths. What is the length of the hypotenuse? What about the other side length? What would be an appropriate substitution for an integral containing $\sqrt{4-x^2}$?

Exercise 7.6.4

Draw a right triangle with $\sqrt{x^2+4}$ as the hypotenuse. What are the lengths of the other two sides? What would be an appropriate substitution for an integral containing $\sqrt{x^2+4}$?

Exercise 7.6.5

Integrate the following using an appropriate trigonometric substitution.

a. $\int \frac{x^2}{\sqrt{16-x^2}}dx$

b. $\int \frac{\sqrt{1-x^2}}{x^2}dx$

c. $\int \frac{1}{(9x^2+1)^{3/2}}dx$

d. $\int \frac{\sqrt{x^2-1}}{x}dx$

e. $\int \sqrt{49-x^2}dx$

f. $\int \frac{1}{x\,(x^2-1)^{3/2}}dx(\text{ for } x>1)$

g. $\int \frac{x^3}{\sqrt{4+x^2}}dx$

h. $\int \frac{x^2}{(x^2+81)^2}dx$

Exercise 7.6.6

Complete the square and then integrate

a. $\int \frac{1}{x^2-8x+62}dx$

b. $\int \frac{x^2-8x+16}{(-x^2+8x+9)^{3/2}}dx$

Answers

Activity 7.6.1: Difference of Squares.

e.
$$\begin{aligned} x = \sin(\theta) dx = \cos(\theta) d\theta \\ \int \sqrt{1-x^2} dx &= \int \sqrt{\frac{1-\sin^2(\theta)}{\cos^2(\theta)}} \cos(\theta) d\theta \\ &= \int \sqrt{\cos^2(\theta)} \cos(\theta) d\theta \\ &= \int \cos(\theta) \cos(\theta) d\theta \\ &= \int \cos^2(\theta) d\theta \end{aligned}$$

Activity 7.6.2: Trig Substitution Geoemtry.

b. If one of the side lengths is $\sqrt{x^2-a^2}$, then we know that $x > a$ (otherwise the square root is a non-real number). We also know that $x > \sqrt{x^2-a^2}$ (because $a > 0$). This means that$x$ is the length of the hypotenuse.

If one of the side lengths is $\sqrt{a^2-x^2}$, then we know that $a > x$ (otherwise the square root is a non-real number). We also know that $a > \sqrt{a^2-x^2}$ (because $x > 0$). This means that a is the length of the hypotenuse.

If one of the side lengths is $\sqrt{x^2+a^2}$, then we know that $\sqrt{x^2+a^2} > a$ and $\sqrt{x^2+a^2} > x$ (because x, $a > 0$). This means that $\sqrt{x^2+a^2}$ is the length of the hypotenuse.

Activity 7.6.6: Practicing Trigonometric Substitution.

b.
1. We know that $x > 3$, since $x^2 - 9 > 0$. Similarly, we know that $x > \sqrt{x^2-9}$. So x has to be the length of the hypotenuse.
2. We know that $2 > x$, since $4 - x^2 > 0$. Similarly, we know that $2 > \sqrt{4-x^2}$. So 2 has to be the length of the hypotenuse.
3. We know that $\sqrt{x^2+1} > x$, since we are adding 1 to x^2 under the square root. This also is the reason that $\sqrt{x^2+1} > 1$ (we are adding x^2 under the root). So $\sqrt{x^2+1}$ has to be the length of the hypotenuse.

c.
1. For the integral $\int \frac{\sqrt{x^2-9}}{x} dx$, we have the following triangle.

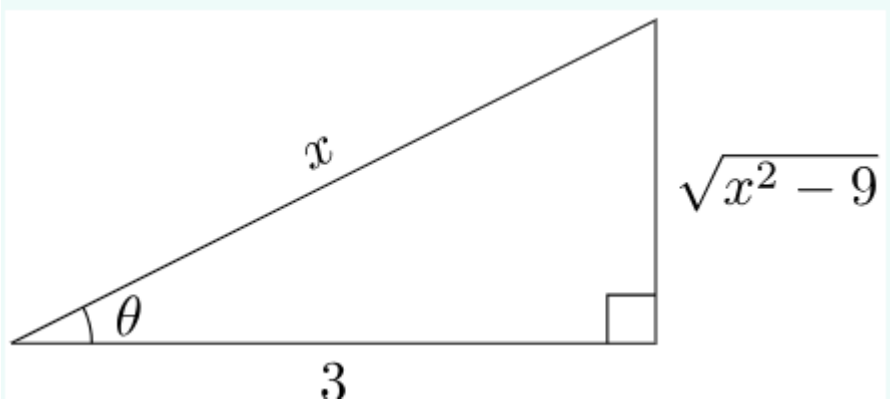


Figure 7.6.7.

This gives us $\cos(\theta) = \frac{3}{x}$ or, equivalently, $\sec(\theta) = \frac{x}{3}$

2. For the integral $\int \frac{x^3}{(4-x^2)^{3/2}} dx$, we have the following triangle.

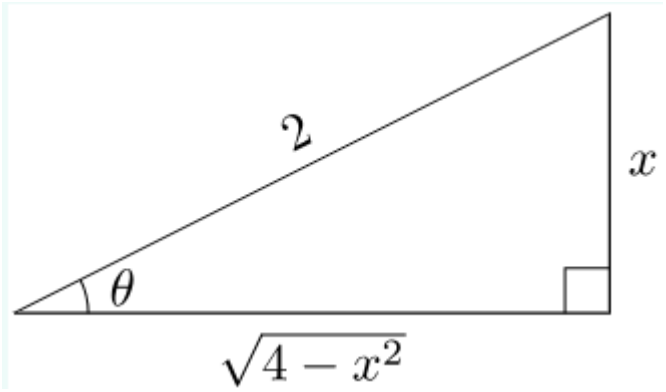


Figure 7.6.8.

This gives us $\sin(\theta) = \frac{x}{2}$.

3. For the integral $\int \frac{1}{x^2\sqrt{x^2+1}} dx$, we have the following triangle.

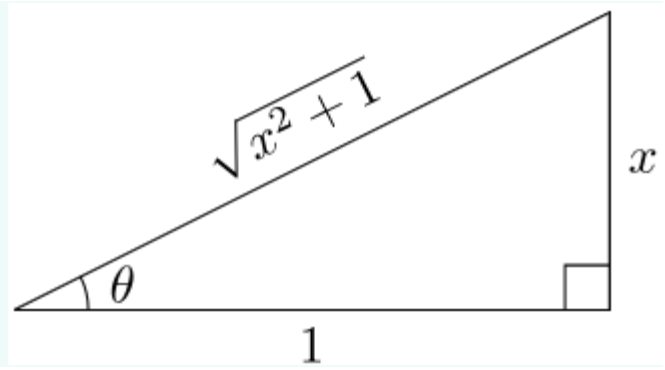


Figure 7.6.9

This gives us $\tan(\theta) = \frac{x}{1}$ or, equivalently, $\tan(\theta) = x$.

d. 1. We will use $x = 3\sec(\theta)$ and $dx = 3\sec(\theta)\tan(\theta)d\theta$. Then, we can expect to find the Pythagorean Identity $9\sec^2(\theta) - 9 = 9\tan^2(\theta)$.
2. We will use $x = 2\sin(\theta)$ and $dx = 2\cos(\theta)d\theta$. Then, we can expect to find the Pythagorean Identity $4 - 4\sin^2(\theta) = 4\cos^2(\theta)$.
3. We will use $x = \tan(\theta)$ and $dx = \sec^2(\theta)d\theta$. Then, we can expect to find the Pythagorean Identity $\tan^2(\theta) + 1 = \sec^2(\theta)$.

e. 1. Using $x = 3\sec(\theta)$, we get:

$$\begin{aligned}\int \frac{\sqrt{x^2-9}}{x}dx &= \int \frac{\sqrt{9\sec^2(\theta)-9}}{3\sec(\theta)}(3\sec(\theta)\tan(\theta))d\theta \\ &= \int \frac{3\sqrt{9\tan^2(\theta)}\sec(\theta)\tan(\theta)}{3\sec(\theta)}d\theta \\ &= 3\int \tan^2(\theta)d\theta \\ &3\int \sec^2(\theta) - 1d\theta \\ &3(\tan(\theta) - \theta) + C \\ &3\tan(\theta) - 3\theta + C\end{aligned}$$

2. Using $x = 2\sin(\theta)$, we get:

$$\begin{aligned}\int \frac{2}{(4-x^2)^{3/2}}dx &= \int \frac{2}{\left(\sqrt{4-4\sin^2(\theta)}\right)^3}(2\cos(\theta))d\theta \\ &= \int \frac{4\cos(\theta)}{\left(\sqrt{4\cos^2(\theta)}\right)^3}d\theta \\ &= \frac{1}{2}\int \frac{1}{\cos^2(\theta)}d\theta \\ &= \frac{1}{2}\int \sec^2(\theta)d\theta \\ &= \frac{1}{2}\tan(\theta) + C\end{aligned}$$

3. Using $x = \tan(\theta)$, we get:

$$\begin{aligned}\int \frac{1}{x^2\sqrt{x^2+1}}dx &= \int \frac{1}{\tan^2(\theta)\sqrt{\tan^2(\theta)+1}}\left(\sec^2(\theta)\right)d\theta \\ &= \int \frac{\sec^2(\theta)}{\tan^2(\theta)\sqrt{\sec^2(\theta)}}d\theta \\ &= \int \frac{\sec(\theta)}{\tan^2(\theta)} \\ &= \int \frac{\cos(\theta)}{\sin^2(\theta)}d\theta \\ &= \int \csc(\theta)\cot(\theta)d\theta \\ &= -\csc(\theta) + C\end{aligned}$$

f. 1. From the triangle, we get $\tan(\theta) = \frac{\sqrt{x^2-9}}{3}$. Then, since $\sec(\theta = \frac{x}{3}$, we can use $\theta = \sec^{-1}\left(\frac{x}{3}\right)$. This gives us:

$$\begin{aligned}\int \frac{\sqrt{x^2-9}}{x}dx &= 3\int \sec^2(\theta) - 1d\theta \\ &= 3\tan(\theta) - 3\theta + C \\ &= \sqrt{x^2-9} - 3\sec^{-1}\left(\frac{x}{3}\right) + C\end{aligned}$$

2. From the triangle, we get $\tan(\theta) = \frac{x}{\sqrt{4-x^2}}$. This gives us:

$$\begin{aligned}\int \frac{2}{(4-x^2)^{3/2}}dx &= \frac{1}{2}\int \sec^2(\theta)d\theta \\ &= \frac{1}{2}\tan(\theta)+C \\ &= \frac{x}{2\sqrt{4-x^2}}+C\end{aligned}$$

3. From the triangle, we get $\csc(\theta)=\frac{\sqrt{x^2+1}}{x}$. This gives us:

$$\begin{aligned}\int \frac{1}{x^2\sqrt{x^2+1}}dx &= \int \csc(\theta)\cot(\theta)d\theta \\ &= -\csc(\theta)+C \\ &= -\frac{\sqrt{x^2+1}}{x}+C\end{aligned}$$

This page titled 7.6: Trigonometric Substitution is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Peter Keep.

7.7: Partial Fractions

In this last integration technique, we'll consider more rational functions. We've already thought about rational functions a bit (Integrating Rational Functions), but here we'll add some more detail to a special type of rational function. Let's not spoil anything. Instead, we'll just jump into a quick comparison.

Activity 7.7.1: Comparing Rational Integrands.

We're going to compare three integrals:

$$\int \frac{2}{x^2+4x+5}dx$$
$$\int \frac{2}{x^2+4x+3}dx$$
$$\int \left(\frac{1}{x+1}-\frac{1}{x+3}\right)dx$$

a. Start with the first integral:

$$\int \frac{2}{x^2+4x+5}dx$$

How would you approach integrating this? **Hint**. This is a constant over a quadratic: can you complete the square, and then connect the result to an inverse tangent function?

b. Try the same tactic on the second integral:

$$\int \frac{2}{x^2+4x+3}dx$$

You don't need to complete this integral, but think about how you might proceed. **Hint**. Check the structure of the denominator when you complete the square: since you don't end up with a sum of squares, you can't use the inverse tangent setup. You'll need to do a trig substitution instead...

c. Think about the third integral:

$$\int \left(\frac{1}{x+1}-\frac{1}{x+3}\right)dx$$

How would you integrate this?

d. The third integral is unique from the other two in that it is has two terms. Let's combine them together to see how we could write this integral to compare it more closely to the other two. Subtract $\frac{1}{x+1}-\frac{1}{x+3}$ using common denominators and compare your rewritten integral to the other two.

e. Which of these integrals and/or representations of an integral is easiest to work with? Which one is most annoying to work with? Why?

We're going to try to take advantage of the rewritten structure in Activity 7.7.1: when we can decompose a "large" rational function into a sum of "smaller" rational functions, we'll be more likely to be able to anti-differentiate the "smaller" pieces.

The real trick, here, is going to be recognizing when we can do this and building a process for how we do this.

When?

In order to recognize when we will employ this strategy, we should think about what we're doing: we are attempting to write one fraction as a sum or difference of others. We're "undo-ing" fraction addition, here. We can remember that when we add (or subtract) fractions, we need to find a common denominator and rewrite our fraction terms as equivalent versions with this same denominator. This typically is done by just "scaling" the numerator and denominator of a fraction by a factor (often the other denominator). For instance:

$$\begin{aligned}\frac{3}{5}-\frac{1}{3} &= \frac{3}{5}\left(\frac{3}{3}\right)-\frac{1}{3}\left(\frac{5}{5}\right)\\ &= \frac{9}{15}-\frac{5}{15}\\ &= \frac{4}{15}\end{aligned}$$

This same thing happens when we think about rational functions:

$$\begin{aligned}\frac{1}{x}+\frac{2}{x-1} &= \frac{1}{x}\left(\frac{x-1}{x-1}\right)+\frac{2}{x-1}\left(\frac{x}{x}\right)\\ &= \frac{x-1}{x(x-1)}+\frac{2x}{x(x-1)}\\ &= \frac{3x-1}{x(x-1)} \quad \text{or} \quad \frac{3x-1}{x^2-x}\end{aligned}$$

Now, we can notice something about these (common) denominators: since we scaled each fraction before adding/subtracting them, the denominator is a product of factors.

So right away we know that this method only will work well when we can factor the denominator of a rational integrand.

We also know (from Integrating Rational Functions) that once the degree of the numerator is at least the same (or larger) than the degree of the denominator, we can rewrite things using division.

So, we will use this strategy of rewriting rational function integrands as sums or differences of "smaller" rational functions when:

1. The denominator can be factored.
2. The degree of the denominator is larger than the degree of the numerator.

How?

The actual process for finding the smaller "partial" fractions is not complicated. Once we have the denominator of our rational function factored, we can see what the possible denominators that had to "combine" to make a "common" denominator were.

Our general process will be simple: set up these possible rational functions and put unknown placeholder numerators in place, making sure that the degree is smaller than the degree of the denominator. Then, we add these possible rational functions up and compare the numerator (with the unknown placeholders) to the numerator that we want (from our actual function that we're integrating).

There are more tricks along the way, but this process is simple to think about. What we'll find, though, is that the process is full of algebra, which can be tedious.

Let's look at two small examples to see how this could work.

Activity 7.7.2: First Examples of Partial Fractions.

a. Consider the integral:

$$\int \frac{6x-16}{x^2-4x+3}\,dx$$

First, confirm that this would be very annoying to try to use u-substitution on, even though we have a linear numerator and quadratic denominator.

b. Notice that the denominator can be factored:

$$\int \frac{6x-16}{(x-3)(x-1)}\,dx$$

If this integrand function were a sum of two "smaller" rational functions, what would their denominators be? What do we know about their numerators? **Hint**. The denominators would need to multiply to get $(x-3)(x-1)$, and each numerator would have to have a degree smaller than its respective denominator.

c. Use some variables (it's typical to use capital letters like A, B, C, etc.) to represent the numerators, and then add the partial fractions together. What do you get? How does this rational function compare to $\frac{6x-16}{(x-3)(x-1)}$? **Hint**. We want to try to add up $\frac{A}{x-3}+\frac{B}{x-1}$ and compare the result to $\frac{6x-16}{(x-3)(x-1)}$.

d. Set up an equation connecting the numerators, and solve for your unknown variables. What are the two fractions that added together to get $\frac{6x-16}{(x-3)(x-1)}$?

e. Anti-differentiate to solve the integral $\int \frac{6x-16}{(x-3)(x-1)}dx$?

f. Let's do the same thing with a new integral:

$$\int \frac{3x^2-2x+3}{x\left(x^2+1\right)}dx$$

What are the partial fraction forms? What kinds of numerators should we expect to see? Use variables to represent these.

g. Add the partial fractions together and set up an equation for the numerators to solve. What are the two partial fractions after you solve for the unknown coefficients? **Hint**. We want to find values for A, B,and C where

$$\frac{A}{x}+\frac{Bx+C}{x^2+1}=\frac{3x^2-2x+3}{x\left(x^2+1\right)}$$

Use common denominators to add

$$\frac{A}{x}+\frac{Bx+C}{x^2+1}$$

and then set the resulting numerator equal to $3x^2-2x+3$ and try to solve.

h. Anti-differentiate and solve the integral $\int \frac{3x^2-2x+3}{x\left(x^2+1\right)}dx$.

Answer available at the end of this section.

So we can see the basics of how this will work:

1. Figure out the denominators of the fractions we could add to get the function we're integrating.
2. Construct the partial fractions using placeholders for the numerators (making sure to keep the degree of the numerators smaller than the degree of the denominators).
3. Add these placeholder fractions up, and see what the coefficients would have to be in order to make them add to the function we're integrating.
4. Anti-differentiate the smaller rational functions.

More Specific Strategies

We're going to investigate these partial fractions a bit more, and focus on two cases: linear denominators, and quadratic denominators. This will limit the scope of our work enough that this doesn't get too wild, and also works well with the kinds of things we know how to antidifferentiate: we can antidifferentiate any rational function with a linear denominator and we know how to anti-differentiate many rational functions with quadratic denominators.

Partial Fraction Type: Simple Linear Factors

For rational functions where the denominator has some linear factor like $(x-k)$, we can set up a partial fraction with just a constant coefficient in the numerator:

$$\frac{A}{x-k}$$

We have outlined a pretty reasonable approach for these in Activity 7.7.2 when we worked on the integral

$$\int \frac{6x-16}{(x-3)(x-1)}dx$$

We can make note of something useful, though. At one point during the process, we had set up the two partial fractions, added them together, and then we said that we wanted to find the values of A and B that made the general partial fraction forms we had set up match with the actual function we were hoping to integrate.

$$\begin{aligned}\frac{A}{x-3}+\frac{B}{x-1} &= \frac{A(x-1)+B(x-3)}{(x-3)(x-1)}\\ &= \frac{6x-16}{(x-3)(x-1)}\end{aligned}$$

What this meant for us was that we can set the numerators equal to each other and solve for A and B:

$$A(x-1)+B(x-3)=6x-16$$

Let's pause here.

We are hoping to find the values of A and B that will make this equation work out for all values of x. But we also can make use of the fact that this equation will be true for some specific values of x. For instance, we can evaluate this equation at $x=2$:

$$\begin{aligned}A(2-1)+B(2-3)&=6(2)-16\\ A-B&=-4\end{aligned}$$

This just generates another equation connecting A and B that we could use like we did in in Activity 7.7.2.

But is there a nicer, more convenient x-value to use? Can we force this to generate a nicer equation involving our unknown coefficients? Can we find some x-values that will make the factors $(x-1)$ and $(x-3)$ go to 0, for instance?

$$A\underbrace{(x-1)}_{x=1}+B\underbrace{(x-3)}_{x=3}=6x-16$$

Let's try it!

$$\begin{aligned}x=1: \quad A(1-1)+B(1-3) &= 6(1)-16\\ -2B &= -10\\ B &= 5\\ x=3: \quad A(3-1)+B(3-3) &= 6(3)-16\\ 2A &= 2\\ A &= 1\end{aligned}$$

These are the same values we had in Activity 7.7.2! This strategy works well for any simple linear factors, and definitely helps to reduce the amount of algebra required.

Simple Linear Factors.

If a rational function has a simple linear factor in the denominator in the form $(x-k)$, then the corresponding partial fraction is

$$\frac{A}{x-k}$$

where A is a real number constant with $A\neq 0$.

When we anti-differentiate these, we will end up with a logarithmic function.

Partial Fraction Type: Irreducible Quadratic Factors

So what about our other example in Activity 7.7.2? We had a quadratic factor (that couldn't be factored nicely). We'll call these irreducible quadratic factors. The real difference just was the structure of the numerator, where we accounted for the option of the numerator being some linear function, $Ax + B$.

There isn't much more to say here, since the algebra can be frustrating to deal with. It can be helpful to save some of these coefficients for last: that way, we can find some of the "easier" ones (from the Simple Linear Factors cases) first, and hopefully the remaining coefficients won't be too difficult to find.

Irreducible Quadratic Factors.

If a rational function has an irreducible quadratic factor in the denominator in the form $(ax^2 + bx + c)$, then the corresponding partial fraction is

$$\frac{Ax + B}{ax^2 + bx + c}$$

where A and B are real number constants, and A and B cannot both be 0. That is, if one of A or B is 0, then the other cannot be. When we anti-differentiate these, we can expect a logarithmic function, an inverse tangent function, or some combination of the two. We can get other antiderivatives, but those will be beyond the scope of this introductory text.

Partial Fraction Type: Repeated Linear Factors

Let's look at one last case. It's a bit of a weird one, so we'll explore it in its own example.

Activity 7.7.1: Fiddling with Repeated Factors.

Let's sit with the following integral:

$$\int \frac{3x + 5}{x^2 + 2x + 1} dx$$

Before we start, we can think about how annoying it would be to try to start with a u-substitution where $u = x^2 + 2x + 1$.

a. Factor the denominator! What's different about the factors in this denominator compared to the ones in Activity 7.7.2?
b. Why can't we use these two factors to create two partial fractions with Simple Linear Factors?
c. Ok, instead, we're going to do some algebra that is reminiscent of what we have done before in Section 7.3. Can you write the numerator, $3x + 5$, as some coefficient on the factor $(x + 1)$ with some constant "remainder?"

$$3x + 5 = \square(x + 1) + \square$$

d. Why is this reforming of the numerator useful? What does that do, when we write it over the factored denominator? Why did we choose $(x + 1)$ as the factor to use for our rewriting? Feel free to show why this is helpful! **Hint**. Split the fraction up across the sum in the numerator!
e. Integrate $\int \frac{3x + 5}{x^2 + 2x + 1} dx$ using your clever rewriting. Explain why this is a friendlier form.

This is something we can do, algebraically, for every fraction with a "repeated" factor like this. But, more importantly, we can incorporate this idea into how we think about partial fractions.

Answer available at the end of this section.

Repeated Linear Factors.

If a rational function has a repeated linear factor in the denominator in the form $(x-k)^n$, where mn is some integer greater than 1, then the corresponding partial fractions are

$$\frac{A_1}{x-k}+\frac{A_2}{(x-k)^2}+\ldots+\frac{A_n}{(x-k)^n}$$

where $A_1, A_2, \ldots, A_n$ are real numbers and $A_n \neq 0$.

When we antidifferentiate these, we can expect to use the rule and maybe find a logarithm function.

At this point, we can spend our time going through one or two examples where we put this all together.

Example 7.7.4

a. $\int \frac{5x^2+27x+51}{(x-2)(x+3)^3}dx$ **Hint 1**. Your partial fraction forms will look like this:

$$\frac{A}{x-2}+\frac{B}{x+3}+\frac{C}{(x+3)^2}+\frac{D}{(x+3)^3}$$

Hint 2. You can find the values for A and D pretty easily by thinking about convenient x-values, like we talked about in Partial Fraction Type: Simple Linear Factors.

b. $\int \frac{6x^2+25x+27}{(x+1)\left(x^2+4x+5\right)}dx$ **Hint 1**. Your partial fraction forms will look like this:

$$\frac{A}{x+1}+\frac{Bx+C}{x^2+4x+5}$$

Hint 2. You'll be able to easily find A by thinking about convenient x-values, but not B or C.

Solution

a. We'll rewrite our integral using the partial fraction forms:

$$\int \frac{5x^2+27x+51}{(x-2)(x+3)^3}dx=\int \frac{A}{x-2}+\frac{B}{x+3}+\frac{C}{(x+3)^2}+\frac{D}{(x+3)^3}dx$$

When we combine these fractions to compare the numerators, we end up with the following equation:

$$A(x+3)^3+B(x-2)(x+3)^2+C(x-2)(x+3)+D(x-2)=5x^2+27x+51$$

We can evaluate this at $x=2$ and $x=-3$ to reveal the values of A and D:

$$\begin{aligned} x=2: \quad A(5)^3 &= 5(4)+27(2)+51 \\ A &= \frac{125}{125}=1 \\ x=-3: \quad D(-5) &= 5(9)+27(-3)+51 \\ D &= \frac{15}{-5}=-3 \end{aligned}$$

Now that we know $A=1$ and $D=-3$, we can put those into our equation connecting the numerators, and solve for $BandC$.

$$(x+3)^3+B(x-2)(x+3)^2+C(x-2)(x+3)-3(x-2)=5x^2+27x+51$$

If we just consider the cubic terms, then on the left side of the equation we have x^3+Bx^3, and there are no cubic terms on the right side. This means that $1+B=0$ and so $B=-1$. Similarly, we can consider just the constant terms of the (updated) equation:

$$(x+3)^3-(x-2)(x+3)^2+C(x-2)(x+3)-3(x-2)=5x^2+27x+51$$

We can see that on the left side, we'll have $(3)^3-(-2)(3)^2+C(-2)(3)-3(-2)$ and on the right side, the constant term is 51.

$$\begin{aligned}27+18-6C+6&=51\\51-6C&=51\\C&=0\end{aligned}$$

Finally, we have our new, rewritten, integral. We can antidifferentiate.

$$\begin{aligned}\int\frac{5x^2+27x+51}{(x-2)(x+3)^3}dx&=\int\frac{1}{x-2}-\frac{1}{x+3}-\frac{3}{(x+3)^3}dx\\&=\ln|x-2|-\ln|x+3|+\frac{3}{2(x+3)^2}+C\end{aligned}$$

b. Let's, again, rewrite our integral using the partial fraction forms we set up:

$$\int\frac{6x^2+25x+27}{(x+1)\left(x^2+4x+5\right)}dx=\int\text{ fracA }x+1+\frac{Bx+C}{x^2+4x+5}dx$$

Our equation for the combined numerator is:

$$A\left(x^2+4x+5\right)+(Bx+C)(x+1)=6x^2+25x+27$$

We can find A be evaluating at $x=-1$.

$$\begin{aligned}x=-1:\quad A(1-4+5)&=6(1)-25+27\\A&=\frac{8}{2}=4\end{aligned}$$

Now, knowing that $A=2$, we can rewrite our equation to solve for B and C.

$$4\left(x^2+4x+5\right)+(Bx+C)(x+1)=6x^2+25x+27$$

We can collect the quadratic terms, and see the following equation:

$$4x^2+Bx^2=6x^2$$

So $B=2$. Similarly, we can collect the constant terms:

$$4(5)+C=27$$

It is easy to see that $C=7$. So we have our newly rewritten integral:

$$\int\frac{6x^2+25x+27}{(x+1)\left(x^2+4x+5\right)}dx=\int\frac{4}{x+1}+\frac{2x+7}{x^2+4x+5}dx$$

The first term is pretty straightforward to integrate: we'll get a log. The second one, though, will take some work. Let's consider it by itself:

$$\int\frac{2x+7}{x^2+4x+5}dx$$

We can start with a u-substitution of $u=x^2+4x+5$, giving us $dx=2x+4dx$. Let's rewrite the numerator as $2x+4+3$ in order to make this work:

$$\begin{aligned}\int\frac{2x+7}{x^2+4x+5}dx&=\int\frac{(2x+4)+3}{x^2+4x+5}dx\\&=\int\frac{2x+4}{x^2+4x+5}dx+\int\frac{3}{x^2+4x+5}dx\end{aligned}$$

Now, the first of these will work with our stated substitution. The second one, though, will require a different strategy. Let's complete the square to get the inverse tangent form (Theorem 7.3.6). For all three integrals, then, we get:

$$\begin{aligned}\int\frac{6x^2+25x+27}{(x+1)\left(x^2+4x+5\right)}dx&=\int\frac{4}{x+1}+\frac{2x+7}{x^2+4x+5}dx\\&=\int\frac{4}{x+1}+\frac{2x+4}{x^2+4x+5}+\frac{3}{x^2+4x+5}dx\\&=\int\frac{4}{x+1}+\frac{2x+4}{x^2+4x+5}+\frac{3}{(x+2)^2+1}dx\\&=4\ln|x+1|+\ln\left|x^2+4x+5\right|+3\tan^{-1}(x+2)+C\end{aligned}$$

There are more things that we can think about, but it really ends up being just extensions of what we've done. For instance, we could think about repeated quadratic factors or irreducible polynomials that have larger degrees, but the general principles are the same: we set up a placeholder numerator that has a degree less than the denominator and try to solve for the unknown coefficients.

There really are only two limitations for us:

1. As we increase the number of coefficients, it becomes very tedious to solve for them. It isn't difficult, really: just a lot of algebra.
2. As we increase the degree of the kinds of denominators we see, we run out of approaches for anti-differentiation. We could spend much more time talking about integrating more rational functions or dive into the world of irrational coefficients (or even non-real ones), but this serves as a good stopping point for our purposes.

Practice Problems

Exercise 7.7.1

Why do we use partial fraction decomposition on some integrals of rational functions? Give an example and explain why it is helpful in your example.

Exercise 7.7.2

For each rational function described, write out the corresponding partial fraction forms.

a. $\dfrac{p(x)}{(x-4)(x+2)(x-1)}$ where $p(x)$ is some polynomial with degree less than 3.
b. $\dfrac{p(x)}{(x+1)^2(3x-5)^3}$ where $p(x)$ is some polynomial with degree less than 5.
c. $\dfrac{p(x)}{(x^2+1)\,(x^2+2x+5)}$ where $p(x)$ is some polynomial with degree less than 4.
d. $\dfrac{p(x)}{x^4-1}$ where $p(x)$ is some polynomial with degree less than 4.

Exercise 7.7.3

Consider the following integral, with the partial fraction forms written out:

$$\int \frac{x^3+6x^2-x}{(x-2)(x+1)\,(x^2+1)}dx = \int \left(\frac{A}{x-2}+\frac{B}{x+1}+\frac{Cx+D}{x^2+1}\right)dx$$

a. Write an equation connecting the numerators.
b. Find (and use) a specific x-value to input into the equation to solve for A.
c. Find (and use) a specific x-value to input into the equation to solve for B.
d. Why can you not use this strategy to solve for coefficients C or D?
e. Find the cubic terms (you will need to do some multiplication) on both sides of your equation. Use these to solve for C.
f. Find the constant terms (you will need to do some multiplication) on both sides of your equation. Use these to solve for D.
g. Integrate!

Exercise 7.7.4

Explain why partial fractions is not an appropriate technique for the following integral:

$$\int \frac{x^2+x}{x^2-x+1}dx$$

How should we approach this integral, instead?

Exercise 7.7.5

Integrate the following.

a. $\int \frac{2}{(x-1)(x+3)}dx$
b. $\int \frac{4x+1}{(x-4)(x+5)}dx$
c. $\int \frac{2x^2-15x+32}{x\left(x^2-8x^2+64\right)}$
d. $\int \frac{1}{(x+2)(x-2)}dx$
e. $\int \frac{20x}{(x-1)\left(x^2+4x+5\right)}dx$
f. $\int \frac{x^2}{(x-2)^3}dx$

Exercise 7.7.6

In the problems we are looking at in this section, we're limiting ourselves to, at most, irreducible quadratic factors in the denominator. In problems with simple linear factors, repeated linear factors, or irreducible quadratic factors, what types of antiderivative functions do you expect to see? Explain.

Exercise 7.7.7

For each of the following integrals, we will do some preliminary work before using partial fractions to integrate. Really, we'll perform a specific u-substitution that will give us some resulting integral to use partial fractions on.

a. $\int \frac{4e^{2x}}{\left(e^{2x}+3\right)\left(e^{2x}-5\right)}dx$ where we use $u=e^{2x}$.
b. $\int \sqrt{e^x+1}dx$ where we use $u=\sqrt{e^x+1}$.
c. $\int \frac{\sqrt{x}+3}{\sqrt{x}(x-1)}dx$ where we use $u=\sqrt{x}$.

Answers

Activity 7.7.2: First Examples of Partial Fractions.

b. We might have something like $\frac{\square}{x-3}+\frac{\square}{x-1}=\frac{6x-16}{(x-3)(x-1)}$ where the numerators are constant terms (since their degree must be smaller than 1).

c.
$$\begin{aligned}\frac{A}{x-3}+\frac{B}{x-1} &= \frac{A}{x-3}\left(\frac{x-1}{x-1}\right)+\frac{B}{x-1}\left(\frac{x-3}{x-3}\right)\\ &= \frac{A(x-1)}{(x-3)(x-1)}+\frac{B(x-3)}{(x-3)(x-1)}\\ &= \frac{A(x-1)+B(x-3)}{(x-3)(x-1)}\end{aligned}$$

When we compare this to $\frac{6x-16}{(x-3)(x-1)}$, we can see that the denominators are the same, and so, if these are equal to each other, the numerators must be as well.

d.
$$\begin{aligned}A(x-1)+B(x-3) &= 6x-16\\ (A+B)x-(A+3B) &= 6x-16\end{aligned}$$

This means that $A+B=6$ and $A+3B=6$. There are a variety of ways to solve this, but we can say that $A=6-B$, and so then

$$\begin{aligned} A+3B &= 16 \\ (6-B)+3B &= 16 \\ 2B &= 10 \\ B &= 5 \end{aligned}$$

Since $A = 6 - B$, we have $A = 6 - 5 = 1$.

e.
$$\frac{6x-16}{(x-3)(x-1)} = \frac{1}{x-3} + \frac{5}{x-1}$$
$$\begin{aligned} \int \frac{6x-16}{(x-3)(x-1)}dx &= \int \frac{1}{x-3} + \frac{5}{x-1}dx \\ &= \ln|x-3| + 5\ln|x-1| + C \end{aligned}$$

f. We expect to see partial fractions $\frac{\square}{x}$ and $\frac{\square}{x^2+1}$. We know the numerator over x needs to be constant, but we could have a linear numerator over x^2+1. So our partial fractions are:

g.
$$\frac{A}{x} + \frac{Bx+C}{x^2+1}$$
$$\begin{aligned} \frac{A}{x} + \frac{Bx+C}{x^2+1} &= \frac{A}{x}\left(\frac{x^2+1}{x^2+1}\right) + \frac{Bx+C}{x^2+1}\left(\frac{x}{x}\right) \\ &= \frac{A\left(x^2+1\right) + (Bx+C)(x)}{x\left(x^2+1\right)} \end{aligned}$$

Then, we can set the numerators equal to get:

$$\begin{aligned} A\left(x^2+1\right) + (Bx+C)(x) &= 3x^2 - 2x + 3 \\ (A+B)x^2 + Cx + A &= 3x^2 - 2x + 3 \end{aligned}$$

So, we now can see that since $Cx = -2x$, we have $C = -2$. We also can look at the constant terms and see that $A = 3$. Now we know that $(A+B)x^2 = 3x^2$, but since $A = 3$ we have $B = 0$. So we have

h.
$$\int \frac{3x^2-2x+3}{x\left(x^2+1\right)}dx = \int \frac{3}{x} - \frac{2}{x^2+1}dx$$
$$\begin{aligned} \int \frac{3x^2-2x+3}{x\left(x^2+1\right)}dx &= \int \frac{3}{x} - \frac{2}{x^2+1}dx \\ &= 3\ln|x| - 2\tan^{-1}(x) + C \end{aligned}$$

Activity 7.7.1**: Fiddling with Repeated Factors.**

a.
$$x^2 + 2x + 1 = (x+1)(x+1)$$

b.

c.
$$3x + 5 = 3(x+1) + 2$$

d.
$$\begin{aligned} \frac{3x+5}{(x+1)^2} &= \frac{3(x+1)+2}{(x+1)^2} \\ &= \frac{3(x+1)}{(x+1)^2} + \frac{2}{(x+1)^2} \\ &= \frac{3}{x+1} + \frac{2}{(x+1)^2} \end{aligned}$$

e. We'll use our rewritten integral:

$$\int \frac{3x+5}{x^2+2x+1}dx = \int \frac{3}{x+1} + \frac{2}{(x+1)^2}dx$$

Now, we can split this into two integrals to deal with separately. For both, we can use the u-substitution $u = x + 1$.

$$\begin{aligned} \int \frac{3}{x+1} + \frac{2}{(x+1)^2}dx &= \int \frac{3}{u}du + \int \frac{2}{u^2}du \\ &= 3\ln|u| - \frac{2}{u} + C \\ &= 3\ln|x+1| - \frac{2}{x+1} + C \end{aligned}$$

This page titled 7.7: Partial Fractions is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Peter Keep.

8: INFINITE SERIES

8.1: Introduction to Infinite Sequences
8.2: Introduction to Infinite Series
8.3: The Divergence Test and the Harmonic Series
8.4: The Integral Test
8.5: Alternating Series and Conditional Convergence
8.6: Common Series Types
8.7: Comparison Tests
8.8: The Ratio and Root Tests
8.9: Exploration- Infinite Series

This page titled 8: Infinite Series is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Peter Keep.

8.1: Introduction to Infinite Sequences

Sequences as Functions

Before we move on to our actual goal of analyzing infinite series, we will construct infinite sequences. The big thing to remember here is that, when we build and analyze these sequences, we are are really building and analyzing functions. We want to keep this idea of sequences as functions in the forefront, since it will help us as we think about accumulating these function values into infinite series.

Activity 8.1.1: Building our First Sequences.

We might already have some familiarity with sequences. Here, we'll focus less on some of the detailed mechanics and just think about these sequences as functions.

a. Describe a sequence of numbers where you use a consistent rule/ function to build each term (each number) based only on the previous term in the sequence. You will need to decide on some first term to start your sequence.
b. Describe a different sequence of numbers using the same rule to generate new terms/numbers from the previous one. What do you need to do to make these two sequences different from each other?
c. Describe a new sequence of numbers where you use a consistent rule/function to build each term based on its position in the sequence (i.e. the first term will be some rule/function based on the input 1, the second will be based on 2, you'll use 3 to get the third term, etc.). We will call the position of each term in the sequence the index.
d. Describe another, new, sequence of numbers where you use a consistent rule/function to build each term based on its index. This time, make the terms get smaller in size as the index increases.

Definition 8.1.2: Explicit Formula.

An infinite sequence defined using an **explicit formula** is one where the kth term of the sequence is defined as a function output of k, the term's index. Using notation, we might say that $a_k = f(k)$ where:

- a is the "name" of the sequence (similar to how f and g are common names of functions).
- k is the index of the term, typically a non-negative integer.
- $f(k)$ is the function that we use to generate the terms.

Definition 8.1.3: Recursion Relation.

A sequence is defined using a **recursion relation** is one where the kth term of the sequence is defined as a function output of the previous term, the $(k-1)$st term. The sequence also needs some initial term to base the subsequent terms from. Using notation, we might say that $a_k = f(a_{k-1})$.

These definitions are relatively limited. You might, for instance, know of a very famous sequence that is defined recursively by having each term being the sum of the two previous terms. Our study of sequences will be brief and will be leading us towards infinite series, so there are a lot of nuances about sequences that we will skip.

Exercise 8.1.4: Returning to our First Sequences.

Let's return back to the four sequences we created in Activity 8.1.1.

a. For each of the sequences, how are we going to define them? Explicit formulas? Recursion relations? How do you know?

b. Now, for each sequence, define the sequence formally using either an explicit formula or recursion relation, whichever matches with how you described the sequence in Activity 8.1.1.

Example 8.1.5: Practice Writing some Terms.

For each of the following sequences, write out the first handful of terms. There isn't a set amount, but you should write out enough to get a feel for the sequence structure and how the different ways of defining the sequences work. In each, you can start the index k at 1 and count upwards $(k = 1, 2, 3, ...)$.

a. $a_1 = \frac{1}{3}$ and $a_k = 2\left(a_{k-1}\right)^2$
b. $b_k = \frac{\sin(k)}{k^2}$
c. $c_k = \sqrt[k]{k+1}$
d. $d_k = \frac{k + e^k}{e^k}$

Activity 8.1.6: Describing These Sequences.

Let's look at the sequences from Example 8.1.5. Go through the following tasks for each sequence.

a. What do you think each sequence is "counting towards" (if anything)?
b. Can you show that the sequence is counting towards what you think it is with a limit (or show that it's not counting towards anything)?

Activity 8.1.7: Write the Sequence Rules.

We'll look at some sequences by writing out the first handful of terms. From there, our goal is to write out more terms and eventually define each sequence fully. For each sequence, write an explicit formula and a recursion relation to define the sequence. You can choose whether to start your index at $k = 0$ or $k = 1$.

a. $\{a_k\} = \left\{4, \frac{2}{3}, \frac{1}{9}, \frac{1}{54}, ...\right\}$
b. $\{b_k\} = \left\{\frac{3}{5}, \frac{2}{5}, \frac{5}{17}, \frac{3}{13}, \frac{7}{37}, \cdots,\right\}$
c. $\{c_k\} = \left\{\frac{1}{5}, \frac{3}{5}, 1, \frac{7}{5}, ...\right\}$
d. What kinds of connections do you notice between the explicit formulas and the recursion relations for these sequences?

Before moving on, we should make a couple of notes:

- When we add something recursively (where we add the same thing repetitively to get from the kth term to the $(k+1)$st term), this is the same thing as multiplication in an explicit formula!

$$\underbrace{3 + 3 + ... + 3}_{k \text{ times}} = 3k$$

- Similarly, when we multiply something recursively, we can think about this as an exponential in the explicit formula!

$$\underbrace{3 \cdot 3 \cdot 3 \cdots 3}_{k \text{ times}} = 3^k$$

- In general, it can be pretty difficult to find either of these formulas for a given sequence of numbers. In fact, in any sequence where only the first few terms are given, we can find an infinite number of formulas that provide those first few terms and then deviate to any other numbers. We cannot easily extrapolate onto only one "pattern" or formula. Because of this, we'll try to limit our work as much as we can to situations where we have the formula defining the sequences.

Graphing Sequences

We have tried introducing and talking about sequences as special types of functions, mapping natural number inputs to real number outputs. If we are committed to thinking about sequences as functions, with maybe some special context, then we should really investigate how one of our primary representations of functions (graphs) manifests itself in this new context.

There really is not too much to think about here! We can focus on the domain of these functions. If we define a sequence $\{a_k\}$ explicitly, then we have some function $a_k = f(k)$, and we can plot this sequence function in the same way that we normally would any other function $g(x)$. We will use the horizontal axis for the inputs and the vertical axis to represent the outputs, and try to visualize the graph as the set of all of the pairs of inputs with their (single) corresponding output.

The only new feature, then, is that these functions have only non-negative integer inputs. So when we plot the points, we do not get some nice curve acting as a visual representation of the function: we get discrete points floating on the 2-dimensional plane, each with some consistent horizontal spacing between them.

Consider, for example the following function and sequence:

$$f(x) = \frac{x}{x^2+1} \text{ for } x \geq 0$$

and

$$a_k = \frac{k}{k^2+1} \text{ for } k = 0, 1, 2, ...$$

We can graph the curve $y = f(x)$ in the normal way, as a smooth curve starting at the point at $x = 0$.

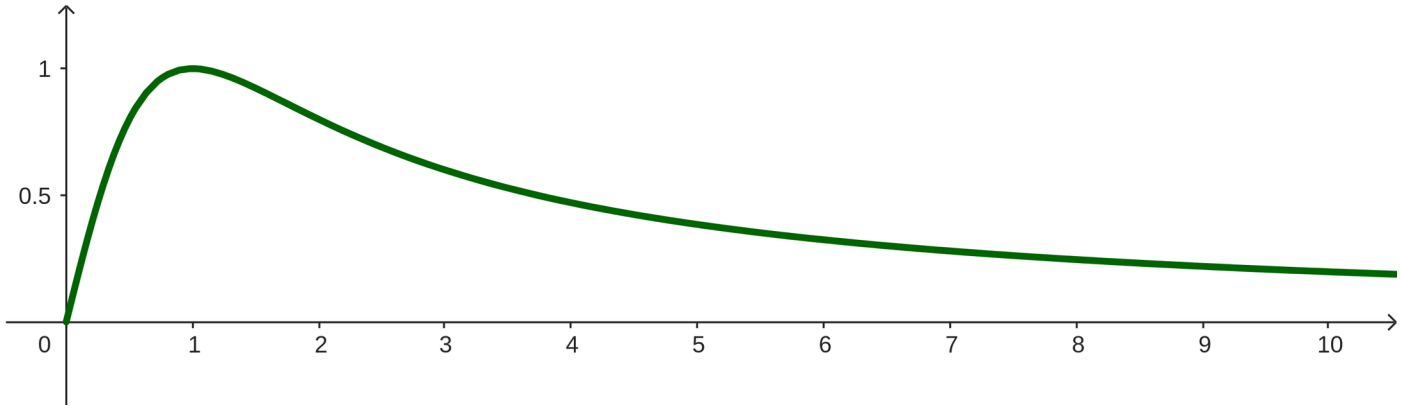


Figure 8.1.8 The function $f(x) = \frac{x}{x^2+1}$ plotted on the interval $[0, \infty)$.

When we plot the sequence $\{a_k\}$, though, we will visualize just the points on $f(x)$ at non-negative integer inputs.

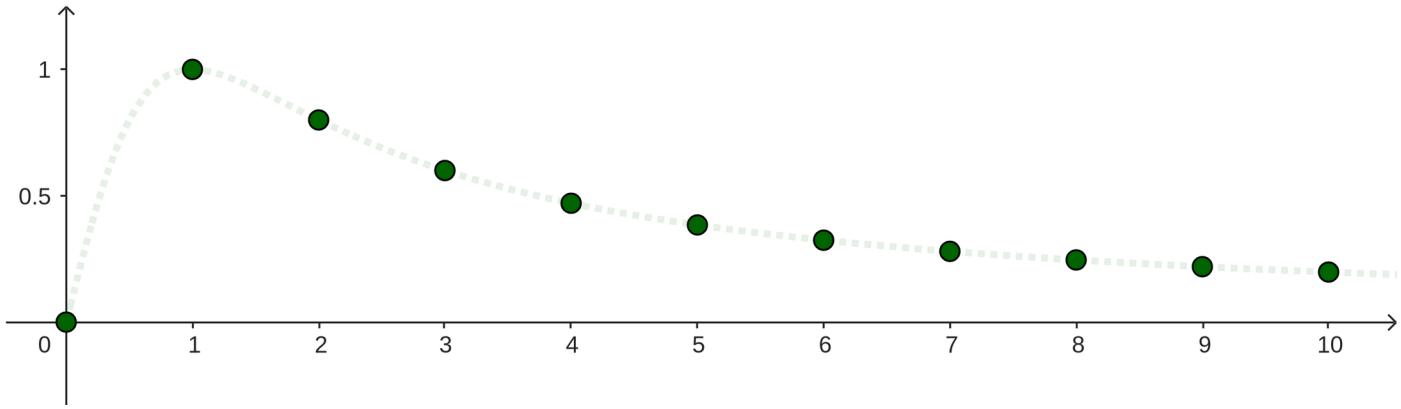


Figure 8.1.9 The sequence $\{a_k\}_{k=0}^{\infty}$ for $a_k = \dfrac{k}{k^2+1}$

We will typically not plot these with the smooth curve of the corresponding continuous function plotted, but in this first example it is useful to highlight how we think about this sequence as a function.

Let's continue to think about these sequences as just functions in a special kind of context. How does this discrete context change how we talk about functions and what kinds of terminology we use?

Sequence Terminology

If a sequence is a function (and we're saying in this introductory section that it is), then we can think of all of the different terminology and adjectives that we use to describe functions. How many of them are relevant to sequences?

- Continuous?
- Differentiable?
- Integrable?
- Increasing?
- Decreasing?

For now, we'll talk about sequences in two ways: their direction and the size of their terms.

Definition 8.1.10: Direction of a Sequence.

We say that a sequence $\{a_k\}_{k=1}^{\infty}$ is **increasing** if, for all $k = 1, 2, 3 \ldots, a_{k+1} > a_k$. If $a_{k+1} \geq a_k$ for all $k = 1, 2, 3, \ldots$ then we say that $\{a_k\}$ is **non-decreasing**.

We say that a sequence $\{a_k\}_{k=1}^{\infty}$ is **decreasing** if, for all $k = 1, 2, 3 \ldots, a_{k+1} < a_k$. If $a_{k+1} \leq a_k$ for all $k = 1, 2, 3, \ldots$ then we say that $\{a_k\}$ is **non-increasing**.

We say that a_k is constant if $a_{k+1} = a_k$, but this is a very boring sequence and we will likely not think terribly hard about these kinds of sequences.

Sometimes we might say that a sequence is **eventually decreasing** (or **eventually non-increasing**) if there is some $K > 1$, and the sequence is decreasing (or non-increasing) for $k = K, K+1, K+2, \ldots$, and similarly for **eventually increasing** or **eventually non-decreasing**.

A good example of a sequence that is eventually decreasing is the one we plotted in Figure 8.1.9. We can see that the sequence increases from $k = 0$ to $k = 1$ (since $a_0 < a_1$), but then decreases after that.

We could think about the corresponding continuous function (the one plotted in Figure 8.1.9) and find the point at which our function starts decreasing: we just need to refer back to Theorem 4.2.7 First Derivative Test to find the interval(s) for which $f(x) = \dfrac{x}{x^2+1}$ decreases.

Definition 8.1.11: Monotonic Sequences.

For the sequence $\{a_k\}_{k=1}^{\infty}$, we say that it is **monotonic** if the sequence is either non-increasing or non-decreasing.

We can, again, include a little modifier to talk about a sequence being eventually monotonic.

Definition 8.1.12: Bounded Sequences.

We say that a sequence $\{a_k\}_{k=1}^{\infty}$ is **bounded below** if there is some real number M such that $a_k \geq M$ for all $k = 1, 2, 3, \ldots$.

Similarly we say that a sequence $\{a_k\}_{k=1}^{\infty}$ is **bounded above** if there is some real number N such that $a_k \leq N$ for all $k = 1, 2, 3, \ldots$.

If a sequence has both an upper bound and a lower bound, then we often just say that the sequence is **bounded**.

Lastly, we'll focus on the end-behavior of a sequence. We'll think about convergence of a sequence in the same way that we did for Improper Integrals: does the limit exist?

Definition 8.1.13: Sequence Convergence.

For the sequence $\{a_k\}$, if L is some real number and $\lim_{k\to\infty} a_k = L$, the we say that the sequence $\{a_k\}$ **converges** to L. If this limit does not exist, we say that the sequence $\{a_k\}$ **diverges**.

Definition 8.1.14: Monotone Convergence Theorem.

If $\{a_k\}$ is a sequence that is both monotonic and bounded, then it must converge.

This theorem seems to be a bit obvious to many students: why would we care about this, when we can just find a limit of the explicit formula for a sequence? We'll see throughout the rest of this chapter that this theorem is one of the most important and most useful results in our study of infinite sequences and infinite series. For now, though, let's use it to find the limits of some recursively defined sequences.

Some Cool Recursive Examples

Let's revisit one of the recursively defined sequences that we've seen already and then think about a couple of other interesting ones. Before we do that, though, we should recognize why we need to treat recursively defined sequences a bit differently than ones defined explicitly.

In an explicit formula, the terms themselves are a function of k, the index. This means that we can simply apply a limit as $k \to \infty$ to understand whether or not the sequence converges and what it might converge to. These limits could be tricky, but we have the tools to evaluate them! In a recursion relation, though, each term is not a function of the index, which means we can't easily apply a limit as $k \to \infty$ to the term definition.

We'll be able to apply a limit, but it will feel a bit different: we're going to go into the computational work for the limit already assuming that the limit exists. Let's see how it goes.

Example 8.1.15

Let's revisit the first sequence from Example 8.1.5: $\{a_k\}_{k=1}^{\infty}$ where $a_1 = \frac{1}{3}$ and $a_k = 2\left(a_{k-1}\right)^2$.

a. Let's start by assuming that the sequence converges. That means that there exists some real number L such that
$$\lim_{k\to\infty} a_k = L$$
What would this L be, if it exists? A key thing to note is that if $\lim_{k\to\infty} a_k$ exists (and we have a symbol, L, for it) then we can say that
$$\lim_{k\to\infty} a_{k-1} = L$$
Whether or not this is obvious to you is not a mark of your understanding, but we need to make sure that this ends up being obvious to you. If it's not, that's ok! But it is an indicator that you should take a couple of minutes to think about this. Once you are convinced that these two limits are the same thing, move on to the next part.

b. Let's now apply a limit to the sequence definition:

$$\begin{aligned} a_k &= 2\left(a_{k-1}\right)^2 \\ \lim_{k\to\infty} a_k &= \lim_{k\to\infty} 2\left(a_{k-1}\right)^2 \\ \underbrace{\lim_{k\to\infty} a_k}_{L} &= 2(\underbrace{\lim_{k\to\infty} a_{k-1}}_{L})^2 \\ L &= 2L^2 \\ 0 &= 2L^2 - L \\ 0 &= L(2L-1) \end{aligned}$$

And so we have two solutions to this equation: $L = 0$ and $L = \frac{1}{2}$. This is strange: how can a sequence have more than one value that it converges to? It's because we have yet to take into account the initial term, a_1! Depending on this value, the sequence might converge or not, and if it does converge, then there are two options for what the sequence can converge to, based on the value of a_1.

c. You can do the next part on your own, but I want you to pick different numbers for a1 and write out some terms of the resulting sequence. You should find that some of them look like they're converging to 0, one of them will converge to $\frac{1}{2}$ (it's a fun hunt to find which one), and some will diverge.

d. Now it is up to us to show that this sequence, with $a_1 = \frac{1}{3}$, does converge. Sure, we have some evidence and a good conjecture that it converges to $\frac{1}{2}$, but that is just our good guess based on what we have seen in the first handful of numbers. We will attempt to convince ourselves that this sequence is both monotonic and bounded. We'll begin with boundedness. It should be clear that $a_k > 0$, since as long as $a_{k-1} \neq 0$, then $\left(a_{k-1}\right)^2 > 0$. Since we start with $a_1 \neq 0$, we are guaranteed to get non-zero values from the formula for a new term! Great news, we have a lower bound. Let's show that $\frac{1}{2}$ is an upper bound: $a_k < \frac{1}{2}$ when

$$\begin{aligned} 2\left(a_{k-1}\right)^2 &< \frac{1}{2} \\ \left(a_{k-1}\right)^2 &< \frac{1}{4} \\ a_{k-1} &< \frac{1}{2} \end{aligned}$$

Since $a_1 < \frac{1}{2}$, we know that each successive term will also be less than $\frac{1}{2}$. So we have an upper bound! So the sequence $\{a_k\}$ is bounded. Now we just need to convince ourselves that this sequence is monotonic. We know that our terms are bounded above by $\frac{1}{2}$, and I hope that this means we can convince ourselves that since our terms are smaller than this, which would produce a constant sequence, then all of our terms are probably decreasing. Let's show this by showing that $a_{k+1} - a_k < 0$:

$$\begin{aligned} a_{k+1} - a_k &= 2\left(a_k\right)^2 - a_k \\ &= a_k\left(2a_k - 1\right) \end{aligned}$$

We can solve for when this is negative! It shouldn't be hard to show that $a_k\left(2a_k - 1\right) < 0$ when $0 < a_k < \frac{1}{2}$. And we've already shown this is true in our case! So $\{a_k\}$ is bounded and monotonic and must therefore converge because of the Monotone Convergence Theorem. Because $a_1 < \frac{1}{2}$, we know that this sequence doesn't converge to $\frac{1}{2}$, and so must converge to the only other option: 0.

Solution

c. You should find that if $a_1 = \frac{1}{2}$, then the sequence is constant and converges to $\frac{1}{2}$. If $0 \leq |a_1| < \frac{1}{2}$ then the sequence seems like it'll converge to 0. And if $|a_1| > \frac{1}{2}$, then it looks like the sequence diverges.

There are some other fun ways of doing this same thing for other recursive examples. The argument above is relatively bulky to use, and so we understandably will not think about recursively defined sequences very much: we'll leave that topic for another course where we have more time to really explore them. If you are interested in trying this same argument with other sequences though, we'll end this section with two more fun examples.

Example 8.1.16

a. Consider the sequence defined by $b_k = \sqrt{2 + b_{k-1}}$ with $b_1 = \sqrt{2}$. Does this sequence converge? To what? **Hint**. Write out some terms to get a feel for things! Then, assuming that the sequence converges to some real number, L, think about what happens when you apply a limit as $k \to \infty$: we should get the equation $L = \sqrt{2} + L$.

b. Consider the sequence defined by $c_k = \dfrac{1}{2\left(c_{k-1}\right) + 1}$ with $c_1 = 1$. Does this sequence converge? To what? **Hint**. Write out some terms to get a feel for things! Then, assuming that the sequence converges to some real number, L, think about what happens when you apply a limit as $k \to \infty$: we should get the equation $L = \dfrac{1}{2L + 1}$.

Exercise 8.1.1

What is the index of an infinite sequence? How does the index relate to terms of the infinite sequence?

Exercise 8.1.2

Explain the differences between how an explicit formula and a recursion relation can define an infinite sequence.

Exercise 8.1.3

Why does a recursion relation defining a sequence rely on an initial value?

Exercise 8.1.4

Use the explicit formulas write out the first 8 terms of each sequence.

a. $\{a_k\} = \left\{\dfrac{1}{k}\right\}_{k=1}^{\infty}$

b. $\{b_k\} = \left\{\dfrac{k+2}{k^2+1}\right\}_{k=1}^{\infty}$

c. $\{c_k\} = \left\{\sin\left(\dfrac{k\pi}{4}\right)\right\}_{k=0}^{\infty}$

d. $\{d_k\} = \left\{\dfrac{1}{2^k}\right\}_{k=0}^{\infty}$

e. $\{j_k\} = \left\{\dfrac{(-1)^k}{k^2}\right\}_{k=1}^{\infty}$

f. $\{\ell_k\} = \left\{\dfrac{k+1}{k!}\right\}_{k=0}^{\infty}$

Exercise 8.1.5

For each of the following recursion relations, write out the first 8 terms of the sequence.

a. $a_k = \dfrac{a_{k-1} + 1}{2}; a_1 = 3$

b. $b_k = \dfrac{2}{b_{k-1}}; b_1 = 4$

c. $c_k = 2c_{k-1} - 2; c_1 = 3$

d. $d_k = \dfrac{d_{k-1}}{3}; d_1 = 2$

e. $j_k = 3\left(j_{k-1}\right)^2 - 1; j_1 = 1$

f. $\ell_k = \dfrac{1}{(\ell_{k-1})^2}; \ell_1 = 2$

Exercise 8.1.6

For each of the following sequences, write down a possibility for an explicit formula defining the sequence as well as a recursion relation (with initial value) defining the sequence.

a. $\{a_k\} = \{4, 5, 6, 7, ...\}$
b. $\{b_k\} = \{0, 2, 4, 6, ...\}$
c. $\{c_k\} = \left\{\dfrac{3}{4}, \dfrac{1}{2}, \dfrac{1}{3}, \dfrac{2}{9}, ...\right\}$
d. $\{d_k\} = \left\{\dfrac{1}{8}, \dfrac{3}{16}, \dfrac{1}{4}, \dfrac{5}{16}, ...\right\}$

Exercise 8.1.7

For each of the following sequences, describe whether they converge or diverge. If they converge, show what they converge to using limit work.

a. $\{a_k\} = \left\{3 + \dfrac{(-1)^{k+1}}{k}\right\}_{k=1}^{\infty}$
b. $\{b_k\} = \{\sqrt[k]{k}\}_{k=1}^{\infty}$
c. $\{c_k\} = \left\{\dfrac{k!}{e^k}\right\}_{k=0}^{\infty}$
d. $\{d_k\} = \left\{\left(1 + \dfrac{2}{k}\right)^k\right\}_{k=1}^{\infty}$
e. $\{j_k\} = \left\{\dfrac{k^3 + 1}{k^2 + \ln(k)}\right\}_{k=1}^{\infty}$
f. $\{\ell_k\} = \{k \ln(1/k)\}_{k=1}^{\infty}$

Exercise 8.1.8

Compare a sequence to a function. What are some similarities? What are some differences? Carefully discuss this in terms of different representations of sequences/functions.

Exercise 8.1.9

Choose through the sequences in any of the problems above to find one to fit each of the descriptions below. Explain your choices for each.

a. Monotonic non-decreasing.
b. Monotonic non-increasing.
c. Bounded. What are the upper and lower bounds?

This page titled 8.1: Introduction to Infinite Sequences is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Peter Keep.

8.2: Introduction to Infinite Series

Let's try to introduce the idea of an infinite series using a framework that we know and are (maybe) comfortable with: integrals!

With an integral, we have a nice way of evaluating integrals of nicely behaved functions with finite limits of integration (Fundamental Theorem of Calculus (Part 2)).

Then, when we talked about improper integrals, we built a nice way to think about evaluating integrals with unbounded limits of integration (Evaluating Improper Integrals (Infinite Width)). How will we use this to think about infinite series, a sum of the infinitely many terms from an infinite sequence?

Partial Sums

If we approach infinite series in a manner similar to improper integrals, then we will need to do a couple of things.

1. Truncate the infinite series at some finite ending point. This is what we did with the integral, when we replaced the infinity with some real number variable t. We might use n for the series "ending index."
2. Find a formula for this truncated/finite version. For the integrals, we could use the Fundamental Theorem of Calculus (Part 2) for this! For series, we'll need to do something else.
3. Apply a limit as t (or n in the case of infinite series) goes off to infinity!

Activity 8.2.1: How Do We Think About Infinite Series?

Let's consider the following sequence:

$$\left\{\frac{9}{10^k}\right\}_{k=1}^{\infty}$$

a. Write out the first 5 terms of the sequence.
b. What does this sequence converge to? Show this with a limit!
c. Now we'll construct a new sequence, this time by adding things up. We're going to be working with the sequence $\{S_n\}_{n=1}^{\infty}$ where

$$S_n = \sum_{k=1}^{k=n}\left(\frac{9}{10^k}\right)$$

Write out the first five terms of this sequence: S_1, S_2, S_3, S_4, S_5.
d. Can you come up with an explicit formula for S_n?
e. Does $\{S_n\}$ converge or diverge? Use a limit to find what it converges to!
f. What do you think this means for the infinite series $\sum_{k=1}^{\infty}\left(\frac{9}{10^k}\right)$? Does the infinite series converge or diverge?

This is hopefully a nice little introduction to how we'll think about infinite series: we'll consider, instead, the sequence of sums where we add up more and more terms. This is also a nice first example, because we really just showed that

$$0.999\ldots = 1$$

since

$$\begin{aligned}\sum_{k=1}^{\infty}\left(\frac{9}{10^k}\right) &= \frac{9}{10} + \frac{9}{100} + \frac{9}{1000} + \ldots \\ &= 0.9 + 0.09 + 0.009 + \ldots \\ &1.\end{aligned}$$

But more importantly, we now have a good strategy for thinking about infinite series as sequences of partial sums.

Definition 8.2.2: Partial Sum.

For an infinite series $\sum_{k=1}^{\infty} a_k$, we call $S_n = \sum_{k=1}^{n} a_k$ the nth Partial Sum of the infinite series.

Definition 8.2.3: Series Convergence.

We say that the infinite series $\sum_{k=1}^{\infty} a_k$ **converges** to the real number L if the sequence $\{S_n\}_{n=1}^{\infty}$ converges to L (where $\lim_{n\to\infty} S_n = L$), where $S_n = \sum_{k=1}^{n} a_k$ is the nth partial sum of the infinite series.

If the sequence of partial sums $\{S_n\}_{n=1}^{\infty}$ diverges (the limit $\lim_{n\to\infty} S_n$ does not exist), then we say that the infinite series $\sum_{k=1}^{\infty} a_k$ **diverges**.

Visualizing the Sequence of Partial Sums

Since we'll think about an infinite series $\sum_{k=0}^{\infty} a_k$ as the sequence of its partial sums, $\{\sum_{k=0}^{n} a_k\}_{n=0}^{\infty}$, then we can think about visualizing an infinite series as really the same thing as visualizing a sequence in general (Graphing Sequences).

Let's consider, as a first (visual) example, the infinite series:

$$\sum_{k=0}^{\infty} \frac{3}{k+1}$$

We can think about the two important sequences that we'll consider:

$$\underbrace{\left\{\frac{3}{n+1}\right\}_{n=0}^{\infty}}_{\{a_n\},\text{ the sequence of terms}} \quad \text{and} \quad \underbrace{\left\{\sum_{k=0}^{n} \frac{3}{k+1}\right\}_{n=0}^{\infty}}_{\{S_n\},\text{ the sequence of partial sums}}$$

We can plot the sequence of terms, $\left\{\frac{3}{n+1}\right\}_{n=0}^{\infty}$, and visualize the limit $\lim_{n\to\infty} a_n = 0$. This sequence of terms converges to 0.

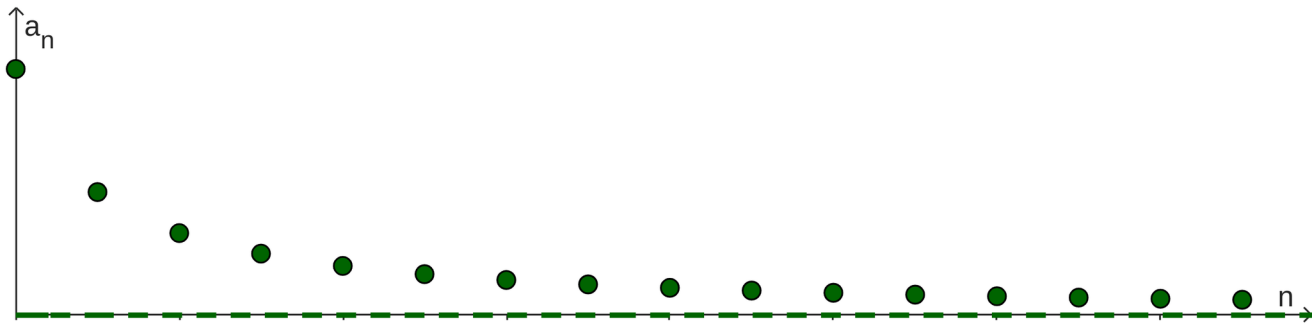


Figure 8.2.4 $\{a_n\}$, the sequence of terms in the series.

Then, we can compare this with the plot of the partial sums, $\{S_n\}$ where:

$$S_n = \sum_{k=0}^{n} \frac{3}{k+1}$$

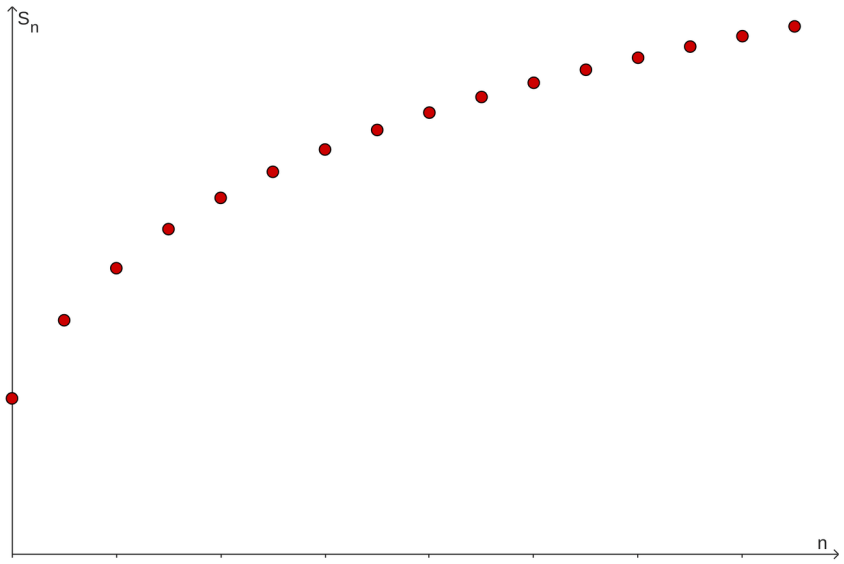


Figure 8.2.5 $\{S_n\}$, the sequence of partial sums for the series.

This image is fine, but not very good at showing how the sequence of terms and the sequence of partial sums are related to each other. We should note that each point in Figure 8.2.5 is the accumulation of the heights of the preceding points in Figure 8.2.4. We can visualize this to make it easier by overlaying some information onto the plot of partial sums in Figure 8.2.5.

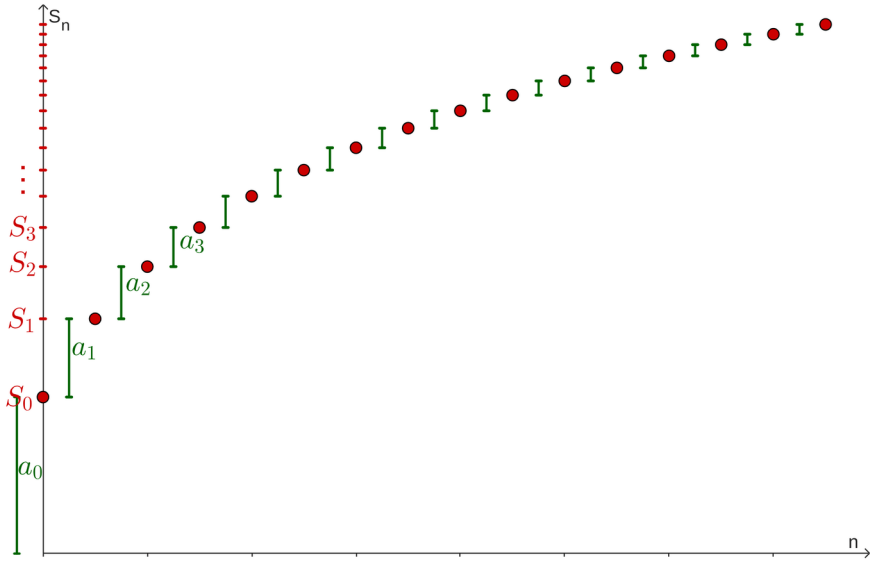
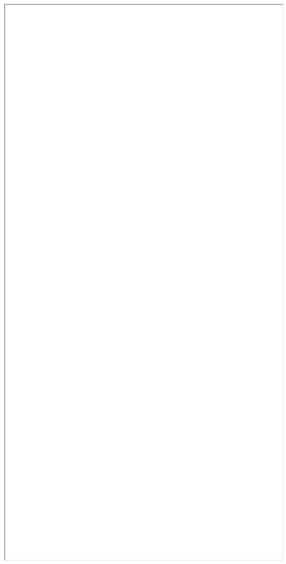


Figure 8.2.6 $\{S_n\}$ and $\{a_n\}$ visualized together.

Hopefully this does a good job of illustrating the connections between the two:

$$
\begin{aligned}
S_0 &= a_0 \\
S_1 &= a_0 + a_1 \\
&= S_0 + a_1 \\
S_2 &= a_0 + a_1 + a_2 \\
&= S_1 + a_2 \\
S_3 &= a_0 + a_1 + a_2 + a_3 \\
&= S_2 + a_3 \\
&\vdots \\
S_n &= a_0 + a_1 + \ldots + a_n \\
&= S_{n-1} + a_n
\end{aligned}
$$

In order to further examine the connection between sequences and partial sums, let's look at an interactive example, where you can create the terms of a sequence (however you'd like) and examine the corresponding partial sums.



Standalone
In Context
Embed

Note 8.2.7: Finding Explicit Formulas.

We had noted earlier (in Section 8.1) that it was hard to find explicit formulas (or recursion relations) for sequences where we had the first few terms.

This remains true when we think about finding formulas for the sequences of partial sums. Notice that it is easy to find the location of the horizontal asymptote in Figure 8.2.4 (by evaluating $\lim_{n\to\infty} a_n$), but that we did not attempt to find one to for the partial sums in Figure 8.2.5 or Figure 8.2.6.

If you'd like to try this, then we need to find a formula for S_n. Try to find the first several partial sums by adding up terms in the series. Then try to find a formula to predict the next partial sum. This will definitely not be easy!

Ok, actually, this will be an impossible task. There is no closed-form formula for this. We cannot simply find $\lim_{n\to\infty} S_n$ in the way that we've found the limit of the sequence of terms.

Special Series

Let's look at three examples where we can think about partial sums and play with our new idea of series convergence.

Example 8.2.2

For each of the following series, write out a few of the terms of the series. Then write out the corresponding partial sums. Use these to find a formula for S_n, the nth partial sum. Then make a claim about whether or not the series converges and what it converges to.

a. $\sum_{k=0}^{\infty}\left(\frac{1}{2^k}\right)$

b. $\sum_{k=2}^{\infty}\left(\frac{k}{\ln(k)}-\frac{k+1}{\ln(k+1)}\right)$

c. $\sum_{k=2}^{\infty}\left(\frac{2}{k^2-1}\right)$ **Hint**. This one is tricky! It's hard to notice anything unless we write out the series term formula a bit differently. Use Partial Fractions to rewrite $\frac{2}{k^2-1}$ as $\frac{1}{k-1}-\frac{1}{k+1}$

Solution

a.

$$\begin{aligned}
S_0 &= 1\\
S_1 &= 1+\frac{1}{2}=\frac{3}{2}\\
S_2 &= 1+\frac{1}{2}+\frac{1}{4}=\frac{7}{4}\\
S_3 &= 1+\frac{1}{2}+\frac{1}{4}+\frac{1}{8}=\frac{15}{8}\\
S_n &= 1+\frac{1}{2}+\frac{1}{4}+\ldots+\frac{1}{2^n}=\frac{2^{n+1}-1}{2^n}\\
&= 2-\frac{1}{2^n}\\
S_n &= \lim_{n\to\infty}\left(2-\frac{1}{2^n}\right)\\
&= 2
\end{aligned}$$

The series converges to 2.

b.

$$\begin{aligned}
S_2 &= \frac{2}{\ln(2)}-\frac{3}{\ln(3)}\\
S_3 &= \frac{2}{\ln(2)}\underbrace{-\frac{3}{\ln(3)}+\frac{3}{\ln(3)}}_{0}-\frac{4}{\ln(4)}\\
&= \frac{2}{\ln(2)}-\frac{4}{\ln(4)}\\
S_4 &= \frac{2}{\ln(2)}\underbrace{-\frac{3}{\ln(3)}+\frac{3}{\ln(3)}}_{0}\underbrace{-\frac{4}{\ln(4)}+\frac{4}{\ln(4)}}_{0}-\frac{5}{\ln(5)}\\
&= \frac{2}{\ln(2)}-\frac{5}{\ln(5)}\\
S_5 &= \frac{2}{\ln(2)}\underbrace{-\frac{3}{\ln(3)}+\frac{3}{\ln(3)}}_{0}\underbrace{-\frac{4}{\ln(4)}+\frac{4}{\ln(4)}}_{0}\underbrace{-\frac{5}{\ln(5)}+\frac{6}{\ln(6)}}_{0}-\frac{6}{\ln(6)}\\
&= \frac{2}{\ln(2)}-\frac{6}{\ln(6)}\\
S_n &= \frac{2}{\ln(2)}\underbrace{-\frac{3}{\ln(3)}+\frac{3}{\ln(3)}}_{0}\underbrace{-\frac{4}{\ln(4)}+\frac{4}{\ln(4)}}_{0}\underbrace{-\frac{5}{\ln(5)}+\ldots+\frac{n}{\ln(n)}}_{0}-\frac{n+1}{\ln(n+1)}\\
&= \frac{2}{\ln(2)}-\frac{n+1}{\ln(n+1)}\\
\lim_{n\to\infty}S_n &= \lim_{n\to\infty}\frac{2}{\ln(2)}-\frac{n+1}{\ln(n+1)}\\
&= -\infty
\end{aligned}$$

So the series $\sum_{k=2}^{\infty}\left(\frac{k}{\ln(k)}-\frac{k+1}{\ln(k+1)}\right)$ diverges.

c.

$$\sum_{k=2}^{\infty}\left(\frac{2}{k^2-1}\right)=\sum_{k=2}^{\infty}\left(\frac{1}{k-1}-\frac{1}{k+1}\right)$$

$$\begin{aligned}
S_2 &= 1-\frac{1}{3}\\
S_3 &= 1-\frac{1}{3}+\frac{1}{2}-\frac{1}{4}\\
S_4 &= 1-\frac{1}{3}+\frac{1}{2}-\frac{1}{4}+\frac{1}{3}-\frac{1}{5}\\
&= 1+\frac{1}{2}-\frac{1}{4}-\frac{1}{5}\\
S_5 &= 1-\frac{1}{3}+\frac{1}{2}-\frac{1}{4}+\frac{1}{3}-\frac{1}{5}+\frac{1}{4}-\frac{1}{6}\\
&= 1+\frac{1}{2}-\frac{1}{5}-\frac{1}{6}\\
S_n &= 1-\frac{1}{3}+\frac{1}{2}-\frac{1}{4}+\frac{1}{3}-\frac{1}{5}+\frac{1}{4}-\frac{1}{6}+\ldots+\underbrace{\frac{1}{n-2}-\frac{1}{n}}_{a_{n-1}}+\underbrace{\frac{1}{n-1}-\frac{1}{n+1}}_{a_n}\\
&1+\frac{1}{2}-\frac{1}{n}-\frac{1}{n+2}\\
S_n &= \lim_{n\to\infty}\left(1+\frac{1}{2}-\frac{1}{n}-\frac{1}{n+2}\right)\\
&= \frac{3}{2}
\end{aligned}$$

These examples are a bit misleading: we often won't be able to do this kind of thing! For most infinite series, we will struggle to find an explicit formula for the nth partial sum. In these examples, though, we took advantage of some specific structure.

In this first example (as well as the example in Activity 8.2.1), we noticed that because of the exponential function defining the terms, we were able to find some nice patterns in the partial sums. We'll explore this a bit more later in Section 8.6.

Then in these other two examples, we noticed that once we could write each term as a difference of two fractions that have a really similar structure, we got these "repeat" values from term to term where the opposite signs made things add up to 0. These are called "telescoping series," and they're mostly fun examples to think about partial sum formulas. We'll see some pop up later though, and Partial Fraction Decomposition is a nice trick to keep in mind for these kinds of things.

Practice Problems

Exercise 8.2.1

What does it mean for an infinite series $\sum_{k=1}^{\infty}a_k$ to converge? What does it mean for this series to diverge?

Exercise 8.2.2

What is a partial sum, S_n, of the infinite series $\sum_{k=1}^{\infty}a_k$? Why do we care about this?

Exercise 8.2.3

For some infinite series $\sum_{k=1}^{\infty}b_k$, we consider the partial sum

$$S_n=\sum_{k=1}^{n}b_k$$

For each of the following descriptions of S_n, explain what conclusions we can make about the infinite series. Be specific, and justify your answers.

a. $S_n = \dfrac{3n^2}{2n^2+1}$
b. $\{S_n\}$ is increasing and bounded.
c. $0 \le S_n \le 1$ for all $n \ge 1$.
d. $4 - \dfrac{1}{n} \le S_n \le 4 + \dfrac{1}{n}$ for all $n \ge 1$

Exercise 8.2.4

Consider the infinite series $\sum_{k=1}^{\infty} c_k = 0.6 + 0.06 + 0.006 + 0.0006 + ...$

a. Write out an explicit formula for the sequence of terms $\{c_k\} = \{0.6, 0.06, 0.006, 0.0006, ...\}$.
b. Write out the partial sums S_1, S_2, S_3, and S_4.
c. Write out an explicit formula for the nth partial sum, S_n.
d. Does the sequence $\{S_n\}$ converge or diverge? Explain. If it converges, what does it converge to?
e. Does the series $\sum_{k=1}^{\infty} c_k$ converge or diverge? Explain. If it converges, what does it converge to.

Exercise 8.2.5

Consider the infinite series $\sum_{k=1}^{\infty} \dfrac{1}{k(k+1)}$.

a. Use Partial Fraction Decomposition to rewrite this series.
b. Write out the first five terms, and the partial sum S_5.
c. Write out the explicit formula for the nth partial sum, S_n.
d. Does the sequence $\{S_n\}$ converge or diverge? Explain. If it converges, what does it converge to?
e. Does the series $\sum_{k=1}^{\infty} \dfrac{1}{k(k+1)}$ converge or diverge? Explain. If it converges, what does it converge to?

This page titled 8.2: Introduction to Infinite Series is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Peter Keep.

8.3: The Divergence Test and the Harmonic Series

The Relationship Between a Sequence and Series

We have looked at both infinite sequences and infinite series so far, and, to make things complicated, we're really thinking about an infinite series (of terms from an infinite sequence) as an infinite sequence (of partial sums of the series). We've looked at how to visualize these (in both Graphing Sequences and Visualizing the Sequence of Partial Sums).

Let's first start with defining a new series. This is a relatively important one by itself (it does have its own name), but it's mostly an important series because it leads us into some new and interesting ways of thinking about series in general.

Definition 8.3.1: Harmonic Series.

We call the series

$$\sum_{k=1}^{\infty} \frac{1}{k}$$

the **Harmonic Series**.

You might not recognize this, but we've worked with a version of this before. The example series that we plotted in Figure 8.2.6 was

$$\sum_{k=0}^{\infty} \frac{3}{k+1}$$

We can notice that if we re-index this by starting at $k = 1$ instead of $k = 0$, we were really just looking at a scaled version of the harmonic series.

$$\begin{aligned} \sum_{k=0}^{\infty} \frac{3}{k+1} &= \sum_{k=1}^{\infty} \frac{3}{k} \\ &= 3\left(\sum_{k=1}^{\infty} \frac{1}{k}\right) \end{aligned}$$

Activity 8.3.2: Investigating the Harmonic Series.

a. Write out the first several terms of the harmonic series, terms from $\left\{\frac{1}{k}\right\}_{k=1}^{\infty}$. Write however many you need to get a feel for how the terms work.

b. Can you find out how many terms you would have to go "into" the series before the term was less than 0.000000001? **Hint**. When is $\frac{1}{k} < \frac{1}{10^8}$?

c. Can you do this same kind of thing, no matter how small? For instance, how many terms would you have to go into the series before the term was less than some real number ε where $\varepsilon > 0$? **Hint**. When is $\frac{1}{k} < \varepsilon$?

d. Remind/explain/convince yourself that what we've really done is show that $\lim_{k\to\infty} \frac{1}{k} = 0$. This isn't a new or terribly interesting fact, but make sure that you understand why the argument above shows this.

e. Let's do something very similar, but with $\left\{\sum_{k=1}^{n} \frac{1}{k}\right\}_{n=1}^{\infty}$, the sequence of partial sums, instead. Write out the first few partial sums. There's no specific number that you need to write, but make sure to write enough partial sums to get a feel for how the partial sums work.

f. Can you find out how many terms you need to add up until the partial sum is larger than 1? **Hint**. Find a value for n to give $\sum_{k=1}^{n} \frac{1}{k} > 1$.

g. Can you find out how many terms you need to add up until the partial sum is larger than 5? **Hint**. Find a value for n to give $\sum_{k=1}^{n} \frac{1}{k} > 5$.

h. Can you find out how many terms you need to add up until the partial sum is larger than 10? **Hint**. Find a value for n to give $\sum_{k=1}^{n} \frac{1}{k} > 10$. This will be absolutely awful to try calculating by hand! Use some piece of technology!

i. Do you think that for any positive number S, we can always find some partial sum $\sum_{k=1}^{n} \frac{1}{k} > S$? What do you think this would mean about

$$\lim_{n\to\infty} S_n = \lim_{n\to\infty} \sum_{k=1}^{n} \frac{1}{k}?$$

To actually show that for any $S > 0$ we could always find an $n > 1$ where

$$\sum_{k=1}^{n} \frac{1}{k} > S$$

is an extremely difficult task! We will show that the Harmonic Series diverges in a different way, but for now I want us to notice these contradictory results: we have a series whose terms get small, but whose partial sums do not seem to converge. We have $\frac{1}{k} \to 0$ but it seems like $\lim_{n\to\infty} S_n$ does not exist. Is this behavior special to the Harmonic Series? Is this something we should make note of? Is there some other connection between the terms of a series and the behavior of the partial sums of the series that we need to note?

Answer available at the end of this section.

Let's continue to think about this strange series, but actually prove that the series itself diverges.

Theorem 8.3.3: The Harmonic Series Diverges.

The **Harmonic Series**,

$$\sum_{k=1}^{\infty} \frac{1}{k},$$

diverges.

Proof. available at the end of this section.

This is a strange result, and one that has been brought up again and again by mathematicians throughout history. We'll see that this series is notable because of its use later on in this chapter, but for now we can simply note that it is strange to see a series of terms that get so small (and so quickly) and yet the sum of those terms diverges.

The Divergence Test

The connection between the terms of a series and the behavior of the infinite series itself is maybe more mysterious than we initially thought. Since we will likely not have "access" to the formula for the partial sums (Note 8.2.7), we will want to explore these kinds of connections as much as we can. They will be the things to help us analyze an infinite series.

Activity 8.3.4: Thinking about the Terms of a Series.

a. What does it mean for the series $\sum_{k=1}^{\infty} a_k$ to converge? Hint. If you need a reminder, take a look back at Definition 8.2.3 Series Convergence

b. Calculate the difference between $S_{n+1} = \sum_{k=1}^{n+1} a_k$ and $S_n = \sum_{k=1}^{n} a_k$, $S_{n+1} - S_n$.
c. If the sequence of partial sums for a series, $\{S_n\}_{n=1}^{\infty}$ converges, then what should we expect to happen with the values of S_{n+1} and S_n? What does this mean for the difference $S_{n+1} - S_n$?
d. Now consider a series that we know diverges, the Harmonic Series. Find the limit,

$$\lim_{k\to\infty} \frac{1}{k}$$

We have a pretty soft result here: we have discovered a connection between the terms of a series and the series behavior itself, but it's not as strong as we'd like.

Theorem 8.3.5: Divergence Test.

For an infinite series $\sum_{k=0}^{\infty} a_k$, if the infinite series converges then

$$\lim_{k\to\infty} a_k = 0.$$

This is equivalent to saying that if

$$\lim_{k\to\infty} a_k \neq 0$$

then the infinite series $\sum_{k=0}^{\infty} a_k$ diverges.

Proof. available at the end of this section.

We should notice that the Divergence Test never tells us that a series converges. The only conclusions are that a series diverges (if $\lim_{k\to\infty} a_k \neq 0$) or the test fails/is inconclusive (if $\lim_{k\to\infty} a_k = 0$). In these cases, we've seen instances where a series could converge or diverge and have terms that approach 0.

Example 8.3.6

Apply the Divergence Test to the following series and interpret the results.

a. $\sum_{k=0}^{\infty} \frac{k^{15} - 4k^{10} + 10k^4}{e^{2k}}$ **Hint**. We can do a couple of things here! There is a nice result about limits of polynomials that we can use in the numerator (Polynomial End Behavior Limits). We could also get this same result using some other techniques, like what we use to prove that theorem. Or we can use L'Hôpital's Rule to evaluate the limit, since we have a $\frac{\infty}{\infty}$ indeterminate form.
b. $\sum_{k=1}^{\infty} \frac{(-1)^{k+1}}{k^2+1}$ **Hint**. These terms are strange! The $(-1)^{k+1}$ part really just impacts the sign of the terms, since it is either 1 or –1 depending on if k is even or odd. We can consider only one sign (maybe the positive), and then try to make a conclusion about the alternating terms. Do they go to 0?
c. $\sum_{k=1}^{\infty} \frac{(-1)^{k+1}}{\sqrt[k]{k}}$ **Hint**. This is similar: focus on only the positive terms for now. But that denominator is also strange! If you want to focus only on the denominator, you can use the following friendly rearrangement:

$$\begin{aligned}\lim_{k\to\infty} \sqrt[k]{k} &= \lim_{k\to\infty} k^{\frac{1}{k}} \\ &= \lim_{k\to\infty} e^{\ln(k^{1/k})} \\ &= \lim_{k\to\infty} e^{\frac{1}{k}\ln(k)} \\ &= e^{\lim_{k\to\infty} \frac{\ln(k)}{k}}\end{aligned}$$

Now you can use L'Hôpital's Rule to evaluate this part!

Squeeze Theorem and Growth Rates

In order for us to calculate the limits of the terms of an infinite series, we are often using results like The Squeeze Theorem and L'Hôpital's Rule. Let's observe two facts that will help us out.

We don't need to apply the Squeeze Theorem in "full," since we're only concerned with whether the limit is 0 or not. Let's clarify with an example:

$$\sum_{k=1}^{\infty} \frac{(-1)^{k+1}2k^2}{7k^2+k+1} = \frac{2}{9} - \frac{8}{31} + \frac{18}{67} - \frac{32}{117} + - \ldots$$

For us to apply the Divergence Test, we really should note that:

$$-\frac{2k^2}{7k^2+k+1} \leq \frac{(-1)^{k+1}2k^2}{7k^2+k+1} \leq \frac{2k^2}{7k^2+k+1}$$

We could look at the limits of both of the bounds, but since we're concerned with whether they are 0 or not, we can simply check one:

$$\lim_{k\to\infty} \left(\frac{2k^2}{7k^2+k+1} \right) = \frac{2}{7}$$

Since we know that the negative version of this limit isn't 0 either, we can see that the limit $\lim_{k\to\infty} \frac{(-1)^{k+1}2k^2}{7k^2+k+1}$ doesn't exist. Thus, the series diverges.

We can test just the positive version of the limit in each of these cases, then!

Lastly, we can notice now that L'Hôpital's Rule can be difficult to apply, especially since the functions we're working with are not required to be differentiable. For instance, how could we consider the series

$$\sum_{k=0}^{\infty} \frac{k!}{3^k}$$

by thinking about the limit

$$\lim_{k\to\infty} \frac{k!}{3^k}?$$

Aside.

The factorial function, $k!$, is formally defined as a recursive function, where $0! = 1$ and $n! = n(n-1)!$. More informally, this really means that a factorial really tells us to multiply the integer by all of the positive integers less than it. So:

$$7! = 7(6)(5)(4)(3)(2)(1)$$

We define $0! = 1$ for convenience, but a nice explanation of why we might do that is based on counting arrangements.

If we were to arrange 3 items on a shelf, we could count the number of total arrangements by thinking about how many options we have to place the first item on the shelf (3) multiplied by how many options we have for the second item (only 2 remaining) and then how many options we have for the last item (once we've placed the other two, we only have 1 option). So the number of arrangements is 3! = 3(2)(1).

How many ways can you arrange an empty shelf?

Because we can't think about the derivative of k! (since the function is not continuous, and thus not differentiable in the traditional sense), we need to consider this limit in a different way. We can instead appeal to **growth rates** of these functions: that is, the rate at which these functions approach infinity.

Growth Rates of Functions.

If m and n are real numbers with $0 < m < n$, then we can arrange the growth rates of the following functions from "slow" to "fast:"

$$\ln(k) \ll k^m \ll k^m \ln(k) \ll k^n \ll m^k \ll n^k \ll k! \ll k^k$$

So, returning to our series $\sum_{k=0}^{\infty} \frac{k!}{3^k}$, we can see that

$$\lim_{k\to\infty} \frac{k!}{3^k} = \infty$$

since $k!$ has a "faster" growth rate than 3^k.

Note 8.3.7

It's worth investigating this yourself! Try to convince yourself that you understand why exponentials grow "slower" than factorials. Why does k^k have a "faster" growth rate than $k!$?

Exercise 8.3.1

Give an example of a series that diverges due to the Divergence Test. Explain.

Exercise 8.3.2

Give an example of a series that fails the Divergence Test. Do you know if this series converges or diverges? Explain.

Exercise 8.3.3

Apply the Divergence Test to each of the following series. Explain the conclusions of the test for each series.

a. $\sum_{k=1}^{\infty} \left(\frac{k}{k^2+1} \right)$
b. $\sum_{k=1}^{\infty} \left(\frac{k}{e^k} \right)$
c. $\sum_{k=2}^{\infty} \left(\frac{1}{k\ln(k)} \right)$
d. $\sum_{k=2}^{\infty} \left(\frac{k}{\ln(k)} \right)$
e. $\sum_{k=1}^{\infty} (\sqrt[k]{k})$
f. $\sum_{k=0}^{\infty} \left(\frac{k^2+k}{e^k} \right)$
g. $\sum_{k=1}^{\infty} \left(\frac{1}{k^2+1} \right)$
h. $\sum_{k=1}^{\infty} \left(\frac{k!}{k^2} \right)$
i. $\sum_{k=1}^{\infty} \left(\frac{(-1)^k(k+1)}{k+\frac{1}{k}} \right)$

Answers

Activity 8.3.2: Investigating the Harmonic Series.

g. $\sum_{k=1}^{83} \frac{1}{k} \approx 5.00207\ldots$

This is the first partial sum greater than 5.

h. $\sum_{k=1}^{12367} \frac{1}{k} \approx 10.000043\ldots$

This is the first partial sum greater than 10.

Theorem 8.3.3: The Harmonic Series Diverges.

Let's assume, for the sake of eventual contradiction, that the harmonic series converges. Our goal in this proof is to show that this assumption (convergence) logically leads to an internal contradiction. This would mean that the assumption (convergence) cannot be true.

So, let's assume that the harmonic series converges.

Based on our definition of series convergence (Definition 8.2.3), there exists some real number S such that:

$$\lim_{n\to\infty} \sum_{k=1}^{n} \frac{1}{k} = S$$

We're going to think about this number, S, and show that there cannot be such a number.

First, let's write out what S is:

$$S = 1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \frac{1}{5} + \frac{1}{6} + \ldots$$

We're now going to systematically change the numbers being added together in order to create some number that is smaller than S: we're going to take all of the odd terms and make them as small as the next term after it:

$$S = 1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \frac{1}{5} + \frac{1}{6} + \ldots$$
$$S > \frac{1}{2} + \frac{1}{2} + \frac{1}{4} + \frac{1}{4} + \frac{1}{6} + \frac{1}{6} + \ldots$$

Note, though, that we can group together these duplicate terms and add them. Let's do that!

$$S > \left(\frac{1}{2} + \frac{1}{2}\right) + \left(\frac{1}{4} + \frac{1}{4}\right) + \left(\frac{1}{6} + \frac{1}{6}\right) + \ldots$$
$$S > 1 + \frac{1}{2} + \frac{1}{3} + \ldots$$

But we should recognize this new series that is smaller than S...it's the harmonic series! Which, by our initial assumption, is also S!

Ok, so what we have shown is that if the harmonic series converges, then it converges to some number S that has the contradictory property of being smaller than itself.

There is no such number.

This is a contradiction, then, and so the harmonic series must diverge.

Theorem 8.3.5: Divergence Test.

We will prove the claim that if an infinite series converges, then its sequence of terms must converge to 0.

This result will fall out of a simple exploration of what partial sums are. We noted in Section 8.2 that we can write any partial sum as the sum of the previous partial sum and the next term:

$$\begin{aligned} S_0 &= a_0 \\ S_1 &= a_0 + a_1 \\ &= S_0 + a_1 \\ S_2 &= a_0 + a_1 + a_2 \\ &= S_1 + a_2 \\ S_3 &= a_0 + a_1 + a_2 + a_3 \\ &= S_2 + a_3 \\ &\vdots \\ S_n &= a_0 + a_1 + \ldots + a_n \\ &= S_{n-1} + a_n \end{aligned}$$

Let's now say that the series we are dealing with converges. This means that $\lim_{n\to\infty} S_n = S$ for some real number S.

What, then, would the limit of S_{n-1} be as $n \to \infty$?

It has to also be S! If the partial sums converge, then these two partial sums must converge to each other n increases:

$$\lim_{n\to\infty} S_{n-1} = \lim_{n\to\infty} S_n$$

So, since $S_n = S_{n-1} + a_n$, we can investigate the limit of a_n:

$$\begin{aligned} a_n &= S_n - S_{n-1} \\ \lim_{n\to\infty} a_n &= \lim_{n\to\infty} (S_n - S_{n-1}) \\ &= 0 \end{aligned}$$

So of course the nth term has to converge to 0 in the limit!

This page titled 8.3: The Divergence Test and the Harmonic Series is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Peter Keep.

8.4: The Integral Test

Infinite Series As a Kind of Integral

Let's start here with a connection between objects. We've already thought about the connection between an infinite series and the sequence of terms of the series (Theorem 8.3.5 Divergence Test). Now we'll think about the connections between two objects that are similar to each other in that they both represent an accumulation of function values.

Earlier (in Sequences as Functions) we tried to describe sequences as just a special kind of function: the domain is the set of non-negative integer (or positive integers, depending on whether we start our index at $n = 0$ or $n = 1$) and we map these inputs to real number outputs. And now we want to think about what it might mean to accumulate the values of these kinds of functions.

Function value accumulation is what we've been looking at lately! That's what integration is! We are trying to accumulate all of the function values and weigh them based on their "width." In the context of continuous functions, that means we start approximating this accumulation by looking at some finite number of function values that we pick, and we give them some Δx width between them. That's our Riemann sum:

$$\sum_{k=1}^{n} f\left(x_k^*\right) \Delta x$$

And from there, we work on making that space between function values get smaller (as the number of function values we use gets higher). So when n is the number of function values, we can let $n \to \infty$ and correspondingly we get $\Delta x \to dx$, the differential in our integral:

$$\lim_{n\to\infty} \sum_{k=1}^{n} f\left(x_k^*\right) \Delta x = \int_{x=a}^{x=b} f(x)dx$$

And this is how we've talked about infinite series so far, even adopting the same notions of convergence and thinking about how we extend a familiar idea (in this case adding numbers, compared to integrating a function) out to infinity: we just keep walking our (finite) ending point out to infinity using a limit!

So this brings us to this comparison of the same types of objects across these two different contexts.

Table 8.4.1 Comparisons of Calculus Objects in Continuous and Discrete Contexts

Object	Continuous Context	Discrete Context
Function	$f(x)$	a_k
Graph	Figure 8.1.8	Figure 8.1.9
Finite Accumulation	Definite Integral $A(t) = \int_{x=0}^{x=t} f(x)dx$	Partial Sum $S_n = \sum_{k=0}^{n} a_k$
Infinite Accumulation	Improper Integral $\int_{x=0}^{\infty} f(x)dx$	Infinite Series $\sum_{k=0}^{\infty} a_k$

So in this section, we'll investigate this link between infinite series and improper integrals as the same kind of object occurring in different contexts. Intuitively, then, they'll be related to each other, under the right conditions.

The Integral Test

The Integral Test is really about connecting the behavior of an integral and a series, and the way that we'll do it is by trying to visualize what an infinite series is (a sum of function values, where the function inputs are spaced apart by 1) and linking that to a Riemann sum. From there, we'll use the Monotone Convergence Theorem on the sequence of partial sums to show that the series converges.

Activity 8.4.2: Integrals and Infinite Series

We're going to work with a graph of a continuous function, and we're going to start with a couple of conditions:

1. Our function will be continuous wherever it's defined.
2. Our function will be decreasing on its domain.
3. All of the function outputs will be positive.

Let's not worry about picking a specific function for this, but we will visualize a graph of one that meets these three requirements.

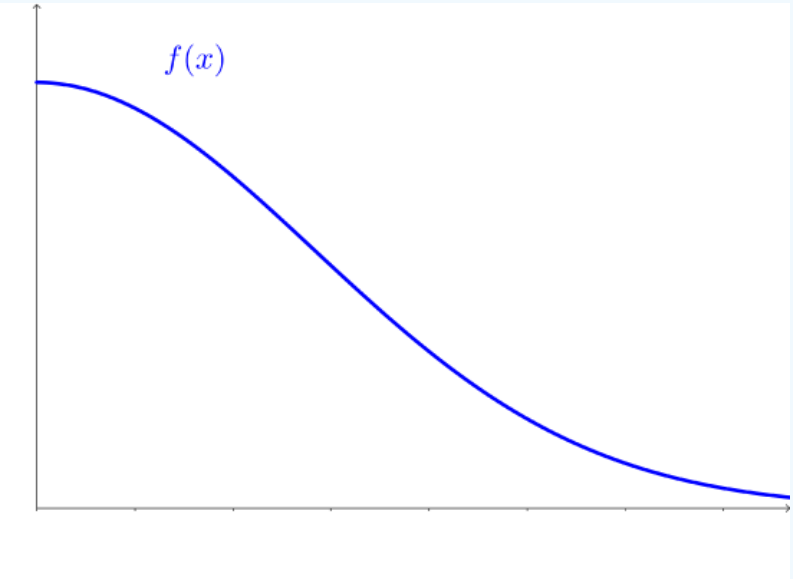


Figure 8.4.3

We can then visualize the sequence of terms, $a_k = f(k)$ for k = 0, 1, 2, ...,.

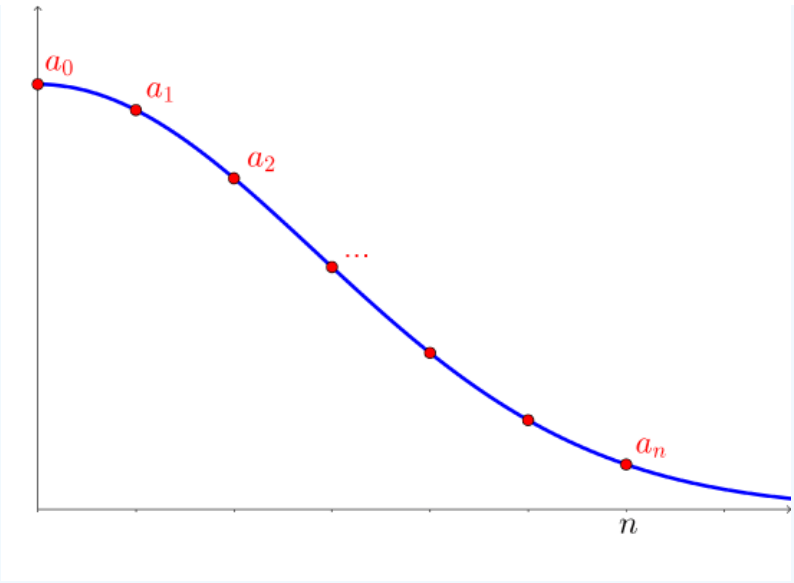


Figure 8.4.4

a. How does the partial sum, $\sum_{k=0}^{n} a_k$ compare to the Riemann sum for $f(x)$ from $x = 0$ to $x = n$ with $n + 1$ rectangles? **Hint**. It might help to visualize the Left Riemann sum, specifically!
b. We're going to visualize the accumulation of $f(x)$ from $x = 0$ to $x = n$ by thinking about the integral:

$$\int_{x=0}^{x=n} f(x)dx.$$

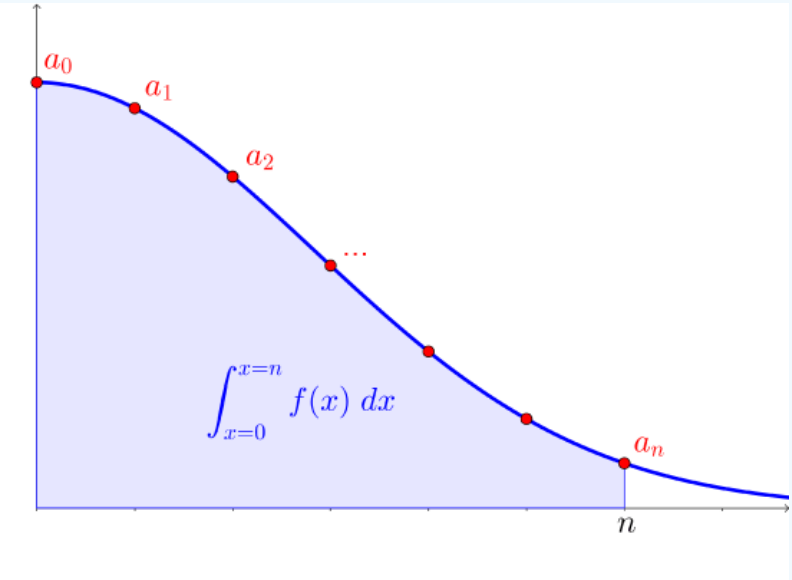


Figure 8.4.5

How does this area compare to the Riemann sum you thought of above? Compare them with an inequality and make sure you can explain why this has to be true. **Hint**. Here's a picture of the left Riemann sum!

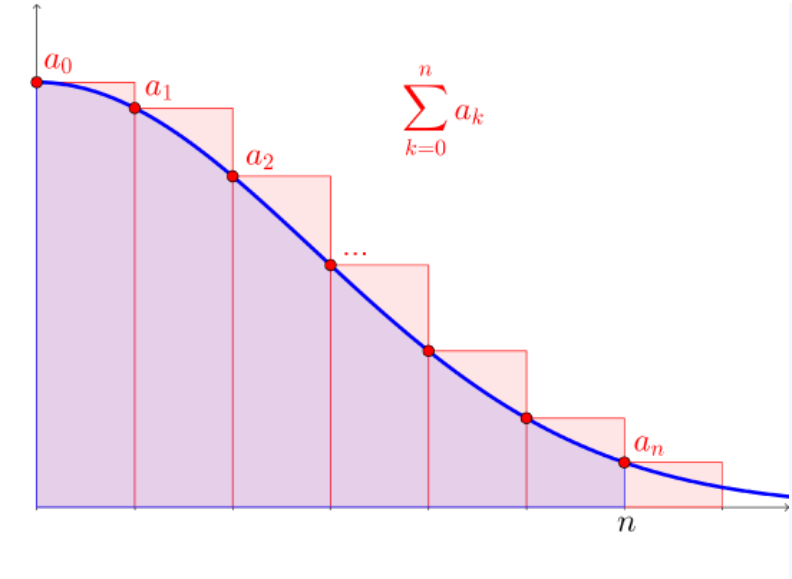


Figure 8.4.6

c. Remove the first term of the series, a_0, and instead think of the sum $\sum_{k=1}^{n} a_k$. Can you still think of this as a Riemann sum to approximate the area from the integral $\int_{x=0}^{x=n} f(x)dx$? How does this new Riemann sum compare to the area formed by the integral? Compare them with an inequality and make sure you can explain why this has to be true. **Hint**. Notice, now, that we need a_1 to be the representative height for the first rectangle. How does that change the formation of the rectangles?

d. We have thought about two sums, and we can connect them:

$$\sum_{k=0}^{n} a_k = a_0 + \sum_{k=1}^{n} a_k.$$

Use the sums to bound the integral:

$$\square \quad < \int_{x=0}^{x=n} f(x)dx < \quad \square$$

e. Similarly, use the integral to bound the sum:

$$\square \quad < \sum_{k=0}^{n} a_k < \quad \square$$

These bounds are going to be super useful! Discovering them is the main task for finding the connections between improper integrals and infinite series. These inequalities might seem kind of strange at first, but we're going to apply a limit to everything as $n \to \infty$, and then think about our definitions of convergence (Definition 7.1.7 and Definition 8.2.3).

Answer available at the end of this section.

Theorem 8.4.6: Integral Test.

If $\sum_{k=0}^{\infty} a_k$ is an infinite series with $a_k > 0$ for all $k \geq 0$ and $f(x)$ is a continuous and decreasing function with $f(k) = a_k$ for all $k \geq 0$, then we can compare the behaviors of $\sum_{k=0}^{\infty} a_k$ and $\int_{x=0}^{\infty} f(x)dx$: the integral and the series are guaranteed to either both diverge or both converge.

Proof. available at the end of this section.

So now we have a more formal link between these two objects. We have some intuition from Table 8.4.1 that these are pretty similar objects: one of them is an accumulation of function values in a continuous context, while the other is an accumulation of function values in a discrete context.

We can think about this even more formally! We reminded ourselves at the beginning of this section that, when we think about the

limit of a Riemann sum (in the continuous context) that as $n \to \infty$, we get $\Delta x \to dx$, where dx is essentially the distance between inputs. This produces our integral.

But now, in the discrete context, we have something different. Without being too formal, we can think about the corresponding limit of Δk as we slice this up further. Because our functions are in the discrete context, our inputs have a minimum distance between each other: they're all 1 unit apart! So here, in the limit where we expect $\Delta x \to dx$, we get $\Delta k \to 1$. And similarly, in a typical definite integral, we are adding up an infinite number of function outputs between some starting input and stopping input. Now in the discrete context, we don't get that! We get our normal partial sum:

$$\sum_{k=0}^{n} a_k \underbrace{\Delta k}_{1} = \sum_{k=0}^{n} a_k$$

So the Integral Test is pretty obvious, really: these corresponding objects retain the same type of behavior when we translate them back and forth between the continuous context and the discrete context.

Great! Let's apply this, now.

Example 8.4.9

For each of the following infinite series, decide whether it is possible (and reasonable) to use the Integral Test. If it is, apply the test and interpret the conclusions.

a. $\sum_{k=0}^{\infty} \frac{1}{k^2+1}$ **Hint**. This would connect with the integral $\int_{x=0}^{\infty} \frac{1}{x^2+1} dx$. Does the function, $f(x) = \frac{1}{x^2+1}$, meet the requirements of the Integral Test? Does it look like something you could antidifferentiate?

b. $\sum_{k=0}^{\infty} \frac{1}{e^{k^2}}$ **Hint**. We can rewrite $\frac{1}{e^{k^2}}$ to e^{-k^2}, and so we're thinking about the integral $\int_{x=0}^{\infty} e^{-x^2} dx$. Does this function meet the conditions of the Integral Test? Can we antidifferentiate?

c. $\sum_{k=1}^{\infty} \frac{k}{e^{k^2+1}}$ **Hint**. The Integral Test would connect this series to $\int_{x=1}^{\infty} \frac{x}{e^{x^2+1}} dx$. Does the function $f(x) = \frac{x}{e^{x^2+1}}$ meet the requirements of the Integral Test? Could we antidifferentiate?

d. $\sum_{k=2}^{\infty} \frac{1}{k \ln(k)}$ **Hint**. We're considering the function $f(x) = \frac{1}{x \ln(x)}$ and the integral $\int_{x=2}^{\infty} \frac{1}{x \ln(x)} dx$. Does this work for the Integral Test?

Solution

a. This is a fine opportunity to apply the Integral Test. The Integral Test says that we can link the behavior of the integral and the series:

$$\sum_{k=0}^{\infty} \frac{1}{k^2+1} \sim \int_{x=0}^{\infty} \frac{1}{x^2+1} dx$$

Let's think about the integral!

$$\begin{aligned} \int_{x=0}^{\infty} \frac{1}{x^2+1} dx &= \lim_{t \to \infty} \int_{x=0}^{x=t} \frac{1}{x^2+1} dx \\ &= \lim_{t \to \infty} \left(\tan^{-1}(x) \right) \Big|_{x=0}^{x=t} \\ &= \lim_{t \to \infty} \tan^{-1}(t) - \tan^{-1}(0) \\ &= \frac{\pi}{2} - 0 \\ &= \frac{\pi}{2} \end{aligned}$$

This integral converges. *Conclusion*: Since the integral $\int_{x=0}^{\infty} \frac{1}{x^2+1} dx$ converges, we know that the infinite series $\sum_{k=0}^{\infty} \frac{1}{k^2+1}$ also converges.

b. Unfortunately, this integral is going to be very difficult for us! The function $f(x) = e^{-x^2}$ has an antiderivative on the interval $[0, \infty)$ (it's a continuous function, and so it is integrable according to the Fundamental Theorem of Calculus (Part 1)). This function, though, doesn't have what we call an **elementary antiderivative**: any antideriative of e^{-x^2} can't be written as a combination of our basic function types. This means that we're unable to integrate this using our typical techniques, and (for now), we don't know if the integral converges or not. *Conclusion*: We won't apply the Integral Test, and so we don't know whether the series $\sum_{k=0}^{\infty} \frac{1}{e^{k^2}}$ converges or not.

c. Let's apply the Integral Test. we'll connect the behavior of the integral and the series:

$$\sum_{k=1}^{\infty} \frac{k}{e^{k^2+1}} \sim \int_{x=1}^{\infty} \frac{x}{e^{x^2+1}} dx$$

We'll consider the integral, and use a u-substitution where $u = -(x^2+1)$ and $du = -2x dx$.

$$\begin{aligned}
\int_{x=1}^{\infty} \frac{x}{e^{x^2+1}} dx &= \lim_{t\to\infty} \int_{x=1}^{x=t} \frac{x}{e^{x^2+1}} dx \\
&= \lim_{t\to\infty} -\frac{1}{2} \int_{x=1}^{x=t} \frac{-2x}{e^{x^2+1}} dx \\
&= -\frac{1}{2} \lim_{t\to\infty} \int_{u=-2}^{u=-(t^2+1)} e^u du \\
&= -\frac{1}{2} \lim_{t\to\infty} \left(e^u\right)\Big|_{u=-2}^{u=-(t^2+1)} \\
&= -\frac{1}{2} \lim_{t\to\infty} e^{-t^2+1} - e^{-2} \\
&= 0 + \frac{1}{2e^2}
\end{aligned}$$

This integral converges. *Conclusion*: The integral $\int_{x=1}^{\infty} \frac{x}{e^{x^2+1}} dx$ converges, and so we know that the infinite series $\sum_{k=1} \frac{k}{e^{k^2+1}}$ also converges.

d. If we apply the Integral Test, we're connecting the following series and integral:

$$\sum_{k=2}^{\infty} \frac{1}{k\ln(k)} \sim \int_{x=2}^{\infty} \frac{1}{x\ln(x)} dx$$

We'll consider the integral and use the substitution $u = \ln(x)$ so that $du = \frac{1}{x} dx$.

$$\begin{aligned}
\int_{x=2}^{\infty} \frac{1}{x\ln(x)} dx &= \lim_{t\to\infty} \int_{x=2}^{x=t} \frac{1}{x\ln(x)} dx \\
&= \lim_{t\to\infty} \int_{u=\ln(2)}^{u=\ln(t)} \frac{1}{u} du \\
&= \lim_{t\to\infty} \left(\ln|u|\right)\Big|_{u=\ln(2)}^{u=\ln(t)} \\
&= \lim_{t\to\infty} \ln|\ln(t)| - \ln(\ln(2)) \\
&= \infty
\end{aligned}$$

This integral diverges. *Conclusion*: We found that the integral $\int_{x=2}^{\infty} \frac{1}{x\ln(x)} dx$ diverges, which means that the series $\sum_{k=2}^{\infty} \frac{1}{k\ln(k)}$ also diverges.

As we develop more strategies and tests for series convergence, we should pause and summarize our test.

Integral Test Strategy.

We want to use this for functions that are relatively easy to antidifferentiate. Looking for u-substitution is a good idea, and sometimes we can straightforwardly apply integration by parts.

We'll find, though, that this test will mostly be used to introduce a family of infinite series and build up our intuition about partial sums (since we're really using the integrals to find a bound on our monotonic sequence of partial sums).

Why Do We Need These Conditions?

Before we wrap things up, let's just add some commentary on the conditions or requirements of the Integral Test. This is really a part of a broader discussion on conditions/requirements in mathematical results in general, but we'll limit ourselves to just this specific test. There are three conditions that we can consider: positive, decreasing, and continuous.

1. *Positive*: We need $f(x)$ and a_k to be positive in order for us to get monotonic sequences. Since we know that $a_k > 0$, we guarantee ourselves that $S_{n+1} > S_n$, since we're really adding another positive term. The same thing is true for the integrals, where we guarantee that
$$\int_{x=0}^{x=n+1} f(x)dx > \int_{x=0}^{x=n} f(x)dx$$
2. *Decreasing*: This one serves two purposes. First, we use the direction of the function to get some ideas on how the Left and Right Riemann sum compare to the areas: with a decreasing function, the Left sum will always overestimate the integral while the Right sum will underestimate it. We could have just required our function to be **monotonic**, though, since that would guarantee that one of those Riemann sums overestimated the integral while the other underestimated it: we really don't care about the order. So why would we need it, specifically, to be decreasing? Easy: if $\{a_k\}$ is positive and increasing, then $\lim_{k\to\infty} a_k \neq 0$, and we can just apply the Divergence Test.
3. *Continuous*: This one is pretty simple. Continuity guarantees that an antiderivative of our function exists on the interval we're looking at. We might not be able to actually find it easily, but at least we know there is one! Without continuity, we can't antidifferentiate easily and it doesn't make sense to think about the integral.

Note 8.4.8

We could pretty easily make a version of the Integral Test that would work for functions that are negative and increasing.

We talked about why it's nice to have a sequence of terms that correspond with a monotonic function, but what happens when they don't? Briefly, we can say that there are plenty of examples where the series and the integral may not behave similarly. For an easy to see example, let's consider the following series:

$$\sum_{k=0}^{\infty} \sin^2(\pi \cdot k)$$

If we try to think about the corresponding integral, we're considering:

$$\int_{x=0}^{\infty} \sin^2(\pi \cdot x)dx$$

Can you see what the problem is? The issue becomes evident when we plot the sequence of terms and the continuous function on the same axes.

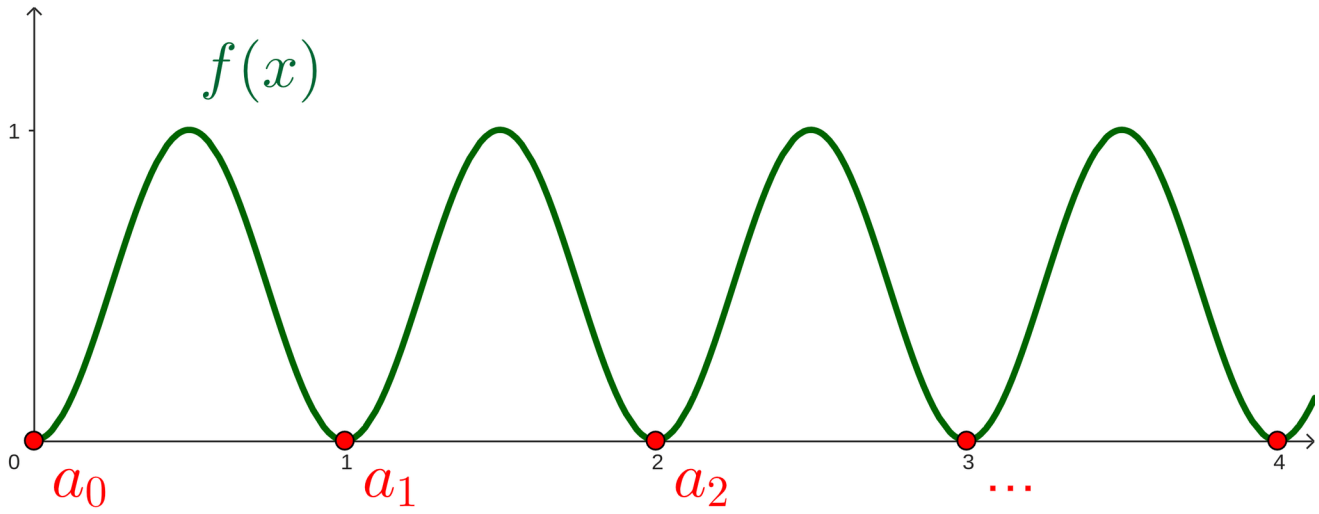


Figure 8.4.9

We can see that if we define the sequence of terms using $a_k = \sin^2(\pi \cdot k)$, then $a_k = 0$ for $k = 0, 1, 2, \ldots$. This means that

$$\sum_{k=0}^{\infty} \sin^2(\pi \cdot k) = 0 + 0 + 0 + \ldots = 0$$

Meanwhile, we can see that the integral $\int_{x=0}^{\infty} \sin^2(\pi \cdot x) dx$ will diverge to ∞, since every oscillation adds the same positive area over and over. The limit will not exist!

In the Integral Test, and in general, we want to (and need to) be careful about the conditions we apply: we want them to be general enough that we can actually use the test, but specific enough to protect against strange counterexamples.

Overall, we're really just trying to connect an integral in one context to an integral in another context. We have talked, briefly, about how we can think of an infinite series as a kind of improper integral for some discrete function (since we are summing up the function values multiplied by the minimal distance between inputs).

Earlier (in Section 5.6 Introduction to u-Substitution and Section 7.6 Trigonometric Substitution), we tried to visualize the variable substitutions (with Figure 5.6.3, Figure 5.6.5, and Figure 7.6.2) as a translation of functions between one context and another. We can visualize the Integral Test in a similar way: we're translating our integral from a Discrete context into a Continuous context, figuring out whether the integral converges in this context, and then taking the conclusions and applying them to the integral in the Discrete context.

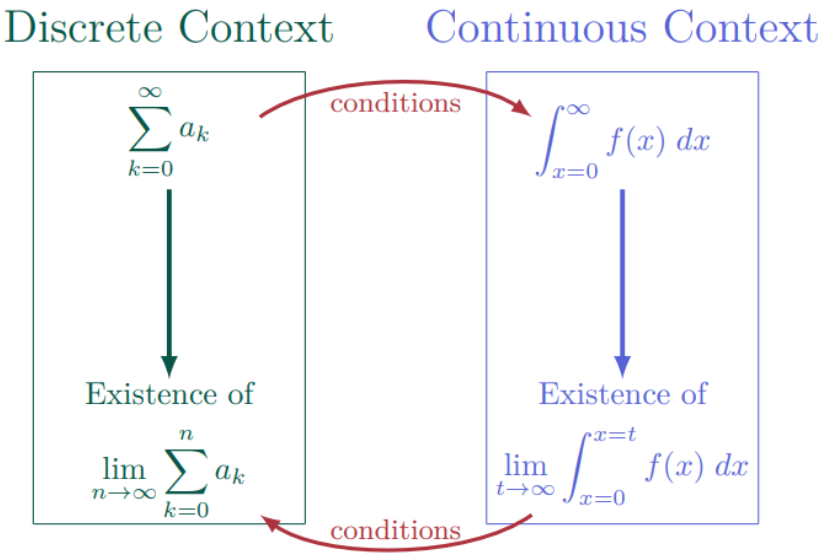


Figure 8.4.10 Integral Test, visualized as a transformation between contexts.

In both the discrete context and the continuous context, we care about whether the limit of the partial sum (or the limit of the definite integral) exists. Under the conditions of the Integral Test, though, we can translate our infinite series to the continuous context and represent it as a traditional improper integral. If we can get conclusions about whether the improper integral converges or not (based on the behavior of the limit of the definite integral), then we can bring that conclusion back to the discrete context and apply it to the limit of the partial sum.

This process works both ways! We could use information about an infinite series to learn about whether the corresponding improper integral converges or not! But for now, we are likely much more confident working with the integral in the continuous context, and so it makes most sense to think about the Integral Test using the directions pictured in Figure 8.4.10.

Practice Problems

Exercise 8.4.1

What are the conditions of the Integral Test? What are the conclusions of the Integral Test?

Exercise 8.4.2

For each of the following series, apply the Integral Test and interpret the conclusion of the test.

a. $\sum_{k=1}^{\infty} \left(\frac{k}{k^2+1} \right)$
b. $\sum_{k=1}^{\infty} \left(\frac{k}{e^k} \right)$
c. $\sum_{k=2}^{\infty} \left(\frac{1}{k\ln(k)} \right)$
d. $\sum_{k=1}^{\infty} \left(\frac{k^2+k}{e^k} \right)$
e. $\sum_{k=1}^{\infty} \left(\frac{1}{k^2+1} \right)$

Exercise 8.4.3

For what values of p (with $p > 0$) will the following series converge?

$$\sum_{k=1}^{\infty} \frac{1}{k^p}$$

Exercise 8.4.4

For what values of r (with $r > 0$) will the following series converge?

$$\sum_{k=0}^{\infty} r^k$$

Answers

Activity 8.4.2: Integrals and Infinite Series

a. Hopefully we can see that the partial sum, $\sum_{k=0}^{n} a_k$ is the exact same thing as the Left Riemann sum!

b. Since $\sum_{k=0}^{n} a_k$ is a left Riemann sum for $f(x)$, and since $f(x)$ is decreasing, then we know that each rectangle is formed from the highest point on each subinterval. That means that each rectangle's area overestimates the area under the curve on that subinterval. Note, also, that since this is a left Riemann sum, the nth rectangle is hanging past the end of the definite integral. This means that:

$$\sum_{k=0}^{n} a_k > \int_{x=0}^{x=n} f(x)dx$$

c. We can form a Right Riemann sum! Note that we don't "overhang" the interval anymore.

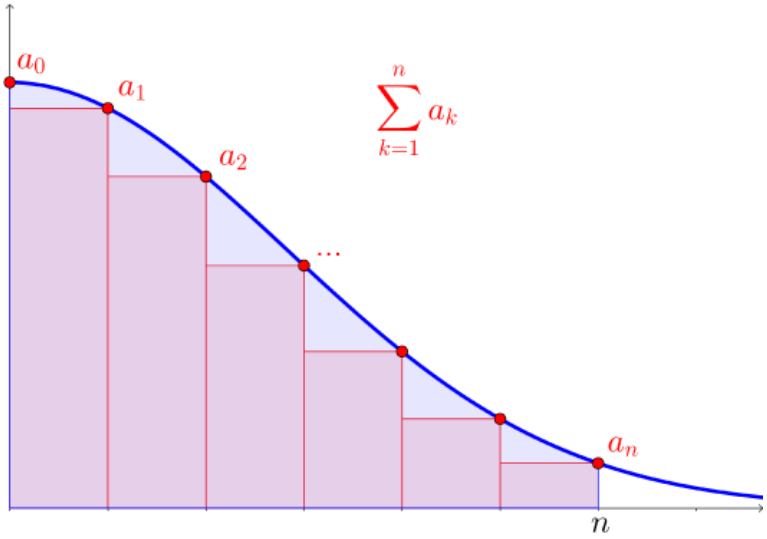


Figure 8.4.7

Note, now, that we are using the lowest point on each subinterval to form the rectangle. This means that:

$$\sum_{k=1}^{n} a_k < \int_{x=0}^{x=n} f(x)dx$$

d.
$$\sum_{k=1}^{n} a_k < \int_{x=0}^{x=n} f(x)dx < \sum_{k=0}^{n} a_k$$

$$\left(\sum_{k=0}^{n} a_k\right) - a_0 < \int_{x=0}^{x=n} f(x)dx$$

e.
$$\int_{x=0}^{x=n} f(x)dx < \sum_{k=0}^{n} a_k < a_0 + \int_{x=0}^{x=n} f(x)dx$$

Theorem 8.4.6: Integral Test.

The proof of this will come in two parts. First, we'll prove that $\sum_{k=0}^{\infty} a_k$ converges when $\int_{x=0}^{\infty} f(x)dx$ converges. Then, we'll prove that $\sum_{k=0}^{\infty} a_k$ diverges when

$\int_{x=0}^{\infty} f(x)dx$ diverges.

1. Let's start with the assumption that $\int_{x=0}^{\infty} f(x)dx$ converges. We know, based on Definition 7.1.7, that this means that $\lim_{n\to\infty} \int_{x=0}^{x=n} f(x)dx$ exists. We also know, since $f(x) > 0$, that

$$\lim_{n\to\infty} \int_{x=0}^{x=n} f(x)dx > \int_{x=0}^{x=n} f(x)dx$$

This means:

$$\sum_{k=0}^{n} a_k < a_0 + \int_{x=0}^{x=n} f(x)dx$$
$$\sum_{k=0}^{n} a_k < a_0 + \lim_{n\to\infty} \int_{x=0}^{x=n} f(x)dx$$
$$\sum_{k=0}^{n} a_k < a_0 + \int_{x=0}^{\infty} f(x)dx$$

This means that the partial sum, $S_n = \sum_{k=0}^{n} a_k$ has an upper bound. We also know that, since $a_k > 0$ for all $k = 0, 1, 2, ...$, then $S_{n+1} > S_n$. This means that the sequence of partial sums, $\{S_n\}_{n=0}^{\infty}$ is both monotonic and bounded, and therefore must converge (by the Monotone Convergence Theorem). Thus, $\sum_{k=0}^{\infty} a_k$ converges.

2. Now, we can start with the assumption that the integral $\int_{x=0}^{\infty} f(x)dx$ diverges. Since we know that $f(x)$ is positive, then we know that

$$\lim_{n\to\infty} \int_{x=0}^{x=n} f(x)dx = \infty$$

We can reconsider the inequalities from Activity 8.4.2:

$$\int_{x=0}^{x=n} f(x)dx < \sum_{k=0}^{n} a_k$$
$$\lim_{n\to\infty} \int_{x=0}^{x=n} f(x)dx < \lim_{n\to\infty} \sum_{k=0}^{n} a_k$$
$$\infty < \sum_{k=0}^{\infty} a_k$$

So then $\sum_{k=0}^{\infty} a_k$ must also diverge.

This is everything we need to prove! Note that we could replicate this proof by swapping the role of the series and the integral to get the same conclusion.

This page titled 8.4: The Integral Test is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Peter Keep.

8.5: Alternating Series and Conditional Convergence

Before we move too far forward, let's circle back to a point made in Why Do We Need These Conditions?. In the Integral Test, we required the terms of our series (and the continuous function we connected it with) to be positive. This was really just a mechanism that allowed us to say, in our proof, that the sequence of partial sums was monotonic. When we accumulate more of a positive thing, the total gets bigger. This is half of what we needed for us to employ the Monotone Convergence Theorem. And because this is such a useful tool, we'll see more of this "positive term series"condition showing up in the tools we use to see if a series converges.

But that makes this a perfect time to stop and ask a hallowed mathematical question: What happens if that property isn't there? What happens when our series does not only have positive terms?

We definitely have fewer tools to use, since we don't get anything that relies on applying the Monotone Convergence Theorem to partial sums. So instead, we'll take a brief detour into something we call Alternating Series (a series whose terms alternate in sign).

Activity 8.5.1: Which is More Likely to Converge?

We're going to try to think about what might be different when we analyze an alternating series compared to a series with only positive (or non-negative) terms

Let's say that {ak}∞ k=0 is some sequence of positive real numbers. Now let's consider the two series:

$$\sum_{k=0}^{\infty} a_k \quad \text{vs} \quad \sum_{k=0}^{\infty} (-1)^k a_k$$

a. Let's first consider the sequences of terms: $\{a_k\}$ compared with $\{(-1)^k a_k\}$. Is either of these more or less likely to converge? Does this tell us anything about whether or not the corresponding series converges? **Hint 1**. Try thinking about how we might find $\lim_{k\to\infty} (-1)^k a_k$ especially using the The Squeeze Theorem. **Hint 2**. What does the Divergence Test say? Is either of these sequences more or less likely to converge to 0 (or not)?

b. Now let's think of the partial sums:

$$\left\{\sum_{k=0}^{n} a_k\right\}_{n=0}^{\infty} \quad \text{vs} \quad \left\{\sum_{k=0}^{n} (-1)^k a_k\right\}_{n=0}^{\infty}$$

Is either of these sequences more or less likely to converge? Does this tell us anything about whether or not the corresponding series converges? **Hint**. It might be helpful to plot these partial sums! Referencing a picture like Figure 8.2.6 and think about how this will change the plot of partial sums for terms that alternate in sign.

c. Now make a conjecture about which infinite series is more likely to converge:

$$\sum_{k=0}^{\infty} a_k \quad \text{vs} \quad \sum_{k=0}^{\infty} (-1)^k a_k$$

Remember that $a_k > 0$ for $k = 0, 1, 2, \ldots$, so the only differences are the changes in sign. **Hint**. Remember that series convergence is defined by the behavior of the sequence of partial sums!

Defining Alternating Series, and the Main Result

Ok so hopefully we have convinced ourselves that a series that has terms that alternate in sign might have an "easier"time converging than a series with only positive terms. Let's start with a definition (so that we can continue to refer to these types of series easily), and then move towards the main result.

Definition 8.5.2: Alternating Series.

An infinite series $\sum_{k=0}^{\infty} a_k$ is called an **Alternating Series** when $a_k = (-1)^k |a_k|$ or$a_k = (-1)^{k+1} |a_k|$ for all $k = 0, 1, 2, ...$ That is, the sign of the terms alternates:

$$\sum_{k=0}^{\infty} a_k = |a_0| - |a_1| + |a_2| - + ...$$

So this is the type of series we're thinking of: the terms perfectly alternate in sign.

Aside.

At the beginning of this section, we talked about removing conditions and seeing what happens. We are not fully generalizing to terms that are just different in sign, since we still have a specific format that the terms need to hold to: switching between positive, negative, positive negative, etc. Series that have a less consistent pattern between positive and negative signs are harder to think about, and so we will have to be satisfied with loosening the restrictions on signs of the terms without fully removing any conditions.

Now, we will bring into focus one of the main results about alternating series before investigating these types of series (and how they converge) further. To lead into this new result, let's remind ourselves of a few things:

1. Remember the Divergence Test. We, specifically, want to remember that, in general, we don't know anything about a series if $\lim_{k\to\infty} a_k = 0$.
2. We have typically been looking at infinite series where we impose a further restriction on the terms: we looked at infinite series with only positive terms. We can get a lot more information about these series!
3. Now we are looking at infinite series with a different kind of structure on their terms: the signs alternate. So it won't be a surprise when we get to add information about them!

The big idea that we'll use the extra information about these alternating series (based on how we defined them) to get more information from the limit of the terms. Normally we don't get any information when $\lim_{k\to\infty} a_k = 0$: the series of those terms could converge or diverge, and we can't tell! But with this new structure of the terms (the alternating signs), we'll actually be able to tell something from the limit of the terms being 0.

Let's look at it! We're going to think about visualizing the partial sums.

In Figure 8.2.6, we looked at the partial sums of an infinite series and saw how the terms made up the differences between those partial sums. We're going to think about this same picture, but think about it through the lense of:

1. an alternating series (with terms that alternate in sign) where...
2. the terms of the alternating series approach 0 in the limit: $\lim_{k\to\infty} a_k = 0$. Note that this means that the size of the terms must go to 0 in the limit, $|a_k| \to 0$.

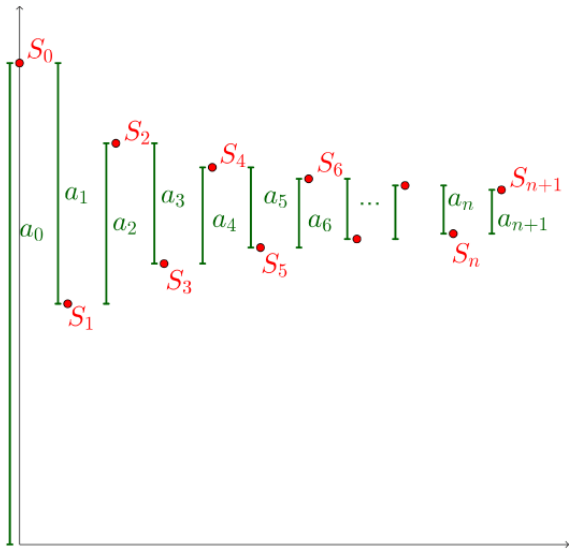


Figure 8.5.3 Partial sums of an alternating series.

Let's note a couple of things:

- All of the "even-indexed"terms (a_0, a_2, a_4, etc.) are positive, while all of the "odd-indexed"terms (a_1, a_3, a_5, etc.) are negative.
- This means that all of the "even-indexed"partial sums are big, while all of the "odd-indexed partial sums"are small. Our sequence of partial sums bounces up to an even-index and bounces down to an odd-index. As long as the size of the terms (the size of the differences between partial sums) is decreasing like we have pictured, then each "next"even-indexed partial sum is a bit smaller than the "previous"even-indexed partial sum. The same thing is true for the odd-indexed partial sums.
- The terms themselves represent the distance between these successive partial sums: the difference between S_n and S_{n+1} is the term a_{n+1}.

So, as long as the sizes of the terms are decreasing (or, as long as the distance between partial sums is decreasing consistently), then what happens when the terms (the distance between partial sums) goes to 0?

The even-indexed partial sums and the odd-indexed partial sums approach each other!

Theorem 8.5.4: Alternating Series Test.

If $\sum_{k=0}^{\infty} a_k$ is an alternating series and the size of the terms $|a_k|$ is decreasing, then if $\lim_{k\to\infty} a_k = 0$ then $\sum_{k=0}^{\infty} a_k$ converges.

Proof. available at the end of this section.

We can actually get another result really easily from this one! It's about how we might approximate the value that an infinite series converges to.

We know that if a series converges, it's because the limit of the partial sums exists. So we "just"need to find what the real number, S, is when

$$\lim_{n\to\infty} \sum_{k=0}^{n} a_k = S$$

We said earlier, though, that this is pretty hard to do! We can always approximate this limit, though, by looking at a big partial sum: if the limit exists, then probably adding up the first 10,000 terms will be a pretty good approximation (since we don't expect the partial sums to change, much, as n gets bigger). But how many terms is enough to give us a good enough approximation?

That's a hard question to answer, until we think about alternating series.

Activity 8.5.5: Approximating an Alternating Series.

Let's look, again, at the picture of the partial sums of an alternating series in Figure 8.5.3. We're going to assume that the series converges, which means that:

- $\lim_{n\to\infty} S_n$ exists.
- $\lim_{n\to\infty} a_n = 0$.

Let's add to our figure.

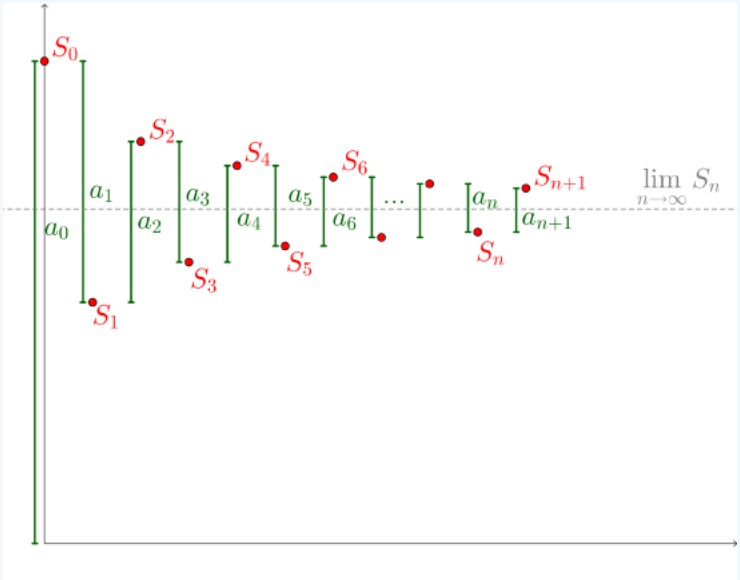


Figure 8.5.6

a. Why are the even-indexed partial sums sitting above the odd-indexed partial sums? **Hint**. We started, in this case, with a positive term a_0, and then the rest of the signs alternate. How does this change the partial values of partial sums?
b. Why are the even-indexed partial sums sitting above the horizontal line, $\lim_{n\to\infty} S_n$? **Hint**. We know that the sizes of the terms, $|a_n|$, decreases. This means that the sequence $\{S_0, S_2, S_4, ...\}$ is decreasing. What does that mean about how these partial sums relate to the limit of partial sums?
c. Why are the odd-indexed partial sums sitting below the horizontal line, $\lim_{n\to\infty} S_n$? **Hint**. We know that the sizes of the terms, $|a_n|$, decreases. This means that the sequence $\{S_1, S_3, S_5, ...\}$ is increasing. What does that mean about how these partial sums relate to the limit of partial sums?
d. If we were trying to approximate the value of $\lim_{n\to\infty} S_n$, how can we use the partial sums to build an interval that approximates the value?

Answer available at the end of this section.

Theorem 8.5.7: Approximations of Alternating Series.

If $\sum_{k=0}^{\infty} a_k$ is a converging alternating series, then the value S that the series converges to is bound between consecutive partial sums. Another way of saying this is that the partial sum S_n approximates the actual value of $\sum_{k=0}^{\infty} a_k$ with a maximum error of $|a_{n+1}|$.

Convergence, More Carefully

Let's circle back to an important point from Activity 8.5.1: an alternating series is more likely to converge than its positive-term counterpart.

Let's look at a classic example of this: the alternating harmonic series.

Activity 8.5.8: The Alternating Harmonic Series Converges.

Let's consider the alternating harmonic series, as written below:

$$\sum_{k=1}^{\infty} \frac{(-1)^{k+1}}{k} = 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \frac{1}{5} - \frac{1}{6} + - \ldots$$

a. First, confirm that this series does converge, using the Alternating Series Test.
b. Find $S_6 = \sum_{k=1}^{k=6} \frac{(-1)^{k+1}}{k}$.
c. Use Theorem 8.5.7 Approximations of Alternating Series to create a bound on this estimation. Report an interval that you know that the actual value that the alternating harmonic series converges to is in.
d. Use technology to find S_{1000}. Compare this value to $\ln(2)$. Would it surprise you for someone to claim that the alternating harmonic series converges to $\ln(2)$?

So we have some evidence that says that the alternating harmonic series converges to $\ln(2)$. Let's look at another "version" of this "same" series. You'll notice that we're using scare-quotes on "version" and "same," and that's what we're going to investigate.

Activity 8.5.9: The Alternating Harmonic Series Converges (Again).

Now let's consider a new, rearrranged, version of the alternating harmonic series:

$$1 - \frac{1}{2} - \frac{1}{4} + \frac{1}{3} - \frac{1}{6} - \frac{1}{8} + \frac{1}{5} - \frac{1}{10} - \frac{1}{12} + \ldots$$

We can write this in summation notation as:

$$\sum_{k=1}^{\infty} \left(\frac{1}{2k-1} - \frac{1}{4k-2} - \frac{1}{4k} \right)$$

a. First, confirm that all of the terms from the alternating harmonic series will eventually show up in this series. Convince yourself that we are truly adding (and subtracting) all of the same values with just a different order.
b. Does this new series converge? Check, using the Alternating Series Test.
c. Add up the first few terms of this series to find the value of a partial sum. You can choose how many terms you add. Does it look like it will also converge to $\ln(2)$?
d. Use technology to add up many terms of this series. Can you convince yourself what this series converges to?

This is strange! We have two series that seem to be the same thing (one is just a rearranged version of the other) that both converge, but they converge to different things! One seemingly converges to $\ln(2)$ while the other converges to half of that: $\frac{\ln(2)}{2}$. Very strange!

This doesn't seem to follow the normal rules of addition: we lose the normal associative property of addition, where the order or the way that we group terms to add typically doesn't matter. Here it does!

It turns out (and we won't prove this) that this type of convergence happens only for alternating series, and further it only happens for alternating series whose positive-term counterpart diverges.

Definition 8.5.10: Conditional (and Absolute) Convergence.

If $\sum_{k=0}^{\infty} a_k$ is a converging alternating series, then we say that the alternating series $\sum_{k=0}^{\infty} a_k$ **converges conditionally** if $\sum_{k=0}^{\infty} |a_k|$ diverges. If, instead, the positive-term series $diverges. I \sum_{k=0}^{\infty} |a_k|$ converges as well, then we say that the alternating series $\sum_{k=0}^{\infty} a_k$ **converges absolutely**.

Aside.

We won't go into this more, but a very cool result is that if an alternating series converges conditionally, then we can find a rearrangement of the terms that will converge to literally any real number that you'd like. We can rearrange the terms in such a way that the series converges to 0 or to π or to anything else you'd like.

But wait, it gets even stranger: you can also rearrange the terms in such a way that the resulting series diverges to ∞ or to −∞. These series really test the limits of what we mean when we say that a series converges, and so I hope that it seems reasonable to classify them separately of the "normally behaved" series.

So now, for a series whose terms alternate in size, we can:

- Classify whether the series converges or not (Theorem 8.5.4 Alternating Series Test).
- Further classify any converging alternating series to see whether the value it converges to is invariant to rearrangement. (Definition 8.5.10 Conditional (and Absolute) Convergence).
- Approximate the value that the series converges to using the nth partial sum, and also include an error bound on that approximation using the $|a_{n+1}|$ (Theorem 8.5.7).

Practice Problems

Exercise 8.5.1

What is an "alternating series?" What are some reasons that we need to consider these as a separate class of infinite series?

Exercise 8.5.2

Explain the differences between the Divergence Test and the Alternating Series Test. What do each tell us about the infinite series? How do we do each of the tests?

Exercise 8.5.3

Give an example of an alternating series that converges conditionally. Explain your reasoning, and show that it converges conditionally.

Exercise 8.5.4

Give an example of an alternating series that converges absolutely. Explain your reasoning, and show that it converges absolutely.

Exercise 8.5.5

Explain how an alternating series could converge conditionally. What about alternating series makes this even possible? Can a positive-term series converge conditionally? Why or why not?

Exercise 8.5.6

For each of the following alternating series, determine whether the series converges or diverges.

a. $\sum_{k=1}^{\infty}\left(\frac{(-1)^{k+1}\tan^{-1}(k)}{k^4}\right)$
b. $\sum_{k=1}^{\infty}\left(\frac{(-1)^{k+1}}{\sqrt{k^2+3}}\right)$
c. $\sum_{k=1}^{\infty}\left(\frac{(-1)^{k}}{\sqrt[k]{k}}\right)$
d. $\sum_{k=2}^{\infty}\left(\frac{(-1)^{k+1}}{\sqrt[3]{k^2-k}}\right)$
e. $\sum_{k=0}^{\infty}\left(\frac{(-1)^{k}}{\sqrt{k^3+1}}\right)$
f. $\sum_{k=1}^{\infty}\left(\frac{(-2)^{k+1}}{k^k}\right)$
g. $\sum_{k=1}^{\infty}\left(\frac{(-1)^{k+1}k}{k^5-k^2+1}\right)$
h. $\sum_{k=1}^{\infty}\left(\frac{\cos(\pi(k-1))}{k^2\ln(k+1)}\right)$
i. $\sum_{k=0}^{\infty}\left(\frac{\cos(\pi k)}{e^k+\pi^k}\right)$

Exercise 8.5.7

In Section 8.4, Practice Problem 2, you used the Integral Test to see if some series converged or diverged. Now, consider the alternating series below: test to see if they diverge, converge conditionally, or converge absolutely.

a. $\sum_{k=1}^{\infty}\left(\frac{(-1)^{k+1}(k)}{k^2+1}\right)$
b. $\sum_{k=1}^{\infty}\left(\frac{(-1)^{k+1}(k)}{e^k}\right)$
c. $\sum_{k=2}^{\infty}\left(\frac{(-1)^{k}}{k\ln(k)}\right)$
d. $\sum_{k=1}^{\infty}\left(\frac{(-1)^{k+1}(k^2+k)}{e^k}\right)$
e. $\sum_{k=1}^{\infty}\left(\frac{(-1)^{k+1}}{k^2+1}\right)$

Exercise 8.5.8

In Section 8.4, Practice Problem 3, you found the values of p (with $p>0$) such that the series $\sum_{k=1}^{\infty}\frac{1}{k^p}$ converges. What are the values of p (with $p>0$) such that

$$\sum_{k=1}^{\infty}\frac{(-1)^{k+1}}{k^p}$$

converges conditionally?

Exercise 8.5.9

In Section 8.4, Practice Problem 4, you found the values of r (with $r>0$) such that the series $\sum_{k=0}^{\infty}\frac{1}{r^k}$converges. Are there any values of r (with $r>0$) such that

$$\sum_{k=0}^{\infty}\frac{(-1)^{k}}{r^k}$$

converges conditionally?

Answers

Theorem $8.5.4$**: Alternating Series Test.**

This proof will follow the discussion before the statement of the theorem. Mostly, we will just fill in some details and provide some further justification for why what we were noticing must be true.

Let's start with the conditions of the test:

- We are considering an Alternating Series, $\sum_{k=0}^{\infty} a_k$. For our purposes, we'll assume that we have something like

$$\sum_{k=0}^{\infty} a_k = |a_0| - |a_1| + |a_2| - |a_3| + - ...$$

  where the even-indexed terms are the positive ones. This could be flipped and it wouldn't make a difference.
- The size of the terms are decreasing. That is, $|a_k| <| a_{k+1}$ for all $k = 0, 1, 2, ...$
- The limit lim $\lim_{k\to\infty} a_k = 0$. Note that this also means that $\lim_{k\to\infty} |a_k| = 0$.

We're going to show that, under these conditions, the alternating series we're considering must converge. The way that we'll do this is, no surprise, by invoking Theorem 8.1.14 Monotone Convergence Theorem. We're going to do it by considering the partial sums in halves: the even-indexed ones and the odd-indexed ones.

First, consider the sequence $\{S_{2n}\}_{n=0}^{\infty} = \{S_0, S_2, S_4, ...\}$. The difference between successive terms in this sequence (successive even-indexed partial sums) is:

$$S_{2n+2} - S_{2n} = -|a_{2n+1}| + |a_{2n+2}|$$

Since the terms of the alternating series are decreasing in size, we know that $|a_{2n+2}| < |a_{2n+1}|$, which means that $S_{2n+2} - S_{2n} < 0$, and so $S_{2n} > S_{2n+2}$.

All of this which is to say, $\{S_{2n}\}_{n=0}^{\infty}$ is a decreasing sequence.

We can apply the same reasoning to the sequence $\{S_{2n+1}\}_{n=0}^{\infty} = \{S_1, S_3, S_5, ...\}$.

We know the differences between successive odd-indexed partial sums is:

$$S_{2n+3} - S_{2n+1} = -|a_{2n+3}| + | \ a_{2n+2}$$

This time, though, $|a_{2n+2}| > |a_{2n+3}|$ and so the difference is positive: $S_{2n+3} - S_{2n+1} > 0$ which means that $S_{2n+3} > S_{2n+1}$. The sequence $\{S_{2n+1}\}_{n=0}^{\infty}$ is an increasing sequence.

We're getting close! We have monotonic sequences. Now we just need to show bounds, and then we'll show that each of these sequences converges. Then, we'll show that they converge to the same thing.

Getting an upper bound on the odd-indexed partial sums and a lower bound on the even-indexed partial sums is pretty easy. Let's consider subsequent partial sums, S_{2n} and S_{2n+1}.

$$\begin{aligned} S_{2n+1} - S_{2n} &= -|a_{2n+1}| \\ S_{2n+1} - S_{2n} &< 0 \\ S_{2n+1} &< S_{2n} \end{aligned}$$

Ok so this is easy: we can just pick any odd-indexed partial sum to be the lower bound on the even-indexed partial sums, and vice versa. So S_0 is an upper bound on $\{S_{2n+1}\}_{n=0}^{\infty}$, since every other S_{2n} is less than S_0, and all of S_{2n+1} partial sums are less than S_{2n}:

$$\begin{aligned} S_0 &\geq S_{2n} && \text{for } n = 0, 1, 2, ... \\ S_0 &\geq S_{2n} > S_{2n+1} && \text{for } n = 0, 1, 2, ... \\ S_0 &> S_{2n+1} && \text{for } n = 0, 1, 2, ... \end{aligned}$$

Similarly, we can say the same thing about S1 being a lower bound for the evenindexed partial sums:

$$\begin{aligned} S_1 &\leq S_{2n+1} && \text{for } n = 0, 1, 2, ... \\ S_1 &\leq S_{2n+1} < S_{2n} && \text{for } n = 0, 1, 2, ... \\ S_1 &< S_{2n} && \text{for } n = 0, 1, 2, ... \end{aligned}$$

So we have shown that the sequences $\{S_{2n}\}_{n=0}^{\infty}$ and $\{S_{2n+1}\}_{n=0}^{\infty}$ are both monotonic and bounded and so both of these sequences must converge.

Now we can show that they converge to the same thing.

Since $\{S_{2n}\}_{n=0}^{\infty}$ converges, let's say that there is some number S_E where

$$\lim_{n\to\infty} S_{2n} = S_E$$

Similarly, there is a number S_O where

$$\lim_{n\to\infty} S_{2n+1} = S_O$$

Now we can use the fact that the limit of the terms is 0:

$$\begin{aligned} \lim_{k\to\infty} a_k &= 0 \\ \lim_{n\to\infty} a_{2n+1} &= 0 \\ \lim_{n\to\infty} (S_{2n+1} - S_{2n}) &= 0 \\ S_O - S_E &= 0 \end{aligned}$$

So the numbers that these two sequences of partial sums converge to are actually equal to each other. So, finally, we know that under the conditions we started with, the alternating series must converge.

Activity $8.5.5$**: Approximating an Alternating Series.**

d. Since all of the even-indexed partial sums are upper bounds of any approximation and the odd-indexed partial sums are lower bounds. This means that any pair of partial sums serves as an approximation:

$$\begin{aligned} S_1 < \lim_{n\to\infty} S_n < S_0 \\ S_{2n+1} < \lim_{n\to\infty} S_n < S_{2n} \end{aligned}$$

This page titled 8.5: Alternating Series and Conditional Convergence is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Peter Keep.

8.6: Common Series Types

In this section, we'll stop and recap some of the common series types that we should recognize moving forward. We'll look at the structure of these series (mainly the functions defining the terms of the series) as well as the convergence criteria for them.

Look back to Activity 8.2.1. We noticed that we were able to find an explicit formula for the nth partial sum, which allowed us to evaluate $\lim_{n\to\infty} S_n$. We noticed this again in Example 8.2.8.

But there are some differences between why we were able to find formulas for the nth partial sum in each of these examples. Let's first focus on the infinite series with terms defined by exponential functions.

Geometric Series

We're going to name these series and define them explicitly. The name, **geometric**, comes from the idea of a geometric mean: each term in the series is the geometric mean of the term before it and after it.

Aside.

The geometric mean is a way of finding a kind of average. Instead of adding up the values and dividing by the number of values, we multiply the values and then take an nth root, where n is the number of things averaged.

For instance, a geometric mean of the numbers 1, 4, and 5 is:

$$\sqrt[3]{1 \cdot 4 \cdot 5} = \sqrt[3]{20}$$

This is approximately 2.714. We can compare that to the arithmetic mean:

$$\frac{1+4+5}{3} = \frac{10}{3}$$

These are just different kinds of measures of "center" of a list of values, although you might be most familiar with the arithmetic mean.

Definition 8.6.1: Geometric Series.

For real numbers a and r with $a, r \neq 0$, we say that the series

$$\sum_{k=0}^{\infty} ar^k = a + ar + ar^2 + ar^3 + \ldots$$

is a **geometric series**. We call r the constant ratio and a the initial term.

We noticed in Section 8.2 Introduction to Infinite Series that these kinds of series, with their exponential structure in the terms, make it relatively easy to find patterns or explicit formulas for the partial sums. Since we can find a formula for S_n based on n, we can find $\lim_{n\to\infty} S_n$. This was something we noted then, and said that it was a pretty rare property.

Let's generalize this a bit, and come up with a general formula for these partial sums. Once we have that, we will be able to find out what any geometric series converges to (if it does converge).

Activity 8.6.2: Building a Convergence Formula for Geometric Series.

We're going to think of constructing two different ways of thinking about how much area of of a circle has been shaded. We can pretend we have a circle with area that is 1, where the radius is $r = \sqrt{\frac{1}{\pi}}$, giving

$$A = \pi \left(\sqrt{\frac{1}{\pi}} \right)^2 = 1$$

Then we can describe the areas we're looking at as almost a percentages of the total area.

a. We are going to split our circle into two parts, with r amount of the area left unshaded and so $1 - r$ area shaded. We'll shade in some angular sector.

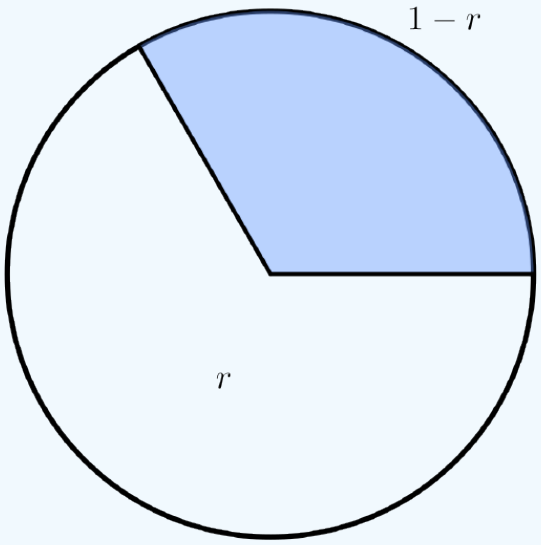


Figure 8.6.3

This part is easy: how much of the area is shaded?

b. This next step will set the stage for how we think about this problem now: we're going to divide the remaining white area up into the same proportional pieces: we'll shade in a ratio of $1 - r$ of the remaining white space and leave a ratio of r of the white space unshaded.

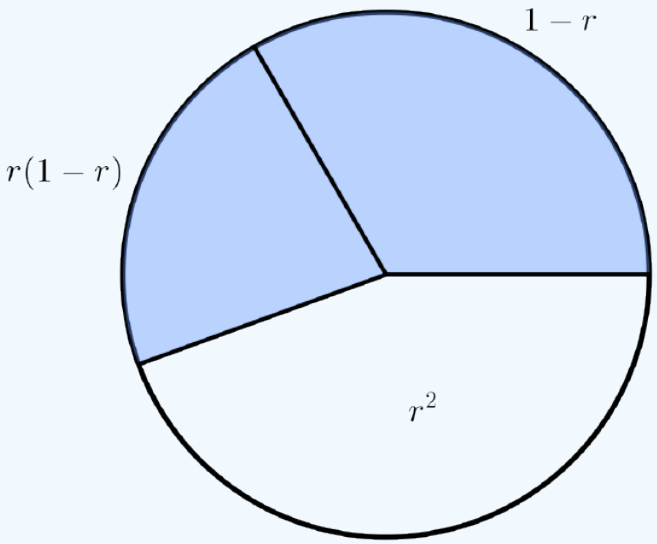


Figure 8.6.4

Can you describe two ways of calculating the total amount of shaded area? **Hint**. It could be helpful to consider how much area has been shaded or to not consider how much unshaded area is left over.

c. We'll repeat the process: shade in more, where the ratio of shaded area to unshaded area is $1-r$ to r.

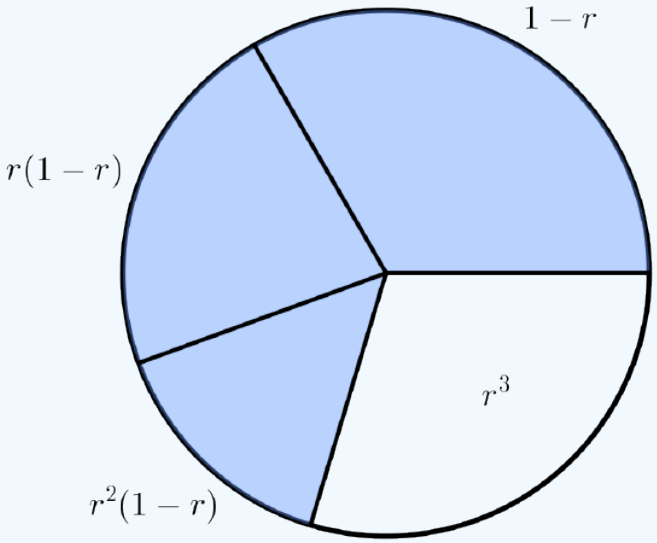


Figure 8.6.5

Can you describe two ways of calculating the total amount of shaded area? **Hint**. It could be helpful to consider how much area has been shaded or to not consider how much unshaded area is left over.

d. Now we're going to repeat this process until we've done it a total of $n+1$ times.

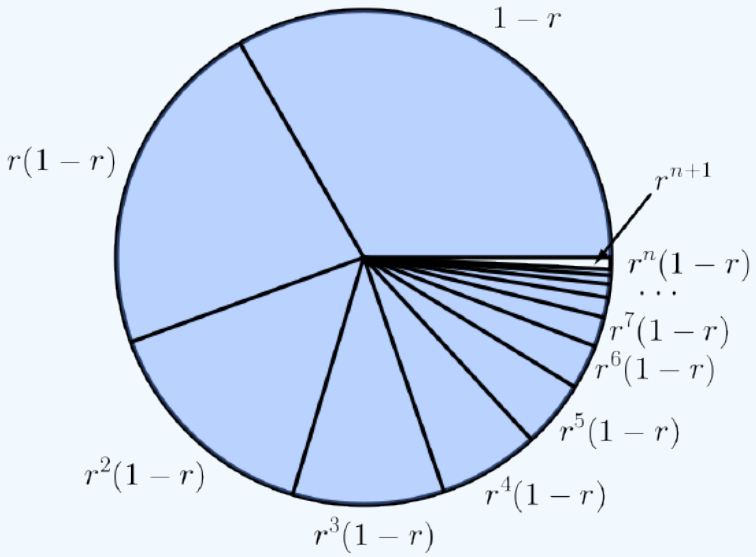


Figure 8.6.6

Can you describe two ways of calculating the total amount of shaded area? **Hint**. It could be helpful to consider how much area has been shaded or to not consider how much unshaded area is left over.

e. In the limit as $n \to \infty$, how much of the area is shaded in? Notice that $(1-r)$ is likely a common factor in one of your ways of calculating this area. Convince yourself, then, that:

$$\lim_{n\to\infty} \sum_{k=0}^{n} r^k = \lim_{n\to\infty} \left(1 + r + r^2 + ... + r^n\right)$$
$$= \frac{1}{1-r}$$

In order for us to move towards a formal statement of the geometric series convergence criteria, we can note that in the above activity, $0 \leq r < 1$. We can extend this to negative values, with $|r| < 1$. We also can note that we could scale the area, maybe call it a, to get a similar formula.

Theorem 8.6.7: Geometric Series Convergence Criteria.

A geometric series $\sum_{k=0}^{\infty} ar^k$ converges to $\frac{a}{1-r}$ when $|r| < 1$ and diverges if $|r| \geq 1$.

p-Series

Another type of structure that we can take advantage of is power functions. This way, we can leverage the Integral Test (since anti-differentiating using the Power Rule for Antiderivatives will be easy) to classify whether or not these series converge.

Let's start by naming these. We'll focus on power functions with negative exponents, or reciprocal power functions.

Definition 8.6.8: p-Series.

For a real number p, we say that the series

$$\sum_{k=1}^{\infty} \frac{1}{k^p}$$

is a p-series. We mostly will be concerned about the case where $p > 0$, making the terms of the series be reciprocal power functions.

Now, we just need to think about integration, and the convergence classification comes quickly from there.

Theorem 8.6.9: p-Series Convergence Criteria.

A p-series $\sum_{k=1}^{\infty} \frac{1}{k^p}$ converges when $p > 1$ and diverges when $p \leq 1$.

Proof. available at the end of this section.

Recapping Our Mathematical Objects

It's a good idea to pause and try to make sure we understand what these infinite series are. We have talked a lot about a whole bunch of objects in this chapter so far: infinite sequences, partial sums, sequences of partial sums, infinite series, integrals, limits, etc. We want to make sure that we can keep track of the ways in which we use these and talk about them. The following activity is brief, but can help make sure we understand some of the interactions we've talked about so far.

Activity 8.6.10: (Im)Possible Combinations.

When we have thought about infinite series, we have thought about three different mathematical objects: the sequence of terms of the series, the sequence of partial sums of the series, and the infinite series itself. As a reminder, if we had an infinite series

$$\sum_{k=1}^{\infty} a_k$$

we can say that:

- $\{a_k\}_{k=1}^{\infty}$ is the sequence of terms of the series
- $S_n = \sum_{k=1}^{n} a_k$ is a partial sum and $\{S_n\}_{n=1}^{\infty}$ is the sequence of partial sums of the series

For each of these three objects—the terms, the partial sums, and the series—we have some notion of what it means for that object to converge or diverge.

Consider the following table of all of the different combinations of convergence and divergence of the three objects. For each combination, decide whether this combination is possible or impossible. If it is possible, give an example of an infinite series whose terms, partial sums, and the series itself converge/diverge appropriately. If it is impossible, give an explanation of why.

Table 8.6.11 (Im)Possible Combinations

$\{a_k\}_{k=1}^{\infty}$	$\{S_n\}_{n=1}^{\infty}$	$\sum_{k=1}^{\infty} a_k$	(Im)Possible?	Example or Explanation
Converges	Converges	Converges		
Converges	Converges	Diverges		
Converges	Diverges	Converges		
Converges	Diverges	Diverges		
Diverges	Converges	Converges		
Diverges	Converges	Diverges		
Diverges	Diverges	Converges		

$\{a_k\}_{k=1}^{\infty}$	$\{S_n\}_{n=1}^{\infty}$	$\sum_{k=1}^{\infty} a_k$	(Im)Possible?	Example or Explanation
Diverges	Diverges	Diverges		

Hint 1. We can think back to some results or definitions that connect pairs of these objects. Can you think of any result or definition that connects an infinite series and a sequence of partial sums? What about a result or definition that connects the sequence of terms with the infinite series?

Hint 2. Look back at Definition 8.2.3 and Theorem 8.3.5.

Answer available at the end of this section.

Moving forward, we'll want to commit these families of series to memory, as well as their convergence criteria.

Here's some justification: for an infinite series like $\sum_{k=0}^{\infty} \frac{1}{k^2+1}$, we previously (in Example 8.4.7) compared this series to the integral $\int_{x=0}^{\infty} \frac{1}{x^2+1} dx$. This worked well, since we could pretty easily anti-differentiate and conclude that the integral converged.

Now, though, we have another connection to make: doesn't this series almost look like a p-Series? It's very close to a reciprocal power function, where the only thing that's "off" is the "+1" in the denominator.

We can hopefully think about changing this example slightly: what about the series $\sum_{k=0}^{\infty} \frac{1}{k^3+1}$? This one could still be compared to the integral $\int_{x=0}^{\infty} \frac{1}{x^3+1} dx$, but this integral will be harder for us to integrate. But if we think about this as almost a p-Series, then we might still be able to have some intuition about it's behavior: it looks kind of like a converging p-series, so shouldn't it also converge?

Our next section will develop this kind of comparison, where instead of comparing an "integral" in the discrete context to one in the continuous context (like we do in the Integral Test), we can compare an "integral" in the discrete context to a similar one in a similar context.

Practice Problems

Exercise 8.6.1

For each of the following series, classify each of the series by type (geometric, p-series, or neither) and whether it converges or diverges (as well as why).

For any geometric series that converges, find what it converges to.

a. $\sum_{k=1}^{\infty} \left(\frac{1}{3^k}\right)$
b. $\sum_{k=1}^{\infty} \left(\frac{1}{k^3}\right)$
c. $\sum_{k=1}^{\infty} \left(\frac{1}{4^k}\right)$
d. $\sum_{k=0}^{\infty} \left(e^{-2k}\right)$
e. $\sum_{k=1}^{\infty} \left(\frac{1}{\sqrt{k-2}}\right)$
f. $\sum_{k=1}^{\infty} \left(\frac{1}{\sqrt[5]{k^7}}\right)$
g. $\sum_{k=1}^{\infty} \left(\frac{1}{\sqrt[7]{k^5}}\right)$
h. $\sum_{k=0}^{\infty} \left(\frac{(-1)^k 3^{k+8}}{5^{k+2}}\right)$

Exercise 8.6.2

This is a fun problem! We'll consider a special cake with a weird twist. The cake is made of stacked, cylindrical tiers, each with a height of 1 unit. The radius of the tiers is $\frac{1}{k}$ for $k = 1, 2, ...$, meaning the bottom tier has a radius of 1, the second tier has a radius of 1/2, the third has a radius of 1/3, etc. The cake, somehow, has an infinite number of tiers.

a. We want to ice the sides of the cake, and want to know how much icing to use to cover the sides of each tier. Create an infinite series to represent the area of the outside edges. Does the series converge or diverge? Explain.
b. We want to figure out how much batter needed to bake the cake. Create an infinite series representing the volume of the tiers. Does the series converge or diverge? Explain.

Answers

Theorem 8.6.9: p-Series Convergence Criteria.

Let's divide this into four cases: when $p \leq 0$, when $0 < p < 1$, when $p = 1$, and when $p > 1$.

Case 1: $p \leq 0$

Note that for $\frac{1}{k^p}$ with $p < 0$, we can write this as $k^{|p|}$. Now we can consider the limit of the terms, in order to use the Divergence Test.

$$\lim_{k\to\infty} \frac{1}{k^p} = \lim_{k\to\infty} k^{|p|}$$

Since this limit is non-zero (since it is either ∞ or 1, depending on whether $p = 0$ or not), the series diverges by the Divergence Test.

Case 2: $0 < p < 1$

When $0 < p < 1$, we can apply the Integral Test to the series. It is worth showing that the conditions of the test are met, but this is left up to the reader.

So now we will consider the integral $\int_{x=1}^{\infty} \frac{1}{x^p} dx$ as a way of seeing whether the series $\sum_{k=1}^{\infty} \frac{1}{k^p}$ converges or diverges.

$$\begin{aligned}\int_{x=0}^{\infty} \frac{1}{x^p} dx &= \lim_{t\to\infty} \int_{x=1}^{x=t} \frac{1}{x^p} dx \\ &= \lim_{t\to\infty} \left(\frac{x^{1-p}}{(1-p)} \right) \Bigg|_{x=1}^{x=t} \\ &= \lim_{t\to\infty} \frac{t^{1-p}}{1-p} - \frac{1}{1-p}\end{aligned}$$

We can note that since $0 < p < 1$, that $1 - p > 0$. This means that when $t \to \infty$,, $t^{1-p} \to \infty$ as well.

$$\int_{x=0}^{\infty} \frac{1}{x^p} dx = \lim_{t\to\infty} \frac{t^{1-p}}{1-p} - \frac{1}{1-p} = \infty$$

This integral diverges, and so then does the series.

Case 3: $p = 1$

This is the Harmonic Series! This series diverges (Theorem 8.3.3).

Case 4: $p > 1$

We can repeat the proof from Case 2, but we will end with a different conclusion based on the sign of the exponent! Let us, again, apply the Integral Test.

Consider the integral $\int_{x=1}^{\infty} \frac{1}{x^p} dx$ as a way of seeing whether the series $\sum_{k=1}^{\infty} \frac{1}{k^p}$ converges or diverges.

$$\begin{aligned}\int_{x=0}^{\infty} \frac{1}{x^p} dx &= \lim_{t\to\infty} \int_{x=1}^{x=t} \frac{1}{x^p} dx \\ &= \lim_{t\to\infty} \left(\frac{x^{1-p}}{(1-p)} \right) \Bigg|_{x=1}^{x=t} \\ &= \lim_{t\to\infty} \frac{t^{1-p}}{1-p} - \frac{1}{1-p}\end{aligned}$$

Now, though, we have $p > 1$ which means that $1 - p < 0$. This means that $t^{1-p} = \frac{1}{t^{|p-1|}}$. So now we will consider the limit, and note that as $t \to \infty$, we get $\frac{1}{t^{|p-1|}} \to 0$.

$$\int_{x=0}^{\infty} \frac{1}{x^p} dx = \lim_{t\to\infty} \frac{1}{(1-p)t^{|p-1|}} - \frac{1}{1-p} = -\frac{1}{1-p}$$

This integral converges, and so then does the series. We remember, though, that the series converges to something different than the integral, and so we do not know what the series converges to.

Activity 8.6.10: (Im)Possible Combinations.

(Im)Possible?	Example or Explanation
Possible	Any converging series serves as an example.
Impossible	The sequence of partial sums and the infinite series are the same object, and so must behave in the same way. Definition 8.2.3

(Im)Possible?	Example or Explanation
Impossible	The sequence of partial sums and the infinite series are the same object, and so must behave in the same way. Definition 8.2.3
Possible	
Impossible	If the infinite series converges, then the sequence of terms must converge to 0. Theorem 8.3.5
Impossible	The sequence of partial sums and the infinite series are the same object, and so must behave in the same way. Definition 8.2.3
Impossible	Both of the reasons, Definition 8.2.3 and Theorem 8.3.5 apply here!
Possible	

This page titled 8.6: Common Series Types is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Peter Keep.

8.7: Comparison Tests

So far, our strategies for thinking about infinite series have been focused around drawing a connection between the infinite series we care about and some other mathematical object:

- The Divergence Test draws a connection (even thought it's a limited one) between the terms of the series and the series itself.
- The Alternating Series Test draws a (stronger) connection between the terms of, specifically, an Alternating Series and the series itself.
- The Integral Test draws a connection between the series and a corresponding integral.

Now we'll work on building the most important series convergence test mechanism: we'll draw a link between the series we care about and some other series that we already know about.

This is helpful for three reasons:

1. We already have a couple of types of series that we know about (Section 8.6), and we can keep adding to that list.
2. We can take advantage of similar structure or common term formulas when we see them to essentially say, "This series kind of looks like one that I recognize. I wonder if they act the same?"
3. We don't always have to integrate things using the Integral Test! Integrating can be hard!

Comparing Partial Sums

We're going to start by trying to do the same thing we did when we build the Integral Test: show that the partial sums are monotonic and bounded and then make use of the Monotone Convergence Theorem.

Activity 8.7.1: Comparisons to Bound Partial Sums.

This activity is mostly going to be thinking about proof mechanisms, and so it might be helpful to review Activity 8.4.2 Integrals and Infinite Series. If you want to see more, then the proof of Integral Test will provide some further details on why the inequalities we built were useful.

a. In the Integral Test, how did we guarantee that our sequence of partial sums was monotonic?
b. How did we know that, as long as the corresponding integral converged, then our sequence of partial sums was bounded?
c. How did we know that, as long as our integral diverged, then our sequence of partial sums had to diverge as well?
d. What happens if we swap out the integral we're connecting our series to with a different series? For these inequalities to be useful, what do we need to be true?

Hint 1. For our sequence of partial sums to be bounded, we originally had that our partial sums were smaller than an expression involving an integral. Then, those integrals eventually converged to something, and so our sequence of partial sums was bounded by a value. Can we swap out an integral for a different partial sum? What has to happen in the limit to guarantee that our series converges?

Hint 2. Similarly, when we showed that our partial sums were always greater than something involving an integral, then we were able to show that if that integral diverged, then our partial sums had no choice but to diverge to infinity as well. What kind of partial sum do you need to change out the integral for?

This is it! We have everything we need to construct another link: this time between the infinite series we care about and another infinite series that we think acts the same.

To illustrate what we've done, let's think about two sequences of terms: $\{a_k\}_{k=0}^{\infty}$ and $\{b_k\}_{k=0}^{i} nfty$ where $a_k \leq b_k$ for $k \geq 0$. We can think of the graphs, pictured below.

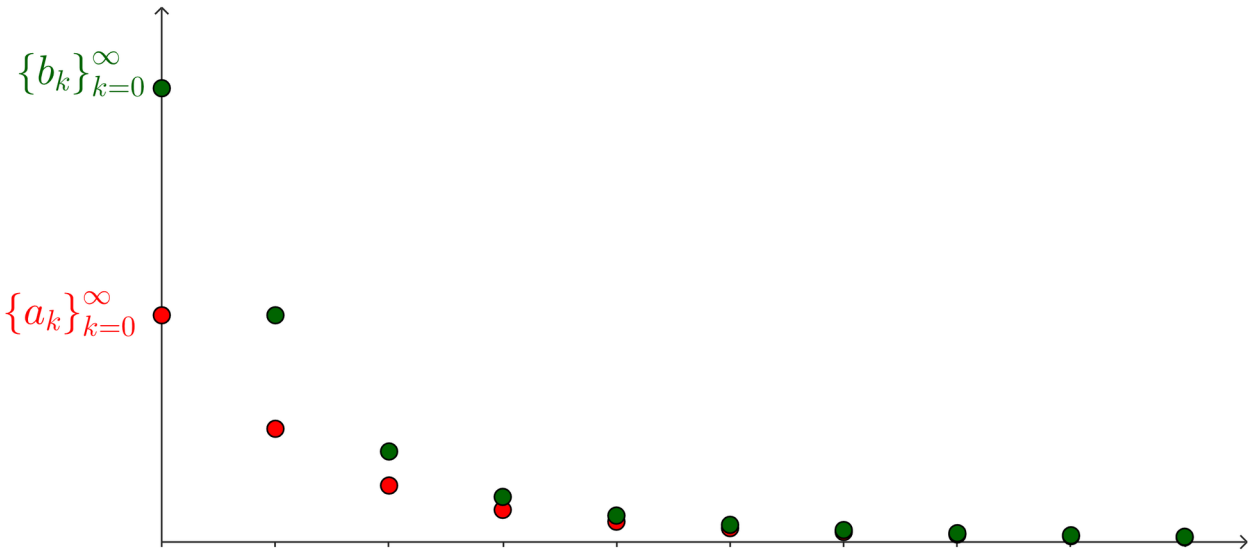


Figure 8.7.2 The “smaller” sequence of terms $\{a_k\}_{k=0}^{\infty}$ graphed alongside the “bigger” sequence of terms $\{b_k\}_{k=0}^{\infty}$.

Now we can think about the graphs of these partial sums. Let’s first think about what might happen if the series $\sum_{k=0}^{\infty} b_k$ converges. We’ll plot the partial sums, and the sequence of partial sums has to converge to something. But then we can think about the sequence of partial sums from the ak terms: the smaller terms, of course, will build smaller partial sums.

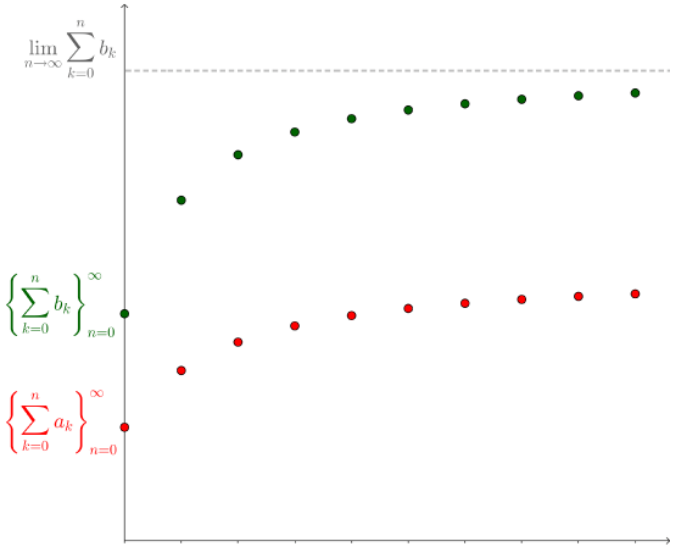


Figure 8.7.3 Comparison of partial sums when $\sum_{k=0}^{\infty} b_k$ converges.

So we can pretty easily use $\lim_{n\to\infty} \sum_{k=0}^{n} b_k$ as an upper bound on the sequence $\{\sum_{k=0}^{n} a_k\}_{n=0}^{\infty}$! And now we can say that the sequence of partial sums for $\sum_{k=0}^{\infty} a_k$ is monotonic and bounded, and so it must converge.

Then we can ask about the diverging case. This time, we’ll say that the smaller series, $\sum_{k=0}^{\infty} a_k$ diverges.

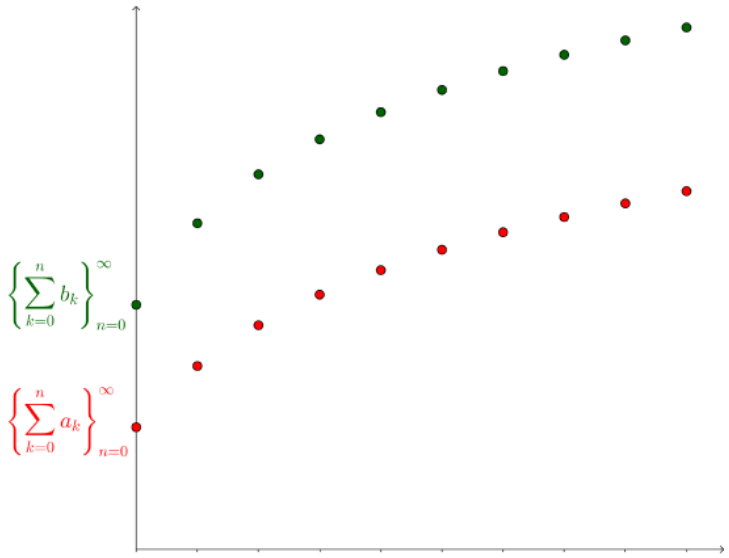


Figure 8.7.4 Comparison of partial sums when $\sum_{k=0}^{\infty} a_k$ diverges.

We can think that this "smaller" sequence of partial sums is "pushing" the "larger" sequence of partial sums up to infinity with it.

These two cases make up our first comparison mechanic between two infinite series.

Theorem 8.7.5: Direct Comparison Test.

If $\sum_{k=0}^{\infty} a_k$ and $\sum_{k=0}^{\infty} b_k$ are infinite series with positive terms ($a_k > 0$ and $b_k > 0$ for $k \geq 0$) with the ordering $a_k \leq b_k$ for $k \geq 0$, then:

- If $\sum_{k=0}^{\infty} a_k$ diverges, then $\sum_{k=0}^{\infty} b_k$ also diverges.
- If $\sum_{k=0}^{\infty} b_k$ converges, then $\sum_{k=0}^{\infty} a_k$ also converges.

At the end of Section 8.6 Common Series Types, we referenced the series:

$$\sum_{k=0}^{\infty} \frac{1}{k^3+1}$$

Let's use our new test!

Example 8.7.6

Does the series $\sum_{k=0}^{\infty} \frac{1}{k^3+1}$ converge or diverge? How do you know?

Hint 1. If we're going to use our new Direct Comparison Test, we need to identify two things:

1. Some intuition on whether we think our series converges or not.
2. An appropriate series to compare to. Likely, this will be either a Geometric Series or a p-Series, since we have clear convergence criteria for each of those.

We can do this in any order: sometimes we might use the structure of the series we're looking at to give us a good candidate to compare to, and that might tell us the behavior we think we're looking for. Other times we might have good intuition about convergence/divergence of the series which will tell us whether we need to find a series that is smaller or larger to compare to. What do you think? Do we have a suitable comparison?

Hint 2. Compare to the p-Series $\sum_{k=1}^{\infty} \frac{1}{k^3}$. Here, $p = 3$. Based on this, do we need to show that $\frac{1}{k^3+1}$ is greater than or less than $\frac{1}{k^3}$? Show this! Does the change in the starting index matter?

Solution

Let's compare our series to the converging p-series with $p = 3$:

$$\sum_{k=0}^{\infty} \frac{1}{k^3+1} \sim \sum_{k=1}^{\infty} \frac{1}{k^3}$$

We want to show that $\frac{1}{k^3+1} \leq \frac{1}{k^3}$ for $k \geq 1$.

Let's start with comparing the denominators, and move from there.

$$k^3 + 1 > k^3 \text{ for any value of } k$$

Now we can think about reciprocals:

$$\frac{1}{k^3+1} < \frac{1}{k^3} \text{ for all } k \geq 1$$

This means that, since $\sum_{k=1}^{\infty} \frac{1}{k^3}$ converges, then $\sum_{k=1}^{\infty} \frac{1}{k^3+1}$ must also converge.

We can just add the term where $k = 0$ to get

$$\begin{aligned} \sum_{k=0}^{\infty} \frac{1}{k^3+1} &= \overbrace{\frac{1}{0^3+1}}^{k=0} + \sum_{k=1}^{\infty} \frac{1}{k^3+1} \\ &= 1 + \sum_{k=1}^{\infty} \frac{1}{k^3+1} \end{aligned}$$

So then we know that our series, $\sum_{k=0}^{\infty} \frac{1}{k^3+1}$ must converge.

A quick note: in this example, we thought about building the link

$$\sum_{k=1}^{\infty} \frac{1}{k^3+1} \sim \sum_{k=1}^{\infty} \frac{1}{k^3}$$

and then arguing that changing the index to start at $k = 0$ by adding in a single term wouldn't change the behavior of the infinite series.

In general, we don't need to add in that last argument. We can assume from here on out that changing our starting index won't impact the behavior of the series (as long as we're avoiding things like division by 0 and such). So showing that $\frac{1}{k^3+1} < \frac{1}{k^3}$ for $k \geq 1$ is enough for us!

(Un)Helpful Comparisons

Activity 8.7.7: (Un)Helpful Comparisons.

We're going to consider a handful of infinite series, here:

1. $\sum_{k=1}^{\infty} \frac{2}{k\,(3^k)}$

2. $\sum_{k=3}^{\infty} \frac{\sqrt{k+1}}{k-2}$

a. Pick a series that is reasonable to use as a comparison for each of the series listed. Remember, we want:
   - A series that is recognizable (probably a Geometric Series or a p-Series), so that we know the behavior of it: we need to know whether the series we're comparing to converges or diverges!
   - A series that is similar enough to the series we're working with that we can construct an inequality comparing the term structure. It can be hard to compare functions that are seemingly unrelated to each other!
   - A series that has terms that are either larger or smaller than our series, depending on whether we are trying to show that our series converges or diverges.

   Hint 1. It can be useful to think about what the dominating pieces of each term structure are. What parts of the function will be most important, especially as the index, k, gets larger?

   Hint 2.

   1. In the denominator, think about k contrasted with 3^k. Which of these will be more influential in determining the value of the term when k gets large?
   2. In the numerator, think about k contrasted with +1 under the square root. Which of these will be more influential? In the denominator, contrast k and –2. Which of these will be more influential? What does this fraction look like when we just consider the numerator and denominator's most dominant pieces?

b. Build the comparison from the series we start with to the one you picked. What kinds of conclusions can you make?

   Hint.

   1. When we compare $\frac{2}{k\,(3^k)}$ with $\frac{2}{3^k}$, we can start with the denominator. What does multiplying by $\diamondsuit$ do to the value of 3? It might help to write some of the values out for $k = 1, 2, 3, \ldots$
   2. When we compare $\frac{\sqrt{k+1}}{k-2}$ with $\frac{\sqrt{k}}{k}$, we can think of the numerator and denominator separately. What does adding 1 under the square root do to the value of $\sqrt{k}$? What does subtracting 2 in the denominator do to k? How do these impact the fraction?

c. We're going to change the series we're considering to two slightly different series:
   i. $\sum_{k=1}^{\infty} \frac{2k}{3^k}$
   ii. $\sum_{k=3}^{\infty} \frac{\sqrt{k-1}}{k+2}$

   How do these small changes impact the inequalities you built?

   Hint.

   1. We moved k from the denominator to the numerator. If the numerator is the thing getting bigger, how does that impact the size of this fraction in the same comparison as before?
   2. We changed the signs: we're subtracting in the numerator and adding in the denominator. How does that impact the size of this fraction in the same comparison as before?

d. How do these changes in the inequalities change the conclusions we can draw from the Direct Comparison Test?
e. What do you think is happening with these series: do you think that these small changes are enough to change the behavior of the series (i.e. whether it converges or diverges)?

Answer available at the end of this section.

There are some ways around this. When we find a comparison that we think is reasonable, but has the wrong inequality for us to make a conclusion, we can do one of two things:

1. Try a different comparison.
2. Try a different test.

We'll address the second strategy momentarily.

Note 8.7.8

In fact, we might argue that we could start with a different test! It might be more useful to skip using the Direct Comparison Test in favor of just using the Limit Comparison Test, which we'll see soon, or the Ratio Test or Root Test, which we'll talk about in the next section. I think that the only real reason to think about the Direct Comparison Test is that

- Sometimes it can be pretty quick to use. Sometimes.
- It is much easier to see why it works, since we can make a nice argument about monotonic and bounded sequences of partial sums. From there, the Limit Comparison Test (coming up next) isn't a big jump conceptually. But it might be harder to start with.

So let's think about one of the series from Activity 8.7.7:

$$\sum_{k=1}^{\infty} \frac{2k}{3^k}$$

We, intuitively, tried to compare this a related geometric series:

$$\sum_{k=1}^{\infty} \frac{2k}{3^k} \sim \underbrace{\sum_{k=1}^{\infty} \frac{2}{3^k}}_{\text{converges}}$$

Unfortunately, we weren't able to make this connection, since $\frac{2k}{3^k} \geq \frac{2}{3^k}$ for $k = 1, 2, 3 \ldots$

We could, instead, still think that this is a converging series but compare it to the converging p-series, $\sum_{k=1}^{\infty} \frac{1}{k^2}$. We'll actually compare it to $\sum_{k=1}^{\infty} \frac{2}{k^2}$, but this will converge as well (since the coefficient will just scale the value that the series converges to).

Note, first, that $3^k \geq k^3$ for $k \geq 3$. At $k = 3$ we have $3^3 = 3^3$, but after this intersection point, the exponential will be larger than the power function.

$$\begin{aligned} 3^k &\geq k^3 \text{ for } k = 3, 4, 5, \ldots \\ \frac{1}{3^k} &\leq \frac{1}{k^3} \text{ for } k = 3, 4, 5 \ldots \\ \frac{2k}{3^k} &\leq \frac{2k}{k^3} \text{ for } k = 3, 4, 5 \ldots \\ \frac{2k}{3^k} &\leq \frac{2}{k^2} \text{ for } k = 3, 4, 5 \ldots \end{aligned}$$

And there we go! We have $\frac{2k}{3^k} \leq \frac{2}{k^2}$ for $k = 3, 4, 5 \ldots$ and $\sum_{k=3}^{\infty} \frac{2}{k^2}$ converges, so then $\sum_{k=3}^{\infty} \frac{2k}{3^k}$ must also converge.

Of course, we can add two more real numbers (when $k = 1$ and $k = 2$), and the resulting series will still converge. So, $\sum_{k=1}^{\infty} \frac{2k}{3^k}$ converges, just like we originally thought.

Let's not get ahead of ourselves: this quick change from one comparison to another is rarely easy! Here are some questions we can ask:

- Where did $\sum_{k=1}^{\infty} \frac{1}{k^2}$ come from? How could we have anticipated that as being useful?
- Didn't we think, originally, that this was almost a geometric series? Why are we switching to comparing to a p-series?
- Is there some systematic way for us to think about what to compare to? Something other than appealing to growth rates?

These questions will remain largely unanswered in this text, other than the following (unhelpful) solution: getting good at thinking about inequalities and comparisons means that inequalities and comparisons become easier.

This is not intended to be a self-indulged brag: I am not that good at thinking about inequalities and comparisons either! It's just meant to have you confront the fact that this test, while understandable, is not always useful in practice. There are better paths forward!

Let's follow those instead.

Limit Comparison

Let's revisit some of our intuition from earlier. When we talked about the series in Example 8.7.6, we made the claim that it probably acted like the p-series $\sum_{k=1}^{\infty} \frac{1}{k^3}$. Why did we choose this series?

Similarly, how did we pick our comparisons in Activity 8.7.7?

We had thought about what parts of the functions defining the terms of the series would take over as $k \to \infty$. We were thinking about which parts of these terms ends up dominating, in behavior, over the other parts.

We're thinking about limits, really!

We're going to use this intuition that we have about limits and relative growth rates to come up with another way of thinking about whether two things act similarly. There are a few ways that we could approach this, but it will be useful to think about this comparison of relative growth rates as ratios.

What if we think about the ratio of the functions defining the terms of the series we're comparing, and see how this ratio acts in the limit as $k \to \infty$?

Activity 8.7.9: Ratios for Comparison.

Let's start with some functions: we'll consider $f(x)$ and $g(x)$ as two functions that are continuous when $x \geq 0$ with $f(x) \to 0$ and $g(x) \to 0$ as $x \to \infty$.

All of this is so that we can think about $\frac{f(x)}{g(x)}$ and know that we have an indeterminate form. We could put the requirement of differentiability on these functions (so that we could think about L'Hôpital's Rule), but we don't need to do that. We're going to now consider the limit: $\lim_{x\to\infty} \frac{f(x)}{g(x)}$.

a. What would the limit $\lim_{x\to\infty} \frac{f(x)}{g(x)}$ look like if $g(x) \to 0$ at a faster rate than $f(x)$ does? In this case, we might say that:

$$f(x) \gg g(x)$$

Hint. What does it normally look like when a fraction of numbers has a very small denominator compared to the numerator?

b. What would the limit $\lim_{x\to\infty} \frac{f(x)}{g(x)}$ look like if $f(x) \to 0$ at a faster rate than $g(x)$ does? In this case, we might say that:

$$g(x) \gg f(x)$$

Hint. What does it normally look like when a fraction of numbers has a very large denominator compared to the numerator?

c. If the functions $f(x)$ and $g(x)$ eventually act equivalently, then what does the limit $\lim_{x\to\infty} \frac{f(x)}{g(x)}$ look like? **Hint 1**. When we say that these two functions act equivalently, we might mean that while they both approach 0 in the limit, they do it in the same way, or with identical growth rates. **Hint 2**. What does it normally look like when a fraction of numbers is something over an equivalent thing?

d. If the function $f(x)$ eventually acts like some scaled version of $g(x)$, then what does the limit $\lim_{x\to\infty} \frac{f(x)}{g(x)}$ look like? **Hint**. We can think that $f(x) \to L(g(x))$ as $x \to \infty$ for some real number L.

So we have three outcomes of our comparisons using ratios of functions in the limit:

1. The numerator function could, eventually, be so much larger than the denominator function that the limit of the ratio is infinite.
2. The denominator could, eventually, be so much larger than the numerator function that the limit of the ratio is zero.
3. The numerator and denominator could, eventually, act so similarly to each other (or like scaled versions of each other) that the limit is some real number that isn't 0.

This is the motivation for the next comparison test! We'll just think about the functions defining the terms of a series instead, and we'll make conclusions about the series themselves instead of the functions defining the terms.

Theorem 8.7.10: Limit Comparison Test.

If $\sum_{k=0}^{\infty} a_k$ and $\sum_{k=0}^{\infty} b_k$ are infinite series with positive terms ($a_k > 0$ and $b_k > 0$ for $k \geq 0$), then we can consider $\lim_{k\to\infty} \frac{a_k}{b_k}$.

- If $\lim_{k\to\infty} \frac{a_k}{b_k} = 0$, then:
  - If $\sum_{k=0}^{\infty} a_k$ diverges, then $\sum_{k=0}^{\infty} b_k$ diverges as well.
  - If $\sum_{k=0}^{\infty} b_k$ converges, then $\sum_{k=0}^{\infty} a_k$ converges as well.
- If $\lim_{k\to\infty} \frac{a_k}{b_k} = \infty$, then:
  - If $\sum_{k=0}^{\infty} a_k$ converges, then $\sum_{k=0}^{\infty} b_k$ converges as well.
  - If $\sum_{k=0}^{\infty} b_k$ diverges, then $\sum_{k=0}^{\infty} a_k$ diverges as well.
- If $\lim_{k\to\infty} \frac{a_k}{b_k} = L$ where L is some non-zero real number, then $\sum_{k=0}^{\infty} a_k$ and $\sum_{k=0}^{\infty} b_k$ will either both converge or both diverge.

Example 8.7.11

For each of the following infinite series, try to select an appropriate comparison series, and then apply a comparison test to make conclusions about whether the series converges or diverges.

a. $\sum_{k=1}^{\infty} \frac{e^{1/k}}{\sqrt{k^2+k}}$ **Hint 1**. What happens, in the limit as $k \to \infty$ to $e^{1/k}$? How does $\sqrt{k^2+k}$ act in the limit: does the k term influence much, compared to k^2? **Hint 2**. One of the comparisons we can try is to link the behavior of these two series:

$$\sum_{k=1}^{\infty} \frac{e^{1/k}}{\sqrt{k^2+k}} \sim \sum_{k=1}^{\infty} \frac{1}{\sqrt{k^2}} = \sum_{k=1}^{\infty} \frac{1}{k}$$

b. $\sum_{k=0}^{\infty} \frac{1}{k!}$ **Hint**. This is a hard one to come up with a reasonable comparison. Try writing out some terms and getting a feel for what kinds of things are happening structurally: • Is there something you can describe, recursively, about how we get from one term in the series to the next? • Are there consistent operations that we're applying to terms? • Does this remind you of anything?

Solution

a. In order to link these series, we'll apply a limit comparison test:

$$\sum_{k=1}^{\infty} \frac{e^{1/k}}{\sqrt{k^2+k}} \sim \sum_{k=1}^{\infty} \frac{1}{\sqrt{k^2}} = \sum_{k=1}^{\infty} \frac{1}{k}$$

So let's investigate the limit of the ratio of the terms.

$$\lim_{k\to\infty}\frac{\left(\frac{e^{1/k}}{\sqrt{k^2+k}}\right)}{\left(\frac{1}{k}\right)} = \lim_{k\to\infty}\left(\frac{e^{1/k}}{\sqrt{k^2+k}}\right)\left(\frac{k}{1}\right)$$
$$= \lim_{k\to\infty}\frac{ke^{1/k}}{\sqrt{k^2+k}}$$
$$= \left(\lim_{k\to\infty}e^{1/k}\right)\left(\lim_{k\to\infty}\frac{k}{\sqrt{k^2+k}}\right)$$
$$= \lim_{k\to\infty}\frac{k}{\sqrt{k^2+k}}$$

This limit is one that we can think about using Theorem 1.4.9: we can apply the limit in the denominator under the root, and notice that the whole thing is really dependent on the behavior of k^2. We'll apply a technique that is used in the proof of the theorem.

$$\lim_{k\to\infty}\frac{k}{\sqrt{k^2+k}} = \lim_{k\to\infty}\frac{k}{\sqrt{k^2\left(1+\frac{k}{k^2}\right)}}$$
$$= \lim_{k\to\infty}\frac{k}{\sqrt{k^2}\sqrt{1+\frac{1}{k}}}$$
$$= \lim_{k\to\infty}\frac{k}{k\sqrt{1+\frac{1}{k}}}$$
$$= \lim_{k\to\infty}\frac{1}{\sqrt{1+\frac{1}{k}}}$$

Now, since $\frac{1}{k}\to 0$ as $k\to\infty$ we end up with:

$$\lim_{k\to\infty}\frac{\left(\frac{e^{1/k}}{\sqrt{k^2+k}}\right)}{\left(\frac{1}{k}\right)} = 1$$

Conclusions: Since we're comparing our series to the diverging p-series $\sum_{k=1}^{\infty}\frac{1}{k}$, and the limit comparison test says that these two infinite series must have the same behavior (since the limit of the ratio of the term functions was 1), then we can conclude that the infinite series $\sum_{k=1}^{\infty}\frac{e^{1/k}}{\sqrt{k^2+k}}$ also diverges.

b. We can compare this to a converging geometric series, $\sum_{k=0}^{\infty}\frac{1}{2^k}$. So, we want to draw the following link:

$$\sum_{k=0}^{\infty}\frac{1}{k!}\sim\sum_{k=0}^{\infty}\frac{1}{2^k}$$

The limit comparison test follows as such:

$$\lim_{k\to\infty}\frac{\left(\frac{1}{k!}\right)}{\left(\frac{1}{2^k}\right)} = \lim_{k\to\infty}\frac{2^k}{k!}$$

We cannot use L'Hopital's Rule here, since $k!$ is not a continuous function for real numbers (since it only takes in non-negative integer inputs) and so it is not differentiable. We can, instead, appeal to growth rates: $k!$ approaches infinity much faster than 2^k, and so this fraction has a much larger denominator.

$$\lim_{k\to\infty}\frac{2^k}{k!} = 0$$

Conclusions: By the Limit Comparison Test, this means that $\frac{1}{k!} \ll \frac{1}{2^k}$, and so, since $\sum_{k=0}^{\infty} \frac{1}{2^k}$ is a converging geometric series, then the infinite series $\sum_{k=0}^{\infty} \frac{1}{k!}$ must also converge.

There isn't anything special about the series we're comparing to: We could have used any mix of the following comparisons to find the same conclusions:

- $\sum_{k=1}^{\infty} \frac{e^{1/k}}{\sqrt{k^2+k}} \sim \sum_{k=1}^{\infty} \frac{1}{\sqrt{k}}$
- $\sum_{k=0}^{\infty} \frac{1}{k!} \sim \sum_{k=1}^{\infty} \frac{1}{k^2}$
- $\sum_{k=0}^{\infty} \frac{1}{k!} \sim \sum_{k=1}^{\infty} \frac{1}{10^k}$

There's no magic series to compare to: one of the nice things about the limit comparison test is that we can compare a series to another one not based on similar structure, but based on intuited behavior: if you think a series converges, compare it to a converging p-series if you'd like!

There certainly are some instances where it is reasonable to pick a specific series to compare to based on the structure.

Example 8.7.12

Consider the following series:

$$\sum_{k=1}^{\infty} \frac{2k^2 - k + 3\sqrt{k}}{3k^{7/3} + 4k^{5/3} - k^{2/5}}$$

Perform a test and state a conclusion about whether or not this series converges.

Hint 1. There are a lot of power functions here! Which ones do you think are most important in deciding how quickly the terms approach 0?

Hint 2. This isn't a p-Series, but it might act like one. Compare it to a relevant p-Series!

Solution

We can note that the numerator is really driven by the quadratic term, k^2, while the denominator's behavior is determined by $k^{7/3}$, the power function with the highest exponent. We can make the following comparison:

$$\begin{aligned}\sum_{k=1}^{\infty} \frac{2k^2 - k + 3\sqrt{k}}{3k^{7/3} + 4k^{5/3} - k^{2/5}} &\sim \sum_{k=1}^{\infty} \frac{k^2}{k^{7/3}} \\ &\sim \sum_{k=1}^{\infty} \frac{1}{k^{1/3}}\end{aligned}$$

Note that this is a diverging p-series, and so, in performing this comparison, we think that our series also converges. Let's show this using a limit comparison.

$$\begin{aligned}\lim_{k\to\infty} \frac{\left(\frac{1}{k^{1/3}}\right)}{\left(\frac{2k^2 - k + 3\sqrt{k}}{3k^{7/3} + 4k^{5/3} - k^{2/5}}\right)} &= \lim_{k\to\infty} \left(\frac{1}{k^{1/3}}\right)\left(\frac{3k^{7/3} + 4k^{5/3} - k^{2/5}}{2k^2 - k + 3\sqrt{k}}\right) \\ &= \lim_{k\to\infty} \frac{3k^{7/3} + 4k^{5/3} - k^{2/5}}{2k^{7/3} - k^{4/3} + 3k^{5/6}} \\ &= \frac{3}{2}\end{aligned}$$

Conclusions : Since this limit is a non-zero real number, we can conclude that our two series have the same behavior. So, since $\sum_{k=1}^{\infty} \frac{1}{k^{1/3}}$ diverged, then we know that the series $\sum_{k=1}^{\infty} \frac{2k^2 - k + 3\sqrt{k}}{3k^{7/3} + 4k^{5/3} - k^{2/5}}$ must also diverge.

This example is not unique! Note that out selection of the p-series to use as comparison is based on what we know will happen in the limit that we eventually use in the test itself!

Theorem 8.7.13: Rational Comparison Theorem.

If a_k is a rational function of k, $a_k = \frac{p(k)}{q(k)}$ where $p(k)$ is a degree m polynomial function and $q(k)$ is a degree n polynomial function, then:

- If $n - m > 1$, then $\sum_{k=1}^{\infty} a_k$ converges.
- If $n - m \leq 1$, then $\sum_{k=1}^{\infty} a_k$ diverges.

Note 8.7.14

We can extend this result pretty easily by loosening up the "rational function" requirement. If we have combinations of power functions (even with non-integer exponents), this works as well!

So we have a great tool for analyzing series that look similar to p-series: as long as the most aggressive pieces of our terms are defined by power functions, then we can connect the series to the relevant p-series and use the same convergence criteria through arguments about the degrees of these power functions.

What about infinite series that are seemingly not connected to a p-series? We saw other series in this section that acted more like a geometric series:

- $\sum_{k=1}^{\infty} \frac{2}{k\,(3^k)}$
- $\sum_{k=1}^{\infty} \frac{2k}{3^k}$
- $\sum_{k=1}^{\infty} \frac{1}{k!}$

We can set up some comparisons, like we did in this section, for each of these series individually. But is there a result similar to Theorem 8.7.13 for series that act like a geometric series?

It turns out that the answer is a resounding, "Yes!" We just need to think about what aspects of a geometric series we're looking for.

Practice Problems

Exercise 8.7.1

For a series $\sum a_k$ where all of the terms a_k are positive, explain whether the series converges, diverges, or if we cannot tell based on the following comparisons.

a. $a_k < \frac{1}{k^2}$ for all k.
b. $a_k > \frac{1}{k^2}$ for all k.
c. $a_k < \frac{1}{k}$ for all k.
d. $a_k > \frac{1}{k}$ for all k.

e. $a_k < 1$ for all k.
f. $a_k > 1$ for all k.

Exercise 8.7.2

For each pair of series, compare the terms of each with either < or >. What does the Direct Comparison Test say about the first series in each pair (if anything)?

a. $\sum_{k=1}^{\infty}\left(\frac{1}{\sqrt{k^3+1}}\right)$ compared with $\sum_{k=1}^{\infty}\left(\frac{1}{k^{3/2}}\right)$
b. $\sum_{k=1}^{\infty}\left(\frac{1}{k^2 4^k}\right)$ compared with $\sum_{k=1}^{\infty}\left(\frac{1}{4^k}\right)$
c. $\sum_{k=2}^{\infty}\left(\frac{k+1}{(k-1)^2}\right)$ compared with $\sum_{k=2}^{\infty}\left(\frac{1}{k}\right)$
d. $\sum_{k=1}^{\infty}\left(\frac{1}{k!}\right)$ compared with $\sum_{k=1}^{\infty}\left(\frac{1}{k^2}\right)$

Exercise 8.7.3

What is an advantage of using the Limit Comparison Test instead of the Direct Comparison Test? Demonstrate this with a specific example.

Exercise 8.7.4

For each pair of series, attempt to compare the terms with either < or > . Explain why the Direct Comparison Test fails (or explain why the comparison is difficult), and then apply the Limit Comparison Test, and explain the conclusions.

a. $\sum_{k=1}^{\infty}\left(\frac{1}{\sqrt{k+1}}\right)$ compared with $\sum_{k=1}^{\infty}\left(\frac{1}{k^{1/2}}\right)$
b. $\sum_{k=1}^{\infty}\left(\frac{k-1}{(k+1)^2}\right)$ compared with $\sum_{k=1}^{\infty}\left(\frac{1}{k}\right)$
c. $\sum_{k=1}^{\infty}\left(\sqrt{\frac{k}{k^2+1}}\right)$ compared with $\sum_{k=1}^{\infty}\left(\frac{1}{\sqrt{k}}\right)$

Exercise 8.7.5

For each series below, determine whether the series converges or diverges using either the Direct or Limit Comparison Test.

a. $\sum_{k=2}^{\infty}\left(\frac{4k}{k^3-1}\right)$
b. $\sum_{k=1}^{\infty}\left(\frac{k}{2^k}\right)$
c. $\sum_{k=1}^{\infty}\left(\frac{3}{\sqrt{k}+\sqrt[3]{k}}\right)$
d. $\sum_{k=2}^{\infty}\left(\sqrt[4]{\frac{k}{k^4-1}}\right)$
e. $\sum_{k=1}^{\infty}\left(\frac{1}{\ln(k)}\right)$
f. $\sum_{k=1}^{\infty}\left(\frac{2}{k^k}\right)$
g. $\sum_{k=1}^{\infty}\left(\frac{k^3-k+1}{k^{9/2}+4k^2+1}\right)$
h. $\sum_{k=1}^{\infty}\left(\frac{k^{3/2}+4}{k^2-k+\sqrt{k}}\right)$

Answers

Activity 8.7.7: (Un)Helpful Comparisons.

a. 1. Let's link the following two series using a comparison:

$$\sum_{k=1}^{\infty} \frac{2}{k\left(3^k\right)} \sim \sum_{k=1}^{\infty} \frac{2}{3^k}$$

We can notice that $\sum_{k=1}^{\infty} \frac{2}{3^k}$ is a geometric series with $a = \frac{2}{3}$ and $r = \frac{1}{3}$. It converges!

(Specifically, this series converges to 1.)

2. Let's link the following two series using a comparison:

$$\sum_{k=3}^{\infty} \frac{\sqrt{k+1}}{k-2} \sim \sum_{k=3}^{\infty} \frac{\sqrt{k}}{k} = \sum_{k=3}^{\infty} \frac{1}{\sqrt{k}}$$

Note that $\sum_{k=3}^{\infty} \frac{1}{\sqrt{k}}$ is a diverging p-series, since $p = \frac{1}{2}$.

b. 1. Since $k \geq 1$, we can say that

$$k\left(3^k\right) \geq 3^k \text{ for } k = 1, 2, 3, \ldots$$

This means that

$$\frac{1}{k\left(3^k\right)} \leq \frac{1}{3^k} \text{ for } k = 1, 2, 3, \ldots$$

since a big denominator leads to a small fraction. Then, we can scale both by a factor of 2:

$$\frac{2}{k\left(3^k\right)} \leq \frac{2}{3^k} \text{ for } k = 1, 2, 3, \ldots$$

Since $\sum_{k=1}^{\infty} \frac{2}{3^k}$ is a converging geometric series (since $r = \frac{1}{3}$ and so $|r| < 1$), then we can say that $\sum_{k=1}^{\infty} \frac{2}{k\left(3^k\right)}$ must also converge.

2. We can see that $k - 2 < k$. Then, when we think about reciprocals, we see that

$\frac{1}{k-2} > \frac{1}{k}$ for $k > 2$

since a small denominator leads to a big fraction. Then, we can multiply both by $\sqrt{k}$:

$\frac{\sqrt{k}}{k-2} > \frac{\sqrt{k}}{k}$

Finally, since $k + 1 > k$,, we know that $\sqrt{k+1} > \sqrt{k}$.

This means that:

$\frac{\sqrt{k+1}}{k-2} > \frac{\sqrt{k}}{k-2} > \frac{\sqrt{k}}{k}$

So, we have that $\frac{\sqrt{k+1}}{k-2} > \frac{1}{\sqrt{k}}$ for $k = 3, 4, 5, \ldots$

Since $\sum_{k=3}^{\infty} \frac{1}{\sqrt{k}}$ is a diverging p-series (since $p = \frac{1}{2} \leq 1$), then $\sum_{k=3}^{\infty} \frac{\sqrt{k+1}}{k-2}$ must also diverge.

c. 1. Since $k \geq$, we know that $2k \geq 2$. This means that

$$\frac{2k}{3^k} \geq \frac{2}{3^k} \text{ for } k = 1, 2, 3, \ldots$$

2. We have $k + 2 > k$, and so:

$$\frac{1}{k+2} < \frac{1}{k} \text{ for } k = 3, 4, 5 \ldots$$

This means that:

$$\frac{\sqrt{k}}{k+2} < \frac{1}{\sqrt{k}} \text{ for } k = 3, 4, 5 \ldots$$

Finally, since $\sqrt{k-1} < \sqrt{k}$, we have:

$$\frac{\sqrt{k-1}}{k+2} < \frac{1}{\sqrt{k}} \text{ for } k = 3, 4, 5 \ldots$$

d. In both cases, the Direct Comparison Test is inconclusive. This means that we cannot conclude that either of these series converges or diverges: we don't have enough information to justify any claim we might make.

This page titled 8.7: Comparison Tests is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Peter Keep.

8.8: The Ratio and Root Tests

We have just learned about one of the big and important tools that are used for testing convergence. The Limit Comparison Test is super useful for rational functions and things that act like p-Series, since these power functions and related function types behave so nicely, and are so friendly to work with, in the end behavior limit. The algebra typically works well, and we can analyze the limits pretty easily.

In this section, we're going to try to draw a similar connection between our other family of common series, the Geometric Series, and series that act similarly. So for us to begin, we want to think about what it might mean for a series to have terms that act similarly to the terms of a geometric series.

Activity 8.8.1: Reminder about Geometric Series.

We are going to build some convergence tests that try to link some infinite series to the family of geometric series and show that even though a series is not geometric, it might act enough like one to be considered "eventually geometric-ish." But first, what does it mean for a series to be a geometric series?

a. Describe a defining characteristic of a geometric series. What makes it geometric?
b. Can you describe this characteristic in another way? For instance, if you described a geometric series using a characteristic about the Explicit Formula, can you describe the same characteristic in the context of the Recursion Relation instead? Or vice versa?
c. Write out a generalized and simplified form of the term a_k of a geometric series explicitly and recursively. In each case, solve for r, the ratio between terms.

Eventually Geometric-ish

We're going to use these two guiding features of a geometric series to determine if a series is geometric-ish. That is, if a series is not actually a geometric series, do the terms act like terms from a geometric series in the limit? Is there some eventually (almost) constant ratio between consecutive terms? Is there one in the limit as $k \to \infty$? If the terms aren't actually exponential functions of k, do they kind of act like that in the limit as $k \to \infty$? If so, shouldn't they act like geometric series and converge with the same criteria?

Theorem 8.8.2: Root Test.

Let $\sum_{k=0}^{\infty} a_k$ be an infinite series with $a_k > 0$ for $k \geq 0$ and consider $\lim_{k\to\infty} \sqrt[k]{a_k}$.

- If there is some real number r with $r = \lim_{k\to\infty} \sqrt[k]{a_k}$ and $0 \leq r < 1$, then the series $\sum_{k=0}^{\infty} a_k$ converges.
- If there is some real number r with $r = \lim_{k\to\infty} \sqrt[k]{a_k}$ and $r > 1$ or if $\lim_{k\to\infty} \sqrt[k]{a_k}$ does not exist, then the series $\sum_{k=0}^{\infty} a_k$ diverges.
- If $\lim_{k\to\infty} \sqrt[k]{a_k} = 1$ then the Root Test fails and is inconclusive.

Theorem 8.8.3: Ratio Test.

Let $\sum_{k=0}^{\infty} a_k$ be an infinite series with $a_k > 0$ for $k \geq 0$ and consider $\lim_{k\to\infty} \frac{a_{k+1}}{a_k}$.

- If there is some real number r with $r = \lim_{k\to\infty} \frac{a_{k+1}}{a_k}$ and $0 \leq r < 1$, then the series $\sum_{k=0}^{\infty} a_k$ converges.
- If there is some real number r with $r = \lim_{k\to\infty} \frac{a_{k+1}}{a_k}$ and $r > 1$ or if $\lim_{k\to\infty} \frac{a_{k+1}}{a_k}$ does not exist, then the series $\sum_{k=0}^{\infty} a_k)$ diverges.
- If $\lim_{k\to\infty} \frac{a_{k+1}}{a_k} = 1$ then the Ratio Test fails and is inconclusive.

So these tests are good and fine, but when do we use them? How do we know that they can be helpful? The key is to notice behavior in the terms of a series that look "kind of" geometric: we're looking for k up in the exponent on things and we're looking for repeated multiplication.

Activity 8.8.4: When Are These Tests Useful?

We're going to look at a couple of small examples where we can rewrite some expressions into friendlier forms, and try to connect these rewriting strategies to the Ratio and Root Tests.

a. Rewrite the following expression into a friendlier form. Explain why this new form is friendlier.

$$\sqrt[k]{\frac{2^{k+1}}{7^{3k}}}$$

b. Rewrite the following expression into a friendlier form. Explain why this new form is friendlier.

$$\sqrt[k]{\frac{(k+1)^k}{4^{2k+1}}}$$

c. Rewrite the following expression into a friendlier form. Explain why this new form is friendlier.

$$\frac{(5^{k+2})\,(6^{k-3})}{(5^{k-1})\,(6^{k+1})}$$

d. Rewrite the following expression into a friendlier form. Explain why this new form is friendlier.

$$\frac{103!}{99!}$$

e. Rewrite the following expression into a friendlier form. Explain why this new form is friendlier.

$$\frac{(2k+4)!}{(2k+2)!}$$

f. Why do you think the Ratio Test especially will be useful for series whose terms include factorials and exponentials? Why do you think the Root Test will be useful for series whose terms include exponentials and functions raised to functions (of k)?

Example 8.8.5

For each infinite series, apply one of the Ratio or Root tests and interpret the conclusions of the test.

a. $\sum_{k=0}^{\infty} \frac{2^k}{k!}$

b. $\sum_{k=0}^{\infty} \frac{k^2}{(k+1)^k}$

c. $\sum_{k=1}^{\infty} \frac{k\ln(k+1)}{e^{2k}}$

d. $\sum_{k=0}^{\infty} \frac{(-1)^k(k+2)}{(k!)\,(3^k)}$

Hint. This is an alternating series! We can show that this series converges using the Alternating Series Test, and so we really need to test for absolute convergence. That works perfectly, though, since the Ratio and Root tests only test series with positive terms. So test the series:

$$\sum_{k=0}^{\infty} \frac{(k+2)}{(k!)\,(3^k)}$$

Inconclusive Results

We've seen some of our series tests have some inconclusive results:

- The Divergence Test is inconclusive when the limit of the terms is 0.
- The Direct Comparison Test is inconclusive when the inequality is the "wrong" way around (like when the series we're interested in has terms smaller than a diverging series, or has terms larger than a converging series). The Limit Comparison Test has the same issue (where the order of the inequality matters) when the limit we get is either 0 or ∞.
- And now we find that the Ratio Test and Root Test are inconclusive when the limit we consider in the test is equal to 1.

So what does it really mean when the conclusion from the series convergence test is "we actually can't tell if our series converges or now?"

In general, this means that we can find an example of a series that converges and an example of a series that diverges, where both give the same result of the test. We know, for example, that every converging series has terms where the limit of the terms is 0, but we have also see examples of series whose terms approach 0 in the limit that diverge. The first example we looked at that did this was the Harmonic Series.

For the Ratio and the Root Tests, specifically, though, we can glean some more information.

We introduced these convergence tests by thinking about whether or not a series could act like a Geometric Series in the long run (in the limit). And the tests do a great job of uncovering this eventually-geometric-ish behavior! So good, in fact, that (almost) any series that eventually acts kind of like a geometric series will be correctly classified as converging or diverging.

> **Aside**.
>
> This claim, that the Ratio and Root test can classify every series that acts kind of like a geometric series is intentionally ambiguous. What does it mean to act "kind of" like a geometric series? We cannot precisely define this term.
>
> We can say, though, that these tests do a great job, but are not perfect. For instance, there are examples of carefully created infinite series where the Root Test classifies it correctly as a converging or diverging series, but the Ratio Test is inconclusive. We will not study this in this text, but it's interesting!

So what, then, does it mean when the Ratio or Root Test is inconclusive? Simply this: it means that the series we're considering does not act like a geometric series.

Let's consider the general p-Series, and apply the Ratio Test:

$$\begin{aligned} r &= \lim_{k\to\infty} \frac{\left(\frac{1}{(k+1)^p}\right)}{\left(\frac{1}{k^p}\right)} \\ &= \lim_{k\to\infty} \frac{k^p}{(k+1)^p} \end{aligned}$$

This limit is 1.

Note that this limit does not depend on the exponent, p, which means that the limit will be 1 when the p-series converges and also when the p-series diverges.

So, in a sense, when we get an inconclusive result from the Ratio or Root Tests, we are learning a lot about our series: it doesn't act like a geometric series, and that's likely because it does act like a p-series! We should try to approach this series again, this time looking for a suitable p-series to compare it to (whether directly or using the Limit Comparison Test).

Practice Problems

Exercise 8.8.1

For the Ratio Test, we note that if $\lim_{k\to\infty}\left|\frac{a_{k+1}}{a_k}\right| < 1$, then the series $\sum a_k$ converges. Why is this? Explain some general intuition for why the Ratio Test guarantees convergence in this case.

Exercise 8.8.2

For the Root Test, we note that if $\lim_{k\to\infty}\sqrt[k]{|a_k|} < 1$, then the series $\sum a_k$ converges. Why is this? Explain some general intuition for why the Root Test guarantees convergence in this case.

Exercise 8.8.3

Note that in both the Ratio and Root Tests above, that we use the absolute value of the terms. Is it possible, then, for the Ratio or Root Test to tell us that a series converges conditionally?

Exercise 8.8.4

Simplify each of the following expressions.

a. $\frac{104!}{103!}$
b. $\frac{22!}{20!}$
c. $\sqrt[k]{\frac{2^k}{e^{k^2}}}$
d. $\frac{(k+4)!}{(k+3)!}$
e. $\frac{3^{k+2}}{3^k}$
f. $\frac{2^{k+1}(k)!}{2^k(k+1)!}$

Exercise 8.8.5

For the following series, use the apply the Ratio or Root Test (whichever is more applicable) and interpret the results.

a. $\sum_{k=1}^{\infty}\frac{(k+1)^2}{k!}$
b. $\sum_{k=1}^{\infty}\left(\frac{1}{k!}\right)^k$
c. $\sum_{k=2}^{\infty}\frac{k^2}{e^k}$
d. $\sum_{k=1}^{\infty}\frac{(-1)^k}{(3k)^k}$
e. $\sum_{k=1}^{\infty}\left(\frac{1}{\ln(k+1)}\right)^k$
f. $\sum_{k=1}^{\infty}\frac{k}{\pi^k}$
g. $\sum_{k=1}^{\infty}\left(\frac{k}{\pi}\right)^k$

h. $\sum_{k=1}^{\infty} \frac{(-1)^{k+1} k^k}{k!}$

i. $\sum_{k=1}^{\infty} \frac{k!}{(2k+1)!}$

Exercise 8.8.6

For what values of p (with $p > 0$) will the following series converge?

$$\sum_{k=1}^{\infty} \frac{k^p}{k!}$$

Exercise 8.8.7

For each of the following infinite series, use any methods we've learned about so far to determine whether the series converges or diverges. If it converges and you are able to tell what it converges to, say so and explain. If you cannot tell what it converges to, say so and explain. Clearly state the test or method you use or the type of series.

a. $\sum_{k=2}^{\infty} \left(\frac{k^2 - 1}{(k^3 - k + 1)^4} \right)$

b. $\sum_{k=2}^{\infty} \left(\frac{1}{k \ln(k^4)} \right)$

c. $\sum_{k=0}^{\infty} \left(\frac{(-1)^k + 2^k}{e^{2k}} \right)$

d. $\sum_{k=2}^{\infty} \left(\frac{1}{k^2 - 1} \right)$

e. $\sum_{k=2}^{\infty} \left(\frac{1}{k\sqrt{k^2 - 1}} \right)$

f. $\sum_{k=0}^{\infty} \left(\frac{4^k (k^2)}{k!} \right)$

g. $\sum_{k=2}^{\infty} \left(\frac{\ln(k)}{k + 2} \right)$

This page titled 8.8: The Ratio and Root Tests is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Peter Keep.

8.9: Exploration- Infinite Series

Interesting Infinite Series

We've seen a lot of infinite series examples, and we have a lot of tools for how to think about infinite series. Take a look at some more examples that are a bit strange to think about.

Exploration 8.9.1: Interesting Infinite Series.

a. Consider the series

$$\sum_{k=1}^{\infty} \frac{1}{D(k)k^p}$$

where $D(k)$ is the function that returns the number of digits in the number k (so $D(9) = 1$ and $D(435) = 3$), and p is a real number. For what values of p does this series converge?

b. Consider the series

$$\sum_{k=1}^{\infty} \frac{1}{F(k)k^p}$$

where $F(k)$ is the function that returns the kth Fibbonacci number, and p is a real number. As a reminder, the Fibbonacci sequence is:

$$1, 1, 2, 3, 5, 8, 13, 21 \ldots$$

where $F(k) = F(k-1) + F(k-2)$. For what values of p does this series converge?

This page titled 8.9: Exploration- Infinite Series is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Peter Keep.

9: POWER SERIES

9.1: Polynomial Approximations of Functions
9.2: Power Series Convergence
9.3: How to Build Taylor Series
9.4: How to Use Taylor Series
9.5: Exploration- Power Series

This page titled 9: Power Series is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Peter Keep.

9.1: Polynomial Approximations of Functions

Before we start, it might be helpful to remind ourselves of the way we have used linear functions to approximate other functions in Section 4.6 Linear Approximations.

What Do We Want From a Polynomial Approximation?

It's always good to lay out our expectations clearly. We want to make sure that, when we create this new thing, that we have some clear idea of what we're trying to accomplish in its creation. So let's start with our Linear Approximation of a Function. What were the important properties of this linear function we created?

1. The function's output, $f(x)$, matched the output from the linear approximation, $L(x)$, at the center $x = a$.
$$\begin{aligned} L(x) &= f'(a)(x-a) + f(a) \\ L(a) &= f'(a)(a-a) + f(a) \\ &f(a) \end{aligned}$$
2. The function's first derivative, $f'(x)$, matched the slope of the linear approximation, $L'(x)$, at $x = a$.
$$\begin{aligned} L'(x) &= \frac{d}{dx}\left(f'(a)(x-a) + f(a)\right) \\ &= f'(a) \\ L'(a) &= f'(a) \end{aligned}$$

This seems like a good structure! We have $f(a) = L(a)$, giving us that nice intersection between our approximation and the function we're approximating at the center. Then we used the slope of the function to estimate how our approximating function should move away from that center. The only real problem is that our function probably doesn't have a constant slope, while a constant slope is the defining characteristic of the line we're using.

So how do we extend this, then, into a better approximation? Well, an easy next step is to make the slope of our approximating function change as we move away from the center, $x = a$. That way, maybe the slopes change in a way that's similar to the slopes of the actual function, and our approximating curve (not a line anymore) can follow the function for a bit longer.

What we're saying is that we want a function where the second derivative is $f''(a)$, just like our first derivative matched $f'(a)$. Instead of using letters like L for linear and Q for quadratic, let's just call these approximations by their function type (polynomials!) and degree. So $p_1(x) = f'(a)(x-a) + f(a)$, the first-degree polynomial approximation.

Now, to find $p_2(x)$, we're going to add a term to the first-degree polynomial:

$$p_2(x) = \square\,(x-a)^2 + f'(a)(x-a) + f(a)$$

Let's differentiate this function twice, and force it to match $f''(a)$ at $x = a$.

$$\begin{aligned} p_2(x) &= \square\,(x-a)^2 + f'(a)(x-a) + f(a) \\ p_2'(x) &= 2(\square\,(x-a) + f'(a) \\ p_2''(x) &= 2(\square\,) \end{aligned}$$

What do we need to fill in the blank to make this match $f''(a)$?

$$\begin{aligned} p_2''(x) &= 2\left(\frac{f''(a)}{2}\right) \\ &= f''(a) \end{aligned}$$

So we get:

$$p_2(x) = \frac{f''(a)}{2}(x-a)^2 + f'(a)(x-a) + f(a)$$

What if we wanted a higher degree? Like, $p_3(x)$? Let's repeat the same process and see what happens!

$$p_3(x) = \square\,(x-a)^3 + \frac{f''(a)}{2}(x-a)^2 + f'(a)(x-a) + f(a)$$
$$p_3'(x) = 3(\square\,)(x-a)^2 + f''(a)(x-a) + f'(a)$$
$$p_3''(x) = 3(2)(\square\,)(x-a) + f''(a)$$
$$p_3'''(x) = 3(2)(\square\,)$$

If we, again, want this third derivative to match $f'''(a)$ (so that the rate at which the slope changes as we move away from the center changes in the same way that it does on f... whew, that is going to be hard to interpret!), then we need to fill the blank in with $\frac{f'''(a)}{3(2)}$.

How Do We Build a Polynomial Approximation?

Let's jump into a generalization of what we've just done. Answer these few questions with yourself, just to make sure you can see where we're going:

1. Why, in the coefficients for each term, do we have the derivative that matches the degree of the term? (First derivative for the first degree term, second derivative for the second degree term, and third derivative for the third degree term.)
2. Why, in the coefficients for each term, do we divide? What are we dividing by, and why do we need these numbers specifically?
3. If we add a 4th, 5th, and 6th term, what will we divide those 4th, 5th, and 6th derivatives ($f^{(4)}(a)$, $f^{(5)}(a)$, and $f^{(6)}(a)$) by?

Definition 9.1.1: Polynomial Approximation.

If $f(x)$ is a function that is n-times differentiable at $x = a$ (that is, the function/derivative values $f(a)$, $f'(a)$, $f''(a)$, ..., $f^{(n)}(a)$ all exist), then the nth degree polynomial approximation of $f(x)$ centered at $x = a$ is:

$$p_n(x) = f(a) + f'(a)(x-a) + \frac{f''(a)}{2}(x-a)^2 + \ldots + \frac{f^{(n)}(a)}{n!}(x-a)^n$$
$$= \sum_{k=0}^{n} \frac{f^{(k)}(a)}{k!}(x-a)^k$$

Activity 9.1.2: Build a Polynomial.

We're going to use the formula in Definition 9.1.1 to construct two different polynomials that approximate two different approximations. Then, we'll use them to approximate things!

a. We're going to start with approximating the function $f(x) = \sin(x)$ centered at $x = 0$. Let's choose to look at a 5th degree polynomial. This means we'll need to find the first five derivatives of $f(x) = \sin(x)$. Then, we'll evaluate our function and the five derivatives at the center. After that, we can divide by the relevant factorial in order to create the coefficients of our polynomial. Fill out the following chart to produce these coefficients.

Table 9.1.3 Coefficients for Polynomial Approximation

k	$f^{(k)}(x)$	$f^{(k)}(a)$	$\frac{f^{(k)}(a)}{k!}$
$k = 0$	$f(x) = \sin(x)$	$f(0) =$	
$k = 1$	$f'(x) =$	$f'(0) =$	
$k = 2$	$f''(x) =$	$f''(0) =$	
$k = 3$	$f'''(x) =$	$f'''(0) =$	
$k = 4$	$f^{(4)}(x) =$	$f^{(4)}(0) =$	

k	$f^{(k)}(x)$	$f^{(k)}(a)$	$\frac{f^{(k)}(a)}{k!}$
$k=5$	$f^{(5)}(x)=$	$f^{(5)}(0)=$	

b. Now we can use these coefficients to construct the polynomial! These coefficients should all be on power functions in the form $(x-a)^k$ for $k=0,1,\ldots,5$. These (added together) will form your polynomial, $p_5(x)$.
c. Approximate $f(1)=\sin(1)$ using your polynomial.
d. Let's repeat this for another function. Let's build a 5th degree polynomial approximation of $g(x)=e^x$ centered at $x=0$. We can construct the coefficients in the same way.

Table 9.1.4 Coefficients for Polynomial Approximation

k	$g^{(k)}(x)$	$g^{(k)}(a)$	$\frac{g^{(k)}(a)}{k!}$
$k=0$	$g(x)=e^x$	$g(0)=$	
$k=1$	$g'(x)=$	$g'(0)=$	
$k=2$	$g''(x)=$	$g''(0)=$	
$k=3$	$g'''(x)=$	$g'''(0)=$	
$k=4$	$g^{(4)}(x)=$	$g^{(4)}(0)=$	
$k=5$	$g^{(5)}(x)=$	$g^{(5)}(0)=$	

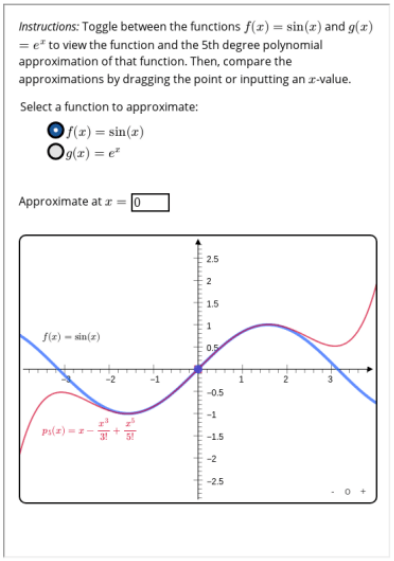
e. And now, again, we can use these coefficients to construct the polynomial! These coefficients should all be on power functions in the form $(x-a)^k$ for $k=0,1,\ldots,5$. These (added together) will form your polynomial, $p_5(x)$.
f. Approximate $g(-3)=\dfrac{1}{e^3}$ using your polynomial.

Now that we know how to build and use these, we should probably think about accuracy. Are the estimations coming from these polynomials even accurate? How do we talk about that?

We won't formally define this too much for now: instead, we'll just look at things visually and see if we can figure out what might go into how we talk about accuracy of our estimations.

Activity 9.1.5How Good Are Our Approximations?:

We're going to think more carefully about our approximations of $\sin(1)$ and e^{-3} from Activity 9.1.2. In order for us to do this, let's visualize the function and the 5th degree polynomial for it.



Standalone
In Context
Embed

a. How good of a job did the polynomial approximation do when approximating $\sin(1)$? How can you tell, visually?
b. How good of a job did the polynomial approximation do when approximating e^{-3}? How can you tell, visually?
c. How does the relationship between the "center" and the x-value that we're approximating at impact the accuracy of our approximation?
d. How do you think you could make these approximations better (without changing the center)?

So we have a couple of main ideas about the accuracy of our approximations. We don't need to formalize them, but we can use them as a guiding rule for how we talk about these polynomial approximations.

Accuracy in Polynomial Approximations.

- Approximations using x-values closer to the center are likely to be more accurate than approximations using x-values farther away from the center.
- Polynomials with larger degrees give more accurate approximations than polynomials with smaller degrees at the same x-values.

Are These Partial Sums?

We have been using some familiar language here...we're talking about these "approximations" improving as we increase some parameter, n. We have some intuition that when we increase n, these approximations "approach" an object (in this case, a function) in some sense.

We're adding more and more terms to this sum as we increase n. Is a polynomial approximation just a partial sum? Are we just going to look at a limit as $n \to \infty$ and see what happens?

Activity 9.1.6: Partial Sums of What?

Let's revisit our 5th degree polynomial approximations from Activity 9.1.2.

$$\sin(x) \approx x - \frac{x^3}{3!} + \frac{x^5}{5!}$$
$$e^x \approx 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \frac{x^4}{4!} + \frac{x^5}{5!}$$

These approximations work well for x-values that are close to 0, but we will not be more formal than that.

a. Make a conjecture about what the 7th degree polynomial approximations are for each of these functions. What about the 15th degree?
b. Make a conjecture about what the general formula would be for these terms. If you were to write these out using summation notation, what would they look like?
c. Why does the polynomial approximation for the sine function only have odd-exponent terms?
d. Make a conjecture about what a polynomial approximation for $f(x) = \cos(x)$ centered at $x = 0$ would be.

We'll explore these polynomials more, but we're specifically interested in the infinite-series version of these things. We'll give these polynomials a name, for easy reference: **Taylor polynomials**.

Aside.

These polynomials are named after the 17th century British mathematician, Brook Taylor. Taylor is mostly famous for his contributions in this area of calculus, but he also sat on the committee that mediated a major dispute between Isaac Newton and Gottfried Leibniz centered around their individual claims to get credit for the "creation" of calculus.

Taylor was not able to give an unbiased judgment, due to his public support of Isaac Newton.

The main result to come out of this has to do with the relationship between these polynomials and the function they are approximating.

Theorem 9.1.7: Taylor's Theorem.

If $f(x)$ is a function that is n-times differentiable at $x = a$, then there is some remainder function, $R_n(x)$, such that:

$$f(x) = \left(\sum_{k=0}^{n} \frac{f^{(k)}(a)}{k!}(x-a)^k \right) + R_n(x)(x-a)^n$$

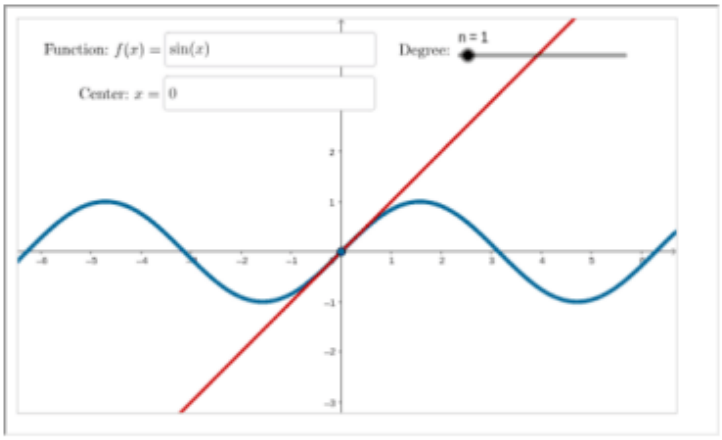
That is, the remainder is small for x-values close to the center, $x = a$.

This result is great, in that it gives us the confidence to approximate these functions. We can add on to this the idea that we expect $R_n(x)$ to get smaller as $n \to \infty$. This should lead us to some idea of convergence, which we can think about here.

Activity 9.1.8: How Do These Polynomials Converge?

We're going to end here by thinking about these polynomials as some partial sum from an infinite series. If there is an infinite series, we should be prepared to think about convergence!

We're going to think about convergence in the same way that we have already: as an end behavior limit of the partial sums. So let's spend our time investigating this end behavior by visualizing polynomial approximations as the degree increases.



Standalone
In Context
Embed

a. What happens to the polynomial approximation of $\sin(x)$ centered at $x = 0$ as the degree $n \to \infty$?
b. Does this behavior change if we centered our approximation elsewhere?
c. What happens to the polynomial approximation of e x centered at $x = 0$ as the degree $n \to \infty$?
d. Does this behavior change if we centered our approximation elsewhere?
e. What happens to the polynomial approximation of $\cos(x)$ centered at $x = 0$ as the degree $n \to \infty$?
f. Does this behavior change if we centered our approximation elsewhere?
g. What happens to the polynomial approximation of $\ln(x)$ centered at $x = 2$ as the degree $n \to \infty$?
h. Describe the difference in what you're seeing with the log function compared to the other functions we've thought about. Describe how the polynomial approximations converge: do they converge to the log function? How? More importantly, where?
i. Does this behavior change if we centered our approximation elsewhere?

Ok, so this is pretty interesting! For some functions, like $\sin(x), e^x$, and $\cos(x)$, it seems like the polynomials that we built will end up matching (converging to) the functions pretty much everywhere: as $n \to \infty$ we can get any function value from the polynomial, approximated to whatever accuracy we'd like!

But that's different for the log function. We were only able to get our polynomial to match the function's behavior on a specific interval of x-values. No matter how high the degree of the polynomial was, we weren't able to get it close to approximating something like $\ln(10)$, for instance. Unless we changed the center, of course!

This is going to bring up some great questions about how these partial sums converge to functions. We'll talk all about that in the next section!

Practice Problems

Exercise 9.1.1

Consider a polynomial approximation of some function $f(x)$ centered at 0. Explain how the degree of the polynomial approximation impacts the accuracy of the approximation.

Exercise 9.1.2

Consider a function $f(x)$ with a a second degree polynomial, $p(x) = \frac{f''(a)}{2}(x-a)^2 + f'(a)(x-a) + f(a)$. We want to approximate $f(b)$ using the polynomial by evaluating $p(b)$. Explain how the location of b in relation to the center a impacts the accuracy of the approximation.

Exercise 9.1.3

Say we want to approximate $\sqrt[3]{7}$ using a third degree polynomial.

a. Which function, $f(x)$, will you use to build the polynomial?
b. What is an appropriate value for the center, a, for your polynomial? Why is this value appropriate?
c. Build the polynomial approximation, $p(x)$, to approximate $f(x)$.
d. Evaluate your polynomial at $x = 7$. What does this tell you about the value of $\sqrt[3]{7}$?

Exercise 9.1.4

Consider the function $f(x) = \sin(x)$. Our goal is to approximate $\sin(0.023)$.

a. Build a third degree polynomial approximation of $f(x)$ centered at 0 and use it to approximate $\sin(0.023)$.
b. Build a fifth degree polynomial approximation of $f(x)$ centered at 0 and use it to approximate $\sin(0.023)$.
c. What would you do to get a better approximation of the value $\sin(0.023)$?
d. What could you do to get the exact value of $\sin(0.023)$ using polynomials?

Exercise 9.1.5

For each of the following functions, find a fourth degree polynomial to approximate the given value.

a. Approximate $\sqrt{1.05}$ using $f(x) = \sqrt{1+x}$.
b. Approximate $\ln(0.95)$ using $f(x) = \ln(x)$.
c. Approximate $e^{-0.12}$ using $f(x) = e^{-x}$.
d. Approximate $\cos(-0.15)$ using $f(x) = \cos(x)$.
e. Approximate $\frac{1}{1.02}$ using $f(x) = \frac{1}{x+1}$.

This page titled 9.1: Polynomial Approximations of Functions is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Peter Keep.

9.2: Power Series Convergence

At the end of Section 9.1 Polynomial Approximations of Functions, we saw that these polynomial approximations can be thought of as partial sums of some larger infinite series. These infinite series are begging us to think about different notions of convergence, and at the end of the section, we saw that the polynomials, as the degree increases off to infinity, converge to match the function they are approximating, but this might be dependent on some interval of x-values. A domain, in a sense.

In this section, we'll investigate what it means for a power series, these infinite series of power functions, to converge. Let's define a power series, and then we can think about convergence from there.

Definition 9.2.1: Power Series.

A power series centered at $x = a$ is an infinite series in the form:

$$\sum_{k=0}^{\infty} c_k(x-a)^k = c_0 + c_1(x-a) + c_2(x-a)^2 + \ldots$$

where $\{c_k\}_{k=0}^{\infty}$ is a sequence of real numbered coefficients.

We have a good idea of how we can build these sequences of coefficients in order for the power series we construct to converge to specific functions that we are interested in.

Before we state this formally, let's write down what we mean when we talk about convergence of power series.

One last thing: we have a kind of closure, so far, in our series. If we add up an infinite amount of numbers, then the infinite series might converge. If it does, it converges to a number (since sums of numbers are numbers). In a power series, though, we are adding up an infinite number of functions of x. If this series converges, it will converge to a function of x (since sums of functions are functions). So every power series is really a function.

Interval of Convergence

Activity 9.2.2: Polynomial Division.

We're going to do some fiddling with polynomials, and hopefully use this as a bridge to connect how we think of polynomials and power series with how we think about our traditional infinite series and the notions of convergence that we've already built.

a. We're going to factor some polynomials, but we might end up using some division. First, we'll confirm some factors that we already know.

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

We'll confirm this by using division.

$$x-1 \;\overline{)\; x^2 \qquad -1}$$

b. Now let's factor $x^3 - 1$. If the factors for this polynomial isn't as familiar, it might be helpful to know that $(x-1)$ is also a factor of $x^3 - 1$.

$$x-1 \;\overline{)\; x^3 \qquad -1}$$

c. Let's try another one. Complete the following division.

$$x-1 \;\overline{)\; x^4 \qquad\qquad -1}$$

d. Can you generalize this? Find the formula for $\frac{x^n-1}{x-1}$ for some positive integer n.

$$x-1 \;\overline{)\; x^n \qquad\qquad -1}$$

e. Now that we have good evidence that

$$\sum_{k=0}^{\infty} x^k = \frac{x^n-1}{x-1}$$

We can apply a limit as $n \to \infty$.

$$\begin{aligned}\sum_{k=0}^{\infty} x^k &= \lim_{n\to\infty} \sum_{k=0}^{n} x^k \\ &= \lim_{n\to\infty} \frac{x^n-1}{x-1}\end{aligned}$$

For what values of x will this limit exist?

f. Revisit Theorem 8.6.7 Geometric Series Convergence Criteria. Is there any difference for what we've just done compared to this result that we already know?

In this activity, we see that we can rethink about our Geometric Series family of series as a power series! Then, instead of saying that we have some requirements on the "ratio" for the geometric series to converge, we can say that the power series $g(x) = \sum_{k=0}^{\infty} x^k$ converges for x-values in the interval (–1, 1).

Aside.

If we do this same thing with our other common series, the p-series, then we'd not have a power series, but something slightly different:

$$f(x) = \sum_{k=1}^{\infty} \frac{1}{k^x}$$

This series definitely converges for real x-values in the interval $(1, \infty)$.

This function is called the **Riemann zeta function**, and is hugely important to many different fields of mathematics. We often care about this function when it has complex-number inputs (instead of just real-number inputs). It is also the focal point of one of the most famous unsolved mathematical questions, the Riemann hypothesis.

We noticed in Activity 9.1.8 How Do These Polynomials Converge? that the polynomials built to approximate the natural log function does seem to converge to the function as $n \to \infty$, but only for specific x-values. Hopefully we have some nice ideas as to why that happened: there is a vertical asymptote, and so maybe the "distance" from the center that this polynomial could approximate $\ln(x)$ at is limited!

In general, we can notice that this isn't new: we have families of infinite series that have specific values of variables for which they converge.

A big thing to notice is that these power series, by definition, include exponentials in them (since x^k is a power function of x but is also an exponential function of k). This means that they're great candidates to use the Ratio Test. Since we have a variable x (and

we don't know if this variable is taking on positive or negative values), we'll need to test these series for absolute convergence.

Activity 9.2.3: Some Power Series and their Convergence.

Let's consider a couple of power series and apply some convergence tests to them in order for us to find out how it might converge.

a. Consider the power series:

$$\sum_{k=1}^{\infty} \frac{(x-1)^k}{k^2}$$

In order for us to apply the Ratio Test, we'll actually need to consider the positive-term version:

$$\sum_{k=1}^{\infty} \frac{|x-1|^k}{k^2}$$

Apply the Ratio Test. What do you get in the limit of the ratio between terms?

b. What kind of result from the Ratio Test guarantees convergence for the series? What are the x-values that guarantee convergence?

c. The Ratio Test is inconclusive when the limit is equal to 1. What x-values does this happen at? Consider the power series evaluated at each of these x-values. Do these series converge or diverge?

d. Consider the power series:

$$\sum_{k=1}^{\infty} \frac{(x-1)^k}{\sqrt{k}}$$

Find the interval of x-values for which this series converges and test the end points of the interval in the same way as earlier. If it differs, explain why.

Definition 9.2.4: Interval of Convergence.

For a power series $\sum_{k=0}^{\infty} c_k(x-a)^k$ centered at $x=a$, the interval of x-values for which the power series converges is called the **Interval of Convergence**. The distance from the center to endpoints of the interval is called the **Radius of Convergence**

So this is how we'll think about convergence! When we think about power series and their convergence, we're specifically thinking about convergence for specific inputs. These series are really families of infinite series, and we can try to explain their convergence criteria.

And for a power series, there is always some x–value for which if converges. For the power series

$$f(x) = \sum_{k=0}^{\infty} c_k(x-a)^k$$

centered at $x=a$, then as long as $\{c_k\}$ is a sequence of real numbers, then the series converges at $x=a$. We get $f(a)=c_0$, the constant term. This should match with how we thought about these series originally! They came from polynomial approximations of our functions, where of course the polynomials needed to match the function value at the center. We were thinking about tangent lines, tangent quadratics, tangent cubics, etc. They need to be tangent, and so they "converge" at that single center x -value at least.

Example 9.2.5

For the power series $f(x) = \sum_{k=0}^{\infty} \frac{3^k\,(x^k)}{2k+1}$, find the interval of convergence. Along the way, it will likely be helpful to identify the center and the radius of convergence.

Solution

Let's apply the Ratio Test. We'll technically be applying this to $\sum_{k=0}^{\infty} \frac{3^k|x|^k}{2k+1}$.

$$\begin{aligned}\lim_{k\to\infty} \frac{3^{k+1}|x|^{k+1}}{2(k+1)+1} \cdot \frac{2k+1}{3^k|x|^k} &= \lim_{k\to\infty} \frac{3|x|(2k+1)}{2k+3} \\ &= 3|x|\end{aligned}$$

For us to conclude that this series converges, we need the limit from the ratio test to be less than 1.

$$3|x| < 1 \to |x| < \frac{1}{3}$$

This series is centered at 0 with a radius of convergence of $\frac{1}{3}$. So we know that this series converges for $-\frac{1}{3} < x < \frac{1}{3}$.

Since the Ratio Test is inconclusive when $x = -\frac{1}{3}$ and $x = \frac{1}{3}$ we'll test those individually.

When $x = -\frac{1}{3}$

$$\begin{aligned}&\sum_{k=0}^{\infty} \frac{3^k\left(-\frac{1}{3}\right)^k}{2k+1} \\ &= \sum_{k=0}^{\infty} \frac{(-1)^k}{2k+1}\end{aligned}$$

We can apply the Alternating Series Test!

$$\begin{aligned}&\lim_{k\to\infty} \frac{1}{2k+1} \\ &= 0\end{aligned}$$

This series converges!

When $x = \frac{1}{3}$

$$\begin{aligned}&\sum_{k=0}^{\infty} \frac{3^k\left(\frac{1}{3}\right)^k}{2k+1} \\ &= \sum_{k=0}^{\infty} \frac{1}{2k+1}\end{aligned}$$

We can apply the Rational Comparison Theorem! Since the difference in degrees is 1, we know that this series could be compared to the Harmonic series, and so it diverges.

So the interval of convergence is $\left[-\frac{1}{3}, \frac{1}{3}\right)$.

Operations on Power Series

Theorem 9.2.6: Operations on Power Series.

For two power series $\sum_{k=0}^{\infty} c_k x^k$ and $\sum_{k=0}^{\infty} d_k x^k$ that converge to $f(x)$ and $g(x)$ (respectively) on the interval of convergence I, we can consider the following operations to combine power series.

- Sum: $\sum_{k=0}^{\infty} (c_k + d_k) x^k$ converges to $f(x) + g(x)$ on I.

- Difference: $\sum_{k=0}^{\infty}(c_k - d_k)\,x^k$ converges to $f(x) - g(x)$ on I.
- Product: If bx^n is a power function, then $\sum_{k=0}^{\infty} b\,(c_k)\,x^{k+n}$ converges to $(bx^n)\,f(x)$ on I.
- Composition: If bx^n is a power function, then $\sum_{k=0}^{\infty} c_k\,(b^k)\,(x^{nk})$ converges to $f\,(bx^n)$ when bx^n is in I.

We can do something similar with some of our calculus operations: differentiation and integration.

Theorem 9.2.7: Differentiating and Integrating Power Series.

If $\sum_{k=0}^{\infty} c_k(x-a)^k$ converges to $f(x)$ on an interval of convergence with a radius $R > 0$, then:

- $\sum_{k=0}^{\infty} kc_k(x-a)^{k-1}$ converges to $f'(x)$.
- $\sum_{k=0}^{\infty} \dfrac{c_k(x-a)^{k+1}}{k+1}$ converges to $F(x)$, an antiderivative of $f(x)$.

Both of these converge on an interval of convergence centered at $x = a$ with radius R.

Note 9.2.8

We're being weird about naming the interval of convergence for these. The issue is that when we differentiate, we might lose closed endpoints of an interval. Similarly, when we antidifferentiate, we could add endpoints to an open interval.

We can see this in some of the examples that follow, but the intervals of convergence are going to be identical except for possibly at the endpoints.

Let's finish this here by revisiting the power series from Example 9.2.5.

Example 9.2.9

Let's reconsider the power series $\sum_{k=0}^{\infty} \dfrac{3^k\,(x^k)}{2k+1}$. We know, from Example 9.2.5, that this power series converges to $f(x)$ for x -values in the interval $\left[-\frac{1}{3}, \frac{1}{3}\right)$.

a. Find a power series that converges to $x^7 f(x)$. What is the interval of convergence?
b. Find a power series that converges to $f\,(x^2)$. What is the interval of convergence?
c. Find a power series that converges to $f'(x)$. What is the interval of convergence?

Solution

a.

$$\begin{aligned} f(x) &= \sum_{k=0}^{\infty} \frac{3^k\,(x^k)}{2k+1} \\ x^7 f(x) &= x^7 \sum_{k=0}^{\infty} \frac{3^k\,(x^k)}{2k+1} \\ &= \sum_{k=0}^{\infty} \frac{3^k\,(x^k)\,(x^7)}{2k+1} \\ &= \sum_{k=0}^{\infty} \frac{3^k x^{k+7}}{2k+1} \end{aligned}$$

The interval of convergence doesn't change: $\left[-\frac{1}{3}, \frac{1}{3}\right)$.

b.

$$f(x) = \sum_{k=0}^{\infty} \frac{3^k (x^k)}{2k+1}$$

$$f\left(x^2\right) = \sum_{k=0}^{\infty} \frac{3^k \left(\left(x^2\right)^k\right)}{2k+1}$$

$$= \sum_{k=0}^{\infty} \frac{3^k \left(x^{2k}\right)\left(x^7\right)}{2k+1}$$

The interval of convergence is $\left(-\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}\right)$.

c.

$$f(x) = \sum_{k=0}^{\infty} \frac{3^k (x^k)}{2k+1}$$

$$f'(x) = \frac{d}{dx}\left(\sum_{k=0}^{\infty} \frac{3^k (x^k)}{2k+1}\right)$$

$$= \sum_{k=0}^{\infty} \frac{d}{dx}\left(\frac{3^k (x^k)}{2k+1}\right)$$

$$= \sum_{k=0}^{\infty} \frac{3^k \frac{d}{dx}(x^k)}{2k+1}$$

$$= \sum_{k=0}^{\infty} \frac{k3^k (x^{k-1})}{2k+1}$$

Since k is in the coefficient, starting the index at $k = 0$ means that the "first" term is 0: we can re-index if we'd like, but it is not necessary.

$$\sum_{k=1}^{\infty} \frac{k3^k (x^{k-1})}{2k+1} \quad \text{or} \quad \sum_{k=0}^{\infty} \frac{(k+1)3^{k+1} (x^k)}{2k+3}$$

In any of these cases, the interval of convergence is $\left(-\frac{1}{3}, \frac{1}{3}\right)$.

These operations aren't that useful without a purpose. Similarly, these power series aren't that interesting without knowing what functions they converge to. We'll now start putting some things we've learned together.

- We have ideas of power series representations for $\sin(x)$, $\cos(x)$, and e^x.
- We know the geometric series, which we can use to get a power series representation for $\frac{1}{1-x}$.
- We have a general way of building terms of these power series, using the polynomial approximations and the formula for building those terms.
- We have ways of combining, differentiating, and anti-differentiating power series.

We'll move forward and demonstrate some really great ways of constructing power series representations of some new functions. And then we'll show some very fun uses of these power series representations to illustrate how friendly they are to work with.

Practice Problems

Exercise 9.2.1

Describe the following terms in your own words, and explain how they relate to a power series.

a. Center.
b. Interval of Convergence.
c. Radius of Convergence.

Exercise 9.2.2

Does a power series ever diverge everywhere? Explain why not, using an example.

Exercise 9.2.3

For a power series that converges for $\left|\frac{x-3}{4}\right| < 1$ and diverges elsewhere, find the center of the power series, the radius of convergence, and interval of convergence.

Exercise 9.2.4

For a power series that converges for $|4(x+1)| < 1$ and diverges elsewhere, find the center of the power series, the radius of convergence, and interval of convergence.

Exercise 9.2.5

For each of the following power series, determine the interval of convergence. Be specific on which convergence test you are applying and how you are finding the interval of convergence.

a. $\sum_{k=0}^{\infty} (3x)^k$
b. $\sum_{k=0}^{\infty} \left(\frac{x}{3}\right)^k$
c. $\sum_{k=0}^{\infty} \frac{(x-3)^k}{k!}$
d. $\sum_{k=0}^{\infty} \frac{(x-3)^k}{k}$
e. $\sum_{k=1}^{\infty} \frac{2^k(x-1)^k}{k}$
f. $\sum_{k=1}^{\infty} \left(\frac{x}{k}\right)^k$
g. $\sum_{k=0}^{\infty} (-1)^k(3x-4)^k$
h. $\sum_{k=0}^{\infty} \left(\frac{3x}{5}\right)^{k/2}$
i. $\sum_{k=1}^{\infty} \frac{(x-1)^k 3^k}{(k+1)^k}$

This page titled 9.2: Power Series Convergence is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Peter Keep.

9.3: How to Build Taylor Series

Let's first clear the air: we've been talking about Taylor Polynomials that are partial sums of some larger infinite series. Then we talked about Power Series, a generic structure of an infinite series of power functions of x. We want to combine these together, and name the specific power series that these Polynomial Approximations are partial sums of.

Definition 9.3.1: Taylor Series.

If $f(x)$ is an infinitely differentiable function at $x = a$, then the Taylor series expansion of $f(x)$ centered at $x = a$ is:

$$\sum_{k=0}^{\infty} \frac{f^{(k)}(a)}{k!}(x-a)^k$$

Aside.

If the Taylor series expansion of $f(x)$ is centered at $x = 0$, some people name these specifically, calling them **Maclaurin series**. Colin Maclaurin made good use of Taylor series, and so is credited with having this special case named after him.

Now that we have named these Taylor series, we want to work on constructing these Taylor series for a given function. In this section, we'll investigate a couple of ways that we can do this.

Constructing Directly

We saw in Section 9.1 Polynomial Approximations of Functions that we can use the idea of a polynomial partial sum to build out a Taylor series representation for a function.

We'll need to build a bunch of terms for some Taylor polynomial, and then we can hopefully extrapolate afterwards. We want to come up with a general formula to define these terms, so that we can write the explicit formula in the infinite series, but we know that this, in general, is a difficult task! It's hard to find these explicit formulas for the terms in these series.

It's not impossible, though! You can remind yourself of how we've done this! We found (and speculated about) the Taylor series for $f(x) = e^x$, $f(x) = \sin(x)$, and $f(x) = \cos(x)$ (all centered at $x = 0$) in Activity 9.1.2 Build a Polynomial and Activity 9.1.6 Partial Sums of What?.

Activity 9.3.2: Constructing a Taylor Series Directly.

Let's work on this! To be honest, the goal here is to realize how difficult this can be. Let's consider the function $f(x) = \tan(x)$. We're going to build a Taylor polynomial approximating $f(x)$ centered at $x = 0$.

a. We're going to build a 4th degree polynomial approximating $f(x) = \tan(x)$. In order to do this, we'll need to find the coefficients for the 5 terms.

Table 9.3.3 Coefficients for Polynomial Approximation

k	$f^{(k)}(x)$	◈(◈)(◈)	$\frac{f^{(k)}(a)}{k!}$
$k = 0$	$f(x) = \tan(x)$	$f(0) =$	
$k = 1$	$f'(x) =$	$f'(0) =$	
$k = 2$	$f''(x) =$	$f''(0) =$	

k	$f^{(k)}(x)$	$f^{(k)}(0)$	$\frac{f^{(k)}(a)}{k!}$
$k=3$	$f'''(x)=$	$f'''(0)=$	
$k=4$	$f^{(4)}(x)=$	$f^{(4)}(0)=$	
$k=5$	$f^{(5)}(x)=$	$f^{(5)}(0)=$	

b. Now we can use these coefficients to construct the polynomial! These coefficients should all be on power functions in the form $(x-a)^k$ for $k=0,1,\ldots,5$. These (added together) will form your polynomial, $p_5(x)$.
c. Can you find a general explicit formula for the terms? If you're having a hard time with this, what could you do to try to make this easier?

Why is this difficult? We're going to stop there. This series is famously difficult to work with. We won't generate the whole Taylor series this way, and this strategy really doesn't work well for functions like this (where the derivatives are pretty annoying to find).

Connections To Other Taylor Series

A (hopefully) more useful strategy is to collect the Taylor series that we know, and then try to build a series for a function based on how that function relates. We saw this in Theorem 9.2.6 Operations on Power Series and Theorem 9.2.7 Differentiating and Integrating Power Series.

Activity 9.3.4: Connecting to Another Series.

Let's start with some "known" series.

$$\begin{aligned}
\frac{1}{1-x} &= \sum_{k=0}^{\infty} x^k && \text{on } (-1,1)\\
e^x &= \sum_{k=0}^{\infty} \frac{x^k}{k!} && \text{on } (-\infty,\infty)\\
\sin(x) &= \sum_{k=0}^{\infty} \frac{(-1)^k x^{2k+1}}{(2k+1)!} && \text{on } (-\infty,\infty)\\
\cos(x) &= \sum_{k=0}^{\infty} \frac{(-1)^k x^{2k}}{(2k)!} && \text{on } (-\infty,\infty)
\end{aligned}$$

We're going to build a Taylor series for a logarithmic function:

$$y=\ln(1+x)$$

a. In order to start, we're going to find a Taylor series for $g(x)=\frac{1}{1+x}$. Which of the "known" series above is this most similar to? How is it different?
b. Find a Taylor series for $g(x)=\frac{1}{1+x}$ by using Theorem 9.2.6. Can you connect this series to the known series by multiplying or composing something? **Hint**. Note that $1+x=1-(-x)$.
c. Notice that the function we're trying to get a Taylor series for, $y=\ln(1+x)$, is an antiderivative of $g(x)=\frac{1}{1+x}$. Antidifferentiate the series you found above, using Theorem 9.2.7 Differentiating and Integrating Power Series. We know that at $x=0, y=0$, so use this information to find specific constant that we need when we antidifferentiate.
d. What is the interval of convergence for this new series? How does that relate to the interval of convergence of the "known" series you started with?

Answer available at the end of this section.

We're going to try this again with a different function as the goal!

Activity 9.3.5: Connecting to Yet Another Series.

Here, again, are some updated "known" series

$$\begin{aligned}
\frac{1}{1-x} &= \sum_{k=0}^{\infty} x^k && \text{on } (-1,1)\\
e^x &= \sum_{k=0}^{\infty} \frac{x^k}{k!} && \text{on } (-\infty,\infty)\\
\sin(x) &= \sum_{k=0}^{\infty} \frac{(-1)^k x^{2k+1}}{(2k+1)!} && \text{on } (-\infty,\infty)\\
\cos(x) &= \sum_{k=0}^{\infty} \frac{(-1)^k x^{2k}}{(2k)!} && \text{on } (-\infty,\infty)\\
\ln(1+x) &= \sum_{k=0}^{\infty} \frac{(-1)^k x^{k+1}}{k+1} && \text{on } (-1,1]
\end{aligned}$$

We're going to build a Taylor series for the inverse tangent function:

$$y = \tan^{-1}(x)$$

a. In order to start, we're going to find a Taylor series for $g(x) = \dfrac{1}{1+x^2}$. Which of the "known" series above is this most similar to? How is it different?

b. Find a Taylor series for $g(x) = \dfrac{1}{1+x^2}$ by using Theorem 9.2.6. Can you connect this series to the known series by multiplying or composing something? **Hint**. Note that $1+x^2 = 1-(-x^2)$.

c. Notice that the function we're trying to get a Taylor series for, $y = \tan^{-1}(x)$, is an antiderivative of $g(x) = \dfrac{1}{1+x^2}$. Antidifferentiate the series you found above, using Theorem 9.2.7 Differentiating and Integrating Power Series. We know that at $x = 0, y = 0$, so use this information to find specific constant that we need when we antidifferentiate.

d. What is the interval of convergence for this new series? How does that relate to the interval of convergence of the "known" series you started with?

Answer available at the end of this section.

Now we can update our list of "known" series!

Known Taylor Series.

$$\begin{aligned}
\frac{1}{1-x} &= \sum_{k=0}^{\infty} x^k && \text{on } (-1,1)\\
e^x &= \sum_{k=0}^{\infty} \frac{x^k}{k!} && \text{on } (-\infty,\infty)\\
\sin(x) &= \sum_{k=0}^{\infty} \frac{(-1)^k x^{2k+1}}{(2k+1)!} && \text{on } (-\infty,\infty)\\
\cos(x) &= \sum_{k=0}^{\infty} \frac{(-1)^k x^{2k}}{(2k)!} && \text{on } (-\infty,\infty)\\
\ln(1+x) &= \sum_{k=0}^{\infty} \frac{(-1)^k x^{k+1}}{k+1} && \text{on } (-1,1]\\
\tan^{-1}(x) &= \sum_{k=0}^{\infty} \frac{(-1)^k x^{2k+1}}{2k+1} && \text{on } [-1,1]
\end{aligned}$$

Most importantly, we can use these two activities as a framework for a more general way of constructing Taylor series representations of functions. If we can link the function we're interested in to one of the functions with a "known" Taylor series,

then we can use the link between functions as a way of linking the respective Taylor series.

That is, if we know the Taylor series for $f(x)$ and we want to find the Taylor series for a different function $g(x)$ where $g(x)$ is some sort of transformed version of $f(x)$, then we can list the transformations or changes to $f(x)$ in order to produce $g(x)$, and apply those transformations or changes to the Taylor series for $f(x)$.

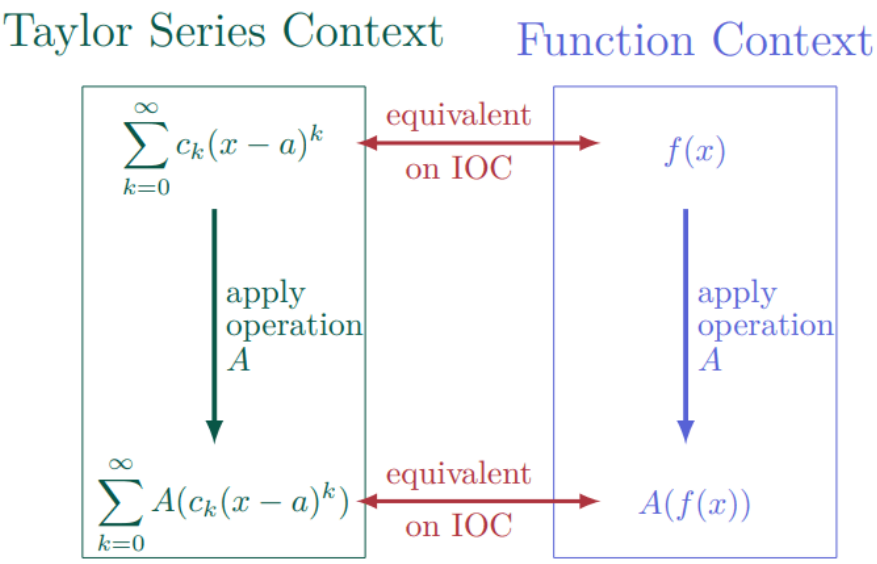


Figure 9.3.6 Constructing a Taylor series from a related series.

So when the operation A is something from either Theorem 9.2.6 Operations on Power Series or Theorem 9.2.7 Differentiating and Integrating Power Series, then we can apply these operations to the Taylor series of $f(x)$ in order to construct the Taylor series of $A(f(x))$. Interestingly enough, we can note that the intervals of convergence can be mostly constructed from these same transformations: there might be some changes to the inclusion of the endpoints based on differentiation and integration.

Now that we have a convenient way of constructing Taylor series for conveniently linked functions, we should think about how we might use these Taylor series representations of functions.

Practice Problems

Exercise 9.3.1

What is the difference between a Taylor Polynomial approximation for $f(x)$ centered at $x = a$ and the Taylor Series for $f(x)$ centered at $x = a$?

Exercise 9.3.2

Given the geometric power series $\sum_{k=0}^{\infty} x^k = \dfrac{1}{1-x}$ for $|x| < 1$, find the Taylor series representations for the following functions. Make sure to explain, for each, how it connects to the geometric Taylor series. Also, report the interval of convergence for each.

a. $\dfrac{x^3}{1-x}$
b. $\dfrac{1}{1+x^4}$
c. $\dfrac{1}{x^2-1} = -\dfrac{1}{1-x^2}$

d. $\dfrac{1}{(1-x)^2}$

Exercise 9.3.3

or each of the following functions, explain how they are related to $f(x) = \sin(x)$. Then, use the connections you describe to find the Taylor Series for the function given by relating it to the Taylor Series for $\sin(x)$.

a. $\sin\left(x^2\right)$
b. $2x\sin(x)$
c. $2x\cos\left(x^2\right)$

Exercise 9.3.4

Find a Taylor Series for each of the following functions. Report the center and interval of convergence for each.

a. $f(x) = xe^x)$
b. $f(x) = e^{\sqrt{x}}$
c. $f(x) = e^{x+2}$
d. $f(x) = \cos(2x)$
e. $f(x) = x^3\tan^{-1}\left(x^2\right)$
f. $f(x) = \sqrt[5]{x^7}\ln(x+1)$

Answers

Activity 9.3.4: Connecting to Another Series.

b. We're going to connect $g(x) = \dfrac{1}{1+x}$ to the geometric Taylor series for $f(x) = \dfrac{1}{1-x}$. We can note that $g(x) = f(-x)$.

c.
$$\begin{aligned} f(-x) &= \sum_{k=0}^{\infty}(-x)^k \\ &= \sum_{k=0}^{\infty}(-1)^k x^k \end{aligned}$$

$$\begin{aligned} G(x) &= \int g(x)dx \\ G(x) &= \int\left(\sum_{k=0}^{\infty}(-1)^k x^k\right)dx \\ &= \sum_{k=0}^{\infty}(-1)^k\int x^k dx \\ &= \sum_{k=0}^{\infty}\frac{(-1)^k x^{k+1}}{k+1} + C \end{aligned}$$

We want $G(0) = 0$, so we need to let $C = 0$. So:

$$\ln(1+x) = \sum_{k=0}^{\infty}\frac{(-1)^k x^{k+1}}{k+1}$$

Sometimes this is written as:

$$\ln(1+x) = \sum_{k=1}^{\infty}\frac{(-1)^{k+1}x^k}{k}$$

d. We can apply the Ratio Test:

$$\begin{aligned} \lim_{k\to\infty}\frac{|x|^{k+2}}{k+2}\cdot\frac{k+1}{|x|^{k+1}} &= \lim_{k\to\infty}\frac{|x|(k+1)}{k+2} \\ &= |x| \end{aligned}$$

So we know that the series will converge for $|x| < 1$. Now we can test the endpoints at $x = -1$ and $x = 1$.
Case: $x = -1$

$$\sum_{k=0}^{\infty}\frac{-1}{k+1}$$

This is the negative harmonic series, and so the series diverges at $x = -1$.

Case: $x = 1$

$$\sum_{k=0}^{\infty} \frac{(-1)^k}{k+1}$$

This is the alternating harmonic series, and so the series converges at $x = -1$.

Activity $9.3.5$**: Connecting to Yet Another Series.**

b. We're going to connect $g(x) = \dfrac{1}{1+x^2}$ to the geometric Taylor series for $f(x) = \dfrac{1}{1-x}$. We can note that $g(x) = f\left(-x^2\right)$.

c.
$$\begin{aligned} f\left(-x^2\right) &= \sum_{k=0}^{\infty} \left(-x^2\right)^k \\ &= \sum_{k=0}^{\infty} (-1)^k x^{2k} \end{aligned}$$

$$\begin{aligned} G(x) &= \int g(x)dx \\ G(x) &= \int \left(\sum_{k=0}^{\infty} (-1)^k x^{2k} \right) dx \\ &= \sum_{k=0}^{\infty} (-1)^k \int x^{2k} dx \\ &= \sum_{k=0}^{\infty} \frac{(-1)^k x^{2k+1}}{2k+1} + C \end{aligned}$$

We want $G(0) = 0$, so we need to let $C = 0$. So:

d.
$$\tan^{-1}(x) = \sum_{k=0}^{\infty} \frac{(-1)^k x^{2k+1}}{2k+1}$$

This page titled 9.3: How to Build Taylor Series is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Peter Keep.

9.4: How to Use Taylor Series

In this final topic for this chapter, we'll pause for a moment to figure out what we could actually use these Taylor series for. This next statement is the personal opinion of the author: I do not typically care to focus on the applications of mathematical topics. In this text, you'll maybe notice that there is less of a focus on applications of derivatives and integrals into other areas and more focus on what these calculus objects actually are.

We'll try to focus on the flexible uses and applications of Taylor series as a way of allowing us the freedom to think about calculus topics in hard-to-work with contexts, or to extend our uses of calculus ideas into new contexts.

Approximations

Taylor series have, as a part of their structure, a built-in approximation strategy. If we want to approximate a series, we can use a partial sum! Approximations of functions using Taylor series, then, are easily approximated using the polynomials that build the partial sums.

We can use this as a way of approximating the functions themselves or approximating function outputs (by evaluating our Taylor series at specific x-values).

Activity 9.4.1: Approximating π and Other Values.

Let's pick a couple of values that are just based on functions with Known Taylor Series and approximate them.

a. If we note that $\tan^{-1}(1) = \frac{\pi}{4}$, then we can say that $\pi = 4\tan^{-1}(1)$. Find an infinite series that converges to $\tan^{-1}(1)$, and use that to construct a series that converges to π.
b. If we note that $\frac{1}{e} = e^{-1}$, then find an infinite series that converges to $\frac{1}{e}$.
c. Find an infinite series that converges to $\sin(1)$.
d. Note that each of these three series are alternating series! We can approximate them using a partial sum, and then we can get an error bound for that partial sum by looking at the size of the next term in the infinite series (Theorem 8.5.7 Approximations of Alternating Series). Approximate the value of each infinite series using a partial sum with the same number of terms. You can pick the number of terms you use. Then, compare the margin of error. Which approximations are most/least accurate? Why do you think that is?

Integrals

We have built a whole host of antidifferentiation strategies in Chapter 7, but there are plenty of functions that are resistant to those specific techniques. In fact, most functions do not have what we call elementary antiderivatives: antiderivative functions that can be written as some combination of the basic, named function types.

A typical example is $\sin\left(x^2\right)$. Let's look at an integral involving this function!

Activity 9.4.2: Integrating using Taylor Series.

Let's consider the definite integral:

$$\int_{x=0}^{x=1} \sin\left(x^2\right)\,dx$$

We know this function is continuous everywhere, and so it is continuous on the interval [0, 1]. So the Fundamental Theorem of Calculus applies. All we need to do is find an antiderivative, evaluate it at the end-points of the interval, and subtract.

Great! Easy!

Except that we can't write out the family of antiderivatives for this function.

So let's convert over to the Taylor series context and solve our problem there.

a. Create a Taylor series for the function $f(x) = \sin(x^2)$. **Hint**. Note that if we let $g(x) = \sin(x)$, then we can write our function as $f(x) = g(x^2)$. This can help serve as a blueprint for creating the Taylor series for $f(x)$.
b. Find the interval of convergence for this Taylor series. Specifically, we want to know if this series converges to $\sin(x^2)$ on the interval [0, 1], since that's the interval we're integrating over. **Hint**. Since the Taylor series for $\sin(x)$ converges when x is in $(-\infty, \infty)$, then we expect that the Taylor series for $\sin(x^2)$ will converge when x^2 is in $(-\infty, \infty)$. What kinds of x-values will produce real numbers when we square them?
c. If $f(x) = \sin(x^2)$, find a Taylor series representation of $F(x)$, an antiderivative of $f(x) = \sin(x^2)$.
d. Evaluate this antidervative at the endpoints and subtract:

$$F(1) - F(0)$$

. **Hint**. Since $F(x) = \sum_{k=0}^{\infty} \frac{(-1)^k x^{4k+3}}{(4k+3)(2k+1)!}$, we're really evaluating this power series at $x = 0$ and $x = 1$.
e. Use Theorem 8.5.7 Approximations of Alternating Series to approximate the value of $\int_{x=0}^{x=1} \sin(x^2)\,dx$ with a maximum error of 10^{-5}. **Hint**. We're really tasked with finding a value of k to make

$$\left|\frac{(-1)^k}{(4k+3)(2k+1)!}\right| < 0.000001$$

or

$$\left|\frac{(-1)^k}{(4k+3)(2k+1)!}\right| < \frac{1}{100000}$$

These terms will get small quickly, due to the factorial. It's probably easiest to just do a little guess-and-check process.

Answer available at the end of this section.

We got lucky here, in that the alternating series was reasonable to approximate with some specified error bound (based on Theorem 8.5.7), but this isn't technically unique to alternating series! There are some other ways of building error bounds on our partial sum approximations, but we won't spend our time building them.

Euler's Formula

Three of the Taylor series that we've looked at have been pretty similar:

$$\begin{aligned} e^x &= \sum_{k=0}^{\infty} \frac{x^k}{k!} \\ \sin(x) &= \sum_{k=0}^{\infty} \frac{(-1)^k x^{2k+1}}{(2k+1)!} \\ \cos(x) &= \sum_{k=0}^{\infty} \frac{(-1)^k x^{2k}}{(2k)!} \end{aligned}$$

It's probably worth stopping for a moment to remember why the exponential function's Taylor series has "all" of the terms, while the ones for the sine and cosine functions have only the odd and even terms (respectively), and also why the terms for the sine and cosine function alternate in sign.

If you want to refresh yourself on why these series have these kinds of structures, take a look back at Section 9.1, where we built the partial sums of these series and conjectured about the infinite series for each.

Before we begin, let's remember a fact about i, the imaginary unit where $i^2 = -1$.

Fact 9.4.3 Powers of i.

For $i = \sqrt{-1}$ where $i^2 = -1$, we can think about the following powers of i:

$i^0 = 1$	$i^4 = 1$	$i^8 = 1$	$i^{4k} = 1$
$i^1 = i$	$i^5 = i$	$i^9 = i$	$i^{4k+1} = i$
$i^2 = -1$	$i^6 = -1$	$i^{10} = -1$	$i^{4k+2} = -1$
$i^3 = -i$	$i^7 = -i$	$i^{11} = -i$	i^{4 k+3}=-i\)
			(where k is some integer)

This structure (flipping back and forth between two things, and changing the sign as we do it) might feel familiar! It almost feels like the same thing that happens with derivatives of the sine and cosine functions, right? Hmm....

Activity 9.4.4: Constructing Euler's Formula.

a. Compose (ix) inside of the e^x function and it's Taylor series to come up with a Taylor series representation for e^{ix}.
b. Write out several terms of this new series. You should write out at least 6 of them, but more is fun, too! Now, use Fact 9.4.3 Powers of i to rewrite these terms.
c. Which terms have are real coefficients? Which terms have imaginary coefficients? If you group these together, do you recognize these terms?
d. Write e^{ix} in terms of sines and cosines.

Answer available at the end of this section.

Theorem 9.4.5: Euler's Formula.

For i, the imaginary unit,

$$e^{ix} = \cos(x) + i\sin(x)$$

This is a really powerful theorem for a couple of reasons, but we'll just touch on a few of them. The first, and maybe least important, is a really cool identity, called **Euler's Identity**. It's just Euler's Formula evaluated at $x = \pi$:

$$\begin{aligned} e^{i\pi} &= \cos(\pi) + i\sin(\pi) \\ e^{i\pi} &= -1 \\ e^{i\pi} + 1 &= 0 \end{aligned}$$

This last statement is the one most used. It's often cited as one of the most "beautiful" or interesting equations in mathematics, since it combines 5 of the most important numbers:

- e, the natural exponential base, $e = \lim_{n\to\infty}\left(1 + \frac{1}{n}\right)^n$.
- i, the imaginary unit where $i = \sqrt{-1}$ and $i^2 = -1$.
- π, the constant ratio between the circumference and diameter of every circle.
- 1, the multiplicative identity, where $1(x) = x$.
- 0, the additive identity, where $0 + x = x$.

These 5 numbers are all combined into one connected equation, even though they are seemingly not connected at all in their "origins". It's no wonder so many people love this equation!

Aside.

As a fun scavenger hunt, see if you can find the faculty in the mathematics department at Moraine Valley Community College that has a tattoo of this identity!

Complex Analysis

Euler's Formula is a really nice way of linking some of our well-known functions to their complex-function extensions. We have complex exponentials, and we can use this same formula to create a complex-valued logarithmic function.

From here, mathematicians were able to think about how the concepts of calculus could be applied to functions that took in complex-valued inputs and outputs complex-valued outputs. This area of mathematics proved to be rich in applications and also hugely interesting.

Trigonometric Identities

Trigonometric identities are used in a ton of different areas of mathematics (including in this textbook) to try to rewrite some expressions into friendlier forms. One of the big struggles for mathematics students, though, is trying to remember them when needed. The author of this textbook thinks, as a matter of personal opinion, that memorizing facts and identities like this is overrated in general, but being able to construct these when necessary is more helpful.

A problem with many trigonometric identities is that the construction or explanation of them often relies on tricky geometry involving lots of triangles.

Euler's Formula provides a really nice way of thinking of these identities, since we can construct them from the definition of the complex exponential.

Example 9.4.6

Let's construct the sum of angle identites for sine and cosine. We, eventually, want to have some formula for $\sin(\alpha+\beta)$ and $\cos(\alpha+\beta)$. In order to do this, let's evaluate e^{ix} at $x=\alpha+\beta$ and see what happens.

$$e^{i(\alpha+\beta)} = \cos(\alpha+\beta) + i\sin(\alpha+\beta)$$

So we can see that the "real" part of this is the sum of angles for the cosine function, and the "imaginary" part is the sum of angles for the sine function.

But now we can think about $e^{i(\alpha+\beta)}$ by considering that a sum of exponents could really be a product of exponentials:

$$e^{i(\alpha+\beta)} = e^{i\alpha}e^{i\beta}$$

Let's look at this second version.

$$\begin{aligned} e^{i\alpha}e^{i\beta} &= (\cos(\alpha)+i\sin(\alpha))(\cos(\beta)+i\sin(\beta)) \\ &= \cos(\alpha)\cos(\beta) + i\sin(\alpha)\cos(\beta) + i\cos(\alpha)\sin(\beta) - \sin(\alpha)\sin(\beta) \\ &= (\cos(\alpha)\cos(\beta) - \sin(\alpha)\sin(\beta)) + i(\sin(\alpha)\cos(\beta) + \cos(\alpha)\sin(\beta)) \end{aligned}$$

Now we can remember that this should be equivalent to $\cos(\alpha+\beta)+i\sin(\alpha+\beta)$, from earlier:

$$\cos(\alpha+\beta) + i\sin(\alpha+\beta) = (\cos(\alpha)\cos(\beta) - \sin(\alpha)\sin(\beta)) + i(\sin(\alpha)\cos(\beta) + \cos(\alpha)\sin(\beta))$$

So we can equate the real parts and the imaginary parts separately to get:

$$\cos(\alpha+\beta)=\cos(\alpha)\cos(\beta)-\sin(\alpha)\sin(\beta)$$
$$\sin(\alpha+\beta)=\sin(\alpha)\cos(\beta)+\cos(\alpha)\sin(\beta)$$

This strategy is super useful for creating these identities, since we can often take advantage of exponent properties to write two different equivalent expressions to connect.

These are only a few applications of Taylor series, but hopefully they are enough to show a sufficiently wide range of applications or uses. Taylor series, whether they're used to get polynomial approximations of objects or to translate some function into a different context, are some of the most used applications of calculus, and are a wonderful way to think about connections between some of the calculus concepts that we've talked about.

Practice Problems

Exercise 9.4.1

We can define hyperbolic trigonometric functions in the following way:

$$\sinh(x)=\frac{e^x-e^{-x}}{2}$$
$$\cosh(x)=\frac{e^x+e^{-x}}{2}$$

a. Find a Taylor Series representation for $\sinh(x)$ centered at $x=0$.
b. Find a Taylor Series representation for $\cosh(x)$ centered at $x=0$.
c. Show that $\sinh(x)+\cosh(x)=e^x$ using the definition of the hyperbolic trigonometric functions above.
d. Show that $\sinh(x)+\cosh(x)=e^x$ using the Taylor Series.

Exercise 9.4.2

Use a Taylor series to estimate the value of the integral

$$\int_{x=0}^{x=1} e^{-x^2}dx$$

within a maximum error margin of 0.001.

Answers

Activity 9.4.2: Integrating using Taylor Series.

a.
$$\begin{aligned} g(x)&=\sin(x)\\ &=\sum_{k=0}^{\infty}\frac{(-1)^k x^{2k+1}}{(2k+1)!}\\ f(x)&=g\left(x^2\right)\\ &=\sum_{k=0}^{\infty}\frac{(-1)^k\left(x^2\right)^{2k+1}}{(2k+1)!}\\ &=\sum_{k=0}^{\infty}\frac{(-1)^k x^{4k+2}}{(2k+1)!}\end{aligned}$$

b. The interval of convergence is $(-\infty,\infty)$, so we should be comfortable using this series as a replacement for $\sin\left(x^2\right)$ in the integral we're working on.
c. We can find $F(x)$ by applying the Power Rule for Antiderivatives.

$$\begin{aligned} F(x)&=\sum_{k=0}^{\infty}\frac{(-1)^k\left(\frac{x^{4k+3}}{4k+3}\right)}{(2k+1)!}\\ &=\sum_{k=0}^{\infty}\frac{(-1)^k x^{4k+3}}{(4k+3)(2k+1)!}\end{aligned}$$

Note that we don't need to worry about any constant of integration, since we are simply choosing *one* of the antiderivatives.
d. First, we evaluate at $x=1$:

$$\begin{aligned} F(1) &= \sum_{k=0}^{\infty} \frac{(-1)^k (1)^{4k+3}}{(4k+3)(2k+1)!} \\ &= \sum_{k=0}^{\infty} \frac{(-1)^k}{(4k+3)(2k+1)!} \end{aligned}$$

Now we evaluate at $x = 0$:

$$\begin{aligned} F(0) &= \sum_{k=0}^{\infty} \frac{(-1)^k (0)^{4k+3}}{(4k+3)(2k+1)!} \\ &= \sum_{k=0}^{\infty} (0) \\ &= 0 \end{aligned}$$

So we can notice that $F(1) - F(0) = F(1)$.

$$\int_{x=0}^{x=1} \sin\left(x^2\right) dx = \sum_{k=0}^{\infty} \frac{(-1)^k}{(4k+3)(2k+1)!}$$

e. When $k = 4$ we get

$$\begin{aligned} \left| \frac{(-1)^k}{(4k+3)(2k+1)!} \right| &= \frac{1}{(4(4)+3)(2(4)+1)!} \\ &= \frac{1}{19(9!)} \\ &= \frac{1}{6894720} \end{aligned}$$

Since this is smaller than our margin of error, we know that the actual value of the integral is bounded between S_3 and S_4.

$$\begin{aligned} S_3 &= \sum_{k=0}^{3} \frac{(-1)^k}{(4k+3)(2k+1)!} \\ &= \frac{258019}{831600} \\ &\approx 0.31026816 \\ S_4 &= \sum_{k=0}^{4} \frac{(-1)^k}{(4k+3)(2k+1)!} \\ &= \frac{117656719}{379209600} \\ &\approx 0.31026830 \end{aligned}$$

So we know that the integral is some number in between these! We can actually round this value to six decimal places and be pretty confident in this estimation!

$$\int_{x=0}^{x=1} \sin\left(x^2\right) dx \approx 0.310268$$

Activity 9.4.4**: Constructing Euler's Formula.**

a. $e^{ix} = \sum_{k=0}^{\infty} \frac{(ix)^k}{k!}$

$$e^{ix} = \sum_{k=0}^{\infty} \frac{(ix)^k}{k!}$$

b.

$$\begin{aligned} &= i + ix + \frac{i^2x^2}{2!} + \frac{i^3x^3}{3!} + \frac{i^4x^4}{4!} + \frac{i^5x^5}{5!} + \frac{i^6x^6}{6!} + \frac{i^6x^6}{6!} + \frac{i^7x^7}{7!} + \ldots \\ &= 1 + ix - \frac{x^2}{2!} - \frac{ix^3}{3!} + \frac{x^4}{4!} + \frac{ix^5}{5!} - \frac{x^6}{6!} - \frac{ix^7}{7!} + \ldots \end{aligned}$$

c.

d. $e^{ix} = \cos(x) + i\sin(x)$

This page titled 9.4: How to Use Taylor Series is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Peter Keep.

9.5: Exploration- Power Series

We're going to use Taylor series a bit more here, and we'll look at two interesting problems: one a historical relic, and the other a modern statistical result.

The Basel Problem

The Basel Problem is the name given to a problem about infinite series, solved by Leonard Euler.

Theorem 9.5.1: Basel Problem.

The sum of reciprocal squares converges, and specifically:

$$\sum_{k=1}^{\infty} \frac{1}{k^2} = \frac{\pi^2}{6}$$

This result was incredible at the time of its proof for two reasons:

1. This was a longstanding question in mathematics! Mathematicians knew that this infinite series converged, but were stumped about what the value could be.
2. The proof of the result actually proved what a whole bunch of different p-series converged to.

Let's follow along with Euler's proof of this result!

Exploration 9.5.2: The Basel Problem.

We're going to follow along this proof, and for most of the time it might seem like we're fiddling with unrelated results and functions. By the end, we'll tie what we're doing together to show that:

$$\sum_{k=1}^{\infty} \frac{1}{k^2} = \frac{\pi^2}{6}$$

a. First, consider the function $f(x) = \begin{cases} \frac{\sin(x)}{x} & \text{for } x \neq 0 \\ 1 & \text{for } x = 0 \end{cases}$. Find a Taylor series representation for this function. **Hint**. Try using the fact that:

$$\begin{aligned} \sin(x) &= \sum_{k=1}^{\infty} \frac{(-1)^k x^{2k+1}}{(2k+1)!} \\ &= x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + - \ldots \end{aligned}$$

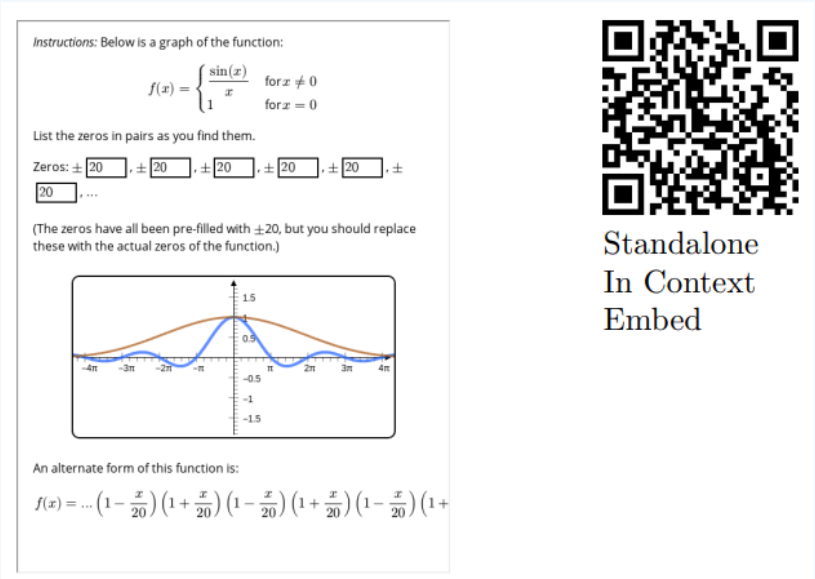
Can you divide by x?

b. We're going to consider this function again, but in a different form. We're going to think about this infinite polynomial, but we'll try to write it by thinking about the factors that we expect to see, specifically when we think of the zeros of the function. So first, note that the zeros of this function, $\frac{\sin(x)}{x}$ must occur when $\sin(x) = 0$. What are the x-values that make $\sin(x) = 0$?

c. We'll use these zeros to write out the function's factors, but we're also going to make use of the fact that $f(0) = 1$. This means that if $x = c$ is a zero of the function, then the corresponding factor should be written as $\left(1 - \frac{x}{c}\right)$. So, if $c_1, c_2, c_3, c_4, \ldots$ are all zeros, then we should write the function as:

$$f(x) = \left(1 - \frac{x}{c_1}\right)\left(1 - \frac{x}{c_2}\right)\left(1 - \frac{x}{c_3}\right)\left(1 - \frac{x}{c_4}\right)\ldots$$

Use the zeros of the $\frac{\sin(x)}{x}$ function that you found earlier to write out the factors of this.



d. Note that we have pairs of factors that are in the form $\left(1-\frac{x}{c}\right)\left(1+\frac{x}{c}\right)$. This is hopefully a very recognizable factoring pattern! It's the difference of squares! What do you get when you multiply these? This should allow you to write your factored function from above in a slightly different way, by combining the pairs of factors using this multiplication. What does this look like?

e. We are almost complete. We now have represented the function

$$f(x)\begin{cases} \frac{\sin(x)}{x} & \text{for } x \neq 0 \\ 1 & \text{for } x = 0 \end{cases}$$

as an infinite sum

$$1-\frac{x^2}{3!}+\frac{x^4}{5!}-\frac{x^6}{7!}+-...$$

and also as an infinite product

$$\left(1-\frac{x^2}{\pi^2}\right)\left(1-\frac{x^2}{4\pi^2}\right)\left(1-\frac{x^2}{9\pi^2}\right)...$$

We can note that the constant term is 1 in each, since it is the product of all of the constant terms in the infinite product. We know that the quadratic terms should be $-\frac{x^2}{6}$ (from the infinite sum). Can you find what the quadratic term from the infinite product would be? **Hint**. This should be the quadratic term in each factor multiplied by the constant term in each of the rest of the factors, summed together.

f. You should have shown that:

$$-\frac{x^2}{6}=-\frac{x^2}{\pi^2}-\frac{x^2}{4\pi^2}-\frac{x^2}{9\pi^2}-...$$

What is the sum of $)1+\frac{1}{4}+\frac{1}{9}+...\frac{1}{k^2}+...$?

g. Can you use this same strategy to find our what the sum $\sum_{k=1}^{\infty}\frac{1}{k^4}$ converges to? Note that we're really asking what the degree 4 terms are in each representation of the function $f(x)$.

The normal distribution (what we typically think of when we imagine a "bell curve") is very important in the study of probability and statistics.

Definition 9.5.3: Standard Normal Distribution.

We say that a random variable follows the standard normal distribution when the mean or expected value of the random variable is 0, the variance of the random variable is 0, and the probability density function is:

$$f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$$

There are a lot of technical details behind some of these terms, but we won't concern ourselves with it too much. For now, we'll state two facts:

1. We use an integral to represent probability. For instance, the probability that a random variable will take on a value that is up to 1 standard deviation above the mean is:

$$\frac{1}{\sqrt{2\pi}} \int_{x=0}^{x=1} e^{-\frac{x^2}{2}} dx$$

   We can change the limits of integration to change the range of values that we are concerned with.
2. The **Empirical Rule**, or the **68-95-99.7** rule states that:
   - About 68% of the values will be within 1 standard deviation of the mean (between $x = -1$ and $x = 1$).
   - About 95% of the values will be within 2 standard deviations of the mean (between $x = -2$ and $x = 2$).
   - Almost all (about 99.7%) of the values will be within 3 standard deviations of the mean (between $x = -3$ and $x = 3$).

Exercise 9.5.4: Normal Distribution.

a. Write an integral that represents the probability of a random variable with a normal distribution taking on a value up to 2 standard deviations above the mean.
b. Based on the Empirical Rule, what do you expect the value of this integral to be?
c. Construct a Taylor series representing the function $e^{-\frac{x^2}{2}}$.
d. In order to estimate your integral, we will need to antidifferentiate $e^{-\frac{x^2}{2}}$. This function does not have an **elementary antiderivative**, meaning we cannot express an antiderivative in terms of the traditional functions we have named and the traditional operations we have defined. Antidifferentiate your Taylor series representation of $e^{-\frac{x^2}{2}}$, and call this $F(x)$.
e. Evaluate your integral using this new representation an antiderivative, $F(x)$.
f. This is an alternating series, and so we can approximate this value to whatever degree of accuracy we'd like using Theorem 8.5.7 Approximations of Alternating Series. How many terms do you need, in order to approximate this probability within 0.01?
g. Construct, evaluate, and approximate the integral representing the probability that a normally distributed random variable takes on a value between 1 and 1.5 standard deviations above the mean (using the same number of terms as above).

This page titled 9.5: Exploration- Power Series is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Peter Keep.

10: PARAMETRIC AND POLAR CURVES

10.1: Parametric Curves
10.2: Polar Curves
10.3: Calculus with Parametric Curves
10.4: Calculus with Polar Curves

This page titled 10: Parametric and Polar Curves is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Peter Keep.

10.1: Parametric Curves

Let's start with a small deviation from calculus to introduce (or re-introduce) some ideas of how we might define functions and curves. We have been working with functions and graphs of functions for so long, but that might mean that some of the ideas are so engrained that might not even think about them. Let's remind ourselves as much as possible.

Activity 10.1.1: A Refresher on Functions.

This is mostly a discussion activity. There's not a lot to do, other than think and share your thoughts with others. If you're not in a classroom setting, this can be hard! Find someone else with a relatively similar math background, and talk about this with them!

a. When you think of a function, what do you think of? What do you imagine in your mind when you hear the word "function" (in a mathematical context)? Is it based on a definition, a visualization, an example, or some other more loosely connected set of phrases?
b. What do you remember about variables? For instance, what is their purpose, their relationship, or their use in defining or visualizing a function?
c. How do you visualize a function? What does the picture represent, or how would you explain it to others? What are the important aspects of the visualization that we might use to control what these pictures look like, or what do we try to keep consistent when we visualize these functions? Why do we do this?
d. How might you define a "point on the function?"

None of this should be new to you, but it probably has been some time since you've thought about it! The reason we're starting this chapter (and this section) with this little reminder is simple: we're going to think about how we could slightly alter some of these basic concepts and change the way that we think about functions, curves, and graphs.

What If...?

We're going to try to change some of the assumptions about curves and graphs and functions by asking some "What if...?" questions. For this section, we'll ask a very short one:

What if, instead of the y variable being an output that depends on the x variable input, we had both x and y as outputs of some third variable, like time?

Instead of thinking of y as a function output of $x, y = f(x)$, we'll think about the vertical component y and the horizontal component x of our curve being defined as functions of t. Now, instead of x being an independent variable and y being dependent on the value of x, we have that x and y are independent of each other, both dependent on t, which we often call a parameter. Typically, we think of this parameter as time: x and y both act independently, but change as time marches forward.

Definition 10.1.2: Parametric Curve

A curve is defined parametrically if

$$x = f(t)$$
$$y = g(t)$$

for values of t in some interval.

Aside.

It's kind of like an etch-a-sketch toy, where you move the two dials independently to trace out some picture! One dial represents the horizontal position while the other represents the vertical position. If you turn both dials at once, you'll create some sort of curve!

This small change in perspective (we're changing how we view or think of the relationship between variables) impacts a lot of what we assumed and recalled in Activity 10.1.1.

Exploring Parametric Curves

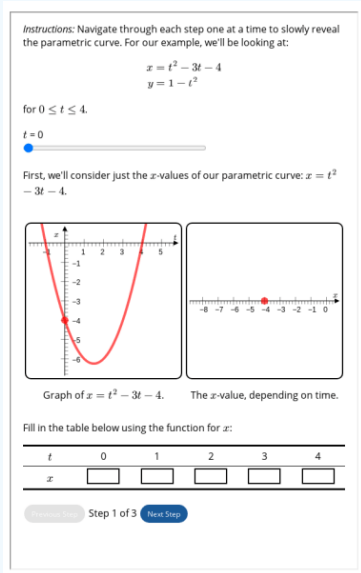
Let's begin with an actual example and see how the parametric definitions can build the curve.

Activity 10.1.3: Our First Parametric Curve.

We'll begin by exploring the parametrically defined curve:

$$x = t^2 - 3t - 4$$
$$y = 1 - t^2$$

for $0 \leq t \leq 4$. For each task, we'll use the following visualization:



Standalone
In Context
Embed

a. In Step 1, how does the graph of $x = t^2 - 3t - 4$ give us information about the horizontal movement of our parametric curve?
b. Fill in the table with the x-values at different values of t. Describe the horizontal motion of the parametric curve using this information.
c. In Step 2, how does the graph of $y = 1 - t^2$ give us information about the vertical movement of our parametric curve?
d. Fill in the table with the y-values at different values of t. Describe the vertical motion of the parametric curve using this information.
e. In Step 3, confirm that the point follows the path you constructed.

This small interactive graph is a nice one, because it highlights one of the important aspects of a parametric curve: motion. A graph of a parametric curve is really a path that a point travels along, which is pretty different from our original interpretation of a graph—the collection of points (x, y) where $y = f(x)$.

Definition 10.1.4: Orientation.

The direction for which a parametric curve is traced when the value of the parameter is increasing is called the **orientation** of the curve.

We often try to include some visual clues depicting the orientation or direction of the curve when we visualize it. Often, this can be achieved with a simple animation (where we can see a point move along the curve), but this is also achieved using arrows along a curve.

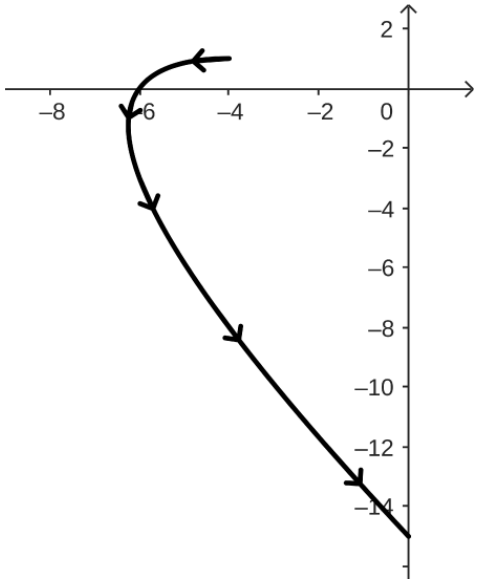


Figure 10.1.5 Orientation of a parametric curve: $\begin{aligned} x &= t^2 - 3t - 4 \\ y &= 1 - t^2 \end{aligned}$ for $0 \leq t \leq 4$

More Parametric Curves

We could end our section with a list of facts or formulas for different curves that we might recognize, but it will be more fun to explore these and discover them ourselves.

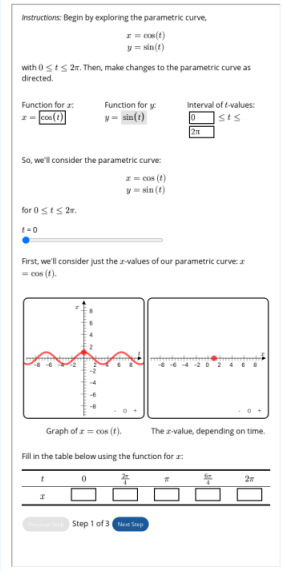
Activity 10.1.6: Explore Some Curves.

We'll start off with a family of parametric curves, where we can change the functions and hopefully predict some of the curve's behavior.

a. Let's start by thinking about the parametric curve:

$$\begin{aligned} x &= \cos(t) \\ y &= \sin(t) \end{aligned}$$

for $0 \leq t \leq 2\pi$. Before you visualize any of this, think about the behavior of the x-values and y-values. What do you think they'll do? What will this curve look like? Visualize the movement of the x-values, the y-values, and the complete parametric curve using the graphing tool below.



Standalone
In Context
Embed

b. What do you think would happen when we multiply the function for x by a factor of 3? In the graphing tool above, change the function for x to:

$$x = 3\cos(t)$$

Explain the changes you're seeing.

c. What do you think would happen when we multiply the input of the cosine function by 3? In the graphing tool above, change the function for x to:

$$x = \cos(3t)$$

Explain the changes you're seeing.

d. Explore the parametric curve:

$$x = \cos(3t)$$
$$y = \sin(2t)$$

for $0 \le t \le 2\pi$. Explain to yourself why the curve looks the way that it does.

e. Create a parametric curve for a circle, but reverse the orientation.

f. Create a parametric curve for a circle centered at the point (2, –1) with radius 3.

g. What do you think the following parametric curve will look like:

$$x = \cos^3(t)$$
$$y = \sin(t)$$

for $0 \le t \le 2\pi$?

Eliminating the Parameter

It is possible that, in some case, we might want to represent our curve in a more traditional "y as a function of x" or "x as a function of y" setting. In this case, we need to manipulate our functions for x and y in order to eliminate the parameter, t, from our function representation.

Typically, this can be done by isolating t in either the function definition for x or y, and then substituting this new expression for t into the other variable's function.

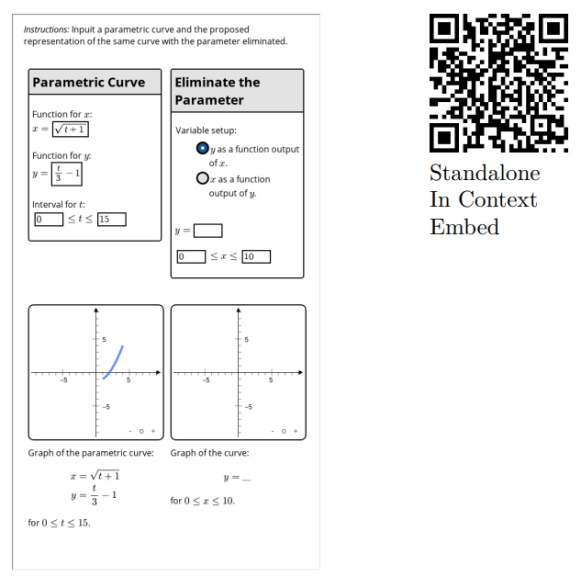
Exercise 10.1.7: Eliminate the Parameter.

Consider the parametric curve

$$x = \sqrt{t+1}$$
$$y = \frac{t}{3} + 1$$

for $0 \le t \le 15$.

a. Solve $x = \sqrt{t+1}$ for t.
b. Substitute this expression for t into $y = \frac{t}{3} - 1$.
c. Start over, and instead solve $y = \frac{t}{3} - 1$ for t.
d. Now substitute this expression for t into $x = \sqrt{t+1}$.
e. Use the graphing tool below to check to see if these curves are all the same. You'll need to find appropriate intervals of x-values or y-values along the way.



Standalone
In Context
Embed

Answer available at the end of this section.

Activity 10.1.8: Eliminate More Parameters.

For each parametrically-defined curve, eliminate the parameter to create a function $y = f(x)$ or $x = f(y)$. Feel free to use the interactive graphing tool in Activity 10.1.7.

a. $x = t^2 - 1$, $y = (3t+1)^3$ for $0 \le t \le 2$.
b. $x = \ln(t+1)$, $y = \sin(t)$ for $0 \le t$.
c. $x = \cos^2(t)$, $y = (t+1)^{3/2}$ for $0 \le t$.
d. $x = \tan^{-1}(x)$, $y = t^2$ for $-2 \le t \le 2$.

We might notice that this strategy depends on us being able to solve for t in at least one of these functions. For instance, the following parametric curve will be hard to employ this strategy:

$$x = t - \sin(t)$$
$$y = t\cos(t)$$

There are no "nice" ways of solving for t in either of these functions!

This is everything we need for now. We're going to make one more quick stop to check out another way of representing functions, and then we'll spend our time thinking about how we might do calculus with these parametric curves.

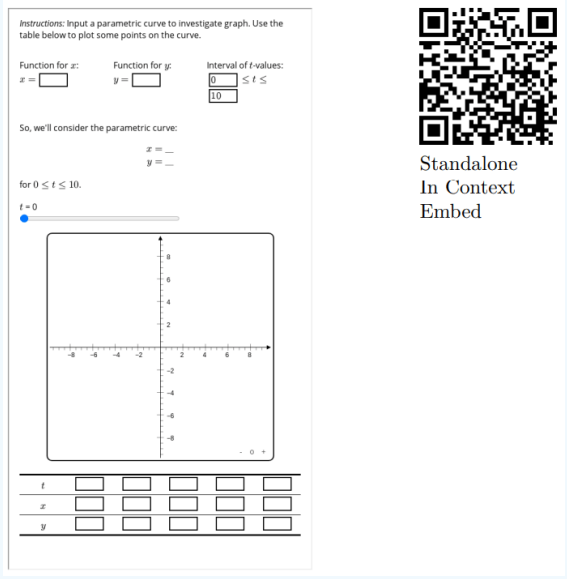
Practice Problems

Exercise 10.1.1

Explain how a curve written in parametric form is different than one written in standard function notation.

Exercise 10.1.2

For each of the following parametric curves, plot the curves in the graphing utility below. Then, find at least 5 different points on the curve, plot them, and describe the curve.



a. $\begin{aligned} x &= t-4 \\ y &= 3+t-t^2 \end{aligned}$ for $0 \le t \le 10$.

b. $\begin{aligned} x &= t+3 \\ y &= 5-2t \end{aligned}$ for $0 \le t \le 5$.

c. $\begin{aligned} x &= 5\cos^2(t) \\ y &= 5\sin^2(t) \end{aligned}$ for $0 \le t \le 2\pi$.

d. $\begin{aligned} x &= \sqrt{t+4} \\ y &= 5t+1 \end{aligned}$ for $0 \le t \le 5$.

Exercise 10.1.3

Eliminate the parameter, t, in each of the parametric curves, expressing the curve either as a function $y = f(x)$ or $x = f(y)$.

a. $\begin{aligned} x &= t-4 \\ y &= 3+t-t^2 \end{aligned}$ for $0 \le t \le 10$.

b. $\begin{aligned} x &= t+3 \\ y &= 5-2t \end{aligned}$ for $0 \le t \le 5$.

c. $\begin{aligned} x &= 5\cos^2(t) \\ y &= 5\sin^2(t) \end{aligned}$ for $0 \le t \le 2\pi$.

d. $\begin{aligned} x &= \sqrt{t+4} \\ y &= 5t+1 \end{aligned}$ for $0 \le t \le 5$.

Exercise 10.1.4

In Activity 10.1.6 Explore Some Curves, we saw the we can define a circle parametrically by

$\begin{aligned} x &= \cos(t) \\ y &= \sin(t) \end{aligned}$ for $0 \le t \le 2\pi$.

a. Remind yourself why this parametric definition makes sense, based on how we think about trig functions on a unit circle.
b. Eliminate the parameter to create an equation with x and y related to each other.
c. How does the function $f(t) = \cos(5t)$ compare to the function $g(t) = \cos(t)$? How about $\ell(t) = \sin(5t)$ compared to the function $m(t) = \sin(t)$?
d. The parametric curve

$$\begin{aligned} x &= \cos(5t) \\ y &= \sin(5t) \end{aligned}$$

for $0 \le t \le 2\pi$ looks exactly the same as the unit circle we defined earlier. How is it different? What does the coefficient on t do?
e. Create a parametric curve that looks like the unit circle but has an opposite orientation (direction).

Answers

Exercise 10.1.7**: Eliminate the Parameter.**

a. $t = x^2 - 1$

b.
$$\begin{aligned} y &= \frac{t}{3} - 1 \\ &= \frac{x^2-1}{3} - 1 \\ &= \frac{x^2}{3} - \frac{4}{3} \end{aligned}$$

c. $t = 3y + 3$

d.
$$\begin{aligned} x &= \sqrt{t+1} \\ &= \sqrt{(3y+3)+1} \\ &= \sqrt{3y+4} \end{aligned}$$

This page titled 10.1: Parametric Curves is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Peter Keep.

10.2: Polar Curves

It's time for another small deviation to think about functions, again. This time, we're going to specifically think about graphs of functions and how we visualize things.

Activity 10.2.1: A Refresher on Graphing Functions.

This is, again, mostly a discussion activity. There's not much to do, but you should think about these questions and share your thoughts with others.

a. How do we plot a typical function (in the form $y = f(x)$) on a coordinate grid? How do you think about the different variables, and how do we use them to plot a point?
b. How do we represent the coordinates of a point? What does each component of a point tell us about its location?
c. Does the "grid" that we use match with the way you might think about locations of things in your life? Where else do humans use a grid like this to describe locations?

We'll move forward in a familiar way. Just like in Section 10.1, we'll ask a "What if...?" question to alter some basic ideas about how we graph things, and think about functions in a slightly different way.

What If...?

What if, instead of x and y representing the horizontal and vertical locations of our point, we thought about a point's location using a direction and a distance?

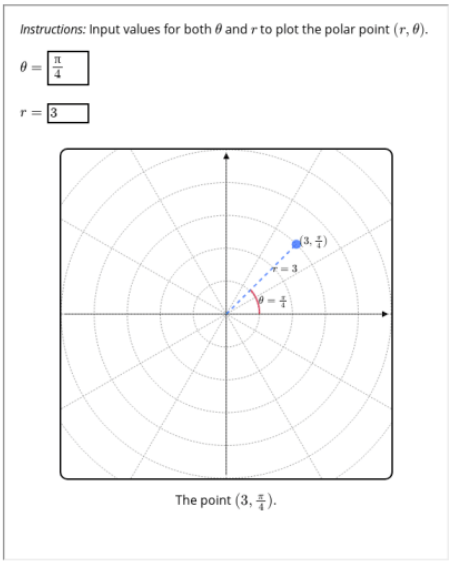
Instead of thinking about a point, (x, y), as the location described on a rectangular grid, we might think about a point described on a circular grid: the coordinates might describe a direction (an angle, measured in radians) and a distance from the origin.

This isn't a new way of describing locations! While we might, in a city like Chicago, describe the location of something using rectangular coordinates ("My favorite sushi place is 7 blocks north and 3 block east of here."), this isn't the only way to describe a location. In the middle of a forest, for instance, these directions wouldn't be helpful. Instead, we might point someone in the right direction and tell them how far away something is ("There's a nice spot to forage for mushrooms a half a mile to the southwest of here.").

Definition 10.2.3: Polar Coordinates.

A point (r, θ) is written in the polar coordinate system, where r represents the distance between the point and the origin (called the pole) and θ represents the angle measured to the positive horizontal axis and a ray from the pole passing through the point.

We can practice visualizing these points using the graphing utility below.



Standalone
In Context
Embed

We will mostly be interested in thinking about functions defined using this new coordinate system.

Definition 10.2.3: Polar Curve.

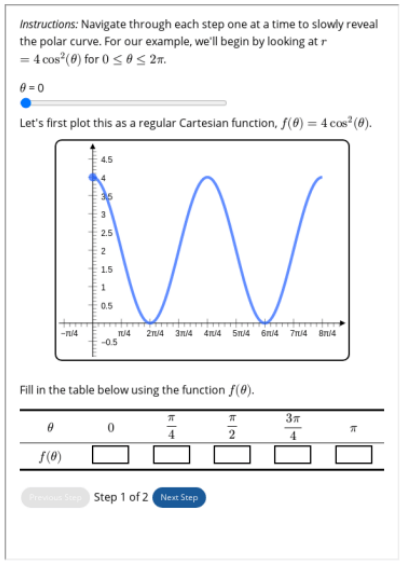
A **polar curve** is a curve defined by the function $r = f(\theta)$, where θ is the input/independent variable and r is the output/dependent variable.

Exploring Polar Coordinates

We'll start with a small example to try to understand more about how we can visualize polar curves.

Activity 10.2.4: First Polar Curve.

We'll begin by exploring the polar curve, $r = 4\cos^2(\theta)$ for $0 \le \theta \le 2\pi$. For each task, we'll use the following visualization:



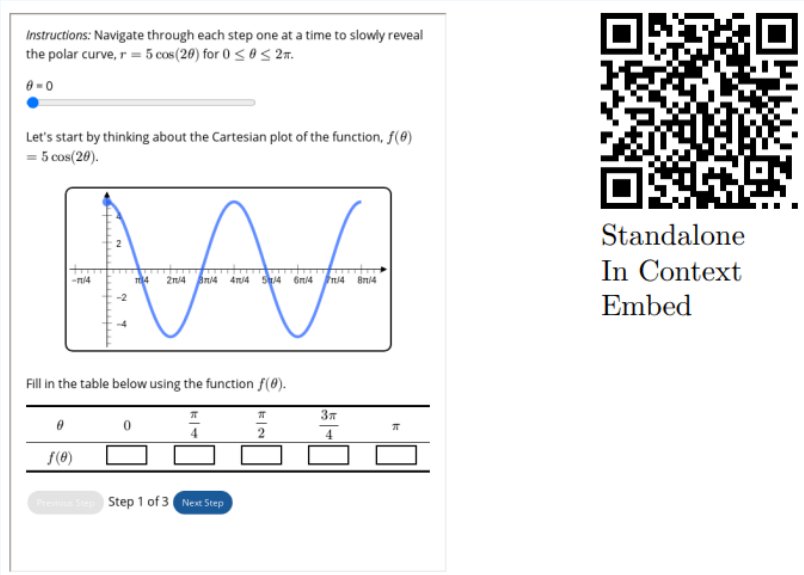
Standalone
In Context
Embed

a. In Step 1, explain the shape and behavior of the function based on your knowledge of trigonometric functions. How do you know, based on the function being specifically $4\cos^2(\theta)$ that it will look like this? **Hint**. Why does this function oscillate between 0 and 4? Why did you know where the zeros would be?
b. Fill in the table with the function outputs $f(\theta)$ for different values of θ.
c. In Step 2, why do the points that we filled in end up where they are on the polar curve? Explain how we use the coordinates you input in the previous step to plot these polar points.
d. Explain why the polar point has this shape/behavior. Refer back to explanation of the Cartesian curve in Step 1, and explain how these graphical features manifest themselves in the polar curve. **Hint**. How do we visualize the oscillation between 0 and 4 in this context? What is happening at the zeros of the function?
e. Fill in the table with the r outputs at different values of θ. Describe why these points are plotted where they are, including the quadrants and distances from the origin/pole.
f. Why do the two points (4, 0) and $(4, 2\pi)$ seem to overlap in the graph? Why does it look like this is only a single point?

One thing to note, and focus on more closely, is our use of the word "distance" when we defined r. Our typical understanding of distance is that this is a non-negative measurement. In Activity 10.2.4 , we looked at a non-negative function, so we didn't have to consider what might happen when our function has negative outputs. In order to give ourselves more flexibility in this coordinate system, let's consider a new function—one with both positive and negative outputs.

Activity 10.2.5: A Polar Flower.

We'll now explore a polar flower, $r = 5\cos(2\theta)$. For each task, we'll use the following visualization:



a. In Step 1, explain how the changes to the cosine function (multiplying the input, θ by 2 and multiplying the cosine function by 5) changes the graph from a standard cosine function's graph. Where are the zeros, compared to where they typically are?
b. Fill in the table with the function outputs $f(\theta)$ for different values of θ. Note that the points have been colored based on the sign on the output.
c. In Step 2, explain why there only looks to be one grey point (when $r = 0$, when there were more plotted in Step 1.)
d. Think about how θ changes. As θ increases, where should the points be traveling? Describe this in terms of quadrants on the graph if you'd like. Does this match the movement of the points on the polar function?
e. Fill in the table with the r outputs at different values of θ. Describe why these points are plotted where they are, including the quadrants and distances from the origin/pole.
f. In Step 3, confirm that the polar curve $r = 5\cos(2\theta)$ doesn't follow the counter-clockwise path that we expect when θ increases. Compare the path that $r = 5\cos(2\theta)$ travels with the path that $r = |5\cos(2\theta)|$ travels as θ increases. What is happening to the curve when $r = 5\cos(2\theta)$ is negative?
g. What, then, do you think happens to a polar point (r, θ) when $r < 0$? How do we plot it?

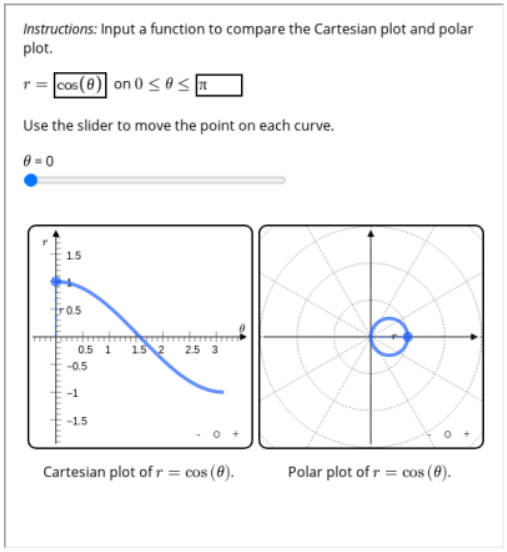
This should be a relatively intuitive extension of our idea of “distance.” When our a polar point has a negative value of r, the point is reflected through the pole (the origin).

Note 10.2.6 Many Representations of One Point.

We’ve seen now that a single point can be represented in many ways!

1. Polar points (r, θ) and $(r, \theta + 2k\pi)$, where k is some integer, will end up looking identical.
2. Polar points (r, θ) and $(-r, \theta + (2k + 1)\pi)$, where k is some integer, will end up looking identical.

Before moving on, it might be nice to get more familiarity with polar plots of curves. Use the graphing utility below to look at different curves and compare the Cartesian and polar plots.



Standalone
In Context
Embed

Converting Between Coordinate Systems

When we thought about Parametric Curves, we used a technique called Eliminating the Parameter to write a parametric curve in a more standard format, $y = f(x)$. It seems reasonable to come up with a way of converting a polar curve to Cartesian one, as well.

Before we tackle this, we’ll focus on converting the coordinates of a single point: we’ll see how we can write the coordinates of a polar point in the Cartesian coordinate system and the coordinates of a Cartesian point in the polar coordinate system.

Activity 10.2.7: Converting the Coordinates of a Point.

We’re going to think about the two coordinate systems (and converting between them) by considering a point and collecting information. For simplicity, we’ll visualize a point in the first quadrant.

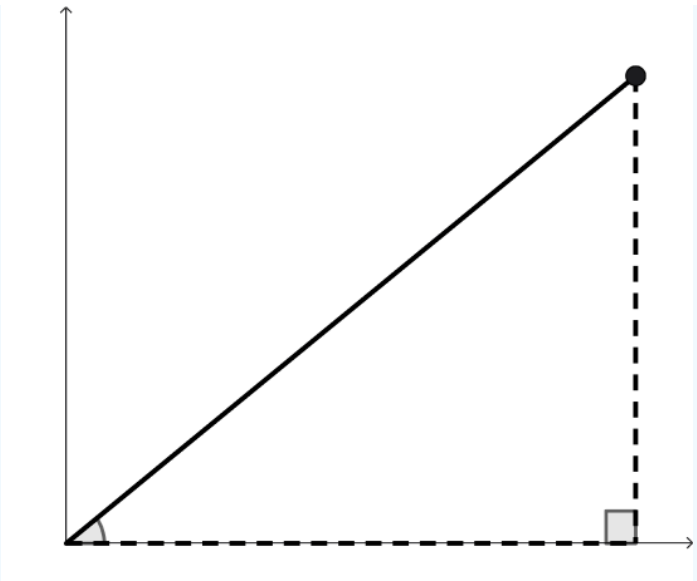


Figure 10.2.8

a. If the point is written in Cartesian coordinates, (x, y), label x and y in the image.
b. If the point is written in polar coordinates, (r, θ), label $\theta$and r in the image.
c. Assume that we know the values of x and y. Come up with an equation linking r, x, and y in order for us to find the value of r. **Hint**. What equations do we have connecting the side lengths of a triangle like this one?
d. Now come up with some equation linking θ, x, and y in order for us to find the value of θ. **Hint**. What kinds of trigonometric functions connect these three variables?
e. Now imagine that we know the values of r and θ instead. Come up with an equation connecting x, r, and θ in order for us to find the value of x. **Hint**. What kinds of trigonometric functions connect these two side lengths of this triangle?
f. Come up with an equation connecting y, r, and θ in order for us to find the value of y. **Hint**. What kinds of trigonometric functions connect these two side lengths of this triangle?

This small triangle exploration gives us a lot of traction! We now have a collection of formulas to help us convert the coordinates of a point back and forth between the Cartesian coordinate system and the polar coordinate system.

Theorem 10.2.9: Polar/Cartesian Coordinate Conversion.

For a point (r, θ) in the polar coordinate system, the corresponding Cartesian point (x, y) can be written where:

$$x = r\cos(\theta)$$
$$y = r\sin(\theta)$$

For a point (x, y) in the Cartesian coordinate system, a corresponding polar point (r, θ) can be written where:

$$\tan(\theta) = \frac{y}{x}$$
$$r^2 = x^2 + y^2$$

Note 10.2.10 Be Careful Finding Polar Coordinates.

We saw in Note 10.2.6 that polar points can be written in an infinite number of ways, and so we need to be careful while solving for θ and r if we are given the Cartesian coordinates x and y.

- Note that $\theta = \tan^{-1}\left(\frac{y}{x}\right)$ will only give values of θ in the interval $\left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$.

- Note that $r = \sqrt{x^2 + y^2}$ will only ever give values of r in $[0, \infty)$.

For a Cartesian point in the second or third quadrant, we can either change the value of θ to make sure it is pointing in the correct direction, or change the sign on r to reflect it back into the correct quadrant.

Activity 10.2.11: Converting Coordinates.

Fill in the following table, converting points written in the Cartesian coordinate system to the polar coordinate system and points in the polar coordinate system to the Cartesian coordinate system. It might be helpful to sketch the point first! Then, find a second polar representation for each point.

Table 10.2.12

Cartesian	Polar	Second Polar
	$\left(3, \frac{7\pi}{4}\right)$	
	$\left(\frac{1}{2}, \frac{\pi}{6}\right)$	
$(-4, 4\sqrt{3})$		
$(-3, -5)$		
	$\left(-1, -\frac{2\pi}{3}\right)$	

Answer available at the end of this section.

Let's finish this section with one last note. In thinking about converting the coordinates of a single point, we've actually come up with a way of converting functions!

For a polar function $r = f(\theta)$, we can use Theorem 10.2.9 to write a parametrically defined curve:

$$x = f(\theta)\cos(\theta)$$
$$y = f(\theta)\sin(\theta)$$

In general, eliminating the parameter to write this as a more standard $y = g(x)$ is a difficult, if even possible, task.

Practice Problems

Exercise 10.2.1

Explain how a curve written in polar coordinates is different than one written in Cartesian coordinates.

Exercise 10.2.2

Convert the following Cartesian coordinates to polar coordinates

a. (4, 0)
b. (−5, 5)
c. (3, −3)
d. $(-\sqrt{3}, -1)$
e. (0, −3)
f. $(-4, 4\sqrt{3})$

Exercise 10.2.3

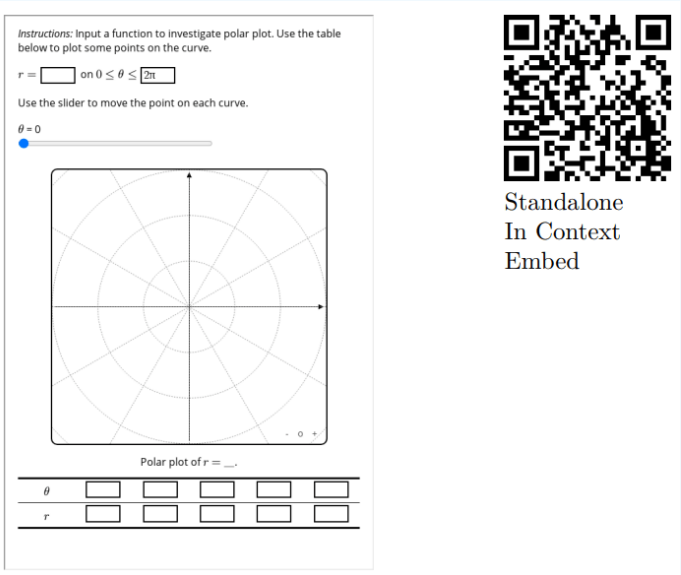
For each point above, find a second, different, representation of the same point. Explain how you did this.

Exercise 10.2.4

Explain why there are an infinite number of possible polar coordinates corresponding with each pair of Cartesian coordinates above

Exercise 10.2.5

For each of the following polar curves, plot the curve in the graphing utility below. Then, find at least 5 different points on the polar curve, plot them, and describe the behavior of the curve.



a. $r = \sin(\theta) + \cos(\theta)$
b. $r = \sin^2(\theta)$
c. $r = \cos(4\theta) + \theta$
d. $r = \sin^2(-\theta)$
e. $r = \theta^2 \sin(\theta)$

Answers

Activity 10.2.11: Converting Coordinates.

Table 10.2.13. One possible solution

Cartesian	Polar	Second Polar
$\left(\frac{3}{\sqrt{2}}, -\frac{3}{\sqrt{2}}\right)$	$\left(3, \frac{7\pi}{4}\right)$	$\left(-3, \frac{3\pi}{4}\right)$
$\left(\frac{\sqrt{3}}{4}, \frac{1}{4}\right)$	$\left(\frac{1}{2}, \frac{\pi}{6}\right)$	$\left(-\frac{1}{2}, \frac{7\pi}{6}\right)$
$(-4, 4\sqrt{3})$	$\left(8, \frac{2\pi}{3}\right)$	$\left(8, \frac{8\pi}{3}\right)$

Cartesian	Polar	Second Polar
$(-3,-5)$	$\left(5,\tan^{-1}\left(\frac{4}{3}\right)+\pi\right)$	$\left(-5,\tan^{-1}\left(\frac{4}{3}\right)\right)$
$\left(\frac{1}{2},\frac{\sqrt{3}}{2}\right)$	$\left(-1,-\frac{2\pi}{3}\right)$	$\left(1,\frac{\pi}{3}\right)$

This page titled 10.2: Polar Curves is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Peter Keep.

10.3: Calculus with Parametric Curves

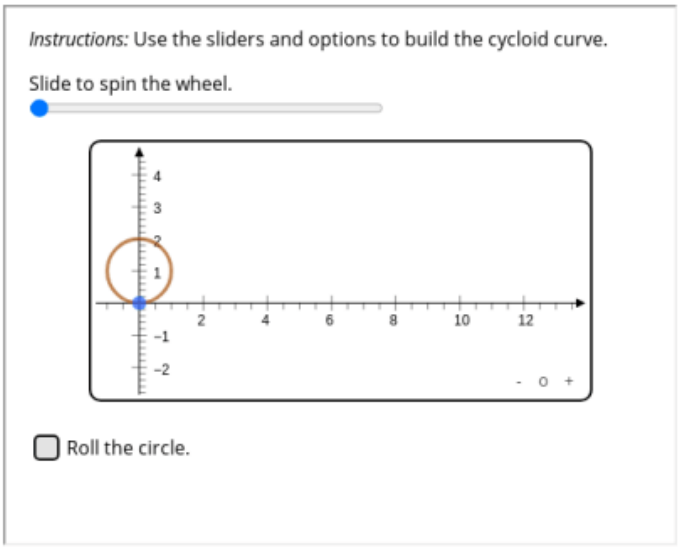
Now that we're a bit more familiar with parametric curves, we'll think about them in the context of calculus. There are only a few results that we'll build, but all of them should end up being relatively intuitive, as long as we refer back to the results in our standard context.

But first, let's introduce a really nice parametric curve that we'll use to think about the calculus results.

The Cycloid

The cycloid is a great example of a parametric curve, because the description of the curve leads to the parametric definition of the curve.

If we paint a dot on the edge of a wheel and roll the wheel in a straight path, what path will the point trace?



Standalone
In Context
Embed

The visual above is helpful in building the parametric definition of the cycloid. We start with a circle. Due to clockwise rotation (instead of the typical counter-clockwise rotation) as well as the point beginning at the bottom of the circle, we parameterize it as:

$$x = -\sin(t)$$
$$y = -\cos(t)$$

In order to shift the circle up to be centered at (0, 1), we can add 1 to the y-value and use the parametric curve:

$$x = -\sin(t)$$
$$y = 1 - \cos(t)$$

Then, in order to shift the circle to the right as the time changes, we add t to the x-value:

$$x = t - \sin(t)$$
$$y = 1 - \cos(t)$$

If we want to scale the circle to have a generic radius, r, we can scale the whole thing!

Definition 10.3.1: The Cycloid.

The cycloid passing through the origin and generated by a circle of radius r rolling along the x-axis from left to right is parameterized as:

$$x = r(t - \sin(t))$$
$$y = r(1 - \cos(t))$$

Slopes of Tangent Lines

Before we begin, let's remind ourselves of a key idea when we were introducing parametric curves: Both x and y are independent functions of some third variable (or parameter), t.

In order for us think about derivatives, we should revisit The Chain Rule.

For us to find $\frac{dy}{dx}$, we can start by finding $\frac{dy}{dt}$ using the Chain Rule:

$$\frac{dy}{dt} = \frac{dy}{dx} \cdot \frac{dx}{dt}$$

Now, we can solve for $\frac{dy}{dx}$.

Theorem 10.3.2: The Chain Rule (Restated).

If x and y are independent variables and both are functions of some third variable t, then:

$$\frac{dy}{dx} = \frac{\left(\frac{dy}{dt}\right)}{\left(\frac{dx}{dt}\right)}$$

when $\frac{dy}{dt}$ and $\frac{dx}{dt}$ exist and $\frac{dx}{dt} \neq 0$.

Now, let's revisit the cycloid!

Activity 10.3.3: Slopes on a Cycloid.

Consider the cycloid formed by a circle with radius $r = 1$:

$$x = t - \sin(t)$$
$$y = 1 - \cos(t)$$

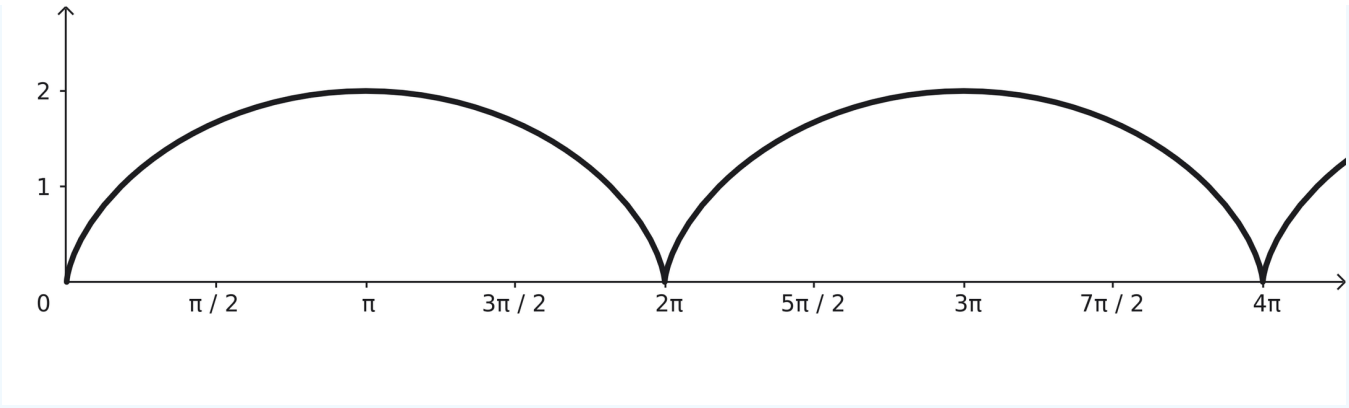


Figure 10.3.4 The cycloid.

a. First, make a conjecture: where do you think the cycloid will have horizontal tangent lines? What about vertical tangent lines, or other points where the derivative doesn't exist?
b. Find $\frac{dy}{dt}$ and $\frac{dx}{dt}$, and construct $\frac{dy}{dx}$.
c. Find the values of t where $\frac{dy}{dx} = 0$.
d. Find the values of t where $\frac{dy}{dx}$ does not exist.
e. Use these values of t to find the points on the cycloid where there are horizontal tangent lines or points where the derivative doesn't exist. Do they match what you conjectured?

Areas Bounded by Curves

We can extend our idea from slopes, where we had to revisit the Chain Rule, towards constructing an integral for a parametric curve. In order to find the area under a curve, we will still use the basic idea of a Riemann Sum: we want to multiply a height (y) by a width (dx) on small and specific intervals, and add them up to construct an integral. Since our height variable, y, is a function of t, the integral will end up with an input variable that doesn't match the differential. We can use a change of variables (like in u -Substitution) to change the differential from dx to dt.

$$dx = x'(t)dt$$

This changes our differential, allowing us to integrate with regard to t, the input variable.

Theorem 10.3.5: Parametric Area.

For a parametric curve

$$x = x(t)$$
$$y = y(t)$$

on the interval where $x(t)$ and $y(t)$ are continuous on the interval $a \leq t \leq b$ and $x(t)$ is differentiable on $a < t < b$, then the signed area bounded between the parametric curve and the x-axis from from $t = a$ until $t = b$ is

$$\int_{t=a}^{t=b} y(t)x'(t)dt$$

Example 10.3.6: Area Trapped in the Cycloid.

We return to the cycloid. We'll again use the cycloid generated from a circle with radius $r = 1$:

$$x = t - \sin(t)$$
$$y = 1 - \cos(t)$$

a. Set up an integral to calculate the area trapped under one arch of the cycloid.
b. Evaluate this integral!
c. How does this change when the cycloid is generated from a circle with radius r?

$$x = r(t - \sin(t))$$
$$y = r(1 - \cos(t))$$

Solution

a.
$$\int_{t=0}^{t=2\pi} (1-\cos(t))^2 dt$$

b.
$$\begin{aligned}\int_{t=0}^{t=2\pi} (1-\cos(t))^2 dt &= \int_{t=0}^{t=2\pi} \left(1 - 2\cos(t) + \cos^2(t)\right) dt \\ &= \int_{t=0}^{t=2\pi} \left(\frac{3}{2} - 2\cos(t) + \frac{1}{2}\cos(2t)\right) dt \\ &= \left(\frac{3t}{2} - 2\sin(t) - \frac{1}{4}\sin(2t)\right)\Big|_{t=0}^{t=2\pi} \\ &= 3\pi\end{aligned}$$

c.
$$\begin{aligned}\int_{t=0}^{t=2\pi} (r - r\cos(t))^2 dt &= r^2 \int_{t=0}^{t=2\pi} (1-\cos(t))^2 dt \\ &= r^2 \int_{t=0}^{t=2\pi} \left(1 - 2\cos(t) + \cos^2(t)\right) dt \\ &= r^2 \int_{t=0}^{t=2\pi} \left(\frac{3}{2} - 2\cos(t) + \frac{1}{2}\cos(2t)\right) dt \\ &= r^2 \left(\frac{3t}{2} - 2\sin(t) - \frac{1}{4}\sin(2t)\right)\Big|_{t=0}^{t=2\pi} \\ &= 3\pi r^2\end{aligned}$$

Arc Lengths

We have built an arc length formula already in this textbook! Remember: Definition 6.5.7 Length of a Curve? Let's rebuild this integral formula in a way that makes sense for a parametric curve.

Activity 10.3.7: Building a Parametric Arc Length Formula.

Before we begin, it might be helpful to remind yourself how we build the arclength formula in Section 6.5 Arc Length and Surface Area. Let's drop back in to that derivation with some small changes. Because we're dealing with a parametric curve, we'll use t as our input variable, and so we'll eventually need a Δt to turn into a differential dt in the integral. This means that we'll have non-uniform Δx distances, since they'll be based on the changes in the input, t. So we'll actually use Δx_k to represent the change in the x-variable on the kth subinterval. So, our length on the kth subinterval is:

$$\ell_k = \sqrt{(\Delta x_k)^2 + (\Delta y_k)^2}$$

a. In order to build the arc length formula in Section 6.5, we factored out $\sqrt{\Delta x^2}$ in order to end up with a Δx that turned into the differential dx in the integral. This time, factor out a Δt^2 under the square root.
b. Now, create a Riemann sum and let $n \to \infty$ (and, correspondingly, $\Delta t \to 0$). Note that you should end up with some differentials instead of deltas!

Answer available at the end of this section.

Definition 10.3.8: Parametric Arc Length.

If the functions $x(t)$ and $y(t)$ are continuous on $[a,b]$ and differentiable on (a,b), the the length of the parametric curve

$$x = x(t)$$
$$y = y(t)$$

on $a \le t \le b$ is

$$\ell = \int_{t=a}^{t=b} \sqrt{(x'(t))^2 + (y'(t))^2}dt$$

Example 10.3.9: Length of a Cycloid Arc.

Let's go back, one more time, to the cycloid. We'll again use the cycloid generated from a circle with radius $r = 1$:

$$x = t - \sin(t)$$
$$y = 1 - \cos(t)$$

a. Set up the integral measuring the length of one arch of the cycloid.

b. Now, evaluate the integral! **Hint**. It might help to remember a trig identity! We know $\sin^2(\theta) = \frac{1}{2}(1 - \cos(2\theta)$. So then we can note that

$$\frac{1}{2}(1 - \cos(t)) = \sin^2\left(\frac{t}{2}\right)$$

c. How does this change when the cycloid is generated from a circle with radius r?

$$x = r(t - \sin(t))$$
$$y = r(1 - \cos(t))$$

Solution

a.
$$\begin{aligned}\ell &= \int_{t=0}^{t=2\pi} \sqrt{(1-\cos(t))^2 + (\sin(t))^2}dt \\ &= \int_{t=0}^{t=2\pi} \sqrt{1 - 2\cos(t) + \cos^2(t) + \sin^2(t)}dt \\ &= \int_{t=0}^{t=2\pi} \sqrt{2 - 2\cos(t)}dt \\ &= \sqrt{2}\int_{t=0}^{t=2\pi} \sqrt{1 - \cos(t)}dt\end{aligned}$$

b.
$$\begin{aligned}\ell &= \sqrt{2}\int_{t=0}^{t=2\pi} \sqrt{1 - \cos(t)}dt \\ &= 2\int_{t=0}^{t=2\pi} \sqrt{\frac{1}{2}(1 - \cos(t))}dt \\ &= 2\int_{t=0}^{t=2\pi} \sqrt{\sin^2\left(\frac{t}{2}\right)}dt \\ &= 2\int_{t=0}^{2\pi} \sin\left(\frac{t}{2}\right)dt \\ &= 2\left(-2\cos\left(\frac{t}{2}\right)\right)\Big|_{t=0}^{t=2\pi} \\ &= 2(-2-2) \\ &= 8\end{aligned}$$

c.

$$\begin{aligned}
\ell &= \int_{t=0}^{t=2\pi} \sqrt{(r(1-\cos(t)))^2 + (r\sin(t))^2}dt \\
&= \int_{t=0}^{t=2\pi} \sqrt{r^2\left(1 - 2\cos(t) + \cos^2(t) + \sin^2(t)\right)}dt \\
&= \int_{t=0}^{t=2\pi} r\sqrt{2 - 2\cos(t)}dt \\
&= r\sqrt{2}\int_{t=0}^{t=2\pi} \sqrt{1-\cos(t)}dt \\
&= 2r\int_{t=0}^{t=2\pi} \sqrt{\frac{1}{2}(1-\cos(t))}dt \\
&= 2r\int_{t=0}^{t=2\pi} \sqrt{\sin^2\left(\frac{t}{2}\right)}dt \\
&= 2r\int_{t=0}^{2\pi} \sin\left(\frac{t}{2}\right)dt \\
&= 2r\left(-2\cos\left(\frac{t}{2}\right)\right)\Big|_{t=0}^{t=2\pi} \\
&= 2r(-2-2) \\
&= 8r
\end{aligned}$$

Answers

Activity 10.3.7: Building a Parametric Arc Length Formula.

a.

$$\begin{aligned}
\ell_k &= \sqrt{(\Delta t)^2\left(\frac{(\Delta x_k)^2}{(\Delta t)^2} + \frac{(\Delta y_k)^2}{(\Delta t)^2}\right)} \\
&= \sqrt{(\Delta t)^2}\sqrt{\left(\frac{\Delta x_k}{\Delta t}\right)^2 + \left(\frac{\Delta y_k}{\Delta t}\right)^2} \\
&= \sqrt{\left(\frac{\Delta x_k}{\Delta t}\right)^2 + \left(\frac{\Delta y_k}{\Delta t}\right)^2}\Delta t
\end{aligned}$$

b.

$$\begin{aligned}
\ell &\approx \sum_{k=1}^{n}\sqrt{\left(\frac{\Delta x_k}{\Delta t}\right)^2 + \left(\frac{\Delta y_k}{\Delta t}\right)^2}\Delta t \\
&= \lim_{n\to\infty}\sum_{k=1}^{n}\sqrt{\left(\frac{\Delta x_k}{\Delta t}\right)^2 + \left(\frac{\Delta y_k}{\Delta t}\right)^2}\Delta t \\
&= \int_{t=a}^{t=b}\sqrt{\left(\frac{dx}{dt}\right)^2 + \left(\frac{dy}{dt}\right)^2}dt
\end{aligned}$$

This page titled 10.3: Calculus with Parametric Curves is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Peter Keep.

10.4: Calculus with Polar Curves

In one sense, we don't have much to do in this section. Since we can think of polar functions as parameterized Cartesian functions, we can rely on the results from Section 10.3 Calculus with Parametric Curves for much of this! In fact, that's exactly what we'll do.

When we consider a polar function $r = f(\theta)$, we'll also think of it as a parametric one:

$$x = f(\theta)\cos(\theta)$$
$$y = f(\theta)\sin(\theta)$$

Slopes and arc lengths will follow pretty directly from our results for parametric curves, but we'll need to think about the areas bounded by polar curves a bit more.

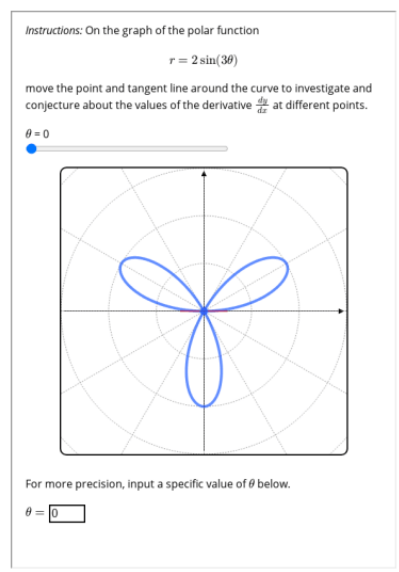
Slopes of Tangent Lines

Activity 10.4.1: Slopes on a Polar Flower.

Consider the polar flower defined by

$$r = 2\sin(3\theta)$$

for $0 \leq \theta \leq \pi$. For this activity, refer to the graphing utility below.



Standalone
In Context
Embed

a. How many points on this polar flower do you think have horizontal tangent lines?
b. How many points on this polar flower do you think have vertical tangent lines?
c. Make a conjecture about the slope of the tangent line when $\theta = \frac{5\pi}{6}$ by looking at the graph only.
d. Find an expression for $\frac{dy}{dx}$.
e. Pick one of the values of θ where you thought this curve might have a horizontal or vertical tangent line. Evaluate $\frac{dy}{dx}$ at that value, and check to see if you were correct.
f. Evaluate $\frac{dy}{dx}$ when $\theta = \frac{5\pi}{6}$. How close was your guess at the slope?

Answer available at the end of this section.

Arc Lengths

We could, similarly, dive right into our Parametric Arc Length formula, but let's pause for a second. It might get a bit messy working with each specific $\frac{dx}{d\theta}$ and $\frac{dy}{d\theta}$, so we'll start with this in a more general sense.

Consider some polar function named r.

We'll use the product rule to find the derivatives needed for the arc length formula:

$$\begin{aligned}\frac{dx}{d\theta} &= \frac{d}{d\theta}(r\cos(\theta)) \\ &= r'\cos(\theta) - r\sin(\theta) \\ \frac{dy}{d\theta} &= \frac{d}{d\theta}(r\sin(\theta)) \\ &= r'\sin(\theta) + r\cos(\theta)\end{aligned}$$

Let's now find use these in our arc length formula!

$$\begin{aligned}\ell &= \int_{\theta=\alpha}^{\theta=\beta} \sqrt{(r'\cos(\theta) - r\sin(\theta))^2 + (r'\sin(\theta) + r\cos(\theta))^2}d\theta \\ &= \int_{\theta=\alpha}^{\theta=\beta} \sqrt{(r'\cos(\theta))^2 - 2rr'\sin(\theta)\cos(\theta) + (r\sin(\theta))^2 + (r'\sin(\theta))^2 + 2rr'\sin(\theta)\cos(\theta) + (r\cos(\theta))^2}d\theta \\ &= \int_{\theta=\alpha}^{\theta=\beta} \sqrt{r^2\left(\sin^2(\theta) + \cos^2(\theta)\right) + (r')^2\left(\sin^2(\theta) + \cos^2(\theta)\right)}d\theta \\ &= \int_{\theta=\alpha}^{\theta=\beta} \sqrt{r^2 + (r')^2}d\theta\end{aligned}$$

Definition 10.4.2: Polar Arc Length.

If $r(\theta)$ is a continuous function on $[\alpha, \beta]$ and differentiable on (α, β), then the length of the polar curve $r(\theta)$ between $\theta = \alpha$, and $\theta = \beta$ is:

$$\ell = \int_{\theta=\alpha}^{\theta=\beta} \sqrt{r^2 + \left(\frac{dr}{d\theta}\right)^2}d\theta$$

Example 10.4.3: Petal Length.

For the polar graphs that we've been calling "Polar Flowers," we can call each loop that the polar curve makes (between the points where it crosses the pole) petals.

We'll consider, again, the polar flower $r = 2\sin(3\theta)$.

a. Find the interval of θ that defines the first petal of the polar flower. **Hint**. We're really thinking about where $2\sin(3\theta) = 0$.
b. Set up an integral for the length of the curve tracing out that first petal.

Solution

b.
$$\begin{aligned}\ell &= \int_{\theta=0}^{\theta=\pi/3} \sqrt{r^2 + \left(\frac{dr}{d\theta}\right)^2}\, d\theta \\ &= \int_{\theta=0}^{\theta=\pi/3} \sqrt{(2\sin(3\theta))^2 + (6\cos(3\theta))^2}\, d\theta \\ &= \int_{\theta=0}^{\theta=\pi/3} \sqrt{4\sin^2(3\theta) + 36\cos^2(3\theta)}\, d\theta\end{aligned}$$

Areas Bounded by Curves

The last calculus topic we'll visit in this section will be one to reinforce, for one last time, an important theme: the theme of accumulation.

We're going to go back to our slice-and-sum process to build a formula for calculating a polar area.

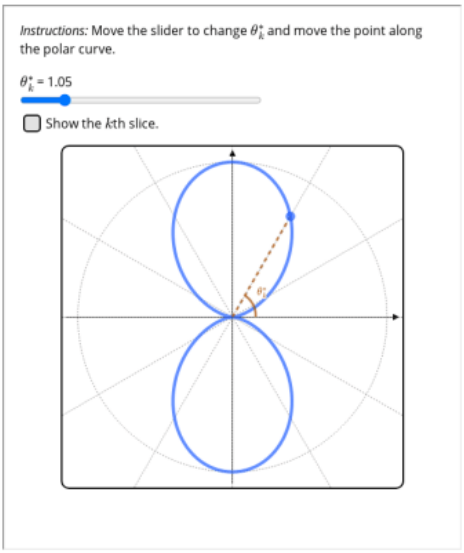
Activity 10.4.4: Slice-and-Sum for Polar Areas.

Remember slice-and-sum? Here's the general procedure:

- We'll divide up the interval of inputs into n (equally-sized) subintervals. We'll pick one value from each subinterval.
- One each slice, we'll use the representative input variable to calculate the measurement we're interested in. In this case, it's an area.
- Add up the areas of the n slices, and the sum should be an approximation of the actual area we're interested in.
- Use a limit as $n \to \infty$ to produce an integral our of a Riemann sum!

a. The area of a sector of a circle with angle α is:
$$A = \frac{1}{2}\alpha r^2$$
Convince yourselves that this is true, using appropriate test angles for:
   - A circle.
   - A semicircle.
   - A quarter of a circle.

b. The reason that we're thinking about sectors of circles is that when we slice an interval of angles in a polar coordinate system, a bunch of rays extending away from the pole. When do this for a polar curve, we'll pick a point in the subinterval and assume that the whole subinterval has a constant function output, just like we did with the rectangle. Take a look at the graphing utility below, and get a feel for how we'll find the kth slice.



Standalone
In Context
Embed

c. Calculate the area of the kth slice, A_k.
d. Add these up to approximate the total area.
e. Introduce a limit to find an integral formula for the area trapped inside a polar curve from $\theta = \alpha$ to $\theta = \beta$.

Answer available at the end of this section.

Example 10.4.5: Petal Area.

Let's consider, one last time, the polar flower $r = 2\sin(3\theta)$. Set up and evaluate an integral calculate the area bounded inside one of the petals of the polar flower.

Solution

$$\begin{aligned} A &= \frac{1}{2}\int_{\theta=0}^{\theta=\pi/3} (2\sin(3\theta))^2 d\theta \\ &= \frac{1}{2}\int_{\theta=0}^{\theta=\pi/3} \left(4\sin^2(3\theta)\right) d\theta \\ &= \int_{\theta=0}^{\theta=\pi/3} (1-\cos(3\theta)) d\theta \\ &= \left(\theta + \frac{1}{3}\sin(3\theta)\right)\Big|_{\theta=0}^{\theta=\pi/3} \\ &= \frac{\pi}{3} \end{aligned}$$

Answers

Activity 10.4.1: Slopes on a Polar Flower.

d.
$$\begin{aligned} \frac{dy}{dx} &= \frac{\frac{d}{d\theta}(2\sin(3\theta)\sin(\theta))}{\frac{d}{d\theta}(2\sin(3\theta)\cos(\theta))} \\ &= \frac{2\sin(3\theta)\cos(\theta) + 6\cos(3\theta)\sin(\theta)}{6\cos(3\theta)\cos(\theta) - 2\sin(3\theta)\sin(\theta)} \\ &= \frac{\sin(3\theta)\cos(\theta) + 3\cos(3\theta)\sin(\theta)}{3\cos(3\theta)\cos(\theta) - \sin(3\theta)\sin(\theta)} \end{aligned}$$

e.
$$\frac{dy}{dx}\Big|_{\theta=\frac{5\pi}{6}} = \frac{\sin\left(\frac{15\pi}{6}\right)\cos\left(\frac{5\pi}{6}\right) + 3\cos\left(\frac{15\pi}{6}\right)\sin\left(\frac{5\pi}{6}\right)}{3\cos\left(\frac{15\pi}{6}\right)\cos\left(\frac{5\pi}{6}\right) - \sin\left(\frac{15\pi}{6}\right)\sin\left(\frac{5\pi}{6}\right)}$$

f.
$$\begin{aligned} &= \frac{-\frac{\sqrt{3}}{2}}{-\frac{1}{2}} \\ &= \sqrt{3} \end{aligned}$$

Activity 10.4.4: Slice-and-Sum for Polar Areas.

c.
$$A_k = \frac{1}{2}\left(r\left(\theta_k^*\right)\right)^2 \Delta\theta$$

d.
$$A \approx \frac{1}{2}\sum_{k=1}^{n} \left(r\left(\theta_k^*\right)\right)^2 \Delta\theta$$

e.
$$\begin{aligned} A &= \lim_{n\to\infty} \frac{1}{2}\sum_{k=1}^{n} \left(r\left(\theta_k^*\right)\right)^2 \Delta\theta \\ &= \frac{1}{2}\int_{\theta=\alpha}^{\theta=\beta} r^2 d\theta \end{aligned}$$

This page titled 10.4: Calculus with Polar Curves is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Peter Keep.

INDEX

No index terms found.

This page titled Index is shared under a license and was authored, remixed, and/or curated by Peter Keep.

GLOSSARY

S

Sample Word 1 | Sample Definition 1

This page titled Glossary is shared under a license and was authored, remixed, and/or curated by Peter Keep.

DETAILED LICENSING

Overview

Title: Discover Calculus

Webpages: 80

All licenses found:

- CC BY 4.0: 91.3% (73 pages)
- Public Domain: 5% (4 pages)
- Undeclared: 3.8% (3 pages)

By Page

- Discover Calculus - *CC BY 4.0*
  - Front Matter - *CC BY 4.0*
    - TitlePage - *CC BY 4.0*
    - InfoPage - *Public Domain*
    - Table of Contents - *Public Domain*
    - Licensing - *Public Domain*
  - 1: Limits - *CC BY 4.0*
    - 1.1: The Definition of the Limit - *CC BY 4.0*
    - 1.2: Evaluating Limits - *CC BY 4.0*
    - 1.3: First Indeterminate Forms - *CC BY 4.0*
    - 1.4: Limits Involving Infinity - *CC BY 4.0*
    - 1.5: The Squeeze Theorem - *CC BY 4.0*
    - 1.6: Continuity and the Intermediate Value Theorem - *CC BY 4.0*
    - 1.7: Exploration- Limits - *CC BY 4.0*
  - 2: Derivatives - *CC BY 4.0*
    - 2.1: Introduction to Derivatives - *CC BY 4.0*
    - 2.2: Interpreting Derivatives - *CC BY 4.0*
    - 2.3: Some Early Derivative Rules - *CC BY 4.0*
    - 2.4: The Product and Quotient Rules - *CC BY 4.0*
    - 2.5: The Chain Rule - *CC BY 4.0*
  - 3: Implicit Differentiation - *CC BY 4.0*
    - 3.1: Implicit Differentiation - *CC BY 4.0*
    - 3.2: Derivatives of Inverse Functions - *CC BY 4.0*
    - 3.3: Logarithmic Differentiation - *CC BY 4.0*
  - 4: Applications of Derivatives - *CC BY 4.0*
    - 4.1: Mean Value Theorem - *CC BY 4.0*
    - 4.2: Increasing and Decreasing Functions - *CC BY 4.0*
    - 4.3: Concavity - *CC BY 4.0*
    - 4.4: Global Maximums and Minimums - *CC BY 4.0*
    - 4.5: Optimization - *CC BY 4.0*
    - 4.6: Linear Approximations - *CC BY 4.0*
    - 4.7: L'Hôpital's Rule - *CC BY 4.0*
    - 4.8: Exploration- Applications of Derivatives - *CC BY 4.0*
  - 5: Antiderivatives and Integrals - *CC BY 4.0*
    - 5.1: Antiderivatives and Indefinite Integrals - *CC BY 4.0*
    - 5.2: Riemann Sums and Area Approximations - *CC BY 4.0*
    - 5.3: The Definite Integral - *CC BY 4.0*
    - 5.4: The Fundamental Theorem of Calculus - *CC BY 4.0*
    - 5.5: More Results about Definite Integrals - *CC BY 4.0*
    - 5.6: Introduction to u Substitution - *CC BY 4.0*
  - 6: Applications of Integrals - *CC BY 4.0*
    - 6.1: Integrals as Net Change - *CC BY 4.0*
    - 6.2: Area Between Curves - *CC BY 4.0*
    - 6.3: Volumes of Solids of Revolution - *CC BY 4.0*
    - 6.4: More Volumes- Shifting the Axis of Revolution - *CC BY 4.0*
    - 6.5: Arc Length and Surface Area - *CC BY 4.0*
    - 6.6: Other Applications of Integrals - *CC BY 4.0*
  - 7: Techniques for Antidifferentiation - *CC BY 4.0*
    - 7.1: Improper Integrals - *CC BY 4.0*
    - 7.2: More on u Substitution - *CC BY 4.0*
    - 7.3: Manipulating Integrands - *CC BY 4.0*
    - 7.4: Integration by Parts - *CC BY 4.0*
    - 7.5: Integrating Powers of Trigonometric Functions - *CC BY 4.0*
    - 7.6: Trigonometric Substitution - *CC BY 4.0*
    - 7.7: Partial Fractions - *CC BY 4.0*
  - 8: Infinite Series - *CC BY 4.0*
    - 8.1: Introduction to Infinite Sequences - *CC BY 4.0*
    - 8.2: Introduction to Infinite Series - *CC BY 4.0*
    - 8.3: The Divergence Test and the Harmonic Series - *CC BY 4.0*
    - 8.4: The Integral Test - *CC BY 4.0*
    - 8.5: Alternating Series and Conditional Convergence - *CC BY 4.0*
    - 8.6: Common Series Types - *CC BY 4.0*
    - 8.7: Comparison Tests - *CC BY 4.0*

- 8.8: The Ratio and Root Tests - *CC BY 4.0*
- 8.9: Exploration- Infinite Series - *CC BY 4.0*

◦ 9: Power Series - *CC BY 4.0*
- 9.1: Polynomial Approximations of Functions - *CC BY 4.0*
- 9.2: Power Series Convergence - *CC BY 4.0*
- 9.3: How to Build Taylor Series - *CC BY 4.0*
- 9.4: How to Use Taylor Series - *CC BY 4.0*
- 9.5: Exploration- Power Series - *CC BY 4.0*

◦ 10: Parametric and Polar Curves - *CC BY 4.0*
- 10.1: Parametric Curves - *CC BY 4.0*
- 10.2: Polar Curves - *CC BY 4.0*
- 10.3: Calculus with Parametric Curves - *CC BY 4.0*
- 10.4: Calculus with Polar Curves - *CC BY 4.0*

◦ Back Matter - *Undeclared*
- Index - *Undeclared*
- Glossary - *Undeclared*
- Detailed Licensing - *Public Domain*

This page titled Discover Calculus is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Peter Keep.