7-2017

APEX Calculus: UND Edition

Gregory Hartman
Virginia Military Institute

Department of Mathematics, University of North Dakota

Recommended Citation

A$^{P}_{E}$X CALCULUS
Late Transcendentals

University of North Dakota

Adapted from A$^{P}_{E}$X Calculus by
Gregory Hartman, Ph.D.
Department of Applied Mathematics
Virginia Military Institute

Revised July 27, 2017

Contributing Authors
Troy Siemers, Ph.D.
Department of Applied Mathematics
Virginia Military Institute

Michael Corral
Mathematics
Schoolcraft College

Brian Heinold, Ph.D.
Department of Mathematics
and Computer Science
Mount Saint Mary's University

Paul Dawkins, Ph.D.
Department of Mathematics
Lamar University

Dimplekumar Chalishajar, Ph.D.
Department of Applied Mathematics
Virginia Military Institute

Editor
Jennifer Bowen, Ph.D.
Department of Mathematics
and Computer Science
The College of Wooster

Contents

Preface

A Note on Using this Text

Thank you for reading this short preface. Allow us to share a few key points about the text so that you may better understand what you will find beyond this page.

This text comprises a three–volume series on Calculus. The first part covers material taught in many "Calculus 1" courses: limits, derivatives, and the basics of integration, found in Chapters 1 through 6. The second text covers material often taught in "Calculus 2": integration and its applications, along with an introduction to sequences, series and Taylor Polynomials, found in Chapters 7 through 10. The third text covers topics common in "Calculus 3" or "Multivariable Calculus": parametric equations, polar coordinates, vector–valued functions, and functions of more than one variable, found in Chapters 11 through 15. All three are available separately for free.

Printing the entire text as one volume makes for a large, heavy, cumbersome book. One can certainly only print the pages they currently need, but some prefer to have a nice, bound copy of the text. Therefore this text has been split into these three manageable parts, each of which can be purchased separately.

A result of this splitting is that sometimes material is referenced that is not contained in the present text. The context should make it clear whether the "missing" material comes before or after the current portion. Downloading the appropriate pdf, or the entire *APEX Calculus LT* pdf, will give access to these topics.

For Students: How to Read this Text

Mathematics textbooks have a reputation for being hard to read. High–level mathematical writing often seeks to say much with few words, and this style often seeps into texts of lower–level topics. This book was written with the goal of being easier to read than many other calculus textbooks, without becoming too verbose.

Each chapter and section starts with an introduction of the coming material, hopefully setting the stage for "why you should care," and ends with a look ahead to see how the just-learned material helps address future problems. Additionally, each chapter includes a section zero, which provides a basic review and practice problems of pre-calculus skills. Since this content is a pre-requisite for calculus, reviewing and mastering these skills are considered your responsibility. This means that it is your responsibility to seek assistance outside of class from your instructor, a math resource center or other math tutoring available on-campus. A solid understanding of these skills are essential to your success in solving calculus problems.

Please read the text; it is written to explain the concepts of Calculus. There are numerous examples to demonstrate the meaning of definitions, the truth

of theorems, and the application of mathematical techniques. When you encounter a sentence you don't understand, read it again. If it still doesn't make sense, read on anyway, as sometimes confusing sentences are explained by later sentences.

You don't have to read every equation. The examples generally show "all" the steps needed to solve a problem. Sometimes reading through each step is helpful; sometimes it is confusing. When the steps are illustrating a new technique, one probably should follow each step closely to learn the new technique. When the steps are showing the mathematics needed to find a number to be used later, one can usually skip ahead and see how that number is being used, instead of getting bogged down in reading how the number was found.

Some proofs have been delayed until later (or omitted completely). In mathematics, *proving* something is always true is extremely important, and entails much more than testing to see if it works twice. However, students often are confused by the details of a proof, or become concerned that they should have been able to construct this proof on their own. To alleviate this potential problem, we do not include the more difficult proofs in the text. The interested reader is highly encouraged to find other proofs online or from their instructor. In most cases, one is very capable of understanding what a theorem *means* and *how to apply it* without knowing fully *why* it is true.

Work through the examples. The best way to learn mathematics is to do it. Reading about it (or watching someone else do it) is a poor substitute. For this reason, every page has a place for *you* to put *your* notes so that *you* can work out the examples. That being said, sometimes it is useful to watch someone work through an example. For this reason, this text also provides links to online videos where someone is working through a similar problem. If you want even more videos, these are generally chosen from

- Khan Academy: `https://www.khanacademy.org/`
- Math Doctor Bob: `http://www.mathdoctorbob.org/`
- Just Math Tutorials: `http://patrickjmt.com/` (unfortunately, they're not well organized)

Some other sites you may want to consider are

- Larry Green's Calculus Videos: `http://www.ltcconline.net/greenl/courses/105/videos/VideoIndex.htm`
- Mathispower4u: `http://www.mathispower4u.com/`
- Yay Math: `http://www.yaymath.org/` (for prerequisite material)

All of these sites are completely free (although some will ask you to donate). Here's a sample one:

> Watch the video:
> Practical Advice for Those Taking College Calculus at
> `https://youtu.be/ILNfpJTZLxk`

Interactive, 3D Graphics

New to Version 3.0 is the addition of interactive, 3D graphics in the .pdf version. Nearly all graphs of objects in space can be rotated, shifted, and zoomed in/out so the reader can better understand the object illustrated.

As of this writing, the only pdf viewers that support these 3D graphics are Adobe Reader & Acrobat (and only the versions for PC / Mac / Unix / Linux computers, not tablets or smartphones). To activate the interactive mode, click on the image. Once activated, one can click/drag to rotate the object and use the scroll wheel on a mouse to zoom in/out. (A great way to investigate an image is to first zoom in on the page of the pdf viewer so the graphic itself takes up much of the screen, then zoom inside the graphic itself.) A CTRL-click/drag pans the object left/right or up/down. By right-clicking on the graph one can access a menu of other options, such as changing the lighting scheme or perspective. One can also revert the graph back to its default view. If you wish to deactivate the interactivity, one can right-click and choose the "Disable Content" option.

Thanks

There are many people who deserve recognition for the important role they have played in the development of this text. First, I thank Michelle for her support and encouragement, even as this "project from work" occupied my time and attention at home. Many thanks to Troy Siemers, whose most important contributions extend far beyond the sections he wrote or the 227 figures he coded in Asymptote for 3D interaction. He provided incredible support, advice and encouragement for which I am very grateful. My thanks to Brian Heinold and Dimplekumar Chalishajar for their contributions and to Jennifer Bowen for reading through so much material and providing great feedback early on. Thanks to Troy, Lee Dewald, Dan Joseph, Meagan Herald, Bill Lowe, John David, Vonda Walsh, Geoff Cox, Jessica Libertini and other faculty of VMI who have given me numerous suggestions and corrections based on their experience with teaching from the text. (Special thanks to Troy, Lee & Dan for their patience in teaching Calc III while I was still writing the Calc III material.) Thanks to Randy Cone for encouraging his tutors of VMI's Open Math Lab to read through the text and check the solutions, and thanks to the tutors for spending their time doing so. A very special thanks to Kristi Brown and Paul Janiczek who took this opportunity far above & beyond what I expected, meticulously checking every solution and carefully reading every example. Their comments have been extraordinarily helpful. I am also thankful for the support provided by Wane Schneiter, who as my Dean provided me with extra time to work on this project. I am blessed to have so many people give of their time to make this book better.

A^P_EX — Affordable Print and Electronic teXts

A^P_EX is a consortium of authors who collaborate to produce high–quality, low–cost textbooks. The current textbook–writing paradigm is facing a potential revolution as desktop publishing and electronic formats increase in popularity. However, writing a good textbook is no easy task, as the time requirements alone are substantial. It takes countless hours of work to produce text, write examples and exercises, edit and publish. Through collaboration, however, the cost to any individual can be lessened, allowing us to create texts that we freely distribute electronically and sell in printed form for an incredibly low cost. Having said that, nothing is entirely free; someone always bears some cost. This text "cost" the authors of this book their time, and that was not enough. *APEX Calculus* would not exist had not the Virginia Military Institute, through a generous Jackson–Hope grant, given the lead author significant time away from teaching so he could focus on this text.

Each text is available as a free .pdf, protected by a Creative Commons Attribution — Noncommercial 4.0 copyright. That means you can give the .pdf to

anyone you like, print it in any form you like, and even edit the original content and redistribute it. If you do the latter, you must clearly reference this work and you cannot sell your edited work for money.

We encourage others to adapt this work to fit their own needs. One might add sections that are "missing" or remove sections that your students won't need. The source files can be found at `https://github.com/APEXCalculus`.

You can learn more at `www.vmi.edu/APEX`.

Greg Hartman

Creating A$^\text{P}_\text{E}$X LT

Starting with the source at `https://github.com/APEXCalculus`, faculty at the University of North Dakota made several substantial changes to create A$^\text{P}_\text{E}$X Late Transcendentals. The most obvious change was to rearrange the text to delay proving the derivative of transcendental functions until Calculus 2. UND added Sections 7.1 and 7.3, adapted several sections from other resources, created the prerequisite sections, included links to videos and Geogebra, and added several examples and exercises. In the end, every section had some changes (some more substantial than others), resulting in a document that is about 20% longer. The source files can now be found at

`https://github.com/teepeemm/APEXCalculusLT_Source`.

Extra thanks are due to Michael Corral for allowing us to use portions of his Vector Calculus, available at `www.mecmath.net/` (specifically, Sections 11.7, 13.9, and 14.7, and Chapter 15) and to Paul Dawkins for allowing us to use portions of his online math notes from `tutorial.math.lamar.edu/` (specifically, Sections 8.5 and 9.7, as well as "Area with Parametric Equations" in Section 10.3). The work on Calculus III was partially supported by the NDUS OER Initiative.

Calculus I

1.0   Chapter Prerequisites

The material in this section provides a basic review of and practice problems for pre-calculus skills essential to your success in Calculus. You should take time to review this section and work the suggested problems (checking your answers against those in the back of the book). Since this content is a pre-requisite for Calculus, reviewing and mastering these skills are considered your responsibility. This means that minimal, and in some cases no, class time will be devoted to this section. When you identify areas that you need help with we strongly urge you to seek assistance outside of class from your instructor or other student tutoring service.

Functions

A **function** f is a rule that assigns each element x from a set (called the domain) to exactly one element, called $f(x)$, in another set. Unless we say otherwise, the **domain** is the set of all real numbers for which the rule makes sense and defines a real number. All possible values of $f(x)$ are called the **range** of f. We use four ways to represent a function.

- By a graph
- By a table of values

- By an explicit formula
- By a verbal description

Throughout the book we will use several representations of any given function to help give us a better understanding of the problem. The graphs in Figure 1.1 contain most of the base functions we can use to build other functions using transformations.

$$y = x^2 \qquad y = x^3 \qquad y = e^x \qquad y = \sin x \qquad y = |x|$$

$$y = \sqrt{x} \qquad y = \sqrt[3]{x} \qquad y = \ln x \qquad y = \cos x \qquad y = \frac{1}{x}$$
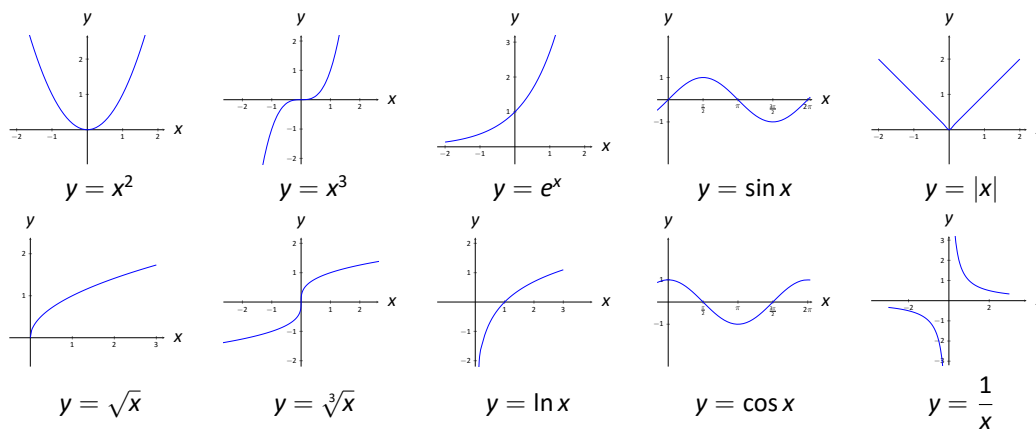
Figure 1.1: Basic Function Graphs

We will often transform these functions into other functions as given in the next two figures.

The function	shifts $f(x)$
$y = f(x) + c$	c units upward
$y = f(x) - c$	c units downward
$y = f(x + c)$	c units left
$y = f(x - c)$	c units right

Figure 1.2: Translations of Basic Functions with $c > 0$

The function	transforms $f(x)$ by
$y = cf(x)$	stretching vertically by a factor of c
$y = \frac{1}{c}f(x)$	shrinking vertically by a factor of c
$y = f(cx)$	shrinking horizontally by a factor of c
$y = f(\frac{x}{c})$	stretching horizontally by a factor of c
$y = -f(x)$	reflecting about the x-axis
$y = f(-x)$	reflecting about the y-axis

Figure 1.3: Scaling Basic Functions with $c > 1$

Domain

We said above that domain is the set of real numbers for which the function (rule) defines a real number and makes sense. Ask yourself, "what values can I put into the function and get a real value out?" There are generally two key expressions that will limit the domain of a function from all real numbers. We may not divide by zero and we may not have a negative number underneath an even root. The following examples illustrate how we restrict the domain when we see these expressions.

Example 1   **Finding a domain**
Find the domain of the function $f(x) = \sqrt{x - 4}$.

   SOLUTION   The square root of a negative number is not defined as a real number so the domain of f will be all real numbers for which $x - 4 \geq 0$ which is $x \geq 4$. In interval notation, this is $[4, \infty)$.

Example 2   **Finding a domain**
Find the domain of the function $g(x) = \dfrac{3}{x^2 - 9}$.

Notes:

SOLUTION   We cannot divide by zero so we factor the denominator of g and exclude those values where the denominator is zero.

$$g(x) = \frac{3}{x^2 - 9} = \frac{3}{(x-3)(x+3)}$$

We see that $x \neq 3, -3$ for g to be defined, which is written in interval notation as $(-\infty, -3) \cup (-3, 3) \cup (3, \infty)$.

Example 3   **Finding a domain**

Find the domain of the function $h(x) = \dfrac{1}{\sqrt{x^2 - 4}}$

SOLUTION   For h to be defined as a real number we must have $x^2 - 4 > 0$. This is equivalent to $(x-2)(x+2) > 0$ and we create a sign chart:
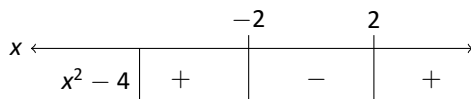
This shows that $x^2 - 4$ will be greater than zero on $(-\infty, -2) \cup (2, \infty)$.

Notes:

Exercises 1.0

Problems

In Exercises 1–10, find the domain of the given function.

1. $g(x) = (x - 3)^2 + 5$

2. $f(x) = \sqrt{x + 7} - 3$

3. $f(x) = \sqrt{x^2 - 6x - 7}$

4. $f(x) = 3|x - 2| + 4$

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

6. $g(x) = \dfrac{x - 3}{x^2 - x + 6}$

7. $h(x) = \sin(x + 3\pi)$

8. $f(x) = \dfrac{4x + 1}{\sqrt{x^2 - 4}}$

9. $h(x) = \dfrac{\cos x}{x}$

10. $g(x) = |x^2 - x - 6|$

In Exercises 11–14, graph the given f.

11. $f(x) = \begin{cases} x^2 - 3 & x < 2 \\ x - 4 & x \geq 2 \end{cases}$

12. $f(x) = \begin{cases} 3 & x \leq -1 \\ 2 - x^2 & -1 < x < 2 \\ -3 & x \geq 2 \end{cases}$

13. $f(x) = \begin{cases} x + 3 & x < -2 \\ x^2 + 4 & -2 \leq x \leq 3 \\ e^{-x} & x > 3 \end{cases}$

14. $f(x) = \begin{cases} \sin x & x \leq 0 \\ \frac{1}{2}x + 1 & x > 0 \end{cases}$

In Exercises 15–17, evaluate the expressions for the given f.

15. $f(x) = 3x^2 - 2x + 6$
    - (a) $f(2)$
    - (b) $f(-1)$
    - (c) $f(a)$
    - (d) $f(x + h)$
    - (e) $\dfrac{f(x + h) - f(x)}{h}$

16. $f(x) = \sqrt{x - 2}$
    - (a) $f(4)$
    - (b) $f(-3)$
    - (c) $f(t)$
    - (d) $f(x + h)$
    - (e) $\dfrac{f(x + h) - f(x)}{h}$

17. $f(x) = \dfrac{1}{x}$
    - (a) $f(-1)$
    - (b) $f(9)$
    - (c) $f(t + 3)$
    - (d) $f(x + h)$
    - (e) $\dfrac{f(x + h) - f(x)}{h}$

In Exercises 18–21, use sign diagrams to find the solutions to the nonlinear inequalities.

18. $\dfrac{(x - 2.13)(x - 2.12)^2}{(2.15 - x)(x - 2.14)^3} \leq 0$

19. $\dfrac{(5.678 - x)^3(x - 5.677)}{(x - 5.679)^2} \leq 0$

20. $\dfrac{1}{x - 0.3} \geq 2$

21. $\dfrac{x}{0.1 - x} \leq -2$

1: LIMITS

Calculus means "a method of calculation or reasoning." When one computes the sales tax on a purchase, one employs a simple calculus. When one finds the area of a polygonal shape by breaking it up into a set of triangles, one is using another calculus. Proving a theorem in geometry employs yet another calculus.

Despite the wonderful advances in mathematics that had taken place into the first half of the 17th century, mathematicians and scientists were keenly aware of what they *could not do.* (This is true even today.) In particular, two important concepts eluded mastery by the great thinkers of that time: area and rates of change.

Area seems innocuous enough; areas of circles, rectangles, parallelograms, etc., are standard topics of study for students today just as they were then. However, the areas of *arbitrary* shapes could not be computed, even if the boundary of the shape could be described exactly.

Rates of change were also important. When an object moves at a constant rate of change, then "distance = rate $\times$ time." But what if the rate is not constant — can distance still be computed? Or, if distance is known, can we discover the rate of change?

It turns out that these two concepts were related. Two mathematicians, Sir Isaac Newton and Gottfried Leibniz, are credited with independently formulating a system of computing that solved the above problems and showed how they were connected. Their system of reasoning was "a" calculus. However, as the power and importance of their discovery took hold, it became known to many as "the" calculus. Today, we generally shorten this to discuss "calculus."

The foundation of "the calculus" is the *limit.* It is a tool to describe a particular behavior of a function. This chapter begins our study of the limit by approximating its value graphically and numerically. After a formal definition of the limit, properties are established that make "finding limits" tractable. Once the limit is understood, then the problems of area and rates of change can be approached.

1.1 An Introduction To Limits

We begin our study of *limits* by considering examples that demonstrate key concepts that will be explained as we progress.

Consider the function $y = \frac{\sin x}{x}$. When x is near the value 1, what value (if any) is y near?

While our question is not precisely formed (what constitutes "near the value 1"?), the answer does not seem difficult to find. One might think first to look

at a graph of this function to approximate the appropriate y values. Consider Figure 1.4(a), where $y = \frac{\sin x}{x}$ is graphed. For values of x near 1, it seems that y takes on values near 0.85. In fact, when $x = 1$, then $y = \frac{\sin 1}{1} \approx 0.84$, so it makes sense that when x is "near" 1, y will be "near" 0.84.

Consider this again at a different value for x. When x is near 0, what value (if any) is y near? By considering Figure 1.4(b), one can see that it seems that y takes on values near 1. But what happens when $x = 0$? We have

$$y \to \frac{\sin 0}{0} \to {}^{\text{``}} \frac{0}{0} {}^{\text{''}} .$$

The expression "0/0" has no value; it is *indeterminate.* Such an expression gives no information about what is going on with the function nearby. We cannot find out how y behaves near $x = 0$ for this function simply by letting $x = 0$.

Finding a limit entails understanding how a function behaves near a particular value of x. Before continuing, it will be useful to establish some notation. Let $y = f(x)$; that is, let y be a function of x for some function f. The expression "the limit of y as x approaches 1" describes a number, often referred to as L, that y nears as x nears 1. We write all this as

$$\lim_{x \to 1} y = \lim_{x \to 1} f(x) = L.$$

This is not a complete definition (that will come in the next section); this is a pseudo-definition that will allow us to explore the idea of a limit.

Above, where $f(x) = \sin(x)/x$, we approximated

$$\lim_{x \to 1} \frac{\sin x}{x} \approx 0.84 \quad \text{and} \quad \lim_{x \to 0} \frac{\sin x}{x} \approx 1.$$

(We *approximated* these limits, hence used the "$\approx$" symbol, since we are working with the pseudo-definition of a limit, not the actual definition.)

Once we have the true definition of a limit, we will find limits *analytically*; that is, exactly using a variety of mathematical tools. For now, we will *approximate* limits both graphically and numerically. Graphing a function can provide a good approximation, though often not very precise. Numerical methods can provide a more accurate approximation. We have already approximated limits graphically, so we now turn our attention to numerical approximations.

Consider again $\lim_{x \to 1} \sin(x)/x$. To approximate this limit numerically, we can create a table of x and $f(x)$ values where x is "near" 1. This is done in Figure 1.5(a).

Notice that for values of x near 1, we have $\sin(x)/x$ near 0.841. The $x = 1$ row is in bold to highlight the fact that when considering limits, we are *not* concerned with the value of the function at that particular x value; we are only concerned with the values of the function when x is *near* 1.

(a)

(b)

Figure 1.4: $\sin(x)/x$ near $x = 1$ (top) and $x = 0$ (bottom).

x	$\sin(x)/x$
0.9	0.870363
0.99	0.844471
0.999	0.841772
1	**0.841471**
1.001	0.841170
1.01	0.838447
1.1	0.810189

(a)

x	$\sin(x)/x$
-0.1	0.9983341665
-0.01	0.9999833334
-0.001	0.9999998333
0	**not defined**
0.001	0.9999998333
0.01	0.9999833334
0.1	0.9983341665

(b)

Figure 1.5: Values of $\sin(x)/x$ with x near 1 and near 0.

Notes:

Now approximate $\lim_{x\to 0} \sin(x)/x$ numerically. We already approximated the value of this limit as 1 graphically in Figure 1.4(b). The table in Figure 1.5(b) shows the value of $\sin(x)/x$ for values of x near 0. Ten places after the decimal point are shown to highlight how close to 1 the value of $\sin(x)/x$ gets as x takes on values very near 0. We include the $x = 0$ row in bold again to stress that we are not concerned with the value of our function at $x = 0$, only on the behavior of the function *near* 0.

This numerical method gives confidence to say that 1 is a good approximation of $\lim_{x\to 0} \sin(x)/x$; that is,

$$\lim_{x\to 0} \sin(x)/x \approx 1.$$

Later we will be able to prove that the limit is *exactly* 1.

Watch the video:
Introduction to limits at
`https://youtu.be/riXcZT2ICjA`

We now consider several examples that allow us to explore different aspects of the limit concept.

Example 1   **Approximating the value of a limit**
Use graphical and numerical methods to approximate

$$\lim_{x\to 3} \frac{x^2 - x - 6}{6x^2 - 19x + 3}.$$

SOLUTION   To graphically approximate the limit, graph

$$y = (x^2 - x - 6)/(6x^2 - 19x + 3)$$

on a small interval that contains 3. To numerically approximate the limit, create a table of values where the x values are near 3. This is done in Figure 1.6.

The graph shows that when x is near 3, the value of y is very near 0.3. By considering values of x near 3, we see that $y = 0.294$ is a better approximation. The graph and the table imply that

$$\lim_{x\to 3} \frac{x^2 - x - 6}{6x^2 - 19x + 3} \approx 0.294.$$
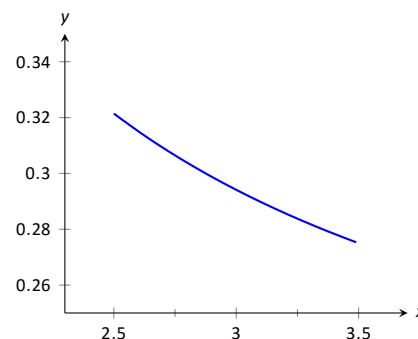
(a)

x	$\frac{x^2 - x - 6}{6x^2 - 19x + 3}$
2.9	0.298780
2.99	0.294569
2.999	0.294163
3	**not defined**
3.001	0.294073
3.01	0.293669
3.1	0.289773

(b)

Figure 1.6: Graphically and numerically approximating a limit in Example 1.

Notes:

This example may bring up a few questions about approximating limits (and the nature of limits themselves).

1. If a graph does not produce as good an approximation as a table, why bother with it?

2. How many values of x in a table are "enough?" In the previous example, could we have just used $x = 3.001$ and found a fine approximation?

Graphs are useful since they give a visual understanding concerning the behavior of a function. Sometimes a function may act "erratically" near certain x values which is hard to discern numerically but very plain graphically. Since graphing utilities are very accessible, it makes sense to make proper use of them.

Since tables and graphs are used only to *approximate* the value of a limit, there is not a firm answer to how many data points are "enough." Include enough so that a trend is clear, and use values (when possible) both less than and greater than the value in question. In Example 1, we used both values less than and greater than 3. Had we used just $x = 3.001$, we might have been tempted to conclude that the limit had a value of 0.3. While this is not far off, we could do better. Using values "on both sides of 3" helps us identify trends.

Example 2 **Approximating the value of a limit**
Graphically and numerically approximate the limit of $f(x)$ as x approaches 0, where

$$f(x) = \begin{cases} x + 1 & x < 0 \\ -x^2 + 1 & x > 0 \end{cases}.$$

SOLUTION Again we graph $f(x)$ and create a table of its values near $x = 0$ to approximate the limit. Note that this is a piecewise defined function, so it behaves differently on either side of 0. Figure 1.7(a) shows a graph of $f(x)$, and on either side of 0 it seems the y values approach 1. Note that $f(0)$ is not actually defined, as indicated in the graph with the open circle.

The table shown in Figure 1.7(b) shows values of $f(x)$ for values of x near 0. It is clear that as x takes on values very near 0, $f(x)$ takes on values very near 1. It turns out that if we let $x = 0$ for either "piece" of $f(x)$, 1 is returned; this is significant and we'll return to this idea later.

The graph and table allow us to say that $\lim_{x \to 0} f(x) \approx 1$; in fact, we are probably very sure it *equals* 1.

(a)

x	$f(x)$
-0.1	0.9
-0.01	0.99
-0.001	0.999
0.001	0.999999
0.01	0.9999
0.1	0.99

(b)

Figure 1.7: Graphically and numerically approximating a limit in Example 2.

Identifying When Limits Do Not Exist

A function may not have a limit for all values of x. That is, we may not be able to say $\lim_{x \to c} f(x) = L$ for some numbers L for all values of c, because there may not

Notes:

be a number that $f(x)$ is approaching. There are three ways in which a limit may fail to exist.

1. The function $f(x)$ may approach different values on either side of c.

2. The function may grow without upper or lower bound as x approaches c.

3. The function may oscillate as x approaches c.

We'll explore each of these in turn.

Example 3 — Different Values Approached From Left and Right
Explore why $\lim_{x\to1} f(x)$ does not exist, where

$$f(x) = \begin{cases} x^2 - 2x + 3 & x \le 1 \\ x & x > 1 \end{cases}.$$

SOLUTION A graph of $f(x)$ around $x = 1$ and a table are given in Figure 1.8. It is clear that as x approaches 1, $f(x)$ does not seem to approach a single number. Instead, it seems as though $f(x)$ approaches two different numbers. When considering values of x less than 1 (approaching 1 from the left), it seems that $f(x)$ is approaching 2; when considering values of x greater than 1 (approaching 1 from the right), it seems that $f(x)$ is approaching 1. Recognizing this behavior is important; we'll study this in greater depth later. Right now, it suffices to say that the limit does not exist since $f(x)$ is not approaching one value as x approaches 1.

Example 4 — The Function Grows Without Bound
Explore why $\lim_{x\to1} 1/(x-1)^2$ does not exist.

SOLUTION A graph and table of $f(x) = 1/(x-1)^2$ are given in Figure 1.9. Both show that as x approaches 1, $f(x)$ grows larger and larger.

We can deduce this on our own, without the aid of the graph and table. If x is near 1, then $(x-1)^2$ is very small, and:

$$\frac{1}{\text{very small number}} = \text{very large number}.$$

Since $f(x)$ is not approaching a single number, we conclude that $\lim_{x\to1} \dfrac{1}{(x-1)^2}$ does not exist.

Notes:

(a)

x	$f(x)$
0.9	2.01
0.99	2.0001
0.999	2.000001
1.001	1.001
1.01	1.01
1.1	1.1

(b)

Figure 1.8: Graphically and numerically observing no limit as $x \to 1$ in Example 3.
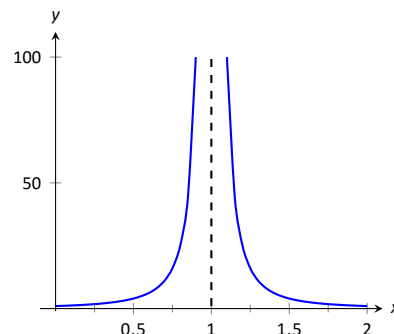
(a)

x	$f(x)$
0.9	100
0.99	10000
0.999	1×10^6
1.001	1×10^6
1.01	10000
1.1	100

(b)

Figure 1.9: Graphically and numerically observing no limit as $x \to 1$ in Example 4.

Example 5   The Function Oscillates

Explore why $\lim\limits_{x\to 0} \sin(1/x)$ does not exist.

SOLUTION   Two graphs of $f(x) = \sin(1/x)$ are given in Figure 1.10. Figure 1.10(a) shows $f(x)$ on the interval $[-1, 1]$; notice how $f(x)$ seems to oscillate near $x = 0$. One might think that despite the oscillation, as x approaches 0, $f(x)$ approaches 0. However, Figure 1.10(b) zooms in on $\sin(1/x)$, on the interval $[-0.1, 0.1]$. Here the oscillation is even more pronounced. Finally, in the table in Figure 1.10(c), we see $\sin(x)/x$ evaluated for values of x near 0. As x approaches 0, $f(x)$ does not appear to approach any value.

It can be shown that in reality, as x approaches 0, $\sin(1/x)$ takes on all values between -1 and 1 infinitely many times! Because of this oscillation, $\lim\limits_{x\to 0} \sin(1/x)$ does not exist.

x	$\sin(1/x)$
0.1	-0.544021
0.01	-0.506366
0.001	0.826880
0.0001	-0.305614
1.0×10^{-5}	0.035749
1.0×10^{-6}	-0.349994
1.0×10^{-7}	0.420548

(a)   (b)   (c)
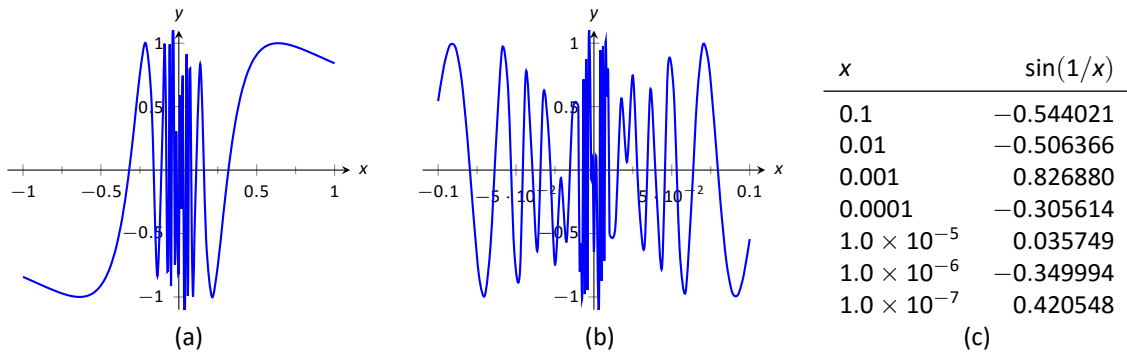
Figure 1.10: Observing that $f(x) = \sin(1/x)$ has no limit as $x \to 0$ in Example 5.

Limits of Difference Quotients

We have approximated limits of functions as x approached a particular number. We will consider another important kind of limit after explaining a few key ideas.

Let $f(x)$ represent the position function, in feet, of some particle that is moving in a straight line, where x is measured in seconds. Let's say that when $x = 1$, the particle is at position 10 ft., and when $x = 5$, the particle is at 20 ft. Another way of expressing this is to say

$$f(1) = 10 \quad \text{and} \quad f(5) = 20.$$

Since the particle traveled 10 feet in 4 seconds, we can say the particle's *average velocity* was 2.5 ft/s. We write this calculation using a "quotient of differences," or, a *difference quotient*:

$$\frac{f(5) - f(1)}{5 - 1} = \frac{10}{4} = 2.5\text{ft/s}.$$
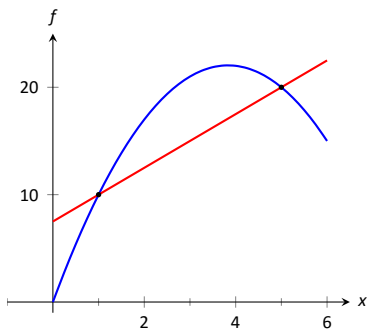
Figure 1.11: Interpreting a difference quotient as the slope of a secant line.

Notes:

This difference quotient can be thought of as the familiar "rise over run" used to compute the slopes of lines. In fact, that is essentially what we are doing: given two points on the graph of f, we are finding the slope of the *secant line* through those two points. See Figure 1.11.

Now consider finding the average speed on another time interval. We again start at $x = 1$, but consider the position of the particle h seconds later. That is, consider the positions of the particle when $x = 1$ and when $x = 1 + h$. The difference quotient is now

$$\frac{f(1+h) - f(1)}{(1+h) - 1} = \frac{f(1+h) - f(1)}{h}.$$

Let $f(x) = -1.5x^2 + 11.5x$; note that $f(1) = 10$ and $f(5) = 20$, as in our discussion. We can compute this difference quotient for all values of h (even negative values!) except $h = 0$, for then we get "0/0," the indeterminate form introduced earlier. For all values $h \neq 0$, the difference quotient computes the average velocity of the particle over an interval of time of length h starting at $x = 1$.

For small values of h, i.e., values of h close to 0, we get average velocities over very short time periods and compute secant lines over small intervals. See Figure 1.12. This leads us to wonder what the limit of the difference quotient is as h approaches 0. That is,

$$\lim_{h \to 0} \frac{f(1+h) - f(1)}{h} = \text{?}$$

As we do not yet have a true definition of a limit nor an exact method for computing it, we settle for approximating the value. While we could graph the difference quotient (where the x-axis would represent h values and the y-axis would represent values of the difference quotient) we settle for making a table. See Figure 1.13. The table gives us reason to assume the value of the limit is about 8.5.

Proper understanding of limits is key to understanding calculus. With limits, we can accomplish seemingly impossible mathematical things, like adding up an infinite number of numbers (and not get infinity) and finding the slope of a line between two points, where the "two points" are actually the same point. These are not just mathematical curiosities; they allow us to link position, velocity and acceleration together, connect cross-sectional areas to volume, find the work done by a variable force, and much more.

In the next section we give the formal definition of the limit and begin our study of finding limits analytically. In the following exercises, we continue our introduction and approximate the value of limits.

Notes:

(a)

(b)

(c)
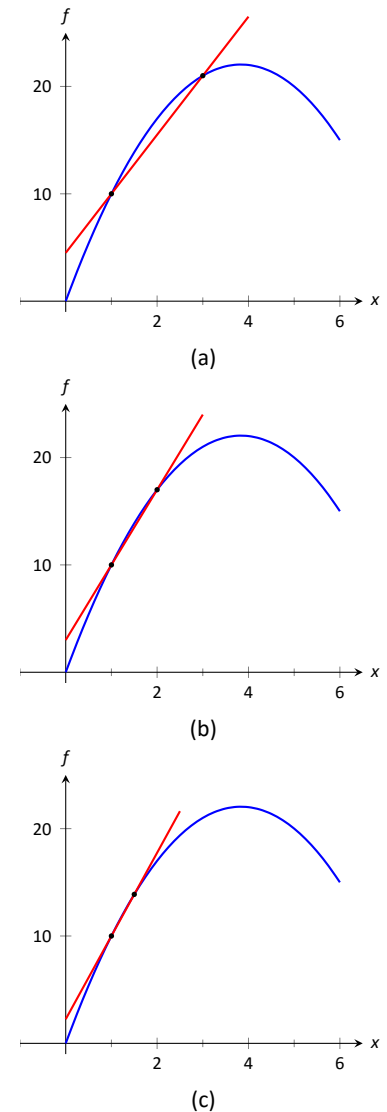
Figure 1.12: Secant lines of $f(x)$ at $x = 1$ and $x = 1 + h$, for shrinking values of h (i.e., $h \to 0$).

h	$\frac{f(1+h)-f(1)}{h}$
-0.5	9.250
-0.1	8.650
-0.01	8.515
0.01	8.485
0.1	8.350
0.5	7.750

Figure 1.13: The difference quotient for $f(x) = -1.5x^2 + 11.5x$ evaluated at values of h near 0.

Exercises 1.1

Terms and Concepts

1. In your own words, what does it mean to "find the limit of $f(x)$ as x approaches 3"?
2. An expression of the form $\frac{0}{0}$ is called ____.
3. T/F: The limit of $f(x)$ as x approaches 5 is $f(5)$.
4. Describe three situations where $\lim\limits_{x \to c} f(x)$ does not exist.
5. In your own words, what is a difference quotient?

Problems

In Exercises 6–15, approximate the given limits both numerically and graphically.

6. $\lim\limits_{x \to 1} x^2 + 3x - 5$

7. $\lim\limits_{x \to 0} x^3 - 3x^2 + x - 5$

8. $\lim\limits_{x \to 0} \dfrac{x + 1}{x^2 + 3x}$

9. $\lim\limits_{x \to 3} \dfrac{x^2 - 2x - 3}{x^2 - 4x + 3}$

10. $\lim\limits_{x \to -1} \dfrac{x^2 + 8x + 7}{x^2 + 6x + 5}$

11. $\lim\limits_{x \to 2} \dfrac{x^2 + 7x + 10}{x^2 - 4x + 4}$

12. $\lim\limits_{x \to 2} f(x)$, where $f(x) = \begin{cases} x + 2 & x \le 2 \\ 3x - 5 & x > 2 \end{cases}$.

13. $\lim\limits_{x \to 3} f(x)$, where $f(x) = \begin{cases} x^2 - x + 1 & x \le 3 \\ 2x + 1 & x > 3 \end{cases}$.

14. $\lim\limits_{x \to 0} f(x)$, where $f(x) = \begin{cases} \cos x & x \le 0 \\ x^2 + 3x + 1 & x > 0 \end{cases}$.

15. $\lim\limits_{x \to \pi/2} f(x)$, where $f(x) = \begin{cases} \sin x & x \le \pi/2 \\ \cos x & x > \pi/2 \end{cases}$.

In Exercises 16–24, a function f and a value a are given. Approximate the limit of the difference quotient, $\lim\limits_{h \to 0} \dfrac{f(a + h) - f(a)}{h}$, using $h = \pm 0.1, \pm 0.01$.

16. $f(x) = -7x + 2, \quad a = 3$

17. $f(x) = 9x + 0.06, \quad a = -1$

18. $f(x) = x^2 + 3x - 7, \quad a = 1$

19. $f(x) = \dfrac{1}{x + 1}, \quad a = 2$

20. $f(x) = -4x^2 + 5x - 1, \quad a = -3$

21. $f(x) = \ln x, \quad a = 5$

22. $f(x) = \sin x, \quad a = \pi$

23. $f(x) = \cos x, \quad a = \pi$

24. $f(x) = \sqrt{x + 4}, \quad a = 0$

1.2 Epsilon-Delta Definition of a Limit

This section introduces the formal definition of a limit. Many refer to this as "the epsilon–delta," definition, referring to the letters ε and δ of the Greek alphabet.

Before we give the actual definition, let's consider a few informal ways of describing a limit. Given a function $y = f(x)$ and an *x*-value, *c*, we say that "the limit of the function *f*, as *x* approaches *c*, is a value *L*":

1. if "*y* tends to *L*" as "*x* tends to *c*."

2. if "*y* approaches *L*" as "*x* approaches *c*."

3. if "*y* is near *L*" whenever "*x* is near *c*."

The problem with these definitions is that the words "tends," "approach," and especially "near" are not exact. In what way does the variable *x* tend to, or approach, *c*? How near do *x* and *y* have to be to *c* and *L*, respectively?

The definition we describe in this section comes from formalizing **3**. A quick restatement gets us closer to what we want:

3′. If *x* is within a certain *tolerance level* of *c*, then the corresponding value $y = f(x)$ is within a certain *tolerance level* of *L*.

The traditional notation for the *x*-tolerance is the lowercase Greek letter delta, or δ, and the *y*-tolerance is denoted by lowercase epsilon, or ε. One more rephrasing of **3′** nearly gets us to the actual definition:

3″. If *x* is within δ units of *c*, then the corresponding value of *y* is within ε units of *L*.

We can write "*x* is within δ units of *c*" mathematically as

$$|x - c| < \delta, \qquad \text{which is equivalent to} \qquad c - \delta < x < c + \delta.$$

Letting the symbol "$\longrightarrow$" represent the word "implies," we can rewrite **3″** as

$$|x - c| < \delta \longrightarrow |y - L| < \varepsilon \qquad \text{or} \qquad c - \delta < x < c + \delta \longrightarrow L - \varepsilon < y < L + \varepsilon.$$

The point is that δ and ε, being tolerances, can be any positive (but typically small) values. Finally, we have the formal definition of the limit with the notation seen in the previous section.

Notes:

Definition 1   The Limit of a Function f

Let I be an open interval containing c, and let f be a function defined on I, except possibly at c. The **limit of $f(x)$, as x approaches c, is** L, denoted by
$$\lim_{x \to c} f(x) = L,$$
means that given any $\varepsilon > 0$, there exists $\delta > 0$ such that for all $x \neq c$, if $|x - c| < \delta$, then $|f(x) - L| < \varepsilon$.

(Mathematicians often enjoy writing ideas without using any words. Here is the wordless definition of the limit:

$$\lim_{x \to c} f(x) = L \iff \forall\, \varepsilon > 0, \exists\, \delta > 0 \text{ s.t. } 0 < |x - c| < \delta \longrightarrow |f(x) - L| < \varepsilon.)$$

Note the order in which ε and δ are given. In the definition, the y-tolerance ε is given *first* and then the limit will exist **if** we can find an x-tolerance δ that works.

Watch the video:
Limits 1b — Delta-Epsilon Formulation at
`https://youtu.be/v5zsbgYrunM`

An example will help us understand this definition. Note that the explanation is long, but it will go through all steps necessary to understand the ideas.

Example 1   Evaluating a limit using the definition
Show that $\lim_{x \to 4} \sqrt{x} = 2$.

SOLUTION   Before we use the formal definition, let's try some numerical tolerances. What if the y tolerance is 0.5, or $\varepsilon = 0.5$? How close to 4 does x have to be so that y is within 0.5 units of 2, i.e., $1.5 < y < 2.5$? In this case, we can proceed as follows:
$$
\begin{aligned}
1.5 &< y < 2.5 \\
1.5 &< \sqrt{x} < 2.5 \\
1.5^2 &< x < 2.5^2 \\
2.25 &< x < 6.25.
\end{aligned}
$$

Notes:

So, what is the desired *x* tolerance? Remember, we want to find a symmetric interval of *x* values, namely $4 - \delta < x < 4 + \delta$. The lower bound of 2.25 is 1.75 units from 4; the upper bound of 6.25 is 2.25 units from 4. We need the smaller of these two distances; we must have $\delta \leq 1.75$. See Figure 1.14.

Given the *y* tolerance $\varepsilon = 0.5$, we have found an *x* tolerance, $\delta \leq 1.75$, such that whenever *x* is within δ units of 4, then *y* is within ε units of 2. That's what we were trying to find.

Let's try another value of ε.

What if the *y* tolerance is 0.01, i.e., $\varepsilon = 0.01$? How close to 4 does *x* have to be in order for *y* to be within 0.01 units of 2 (or $1.99 < y < 2.01$)? Again, we just square these values to get $1.99^2 < x < 2.01^2$, or

$$3.9601 < x < 4.0401.$$

What is the desired *x* tolerance? In this case we must have $\delta \leq 0.0399$, which is the minimum distance from 4 of the two bounds given above.

What we have so far: if $\varepsilon = 0.5$, then $\delta \leq 1.75$ and if $\varepsilon = 0.01$, then $\delta \leq 0.0399$. A pattern is not easy to see, so we switch to general ε try to determine δ symbolically. We start by assuming $y = \sqrt{x}$ is within ε units of 2:

$$
\begin{aligned}
|y - 2| &< \varepsilon & \\
-\varepsilon < y - 2 &< \varepsilon & \text{(Definition of absolute value)} \\
-\varepsilon < \sqrt{x} - 2 &< \varepsilon & (y = \sqrt{x}) \\
2 - \varepsilon < \sqrt{x} &< 2 + \varepsilon & \text{(Add 2)} \\
(2 - \varepsilon)^2 < x &< (2 + \varepsilon)^2 & \text{(Square all)} \\
4 - 4\varepsilon + \varepsilon^2 < x &< 4 + 4\varepsilon + \varepsilon^2 & \text{(Expand)} \\
4 - (4\varepsilon - \varepsilon^2) < x &< 4 + (4\varepsilon + \varepsilon^2) & \text{(Rewrite in the desired form)} \\
-(4\varepsilon - \varepsilon^2) < x - 4 &< (4\varepsilon + \varepsilon^2) & \text{(Rewrite in the desired form)}
\end{aligned}
$$

The "desired form" in the last step is "$-something < x - 4 < something$." Since we want this last interval to describe an *x* tolerance around 4, we have that either $\delta \leq 4\varepsilon - \varepsilon^2$ or $\delta \leq 4\varepsilon + \varepsilon^2$, whichever is smaller:

$$\delta \leq \min\{4\varepsilon - \varepsilon^2, 4\varepsilon + \varepsilon^2\}.$$

Since $\varepsilon > 0$, the minimum is $\delta \leq 4\varepsilon - \varepsilon^2$. That's the formula: given an ε, set $\delta \leq 4\varepsilon - \varepsilon^2$.

Choose $\varepsilon > 0$. Then ...

... choose δ smaller than each of these

width = 1.75   width = 2.25
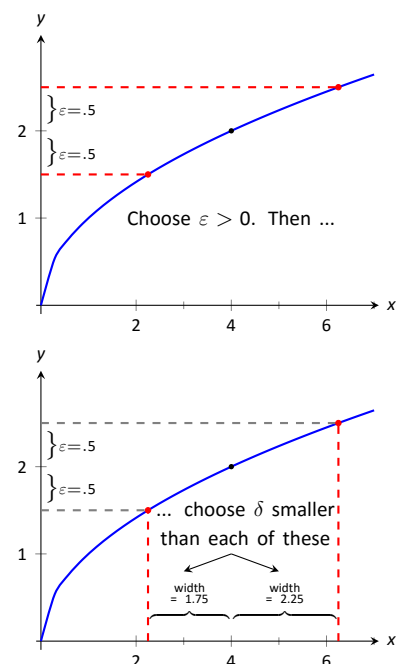
With $\varepsilon = 0.5$, we pick any $\delta < 1.75$.

Figure 1.14: Illustrating the $\varepsilon - \delta$ process.

Notes:

We can check this for our previous values. If $\varepsilon = 0.5$, the formula gives $\delta \le 4(0.5) - (0.5)^2 = 1.75$ and when $\varepsilon = 0.01$, the formula gives $\delta \le 4(0.01) - (0.01)^2 = 0.399$.

So given any $\varepsilon > 0$, set $\delta \le 4\varepsilon - \varepsilon^2$. Then if $|x - 4| < \delta$ (and $x \ne 4$), then $|f(x) - 2| < \varepsilon$, satisfying the definition of the limit. We have shown formally (and finally!) that $\lim\limits_{x \to 4} \sqrt{x} = 2$.

The previous example was a little long in that we sampled a few specific cases of ε before handling the general case. Normally this is not done. The previous example is also a bit unsatisfying in that $\sqrt{4} = 2$; why work so hard to prove something so obvious? Many ε-δ proofs are long and difficult to do. In this section, we will focus on examples where the answer is, frankly, obvious, because the non–obvious examples are even harder. In the next section we will learn some theorems that allow us to evaluate limits *analytically*, that is, without using the ε-δ definition.

We will follow a general pattern to work through δ-ε problems. In some sense, each starts out "backwards." That is, while we want to

1. start with $|x - c| < \delta$ and conclude that

2. $|f(x) - L| < \varepsilon$,

we actually start by assuming

1. $|f(x) - L| < \varepsilon$, then perform some algebraic manipulations to give an inequality of the form

2. $|x - c| < something$.

When we have properly done this, the *something* on the "greater than" side of the inequality becomes our δ. We can refer to this as the "scratch–work" phase of our proof. Once we have δ, we can formally start with $|x - c| < \delta$ and use algebraic manipulations to conclude that $|f(x) - L| < \varepsilon$, usually by using the same steps of our "scratch–work" in reverse order.

We will highlight this process in the following examples.

Example 2   Evaluating a limit using the definition
Show that $\lim\limits_{x \to 1} 3x - 5 = -2$

SOLUTION   Let's do this example symbolically from the start.
Scratch-Work:
We start our scratch-work by considering $|f(x) - (-2)| < \varepsilon$:

Notes:

$$|f(x) - (-2)| < \varepsilon$$
$$|3x - 5 + 2| < \varepsilon$$
$$|3x - 3| < \varepsilon$$
$$3|x - 1| < \varepsilon$$
$$|x - 1| < \frac{\varepsilon}{3}$$

This suggests that we set $\delta = \frac{\varepsilon}{3}$,

Proof
Given $\varepsilon > 0$, choose $\delta = \frac{\varepsilon}{3}$. We assume $|x - 1| < \delta$

$$|x - 1| < \delta$$
$$|x - 1| < \frac{\varepsilon}{3} \qquad \text{(Our choice of } \delta\text{)}$$
$$3|x - 1| < \frac{\varepsilon}{3} \cdot 3 \qquad \text{(Multiply by 3)}$$
$$|3x - 3| < \varepsilon \qquad \text{(Simplify)}$$
$$|3x - 5 + 2| < \varepsilon$$
$$|3x - 5 - (-2)| < \varepsilon,$$

which is what we wanted to show. Thus $\lim\limits_{x \to 1} 3x - 5 = -2$. $\qquad\qquad$ □

Example 3 $\qquad$ **Evaluating a limit using the definition**
Show that $\lim\limits_{x \to 2} 4 - \frac{3}{2}x = 1$.

SOLUTION $\qquad$ Scratch-Work:
We start our scratch-work by considering $|f(x) - 1| < \varepsilon$:

Notes:

$$|f(x) - 1| < \varepsilon$$

$$\left| 4 - \frac{3}{2}x - 1 \right| < \varepsilon$$

$$\left| 3 - \frac{3}{2}x \right| < \varepsilon$$

$$\left| -\frac{3}{2}(-2 + x) \right| < \varepsilon$$

$$\frac{3}{2}|x - 2| < \varepsilon$$

$$|x - 2| < \frac{2\varepsilon}{3}$$

This suggests that we set $\delta = \frac{2\varepsilon}{3}$,

Proof

Given $\varepsilon > 0$, choose $\delta = \dfrac{2\varepsilon}{3}$. We assume $|x - 2| < \delta$

$$|x - 2| < \delta$$

$$|x - 2| < \frac{2\varepsilon}{3}$$

$$\frac{3}{2}|x - 2| < \frac{2\varepsilon}{3} \cdot \frac{3}{2}$$

$$\left| -\frac{3}{2}(x - 2) \right| < \varepsilon$$

$$\left| -\frac{3}{2}x + 3 \right| < \varepsilon$$

$$\left| 4 - \frac{3}{2}x - 1 \right| < \varepsilon,$$

which is what we wanted to show. Thus $\lim\limits_{x \to 2} 4 - \dfrac{3}{2}x = 1$. $\square$

Example 4   **Evaluating a limit using the definition**

Show that $\lim\limits_{x \to 2} x^2 = 4$.

Notes:

SOLUTION   Scratch-Work: We start our scratch-work by considering $|f(x) - 4| < \varepsilon$:

$$|f(x) - 4| < \varepsilon$$
$$|x^2 - 4| < \varepsilon \qquad \text{(Now factor)}$$
$$|(x - 2)(x + 2)| < \varepsilon$$
$$|x - 2| < \frac{\varepsilon}{|x + 2|}. \qquad (1.1)$$

We are at the phase of saying that $|x - 2| < something$, where $something = \varepsilon/|x + 2|$. We want to turn that $something$ into δ. Could we not set $\delta = \dfrac{\varepsilon}{|x + 2|}$?

We are close to an answer, but the catch is that δ must be a *constant* value (so it can't contain x). There is a way to work around this, but we do have to make an assumption. Remember that ε is supposed to be a small number, which implies that δ will also be a small value. In particular, we can (probably) assume that $\delta < 1$. If this is true, then $|x - 2| < \delta$ would imply that $|x - 2| < 1$, giving $1 < x < 3$.

Now, back to the fraction $\dfrac{\varepsilon}{|x + 2|}$. If $1 < x < 3$, then $3 < x + 2 < 5$ (add 2 to all terms in the inequality). Taking reciprocals, we have

$$\frac{1}{5} < \frac{1}{|x + 2|} < \frac{1}{3} \qquad \text{which implies}$$
$$\frac{1}{5} < \frac{1}{|x + 2|} \qquad \text{which implies}$$
$$\frac{\varepsilon}{5} < \frac{\varepsilon}{|x + 2|}. \qquad (1.2)$$

This suggests that we set $\delta \leq \dfrac{\varepsilon}{5}$. This ends our scratch–work, and we begin the formal proof (which also helps us understand why this was a good choice of δ).

Proof
Given ε, let $\delta \leq \varepsilon/5$. We want to show that when $|x - 2| < \delta$, then $|x^2 - 4| < \varepsilon$.

Notes:

We start with $|x - 2| < \delta$:

$$|x - 2| < \delta$$
$$|x - 2| < \frac{\varepsilon}{5}$$
$$|x - 2| < \frac{\varepsilon}{5} < \frac{\varepsilon}{|x + 2|} \qquad \text{(for } x \text{ near 2, from Equation (1.2))}$$
$$|x - 2| \cdot |x + 2| < \varepsilon$$
$$|(x - 2)(x + 2)| < \varepsilon$$
$$|x^2 - 4| < \varepsilon,$$

which is what we wanted to show. Thus $\lim_{x \to 2} x^2 = 4$. □

We have arrived at $|x^2 - 4| < \varepsilon$ as desired. Note again, in order to make this happen we needed δ to first be less than 1. That is a safe assumption; we want ε to be arbitrarily small, forcing δ to also be small.

We have also picked δ to be smaller than "necessary." We could get by with a slightly larger δ, as shown in Figure 1.15. The dashed outer lines show the boundaries defined by our choice of ε. The dotted inner lines show the boundaries defined by setting $\delta = \varepsilon/5$. Note how these dotted lines are within the dashed lines. That is perfectly fine; by choosing x within the dotted lines we are guaranteed that $f(x)$ will be within ε of 4.

In summary, given $\varepsilon > 0$, set $\delta \leq \varepsilon/5$. Then $|x - 2| < \delta$ implies $|x^2 - 4| < \varepsilon$ (i.e. $|y - 4| < \varepsilon$) as desired. This shows that $\lim_{x \to 2} x^2 = 4$. Figure 1.15 gives a visualization of this; by restricting x to values within $\delta = \varepsilon/5$ of 2, we see that $f(x)$ is within ε of 4.

length of ε

length of $\delta = \varepsilon/5$
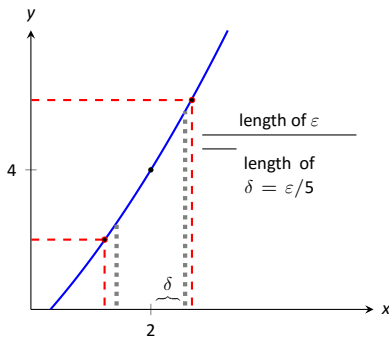
Figure 1.15: Choosing $\delta = \varepsilon/5$ in Example 4.

To better understand the definition of a limit, experiment with the Geogebra app at `http://ggbm.at/RtY27ybS`.

This formal definition of the limit is not an easy concept grasp. Our examples are actually "easy" examples, using "simple" functions like polynomials, square–roots and exponentials. It is very difficult to prove, using the techniques given above, that $\lim_{x \to 0} (\sin x)/x = 1$, as we approximated in the previous section.

There is hope. The next section shows how one can evaluate complicated limits using certain basic limits as building blocks. While limits are an incredibly

Notes:

important part of calculus (and hence much of higher mathematics), rarely are limits evaluated using the definition. Rather, the techniques of the following section are employed.

Notes:

Exercises 1.2

Terms and Concepts

1. What is wrong with the following "definition" of a limit?

   > "The limit of $f(x)$, as x approaches a, is K" means that given any $\delta > 0$ there exists $\varepsilon > 0$ such that whenever $|f(x) - K| < \varepsilon$, we have $|x - a| < \delta$.

2. Which is given first in establishing a limit, the x–tolerance or the y–tolerance?

3. T/F: ε must always be positive.

4. T/F: δ must always be positive.

Problems

5. Use the graph below of f to find a number δ such that if $0 < |x - 2| < \delta$, then $|f(x) - 1| < 0.5$.
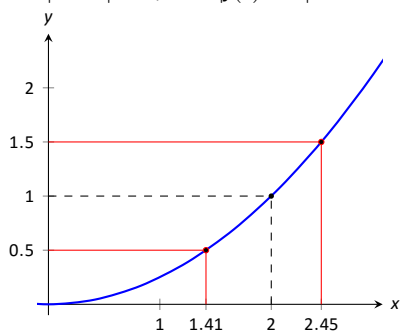
6. Use the graph below of f to find a number δ such that if $0 < |x - 2| < \delta$, then $|f(x) - 1| < 0.3$.
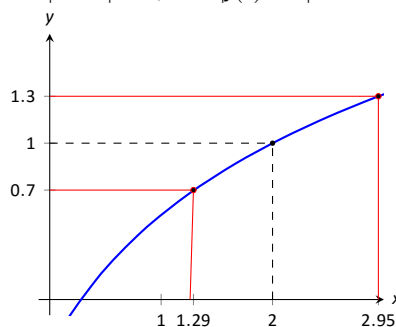
In Exercises 7–14, prove the given limit using an $\varepsilon - \delta$ proof.

7. $\displaystyle\lim_{x \to 5} 3 - x = -2$

8. $\displaystyle\lim_{x \to 5} 4x - 12 = 8$

9. $\displaystyle\lim_{x \to 3} 5 - 2x = -1$

10. $\displaystyle\lim_{x \to 3} x^2 - 3 = 6$

11. $\displaystyle\lim_{x \to 4} x^2 + x - 5 = 15$

12. $\displaystyle\lim_{x \to 2} x^3 - 1 = 7$

13. $\displaystyle\lim_{x \to 2} 5 = 5$

14. $\displaystyle\lim_{x \to 0} \sin x = 0$ (Hint: use the fact that $|\sin x| \leq |x|$, with equality only when $x = 0$.)

1.3 Finding Limits Analytically

In Section 1.1 we explored the concept of the limit without a strict definition, meaning we could only make approximations. In the previous section we gave the definition of the limit and demonstrated how to use it to verify our approximations were correct. Thus far, our method of finding a limit is (1) make a really good approximation either graphically or numerically, and (2) verify our approximation is correct using an ε-δ proof.

Recognizing that ε-δ proofs are cumbersome, this section gives a series of theorems which allow us to find limits much more quickly and intuitively.

Suppose that $\lim\limits_{x \to 2} f(x) = 2$ and $\lim\limits_{x \to 2} g(x) = 3$. What is $\lim\limits_{x \to 2} (f(x) + g(x))$? Intuition tells us that the limit should be 5, as we expect limits to behave in a nice way. The following theorem states that already established limits do behave nicely.

Theorem 1   Basic Limit Properties

Let b, c, L and K be real numbers, let n be a positive integer, and let f and g be functions with the following limits:

$$\lim_{x \to c} f(x) = L \text{ and } \lim_{x \to c} g(x) = K.$$

The following limits hold.

1. Constants: $\quad\quad\quad\quad \lim\limits_{x \to c} b = b$

2. Identity: $\quad\quad\quad\quad\quad \lim\limits_{x \to c} x = c$

3. Sums/Differences: $\quad \lim\limits_{x \to c} (f(x) \pm g(x)) = L \pm K$

4. Scalar Multiples: $\quad\quad \lim\limits_{x \to c} b \cdot f(x) = bL$

5. Products: $\quad\quad\quad\quad \lim\limits_{x \to c} f(x) \cdot g(x) = LK$

6. Quotients: $\quad\quad\quad\quad \lim\limits_{x \to c} f(x)/g(x) = L/K, \ (K \neq 0)$

7. Powers: $\quad\quad\quad\quad\quad \lim\limits_{x \to c} [f(x)]^n = L^n$

8. Roots: $\quad\quad\quad\quad\quad\quad \lim\limits_{x \to c} \sqrt[n]{f(x)} = \sqrt[n]{L} \quad$ (when n is odd or $L \geq 0$)

We will now prove the Sum Property using the formal definition of a limit

---

Notes:

from the previous section. We know that $\lim\limits_{x \to c} f(x) = L$ and $\lim\limits_{x \to c} g(x) = K$. We want to show that $\lim\limits_{x \to c} f(x) + g(x) = L + K$.

Proof
We must show that given any $\varepsilon > 0$, we can find a $\delta > 0$ such that

$$\text{if } 0 < |x - c| < \delta, \text{ then } |f(x) + g(x) - (L + K)| < \varepsilon.$$

We know $\lim\limits_{x \to c} f(x) = L$. So for any $\varepsilon_1 > 0$, we can find $\delta_1 > 0$ such that if $0 < |x - c| < \delta_1$, then $|f(x) - L| < \varepsilon_1$. Similarly we know $\lim\limits_{x \to c} g(x) = K$ so for any $\varepsilon_2 > 0$, we can find $\delta_2 > 0$ such that if $0 < |x - c| < \delta_2$, then $|g(x) - K| < \varepsilon_2$. We will let both ε_1 and ε_2 be $\frac{\varepsilon}{2}$. Now, we have a $\delta_1 > 0$ and a $\delta_2 > 0$ such that:

$$\text{if } 0 < |x - c| < \delta_1, \text{ then } |f(x) - L| < \frac{\varepsilon}{2}$$

$$\text{and}$$

$$\text{if } 0 < |x - c| < \delta_2, \text{ then } |g(x) - K| < \frac{\varepsilon}{2}$$

We will choose $\delta = \min(\delta_1, \delta_2) > 0$. If $0 < |x - c| < \delta$, then $|f(x) - L| < \frac{\varepsilon}{2}$ and $|g(x) - K| < \frac{\varepsilon}{2}$. Add the two inequalities together so that

$$|f(x) - L| + |g(x) - K| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

We will now use the triangle inequality: $|A + B| \leq |A| + |B|$.

$$|f(x) - L + g(x) - K| \leq |f(x) - L| + |g(x) - K| < \varepsilon$$

Thus $|(f(x) + g(x)) - (L + K)| < \varepsilon$, which is what we were trying to show.   □

The other Basic Limit Properties can be proven in a similar way and are left for the reader. Our next theorem requires a few more conditions.

Theorem 2   **Limits of Composition**
Suppose that

$$\lim\limits_{x \to c} f(x) = L \text{ and } \lim\limits_{x \to L} g(x) = g(L) = K.$$

Then $\lim\limits_{x \to c} g(f(x)) = K$.

Notes:

Watch the video:
Limit Laws to Evaluate a Limit, Example 1 at
`https://youtu.be/v_Nz6UUQ4HQ`

We apply the theorem to an example.

Example 1   **Using basic limit properties**
Let

$$\lim_{x \to 2} f(x) = 2, \quad \lim_{x \to 2} g(x) = 3 \quad \text{and} \quad p(x) = 3x^2 - 5x + 7.$$

Find the following limits:

1. $\lim_{x \to 2} \big(f(x) + g(x)\big)$ 3. $\lim_{x \to 2} p(x)$

2. $\lim_{x \to 2} \big(5f(x) + g(x)^2\big)$

SOLUTION

1. Using the Sum/Difference rule, we know that $\lim_{x \to 2} \big(f(x) + g(x)\big) = 2 + 3 = 5$.

2. Using the Scalar Multiple and Sum/Difference rules, we find that
$\lim_{x \to 2} \big(5f(x) + g(x)^2\big) = 5 \cdot 2 + 3^2 = 19$.

3. Here we combine the Power, Scalar Multiple, Sum/Difference and Constant Rules. We show quite a few steps, but in general these can be omitted:

$$\begin{aligned}
\lim_{x \to 2} p(x) &= \lim_{x \to 2} (3x^2 - 5x + 7) \\
&= \lim_{x \to 2} 3x^2 - \lim_{x \to 2} 5x + \lim_{x \to 2} 7 \\
&= 3 \cdot 2^2 - 5 \cdot 2 + 7 \\
&= 9.
\end{aligned}$$

Notes:

Part 3 of the previous example demonstrates how the limit of a quadratic polynomial can be determined using the properties of Theorem 1. Not only that, recognize that

$$\lim_{x \to 2} p(x) = 9 = p(2);$$

i.e., the limit at 2 was found just by plugging 2 into the function. This holds true for all polynomials, and also for rational functions (which are quotients of polynomials), as stated in the following theorem.

Theorem 3   Limits of Polynomial and Rational Functions

Let $p(x)$ and $q(x)$ be polynomials and c a real number. Then:

1. $\lim_{x \to c} p(x) = p(c)$

2. $\lim_{x \to c} \dfrac{p(x)}{q(x)} = \dfrac{p(c)}{q(c)}$, where $q(c) \neq 0$.

Example 2   Finding a limit of a rational function

Using Theorem 3, find

$$\lim_{x \to -1} \frac{3x^2 - 5x + 1}{x^4 - x^2 + 3}.$$

SOLUTION   Using Theorem 3, we can quickly state that

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

$$= \frac{9}{3} = 3.$$

It was likely frustrating in Section 1.2 to do a lot of work to prove that

$$\lim_{x \to 2} x^2 = 4$$

as it seemed fairly obvious. The previous theorems state that many functions behave in such an "obvious" fashion, as demonstrated by the rational function in Example 2.

Polynomial and rational functions are not the only functions to behave in such a predictable way. The following theorem gives a list of functions whose behavior is particularly "nice" in terms of limits. In the next section, we will give a formal name to these functions that behave "nicely."

Notes:

> **Theorem 4   Limits of Basic Functions**
>
> Let c be a real number in the domain of the given function and let n be a positive integer. The following limits hold:
>
> 1. $\lim\limits_{x \to c} \sin x = \sin c$
> 2. $\lim\limits_{x \to c} \cos x = \cos c$
> 3. $\lim\limits_{x \to c} \tan x = \tan c$
>
> 4. $\lim\limits_{x \to c} \csc x = \csc c$
> 5. $\lim\limits_{x \to c} \sec x = \sec c$
> 6. $\lim\limits_{x \to c} \cot x = \cot c$
>
> 7. $\lim\limits_{x \to c} a^x = a^c \ (a > 0)$
> 8. $\lim\limits_{x \to c} \ln x = \ln c$
> 9. $\lim\limits_{x \to c} \sqrt[n]{x} = \sqrt[n]{c}$

Many times, we will combine this theorem with Theorems 1 and 2. If our expression can be built up from the pieces in those theorems, then we can quickly evaluate the limit.

Example 3   Evaluating limits analytically

Evaluate the following limits.

1. $\lim\limits_{x \to \pi} \cos x$
2. $\lim\limits_{x \to 3} (\sec^2 x - \tan^2 x)$
3. $\lim\limits_{x \to \frac{\pi}{2}} \cos x \sin x$
4. $\lim\limits_{x \to 1} e^{\ln x}$
5. $\lim\limits_{x \to 0} \dfrac{\sin x}{x}$

SOLUTION

1. This is a straightforward application of Theorem 4: $\lim\limits_{x \to \pi} \cos x = \cos \pi = -1$.

2. We can approach this in at least two ways. First, by directly applying Theorems 1 and 4, we have:
$$\lim\limits_{x \to 3} (\sec^2 x - \tan^2 x) = \sec^2 3 - \tan^2 3.$$

   Using the Pythagorean Theorem, this last expression is 1; therefore
$$\lim\limits_{x \to 3} (\sec^2 x - \tan^2 x) = 1.$$

   We can also use the Pythagorean Theorem from the start:
$$\lim\limits_{x \to 3} (\sec^2 x - \tan^2 x) = \lim\limits_{x \to 3} 1 = 1,$$

   using the Constant limit rule. Either way, we find the limit is 1.

Notes:

3. Applying the Product limit rule of Theorem 1 and Theorem 4 gives

$$\lim_{x\to\pi/2} \cos x \sin x = \cos(\pi/2)\sin(\pi/2) = 0 \cdot 1 = 0.$$

4. Again, we can approach this in two ways. First, we can use the exponential/logarithmic identity that $e^{\ln x} = x$ and evaluate $\lim_{x\to 1} e^{\ln x} = \lim_{x\to 1} x = 1$.

   We can also use Theorem 2. Using Theorem 4, we have $\lim_{x\to 1} \ln x = \ln 1 = 0$. Applying the Composition rule,

$$\lim_{x\to 1} e^{\ln x} = \lim_{x\to 0} e^x = e^0 = 1.$$

   Both approaches are valid, giving the same result.

5. We encountered this limit in Section 1.1. Applying our theorems, we attempt to find the limit as

$$\lim_{x\to 0} \frac{\sin x}{x} \to \frac{\sin 0}{0} \to \text{``}\frac{0}{0}\text{''}.$$

   This, of course, violates a condition of Theorem 1, as the limit of the denominator is not allowed to be 0. Therefore, we are still unable to evaluate this limit with tools we currently have at hand.

The section could have been titled "Using Known Limits to Find Unknown Limits." By knowing certain limits of functions, we can find limits involving sums, products, powers, etc., of these functions. We further the development of such comparative tools with the Squeeze Theorem, a clever and intuitive way to find the value of some limits.

Before stating this theorem formally, suppose we have functions f, g and h where g always takes on values between f and h; that is, for all x in an interval,

$$f(x) \le g(x) \le h(x).$$

If f and h have the same limit at c, and g is always "squeezed" between them, then g must have the same limit as well. That is what the Squeeze Theorem states, as illustrated in Figure 1.16.
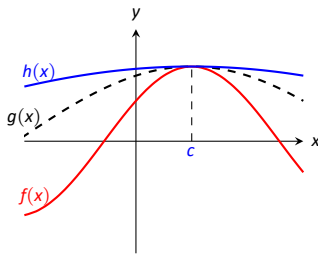
Figure 1.16: The situation of the squeeze theorem

Notes:

> **Theorem 5   Squeeze Theorem**
>
> Let f, g and h be functions on open intervals I and J on either side of c such that for all x in I and J,
>
> $$f(x) \leq g(x) \leq h(x).$$
>
> If
>
> $$\lim_{x \to c} f(x) = L = \lim_{x \to c} h(x),$$
>
> then
>
> $$\lim_{x \to c} g(x) = L.$$

It can take some work to figure out appropriate functions by which to "squeeze" the given function of which you are trying to evaluate a limit. However, that is generally the only place work is necessary; the theorem makes the "evaluating the limit part" very simple.

We use the Squeeze Theorem in the following example to finally prove that $\lim\limits_{x \to 0} \dfrac{\sin x}{x} = 1$.

Example 4   Using the Squeeze Theorem

Use the Squeeze Theorem to show that

$$\lim_{x \to 0} \frac{\sin x}{x} = 1.$$

SOLUTION   We begin by considering the unit circle. Each point on the unit circle has coordinates $(\cos \theta, \sin \theta)$ for some angle θ as shown in Figure 1.17. Using similar triangles, we can extend the line from the origin through the point to the point $(1, \tan \theta)$, as shown. (Here we are assuming that $0 \leq \theta \leq \pi/2$. Later we will show that we can also consider $\theta \leq 0$.)

Figure 1.17 shows three regions have been constructed in the first quadrant, two triangles and a sector of a circle, which are also drawn below. The area of the large triangle is $\frac{1}{2} \tan \theta$; the area of the sector is $\theta/2$; the area of the triangle contained inside the sector is $\frac{1}{2} \sin \theta$. It is then clear from the diagram that
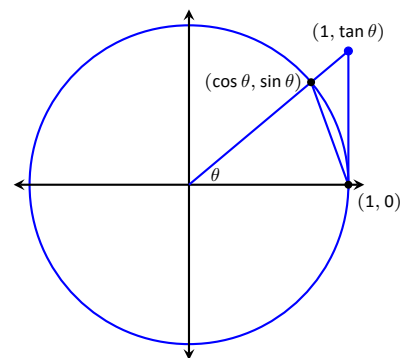
Figure 1.17: The unit circle and related triangles.

Notes:

$$\frac{\tan\theta}{2} \qquad \geq \qquad \frac{\theta}{2} \qquad \geq \qquad \frac{\sin\theta}{2}$$
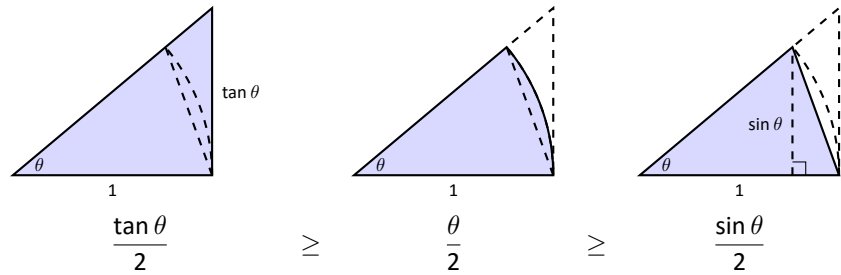
Multiply all terms by $\dfrac{2}{\sin\theta}$, giving

$$\frac{1}{\cos\theta} \geq \frac{\theta}{\sin\theta} \geq 1.$$

Taking reciprocals reverses the inequalities, giving

$$\cos\theta \leq \frac{\sin\theta}{\theta} \leq 1.$$

(These inequalities hold for all values of θ near 0, even negative values, since $\cos(-\theta) = \cos\theta$ and $\sin(-\theta) = -\sin\theta$.)

Now take limits.

$$\lim_{\theta\to0} \cos\theta \leq \lim_{\theta\to0} \frac{\sin\theta}{\theta} \leq \lim_{\theta\to0} 1$$

$$\cos 0 \leq \lim_{\theta\to0} \frac{\sin\theta}{\theta} \leq 1$$

$$1 \leq \lim_{\theta\to0} \frac{\sin\theta}{\theta} \leq 1$$

Clearly this means that $\lim\limits_{\theta\to0} \dfrac{\sin\theta}{\theta} = 1$.

Two notes about the previous example are worth mentioning. First, one might be discouraged by this application, thinking "I would *never* have come up with that on my own. This is too hard!" Don't be discouraged; within this text we will guide you in your use of the Squeeze Theorem. As one gains mathematical maturity, clever proofs like this are easier and easier to create.

Second, this limit tells us more than just that as x approaches 0, $\sin(x)/x$ approaches 1. Both x and $\sin x$ are approaching 0, but the *ratio* of x and $\sin x$ approaches 1, meaning that they are approaching 0 in essentially the same way. Another way of viewing this is: for small x, the functions $y = x$ and $y = \sin x$ are essentially indistinguishable.

We include this special limit, along with three others, in the following theorem.

Notes:

Theorem 6   Special Limits

1. $\lim\limits_{x \to 0} \dfrac{\sin x}{x} = 1$

2. $\lim\limits_{x \to 0} \dfrac{\cos x - 1}{x} = 0$

3. $\lim\limits_{x \to 0} (1 + x)^{\frac{1}{x}} = e$

4. $\lim\limits_{x \to 0} \dfrac{e^x - 1}{x} = 1$

A short word on how to interpret the latter three limits. We know that as x goes to 0, $\cos x$ goes to 1. So, in the second limit, both the numerator and denominator are approaching 0. However, since the limit is 0, we can interpret this as saying that "$\cos x$ is approaching 1 faster than x is approaching 0."

In the third limit, inside the parentheses we have an expression that is approaching 1 (though never equaling 1), and we know that 1 raised to any power is still 1. At the same time, the power is growing toward infinity. What happens to a number near 1 raised to a very large power? In this particular case, the result approaches Euler's number, e, approximately 2.718.

In the fourth limit, we see that as $x \to 0$, e^x approaches 1 "just as fast" as $x \to 0$, resulting in a limit of 1.

Our final theorem for this section will be motivated by the following example.

Example 5   Using algebra to evaluate a limit
Evaluate the following limit:
$$\lim_{x \to 1} \frac{x^2 - 1}{x - 1}.$$

SOLUTION   We would like to apply Theorems 1 and 4 and substitute 1 for x in the quotient. This gives:
$$\lim_{x \to 1} \frac{x^2 - 1}{x - 1} = \frac{1^2 - 1}{1 - 1} = \text{"} \frac{0}{0} \text{"},$$

an indeterminate form. We cannot apply the Theorem 1 because the denominator is 0.

By graphing the function, as in Figure 1.18, we see that the function seems to be linear, implying that the limit should be easy to evaluate. Recognize that the numerator of our quotient can be factored:
$$\text{Let } f(x) = \frac{x^2 - 1}{x - 1} = \frac{(x - 1)(x + 1)}{x - 1}.$$
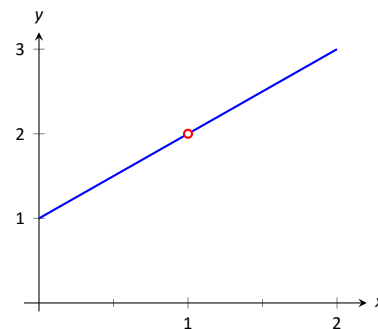
Figure 1.18: Graphing f in Example 5 to understand a limit.

Notes:

The function is not defined when $x = 1$, but for all other x,

$$\frac{x^2 - 1}{x - 1} = \frac{(x - 1)(x + 1)}{x - 1} = \frac{\cancel{(x - 1)}(x + 1)}{\cancel{x - 1}} = x + 1.$$

Clearly $\lim\limits_{x \to 1} x + 1 = 2$. Recall that when considering limits, we are not concerned with the value of the function at 1, only the value the function approaches as x approaches 1. Since $(x^2 - 1)/(x - 1)$ and $x + 1$ are the same at all points except $x = 1$, they both approach the same value as x approaches 1. Therefore we can conclude that

$$\lim_{x \to 1} \frac{x^2 - 1}{x - 1} = \lim_{x \to 1} \frac{(x - 1)(x + 1)}{x - 1} = \lim_{x \to 1} x + 1 = 2.$$

The key to the above example is that the functions $y = (x^2 - 1)/(x - 1)$ and $y = x + 1$ are identical except at $x = 1$. Since limits describe a value the function is approaching, not the value the function actually attains, the limits of the two functions are always equal.

Theorem 7   Limits of Functions Equal At All But One Point

Let $g(x) = f(x)$ for all x in an open interval, except possibly at c, and let $\lim\limits_{x \to c} g(x) = L$ for some real number L. Then

$$\lim_{x \to c} f(x) = \lim_{x \to c} g(x) = L.$$

The Fundamental Theorem of Algebra tells us that when dealing with a rational function of the form $g(x)/f(x)$ and directly evaluating the limit $\lim\limits_{x \to c} \dfrac{g(x)}{f(x)}$ returns "0/0", then $(x - c)$ is a factor of both $g(x)$ and $f(x)$. One can then use algebra to factor this term out, divide, then apply Theorem 7. Some useful algebraic techniques to rewrite functions that return an indeterminate form when evaluating a limit are:

- factoring and dividing out common factors,

- rationalizing the numerator or denominator,

- simplifying the expression, and

- finding a common denominator.

We will demonstrate some of these techniques in the following examples.

Notes:

Example 6   **Evaluating a limit using Theorem 7**

Evaluate $\lim\limits_{x \to 3} \dfrac{x^3 - 2x^2 - 5x + 6}{2x^3 + 3x^2 - 32x + 15}$.

SOLUTION   We begin by attempting to apply Theorems 1 and 4 and substituting 3 for x. This returns the familiar indeterminate form of "0/0". Since the numerator and denominator are each polynomials, we know that $(x-3)$ is factor of each. Using whatever method is most comfortable to you, factor out $(x - 3)$ from each (using polynomial division, synthetic division, a computer algebra system, etc.). We find that

$$\frac{x^3 - 2x^2 - 5x + 6}{2x^3 + 3x^2 - 32x + 15} = \frac{(x - 3)(x^2 + x - 2)}{(x - 3)(2x^2 + 9x - 5)}.$$

We can divide the $(x-3)$ terms as long as $x \ne 3$. Using Theorem 7 we conclude:

$$\begin{aligned}
\lim_{x \to 3} \frac{x^3 - 2x^2 - 5x + 6}{2x^3 + 3x^2 - 32x + 15} &= \lim_{x \to 3} \frac{(x - 3)(x^2 + x - 2)}{(x - 3)(2x^2 + 9x - 5)} \\
&= \lim_{x \to 3} \frac{(x^2 + x - 2)}{(2x^2 + 9x - 5)} \\
&= \frac{10}{40} = \frac{1}{4}.
\end{aligned}$$

Example 7   **Evaluating a limit by rationalizing**

Evaluate $\lim\limits_{x \to 0} \dfrac{\sqrt{x + 4} - 2}{x}$.

SOLUTION   We begin by applying Theorem 4 and substituting 2 for x. This returns the familiar indeterminate form of "0/0". We see the radical in the numerator so we will rationalize the numerator. Using Theorem 7 we find that

$$\begin{aligned}
\lim_{x \to 0} \frac{\sqrt{x + 4} - 2}{x} &= \lim_{x \to 0} \frac{\sqrt{x + 4} - 2}{x} \cdot \frac{\sqrt{x + 4} + 2}{\sqrt{x + 4} + 2} \\
&= \lim_{x \to 0} \frac{(x + 4) - 4}{x(\sqrt{x + 4} + 2)} \\
&= \lim_{x \to 0} \frac{x}{x(\sqrt{x + 4} + 2)} \qquad \text{Simplify the numerator.} \\
&= \lim_{x \to 0} \frac{1}{\sqrt{x + 4} + 2} \qquad \text{Divide out } x. \\
&= \frac{1}{\sqrt{4} + 2} = \frac{1}{4}.
\end{aligned}$$

Notes:

Notice that we didn't distribute the denominator in the second line. Generally speaking, when we are hoping to divide out a factor in a fraction we will need to undo any distributing that we may have prematurely done.

We end this section by revisiting a limit first seen in Section 1.1, a limit of a difference quotient. Let $f(x) = -1.5x^2 + 11.5x$; we approximated the limit $\lim\limits_{h \to 0} \dfrac{f(1+h) - f(1)}{h} \approx 8.5$. We formally evaluate this limit in the following example.

Example 8   **Evaluating the limit of a difference quotient**

Let $f(x) = -1.5x^2 + 11.5x$; find $\lim\limits_{h \to 0} \dfrac{f(1+h) - f(1)}{h}$.

SOLUTION   Since f is a polynomial, our first attempt should be to employ Theorem 4 and substitute 0 for h. However, we see that this gives us "$0/0$." Knowing that we have a rational function hints that some algebra will help. Consider the following steps:

$$\lim_{h \to 0} \frac{f(1+h) - f(1)}{h}$$

$$= \lim_{h \to 0} \frac{-1.5(1+h)^2 + 11.5(1+h) - \left(-1.5(1)^2 + 11.5(1)\right)}{h}$$

$$= \lim_{h \to 0} \frac{-1.5(1 + 2h + h^2) + 11.5 + 11.5h - 10}{h}$$

$$= \lim_{h \to 0} \frac{-1.5h^2 + 8.5h}{h}$$

$$= \lim_{h \to 0} \frac{h(-1.5h + 8.5)}{h}$$

$$= \lim_{h \to 0} (-1.5h + 8.5) \qquad \text{(using Theorem 7, as } h \neq 0)$$

$$= 8.5 \qquad \text{(using Theorem 4)}$$

This matches our previous approximation.

This section contains several valuable tools for evaluating limits. One of the main results of this section is Theorem 4; it states that many functions that we use regularly behave in a very nice, predictable way. In Section 1.6 we give a name to this nice behavior; we label such functions as *continuous.* Defining that term will require us to look again at what a limit is and what causes limits to not exist.

Notes:

Exercises 1.3

Terms and Concepts

1. Explain in your own words, without using ε-δ formality, why $\lim_{x\to c} b = b$.

2. Explain in your own words, without using ε-δ formality, why $\lim_{x\to c} x = c$.

3. What does the text mean when it says that certain functions' "behavior is 'nice' in terms of limits"? What, in particular, is "nice"?

4. Sketch a graph that visually demonstrates the Squeeze Theorem.

5. You are given the following information:

   (a) $\lim_{x\to 1} f(x) = 0$

   (b) $\lim_{x\to 1} g(x) = 0$

   (c) $\lim_{x\to 1} f(x)/g(x) = 2$

   What can be said about the relative sizes of $f(x)$ and $g(x)$ as x approaches 1?

Problems

Use the following limits to evaluate the limits given in Exercises 6–13, where possible. If it is not possible, state so.

$$\lim_{x\to 9} f(x) = f(9) = 6 \qquad \lim_{x\to 6} f(x) = f(6) = 9$$
$$\lim_{x\to 9} g(x) = g(9) = 3 \qquad \lim_{x\to 6} g(x) = g(6) = 3$$

6. $\lim_{x\to 9} \big(f(x) + g(x)\big)$

7. $\lim_{x\to 9} \big(3f(x)/g(x)\big)$

8. $\lim_{x\to 9} \left(\dfrac{f(x) - 2g(x)}{g(x)}\right)$

9. $\lim_{x\to 6} \left(\dfrac{f(x)}{3 - g(x)}\right)$

10. $\lim_{x\to 9} g\big(f(x)\big)$

11. $\lim_{x\to 6} f\big(g(x)\big)$

12. $\lim_{x\to 6} g\big(f(f(x))\big)$

13. $\lim_{x\to 6} f(x)g(x) - f^2(x) + g^2(x)$

Use the following limits to evaluate the limits given in Exercises 14–17, where possible. If it is not possible, state so.

$$\lim_{x\to 1} f(x) = f(1) = 2 \qquad \lim_{x\to 10} f(x) = f(10) = 1$$
$$\lim_{x\to 1} g(x) = g(1) = 0 \qquad \lim_{x\to 10} g(x) = g(10) = \pi$$

14. $\lim_{x\to 1} f(x)^{g(x)}$

15. $\lim_{x\to 10} \cos\big(g(x)\big)$

16. $\lim_{x\to 1} f(x)g(x)$

17. $\lim_{x\to 1} g\big(5f(x)\big)$

In Exercises 18–38, evaluate the given limit.

18. $\lim_{x\to 3} x^2 - 3x + 7$

19. $\lim_{x\to \pi} \left(\dfrac{x-3}{x-5}\right)^7$

20. $\lim_{x\to \pi/4} \cos x \sin x$

21. $\lim_{x\to 0} \ln x$

22. $\lim_{x\to 3} 4^{x^3 - 8x}$

23. $\lim_{x\to \pi/6} \csc x$

24. $\lim_{x\to 0} \ln(1+x)$

25. $\lim_{x\to \pi} \dfrac{x^2 + 3x + 5}{5x^2 - 2x - 3}$

26. $\lim_{x\to \pi} \dfrac{3x+1}{1-x}$

27. $\lim_{x\to 6} \dfrac{x^2 - 4x - 12}{x^2 - 13x + 42}$

28. $\lim_{x\to 0} \dfrac{x^2 + 2x}{x^2 - 2x}$

29. $\lim_{x\to 2} \dfrac{x^2 + 6x - 16}{x^2 - 3x + 2}$

30. $\lim_{x\to 2} \dfrac{x^2 - 10x + 16}{x^2 - x - 2}$

31. $\lim_{x\to -2} \dfrac{x^2 - 5x - 14}{x^2 + 10x + 16}$

32. $\lim_{x\to -1} \dfrac{x^2 + 9x + 8}{x^2 - 6x - 7}$

33. $\lim_{t\to 9} \dfrac{\sqrt{t} - 3}{t - 9}$

34. $\lim_{x\to 0} \dfrac{\sqrt{x^2 + 4} - 2}{x^2}$

35. $\lim_{t\to 3} \dfrac{\frac{1}{t} - \frac{1}{3}}{t - 3}$

36. $\lim_{x\to 0} \dfrac{1}{x} - \dfrac{1}{x^2 + x}$

37. $\lim_{t\to 0} \dfrac{(t-4)^2 - 16}{t}$

38. $\displaystyle\lim_{x\to 13}\frac{\sqrt{x+3}-4}{x-13}$

Use the Squeeze Theorem in Exercises 39–43, where appropriate, to evaluate the given limit.

Hint: $-1 \le \sin x \le 1$ and $-1 \le \cos x \le 1$.

39. $\displaystyle\lim_{x\to 0} x \sin\left(\frac{1}{x}\right)$

40. $\displaystyle\lim_{x\to 0} \sin x \cos\left(\frac{1}{x^2}\right)$

41. $\displaystyle\lim_{x\to 1} f(x)$, where $3x - 2 \le f(x) \le x^3$.

42. $\displaystyle\lim_{x\to 3} f(x)$, where $6x - 9 \le f(x) \le x^2$.

43. $\displaystyle\lim_{x\to 0} x^2 \cos\left(\frac{1}{x}\right)$

Exercises 44–49 challenge your understanding of limits that can be evaluated using the knowledge gained in this section.

44. $\displaystyle\lim_{x\to 0}\frac{\sin 3x}{x}$

45. $\displaystyle\lim_{x\to 0}\frac{\sin 5x}{8x}$

46. $\displaystyle\lim_{x\to 0}\frac{\ln(1+x)}{x}$

47. $\displaystyle\lim_{x\to 0}\frac{\sin x}{x}$, where x is measured in degrees, not radians.

48. $\displaystyle\lim_{x\to 0}\frac{\tan 4x}{\tan 3x}$

49. $\displaystyle\lim_{x\to 0}\frac{\tan 5x}{\sin 7x}$

50. Verify $\displaystyle\lim_{x\to 0}\frac{\cos x - 1}{x} = 0$ Hint: Multiply by $\displaystyle\frac{\cos x + 1}{\cos x + 1}$

1.4   One Sided Limits

In Section 1.1 we explored the three ways in which limits of functions failed to exist:

1. The function approached different values from the left and right,

2. The function grows without bound, and

3. The function oscillates.

In this section we explore in depth the concepts behind #1 by introducing the *one-sided limit*. We begin with formal definitions that are very similar to the definition of the limit given in Section 1.2, but the notation is slightly different and "$x \neq c$" is replaced with either "$x < c$" or "$x > c$." We will consider #2 in more detail in Section 1.5.

Definition 2   One Sided Limits

Left-Hand Limit

Let *I* be an open interval with right endpoint *c*, and let *f* be a function defined on *I*. The **limit of $f(x)$**, **as x approaches c from the left, is** *L*, or, **the left–hand limit of** *f* **at** *c* **is** *L*, denoted by

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

means that given any $\varepsilon > 0$, there exists $\delta > 0$ such that for all $x < c$, if $|x - c| < \delta$, then $|f(x) - L| < \varepsilon$.

Right-Hand Limit

Let *I* be an open interval with left endpoint *c*, and let *f* be a function defined on *I*. The **limit of $f(x)$**, **as x approaches c from the right, is** *L*, or, **the right–hand limit of** *f* **at** *c* **is** *L*, denoted by

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

means that given any $\varepsilon > 0$, there exists $\delta > 0$ such that for all $x > c$, if $|x - c| < \delta$, then $|f(x) - L| < \varepsilon$.

Practically speaking, when evaluating a left-hand limit, we consider only values of *x* "to the left of *c*," i.e., where $x < c$. The admittedly imperfect notation $x \to c^-$ is used to imply that we look at values of *x* to the left of *c*. The notation has nothing to do with positive or negative values of either *x* or *c*. A similar

---

Notes:

statement holds for evaluating right-hand limits; there we consider only values of x to the right of c, i.e., $x > c$. We can use the theorems from previous sections to help us evaluate these limits; we just restrict our view to one side of c.

Watch the video:
One-sided limits from graphs at
`https://youtu.be/nOnd3SiYZqM`

We practice evaluating left and right-hand limits through a series of examples.

Example 1   **Evaluating one sided limits**

Let $f(x) = \begin{cases} 2x & 0 \le x \le 1 \\ 6 - 2x & 1 < x < 2 \end{cases}$, as shown in Figure 1.19. Find each of the following:

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

2. $\displaystyle\lim_{x\to 1^+} f(x)$

3. $\displaystyle\lim_{x\to 1} f(x)$

4. $f(1)$

5. $\displaystyle\lim_{x\to 0^+} f(x)$

6. $f(0)$

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

8. $f(2)$

SOLUTION   For these problems, the visual aid of the graph is likely more effective in evaluating the limits than using f itself. Therefore we will refer often to the graph.

1. As x goes to 1 *from the left*, we see that $f(x)$ is approaching the value of 2. Therefore $\displaystyle\lim_{x\to 1^-} f(x) = 2$.

2. As x goes to 1 *from the right*, we see that $f(x)$ is approaching the value of 4. Recall that it does not matter that there is an "open circle" there; we are evaluating a limit, not the value of the function. Therefore $\displaystyle\lim_{x\to 1^+} f(x) = 4$.

3. *The* limit of f as x approaches 1 does not exist, as discussed in the first section. The function does not approach one particular value, but two different values from the left and the right.

Figure 1.19: A graph of f in Example 1.

Notes:

4. Using the definition and by looking at the graph we see that $f(1) = 2$.

5. As x goes to 0 from the right, we see that $f(x)$ is also approaching 0. There-fore $\lim\limits_{x \to 0^+} f(x) = 0$. Note we cannot consider a left-hand limit at 0 as f is not defined for values of $x < 0$.

6. Using the definition and the graph, $f(0) = 0$.

7. As x goes to 2 from the left, we see that $f(x)$ is approaching the value of 2. Therefore $\lim\limits_{x \to 2^-} f(x) = 2$.

8. The graph and the definition of the function show that $f(2)$ is not defined.

Note how the left and right-hand limits were different at $x = 1$. This, of course, causes *the* limit to not exist. The following theorem states what is fairly intuitive: *the* limit exists precisely when the left and right-hand limits are equal.

Theorem 8   Limits and One Sided Limits

Let f be a function defined on an open interval I containing c. Then

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

if, and only if,

$$\lim_{x \to c^-} f(x) = L \quad \text{and} \quad \lim_{x \to c^+} f(x) = L.$$

The phrase "if, and only if" means the two statements are *equivalent*: they are either both true or both false. If the limit equals L, then the left and right hand limits both equal L. If the limit is not equal to L, then at least one of the left and right-hand limits is not equal to L (it may not even exist).

One thing to consider in Examples $1 - 4$ is that the value of the function may or may not be equal to the value(s) of its left- or right-hand limits, even when these limits agree.

Example 2   Evaluating limits of a piecewise–defined function
Let $f(x) = \begin{cases} 2 - x & 0 < x < 1 \\ (x - 2)^2 & 1 < x < 2 \end{cases}$, as shown in Figure 1.20. Evaluate the following.

Notes:

1. $\lim\limits_{x \to 1^-} f(x)$

2. $\lim\limits_{x \to 1^+} f(x)$

3. $\lim\limits_{x \to 1} f(x)$

4. $f(1)$

5. $\lim\limits_{x \to 0^+} f(x)$

6. $f(0)$

7. $\lim\limits_{x \to 2^-} f(x)$

8. $f(2)$

SOLUTION   Again we will evaluate each using both the definition of f and its graph.

1. As x approaches 1 from the left, we see that $f(x)$ approaches 1. Therefore $\lim\limits_{x \to 1^-} f(x) = 1$.

2. As x approaches 1 from the right, we see that again $f(x)$ approaches 1. Therefore $\lim\limits_{x \to 1^+} f(x) = 1$.

3. *The* limit of f as x approaches 1 exists and is 1, as f approaches 1 from both the right and left. Therefore $\lim\limits_{x \to 1} f(x) = 1$.

4. $f(1)$ is not defined. Note that 1 is not in the domain of f as defined by the problem, which is indicated on the graph by an open circle when $x = 1$.

5. As x goes to 0 from the right, $f(x)$ approaches 2. So $\lim\limits_{x \to 0^+} f(x) = 2$.

6. $f(0)$ is not defined as 0 is not in the domain of f.

7. As x goes to 2 from the left, $f(x)$ approaches 0. So $\lim\limits_{x \to 2^-} f(x) = 0$.

8. $f(2)$ is not defined as 2 is not in the domain of f.

Figure 1.20: A graph of f from Example 2
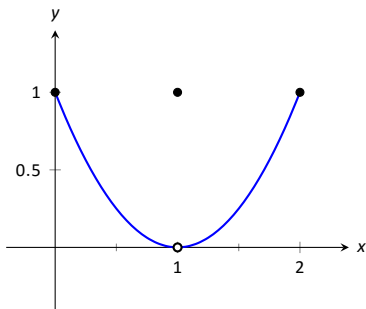
Figure 1.21: Graphing f in Example 3

Example 3   **Evaluating limits of a piecewise–defined function**

Let $f(x) = \begin{cases} (x-1)^2 & 0 \le x \le 2, x \ne 1 \\ 1 & x = 1 \end{cases}$, as shown in Figure 1.21. Evaluate the following.

1. $\lim\limits_{x \to 1^-} f(x)$

2. $\lim\limits_{x \to 1^+} f(x)$

3. $\lim\limits_{x \to 1} f(x)$

4. $f(1)$

Notes:

SOLUTION   It is clear by looking at the graph that both the left and right-hand limits of f, as x approaches 1, is 0. Thus it is also clear that *the* limit is 0; i.e., $\lim\limits_{x \to 1} f(x) = 0$. It is also clearly stated that $f(1) = 1$.

Example 4   **Evaluating limits of a piecewise–defined function**

Let $f(x) = \begin{cases} x^2 & 0 \le x \le 1 \\ 2 - x & 1 < x \le 2 \end{cases}$. Evaluate the following.

1. $\lim\limits_{x \to 1^-} f(x)$

2. $\lim\limits_{x \to 1^+} f(x)$

3. $\lim\limits_{x \to 1} f(x)$

4. $f(1)$

SOLUTION   In this example, we will evaluate the limit by only considering the definition of f.

1. As x approaches 1 from the left, $f(x)$ is defined to be x^2. Therefore

$$\lim\limits_{x \to 1^-} f(x) = \lim\limits_{x \to 1^-} x^2 = 1.$$

2. As x approaches 1 from the right, $f(x)$ is defined to be $2 - x$. Therefore

$$\lim\limits_{x \to 1+} f(x) = \lim\limits_{x \to 1+} 2 - x = 1.$$

3. Since the right and left hand limits are equal at $x = 1$, i.e., $\lim\limits_{x \to 1^-} f(x) = \lim\limits_{x \to 1+} f(x) = 1$, this tells us $\lim\limits_{x \to 1} f(x) = 1$.

4. To find $f(1)$, we use the x^2 piece of our function, so $f(1) = 1$.

Example 5   **Evaluating limits of an absolute value function**

Let $f(x) = \dfrac{|x - 1|}{x - 1}$. Evaluate the following.

1. $\lim\limits_{x \to 1^-} f(x)$

2. $\lim\limits_{x \to 1^+} f(x)$

3. $\lim\limits_{x \to 1} f(x)$

4. $f(1)$

Notes:

SOLUTION   We begin by rewriting $|x - 1|$ as a piecewise function.

$$|x - 1| = \begin{cases} x - 1 & x \geq 1 \\ -(x - 1) & x \leq 1 \end{cases}$$

1. $\displaystyle\lim_{x \to 1^-} f(x) = \lim_{x \to 1^-} \frac{-(x-1)}{x-1} = \lim_{x \to 1^-} -1 = -1$

2. $\displaystyle\lim_{x \to 1^+} f(x) = \lim_{x \to 1^+} \frac{x-1}{x-1} = \lim_{x \to 1^+} 1 = 1$

3. $\displaystyle\lim_{x \to 1} f(x)$ does not exist because the left and right hand limits are not equal.

4. $f(1)$ is undefined.

In Examples $1 - 5$ we were asked to find both $\displaystyle\lim_{x \to 1} f(x)$ and $f(1)$. Consider the following table:

	$\displaystyle\lim_{x \to 1} f(x)$	$f(1)$
Example 1	does not exist	2
Example 2	1	not defined
Example 3	0	1
Example 4	1	1
Example 5	does not exist	not defined

Only in Example 4 do both the function and the limit exist and agree. This seems "nice;" in fact, it seems "normal." This is in fact an important situation which we explore in Section 1.6, entitled "Continuity." In short, a *continuous function* is one in which when a function approaches a value as $x \to c$ (i.e., when $\displaystyle\lim_{x \to c} f(x) = L$), it actually *attains* that value at c. Such functions behave nicely as they are very predictable.

In the next section we examine one more aspect of limits: limits that involve infinity.

Notes:

Exercises 1.4

Terms and Concepts

1. What are the three ways in which a limit may fail to exist?

2. T/F: If $\lim\limits_{x\to 1^-} f(x) = 5$, then $\lim\limits_{x\to 1} f(x) = 5$

3. T/F: If $\lim\limits_{x\to 1^-} f(x) = 5$, then $\lim\limits_{x\to 1^+} f(x) = 5$

4. T/F: If $\lim\limits_{x\to 1} f(x) = 5$, then $\lim\limits_{x\to 1^-} f(x) = 5$

Problems

In Exercises 5–11, evaluate each expression using the given graph of $f(x)$.

5.
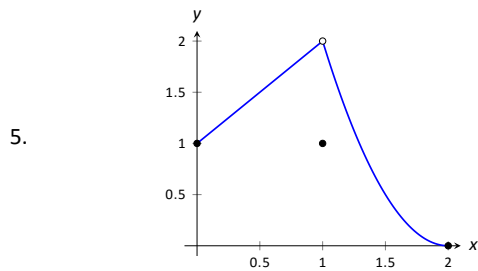
(a) $\lim\limits_{x\to 1^-} f(x)$ (d) $f(1)$

(b) $\lim\limits_{x\to 1^+} f(x)$ (e) $\lim\limits_{x\to 0^-} f(x)$

(c) $\lim\limits_{x\to 1} f(x)$ (f) $\lim\limits_{x\to 0^+} f(x)$

6.
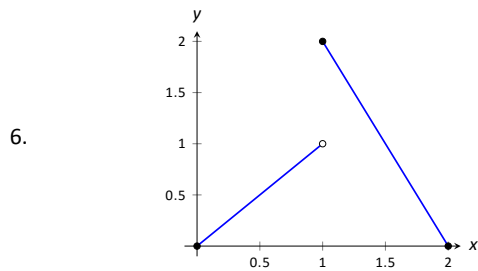
(a) $\lim\limits_{x\to 1^-} f(x)$ (d) $f(1)$

(b) $\lim\limits_{x\to 1^+} f(x)$ (e) $\lim\limits_{x\to 2^-} f(x)$

(c) $\lim\limits_{x\to 1} f(x)$ (f) $\lim\limits_{x\to 2^+} f(x)$

7.

(a) $\lim\limits_{x\to 1^-} f(x)$ (c) $\lim\limits_{x\to 1} f(x)$

(b) $\lim\limits_{x\to 1^+} f(x)$ (d) $f(1)$

8.

(a) $\lim\limits_{x\to 1^-} f(x)$ (c) $\lim\limits_{x\to 1} f(x)$

(b) $\lim\limits_{x\to 1^+} f(x)$ (d) $f(1)$

9.

(a) $\lim\limits_{x\to 0^-} f(x)$ (c) $\lim\limits_{x\to 0} f(x)$

(b) $\lim\limits_{x\to 0^+} f(x)$ (d) $f(0)$

10.

(a) $\lim\limits_{x\to -2^-} f(x)$ (e) $\lim\limits_{x\to 2^-} f(x)$

(b) $\lim\limits_{x\to -2^+} f(x)$ (f) $\lim\limits_{x\to 2^+} f(x)$

(c) $\lim\limits_{x\to -2} f(x)$ (g) $\lim\limits_{x\to 2} f(x)$

(d) $f(-2)$ (h) $f(2)$

11.
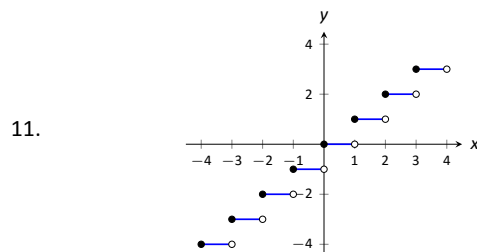
Let $-3 \le a \le 3$ be an integer.

(a) $\lim\limits_{x\to a^-} f(x)$ (c) $\lim\limits_{x\to a} f(x)$

(b) $\lim\limits_{x\to a^+} f(x)$ (d) $f(a)$

In Exercises 12–15, evaluate the given limit.

12. $\lim\limits_{x\to 7^-} \sqrt{x^2 - 49}$.

45

13. $\lim\limits_{x\to 5^+} \sqrt{x^2 - 25}$.

14. $\lim\limits_{x\to -10^-} \sqrt{100 - x^2}$.

15. $\lim\limits_{x\to -8^+} \sqrt{x^2 - 64}$.

In Exercises 16–26, evaluate the given limits of the piecewise defined functions f.

16. $f(x) = \begin{cases} x + 1 & x \le 1 \\ x^2 - 5 & x > 1 \end{cases}$

    (a) $\lim\limits_{x\to 1^-} f(x)$ (c) $\lim\limits_{x\to 1} f(x)$

    (b) $\lim\limits_{x\to 1^+} f(x)$ (d) $f(1)$

17. $f(x) = \begin{cases} 2x^2 + 5x - 1 & x < 0 \\ \sin x & x \ge 0 \end{cases}$

    (a) $\lim\limits_{x\to 0^-} f(x)$ (c) $\lim\limits_{x\to 0} f(x)$

    (b) $\lim\limits_{x\to 0^+} f(x)$ (d) $f(0)$

18. $f(x) = \begin{cases} x^2 - 1 & x < -1 \\ x^3 + 1 & -1 \le x \le 1 \\ x^2 + 1 & x > 1 \end{cases}$

    (a) $\lim\limits_{x\to -1^-} f(x)$ (e) $\lim\limits_{x\to 1^-} f(x)$

    (b) $\lim\limits_{x\to -1^+} f(x)$ (f) $\lim\limits_{x\to 1^+} f(x)$

    (c) $\lim\limits_{x\to -1} f(x)$ (g) $\lim\limits_{x\to 1} f(x)$

    (d) $f(-1)$ (h) $f(1)$

19. $f(x) = \begin{cases} \cos x & x < \pi \\ \sin x & x \ge \pi \end{cases}$

    (a) $\lim\limits_{x\to \pi^-} f(x)$ (c) $\lim\limits_{x\to \pi} f(x)$

    (b) $\lim\limits_{x\to \pi^+} f(x)$ (d) $f(\pi)$

20. $f(x) = \begin{cases} 1 - \cos^2 x & x < a \\ \sin^2 x & x \ge a \end{cases}$,

   where a is a real number.

    (a) $\lim\limits_{x\to a^-} f(x)$ (c) $\lim\limits_{x\to a} f(x)$

    (b) $\lim\limits_{x\to a^+} f(x)$ (d) $f(a)$

21. $f(x) = \begin{cases} x + 1 & x < 1 \\ 1 & x = 1 \\ x - 1 & x > 1 \end{cases}$

    (a) $\lim\limits_{x\to 1^-} f(x)$ (c) $\lim\limits_{x\to 1} f(x)$

    (b) $\lim\limits_{x\to 1^+} f(x)$ (d) $f(1)$

22. $f(x) = \begin{cases} x^2 & x < 2 \\ x + 1 & x = 2 \\ -x^2 + 2x + 4 & x > 2 \end{cases}$

    (a) $\lim\limits_{x\to 2^-} f(x)$ (c) $\lim\limits_{x\to 2} f(x)$

    (b) $\lim\limits_{x\to 2^+} f(x)$ (d) $f(2)$

23. $f(x) = \begin{cases} a(x - b)^2 + c & x < b \\ a(x - b) + c & x \ge b \end{cases}$,

   where a, b and c are real numbers.

    (a) $\lim\limits_{x\to b^-} f(x)$ (c) $\lim\limits_{x\to b} f(x)$

    (b) $\lim\limits_{x\to b^+} f(x)$ (d) $f(b)$

24. $f(x) = \begin{cases} \frac{|x|}{x} & x \ne 0 \\ 0 & x = 0 \end{cases}$

    (a) $\lim\limits_{x\to 0^-} f(x)$ (c) $\lim\limits_{x\to 0} f(x)$

    (b) $\lim\limits_{x\to 0^+} f(x)$ (d) $f(0)$

25. $f(x) = \lim\limits_{x\to 4} \dfrac{|4 - x|}{x - 4}$

    (a) $\lim\limits_{x\to 4^-} f(x)$ (c) $\lim\limits_{x\to 4} f(x)$

    (b) $\lim\limits_{x\to 4^+} f(x)$ (d) $f(4)$

26. $f(x) = \lim\limits_{x\to -2} \dfrac{x + 2}{|x + 2|}$

    (a) $\lim\limits_{x\to -2^-} f(x)$ (c) $\lim\limits_{x\to -2} f(x)$

    (b) $\lim\limits_{x\to -2^+} f(x)$ (d) $f(-2)$

In Exercises 27–30, sketch the graph of a function f that satisfies all of the given conditions.

27. $\lim\limits_{x\to 1^-} f(x) = 2$, $\lim\limits_{x\to 1^+} f(x) = -3$, $f(1) = 0$.

28. $\lim\limits_{x\to -1} f(x) = 3$, $\lim\limits_{x\to 3^-} f(x) = 1$, $\lim\limits_{x\to 3^+} f(x) = -2$, $f(-1) = 1$, $f(3) = -2$.

29. $\lim\limits_{x\to -2^-} f(x) = 1$, $\lim\limits_{x\to -2^+} f(x) = 0$, $\lim\limits_{x\to 0^-} f(x) = 3$, $\lim\limits_{x\to 0^+} f(x) = -1$, $f(-2) = 4$, $f(0) = -3$.

30. $\lim\limits_{x\to 0^-} f(x) = 0$, $\lim\limits_{x\to 0^+} f(x) = 2$, $\lim\limits_{x\to 4^-} f(x) = -2$, $\lim\limits_{x\to 4^+} f(x) = 1$, $f(0) = 2$, $f(4) = -2$.

Review

31. Evaluate the limit: $\lim\limits_{x\to -1} \dfrac{x^2 + 5x + 4}{x^2 - 3x - 4}$.

32. Evaluate the limit: $\lim\limits_{x\to -4} \dfrac{x^2 - 16}{x^2 - 4x - 32}$.

33. Evaluate the limit: $\lim\limits_{h\to 0} \dfrac{\sqrt{3 + h} - \sqrt{3}}{h}$.

34. Approximate the limit numerically: $\lim\limits_{h\to 0} \dfrac{(2 + h)^2 - 4}{h}$.

35. Approximate the limit numerically: $\lim\limits_{x\to 0.2} \dfrac{x^2 + 5.8x - 1.2}{x^2 - 4.2x + 0.8}$.

1.5   Limits Involving Infinity

In Definition 1 we stated that in the equation $\lim\limits_{x\to c} f(x) = L$, both c and L were numbers. In this section we relax that definition a bit by considering situations when it makes sense to let c and/or L be "infinity."

As a motivating example, consider $f(x) = 1/x^2$, as shown in Figure 1.22. Note how, as x approaches 0, $f(x)$ grows very, very large. It seems appropriate, and descriptive, to state that

$$\lim_{x\to 0} \frac{1}{x^2} = \infty.$$

Also note that as x gets very large, $f(x)$ gets very, very small. We could represent this concept with notation such as

$$\lim_{x\to \infty} \frac{1}{x^2} = 0.$$

We explore both types of use of ∞ in turn.

> **Definition 3   Limit of Infinity, ∞**
> We say $\lim\limits_{x\to c} f(x) = \infty$ if for every $M > 0$ there exists $\delta > 0$ such that for all $x \neq c$, if $|x - c| < \delta$, then $f(x) \geq M$.

This is just like the $\varepsilon-\delta$ definition from Section 1.2. In that definition, given any (small) value ε, if we let x get close enough to c (within δ units of c) then $f(x)$ is guaranteed to be within ε of $f(c)$. Here, given any (large) value M, if we let x get close enough to c (within δ units of c), then $f(x)$ will be at least as large as M. In other words, if we get close enough to c, then we can make $f(x)$ as large as we want. We can define limits equal to $-\infty$ in a similar way.

It is important to note that by saying $\lim\limits_{x\to c} f(x) = \infty$ we are implicitly stating that *the* limit of $f(x)$, as x approaches c, *does not exist.* A limit only exists when $f(x)$ approaches an actual numeric value. We use the concept of limits that approach infinity because it is helpful and descriptive.

Watch the video:
Calculus — Infinite Limits at
`https://youtu.be/-vwcLvb9AOs`
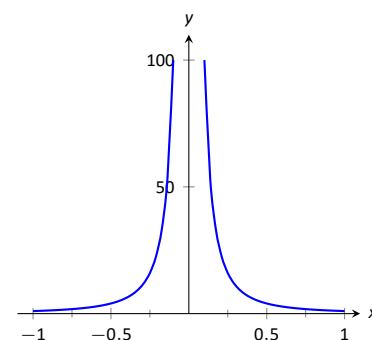
Figure 1.22: Graphing $f(x) = 1/x^2$ for values of x near 0.

Notes:
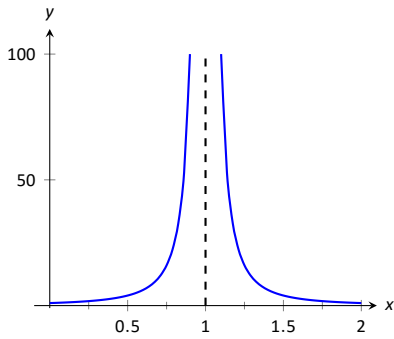
Figure 1.23: Observing infinite limit as $x \to 1$ in Example 1.

Example 1   **Evaluating limits involving infinity**

Find $\lim\limits_{x\to 1} \dfrac{1}{(x-1)^2}$ as shown in Figure 1.23.

SOLUTION   In Example 1.1.4 of Section 1.1, by inspecting values of x close to 1 we concluded that this limit does not exist. That is, it cannot equal any real number. But the limit could be infinite. And in fact, we see that the function does appear to be growing larger and larger, as $f(.99) = 10^4$, $f(.999) = 10^6$, $f(.9999) = 10^8$. A similar thing happens on the other side of 1. In general, let a "large" value M be given. Let $\delta = 1/\sqrt{M}$. If x is within δ of 1, i.e., if $|x-1| < 1/\sqrt{M}$, then:

$$|x-1| < \frac{1}{\sqrt{M}}$$
$$(x-1)^2 < \frac{1}{M}$$
$$\frac{1}{(x-1)^2} > M,$$

which is what we wanted to show. So we may say $\lim\limits_{x\to 1} 1/(x-1)^2 = \infty$.

Example 2   **Evaluating limits involving infinity**

Find $\lim\limits_{x\to 0} \dfrac{1}{x}$, as shown in Figure 1.24.

SOLUTION   It is easy to see that the function grows without bound near 0, but it does so in different ways on different sides of 0. Since its behavior is not consistent, we cannot say that $\lim\limits_{x\to 0} \dfrac{1}{x} = \infty$. However, we can make a statement about one–sided limits. We can state that $\lim\limits_{x\to 0^+} \dfrac{1}{x} = \infty$ and $\lim\limits_{x\to 0^-} \dfrac{1}{x} = -\infty$.

Vertical asymptotes
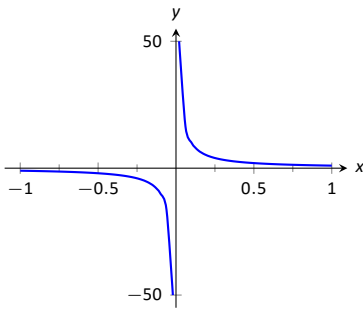
Figure 1.24: Evaluating $\lim\limits_{x\to 0} \dfrac{1}{x}$.

Definition 4   **Vertical Asymptote**

The function $f(x)$ has a **vertical asymptote at $x = c$** if any one of the following is true:

$$\lim_{x\to c^-} f(x) = \pm\infty, \quad \lim_{x\to c^+} f(x) = \pm\infty, \quad \text{or} \quad \lim_{x\to c} f(x) = \pm\infty$$

Notes:

Example 3 **Finding vertical asymptotes**

Find the vertical asymptotes of $f(x) = \dfrac{3x}{x^2 - 4}$.

SOLUTION Vertical asymptotes occur where the function grows without bound; this can occur at values of c where the denominator is 0. When x is near c, the denominator is small, which in turn can make the function take on large values. In the case of the given function, the denominator is 0 at $x = \pm 2$. We will consider the limits as x approaches ± 2 from the left and right to determine the vertical asymptotes.

$$\lim_{x \to 2^+} \frac{3x}{(x-2)(x+2)} = \infty$$

$$\lim_{x \to 2^-} \frac{3x}{(x-2)(x+2)} = -\infty$$

$$\lim_{x \to -2^+} \frac{3x}{(x-2)(x+2)} = \infty$$

$$\lim_{x \to -2^-} \frac{3x}{(x-2)(x+2)} = -\infty$$

We can graphically confirm the limits above by looking at Figure 1.25. Thus the vertical asymptotes are at $x = \pm 2$.

When a rational function has a vertical asymptote at $x = c$, we can conclude that the denominator is 0 at $x = c$. However, just because the denominator is 0 at a certain point does not mean there is a vertical asymptote there. For instance, $f(x) = (x^2 - 1)/(x - 1)$ does not have a vertical asymptote at $x = 1$, as shown in Figure 1.26.

While the denominator does get small near $x = 1$, the numerator gets small too, matching the denominator step for step. In fact, factoring the numerator, we get

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

Dividing out common term, we get that $f(x) = x+1$ for $x \neq 1$. So there is clearly no asymptote, rather a hole exists in the graph at $x = 1$.

The above example may seem a little contrived. Another example demonstrating this important concept is $f(x) = (\sin x)/x$. We have considered this function several times in the previous sections. We found that $\lim_{x \to 0} \dfrac{\sin x}{x} = 1$; i.e., there is no vertical asymptote. No simple algebraic manipulation makes this fact obvious; we used the Squeeze Theorem in Section 1.3 to prove this.

Figure 1.25: Graphing $f(x) = \dfrac{3x}{x^2 - 4}$.
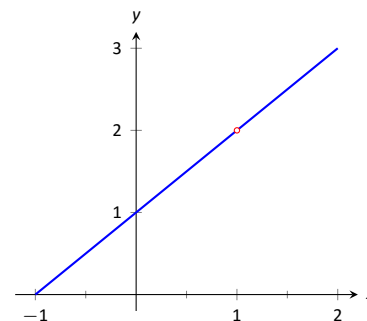
Figure 1.26: Graphically showing that $f(x) = \dfrac{x^2 - 1}{x - 1}$ does not have an asymptote at $x = 1$.

Notes:

If the denominator is 0 at a certain point but the numerator is not, then there will usually be a vertical asymptote at that point. On the other hand, if the numerator and denominator are both zero at that point, then there may or may not be a vertical asymptote at that point. This case where the numerator and denominator are both zero returns us to an important topic.

Indeterminate Forms

We have seen how the limits

$$\lim_{x \to 0} \frac{\sin x}{x} \quad \text{and} \quad \lim_{x \to 1} \frac{x^2 - 1}{x - 1}$$

each return the indeterminate form "0/0" when we blindly plug in $x = 0$ and $x = 1$, respectively. However, 0/0 is not a valid arithmetical expression. It gives no indication that the respective limits are 1 and 2.

With a little cleverness, one can come up with 0/0 expressions which have a limit of ∞, 0, or any other real number. That is why this expression is called *indeterminate*.

A key concept to understand is that such limits do not really return 0/0. Rather, keep in mind that we are taking *limits*. What is really happening is that the numerator is shrinking to 0 while the denominator is also shrinking to 0. The respective rates at which they do this are very important and determine the actual value of the limit.

An indeterminate form indicates that one needs to do more work in order to compute the limit. That work may be algebraic (such as factoring and dividing) or it may require a tool such as the Squeeze Theorem. In a later section we will learn a technique called L'Hôpital's Rule that provides another way to handle indeterminate forms.

Some other common indeterminate forms are $\infty - \infty$, $\infty \cdot 0$, ∞/∞, 0^0, ∞^0 and 1^∞. Again, keep in mind that these are the "blind" results of evaluating a limit, and each, in and of itself, has no meaning. The expression $\infty - \infty$ does not really mean "subtract infinity from infinity." Rather, it means "One quantity is subtracted from the other, but both are growing without bound." What is the result? It is possible to get every value between $-\infty$ and ∞

Note that 1/0 and $\infty/0$ are not indeterminate forms, though they are not exactly valid mathematical expressions, either. In each, the function is growing without bound, indicating that the limit will be ∞, $-\infty$, or simply not exist if the left- and right-hand limits do not match.

Notes:

Limits at Infinity and Horizontal Asymptotes

At the beginning of this section we briefly considered what happens to $f(x) = 1/x^2$ as x grew very large. Graphically, it concerns the behavior of the function to the "far right" of the graph. We make this notion more explicit in the following definition.

Definition 5   Limits at Infinity

1. We say $\lim\limits_{x \to \infty} f(x) = L$ if for every $\varepsilon > 0$ there exists $M > 0$ such that if $x \geq M$, then $|f(x) - L| < \varepsilon$.

2. We say $\lim\limits_{x \to -\infty} f(x) = L$ if for every $\varepsilon > 0$ there exists $M < 0$ such that if $x \leq M$, then $|f(x) - L| < \varepsilon$.

Definition 6   Horizontal Asymptote

The function $f(x)$ has a **horizontal asymptote at y = L** if either

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

We can also define limits such as $\lim\limits_{x \to \infty} f(x) = \infty$ by combining this definition with Definition 3.

Example 4   Approximating horizontal asymptotes

Approximate the horizontal asymptote(s) of $f(x) = \dfrac{x^2}{x^2 + 4}$.

SOLUTION   We will approximate the horizontal asymptotes by approximating the limits

$$\lim_{x \to -\infty} \frac{x^2}{x^2 + 4} \quad \text{and} \quad \lim_{x \to \infty} \frac{x^2}{x^2 + 4}.$$

Figure 1.27(a) shows a sketch of f, and part (b) gives values of $f(x)$ for large magnitude values of x. It seems reasonable to conclude from both of these sources that f has a horizontal asymptote at $y = 1$.

Later, we will show how to determine this analytically.

(a)
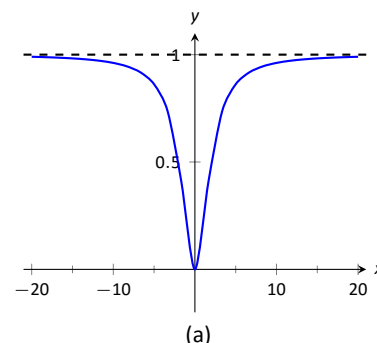
x	$f(x)$
10	0.9615
100	0.9996
10000	0.999996
−10	0.9615
−100	0.9996
−10000	0.999996

(b)

Figure 1.27: Using a graph and a table to approximate a horizontal asymptote in Example 4.

Notes:

Horizontal asymptotes can take on a variety of forms. Figure 1.28(a) shows that $f(x) = x/(x^2 + 1)$ has a horizontal asymptote of $y = 0$, where 0 is approached from both above and below.

Figure 1.28(b) shows that $f(x) = x/\sqrt{x^2 + 1}$ has two horizontal asymptotes; one at $y = 1$ and the other at $y = -1$.

Figure 1.28(c) shows that $f(x) = (\sin x)/x$ has even more interesting behavior than at just $x = 0$; as x approaches $\pm\infty$, $f(x)$ approaches 0, but oscillates as it does this.

(a)   (b)   (c)
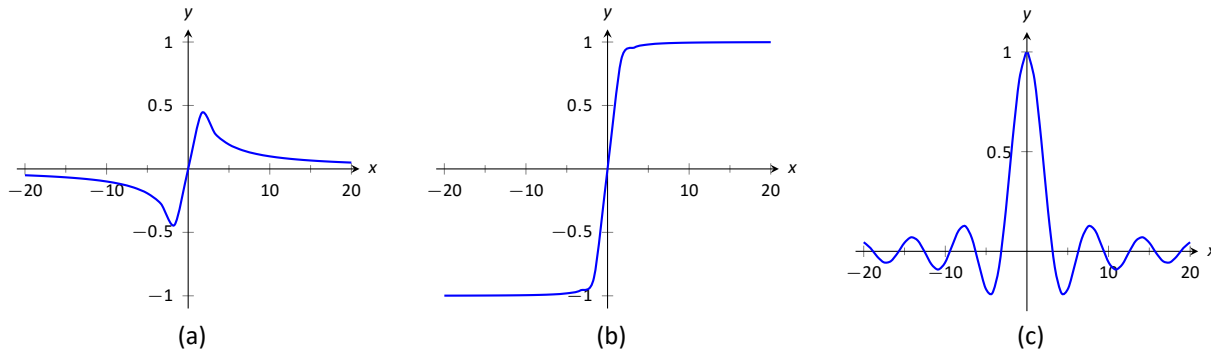
Figure 1.28: Considering different types of horizontal asymptotes.

We can analytically evaluate limits at infinity for rational functions once we understand $\lim\limits_{x\to\infty} 1/x$. As x gets larger and larger, the $1/x$ gets smaller and smaller, approaching 0. We can, in fact, make $1/x$ as small as we want by choosing a large enough value of x. Given ε, we can make $1/x < \varepsilon$ by choosing $x > 1/\varepsilon$. Thus we have $\lim_{x\to\infty} 1/x = 0$. It is now not much of a jump to conclude the following:

Theorem 9   Limits of $\dfrac{1}{x^n}$

For any $n > 0$,

$$\lim_{x\to\infty} \frac{1}{x^n} = 0 \quad \text{and} \quad \lim_{x\to-\infty} \frac{1}{x^n} = 0$$

Now suppose we need to compute the following limit:

$$\lim_{x\to\infty} \frac{x^3 + 2x + 1}{4x^3 - 2x^2 + 9}.$$

A good way of approaching this is to divide through the numerator and denominator by x^3 (hence dividing by 1), which is the largest power of x to appear in

---

Notes:

the function. Doing this, we get

$$\lim_{x\to\infty} \frac{x^3 + 2x + 1}{4x^3 - 2x^2 + 9} = \lim_{x\to\infty} \frac{1/x^3}{1/x^3} \cdot \frac{x^3 + 2x + 1}{4x^3 - 2x^2 + 9}$$

$$= \lim_{x\to\infty} \frac{x^3/x^3 + 2x/x^3 + 1/x^3}{4x^3/x^3 - 2x^2/x^3 + 9/x^3}$$

$$= \lim_{x\to\infty} \frac{1 + 2/x^2 + 1/x^3}{4 - 2/x + 9/x^3}$$

$$= \frac{1 + 0 + 0}{4 - 0 + 0} = \frac{1}{4}.$$

We used the rules for limits (which also hold for limits at infinity), as well as the fact about limits of $1/x^n$. This procedure works for any rational function and is highlighted in the following Key Idea.

Key Idea 1   **Finding Limits of Rational Functions at Infinity**

Let $f(x)$ be a rational function of the following form:

$$f(x) = \frac{a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0}{b_m x^m + b_{m-1} x^{m-1} + \cdots + b_1 x + b_0},$$

where any of the coefficients may be 0 except for a_n and b_m.
To determine $\lim\limits_{x\to\infty} f(x)$ or $\lim\limits_{x\to-\infty} f(x)$:

1. Divide the numerator and denominator by x^m.

2. Simplify as much as possible.

3. Use Theorem 9 to find the limit.

If the highest power of x is the same in both the numerator and denominator (i.e. $n = m$), we will be in a situation like the example above, where we will divide by x^n and in the limit all the terms will approach 0 except for $a_n x^n / x^n$ and $b_m x^m / x^n$. Since $n = m$, this will leave us with the limit a_n/b_m. If $n < m$, then after dividing through by x^m, all the terms in the numerator will approach 0 in the limit, leaving us with $0/b_m$ or 0. If $n > m$, and we try dividing through by x^n, we end up with all the terms in the denominator tending toward 0, while the x^n term in the numerator does not approach 0. This is indicative of some sort of infinite limit.

Intuitively, as x gets very large, all the terms in the numerator are small in comparison to $a_n x^n$, and likewise all the terms in the denominator are small

---

Notes:

compared to $b_n x^m$. If $n = m$, looking only at these two important terms, we have $(a_n x^n)/(b_n x^m)$. This reduces to a_n/b_m. If $n < m$, the function behaves like $a_n/(b_m x^{m-n})$, which tends toward 0. If $n > m$, the function behaves like $a_n x^{n-m}/b_m$, which will tend to either ∞ or $-\infty$ depending on the values of n, m, a_n, b_m and whether you are looking for $\lim_{x \to \infty} f(x)$ or $\lim_{x \to -\infty} f(x)$.

This procedure works for any rational function. In fact, it gives us the following key idea.

Key Idea 2   Limits of Rational Functions at Infinity

Let $f(x)$ be a rational function of the following form:

$$f(x) = \frac{a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0}{b_m x^m + b_{m-1} x^{m-1} + \cdots + b_1 x + b_0},$$

where any of the coefficients may be 0 except for a_n and b_m.

1. If $n = m$, then $\lim\limits_{x \to \infty} f(x) = \lim\limits_{x \to -\infty} f(x) = \dfrac{a_n}{b_m}$.

2. If $n < m$, then $\lim\limits_{x \to \infty} f(x) = \lim\limits_{x \to -\infty} f(x) = 0$.

3. If $n > m$, then $\lim\limits_{x \to \infty} f(x)$ and $\lim\limits_{x \to -\infty} f(x)$ are both infinite.

Example 5   Horizontal Asymptotes Involving Square Roots

Find the horizontal asymptotes of $\dfrac{x}{\sqrt{x^2 + 1}}$.

SOLUTION   We must consider the limits as $x \to \pm\infty$. When x is very large, $x^2 + 1 \approx x^2$ and thus $\sqrt{x^2 + 1} \approx \sqrt{x^2} = |x|$.

$$\lim_{x \to \infty} \frac{x}{\sqrt{x^2 + 1}} = \lim_{x \to \infty} \frac{x/x}{\sqrt{x^2/x^2 + 1/x^2}}$$

$$= \lim_{x \to \infty} \frac{1}{\sqrt{1 + 1/x^2}}$$

$$= 1$$

Notes:

Therefore, $y = 1$ is a horizontal asymptote. Similarly,

$$\lim_{x \to -\infty} \frac{x}{\sqrt{x^2 + 1}} = \lim_{x \to -\infty} \frac{x/(-x)}{\sqrt{x^2/x^2 + 1/x^2}}$$

$$= \lim_{x \to -\infty} \frac{-1}{\sqrt{1 + 1/x^2}}$$

$$= -1$$

Therefore, $y = -1$ is also a horizontal asymptote.

Example 6   **Finding a limit of a rational function**

Confirm analytically that $y = 1$ is the horizontal asymptote of $f(x) = \dfrac{x^2}{x^2 + 4}$, as approximated in Example 4.

SOLUTION   Before using Key Idea 2, let's use the technique of evaluating limits at infinity of rational functions that led to that theorem. The largest power of x in f is 2, so divide the numerator and denominator of f by x^2, then take limits.

$$\lim_{x \to \infty} \frac{x^2}{x^2 + 4} = \lim_{x \to \infty} \frac{x^2/x^2}{x^2/x^2 + 4/x^2}$$

$$= \lim_{x \to \infty} \frac{1}{1 + 4/x^2}$$

$$= \frac{1}{1 + 0}$$

$$= 1.$$

We can also use Key Idea 2 directly; in this case $n = m$ so the limit is the ratio of the leading coefficients of the numerator and denominator, i.e., $1/1 = 1$.

Example 7   **Finding limits of rational functions**

(a) Analytically evaluate the following limits, and (b) Use Key Idea 2 to evaluate each limit.

1. $\displaystyle \lim_{x \to -\infty} \frac{x^2 + 2x - 1}{x^3 + 1}$ 3. $\displaystyle \lim_{x \to \infty} \frac{x^2 - 1}{3 - x}$

2. $\displaystyle \lim_{x \to \infty} \frac{x^2 + 2x - 1}{1 - x - 3x^2}$

Notes:
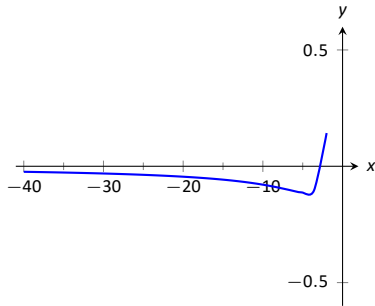
(a)
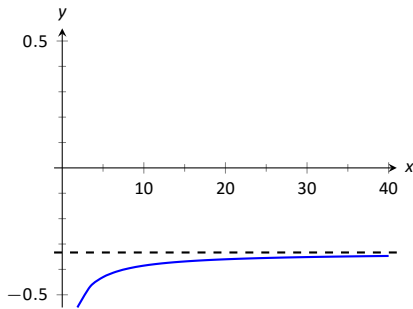
(b)
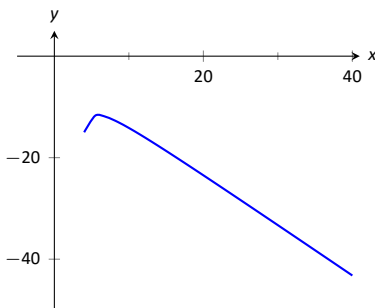
(c)

Figure 1.29: Visualizing the functions in Example 7.

SOLUTION

1. (a) Divide numerator and denominator by x^3.

    $$\lim_{x\to-\infty} \frac{x^2 + 2x - 1}{x^3 + 1} = \lim_{x\to-\infty} \frac{x^2/x^3 + 2x/x^3 - 1/x^3}{x^3/x^3 + 1/x^3}$$

    $$= \lim_{x\to-\infty} \frac{1/x + 2/x^2 - 1/x^3}{1 + 1/x^3}$$

    $$= \frac{0 + 0 + 0}{1 + 0} = 0$$

    (b) The highest power of x is in the denominator. Therefore, the limit is 0; see Figure 1.29(a).

2. (a) Divide numerator and denominator by x^2.

    $$\lim_{x\to\infty} \frac{x^2 + 2x - 1}{1 - x - 3x^2} = \lim_{x\to\infty} \frac{x^2/x^2 + 2x/x^2 - 1/x^2}{1/x^2 - x/x^2 - 3x^2/x^2}$$

    $$= \lim_{x\to\infty} \frac{1 + 2/x - 1/x^2}{1/x^2 - 1/x - 3}$$

    $$= \frac{1 + 0 - 0}{0 - 0 - 3} = -\frac{1}{3}$$

    (b) The highest power of x is x^2, which occurs in both the numerator and denominator. The limit is therefore the ratio of the coefficients of x^2, which is $-1/3$. See Figure 1.29(b).

3. (a) Divide numerator and denominator by x.

    $$\lim_{x\to\infty} \frac{x^2 - 1}{3 - x} = \lim_{x\to\infty} \frac{x^2/x - 1/x}{3/x - x/x}$$

    $$= \lim_{x\to\infty} \frac{x - 1/x}{3/x - 1}$$

    $$= \infty$$

    (b) The highest power of x is in the numerator so the limit will be ∞ or $-\infty$. To see which, consider only the dominant terms from the numerator and denominator, which are x^2 and $-x$. The expression in the limit will behave like $x^2/(-x) = -x$ for large values of x. Therefore, the limit is $-\infty$. See Figure 1.29(c).

Notes:

Exercises 1.5

Terms and Concepts

1. T/F: If $\lim\limits_{x\to 5} f(x) = \infty$, then we are implicitly stating that the limit exists.

2. T/F: If $\lim\limits_{x\to\infty} f(x) = 5$, then we are implicitly stating that the limit exists.

3. T/F: If $\lim\limits_{x\to 1^-} f(x) = -\infty$, then $\lim\limits_{x\to 1^+} f(x) = \infty$

4. T/F: If $\lim\limits_{x\to 5} f(x) = \infty$, then f has a vertical asymptote at $x = 5$.

5. T/F: $\infty/0$ is not an indeterminate form.

6. List 5 indeterminate forms.

7. Construct a function with a vertical asymptote at $x = 5$ and a horizontal asymptote at $y = 5$.

8. Let $\lim\limits_{x\to 7} f(x) = \infty$. Explain how we know that f is/is not continuous at $x = 7$.

Problems

In Exercises 9–14, evaluate the given limits using the graph of the function.

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

   (a) $\lim\limits_{x\to -1^-} f(x)$

   (b) $\lim\limits_{x\to -1^+} f(x)$
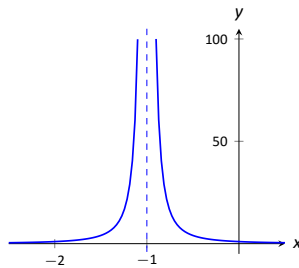
10. $f(x) = \dfrac{1}{(x-3)(x-5)^2}$.

   (a) $\lim\limits_{x\to 3^-} f(x)$ (d) $\lim\limits_{x\to 5^-} f(x)$

   (b) $\lim\limits_{x\to 3^+} f(x)$ (e) $\lim\limits_{x\to 5^+} f(x)$

   (c) $\lim\limits_{x\to 3} f(x)$ (f) $\lim\limits_{x\to 5} f(x)$
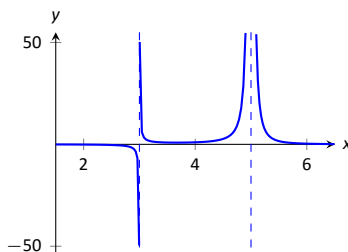
11. $f(x) = \dfrac{1}{e^x + 1}$

   (a) $\lim\limits_{x\to -\infty} f(x)$ (c) $\lim\limits_{x\to 0^-} f(x)$

   (b) $\lim\limits_{x\to\infty} f(x)$ (d) $\lim\limits_{x\to 0^+} f(x)$
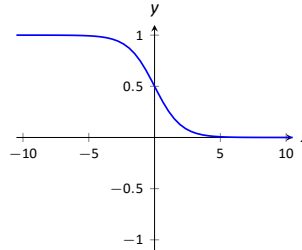
12. $f(x) = x^2 \sin(\pi x)$

   (a) $\lim\limits_{x\to -\infty} f(x)$

   (b) $\lim\limits_{x\to\infty} f(x)$
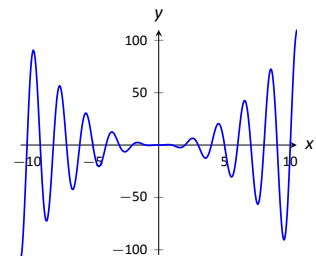
13. $f(x) = \cos(x)$

   (a) $\lim\limits_{x\to -\infty} f(x)$

   (b) $\lim\limits_{x\to\infty} f(x)$
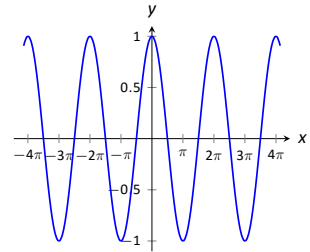
14. $f(x) = 2^x + 10$

   (a) $\lim\limits_{x\to -\infty} f(x)$

   (b) $\lim\limits_{x\to\infty} f(x)$
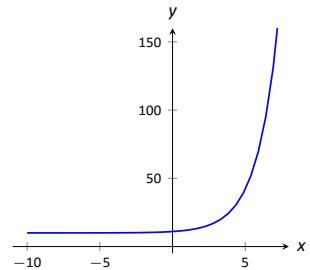
In Exercises 15–18, numerically approximate the following limits:

   (a) $\lim\limits_{x\to 3^-} f(x)$

   (b) $\lim\limits_{x\to 3^+} f(x)$

   (c) $\lim\limits_{x\to 3} f(x)$

15. $f(x) = \dfrac{x^2 - 1}{x^2 - x - 6}$

16. $f(x) = \dfrac{x^2 + 5x - 36}{x^3 - 5x^2 + 3x + 9}$

17. $f(x) = \dfrac{x^2 - 11x + 30}{x^3 - 4x^2 - 3x + 18}$

18. $f(x) = \dfrac{x^2 - 9x + 18}{x^2 - x - 6}$

In Exercises 19–26, identify the horizontal and vertical asymptotes, if any, of the given function.

19. $f(x) = \dfrac{2x^2 - 2x - 4}{x^2 + x - 20}$

20. $f(x) = \dfrac{-3x^2 - 9x - 6}{5x^2 - 10x - 15}$

21. $f(x) = \dfrac{x^2 + x - 12}{7x^3 - 14x^2 - 21x}$

22. $f(x) = \dfrac{x^2 - 9}{9x - 9}$

23. $f(x) = \dfrac{x^2 - 9}{9x + 27}$

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

25. $f(x) = \dfrac{2x^4 + 3}{\sqrt{x^8 + 9}}$

26. $f(x) = \dfrac{3x^3 + 4}{\sqrt{x^6 + 3}}$

In Exercises 27–34, evaluate the given limit.

27. $\displaystyle\lim_{x \to \infty} \dfrac{x^3 + 2x^2 + 1}{x - 5}$

28. $\displaystyle\lim_{x \to \infty} \dfrac{x^3 + 2x^2 + 1}{5 - x}$

29. $\displaystyle\lim_{x \to -\infty} \dfrac{x^3 + 2x^2 + 1}{x^2 - 5}$

30. $\displaystyle\lim_{x \to -\infty} \dfrac{x^3 + 2x^2 + 1}{5 - x^2}$

31. $\displaystyle\lim_{x \to -\infty} \dfrac{\sqrt{4x^2 - 3x + 6}}{3x - 1}$

32. $\displaystyle\lim_{x \to -\infty} \dfrac{\sqrt{9x^6 + 4x^2 + 25}}{3x^3 + 4x + 5}$

33. $\displaystyle\lim_{x \to -\infty} \dfrac{\sqrt{10x^{10} - 4x^4 + 9}}{2x^5 + 2x^2 + 3}$

34. $\displaystyle\lim_{x \to -\infty} \dfrac{\sqrt{25x^4 + 16x^2 + 9}}{10x^2 + 6x + 12}$

Review

35. Use an $\varepsilon - \delta$ proof to show that
$$\lim_{x \to 1} 5x - 2 = 3.$$

36. Let $\displaystyle\lim_{x \to 2} f(x) = 3$ and $\displaystyle\lim_{x \to 2} g(x) = -1$. Evaluate the following limits.

   (a) $\displaystyle\lim_{x \to 2} (f + g)(x)$ (c) $\displaystyle\lim_{x \to 2} (f/g)(x)$

   (b) $\displaystyle\lim_{x \to 2} (fg)(x)$ (d) $\displaystyle\lim_{x \to 2} f(x)^{g(x)}$

37. Evaluate the limit: $\displaystyle\lim_{x \to e} \ln x$.

1.6   Continuity

As we have studied limits, we have gained the intuition that limits measure "where a function is heading." That is, if $\lim_{x \to 1} f(x) = 3$, then as x is close to 1, $f(x)$ is close to 3. We have seen, though, that this is not necessarily a good indicator of what $f(1)$ actually is. This can be problematic; functions can tend to one value but attain another. This section focuses on functions that *do not* exhibit such behavior.

Definition 7   Continuous Function

Let f be a function defined on an open interval I containing c.

1. f is **continuous at** c if $\lim_{x \to c} f(x) = f(c)$.

2. f is **continuous on** I if f is continuous at c for all values of c in I. If f is continuous on $(-\infty, \infty)$, we say f is **continuous everywhere**.

A useful way to establish whether or not a function f is continuous at c is to verify the following three things:

1. $\lim_{x \to c} f(x)$ exists,

2. $f(c)$ is defined, and

3. $\lim_{x \to c} f(x) = f(c)$.

If f is defined near c but is not continuous at c, then we say that f is **discontinuous at c** or f has a **discontinuity at c**. We will discuss three types of discontinuities, as seen in Figure 1.30.
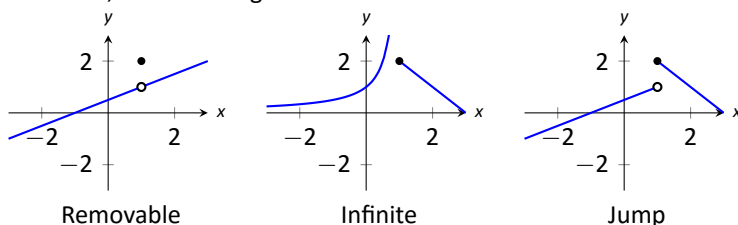
Figure 1.30: Three types of discontinuities

Removable discontinuity   This type of discontinuity is called removable because we could remove the discontinuity by redefining the function at a single point.

Notes:

Infinite discontinuity The function is approaching $\pm\infty$ at some *x* value.

Jump discontinuity The function "jumps" from one value to another.

Watch the video:
Continuity and Limits Made Easy — Part 1 of 2 at
`https://youtu.be/hlorAjSOxWE`

Example 1 **Finding intervals of continuity**
Let *f* be defined as shown in Figure 1.31. Give the interval(s) on which *f* is continuous.

SOLUTION We proceed by examining the three criteria for continuity.

1. The limits $\lim_{x\to c} f(x)$ exists for all *c* between 0 and 3.

2. $f(c)$ is defined for all *c* between 0 and 3, *except for c* $= 1$. We know immediately that *f* cannot be continuous at $x = 1$.

3. The limit $\lim_{x\to c} f(x) = f(c)$ for all *c* between 0 and 3, except, of course, for $c = 1$.

   We conclude that *f* is continuous at every point of $(0, 3)$ except at $x = 1$. Therefore *f* is continuous on $(0, 1) \cup (1, 3)$.

Figure 1.31: A graph of *f* in Example 1.

Example 2 **Finding intervals of continuity**
The *floor function*, $f(x) = \lfloor x \rfloor$, returns the largest integer smaller than or equal to the input *x*. (For example, $f(\pi) = \lfloor \pi \rfloor = 3$.) The graph of *f* in Figure 1.32 demonstrates why this is often called a "step function."
Give the intervals on which *f* is continuous.

SOLUTION We examine the three criteria for continuity.

1. The limits $\lim_{x\to c} f(x)$ do not exist at the jumps from one "step" to the next, which occur at all integer values of *c*. Therefore the limits exist for all *c* except when *c* is an integer.

2. The function is defined for all values of *c*.

Figure 1.32: A graph of the step function in Example 2.

Notes:

3. The limit $\lim_{x \to c} f(x) = f(c)$ for all values of c where the limit exist, since each step consists of just a line.

We conclude that f is continuous everywhere except at integer values of c. So the intervals on which f is continuous are

$$\ldots, (-2, -1), (-1, 0), (0, 1), (1, 2), \ldots.$$

Our definition of continuity on an interval specifies the interval is an open interval. At endpoints or points of discontinuity we may consider continuity from the right or left.

Definition 8   Right and Left Continuity
Right Continuous
Let f be defined on a closed interval with left endpoint a. We say that f is **continuous from the right at** a (or **right continuous at** a) if

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

Left Continuous
Let f be defined on a closed interval with right endpoint b. We say that f is **continuous from the left at** b (or **left continuous at** b) if

$$\lim_{x \to b^-} f(x) = f(b).$$

We can then extend the definition of continuity to closed intervals by considering the appropriate one-sided limits at the endpoints.

Definition 9   Continuity on Closed Intervals
Let f be defined on the closed interval $[a, b]$ for some real numbers a, b. Then f is **continuous on** $[a, b]$ if:

1. f is continuous on (a, b),

2. f is right continuous at a and

3. f is left continuous at b.

Notes:

We can make the appropriate adjustments to talk about continuity on half–open intervals such as $[a, b)$ or $(a, b]$ if necessary.

Continuity is inherently tied to the properties of limits. Because of this, the properties of limits found in Theorems 1 and 3 apply to continuity as well. We will utilize these properties in the following example.

Example 3   **Determining intervals on which a function is continuous**
For each of the following functions, give the domain of the function and the interval(s) on which it is continuous.

1. $f(x) = 1/x$

2. $f(x) = \sin x$

3. $f(x) = \sqrt{x}$

4. $f(x) = \sqrt{1 - x^2}$

5. $f(x) = |x|$

SOLUTION   We examine each in turn.

1. The domain of $f(x) = 1/x$ is $(-\infty, 0) \cup (0, \infty)$. As it is a rational function, we apply Theorem 3 to recognize that f is continuous on all of its domain.

2. The domain of $f(x) = \sin x$ is all real numbers, or $(-\infty, \infty)$. Applying Theorem 4 shows that $\sin x$ is continuous everywhere.

3. The domain of $f(x) = \sqrt{x}$ is $[0, \infty)$. Applying Theorem 4 shows that $f(x) = \sqrt{x}$ is continuous on its domain of $[0, \infty)$.

4. The domain of $f(x) = \sqrt{1 - x^2}$ is $[-1, 1]$. Applying Theorems 1 and 4 shows that f is continuous on all of its domain, $[-1, 1]$.

5. The domain of $f(x) = |x|$ is $(-\infty, \infty)$. We can define the absolute value function as $f(x) = \begin{cases} -x & x < 0 \\ x & x \geq 0 \end{cases}$. Each "piece" of this piecewise defined function is continuous on all of its domain, giving that f is continuous on $(-\infty, 0)$ and $[0, \infty)$. We cannot assume this implies that f is continuous on $(-\infty, \infty)$; we need to check that $\lim\limits_{x \to 0} f(x) = f(0)$, as $x = 0$ is the point where f transitions from one "piece" of its definition to the other. It is easy to verify that this is indeed true, hence we conclude that $f(x) = |x|$ is continuous everywhere.

The following theorem states how continuous functions can be combined to form other continuous functions.

Notes:

Theorem 10   Properties of Continuous Functions
Let f and g be continuous functions on an interval I, let c be a real number and let n be a positive integer. The following functions are continuous on I.

1. Sums/Differences: $f \pm g$

2. Constant Multiples: $c \cdot f$

3. Products: $f \cdot g$

4. Quotients: f/g (as long as $g \neq 0$ on I)

5. Powers: f^n

6. Roots: $\sqrt[n]{f}$ (if $f \geq 0$ on I or n is odd)

The proofs of each of the parts of Theorem 10 follow from the Basic Limit Properties given in Theorem 1. We will prove the product of two continuous functions is continuous now.

Proof
We know that f and g are continuous at c so by definition we have

$$\lim_{x \to c} f(x) = f(c) \quad \text{and} \quad \lim_{x \to c} g(x) = g(c).$$

Therefore,

$$\begin{aligned}
\lim_{x \to c} (f \cdot g)(x) &= \lim_{x \to c} f(x) \cdot g(x) \\
&= \lim_{x \to c} f(x) \cdot \lim_{x \to c} g(x) \\
&= f(c) \cdot g(c) \\
&= (f \cdot g)(c). \qquad \square
\end{aligned}$$

Theorem 11   Continuity of Compositions
Let f be continuous on I, where the range of f on I is J, and let g be continuous on J. Then

$$(g \circ f)(x) = g(f(x))$$

is continuous on I.

Notes:

Now knowing the definition of continuity we can re–read Theorem 4 as giving a list of functions that are continuous on their domains.

Theorem 12   Continuous Functions
The following functions are continuous on their domains.

1. $f(x) = \sin x$ 2. $f(x) = \cos x$

3. $f(x) = \tan x$ 4. $f(x) = \cot x$

5. $f(x) = \sec x$ 6. $f(x) = \csc x$

7. $f(x) = \ln x$ 8. $f(x) = a^x \ (a > 0)$

In the following example, we will show how we apply the previous theorems.

Example 4   Determining intervals on which a function is continuous
State the interval(s) on which each of the following functions is continuous.

1. $f(x) = \sqrt{x - 1} + \sqrt{5 - x}$ 3. $f(x) = \tan x$

2. $f(x) = x \sin x$ 4. $f(x) = \sqrt{\ln x}$

SOLUTION   We examine each in turn, applying Theorems 10 and 12 as appropriate.

1. The square–root terms are continuous on the intervals $[1, \infty)$ and $(-\infty, 5]$, respectively. As f is continuous only where each term is continuous, f is continuous on $[1, 5]$, the intersection of these two intervals. A graph of f is given in Figure 1.33.

2. The functions $y = x$ and $y = \sin x$ are each continuous everywhere, hence their product is, too.

3. Theorem 12 states that $f(x) = \tan x$ is continuous "on its domain." Its domain includes all real numbers except odd multiples of $\pi/2$. Thus $f(x) = \tan x$ is continuous on

$$\ldots \left(-\frac{3\pi}{2}, -\frac{\pi}{2}\right), \ \left(-\frac{\pi}{2}, \frac{\pi}{2}\right), \ \left(\frac{\pi}{2}, \frac{3\pi}{2}\right), \ldots,$$

or, equivalently, on $D = \{x \in \mathbb{R} \mid x \neq \frac{(2n+1)\pi}{2}, \ n \in \mathbb{Z}\}$.
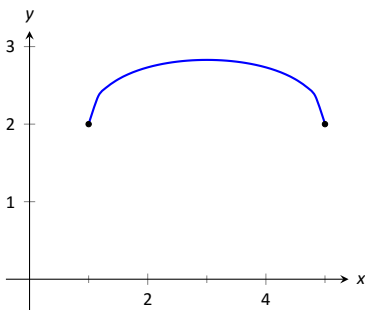
Figure 1.33: A graph of f in Example 4(1).

Notes:

4. The domain of $y = \sqrt{x}$ is $[0, \infty)$. The range of $y = \ln x$ is $(-\infty, \infty)$, but if we restrict its domain to $[1, \infty)$ its range is $[0, \infty)$. So restricting $y = \ln x$ to the domain of $[1, \infty)$ restricts its output is $[0, \infty)$, on which $y = \sqrt{x}$ is defined. Thus the domain of $f(x) = \sqrt{\ln x}$ is $[1, \infty)$.

A common way of thinking of a continuous function is that "its graph can be sketched without lifting your pencil." That is, its graph forms a "continuous" curve, without holes, breaks or jumps. While beyond the scope of this text, this pseudo–definition glosses over some of the finer points of continuity. Very strange functions are continuous that one would be hard pressed to actually sketch by hand.

This intuitive notion of continuity does help us understand another important concept as follows. Suppose f is defined on $[1, 2]$ and $f(1) = -10$ and $f(2) = 5$. If f is continuous on $[1, 2]$ (i.e., its graph can be sketched as a continuous curve from $(1, -10)$ to $(2, 5)$) then we know intuitively that somewhere on $[1, 2]$ f must be equal to -9, and -8, and -7, -6, . . ., 0, 1/2, etc. In short, f takes on all *intermediate* values between -10 and 5. It may take on more values; f may actually equal 6 at some time, for instance, but we are guaranteed all values between -10 and 5.

While this notion seems intuitive, it is not trivial to prove and its importance is profound. Therefore the concept is stated in the form of a theorem and illustrated in Figure 1.34.

Theorem 13 Intermediate Value Theorem

Let f be a continuous function on $[a, b]$ and, without loss of generality, let $f(a) < f(b)$. Then for every value y, where $f(a) < y < f(b)$, there is a value c in (a, b) such that $f(c) = y$

One important application of the Intermediate Value Theorem is root finding. Given a function f, we are often interested in finding values of x where $f(x) = 0$. These roots may be very difficult to find exactly. Good approximations can be found through successive applications of this theorem. Suppose through direct computation we find that $f(a) < 0$ and $f(b) > 0$, where $a < b$. The Intermediate Value Theorem states that there is a c in $[a, b]$ such that $f(c) = 0$. The theorem does not give us any clue as to where that value is in the interval $[a, b]$, just that it exists.

Example 5 Finding roots

Show that $f(x) = x^3 + x + 3$ has at least one real root.
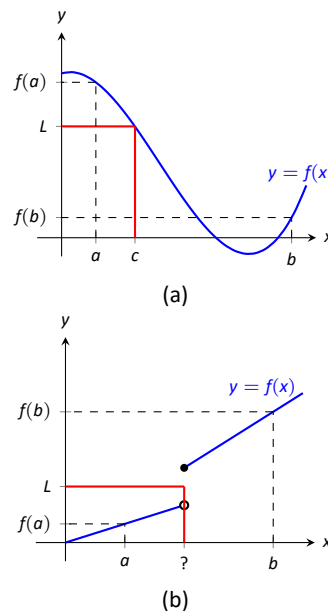
Figure 1.34: A situation where the Intermediate Value Theorem applies (top) and does not (bottom).

Notes:

SOLUTION   We must determine an interval on which the function changes from positive to negative values. We start by evaluating f at different values. We see that $f(0) = 3 > 0$ and $f(1) = 5 > 0$. As we choose larger positive values of x, we can see that $f(x)$ values will continue to grow. Looking at negative x-values, $f(-1) = 1 > 0$ and $f(-2) = -7 < 0$ so we know $f(x)$ must change sign in $[-2, -1]$. Because $f(x)$ is a polynomial, it is continuous on all real numbers so is continuous on $[-2, -1]$. By the Intermediate Value Theorem there is a c in $(-2, -1)$ where $f(x) = 0$. Thus $f(x)$ must have at least one real root on $(-2, -1)$.

Note that in the above example you were not asked to find the root, just to show that the function *had* a root.

There is a technique that produces a good approximation of c. Let d be the midpoint of the interval $[a, b]$ and consider $f(d)$. There are three possibilities:

1. $f(d) = 0$ – we got lucky and stumbled on the actual value. We stop as we found a root.

2. $f(d) < 0$ Then we know there is a root of f on the interval $[d, b]$ – we have halved the size of our interval, hence are closer to a good approximation of the root.

3. $f(d) > 0$ Then we know there is a root of f on the interval $[a, d]$ – again,we have halved the size of our interval, hence are closer to a good approximation of the root.

Successively applying this technique is called the **Bisection Method** of root finding. We continue until the interval is sufficiently small. We demonstrate this in the following example.

Example 6   **Using the Bisection Method**
Approximate the root of $f(x) = x - \cos x$, accurate to three places after the decimal.

SOLUTION   Consider the graph of $f(x) = x - \cos x$, shown in Figure 1.35. It is clear that the graph crosses the x-axis somewhere near $x = 0.8$. To start the Bisection Method, pick an interval that contains 0.8. We choose $[0.7, 0.9]$. Note that all we care about are signs of $f(x)$, not their actual value, so this is all we display.

Iteration 1: $f(0.7) < 0, f(0.9) > 0$, and $f(0.8) > 0$. So replace 0.9 with 0.8 and repeat.

Figure 1.35: Graphing a root of $f(x) = x - \cos x$.

Notes:

Iteration 2: $f(0.7) < 0$, $f(0.8) > 0$, and at the midpoint, 0.75, we see that $f(0.75) > 0$. So replace 0.8 with 0.75 and repeat. Note that we don't need to continue to check the endpoints, just the midpoint. Thus we put the rest of the iterations in Figure 1.36.

Notice that in the 12th iteration we have the endpoints of the interval each starting with 0.739. Thus we have narrowed the zero down to an accuracy of the first three places after the decimal. Using a computer, we have

$$f(0.7390) = -0.00014, \quad f(0.7391) = 0.000024.$$

Either endpoint of the interval gives a good approximation of where f is 0. The Intermediate Value Theorem states that the actual zero is still within this interval. While we do not know its exact value, we know it starts with 0.739.

This type of exercise is rarely done by hand. Rather, it is simple to program a computer to run such an algorithm and stop when the endpoints differ by a preset small amount. One of the authors did write such a program and found the zero of f, accurate to 10 places after the decimal, to be 0.7390851332. While it took a few minutes to write the program, it took less than a thousandth of a second for the program to run the necessary 35 iterations. In less than 8 hundredths of a second, the zero was calculated to 100 decimal places (with less than 200 iterations).

It is a simple matter to extend the Bisection Method to solve problems similar to "Find x, where $f(x) = 0$." For instance, we can find x, where $f(x) = 1$. It actually works very well to define a new function g where $g(x) = f(x) - 1$. Then use the Bisection Method to solve $g(x) = 0$.

Similarly, given two functions f and g, we can use the Bisection Method to solve $f(x) = g(x)$. Once again, create a new function h where $h(x) = f(x) - g(x)$ and solve $h(x) = 0$.

In Section 4.4 another equation solving method will be introduced, called Newton's Method. In many cases, Newton's Method is much faster. It relies on more advanced mathematics, though, so we will wait before introducing it.

This section formally defined what it means to be a continuous function. "Most" functions that we deal with are continuous, so often it feels odd to have to formally define this concept. Regardless, it is important, and forms the basis of the next chapter.

Chapter Summary

In this chapter we:

- defined the limit,

Itera-tion #	Interval	Midpoint Sign
1	$[0.7, 0.9]$	$f(0.8) > 0$
2	$[0.7, 0.8]$	$f(0.75) > 0$
3	$[0.7, 0.75]$	$f(0.725) < 0$
4	$[0.725, 0.75]$	$f(0.7375) < 0$
5	$[0.7375, 0.75]$	$f(0.7438) > 0$
6	$[0.7375, 0.7438]$	$f(0.7407) > 0$
7	$[0.7375, 0.7407]$	$f(0.7391) > 0$
8	$[0.7375, 0.7391]$	$f(0.7383) < 0$
9	$[0.7383, 0.7391]$	$f(0.7387) < 0$
10	$[0.7387, 0.7391]$	$f(0.7389) < 0$
11	$[0.7389, 0.7391]$	$f(0.7390) < 0$
12	$[0.7390, 0.7391]$	

Figure 1.36: Iterations of the Bisection Method of Root Finding

Notes:

- found accessible ways to approximate their values numerically and graphically,

- developed a method of proving the value of a limit (ε-δ proofs),

- explored when limits do not exist,

- considered limits that involved infinity, and

- defined continuity and explored properties of continuous functions.

Why? Mathematics is famous for building on itself and calculus proves to be no exception. In the next chapter we will be interested in "dividing by 0." That is, we will want to divide a quantity by a smaller and smaller number and see what value the quotient approaches. In other words, we will want to find a limit. These limits will enable us to, among other things, determine *exactly* how fast something is moving when we are only given position information.

Later, we will want to add up an infinite list of numbers. We will do so by first adding up a finite list of numbers, then take a limit as the number of things we are adding approaches infinity. Surprisingly, this sum often is finite; that is, we can add up an infinite list of numbers and get, for instance, 42.

These are just two quick examples of why we are interested in limits. Many students dislike this topic when they are first introduced to it, but over time an appreciation is often formed based on the scope of its applicability.

Notes:

Exercises 1.6

Terms and Concepts

1. In your own words, describe what it means for a function to be continuous.
2. In your own words, describe what the Intermediate Value Theorem states.
3. What is a "root" of a function?
4. Given functions f and g on an interval I, how can the Bisection Method be used to find a value c where $f(c) = g(c)$?
5. T/F: If f is defined on an open interval containing c, and $\lim_{x \to c} f(x)$ exists, then f is continuous at c.
6. T/F: If f is continuous at c, then $\lim_{x \to c} f(x)$ exists.
7. T/F: If f is continuous at c, then $\lim_{x \to c^+} f(x) = f(c)$.
8. T/F: If f is continuous on $[a, b]$, then $\lim_{x \to a^-} f(x) = f(a)$.
9. T/F: If f is continuous on $[0, 1)$ and $[1, 2)$, then f is continuous on $[0, 2)$.
10. T/F: The sum of continuous functions is also continuous.

Problems

In Exercises 11–17, a graph of a function f is given along with a value a. Determine if f is continuous at a; if it is not, state why it is not.

11. $a = 1$
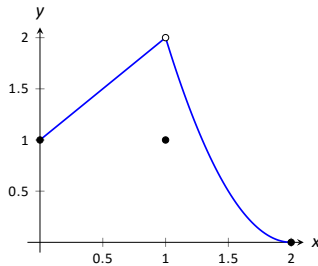
12. $a = 1$
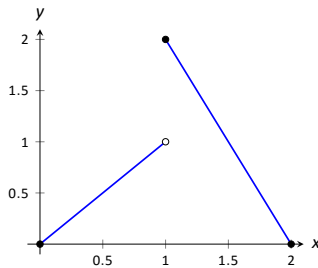
13. $a = 1$
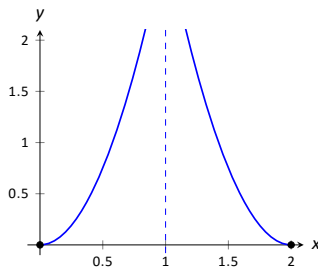
14. $a = 0$
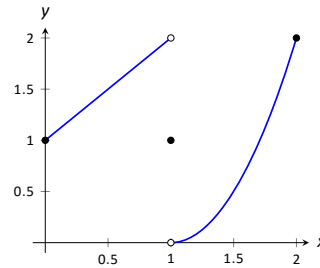
15. $a = 1$
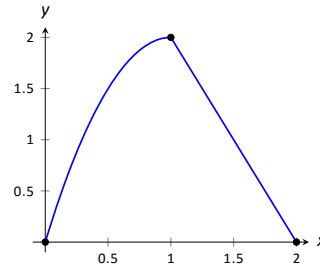
16. $a = 4$
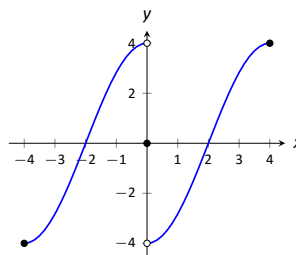
17.
   (a) $a = -2$
   (b) $a = 0$
   (c) $a = 2$
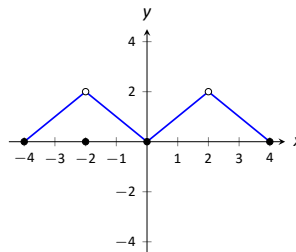
In Exercises 18–21, determine if f is continuous at the indicated values. If not, explain why.

18. $f(x) = \begin{cases} 1 & x = 0 \\ \frac{\sin x}{x} & x > 0 \end{cases}$

   (a) $x = 0$
   (b) $x = \pi$

19. $f(x) = \begin{cases} x^3 - x & x < 1 \\ x - 2 & x \geq 1 \end{cases}$

   (a) $x = 0$
   (b) $x = 1$

20. $f(x) = \begin{cases} \frac{x^2+5x+4}{x^2+3x+2} & x \neq -1 \\ 3 & x = -1 \end{cases}$

    (a) $x = -1$

    (b) $x = 10$

21. $f(x) = \begin{cases} \frac{x^2-64}{x^2-11x+24} & x \neq 8 \\ 5 & x = 8 \end{cases}$

    (a) $x = 0$

    (b) $x = 8$

In Exercises 22–34, give the intervals on which the given function is continuous.

22. $f(x) = x^2 - 3x + 9$

23. $g(x) = \sqrt{x^2 - 4}$

24. $h(k) = \sqrt{1 - k} + \sqrt{k + 1}$

25. $f(t) = \sqrt{5t^2 - 30}$

26. $g(t) = \dfrac{1}{\sqrt{1 - t^2}}$

27. $g(x) = \dfrac{1}{1 + x^2}$

28. $f(x) = e^x$

29. $g(s) = \ln s$

30. $h(t) = \cos t$

31. $f(k) = \sqrt{1 - e^k}$

32. $f(x) = \sin(e^x + x^2)$

33. $f(x) = \begin{cases} \frac{x+1}{x+4} & x < 2 \\ x^2 - 3 & 2 \leq x \leq 5 \\ 6 - 2x & x > 5 \end{cases}$

34. $f(x) = \begin{cases} \frac{1}{x-1} & x < 0 \\ 2x^2 - 3x - 1 & 0 \leq x \leq 2 \\ 5x^2 - 4x & x > 2 \end{cases}$

35. Let $f(x) = \begin{cases} x^2 - 1 & x < 3 \\ x + 5 & x \geq 3 \end{cases}$.

    Is f continuous everywhere?

36. Let f be continuous on $[1, 5]$ where $f(1) = -2$ and $f(5) = -10$. Does a value $1 < c < 5$ exist such that $f(c) = -9$? Why/why not?

37. Let g be continuous on $[-3, 7]$ where $g(0) = 0$ and $g(2) = 25$. Does a value $-3 < c < 7$ exist such that $g(c) = 15$? Why/why not?

38. Let f be continuous on $[-1, 1]$ where $f(-1) = -10$ and $f(1) = 10$. Does a value $-1 < c < 1$ exist such that $f(c) = 11$? Why/why not?

39. Let h be a function on $[-1, 1]$ where $h(-1) = -10$ and $h(1) = 10$. Does a value $-1 < c < 1$ exist such that $h(c) = 0$? Why/why not?

In Exercises 40–43, find the value(s) of a and b so that the function is continuous on $\mathbb{R}$.

40. $g(x) = \begin{cases} ax^2 + 3x & x < 2 \\ x^3 - ax & x \geq 2 \end{cases}$

41. $f(x) = \begin{cases} a^2x - ax & x > 3 \\ 4 & x \leq 3 \end{cases}$

42. $f(x) = \begin{cases} ax - b & x < -1 \\ 2x^2 + 3ax + b & -1 \leq x < 1 \\ 4 & x \geq 1 \end{cases}$

43. $f(x) = \begin{cases} x^2 + 2x & x \leq a \\ -1 & x > a \end{cases}$

In Exercises 44–47, sketch the graph of a function that has the following properties.

44. f is discontinuous at 3, but continuous from the left at 3, and continuous elsewhere.

45. f is discontinuous at -1 and 2, but continuous from the right at -1 and continuous from the left at 2, and continuous elsewhere.

46. f has a jump discontinuity at -2 and an infinite discontinuity at 4 and is continuous elsewhere.

47. f has a removable discontinuity at 2, is continuous only from the left at 5, and is continuous elsewhere.

In Exercises 48–51, show that the functions have at least one real root.

48. $f(x) = x^2 + 2x - 4$

49. $f(x) = \sin x - 1/2$

50. $f(x) = e^x - 2$

51. $f(x) = \cos x - \sin x$

Review

52. Let $f(x) = \begin{cases} x^2 - 5 & x < 5 \\ 5x & x \geq 5 \end{cases}$.

    (a) $\lim\limits_{x \to 5^-} f(x)$ (c) $\lim\limits_{x \to 5} f(x)$

    (b) $\lim\limits_{x \to 5^+} f(x)$ (d) $f(5)$

53. Numerically approximate the following limits:

    (a) $\lim\limits_{x \to -4/5^+} \dfrac{x^2 - 8.2x - 7.2}{x^2 + 5.8x + 4}$

    (b) $\lim\limits_{x \to -4/5^-} \dfrac{x^2 - 8.2x - 7.2}{x^2 + 5.8x + 4}$

54. Give an example of function $f(x)$ for which $\lim\limits_{x \to 0} f(x)$ does not exist.

2.0   Chapter Prerequisites

The material in this section provides a basic review of and practice problems for pre-calculus skills essential to your success in Calculus. You should take time to review this section and work the suggested problems (checking your answers against those in the back of the book). Since this content is a pre-requisite for Calculus, reviewing and mastering these skills are considered your responsibility. This means that minimal, and in some cases no, class time will be devoted to this section. When you identify areas that you need help with we strongly urge you to seek assistance outside of class from your instructor or other student tutoring service.

Rules of Exponents

We will briefly summarize the laws of exponents and equivalent forms of exponent expressions commonly used in this chapter. The laws of exponents are only valid for the values of x and y for which the expression is defined (i.e., nonzero real numbers in the denominator and nonnegative real numbers when roots are even.) Our first is the product of exponents. If m and n are real numbers, then

$$x^m \cdot x^n = x^{m+n}.$$

Example 1   **Product Law of Exponents**

$$x^5 \cdot x^7 = x^{5+7} = x^{12}$$

$$x^{-3} \cdot x^{-4} = x^{-3+(-4)} = x^{-7} = \frac{1}{x^7}$$

$$x^{-1/2} \cdot x^{2/3} = x^{-1/2+2/3} = x^{1/6} = \sqrt[6]{x}.$$

Our next is the quotient of exponents. If m and n are real numbers, then

$$\frac{x^m}{x^n} = x^{m-n}.$$

Example 2   **Quotient Law of Exponents**

$$\frac{x^5}{x^7} = x^{5-7} = x^{-2} = \frac{1}{x^2}$$

$$\frac{x^{-3}}{x^{-4}} = x^{-3-(-4)} = x^1 = x$$

$$\frac{x^{2/3}}{x^{-1/2}} = x^{2/3-(-1/2)} = x^{7/6} = \sqrt[6]{x^7} = x\sqrt[6]{x}.$$

Our third is when a power is raised to a power. Once again, we assume m and n are real numbers. In that case,

$$(x^m)^n = x^{m \cdot n}.$$

Example 3   **Power Law of Exponents**

$$(x^5)^7 = x^{5 \cdot 7} = x^{35}$$

$$(x^{-3})^4 = x^{-3 \cdot 4} = x^{-12} = \frac{1}{x^{12}}$$

$$(x^{-1/2})^{2/3} = x^{(-1/2) \cdot (2/3)} = x^{-1/3} = \frac{1}{\sqrt[3]{x}}.$$

Our final law tells us how to distribute a power over a product and a quotient. If m is a real number, then

$$(xy)^m = x^m y^m \qquad \text{and} \qquad \left(\frac{x}{y}\right)^m = \frac{x^m}{y^m}.$$

Example 4   **Product and Quotient Raised to a Power**

$$(xyz)^7 = x^7 y^7 z^7$$

$$\left(\frac{x}{y}\right)^{-4} = \frac{x^{-4}}{y^{-4}} = \frac{y^4}{x^4}.$$

Factoring and Simplifying Complex Fractions

The following examples demonstrate an efficient factoring technique that can be used to create the various equivalent expressions often needed to complete problems that arise in Calculus. The ability to move flexibly and efficiently among different representations of an expression is an important skill to have.

Example 5   **Factoring out the common factor**
Factor completely to write an equivalent expression:

    1. $x^{7/3} - 4x^{2/3}$ 2. $\frac{1}{2}x(x-3)^{-2/5} + (x-3)^{3/5}$

Notes:

Solution

1. $x^{7/3} - 4x^{2/3} = x^{2/3}(x^{5/3} - 4) = \sqrt[3]{x^2}(\sqrt[3]{x^5} - 4).$

2.

$$\frac{1}{2}x(x-3)^{-2/5} + (x-3)^{3/5} = \frac{1}{2}(x-3)^{-2/5}\big(1 + 2(x-3)\big)$$

$$= \frac{1}{2}(x-3)^{-2/5}(1 + 2x - 6)$$

$$= \frac{1}{2}(x-3)^{-2/5}(2x - 5)$$

$$= \frac{2x - 5}{2(x-3)^{2/5}} \qquad \text{or}$$

$$= \frac{2x - 5}{2\sqrt[5]{(x-3)^2}}$$

Example 6   Simplifying complex fractions

Factor out the lowest power of the common factor to simplify the complex fraction

$$\frac{\frac{2}{3}x(x-2)^{-\frac{1}{3}} + (x-2)^{\frac{2}{3}}}{x^2}.$$

Solution

$$\frac{\frac{2}{3}x(x-2)^{-1/3} + (x-2)^{2/3}}{x^2} = \frac{\frac{1}{3}(x-2)^{-1/3}\big(2x + 3(x-2)\big)}{x^2}$$

$$= \frac{2x + 3x - 6}{3x^2(x-2)^{1/3}}$$

$$= \frac{5x - 6}{3x^2\sqrt[3]{x-2}}$$

Function Composition

Function composition refers to combining functions in a way that the output from one function becomes the input for the next function. In other words, the range (*y*-values) of one function become the domain (*x*-values) of the next function. We denote this as $(f \circ g)(x) = f(g(x))$, where the output of $g(x)$ becomes the input of $f(x)$.

Example 7   Composition of two functions

Given $f(x) = \dfrac{1}{x^2}$ and $g(x) = \sqrt{x+4}$, find $(f \circ g)(x)$ and $(g \circ f)(x)$.

Notes:

SOLUTION   To find $(f \circ g)(x) = f(g(x))$, we substitute the function $g(x)$ into the function $f(x)$. Thus,

$$f(g(x)) = f\left(\sqrt{x+4}\right) = \frac{1}{(\sqrt{x+4})^2} = \frac{1}{x+4}.$$

For $(g \circ f)(x) = g(f(x))$, we substitute the function $f(x)$ into the function $g(x)$. Thus,

$$g(f(x)) = g\left(\frac{1}{x^2}\right) = \sqrt{\frac{1}{x^2} + 4} = \sqrt{\frac{1 + 4x^2}{x^2}} = \frac{\sqrt{1 + 4x^2}}{x}.$$

Example 8   **Composition of three functions**
Given $f(x) = x^2$, $g(x) = \sqrt{4-x}$ and $h(x) = 3x - 5$, find $(f \circ g \circ h)(x)$ and $(g \circ f \circ h)(x)$.

SOLUTION   To find $(f \circ g \circ h)(x)$ we must start with the inside and work our way out.

$$\begin{aligned}
(f \circ g \circ h)(x) &= f(g(h(x))) \\
&= f(g(3x - 5)) \\
&= f\left(\sqrt{4 - (3x - 5)}\right) = f\left(\sqrt{9 - 3x}\right) \\
&= \left(\sqrt{9 - 3x}\right)^2 = 9 - 3x
\end{aligned}$$

For $(g \circ f \circ h)(x)$, we have

$$\begin{aligned}
(g \circ f \circ h)(x) &= g(f(h(x))) \\
&= g(f(3x - 5)) \\
&= g((3x - 5)^2) = g(9x^2 - 30x + 25) \\
&= \sqrt{4 - (9x^2 - 30x + 25)} = \sqrt{30x - 9x^2 - 21}
\end{aligned}$$

In this chapter we will also need to decompose a given function into two or more, less complex functions. For any one function there is often more than one way to write the decomposition. The following examples demonstrate this.

Example 9   **Decomposing a function**
Given $F(x) = \sin(3x^2 + 5)$, find $f(x)$ and $g(x)$ so that $F(x) = f(g(x))$.

SOLUTION   One solution is $f(x) = \sin x$ and $g(x) = 3x^2 + 5$. Another possible solution is $f(x) = \sin(x + 5)$ and $g(x) = 3x^2$.

Notes:

Exercises 2.0

Problems

In Exercises 1–4, simplify each expression. Write your answer so that all exponents are positive.

1. $(5x^4 y^5)(2x^2 y^3)^4$

2. $\left(\dfrac{4a^{3/2}b^3}{a^2 b^{-1/2}}\right)^{-2}$

3. $\dfrac{(-2x^{-3}y^7 z^5)^{-4}}{(x^3 y^{-2} z^5)^3}$

4. $\sqrt[4]{x^8 y^{16} z^{21}}$

In Exercises 5–7, factor to write equivalent expressions.

5. $\dfrac{5}{3}x^{\frac{2}{3}} - \dfrac{5}{3}x^{-\frac{1}{3}}$

6. $\dfrac{\frac{1}{2}x^{-\frac{1}{2}}(x+4) - 3x^{\frac{1}{2}}}{(x+4)^2}$

7. $6x(3x^2 + 2)^4(x^2 - 5)^2 + 24x(3x^2 + 2)^3(x^2 - 5)^3$

8. If $f(x) = x^2 + 2x$ and $g(x) = x - 4$ find
   (a) $(f \circ g)(6)$ (b) $(g \circ f)(6)$
   (c) $(f \circ g)(x)$ (d) $(g \circ f)(x)$

9. If $f(x) = \dfrac{1}{x-5}$ and $g(x) = \sqrt{x-2}$ find
   (a) $(f \circ g)(6)$ (b) $(g \circ f)(6)$
   (c) $(f \circ g)(x)$ (d) $(g \circ f)(x)$

10. If $F(x) = f(g(x))$ identify $f(x)$ and $g(x)$. (a) $F(x) = \frac{5}{x+4}$
    (b) $F(x) = |4 - x^2|$ (c) $F(x) = \sqrt{(x+2)^2 - 5}$

11. If $F(x) = f(g(h(x)))$ identify $f(x), g(x)$ and $h(x)$. (a) $F(x) = \sqrt[3]{(2x+1)^2}$ (b) $F(x) = 2\sqrt[3]{x^2} + 1$

Notes:

2: Derivatives

The previous chapter introduced the most fundamental of calculus topics: the limit. This chapter introduces the second most fundamental of calculus topics: the derivative. Limits describe *where* a function is going; derivatives describe *how fast* the function is going.

2.1 Instantaneous Rates of Change: The Derivative

A common amusement park ride lifts riders to a height then allows them to freefall a certain distance before safely stopping them. Suppose such a ride drops riders from a height of 150 feet. Students of physics may recall that the height (in feet) of the riders, t seconds after freefall (and ignoring air resistance, etc.) can be accurately modeled by $f(t) = -16t^2 + 150$.

Using this formula, it is easy to verify that, without intervention, the riders will hit the ground at $t = 2.5\sqrt{1.5} \approx 3.06$ seconds. Suppose the designers of the ride decide to begin slowing the riders' fall after 2 seconds (corresponding to a height of 86 ft.). How fast will the riders be traveling at that time?

We have been given a *position* function, but what we want to compute is a velocity at a specific point in time, i.e., we want an *instantaneous velocity*. We do not currently know how to calculate this.

However, we do know from common experience how to calculate an *average velocity*. (If we travel 60 miles in 2 hours, we know we had an average velocity of 30 mph.) We looked at this concept in Section 1.1 when we introduced the difference quotient. We have

$$\frac{\text{change in distance}}{\text{change in time}} = \frac{\text{" rise "}}{\text{run}} = \text{average velocity}.$$

We can approximate the instantaneous velocity at $t = 2$ by considering the average velocity over some time period containing $t = 2$. If we make the time interval small, we will get a good approximation. (This fact is commonly used. For instance, high speed cameras are used to track fast moving objects. Distances are measured over a fixed number of frames to generate an accurate approximation of the velocity.)

Consider the interval from $t = 2$ to $t = 3$ (just before the riders hit the ground). On that interval, the average velocity is

$$\frac{f(3) - f(2)}{3 - 2} = \frac{f(3) - f(2)}{1} = -80 \text{ ft/s},$$

where the minus sign indicates that the riders are moving *down*. By narrowing the interval we consider, we will likely get a better approximation of the instantaneous velocity. On $[2, 2.5]$ we have

$$\frac{f(2.5) - f(2)}{2.5 - 2} = \frac{f(2.5) - f(2)}{0.5} = -72 \text{ ft/s.}$$

We can do this for smaller and smaller intervals of time. For instance, over a time span of $1/10^{\text{th}}$ of a second, i.e., on $[2, 2.1]$, we have

$$\frac{f(2.1) - f(2)}{2.1 - 2} = \frac{f(2.1) - f(2)}{0.1} = -65.6 \text{ ft/s.}$$

Over a time span of $1/100^{\text{th}}$ of a second, on $[2, 2.01]$, the average velocity is

$$\frac{f(2.01) - f(2)}{2.01 - 2} = \frac{f(2.01) - f(2)}{0.01} = -64.16 \text{ ft/s.}$$

What we are really computing is the average velocity on the interval $[2, 2+h]$ for small values of h. That is, we are computing

$$\frac{f(2 + h) - f(2)}{h}$$

where h is the change in time after 2 seconds.

What we really want is for $h = 0$, but this, of course, returns the familiar "0/0" indeterminate form. So we employ a limit, as we did in Section 1.1.

We can approximate the value of this limit numerically with small values of h as seen in Figure 2.1. It looks as though the velocity is approaching -64 ft/s. Computing the limit directly gives

	Average Velocity
h	ft/s
1	-80
0.5	-72
0.1	-65.6
0.01	-64.16
0.001	-64.016

Figure 2.1: Approximating the instantaneous velocity with average velocities over a small time period h.

$$\lim_{h \to 0} \frac{f(2 + h) - f(2)}{h} = \lim_{h \to 0} \frac{-16(2 + h)^2 + 150 - (-16(2)^2 + 150)}{h}$$

$$= \lim_{h \to 0} \frac{-64h - 16h^2}{h}$$

$$= \lim_{h \to 0} -64 - 16h$$

$$= -64.$$

Graphically, we can view the average velocities we computed numerically as the slopes of secant lines on the graph of f going through the points $(2, f(2))$ and $(2 + h, f(2 + h))$, as in Figure 2.2. In Figure 2.3, the secant line corresponding to $h = 1$ is shown in three contexts. Figure 2.3(a) shows a "zoomed out" version of f with its secant line. In (b), we zoom in around the points of intersection between f and the secant line. Notice how well this secant line approximates f

Figure 2.2: Computing the difference quotient.

Notes:

between those two points – it is a common practice to approximate functions with straight lines.

As $h \to 0$, these secant lines approach the *tangent line*, a line that goes through the point $(2, f(2))$ with the special slope of -64. In parts (c) and (d) of Figure 2.3, we zoom in around the point $(2, 86)$. In (c) we see the secant line, which approximates f well, but not as well the tangent line shown in (d).

(a)

(b)

(c)

(d)
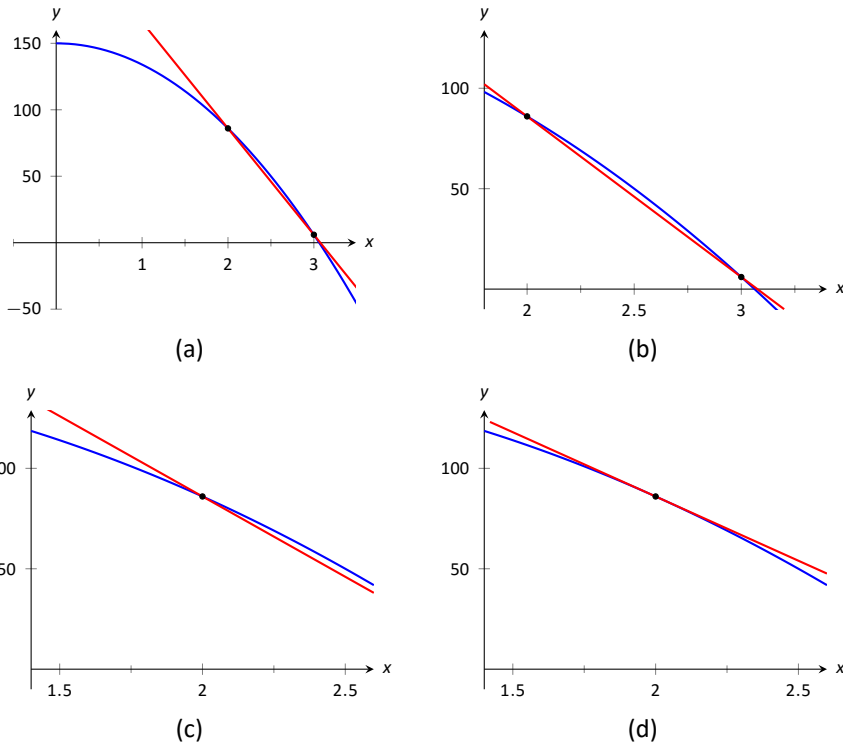
Figure 2.3: Parts (a), (b) and (c) show the secant line to $f(x)$ with $h = 1$, zoomed in different amounts. Part (d) shows the tangent line to f at $x = 2$.

We have just introduced a number of important concepts that we will flesh out more within this section. First, we formally define two of them.

Notes:

Definition 10   Derivative at a Point
Let f be a continuous function on an open interval I and let c be in I. The **derivative of f at c**, denoted $f'(c)$, is

$$\lim_{h \to 0} \frac{f(c + h) - f(c)}{h},$$

provided the limit exists.

If the limit exists, we say that f **is differentiable at** c; if the limit does not exist, then f **is not differentiable at** c. If f is differentiable at every point in I, then f **is differentiable on** I.

Definition 11   Tangent Line
Let f be continuous on an open interval I and differentiable at c, for some c in I. The line with equation $\ell(x) = f'(c)(x - c) + f(c)$ is the **tangent line** to the graph of f at c; that is, it is the line through $(c, f(c))$ whose slope is the derivative of f at c.

Watch the video:
The Difference Quotient — Example 1 at
`https://youtu.be/1O5NEI8UuHM`

Some examples will help us understand these definitions.

Example 1   Finding derivatives and tangent lines
Let $f(x) = 3x^2 + 5x - 7$. Find:

1. $f'(1)$

2. The equation of the tangent line to the graph of f at $x = 1$.

3. $f'(3)$

4. The equation of the tangent line to the graph f at $x = 3$.

Notes:

SOLUTION

1. We compute this directly using Definition 10.

$$\begin{aligned}
f'(1) &= \lim_{h \to 0} \frac{f(1+h) - f(1)}{h} \\
&= \lim_{h \to 0} \frac{3(1+h)^2 + 5(1+h) - 7 - (3(1)^2 + 5(1) - 7)}{h} \\
&= \lim_{h \to 0} \frac{3h^2 + 11h}{h} \\
&= \lim_{h \to 0} 3h + 11 = 11.
\end{aligned}$$

2. The tangent line at $x = 1$ has slope $f'(1)$ and goes through the point $(1, f(1)) = (1, 1)$. Thus the tangent line has equation, in point-slope form, $y = 11(x - 1) + 1$. In slope-intercept form we have $y = 11x - 10$.

3. Again, using the definition,

$$\begin{aligned}
f'(3) &= \lim_{h \to 0} \frac{f(3+h) - f(3)}{h} \\
&= \lim_{h \to 0} \frac{3(3+h)^2 + 5(3+h) - 7 - (3(3)^2 + 5(3) - 7)}{h} \\
&= \lim_{h \to 0} \frac{3h^2 + 23h}{h} \\
&= \lim_{h \to 0} 3h + 23 \\
&= 23.
\end{aligned}$$

4. The tangent line at $x = 3$ has slope 23 and goes through the point $(3, f(3)) = (3, 35)$. Thus the tangent line has equation $y = 23(x - 3) + 35 = 23x - 34$.

A graph of f is given in Figure 2.4 along with the tangent lines at $x = 1$ and $x = 3$.

Linear functions are easy to work with; many functions that arise in the course of solving real problems are not easy to work with. A common practice in mathematical problem solving is to approximate difficult functions with not–so–difficult functions. Lines are a common choice. It turns out that at any given point on the graph of a differentiable function f, the best linear approximation to f is its tangent line. That is one reason we'll spend considerable time finding tangent lines to functions.

One type of function that does not benefit from a tangent–line approximation is a line; it is rather simple to recognize that the tangent line to a line is the line itself. We look at this in the following example.
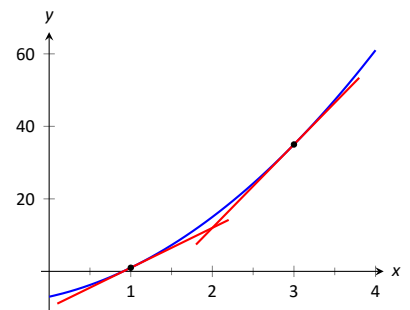
Figure 2.4: A graph of $f(x) = 3x^2 + 5x - 7$ and its tangent lines at $x = 1$ and $x = 3$.

Notes:

Example 2   **Finding the Derivative of a Line**

Consider $f(x) = 3x + 5$. Find the equation of the tangent line to f at $x = 1$ and $x = 7$.

SOLUTION   We find the slope of the tangent line by using Definition 10.

$$\begin{aligned}
f'(1) &= \lim_{h \to 0} \frac{f(1 + h) - f(1)}{h} \\
&= \lim_{h \to 0} \frac{3(1 + h) + 5 - (3 + 5)}{h} \\
&= \lim_{h \to 0} \frac{3h}{h} \\
&= \lim_{h \to 0} 3 \\
&= 3.
\end{aligned}$$

We just found that $f'(1) = 3$. That is, we found the *instantaneous rate of change* of $f(x) = 3x + 5$ is 3. This is not surprising; lines are characterized by being the *only* functions with a *constant rate of change.* That rate of change is called the *slope* of the line. Since their rates of change are constant, their *instantaneous* rates of change are always the same; they are all the slope.

So given a line $f(x) = ax + b$, the derivative at any point x will be a; that is, $f'(x) = a$.

It is now easy to see that the tangent line to the graph of f at $x = 1$ is just f, with the same being true for $x = 7$.

We often desire to find the tangent line to the graph of a function without knowing the actual derivative of the function. In these cases, the best we may be able to do is approximate the tangent line. We demonstrate this in the next example.

Example 3   **Numerical Approximation of the Tangent Line**

Approximate the equation of the tangent line to the graph of $f(x) = \sin x$ at $x = 0$.

SOLUTION   In order to find the equation of the tangent line, we need a slope and a point. The point is given to us: $(0, \sin 0) = (0, 0)$. To compute the slope, we need the derivative. This is where we will make an approximation. Recall that

$$f'(0) \approx \frac{\sin(0 + h) - \sin 0}{h}$$

for a small value of h. We choose (somewhat arbitrarily) to let $h = 0.1$. Thus
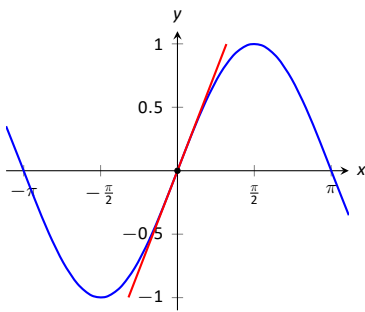
Figure 2.5: $f(x) = \sin x$ graphed with an approximation to its tangent line at $x = 0$.

Notes:

$$f'(0) \approx \frac{\sin(0.1) - \sin 0}{0.1} \approx 0.9983.$$

Thus our approximation of the equation of the tangent line is $y = 0.9983(x - 0) + 0 = 0.9983x$; it is graphed in Figure 2.5. The graph seems to imply the approximation is rather good.

Recall from Section 1.3 that $\lim\limits_{x \to 0} \dfrac{\sin x}{x} = 1$, meaning for values of x near 0, $\sin x \approx x$. Since the slope of the line $y = x$ is 1 at $x = 0$, it should seem reasonable that "the slope of $f(x) = \sin x$" is near 1 at $x = 0$. In fact, since we *approximated* the value of the slope to be 0.9983, we might guess the *actual value* is 1. We'll come back to this later.

Consider again Example 1. To find the derivative of f at $x = 1$, we needed to evaluate a limit. To find the derivative of f at $x = 3$, we needed to again evaluate a limit. We have this process:

input specific number c $\longrightarrow$ | do something to f and c | $\longrightarrow$ return number $f'(c)$
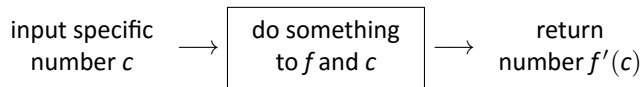
This process describes a *function*; given one input (the value of c), we return exactly one output (the value of $f'(c)$). The "do something" box is where the tedious work (taking limits) of this function occurs.

Instead of applying this function repeatedly for different values of c, let us apply it just once to the variable x. We then take a limit just once. The process now looks like:

input variable x $\longrightarrow$ | do something to f and x | $\longrightarrow$ return function $f'(x)$
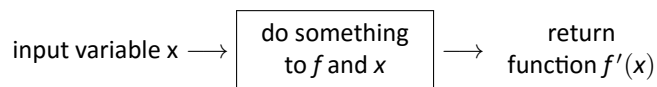
The output is the "derivative function," $f'(x)$. The $f'(x)$ function will take a number c as input and return the derivative of f at c. This calls for a definition.

Notes:

Definition 12   Derivative Function
Let f be a differentiable function on an open interval I. The function

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

is **the derivative of** f.

Notation:
Let $y = f(x)$. The following notations all represent the derivative:

$$f'(x) = y' = \frac{dy}{dx} = \frac{df}{dx} = \frac{d}{dx}(f) = \frac{d}{dx}(y).$$

Important: The notation $\dfrac{dy}{dx}$ is one symbol; it is **not** the fraction "dy/dx". The notation, while somewhat confusing at first, was chosen with care. A fraction–looking symbol was chosen because the derivative has many fraction–like properties. Among other places, we see these properties at work when we talk about the units of the derivative, when we discuss the Chain Rule, and when we learn about integration (topics that appear in later sections and chapters).

Examples will help us understand this definition.

Example 4   **Finding the derivative of a function**
Let $f(x) = 3x^2 + 5x - 7$ as in Example 1. Find $f'(x)$.

SOLUTION   We apply Definition 12.

$$f'(x) = \lim_{h \to 0} \frac{f(x+h) - f(x)}{h}$$
$$= \lim_{h \to 0} \frac{3(x+h)^2 + 5(x+h) - 7 - (3x^2 + 5x - 7)}{h}$$
$$= \lim_{h \to 0} \frac{3h^2 + 6xh + 5h}{h}$$
$$= \lim_{h \to 0} 3h + 6x + 5$$
$$= 6x + 5$$

So $f'(x) = 6x + 5$. Recall earlier we found that $f'(1) = 11$ and $f'(3) = 23$. Note our new computation of $f'(x)$ affirm these facts.

Notes:

Example 5   **Finding the derivative of a function**

Let $f(x) = \dfrac{1}{x+1}$. Find $f'(x)$.

SOLUTION   We apply Definition 12.

$$
\begin{aligned}
f'(x) &= \lim_{h \to 0} \frac{f(x+h) - f(x)}{h} \\
&= \lim_{h \to 0} \frac{\frac{1}{x+h+1} - \frac{1}{x+1}}{h}
\end{aligned}
$$

Now find a common denominator and subtract; factor $1/h$ out front to facilitate reading.

$$
\begin{aligned}
f'(x) &= \lim_{h \to 0} \frac{1}{h} \cdot \left(\frac{x+1}{(x+1)(x+h+1)} - \frac{x+h+1}{(x+1)(x+h+1)} \right) \\
&= \lim_{h \to 0} \frac{1}{h} \cdot \left(\frac{x+1-(x+h+1)}{(x+1)(x+h+1)} \right) \\
&= \lim_{h \to 0} \frac{1}{h} \cdot \left(\frac{-h}{(x+1)(x+h+1)} \right) \\
&= \lim_{h \to 0} \frac{-1}{(x+1)(x+h+1)} \\
&= \frac{-1}{(x+1)(x+1)} \\
&= \frac{-1}{(x+1)^2}
\end{aligned}
$$

So $f'(x) = \dfrac{-1}{(x+1)^2}$. To practice using our notation, we could also state

$$
\frac{d}{dx}\left(\frac{1}{x+1} \right) = \frac{-1}{(x+1)^2}.
$$

Example 6   **Finding the derivative of a function**

Find the derivative of $f(x) = \sin x$.

SOLUTION   Before applying Definition 12, note that once this is found, we can find the actual tangent line to $f(x) = \sin x$ at $x = 0$, whereas we settled

---

Notes:

for an approximation in Example 3.

$$f'(x) = \lim_{h \to 0} \frac{\sin(x+h) - \sin x}{h} \qquad \left(\begin{array}{c} \text{Use trig identity} \\ \sin(x+h) = \sin x \cos h + \cos x \sin h \end{array} \right)$$

$$= \lim_{h \to 0} \frac{\sin x \cos h + \cos x \sin h - \sin x}{h} \qquad \text{(regroup)}$$

$$= \lim_{h \to 0} \frac{\sin x(\cos h - 1) + \cos x \sin h}{h} \qquad \text{(split into two fractions)}$$

$$= \lim_{h \to 0} \left(\frac{\sin x(\cos h - 1)}{h} + \frac{\cos x \sin h}{h} \right) \qquad \left(\text{use } \lim_{h \to 0} \frac{\cos h - 1}{h} = 0 \text{ and } \lim_{h \to 0} \frac{\sin h}{h} = 1 \right)$$

$$= \sin x \cdot 0 + \cos x \cdot 1$$

$$= \cos x \,.$$

We have found that when $f(x) = \sin x$, $f'(x) = \cos x$ (see Figure 2.6).
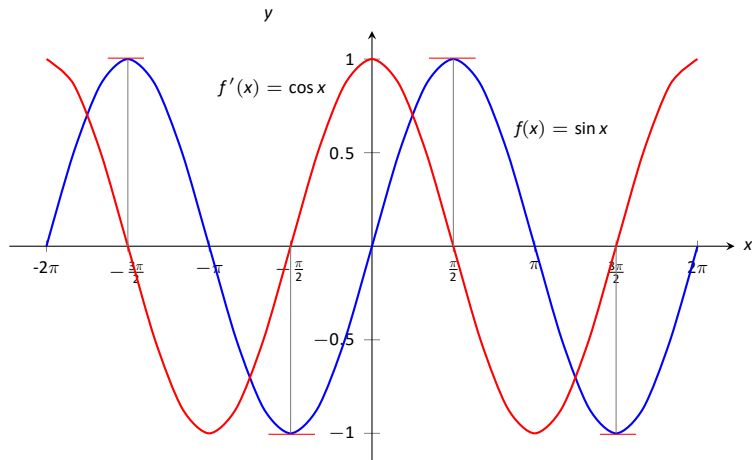
Figure 2.6: The function $f(x) = \sin x$ and its derivative $f'(x) = \cos x$.

Initially, this might be somewhat surprising; the result of a tedious limit process and the sine function is a nice function. Then again, perhaps this is not entirely surprising. The sine function is periodic — it repeats itself on regular intervals. Therefore its rate of change also repeats itself on the same regular intervals. In fact, if we think about $f'(x)$ as the slope of the tangent to the sine curve we notice the following

- when the slope of tangent lines is 0 then $f'(x) = \cos x$ crosses the $x-$axis;

- when the slopes of the tangent lines are positive then f' lies above the $x-$axis; and

- when the slopes of the tangent lines are negative then f' lies below the $x-$axis.

Notes:

We should have known the derivative would be periodic; we now know exactly which periodic function it is.

Thinking back to Example 3, we can find the slope of the tangent line to $f(x) = \sin x$ at $x = 0$ using our derivative. We approximated the slope as 0.9983; we now know the slope is *exactly* $\cos 0 = 1$.

Example 7   **Finding the derivative of a piecewise defined function**
Find the derivative of the absolute value function,

$$f(x) = |x| = \begin{cases} -x & x < 0 \\ x & x \geq 0 \end{cases}.$$

See Figure 2.7.

Solution   We need to evaluate $\lim\limits_{h \to 0} \dfrac{f(x+h) - f(x)}{h}$. As f is piecewise-defined, we need to consider separately the limits when $x < 0$ and when $x > 0$.
When $x < 0$:

$$\begin{aligned} \frac{d}{dx}(-x) &= \lim_{h \to 0} \frac{-(x+h) - (-x)}{h} \\ &= \lim_{h \to 0} \frac{-h}{h} \\ &= \lim_{h \to 0} -1 \\ &= -1. \end{aligned}$$

When $x > 0$, a similar computation shows that $\dfrac{d}{dx}(x) = 1$.

We need to also find the derivative at $x = 0$. By the definition of the derivative at a point, we have

$$f'(0) = \lim_{h \to 0} \frac{f(0+h) - f(0)}{h}.$$

Since $x = 0$ is the point where our function's definition switches from one piece to the other, we need to consider left and right-hand limits. Consider the following, where we compute the left and right hand limits side by side.

$$\begin{aligned} \lim_{h \to 0^-} \frac{f(0+h) - f(0)}{h} &= \\ \lim_{h \to 0^-} \frac{-h - 0}{h} &= \\ \lim_{h \to 0^-} -1 = -1 \end{aligned} \qquad \bigg| \qquad \begin{aligned} \lim_{h \to 0^+} \frac{f(0+h) - f(0)}{h} &= \\ \lim_{h \to 0^+} \frac{h - 0}{h} &= \\ \lim_{h \to 0^+} 1 = 1 \end{aligned}$$

The
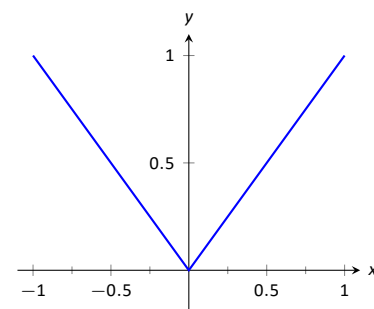
Figure 2.7: The absolute value function, $f(x) = |x|$. Notice how the slope of the lines (and hence the tangent lines) abruptly changes at $x = 0$.
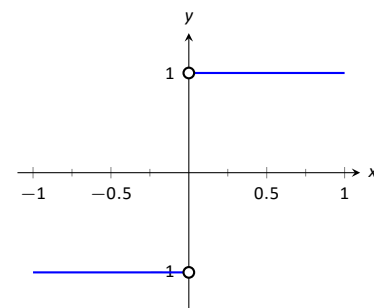
Figure 2.8: A graph of the derivative of $f(x) = |x|$.

Notes:

last lines of each column tell the story: the left and right hand limits are not equal. Therefore the limit does not exist at 0, and f is not differentiable at 0; see Figure 2.8. So we have

$$f'(x) = \begin{cases} -1 & x < 0 \\ 1 & x > 0 \end{cases}.$$

At $x = 0$, $f'(x)$ does not exist; there is a jump discontinuity at 0. So $f(x) = |x|$ is differentiable everywhere except at 0.

The point of non-differentiability came where the piecewise defined function switched from one piece to the other. Our next example shows that this does not always cause trouble.

Example 8 **Finding the derivative of a piecewise defined function**

Find the derivative of $f(x)$, where $f(x) = \begin{cases} \sin x & x \le \pi/2 \\ 1 & x > \pi/2 \end{cases}$. See Figure 2.9.

SOLUTION Using Example 6, we know that when $x < \pi/2$, $f'(x) = \cos x$. It is easy to verify that when $x > \pi/2$, $f'(x) = 0$; consider:

$$\lim_{h \to 0} \frac{f(x+h) - f(x)}{h} = \lim_{h \to 0} \frac{1-1}{h} = \lim_{h \to 0} 0 = 0.$$

So far we have

$$f'(x) = \begin{cases} \cos x & x < \pi/2 \\ 0 & x > \pi/2 \end{cases}.$$

We still need to find $f'(\pi/2)$. Notice at $x = \pi/2$ that both pieces of f' are 0, meaning we can state that $f'(\pi/2) = 0$.

Being more rigorous, we can again evaluate the difference quotient limit at $x = \pi/2$, utilizing again left and right–hand limits:

$$\lim_{h \to 0^-} \frac{f(\pi/2 + h) - f(\pi/2)}{h} =$$
$$\lim_{h \to 0^-} \frac{\sin(\pi/2 + h) - \sin(\pi/2)}{h} =$$
$$\lim_{h \to 0^-} \frac{\sin(\frac{\pi}{2})\cos(h) + \sin(h)\cos(\frac{\pi}{2}) - \sin(\frac{\pi}{2})}{h} =$$
$$\lim_{h \to 0^-} \frac{1 \cdot \cos(h) + \sin(h) \cdot 0 - 1}{h} =$$
$$0$$

$$\lim_{h \to 0^+} \frac{f(\pi/2 + h) - f(\pi/2)}{h} =$$
$$\lim_{h \to 0^+} \frac{1-1}{h} =$$
$$\lim_{h \to 0^+} \frac{0}{h} =$$
$$0$$
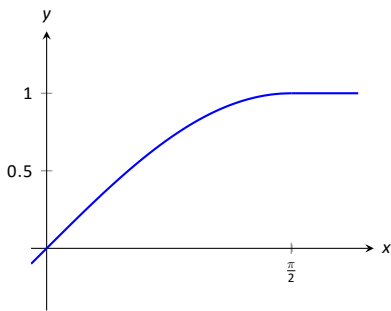
Figure 2.9: A graph of $f(x)$ as defined in Example 8.
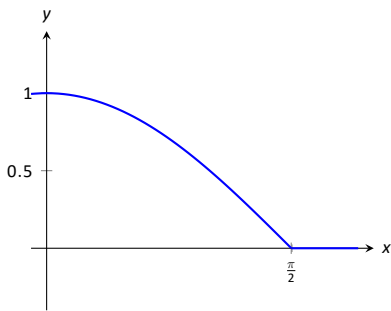
Figure 2.10: A graph of $f'(x)$ in Example 8.

Notes:

Since both the left and right hand limits are 0 at $x = \pi/2$, the limit exists and $f'(\pi/2)$ exists (and is 0). Therefore we can fully write f' as

$$f'(x) = \begin{cases} \cos x & x \leq \pi/2 \\ 0 & x > \pi/2 \end{cases}.$$

See Figure 2.10 for a graph of this function.

Recall we pseudo-defined a continuous function as one in which we could sketch its graph without lifting our pencil. We can give a pseudo–definition for differentiability as well: it is a continuous function that does not have any "sharp corners." One such sharp corner is shown in Figure 2.7. Even though the function f in Example 8 is piecewise–defined, the transition is "smooth" hence it is differentiable. Note how in the graph of f in Figure 2.9 it is difficult to tell when f switches from one piece to the other; there is no "corner."

To better understand the definition of a derivative, experiment with the Geogebra app at `http://mathinsight.org/applet/secant_line_slope`.

This section defined the derivative; in some sense, it answers the question of "What *is* the derivative?" The next section addresses the question "What does the derivative *mean*?"

Notes:

Exercises 2.1

Terms and Concepts

1. T/F: Let f be a position function. The average rate of change on $[a, b]$ is the slope of the line through the points $(a, f(a))$ and $(b, f(b))$.

2. T/F: The definition of the derivative of a function at a point involves taking a limit.

3. In your own words, explain the difference between the average rate of change and instantaneous rate of change.

4. In your own words, explain the difference between Definitions 10 and 12.

5. Let $y = f(x)$. Give three different notations equivalent to "$f'(x)$."

Problems

In Exercises 6–15,

(a) use the definition of the derivative to compute the derivative of the given function.

(b) Find the tangent line to the graph of the given function at $x = c$.

6. $f(x) = 6$ at $x = -2$

7. $f(x) = 2x$ at $x = 3$

8. $f(x) = 4 - 3x$ at $x = 7$

9. $g(x) = x^2$ at $x = -2$

10. $h(x) = 2x - x^2$ at $x = 1$

11. $f(x) = 3x^2 - x + 4$ at $x = -1$

12. $g(x) = \sqrt{x + 3}$ at $x = 1$

13. $r(x) = \dfrac{1}{x}$ at $x = -2$

14. $h(x) = \dfrac{3}{\sqrt{x}}$ at $x = 4$

15. $f(x) = \dfrac{1}{x - 2}$ at $x = 3$

In Exercises 16–19, each limit represents the derivative of some function, f, at some number c. State an appropriate f and c for each.

16. $\lim\limits_{h \to 0} \dfrac{\sqrt{16 + h} - 4}{h}$

17. $\lim\limits_{h \to 0} \dfrac{(3 + h)^4 - 81}{h}$

18. $\lim\limits_{h \to 0} \dfrac{\frac{1}{2+h} - \frac{1}{2}}{h}$

19. $\lim\limits_{h \to 0} \dfrac{\cos(-\pi + h) + 1}{h}$

In Exercises 20–24, a function f and an x–value a are given. Approximate the equation of the tangent line to the graph of f at $x = a$ by numerically approximating $f'(a)$, using $h = 0.1$.

20. $f(x) = x^2 + 2x + 1$, $x = 3$

21. $f(x) = \sqrt{x}$, $x = 4$

22. $f(x) = \dfrac{10}{x + 1}$, $x = 9$

23. $f(x) = e^x$, $x = 2$

24. $f(x) = \cos x$, $x = 0$

25. The graph of $f(x) = x^2 - 1$ is shown.

(a) Use the graph to approximate the slope of the tangent line to f at the following points: $(-1, 0)$, $(0, -1)$ and $(2, 3)$.

(b) Using the definition, find $f'(x)$.

(c) Find the slope of the tangent line at the points $(-1, 0)$, $(0, -1)$ and $(2, 3)$.
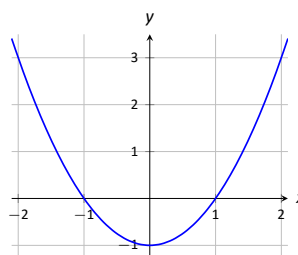
26. The graph of $f(x) = \dfrac{1}{x + 1}$ is shown.

(a) Use the graph to approximate the slope of the tangent line to f at the following points: $(0, 1)$ and $(1, 0.5)$.

(b) Using the definition, find $f'(x)$.

(c) Find the slope of the tangent line at the points $(0, 1)$ and $(1, 0.5)$.
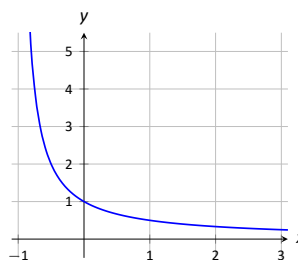
In Exercises 27–30, a graph of a function $f(x)$ is given. Using the graph, sketch $f'(x)$.

27.

28.

29.

30.
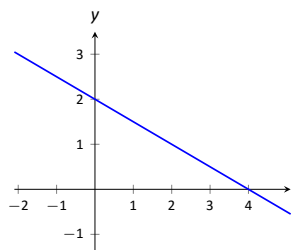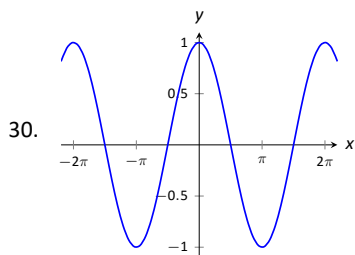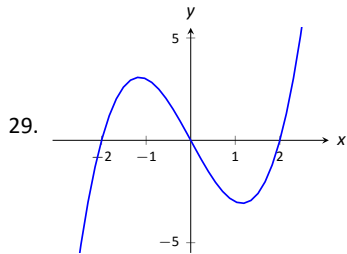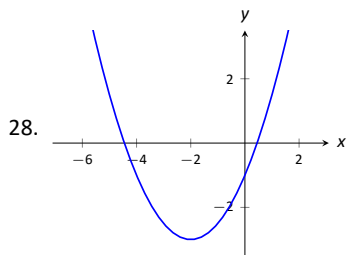
31. Using the graph of $g(x)$ below, answer the following questions.

    (a) Where is $g(x) > 0$?

    (b) Where is $g(x) < 0$?

    (c) Where is $g(x) = 0$?

    (d) Where is $g'(x) < 0$?

    (e) Where is $g'(x) > 0$?

    (f) Where is $g'(x) = 0$?
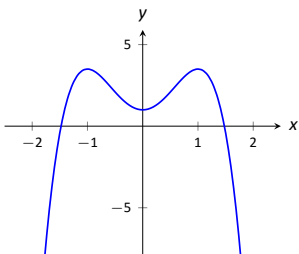
Review

32. Approximate $\lim\limits_{x \to 5} \dfrac{x^2 + 2x - 35}{x^2 - 10.5x + 27.5}$.

33. Use the Bisection Method to approximate, accurate to two decimal places, the root of $g(x) = x^3 + x^2 + x - 1$ on $[0.5, 0.6]$.

34. Give intervals on which each of the following functions are continuous.

    (a) $\dfrac{1}{e^x + 1}$

    (b) $\dfrac{1}{x^2 - 1}$

    (c) $\sqrt{5 - x}$

    (d) $\sqrt{5 - x^2}$

35. Use the graph of $f(x)$ provided to answer the following.

    (a) $\lim\limits_{x \to -3^-} f(x) =$?

    (b) $\lim\limits_{x \to -3+} f(x) =$?

    (c) $\lim\limits_{x \to -3} f(x) =$?

    (d) Where is f continuous?
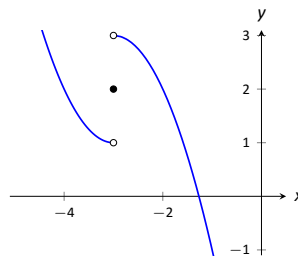
2.2 Interpretations of the Derivative

The previous section defined the derivative of a function and gave examples of how to compute it using its definition (i.e., using limits). The section also started with a brief motivation for this definition, that is, finding the instantaneous velocity of a falling object given its position function. The next section will give us more accessible tools for computing the derivative, tools that are easier to use than repeated use of limits.

This section falls in between the "What is the definition of the derivative?" and "How do I compute the derivative?" sections. Here we are concerned with "What does the derivative mean?", or perhaps, when read with the right emphasis, "What *is* the derivative?" We offer two interconnected interpretations of the derivative, hopefully explaining why we care about it and why it is worthy of study.

Interpretation of the Derivative #1: Instantaneous Rate of Change

The previous section started with an example of using the position of an object (in this case, a falling amusement–park rider) to find the object's velocity. This type of example is often used when introducing the derivative because we tend to readily recognize that velocity is the *instantaneous rate of change of position.* In general, if f is a function of x, then $f'(x)$ measures the instantaneous rate of change of f with respect to x. Put another way, the derivative answers "When x changes, at what rate does f change?" Thinking back to the amusement–park ride, we asked "When time changed, at what rate did the height change?" and found the answer to be "By -64 feet per second."

Now imagine driving a car and looking at the speedometer, which reads "60 mph." Five minutes later, you wonder how far you have traveled. Certainly, lots of things could have happened in those 5 minutes; you could have intentionally sped up significantly, you might have come to a complete stop, you might have slowed to 20 mph as you passed through construction. But suppose that you know, as the driver, none of these things happened. You know you maintained a fairly consistent speed over those 5 minutes. What is a good approximation of the distance traveled?

One could argue the *only* good approximation, given the information provided, would be based on "distance = rate $\times$ time." In this case, we assume a constant rate of 60 mph with a time of $5/60$ hours. Hence we would approximate the distance traveled as 5 miles.

Referring back to the falling amusement–park ride, knowing that at $t = 2$ the velocity was -64 ft/s, we could reasonably assume that 1 second later the riders' height would have dropped by about 64 feet. Knowing that the riders were

---

Notes:

accelerating as they fell would inform us that this is an *under–approximation*. If all we knew was that $f(2) = 86$ and $f'(2) = -64$, we'd know that we'd have to stop the riders quickly otherwise they would hit the ground!

Units of the Derivative

It is useful to recognize the *units* of the derivative function. If y is a function of x, i.e., $y = f(x)$ for some function f, and y is measured in feet and x in seconds, then the units of $y' = f'$ are "feet per second," commonly written as "ft/s." In general, if y is measured in units P and x is measured in units Q, then y' will be measured in units "P per Q", or "P/Q." Here we see the fraction–like behavior of the derivative in the notation:

$$\text{the units of} \quad \frac{dy}{dx} \quad \text{are} \quad \frac{\text{units of } y}{\text{units of } x}.$$

Example 1   **The meaning of the derivative: World Population**
Let $P(t)$ represent the world population t minutes after 12:00 a.m., January 1, 2012. It is fairly accurate to say that $P(0) = 7,028,734,178$ (`www.prb.org`). It is also fairly accurate to state that $P'(0) = 156$; that is, at midnight on January 1, 2012, the population of the world was growing by about 156 *people per minute* (note the units). Twenty days later (or, 28,800 minutes later) we could reasonably assume the population grew by about $(28,800)(156) = 4,492,800$ people.

Example 2   **The meaning of the derivative: Manufacturing**
The term *widget* is an economic term for a generic unit of manufacturing output. Suppose a company produces widgets and knows that the market supports a price of \$10 per widget. Let $P(n)$ give the profit, in dollars, earned by manufacturing and selling n widgets. The company likely cannot make a (positive) profit making just one widget; the start–up costs will likely exceed \$10. Mathematically, we would write this as $P(1) < 0$.

    What do $P(1000) = 500$ and $P'(1000) = 0.25$ mean? Approximate $P(1100)$.

    SOLUTION   The equation $P(1000) = 500$ means that selling 1,000 widgets returns a profit of \$500. We interpret $P'(1000) = 0.25$ as meaning that when producing 1000 widgets, the profit is increasing at rate of \$0.25 per widget (the units are "dollars per widget.") Since we have no other information to use, our best approximation for $P(1100)$ is:

$$P(1100) \approx P(1000) + P'(1000) \times 100 = \$500 + 100 \cdot 0.25 = \$525.$$

We approximate that selling 1,100 widgets returns a profit of \$525.

Notes:

The previous examples made use of an important approximation tool that we first used in our previous "driving a car at 60 mph" example at the beginning of this section. Five minutes after looking at the speedometer, our best approximation for distance traveled assumed the rate of change was constant. In Examples 1 and 2 we made similar approximations. We were given rate of change information which we used to approximate total change. Notationally, we would say that

$$f(c + h) \approx f(c) + f'(c) \cdot h.$$

This approximation is best when h is "small." "Small" is a relative term; when dealing with the world population, $h = 22$ days = 28,800 minutes is small in comparison to years. When manufacturing widgets, 100 widgets is small when one plans to manufacture thousands.

The Derivative and Motion

One of the most fundamental applications of the derivative is the study of motion. Let $s(t)$ be a position function, where t is time and $s(t)$ is distance. For instance, s could measure the height of a projectile or the distance an object has traveled.

Let's let $s(t)$ measure the distance traveled, in feet, of an object after t seconds of travel. Then $s'(t)$ has units "feet per second," and $s'(t)$ measures the *instantaneous rate of distance change* – it measures **velocity**.

Now consider $v(t)$, a velocity function. That is, at time t, $v(t)$ gives the velocity of an object. The derivative of v, $v'(t)$, gives the *instantaneous rate of velocity change* – **acceleration**. (We often think of acceleration in terms of cars: a car may "go from 0 to 60 in 4.8 seconds." This is an *average* acceleration, a measurement of how quickly the velocity changed.) If velocity is measured in feet per second, and time is measured in seconds, then the units of acceleration (i.e., the units of $v'(t)$) are "feet per second per second," or (ft/s)/s. We often shorten this to "feet per second squared," or ft/s^2, but this tends to obscure the meaning of the units.

Perhaps the most well known acceleration is that of gravity. In this text, we use $g = 32$ft/s^2 or $g = 9.8$m/s^2. What do these numbers mean?

A constant acceleration of 32(ft/s)/s means that the velocity changes by 32ft/s each second. For instance, let $v(t)$ measures the velocity of a ball thrown straight up into the air, where v has units ft/s and t is measured in seconds. The ball will have a positive velocity while traveling upwards and a negative velocity while falling down. The acceleration is thus -32ft/s^2. If $v(1) = 20$ft/s, then when $t = 2$, the velocity will have decreased by 32ft/s; that is, $v(2) = -12$ft/s. We can continue: $v(3) = -44$ft/s, and we can also figure that $v(0) = 52$ft/s.

These ideas are so important we write them out as a Key Idea.

Notes:

> **Key Idea 3   The Derivative and Motion**
>
> 1. Let $s(t)$ be the position function of an object. Then $s'(t)$ is the velocity function of the object.
>
> 2. Let $v(t)$ be the velocity function of an object. Then $v'(t)$ is the acceleration function of the object.

We now consider the second interpretation of the derivative given in this section. This interpretation is not independent from the first by any means; many of the same concepts will be stressed, just from a slightly different perspective.

Interpretation of the Derivative #2: The Slope of the Tangent Line

Given a function $y = f(x)$, the difference quotient $\dfrac{f(c + h) - f(c)}{h}$ gives a change in y values divided by a change in x values; i.e., it is a measure of the "rise over run," or "slope," of the line that goes through two points on the graph of f: $\big(c, f(c)\big)$ and $\big(c + h, f(c + h)\big)$. As h shrinks to 0, these two points come close together; in the limit we find $f'(c)$, the slope of a special line called the tangent line.

Lines have a constant rate of change, their slope. Nonlinear functions do not have a constant rate of change, but we can measure their *instantaneous rate of change* at a given x value c by computing $f'(c)$. We can get an idea of how f is behaving by looking at the slopes of its tangent lines. We explore this idea in the following example.

> Watch the video:
> Interpreting slope of a curve exercise at
> `https://youtu.be/CpDfay5NeCg`

Example 3   Understanding the derivative: the rate of change
Consider $f(x) = x^2$ as shown in Figure 2.11 with tangent lines at $x = 1$ and $x = 3$. It is clear that at $x = 3$ the function is growing faster than at $x = 1$, as it is steeper at $x = 3$. How much faster is it growing?

Notes:
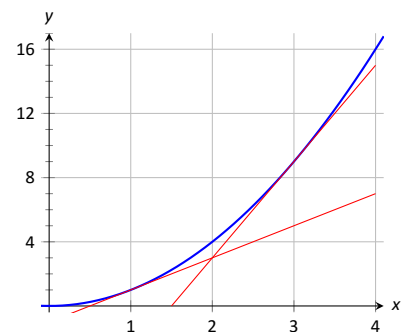
Figure 2.11: A graph of $f(x) = x^2$ and tangent lines.

SOLUTION   We can answer this directly after the following section, where we learn to quickly compute derivatives. For now, we will answer graphically, by considering the slopes of the respective tangent lines.

With practice, one can fairly effectively sketch tangent lines to a curve at a particular point. In Figure 2.11, we have sketched the tangent lines to f at $x = 1$ and $x = 3$, along with a grid to help us measure the slopes of these lines. At $x = 1$, the slope is 2; at $x = 3$, the slope is 6. Thus we can say not only is f growing faster at $x = 3$ than at $x = 1$, it is growing *three times as fast*.

Figure 2.12: Graphs of f and f' in Example 4, along with tangent lines.

Example 4   **Understanding the graph of the derivative**

Consider the graph of $f(x)$ and its derivative, $f'(x)$, in Figure 2.12. Use these graphs to find the slopes of the tangent lines to the graph of f at $x = 1$, $x = 2$, and $x = 3$.

SOLUTION   To find the appropriate slopes of tangent lines to the graph of f, we need to look at the corresponding values of f'.

The slope of the tangent line to f at $x = 1$ is $f'(1)$; this looks to be about -1.

The slope of the tangent line to f at $x = 2$ is $f'(2)$; this looks to be about 4.

The slope of the tangent line to f at $x = 3$ is $f'(3)$; this looks to be about 3.

Using these slopes, the tangent lines to f are sketched in Figure 2.12 as well. Included on the graph of f' in this figure are filled circles where $x = 1$, $x = 2$ and $x = 3$ to help better visualize the y value of f' at those points.

Example 5   **Approximation with the derivative**

Consider again the graph of $f(x)$ and its derivative $f'(x)$ in Example 4. Use the tangent line to f at $x = 3$ to approximate the value of $f(3.1)$.

SOLUTION   Figure 2.13 shows the graph of f along with its tangent line, zoomed in at $x = 3$. Notice that near $x = 3$, the tangent line makes an excellent approximation of f. Since lines are easy to deal with, often it works well to approximate a function with its tangent line. (This is especially true when you don't actually know much about the function at hand, as we don't in this example.)

While the tangent line to f was drawn in Example 4, it was not explicitly computed. Recall that the tangent line to f at $x = c$ is $y = f'(c)(x-c)+f(c)$. While f is not explicitly given, by the graph it looks like $f(3) = 4$. Recalling that $f'(3) = 3$, we can compute the tangent line to be approximately $y = 3(x - 3) + 4$. It is often useful to leave the tangent line in point–slope form.

To use the tangent line to approximate $f(3.1)$, we simply evaluate y at 3.1 instead of f.

$$f(3.1) \approx y(3.1) = 3(3.1 - 3) + 4 = .1 * 3 + 4 = 4.3.$$

We approximate $f(3.1) \approx 4.3$.

Figure 2.13: Zooming in on f at $x = 3$ for the function given in Examples 4 and 5.

Notes:

To demonstrate the accuracy of the tangent line approximation, we now state that in Example 5, $f(x) = -x^3 + 7x^2 - 12x + 4$. We can evaluate $f(3.1) = 4.279$. Had we known f all along, certainly we could have just made this computation. In reality, we often only know two things:

1. What $f(c)$ is, for some value of c, and

2. what $f'(c)$ is.

For instance, we can easily observe the location of an object and its instantaneous velocity at a particular point in time. We do not have a "function f" for the location, just an observation. This is enough to create an approximating function for f.

This last example has a direct connection to our approximation method explained above after Example 2. We stated there that

$$f(c + h) \approx f(c) + f'(c) \cdot h.$$

If we know $f(c)$ and $f'(c)$ for some value $x = c$, then computing the tangent line at $(c, f(c))$ is easy: $y(x) = f'(c)(x - c) + f(c)$. In Example 5, we used the tangent line to approximate a value of f. Let's use the tangent line at $x = c$ to approximate a value of f near $x = c$; i.e., compute $y(c + h)$ to approximate $f(c + h)$, assuming again that h is "small." Note:

$$y(c + h) = f'(c)\big((c + h) - c\big) + f(c) = f'(c) \cdot h + f(c).$$

This is the exact same approximation method used above! Not only does it make intuitive sense, as explained above, it makes analytical sense, as this approximation method is simply using a tangent line to approximate a function's value.

The importance of understanding the derivative cannot be understated. When f is a function of x, $f'(x)$ measures the instantaneous rate of change of f with respect to x and gives the slope of the tangent line to f at x.

Notes:

Exercises 2.2

Terms and Concepts

1. What is the instantaneous rate of change of position called?

2. Given a function $y = f(x)$, in your own words describe how to find the units of $f'(x)$.

3. What functions have a constant rate of change?

Problems

4. Given $f(5) = 10$ and $f'(5) = 2$, approximate $f(6)$.

5. Given $P(100) = -67$ and $P'(100) = 5$, approximate $P(110)$.

6. Given $z(25) = 187$ and $z'(25) = 17$, approximate $z(20)$.

7. Knowing $f(10) = 25$ and $f'(10) = 5$ and the methods described in this section, which approximation is likely to be most accurate: $f(10.1)$, $f(11)$, or $f(20)$? Explain your reasoning.

8. Given $f(7) = 26$ and $f(8) = 22$, approximate $f'(7)$.

9. Given $H(0) = 17$ and $H(2) = 29$, approximate $H'(2)$.

10. Let $V(x)$ measure the volume, in decibels, measured inside a restaurant with x customers. What are the units of $V'(x)$?

11. Let $v(t)$ measure the velocity, in ft/s, of a car moving in a straight line t seconds after starting. What are the units of $v'(t)$?

12. The height H, in feet, of a river is recorded t hours after midnight, April 1. What are the units of $H'(t)$?

13. P is the profit, in thousands of dollars, of building a factory and then producing and selling c cars.

    (a) What are the units of $P'(c)$?

    (b) What is likely true of $P(0)$?

14. T is the temperature in degrees Fahrenheit, h hours after midnight on July 4 in Sidney, NE.

    (a) What are the units of $T'(h)$?

    (b) Is $T'(8)$ likely greater than or less than 0? Why?

    (c) Is $T(8)$ likely greater than or less than 0? Why?

In Exercises 15–18, graphs of functions $f(x)$ and $g(x)$ are given. Identify which function is the derivative of the other.

15.

16.

17.

18.
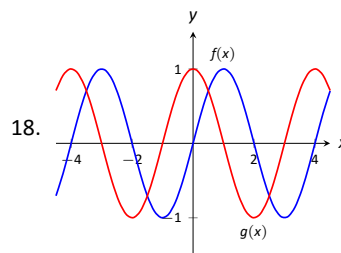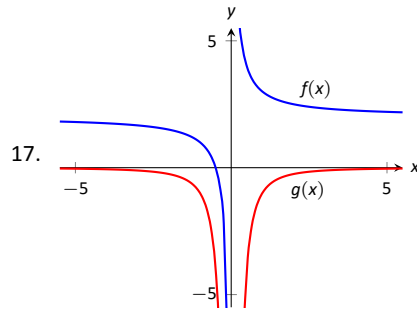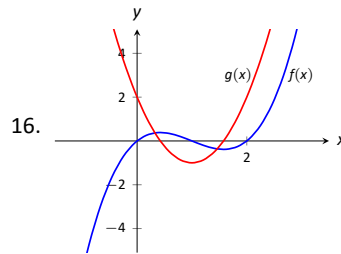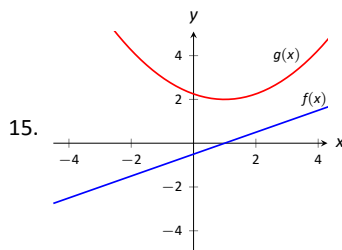
19. If the tangent line to $y = f(x)$ at $(6, 1)$ passes through the point $(2, 4)$, find $f(6)$ and $f'(6)$.

20. Sketch the graph of a function f for which $f(0) = 0, f'(0) > 0, f'(1) = 0$, and $f'(3) < 0$.

21. Sketch the graph of a function h for which $h(1) = 0$, $h'(1) > 0, h'(2) = 0$, and $h'(3) > 0$.

In Exercises 22–24, use the graph of $f(x)$ to sketch $f'(x)$.

22.

23.

24.
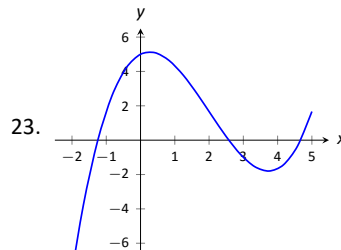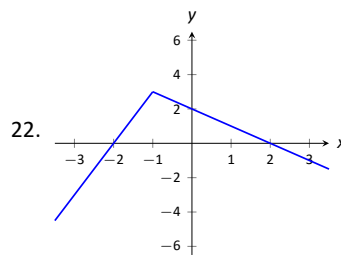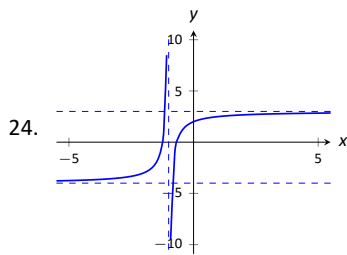
Review

In Exercises 25–26, use the definition to compute the derivatives of the following functions.

25. $f(x) = 5x^2$

26. $f(x) = (x - 2)^3$

In Exercises 27–28, numerically approximate the value of $f'(x)$ at the indicated x value.

27. $f(x) = \cos x$ at $x = \pi$.

28. $f(x) = \sqrt{x}$ at $x = 9$.

2.3   Basic Differentiation Rules

The derivative is a powerful tool but is admittedly awkward given its reliance on limits. Fortunately, one thing mathematicians are good at is *abstraction.* For instance, instead of continually finding derivatives at a point, we abstracted and found the derivative function.

Let's practice abstraction on linear functions, $y = mx + b$. What is y'? Without limits, recognize that linear function are characterized by being functions with a constant rate of change (the slope). The derivative, y', gives the instantaneous rate of change; with a linear function, this is constant, m. Thus $y' = m$.

Let's abstract once more. Let's find the derivative of the general quadratic function, $f(x) = ax^2 + bx + c$. Using the definition of the derivative, we have:

$$f'(x) = \lim_{h \to 0} \frac{a(x+h)^2 + b(x+h) + c - (ax^2 + bx + c)}{h}$$
$$= \lim_{h \to 0} \frac{ah^2 + 2ahx + bh}{h}$$
$$= \lim_{h \to 0} ah + 2ax + b$$
$$= 2ax + b.$$

So if $y = 6x^2 + 11x - 13$, we can immediately compute $y' = 12x + 11$.

In this section (and in some sections to follow) we will learn some of what mathematicians have already discovered about the derivatives of certain functions and how derivatives interact with arithmetic operations. We start with a theorem.

Theorem 14   Derivatives of Common Functions

1. **Constant Rule:** $\dfrac{d}{dx}(c) = 0$,
   where c is a constant.

2. **Power Rule:** $\dfrac{d}{dx}(x^n) = nx^{n-1}$,
   where n is any real number.

3. $\dfrac{d}{dx}(\sin x) = \cos x$

4. $\dfrac{d}{dx}(\cos x) = -\sin x$

5. $\dfrac{d}{dx}(e^x) = e^x$

6. $\dfrac{d}{dx}(\ln x) = \dfrac{1}{x}$

This theorem starts by stating an intuitive fact: constant functions have a rate of change of zero, as they are *constant*. Therefore their derivative is 0. The proof is left as an exercise.

The theorem then states some fairly amazing things.

Notes:

In Part 2, the Power Rule states that the derivatives of functions of the form $y = x^n$ where **n is ANY real number** are very straightforward: multiply by the power, then subtract 1 from the power. This allows us to differentiate Power Functions, Root Functions, and functions with irrational exponents. The work we have done so far only allows us to prove the Power Rule when n is a non-negative integer, which is presented here. We will provide proofs for other values of n as we add the necessary tools to our knowledge of calculus.

Proof of Differentiation Power Rule when n is a non-negative integer

If $n = 0$, then $f(x) = x^0 = 1$ (except when $x = 0$, when the expression is indeterminate). This means that

$$f'(x) = \lim_{h \to 0} \frac{1 - 1}{h} = \lim_{h \to 0} \frac{0}{h} = 0 = 0x^{0-1}$$

as claimed. Now let $f(x) = x^n$, where $n \in \mathbb{Z}^+$. By the definition of derivative,

$$
\begin{aligned}
f'(x) &= \lim_{h \to 0} \frac{(x+h)^n - x^n}{h} \\
&= \lim_{h \to 0} \frac{(x+h)^n - x^n}{h} \qquad \text{use the Binomial Theorem to expand } (x+h)^n \\
&= \lim_{h \to 0} \frac{\binom{n}{0}x^n + \binom{n}{1}hx^{n-1} + \binom{n}{2}h^2x^{n-2} + \cdots + \binom{n}{n-1}h^{n-1}x + \binom{n}{n}h^n - x^n}{h} \\
&= \lim_{h \to 0} \frac{\binom{n}{1}hx^{n-1} + \binom{n}{2}h^2x^{n-2} + \cdots + \binom{n}{n-1}h^{n-1}x + \binom{n}{n}h^n}{h} \\
&= \lim_{h \to 0} \frac{h\left[\binom{n}{1}x^{n-1} + \binom{n}{2}hx^{n-2} + \cdots + \binom{n}{n-1}h^{n-2}x + \binom{n}{n}h^{n-1}\right]}{h}, \text{ divide } h \\
&= \lim_{h \to 0} \binom{n}{1}x^{n-1} + \binom{n}{2}hx^{n-2} + \cdots + \binom{n}{n-1}h^{n-2}x + \binom{n}{n}h^{n-1}, \\
&= nx^{n-1} \qquad \text{since } \binom{n}{1} = n \qquad \qquad \qquad \qquad \square
\end{aligned}
$$

We proved Theorem 14 part 3 in Section 2.1 and part 4 is left as an exercise. In parts 5 and 6 we see something incredible about the functions $y = e^x$ and $y = \ln x$. We will use these rules freely, unfortunately their proofs will have to wait until we know a few more calculus techniques.

Let's practice using this theorem.

Example 1   **Using Theorem 14 to find, and use, derivatives**

Let $f(x) = x^3$.

1. Find $f'(x)$.

Notes:

2. Find the equation of the line tangent to the graph of f at $x = -1$.

3. Use the tangent line to approximate $(-1.1)^3$.

4. Sketch f, f' and the found tangent line on the same axis.

SOLUTION

1. The Power Rule states that if $f(x) = x^3$, then $f'(x) = 3x^2$.

2. To find the equation of the line tangent to the graph of f at $x = -1$, we need a point and the slope. The point is $(-1, f(-1)) = (-1, -1)$. The slope is $f'(-1) = 3$. Thus the tangent line has equation $y = 3(x-(-1)) + (-1) = 3x + 2$.

3. We can use the tangent line to approximate $(-1.1)^3$ as -1.1 is close to -1. We have

$$(-1.1)^3 \approx (-1)^3 + 3(-1.1 - (-1)) = -1 + 3(-.1) = -1.3.$$

We can easily find the actual answer; $(-1.1)^3 = -1.331$.

4. See Figure 2.14.

Figure 2.14: A graph of $f(x) = x^3$, along with its derivative $f'(x) = 3x^2$ and its tangent line at $x = -1$.

Theorem 14 gives useful information, but we will need much more. For instance, using the theorem, we can easily find the derivative of $y = x^3$, but it does not tell how to compute the derivative of $y = 2x^3$, $y = x^3 + \sin x$, nor $y = x^3 \sin x$. The following theorem helps with the first two of these examples (the third is answered in the next section).

Theorem 15   Properties of the Derivative

Let f and g be differentiable on an open interval I and let c be a real number. Then:

1. **Sum/Difference Rule:**
$$\frac{d}{dx}\Big(f(x) \pm g(x)\Big) = \frac{d}{dx}\Big(f(x)\Big) \pm \frac{d}{dx}\Big(g(x)\Big) = f'(x) \pm g'(x)$$

2. **Constant Multiple Rule:**
$$\frac{d}{dx}\Big(c \cdot f(x)\Big) = c \cdot \frac{d}{dx}\Big(f(x)\Big) = c \cdot f'(x).$$

Notes:

Proof of Sum Rule for Differentiation

Let f and g be differentiable on an open interval I and let c be a real number,

$$\frac{d}{dx}(f(x)+g(x)) = \lim_{h\to 0}\frac{[f(x+h)+g(x+h)]-[f(x)+g(x)]}{h}$$
$$= \lim_{h\to 0}\frac{[f(x+h)-f(x)]+[g(x+h)-g(x)]}{h}$$
$$= \lim_{h\to 0}\frac{[f(x+h)-f(x)]}{h}+\lim_{h\to 0}\frac{g(x+h)-g(x)}{h}$$
$$= f'(x)+g'(x) \qquad \square$$

Watch the video:
Basic Derivative Examples at
`https://youtu.be/3dJepii_rJ0`

Theorem 15 allows us to find the derivatives of a wide variety of functions. It can be used in conjunction with the Power Rule to find the derivatives of any polynomial. Recall in Example 2.1.4 that we found, using the limit definition, the derivative of $f(x) = 3x^2 + 5x - 7$. We can now find its derivative without expressly using limits:

$$\frac{d}{dx}\left(3x^2+5x+7\right) = 3\frac{d}{dx}\left(x^2\right)+5\frac{d}{dx}\left(x\right)+\frac{d}{dx}\left(7\right)$$
$$= 3\cdot 2x + 5\cdot 1 + 0$$
$$= 6x + 5.$$

We were a bit pedantic here, showing every step. Normally we would do all the arithmetic and steps in our head and readily find $\frac{d}{dx}\left(3x^2+5x+7\right) = 6x+5$.

Example 2 Using Theorems 14 and 15 to find derivatives

Use Theorems 14 and 15 to differentiate

1. $g(x) = (x^2+1)^3$ 2. $f(x) = \ln\frac{\sqrt{x}}{8}$

Solution Given the differentiation rules we have thus far, our only option for finding $g'(x)$ is to first multiply $g(x)$ out and then apply the sum and power rules. We see that

$$g(x) = x^6 + 3x^4 + 3x^2 + 1$$

Notes:

thus,
$$g'(x) = 6x^5 + 12x^3 + 6x.$$

To differentiate $f(x)$ we will first need to use the Laws of Logarithms to expand f as

$$f(x) = \ln\frac{\sqrt{x}}{8}$$
$$= \ln x^{\frac{1}{2}} - \ln 8$$
$$= \frac{1}{2}\ln x - \ln 8$$

so that,

$$f'(x) = \frac{1}{2}\cdot\frac{1}{x} - 0 = \frac{1}{2}x.$$

Example 3   **Using the tangent line to approximate a function value**
Let $f(x) = \sin x + 2x + 1$. Approximate $f(3)$ using an appropriate tangent line.

SOLUTION   This problem is intentionally ambiguous; we are to *approximate* using an *appropriate* tangent line. How good of an approximation are we seeking? What does appropriate mean?

In the "real world," people solving problems deal with these issues all time. One must make a judgment using whatever seems reasonable. In this example, the actual answer is $f(3) = \sin 3 + 7$, where the real problem spot is $\sin 3$. What is $\sin 3$?

Since 3 is close to π, we can assume $\sin 3 \approx \sin \pi = 0$. Thus one guess is $f(3) \approx 7$. Can we do better? Let's use a tangent line as instructed and examine the results; it seems best to find the tangent line at $x = \pi$.

Using Theorem 14 we find $f'(x) = \cos x + 2$. The slope of the tangent line is thus $f'(\pi) = \cos\pi + 2 = 1$. Also, $f(\pi) = 2\pi + 1 \approx 7.28$. So the tangent line to the graph of f at $x = \pi$ is $y = 1(x-\pi)+2\pi+1 = x+\pi+1 \approx x+4.14$. Evaluated at $x = 3$, our tangent line gives $y = 3 + 4.14 = 7.14$. Using the tangent line, our final approximation is that $f(3) \approx 7.14$.

Using a calculator, we get an answer accurate to 4 places after the decimal: $f(3) = 7.1411$. Our initial guess was 7; our tangent line approximation was more accurate, at 7.14.

The point is *not* "Here's a cool way to do some math without a calculator." Sure, that might be handy sometime, but your phone could probably give you the answer. Rather, the point is to say that tangent lines are a good way of approximating, and many scientists, engineers and mathematicians often face problems too hard to solve directly. So they approximate.

Notes:

Higher Order Derivatives

The derivative of a function f is itself a function, therefore we can take its derivative. The following definition gives a name to this concept and introduces its notation.

Definition 13   Higher Order Derivatives

Let $y = f(x)$ be a differentiable function on I.

1. The *second derivative* of f is:

$$f''(x) = \frac{d}{dx}\left(f'(x)\right) = \frac{d}{dx}\left(\frac{dy}{dx}\right) = \frac{d^2y}{dx^2} = y''.$$

2. The *third derivative* of f is:

$$f'''(x) = \frac{d}{dx}\left(f''(x)\right) = \frac{d}{dx}\left(\frac{d^2y}{dx^2}\right) = \frac{d^3y}{dx^3} = y'''.$$

3. The nth *derivative* of f is:

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

Note: Definition 13 comes with the caveat "Where the corresponding limits exist." With f differentiable on I, it is possible that f' is *not* differentiable on all of I, and so on.

In general, when finding the fourth derivative and on, we resort to the $f^{(4)}(x)$ notation, not $f''''(x)$; after a while, too many ticks is too confusing.

Let's practice using this new concept.

Example 4   Finding higher order derivatives
Find the first four derivatives of the following functions:

$$1.\ f(x) = 4x^2 \qquad 2.\ f(x) = \sin x \qquad 3.\ f(x) = 5e^x$$

SOLUTION

1. Using the Power and Constant Multiple Rules, we have: $f'(x) = 8x$. Continuing on, we have

$$f''(x) = \frac{d}{dx}\left(8x\right) = 8; \qquad f'''(x) = 0; \qquad f^{(4)}(x) = 0.$$

Notice how all successive derivatives will also be 0.

Notes:

2. We employ Theorem 14 repeatedly.

$$f'(x) = \cos x; \quad f''(x) = -\sin x; \quad f'''(x) = -\cos x; \quad f^{(4)}(x) = \sin x.$$

Note how we have come right back to $f(x)$ again. (Can you quickly figure what $f^{(23)}(x)$ is?)

3. Employing Theorem 14 and the Constant Multiple Rule, we can see that

$$f'(x) = f''(x) = f'''(x) = f^{(4)}(x) = 5e^x.$$

Interpreting Higher Order Derivatives

What do higher order derivatives *mean*? What is the practical interpretation?

Our first answer is a bit wordy, but is technically correct and beneficial to understand. That is,

> The second derivative of a function f is the rate of change of the rate of change of f.

One way to grasp this concept is to let f describe a position function. Then, as stated in Key Idea 3, f' describes the rate of position change: velocity. We now consider f'', which describes the rate of velocity change. Sports car enthusiasts talk of how fast a car can go from 0 to 60 mph; they are bragging about the *acceleration* of the car.

We started this chapter with amusement–park riders free–falling with position function $f(t) = -16t^2 + 150$. It is easy to compute $f'(t) = -32t$ ft/s and $f''(t) = -32$ (ft/s)/s. We may recognize this latter constant; it is the acceleration due to gravity. In keeping with the unit notation introduced in the previous section, we say the units are "feet per second per second." This is usually shortened to "feet per second squared," written as "ft/s^2."

It can be difficult to consider the meaning of the third, and higher order, derivatives. The third derivative is "the rate of change of the rate of change of the rate of change of f." That is essentially meaningless to the uninitiated. In the context of our position/velocity/acceleration example, the third derivative is the "rate of change of acceleration," commonly referred to as "jerk."

Make no mistake: higher order derivatives have great importance even if their practical interpretations are hard (or "impossible") to understand. The mathematical topic of *series* makes extensive use of higher order derivatives.

Notes:

Exercises 2.3

Terms and Concepts

1. What is the name of the rule which states that $\frac{d}{dx}(x^n) = nx^{n-1}$, where $n > 0$ is an integer?

2. What is $\frac{d}{dx}(\ln x)$?

3. Give an example of a function $f(x)$ where $f'(x) = f(x)$.

4. Give an example of a function $f(x)$ where $f'(x) = 0$.

5. The derivative rules introduced in this section explain how to compute the derivative of which of the following functions?
   - $f(x) = \frac{3}{x^2}$
   - $g(x) = 3x^2 - x + 17$
   - $h(x) = 5\ln x$
   - $j(x) = \sin x \cos x$
   - $k(x) = e^{x^2}$
   - $m(x) = \sqrt{x}$

6. Explain in your own words how to find the third derivative of a function $f(x)$.

7. Give an example of a function where $f'(x) \neq 0$ and $f''(x) = 0$.

8. Explain in your own words what the second derivative "means."

9. If $f(x)$ describes a position function, then $f'(x)$ describes what kind of function? What kind of function is $f''(x)$?

10. Let $f(x)$ be a function measured in pounds, where x is measured in feet. What are the units of $f''(x)$?

Problems

In Exercises 11–28, compute the derivative of the given function.

11. $f(x) = 7x^2 - 5x + 7$

12. $g(x) = 14x^3 + 7x^2 + 11x - 29$

13. $m(t) = 9t^5 - \frac{1}{8}t^3 + 3t - 8$

14. $f(\theta) = 9\sin\theta + 10\cos\theta$

15. $f(r) = 6e^r$

16. $g(t) = 10t^4 - \cos t + 7\sin t$

17. $f(x) = 2\ln x - x$

18. $p(s) = \frac{1}{4}s^4 + \frac{1}{3}s^3 + \frac{1}{2}s^2 + s + 1$

19. $h(t) = e^t - \sin t - \cos t$

20. $f(x) = \ln(5x^2)$

21. $f(t) = \ln(17) + e^2 + \sin\pi/2$

22. $g(t) = (1 + 3t)^2$

23. $g(x) = (2x - 5)^3$

24. $f(x) = (1 - x)^3$

25. $f(x) = (2 - 3x)^2$

26. $h(x) = \dfrac{x^5 - 2x^3 + x^2}{x^2}$

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

28. $g(\theta) = \dfrac{1 - \sin^2\theta}{\cos\theta}$

29. A property of logarithms is that $\log_a x = \dfrac{\log_b x}{\log_b a}$, for all bases $a, b > 0$ and $a, b \neq 1$.
   (a) Rewrite this identity when $b = e$, i.e., using $\log_e x = \ln x$.
   (b) Use part (a) to find the derivative of $y = \log_a x$.
   (c) Give the derivative of $y = \log_{10} x$.

30. Prove the Constant Rule: $\frac{d}{dx}(c) = 0$, where c is constant.

31. Find the two values of n so that the function $y = x^n$ satisfies the differential equation $x^2 y'' + 2xy' - 6y = 0$.

32. The figure shows the graphs of f, f', and f''. Identify each curve and explain your choices.
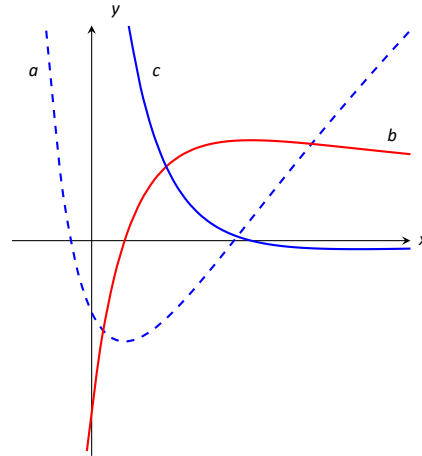
33. The figure shows the graphs of f, f', f'' and f'''. Identify each curve and explain your choices.
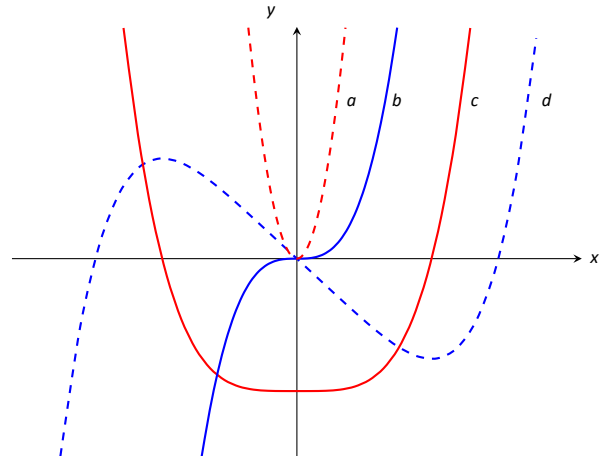
In Exercises 34–39, compute the first four derivatives of the given function.

34. $f(x) = x^6$

35. $g(x) = 2\cos x$

36. $h(t) = t^2 - e^t$

37. $p(\theta) = \theta^4 - \theta^3$

38. $f(\theta) = \sin\theta - \cos\theta$

39. $f(x) = 1,100$

40. The position of a object is described by $s(t) = t^4 - 4t^2$, $t \geq 0$, where s is in feet and t is in seconds. Find

    (a) the velocity and acceleration functions for the object,

    (b) the acceleration after 1.5 seconds, and

    (c) the time(s), in seconds, when the object is at rest.

41. The position of a object is described by $s(t) = 5e^t - 5t$, where s is in inches and t is in seconds. Find

    (a) the velocity and acceleration functions for the object,

    (b) the acceleration after 2 seconds, and

    (c) the acceleration when the object is at rest.

In Exercises 42–47, find the equations of the tangent line to the graph of the function at the given point.

42. $f(x) = x^3 - x$ at $x = 1$

43. $f(t) = e^t + 3$ at $t = 0$

44. $g(x) = \ln x$ at $x = 1$

45. $f(x) = 4\sin x$ at $x = \pi/2$

46. $f(x) = -2\cos x$ at $x = \pi/4$

47. $f(x) = 2x + 3$ at $x = 5$

Review

48. Given that $e^0 = 1$, approximate the value of $e^{0.1}$ using the tangent line to $f(x) = e^x$ at $x = 0$.

49. Approximate the value of $(3.01)^4$ using the tangent line to $f(x) = x^4$ at $x = 3$.

2.4   The Product and Quotient Rules

The previous section showed that, in some ways, derivatives behave nicely. The Constant Multiple and Sum/Difference Rules established that the derivative of $f(x) = 5x^2 + \sin x$ was not complicated. We neglected computing the derivative of things like $g(x) = 5x^2 \sin x$ and $h(x) = \frac{5x^2}{\sin x}$ on purpose; their derivatives are *not* as straightforward. (If you had to guess what their respective derivatives are, you would probably guess wrong.) For these, we need the Product and Quotient Rules, respectively, which are defined in this section.

  We begin with the Product Rule.

Theorem 16   Product Rule

Let f and g be differentiable functions on an open interval I. Then $f \cdot g$ is a differentiable function on I, and

$$\frac{d}{dx}\Big(f(x)g(x)\Big) = f(x)g'(x) + f'(x)g(x).$$

Important: $\frac{d}{dx}\Big(f(x)g(x)\Big) \neq f'(x)g'(x)$. While this answer is simpler than the Product Rule, it is wrong. We can show that this is wrong by considering $f(x) = x^2$ and $g(x) = x^5$.

Using the WRONG rule we get $\frac{d}{dx}[f(x)g(x)] = 2x \cdot 5x^4 = 10x^5$. However, when we simplify the product first and apply the Power Rule, $f \cdot g = x^2 \cdot x^5 = x^7$ and

$$\frac{d}{dx}[f(x)g(x)] = 7x^6 \neq 10x^5.$$

Applying the **real** Product Rule we see that,

$$\begin{aligned}
\frac{d}{dx}[f(x)g(x)] &= x^2 \frac{d}{dx}(x^5) + \frac{d}{dx}(x^2) \cdot x^5 \\
&= x^2 \cdot 5x^4 + 2x \cdot x^5 \\
&= 7x^6
\end{aligned}$$

Watch the video:
The Product Rule for Derivatives at
`https://youtu.be/uPCjqfTOIxg`

Notes:

We practice using this new rule in an example, followed by a proof of the theorem.

Example 1   Using the Product Rule
Use the Product Rule to compute the derivative of $y = 5x^2 \sin x$. Evaluate the derivative at $x = \pi/2$.

SOLUTION   To make our use of the Product Rule explicit, let's set $f(x) = 5x^2$ and $g(x) = \sin x$. We easily compute/recall that $f'(x) = 10x$ and $g'(x) = \cos x$. Employing the rule, we have

$$\frac{d}{dx}\left(5x^2 \sin x\right) = 5x^2 \cos x + 10x \sin x.$$

At $x = \pi/2$, we have

$$y'(\pi/2) = 5\left(\frac{\pi}{2}\right)^2 \cos\left(\frac{\pi}{2}\right) + 10\frac{\pi}{2}\sin\left(\frac{\pi}{2}\right) = 5\pi.$$

We graph y and its tangent line at $x = \pi/2$, which has a slope of 5π, in Figure 2.15. While this does not *prove* that the Product Rule is the correct way to handle derivatives of products, it helps validate its truth.
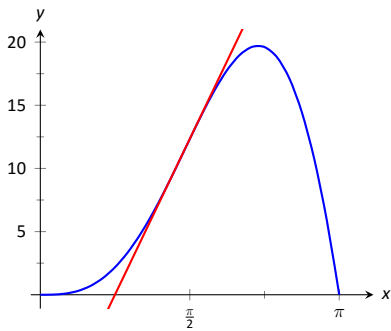
Figure 2.15: A graph of $y = 5x^2 \sin x$ and its tangent line at $x = \pi/2$.

Proof of the Product Rule
By the limit definition, we have

$$\frac{d}{dx}\left(f(x)g(x)\right) = \lim_{h\to 0}\frac{f(x+h)g(x+h) - f(x)g(x)}{h}.$$

We now do something a bit unexpected; add 0 to the numerator (so that nothing is changed) in the form of $-f(x+h)g(x) + f(x+h)g(x)$, and then do some regrouping as shown.

Notes:

$$\frac{d}{dx}\Big(f(x)g(x)\Big) = \lim_{h\to 0}\frac{f(x+h)g(x+h) - f(x)g(x)}{h} \quad \text{(now add 0 to the numerator)}$$

$$= \lim_{h\to 0}\frac{f(x+h)g(x+h) - f(x+h)g(x) + f(x+h)g(x) - f(x)g(x)}{h}$$

$$= \lim_{h\to 0}\frac{\Big(f(x+h)g(x+h) - f(x+h)g(x)\Big) + \Big(f(x+h)g(x) - f(x)g(x)\Big)}{h}$$

$$= \lim_{h\to 0}\frac{f(x+h)g(x+h) - f(x+h)g(x)}{h} + \lim_{h\to 0}\frac{f(x+h)g(x) - f(x)g(x)}{h}$$

$$= \lim_{h\to 0}f(x+h)\frac{g(x+h) - g(x)}{h} + \lim_{h\to 0}\frac{f(x+h) - f(x)}{h}g(x)$$

$$= \lim_{h\to 0}f(x+h)\lim_{h\to 0}\frac{g(x+h) - g(x)}{h} + \lim_{h\to 0}\frac{f(x+h) - f(x)}{h}\lim_{h\to 0}g(x)$$

$$= f(x)g'(x) + f'(x)g(x). \qquad\qquad \square$$

It is often true that we can recognize that a theorem is true through its proof yet somehow doubt its applicability to real problems. In the following example, we compute the derivative of a product of functions in two ways to verify that the Product Rule is indeed "right."

Example 2   Exploring alternate derivative methods
Let $y = (x^2 + 3x + 1)(2x^2 - 3x + 1)$. Find y' two ways: first, by expanding the given product and then taking the derivative, and second, by applying the Product Rule. Verify that both methods give the same answer.

SOLUTION   We first expand the expression for y; a little algebra shows that $y = 2x^4 + 3x^3 - 6x^2 + 1$. It is easy to compute y';

$$y' = 8x^3 + 9x^2 - 12x.$$

Now apply the Product Rule.

$$y' = (x^2 + 3x + 1)\cdot\frac{d}{dx}(2x^2 - 3x + 1) + \frac{d}{dx}(x^2 + 3x + 1)\cdot(2x^2 - 3x + 1)$$

$$= (x^2 + 3x + 1)(4x - 3) + (2x + 3)(2x^2 - 3x + 1)$$

$$= \left(4x^3 + 9x^2 - 5x - 3\right) + \left(4x^3 - 7x + 3\right)$$

$$= 8x^3 + 9x^2 - 12x.$$

The uninformed usually assume that "the derivative of the product is the product of the derivatives." Thus we are tempted to say that $y' = (2x+3)(4x-3) = 8x^2 + 6x - 9$. Obviously this is not correct.

Notes:

Example 3   **Using the Product Rule with a product of three functions**

Let $y = x^3 \ln x \cos x$. Find y'.

SOLUTION   We have a product of three functions while the Product Rule only specifies how to handle a product of two functions. Our method of handling this problem is to simply group the latter two functions together, and consider $y = x^3 \left(\ln x \cos x \right)$. Following the Product Rule, we have

$$y' = (x^3)\left(\ln x \cos x \right)' + 3x^2 \left(\ln x \cos x \right)$$

To evaluate $\left(\ln x \cos x \right)'$, we apply the Product Rule again:

$$= (x^3)\left(\ln x(- \sin x) + \frac{1}{x} \cos x \right) + 3x^2 \left(\ln x \cos x \right)$$

$$= x^3 \ln x(- \sin x) + x^3 \frac{1}{x} \cos x + 3x^2 \ln x \cos x$$

Recognize the pattern in our answer above: when applying the Product Rule to a product of three functions, there are three terms added together in the final derivative. Each term contains only one derivative of one of the original functions, and each function's derivative shows up in only one term. It is straightforward to extend this pattern to finding the derivative of a product of 4 or more functions.

We consider one more example before discussing another derivative rule.

Example 4   **Using the Product Rule**

Find the derivatives of the following functions.

$$1. \, f(x) = x \ln x \qquad 2. \, g(x) = x \ln x - x.$$

SOLUTION   Recalling that the derivative of $\ln x$ is $1/x$, we use the Product Rule to find our answers.

1. $\dfrac{d}{dx}\left(x \ln x \right) = x \cdot 1/x + 1 \cdot \ln x = 1 + \ln x.$

2. Using the result from above, we compute

$$\frac{d}{dx}\left(x \ln x - x \right) = 1 + \ln x - 1 = \ln x.$$

This seems significant; if the natural log function $\ln x$ is an important function (it is), it seems worthwhile to know a function whose derivative is $\ln x$. We have found one. (We leave it to the reader to find others; a correct answer will be *very* similar to this one.)

Notes:

We have learned how to compute the derivatives of sums, differences, and products of functions. We now learn how to find the derivative of a quotient of functions.

Theorem 17   Quotient Rule

Let f and g be functions defined on an open interval I, where $g(x) \neq 0$ on I. Then f/g is differentiable on I, and

$$\frac{d}{dx}\left(\frac{f(x)}{g(x)}\right) = \frac{g(x)f'(x) - f(x)g'(x)}{[g(x)]^2}.$$

Proof of the Quotient Rule

Let the functions f and g be defined and $g(x) \neq 0$ on an open interval I. By the definition of derivative,

$$\frac{d}{dx}\left(\frac{f(x)}{g(x)}\right) = \lim_{h \to 0} \frac{\frac{f(x+h)}{g(x+h)} - \frac{f(x)}{g(x)}}{h}$$

$$= \lim_{h \to 0}\left[\left(\frac{f(x+h)}{g(x+h)} - \frac{f(x)}{g(x)}\right) \cdot \frac{1}{h}\right]$$

$$= \lim_{h \to 0}\left[\left(\frac{f(x+h)g(x) - f(x)g(x+h)}{g(x+h)g(x)}\right) \cdot \frac{1}{h}\right]$$

Adding and subtracting the term $f(x)g(x)$ in the numerator does not change the value of the expression and allows us to separate f and g so that

---

Notes:

$$\frac{d}{dx}\left(\frac{f(x)}{g(x)}\right) = \lim_{h\to 0}\left[\left(\frac{f(x+h)g(x)-f(x)g(x)+f(x)g(x)-f(x)g(x+h)}{g(x+h)g(x)}\right)\cdot\frac{1}{h}\right]$$

$$= \lim_{h\to 0}\left[\frac{f(x+h)g(x)-f(x)g(x)}{hg(x+h)g(x)}+\frac{f(x)g(x)-f(x)g(x+h)}{hg(x+h)g(x)}\right]$$

$$= \lim_{h\to 0}\left[g(x)\frac{f(x+h)-f(x)}{hg(x+h)g(x)}+f(x)\frac{g(x)-g(x+h)}{hg(x+h)g(x)}\right]$$

$$= \lim_{h\to 0}\frac{g(x)\frac{f(x+h)-f(x)}{h}-f(x)\frac{g(x+h)-g(x)}{h}}{g(x+h)g(x)}$$

$$= \frac{\lim\limits_{h\to 0}g(x)\cdot\lim\limits_{h\to 0}\frac{f(x+h)-f(x)}{h}-\lim\limits_{h\to 0}f(x)\cdot\lim\limits_{h\to 0}\frac{g(x+h)-g(x)}{h}}{\lim\limits_{h\to 0}g(x+h)\cdot\lim\limits_{h\to 0}g(x)}$$

$$= \frac{g(x)f'(x)-f(x)g'(x)}{[g(x)]^2} \qquad \qquad \square$$

Let's practice using the Quotient Rule.

Example 5   **Using the Quotient Rule**

Let $f(x) = \dfrac{5x^2}{\sin x}$. Find $f'(x)$.

  SOLUTION   Directly applying the Quotient Rule gives:

$$\frac{d}{dx}\left(\frac{5x^2}{\sin x}\right) = \frac{\sin x\frac{d}{dx}(5x^2)-5x^2\frac{d}{dx}(\sin x)}{(\sin x)^2}$$

$$= \frac{\sin x\cdot 10x-5x^2\cdot\cos x}{\sin^2 x}$$

$$= \frac{10x\sin x-5x^2\cos x}{\sin^2 x}.$$

The Quotient Rule allows us to fill in holes in our understanding of derivatives of the common trigonometric functions. We start with finding the derivative of the tangent function.

Example 6   **Using the Quotient Rule to find $\frac{d}{dx}\left(\tan x\right)$.**

Find the derivative of $y = \tan x$.

Notes:

SOLUTION   At first, one might feel unequipped to answer this question. But recall that $\tan x = \sin x / \cos x$, so we can apply the Quotient Rule.

$$\frac{d}{dx}\left(\tan x\right) = \frac{\cos x \frac{d}{dx}(\sin x) - \sin x \frac{d}{dx}(\cos x)}{(\cos x)^2}$$

$$= \frac{d}{dx}\left(\frac{\sin x}{\cos x}\right)$$

$$= \frac{\cos x \cos x - \sin x(-\sin x)}{\cos^2 x}$$

$$= \frac{\cos^2 x + \sin^2 x}{\cos^2 x}$$

$$= \frac{1}{\cos^2 x}$$

$$= \sec^2 x.$$

This is a beautiful result. To confirm its truth, we can find the equation of the tangent line to $y = \tan x$ at $x = \pi/4$. The slope is $\sec^2(\pi/4) = 2$; $y = \tan x$, along with its tangent line, is graphed in Figure 2.16.

We include this result in the following theorem about the derivatives of the trigonometric functions. Recall we found the derivative of $y = \sin x$ in Example 2.1.6 and stated the derivative of the cosine function in Theorem 14. The derivatives of the cotangent, cosecant and secant functions can all be computed directly using Theorem 14 and the Quotient Rule.

Figure 2.16: A graph of $y = \tan x$ along with its tangent line at $x = \pi/4$.

Theorem 18   Derivatives of Trigonometric Functions

1. $\dfrac{d}{dx}\left(\sin x\right) = \cos x$ 2. $\dfrac{d}{dx}\left(\cos x\right) = -\sin x$

3. $\dfrac{d}{dx}\left(\tan x\right) = \sec^2 x$ 4. $\dfrac{d}{dx}\left(\cot x\right) = -\csc^2 x$

5. $\dfrac{d}{dx}\left(\sec x\right) = \sec x \tan x$ 6. $\dfrac{d}{dx}\left(\csc x\right) = -\csc x \cot x$

The proofs of these derivatives have been presented or left as exercises. To remember the above, it may be helpful to keep in mind that the derivatives of the trigonometric functions that start with "c" have a minus sign in them.

Notes:

Example 7   **Exploring alternate derivative methods**

In Example 5 the derivative of $f(x) = \dfrac{5x^2}{\sin x}$ was found using the Quotient Rule. Rewriting f as $f(x) = 5x^2 \csc x$, find f' using Theorem 18 and verify the two answers are the same.

SOLUTION   We found in Example 5 that $f'(x) = \dfrac{10x \sin x - 5x^2 \cos x}{\sin^2 x}$.
We now find f' using the Product Rule, considering f as $f(x) = 5x^2 \csc x$.

$$
\begin{aligned}
f'(x) &= \frac{d}{dx}\left(5x^2 \csc x\right) \\
&= 5x^2 \frac{d}{dx}(\csc x) - \csc x \frac{d}{dx}(5x^2) \\
&= 5x^2(-\csc x \cot x) + 10x \csc x \qquad \text{(now rewrite trig functions)} \\
&= 5x^2 \cdot \frac{-1}{\sin x} \cdot \frac{\cos x}{\sin x} + \frac{10x}{\sin x} \\
&= \frac{-5x^2 \cos x}{\sin^2 x} + \frac{10x}{\sin x} \qquad \text{(get common denominator)} \\
&= \frac{10x \sin x - 5x^2 \cos x}{\sin^2 x}
\end{aligned}
$$

Finding f' using either method returned the same result. At first, the answers looked different, but some algebra verified they are the same. In general, there is not one final form that we seek; the immediate result from the Product Rule is fine. Work to "simplify" your results into a form that is most readable and useful to you.

When we stated the Power Rule in Section 2.3 we claimed that it worked for all $n \in \mathbb{R}$ but only provided the proof for non-negative integers. The next example uses the Quotient Rule to provide justification of the Power Rule for $n \in \mathbb{Z}$.

Example 8   **Using the Quotient Rule to expand the Power Rule**

Find the derivatives of the following functions.

1. $f(x) = \dfrac{1}{x}$

2. $f(x) = \dfrac{1}{x^n}$, where $n > 0$ is an integer.

SOLUTION   We employ the Quotient Rule.

1. $f'(x) = \dfrac{x \cdot 0 - 1 \cdot 1}{x^2} = -\dfrac{1}{x^2}.$

Notes:

2. $f'(x) = \dfrac{x^n \cdot 0 - 1 \cdot nx^{n-1}}{(x^n)^2} = -\dfrac{nx^{n-1}}{x^{2n}} = -\dfrac{n}{x^{n+1}}.$

The derivative of $y = \dfrac{1}{x^n}$ turned out to be rather nice. It gets better. Consider:

$$\frac{d}{dx}\left(\frac{1}{x^n}\right) = \frac{d}{dx}\left(x^{-n}\right) \qquad \text{(apply result from Example 8)}$$

$$= -\frac{n}{x^{n+1}} \qquad \text{(rewrite algebraically)}$$

$$= -nx^{-(n+1)}$$

$$= -nx^{-n-1}.$$

Thus, for all $n \in \mathbb{Z}$, we can officially apply the Power Rule: multiply by the power, then subtract 1 from the power.

Taking the derivative of many functions is relatively straightforward. It is clear (with practice) what rules apply and in what order they should be applied. Other functions present multiple paths; different rules may be applied depending on how the function is treated. One of the beautiful things about calculus is that there is not "the" right way; each path, when applied correctly, leads to the same result, the derivative. We demonstrate this concept in an example.

Example 9   **Exploring alternate derivative methods**

Let $f(x) = \dfrac{x^2 - 3x + 1}{x}$. Find $f'(x)$ in each of the following ways:

1. By applying the Quotient Rule,

2. by viewing f as $f(x) = \left(x^2 - 3x + 1\right) \cdot x^{-1}$ and applying the Product and Power Rules, and

3. by "simplifying" first through division.

Verify that all three methods give the same result.

    SOLUTION

1. Applying the Quotient Rule gives:

$$f'(x) = \frac{x \cdot \left(2x - 3\right) - \left(x^2 - 3x + 1\right) \cdot 1}{x^2} = \frac{x^2 - 1}{x^2} = 1 - \frac{1}{x^2}.$$

Notes:

2. By rewriting f, we can apply the Product and Power Rules as follows:

$$f'(x) = \left(x^2 - 3x + 1\right) \cdot (-1)x^{-2} + \left(2x - 3\right) \cdot x^{-1}$$
$$= -\frac{x^2 - 3x + 1}{x^2} + \frac{2x - 3}{x}$$
$$= -\frac{x^2 - 3x + 1}{x^2} + \frac{2x^2 - 3x}{x^2}$$
$$= \frac{x^2 - 1}{x^2} = 1 - \frac{1}{x^2},$$

the same result as above.

3. As $x \neq 0$, we can divide through by x first, giving $f(x) = x - 3 + \dfrac{1}{x}$. Now apply the Power Rule to see

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

the same result as before.

Example 9 demonstrates three methods of finding f'. One is hard pressed to argue for a "best method" as all three gave the same result without too much difficulty, although it is clear that using the Product Rule required more steps. Ultimately, the important principle to take away from this is: simplify the answer to a form that seems "simple" and easy to interpret. They are equal; they are all correct. The most appropriate form of f' depends on what we need to do with the function next. For later problems it will be important for us to determine the most appropriate form to use and to move flexibly between the different forms.

In the next section we continue to learn rules that allow us to more easily compute derivatives than using the limit definition directly. We have to memorize the derivatives of a certain set of functions, such as "the derivative of $\sin x$ is $\cos x$." The Sum/Difference, Constant Multiple, Power, Product and Quotient Rules show us how to find the derivatives of certain combinations of these functions. The next section shows how to find the derivatives when we *compose* these functions together.

Notes:

Exercises 2.4

Terms and Concepts

1. T/F: The Product Rule states that $\dfrac{d}{dx}(x^2 \sin x) = 2x \cos x$.

2. T/F: The Quotient Rule states that $\dfrac{d}{dx}\left(\dfrac{x^2}{\sin x}\right) = \dfrac{2x}{\cos x}$.

3. T/F: The derivatives of the trigonometric functions that start with "c" have minus signs in them.

4. What derivative rule is used to extend the Power Rule to include negative integer exponents?

5. T/F: Regardless of the function, there is always exactly one right way of computing its derivative.

6. In your own words, explain what it means to make your answers "clear."

Problems

In Exercises 7–9, use the Quotient Rule to verify these derivatives.

7. $\dfrac{d}{dx}(\cot x) = -\csc^2 x$

8. $\dfrac{d}{dx}(\sec x) = \sec x \tan x$

9. $\dfrac{d}{dx}(\csc x) = -\csc x \cot x$

In Exercises 10–13:

   (a) Use the Product Rule to differentiate the function.

   (b) Manipulate the function algebraically and differentiate without the Product Rule.

   (c) Show that the answers from (a) and (b) are equivalent.

10. $f(x) = x(x^2 + 3x)$

11. $g(x) = 2x^2(5x^3)$

12. $h(s) = (2s - 1)(s + 4)$

13. $f(x) = (x^2 + 5)(3 - x^3)$

In Exercises 14–17:

   (a) Use the Quotient Rule to differentiate the function.

   (b) Manipulate the function algebraically and differentiate without the Quotient Rule.

   (c) Show that the answers from (a) and (b) are equivalent.

14. $f(x) = \dfrac{x^2 + 3}{x}$

15. $g(x) = \dfrac{x^3 - 2x^2}{2x^2}$

16. $h(s) = \dfrac{3}{4s^3}$

17. $f(t) = \dfrac{t^2 - 1}{t + 1}$

In Exercises 18–42, compute the derivative of the given function.

18. $f(x) = x \sin x$

19. $f(t) = \dfrac{1}{t^2}(\csc t - 4)$

20. $H(y) = (y^5 - 2y^3)(7y^2 + y - 8)$

21. $F(y) = \sqrt[3]{y^2}(y^2 + 9y)$

22. $g(x) = \dfrac{x + 7}{x - 5}$

23. $y = \dfrac{\sqrt{x}}{x + 4}$

24. $g(x) = \dfrac{x}{\sqrt{x} + 4}$

25. $g(t) = \dfrac{t^5}{\cos t - 2t^2}$

26. $h(x) = \cot x - e^x$

27. $h(t) = 7t^2 + 6t - 2$

28. $f(x) = \dfrac{x^4 + 2x^3}{x + 2}$

29. $f(x) = \dfrac{x^2 - \sqrt{x}}{x^3}$

30. $y = \left(\dfrac{1}{x^3} + \dfrac{5}{x^4}\right)(2x^3 - x^5)$

31. $g(x) = \dfrac{1}{1 + x + x^2 + x^3}$

32. $p(x) = 1 + \dfrac{1}{x} + \dfrac{1}{x^2} + \dfrac{1}{x^3}$

33. $f(x) = (16x^3 + 24x^2 + 3x)\dfrac{7x - 1}{16x^3 + 24x^2 + 3x}$

34. $f(t) = t^5(\sec t + e^t)$

35. $f(x) = \dfrac{\sin x}{\cos x + 3}$

36. $g(x) = e^2\left(\sin(\pi/4) - 1\right)$

37. $g(t) = 4t^3 e^t - \sin t \cos t$

38. $f(y) = y(2y^3 - 5y - 1)(6y^2 + 7)$

39. $F(x) = (8x - 1)(x^2 + 4x + 7)(x^3 - 5)$

40. $h(t) = \dfrac{t^2 \sin t + 3}{t^2 \cos t + 2}$

41. $f(x) = x^2 e^x \tan x$

42. $g(x) = 2x \sin x \sec x$

In Exercises 43–46, find the equations of the tangent line to the graph of g at the indicated point.

43. $g(s) = e^s(s^2 + 2)$ at $(0, 2)$.

44. $g(t) = t \sin t$ at $\left(\frac{3\pi}{2}, -\frac{3\pi}{2}\right)$

45. $g(x) = \dfrac{x^2}{x - 1}$ at $(2, 4)$

46. $g(\theta) = \dfrac{\cos \theta - 8\theta}{\theta + 1}$ at $(0, 1)$

In Exercises 47–50, find the x–values where the graph of the function has a horizontal tangent line.

47. $f(x) = 6x^2 - 18x - 24$

48. $f(x) = x \sin x$ on $[-1, 1]$

49. $f(x) = \dfrac{x}{x+1}$

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

In Exercises 51–54, find the requested derivative.

51. $f(x) = x \sin x$; find $f''(x)$.

52. $f(x) = x \sin x$; find $f^{(4)}(x)$.

53. $f(x) = \csc x$; find $f''(x)$.

54. $f(x) = (x^3 - 5x + 2)(x^2 + x - 7)$; find $f^{(8)}(x)$.

In Exercises 55–60, f and g are differentiable functions such that $f(2) = 3$, $f'(2) = -1$, $g(2) = -5$, and $g'(2) = 2$. Evaluate the expressions.

55. $(f + g)'(2)$

56. $(f - g)'(2)$

57. $(4f)'(2)$

58. $(f \cdot g)'(2)$

59. $\left(\dfrac{f}{g}\right)'(2)$

60. $\left(\dfrac{g}{f+g}\right)'(2)$

61. If f and g are functions whose graphs are shown, evaluate the expressions.
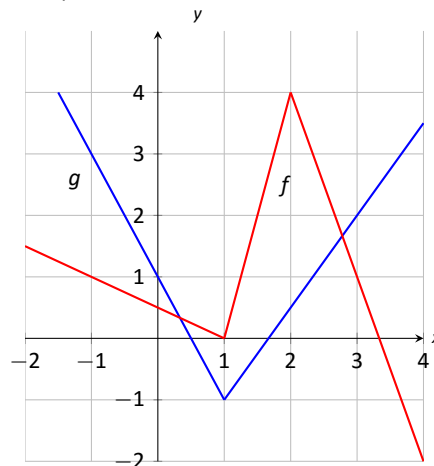
(a) $(fg)'(-1)$ (b) $(f/g)'(-1)$
(c) $(fg)'(3)$ (d) $(g/f)'(3)$

2.5   The Chain Rule

We have covered almost all of the derivative rules that deal with combinations of two (or more) functions. The operations of addition, subtraction, multiplication (including by a constant) and division led to the Sum and Difference rules, the Constant Multiple Rule, the Power Rule, the Product Rule and the Quotient Rule. To complete the list of differentiation rules, we look at the last way two (or more) functions can be combined: the process of composition (i.e. one function "inside" another).

One example of a composition of functions is $f(x) = \cos(x^2)$. We currently do not know how to compute this derivative. If forced to guess, one would likely guess $f'(x) = -\sin(2x)$, where we recognize $-\sin x$ as the derivative of $\cos x$ and $2x$ as the derivative of x^2. However, this is not the case; $f'(x) \neq -\sin(2x)$. In Example 4 we'll see the correct answer, which employs the new rule this section introduces, the **Chain Rule**.

Before we define this new rule, recall the notation for composition of functions. We write $(f \circ g)(x)$ or $f(g(x))$, read as "f of g of x," to denote composing f with g. In shorthand, we simply write $f \circ g$ or $f(g)$ and read it as "f of g." Before giving the corresponding differentiation rule, we note that the rule extends to multiple compositions like $f(g(h(x)))$ or $f(g(h(j(x))))$, etc.

To motivate the rule, let's look at three derivatives we can already compute.

Example 1   **Exploring similar derivatives**
Find the derivatives of $F_1(x) = (1 - x)^2$, $F_2(x) = (1 - x)^3$, and $F_3(x) = (1 - x)^4$. (We'll see later why we are using subscripts for different functions and an uppercase F.)

SOLUTION   In order to use the rules we already have, we must first expand each function as $F_1(x) = 1 - 2x + x^2$, $F_2(x) = 1 - 3x + 3x^2 - x^3$ and $F_3(x) = 1 - 4x + 6x^2 - 4x^3 + x^4$.

It is not hard to see that:

$$F_1'(x) = -2 + 2x,$$
$$F_2'(x) = -3 + 6x - 3x^2 \qquad \text{and}$$
$$F_3'(x) = -4 + 12x - 12x^2 + 4x^3.$$

An interesting fact is that these can be rewritten as

$$F_1'(x) = -2(1 - x),$$
$$F_2'(x) = -3(1 - x)^2 \qquad \text{and}$$
$$F_3'(x) = -4(1 - x)^3.$$

Notes:

A pattern might jump out at you. Recognize that each of these functions is a composition, letting $g(x) = 1 - x$:

$$F_1(x) = f_1(g(x)), \text{ where } f_1(x) = x^2,$$
$$F_2(x) = f_2(g(x)), \text{ where } f_2(x) = x^3,$$
$$F_3(x) = f_3(g(x)), \text{ where } f_3(x) = x^4.$$

We'll come back to this example after giving the formal statements of the Chain Rule; for now, we are just illustrating a pattern.

Theorem 19   The Chain Rule

Let $y = f(u)$ be a differentiable function of u and let $u = g(x)$ be a differentiable function of x. Then $y = f(g(x))$ is a differentiable function of x, and

$$y' = f'(g(x)) \cdot g'(x).$$

We can think of this as taking the derivative of the outer function evaluated at the inner function times the derivative of the inner function. To help understand the Chain Rule, we return to Example 1.

Example 2   Using the Chain Rule

Use the Chain Rule to find the derivatives of the functions given in Example 1.

SOLUTION   Example 1 ended with the recognition that each of the given functions was actually a composition of functions. To avoid confusion, we ignore most of the subscripts here.

$F_1(x) = (1-x)^2$:

We found that

$$y = (1-x)^2 = f(g(x)), \text{ where } f(x) = x^2 \text{ and } g(x) = 1 - x.$$

To find y', we apply the Chain Rule. We need $f'(x) = 2x$ and $g'(x) = -1$.

Part of the Chain Rule uses $f'(g(x))$. This means substitute $g(x)$ for x in the equation for $f'(x)$. That is, $f'(x) = 2(1-x)$. Finishing out the Chain Rule we have

$$y' = f'(g(x)) \cdot g'(x) = 2(1-x) \cdot (-1) = -2(1-x) = 2x - 2.$$

$F_2(x) = (1-x)^3$:

Let $y = (1-x)^3 = f(g(x))$, where $f(x) = x^3$ and $g(x) = (1-x)$. We have $f'(x) = 3x^2$, so $f'(g(x)) = 3(1-x)^2$. The Chain Rule then states

$$y' = f'(g(x)) \cdot g'(x) = 3(1-x)^2 \cdot (-1) = -3(1-x)^2.$$

Notes:

$F_3(x) = (1-x)^4$:

Finally, when $y = (1-x)^4$, we have $f(x) = x^4$ and $g(x) = (1-x)$. Thus $f'(x) = 4x^3$ and $f'(g(x)) = 4(1-x)^3$. Thus

$$y' = f'(g(x)) \cdot g'(x) = 4(1-x)^3 \cdot (-1) = -4(1-x)^3.$$

Example 2 demonstrated a particular pattern: when $f(x) = x^n$ and $y = f(g(x))$, then $y' = n \cdot (g(x))^{n-1} \cdot g'(x)$. This is called the Generalized Power Rule.

> **Theorem 20   Generalized Power Rule**
>
> Let $g(x)$ be a differentiable function. Then
>
> $$\frac{d}{dx}\left(g(x)^n\right) = n \cdot (g(x))^{n-1} \cdot g'(x).$$

This allows us to quickly find the derivative of functions like $y = (3x^2 - 5x + 7 + \sin x)^{20}$. While it may look intimidating, the Generalized Power Rule states that

$$y' = 20(3x^2 - 5x + 7 + \sin x)^{19} \cdot \frac{d}{dx}(3x^2 - 5x + 7 + \sin x)$$
$$= 20(3x^2 - 5x + 7 + \sin x)^{19} \cdot (6x - 5 + \cos x).$$

Treat the derivative–taking process step–by–step. In the example just given, first multiply by 20, then rewrite the inside of the parentheses, raising it all to the 19th power. Then think about the derivative of the expression inside the parentheses, and multiply by that.

Watch the video:
Chain Rule for Finding Derivatives at
`https://youtu.be/6kScLENCXLg`

We now consider more examples that employ the Chain Rule.

Example 3   Using the Chain Rule
Find the derivatives of the following functions:

1. $y = \sin 2x$ 2. $y = \ln(4x^3 - 2x^2)$ 3. $y = e^{-x^2}$

Notes:

Solution

1. Consider $y = \sin 2x$. Recognize that this is a composition of functions, where $f(x) = \sin x$ and $g(x) = 2x$. Thus

$$y' = f'(g(x)) \cdot g'(x) = \cos(2x) \cdot \frac{d}{dx}(2x) = \cos(2x) \cdot 2 = 2\cos 2x.$$

2. Recognize that $y = \ln(4x^3 - 2x^2)$ is the composition of $f(x) = \ln x$ and $g(x) = 4x^3 - 2x^2$. Also, recall that

$$\frac{d}{dx}\left(\ln x\right) = \frac{1}{x}.$$

This leads us to:

$$
\begin{aligned}
y' &= \frac{1}{4x^3 - 2x^2} \cdot \frac{d}{dx}(4x^3 - 2x^2) \\
&= \frac{1}{4x^3 - 2x^2} \cdot (12x^2 - 4x) \\
&= \frac{12x^2 - 4x}{4x^3 - 2x^2} \\
&= \frac{4x(3x - 1)}{2x(2x^2 - x)} \\
&= \frac{2(3x - 1)}{2x^2 - x}.
\end{aligned}
$$

3. Recognize that $y = e^{-x^2}$ is the composition of $f(x) = e^x$ and $g(x) = -x^2$. Remembering that $f'(x) = e^x$, we have

$$y' = e^{x^2} \cdot \frac{d}{dx}(x^2) = e^{-x^2} \cdot (-2x) = (-2x)e^{-x^2}.$$
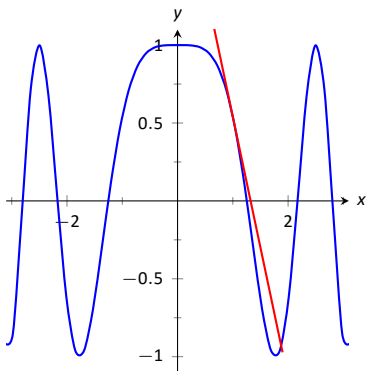
Figure 2.17: $f(x) = \cos x^2$ sketched along with its tangent line at $x = 1$.

Example 4   **Using the Chain Rule to find a tangent line**
Let $f(x) = \cos x^2$. Find the equation of the line tangent to the graph of f at $x = 1$.

Solution   The tangent line goes through the point $(1, f(1)) \approx (1, 0.54)$ with slope $f'(1)$. To find f', we need the Chain Rule.
  $f'(x) = -\sin(x^2) \cdot (2x) = -2x\sin x^2$. Evaluated at $x = 1$, we have $f'(1) = -2\sin 1 \approx -1.68$. Thus the equation of the tangent line is approximately

$$y = -1.68(x - 1) + 0.54.$$

The tangent line is sketched along with f in Figure 2.17.

Notes:

The Chain Rule is used often in taking derivatives. Because of this, one can become familiar with the basic process and learn patterns that facilitate finding derivatives quickly. For instance,

$$\frac{d}{dx}\Big(\ln(\text{anything})\Big) = \frac{1}{\text{anything}} \cdot \frac{d}{dx}(\text{anything}) = \frac{\frac{d}{dx}(\text{anything})}{\text{anything}}.$$

A concrete example of this is

$$\frac{d}{dx}\Big(\ln(3x^{15} - \cos x + e^x)\Big) = \frac{45x^{14} + \sin x + e^x}{3x^{15} - \cos x + e^x}.$$

While the derivative may look intimidating at first, look for the pattern. The denominator is the same as what was inside the natural log function; the numerator is simply its derivative.

This pattern recognition process can be applied to lots of functions. In general, instead of writing "anything", we use u as a generic function of x. We then say

$$\frac{d}{dx}\Big(\ln u\Big) = \frac{u'}{u}.$$

The following is a short list of how the Chain Rule can be quickly applied to familiar functions.

1. $\dfrac{d}{dx}\Big(u^n\Big) = n \cdot u^{n-1} \cdot u'.$ 4. $\dfrac{d}{dx}\Big(\cos u\Big) = -u' \cdot \sin u.$

2. $\dfrac{d}{dx}\Big(e^u\Big) = u' \cdot e^u.$ 5. $\dfrac{d}{dx}\Big(\tan u\Big) = u' \cdot \sec^2 u.$

3. $\dfrac{d}{dx}\Big(\sin u\Big) = u' \cdot \cos u.$

Of course, the Chain Rule can be applied in conjunction with any of the other rules we have already learned. We practice this next.

Example 5   **Using the Product, Quotient and Chain Rules**
Find the derivatives of the following functions.

1. $f(x) = x^5 \sin 2x^3$ 2. $f(x) = \dfrac{5x^3}{e^{-x^2}}.$

SOLUTION

Notes:

1. We must use the Product and Chain Rules. Do not think that you must be able to "see" the whole answer immediately; rather, just proceed step–by–step.

$$f'(x) = x^5 \cdot \frac{d}{dx}(\sin 2x^3) + \frac{d}{dx}(x^5) \cdot \sin 2x^3$$

$$= x^5 \cdot [\cos 2x^3 \cdot \frac{d}{dx}(2x^3)] + 5x^4 \cdot \sin 2x^3$$

$$= x^5 (6x^2 \cos 2x^3) + 5x^4 (\sin 2x^3)$$

$$= 6x^7 \cos 2x^3 + 5x^4 \sin 2x^3.$$

2. We must employ the Quotient Rule along with the Chain Rule. Again, proceed step–by–step.

$$f'(x) = \frac{e^{-x^2} \cdot \frac{d}{dx}(5x^3) - 5x^3 \frac{d}{dx} e^{-x^2}}{(e^{-x^2})^2}$$

$$= \frac{e^{-x^2} \cdot 15x^2 - 5x^3 \cdot e^{-x^2} \cdot \frac{d}{dx}(-x^2)}{(e^{-x^2})^2}$$

$$= \frac{e^{-x^2}(15x^2) - 5x^3((-2x)e^{-x^2})}{(e^{-x^2})^2}$$

$$= \frac{e^{-x^2}(10x^4 + 15x^2)}{e^{-2x^2}}$$

$$= e^{x^2}(10x^4 + 15x^2).$$

A key to correctly working these problems is to break the problem down into smaller, more manageable pieces. For instance, when using the Product and Chain Rules together, just consider the first part of the Product Rule at first: $f(x)g'(x)$. Just rewrite $f(x)$, then find $g'(x)$. Then move on to the $f'(x)g(x)$ part. Don't attempt to figure out both parts at once.

Likewise, using the Quotient Rule, approach the numerator in two steps and handle the denominator after completing that. Only simplify afterward.

We can also employ the Chain Rule itself several times, as shown in the next example.

Example 6   **Using the Chain Rule multiple times**
Find the derivative of $y = \tan^5(6x^3 - 7x)$.

Notes:

SOLUTION   Recognize that we have the function $g(x) = \tan(6x^3 - 7x)$ "inside" the function $f(x) = x^5$; that is, we have $y = \left(\tan(6x^3 - 7x)\right)^5$. We use the Chain Rule multiple times, beginning with the Generalized Power Rule:

$$y' = 5\left(\tan(6x^3 - 7x)\right)^4 \cdot \frac{d}{dx}\tan(6x^3 - 7x)$$

$$= 5\tan^4(6x^3 - 7x) \cdot \sec^2(6x^3 - 7x) \cdot \frac{d}{dx}(6x^3 - 7x)$$

$$= 5\tan^4(6x^3 - 7x) \cdot \sec^2(6x^3 - 7x) \cdot (18x^2 - 7)$$

$$= 5(18x^2 - 7)\tan^4(6x^3 - 7x)\sec^2(6x^3 - 7x)$$

This function is frankly a ridiculous function, possessing no real practical value. It is very difficult to graph, as the tangent function has many vertical asymptotes and $6x^3 - 7x$ grows so very fast. The important thing to learn from this is that the derivative can be found. In fact, it is not "hard;" one must take several small steps and be careful to keep track of how to apply each of these steps.

It is a traditional mathematical exercise to find the derivatives of arbitrarily complicated functions just to demonstrate that it *can be done*. Just break everything down into smaller pieces.

Example 7   Using the Product, Quotient and Chain Rules

Find the derivative of $f(x) = \dfrac{x\cos(x^{-2}) - \sin^2(e^{4x})}{\ln x^2}$.

SOLUTION   This function likely has no practical use outside of demonstrating derivative skills. The answer is given below without simplification. It employs the Quotient Rule, the Product Rule, and the Chain Rule three times.
$f'(x) =$

$$\frac{(\ln x^2)[-x(\sin x^{-2})(-2x^{-3}) + 1 \cdot (\cos(x^{-2})) - 2\sin e^{4x}\cos e^{4x} \cdot (4e^{4x})] - \frac{1}{x^2}(2x) \cdot [x\cos(x^{-2}) - \sin^2(e^{4x})]}{(\ln x^2)^2}.$$

The reader is highly encouraged to look at each term and recognize why it is there. This example demonstrates that derivatives can be computed systematically, no matter how arbitrarily complicated the function is.

Alternate Chain Rule Notation

It is instructive to understand what the Chain Rule "looks like" using "$\frac{dy}{dx}$" notation instead of y' notation. Suppose that $y = f(u)$ is a function of u, where

Notes:

$u = g(x)$ is a function of x, as stated in Theorem 19. Then, through the composition $f \circ g$, we can think of y as a function of x, as $y = f(g(x))$. Thus the derivative of y with respect to x makes sense; we can talk about $\frac{dy}{dx}$. This leads to an interesting progression of notation:

$$y' = f'(g(x)) \cdot g'(x)$$

$$\frac{dy}{dx} = y'(u) \cdot u'(x) \qquad \text{(since } y = f(u) \text{ and } u = g(x))$$

$$\frac{dy}{dx} = \frac{dy}{du} \cdot \frac{du}{dx} \qquad \text{(using "fractional" notation for the derivative)}$$

Here the "fractional" aspect of the derivative notation stands out. On the right hand side, it seems as though the "du" terms divide out, leaving

$$\frac{dy}{dx} = \frac{dy}{dx}.$$

It is important to realize that we *are not* dividing these terms; the derivative notation of $\frac{dy}{dx}$ is one symbol. It is equally important to realize that this notation was chosen precisely because of this behavior. It makes applying the Chain Rule easy with multiple variables. For instance,

$$\frac{dy}{dt} = \frac{dy}{d\square} \cdot \frac{d\square}{d\triangle} \cdot \frac{d\triangle}{dt}.$$

where $\square$ and $\triangle$ are any variables you'd like to use.

One of the most common ways of "visualizing" the Chain Rule is to consider a set of gears, as shown in Figure 2.18. The gears have 36, 18, and 6 teeth, respectively. That means for every revolution of the x gear, the u gear revolves twice. That is, the rate at which the u gear makes a revolution is twice as fast as the rate at which the x gear makes a revolution. Using the terminology of calculus, the rate of u-change, with respect to x, is $\frac{du}{dx} = 2$.

Likewise, every revolution of u causes 3 revolutions of y: $\frac{dy}{du} = 3$. How does y change with respect to x? For each revolution of x, y revolves 6 times; that is,

$$\frac{dy}{dx} = \frac{dy}{du} \cdot \frac{du}{dx} = 2 \cdot 3 = 6.$$

We can then extend the Chain Rule with more variables by adding more gears to the picture.

It is difficult to overstate the importance of the Chain Rule. So often the functions that we deal with are compositions of two or more functions, requiring us to use this rule to compute derivatives. It is often used in practice when actual
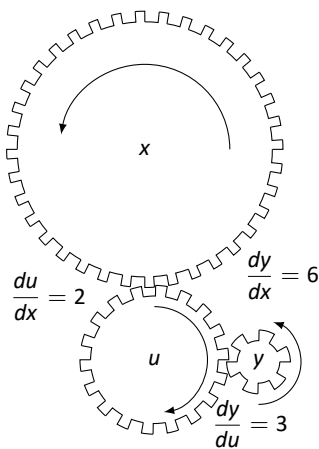
Figure 2.18: A series of gears to demonstrate the Chain Rule. Note how $\frac{dy}{dx} = \frac{dy}{du} \cdot \frac{du}{dx}$

Notes:

functions are unknown. Rather, through measurement, we can calculate $\frac{dy}{du}$ and $\frac{du}{dx}$. With our knowledge of the Chain Rule, finding $\frac{dy}{dx}$ is straightforward.

In the next section, we use the Chain Rule to justify another differentiation technique. There are many curves that we can draw in the plane that fail the "vertical line test." For instance, consider $x^2 + y^2 = 1$, which describes the unit circle. We may still be interested in finding slopes of tangent lines to the circle at various points. The next section shows how we can find $\frac{dy}{dx}$ without first "solving for y." While we can in this instance, in many other instances solving for y is impossible. In these situations, *implicit differentiation* is indispensable.

Notes:

Exercises 2.5

Terms and Concepts

1. T/F: The Chain Rule describes how to evaluate the derivative of a composition of functions.

2. T/F: The Generalized Power Rule states that $\dfrac{d}{dx}\left(g(x)^n\right) = n\big(g(x)\big)^{n-1}$.

3. T/F: $\dfrac{d}{dx}\left(\ln(x^2)\right) = \dfrac{1}{x^2}$.

4. T/F: $\dfrac{d}{dx}(3^x) \approx 1.1 \cdot 3^x$.

5. T/F: $\dfrac{dx}{dy} = \dfrac{dx}{dt} \cdot \dfrac{dt}{dy}$

6. T/F: Taking the derivative of $f(x) = x^2 \sin(5x)$ requires the use of both the Product and Chain Rules.

Problems

In Exercises 7–33, compute the derivative of the given function.

7. $f(x) = (4x^3 - x)^{10}$

8. $f(t) = (3t - 2)^5$

9. $g(\theta) = (\sin\theta + \cos\theta)^3$

10. $h(t) = e^{3t^2 + t - 1}$

11. $f(x) = \left(x + \dfrac{1}{x}\right)^4$

12. $p(x) = \left(x^2 - \dfrac{1}{x^2}\right)^6$

13. $f(x) = \cos(3x)$

14. $g(x) = \tan(5x)$

15. $h(t) = \sin^4(2t)$

16. $p(t) = \cos^3(t^2 + 3t + 1)$

17. $g(x) = \tan^2 x - \tan(x^2)$

18. $w(x) = \sec(e^{x^3})$

19. $f(x) = \ln(\cos x)$

20. $f(x) = \ln(x^2)$

21. $f(x) = 2\ln(x)$

22. $g(t) = 15^2$

23. $r(x) = \dfrac{\sqrt{4x - 3}}{x^2}$

24. $f(x) = \dfrac{(3x^2 - 5)^4}{(2x^3 - 1)^2}$

25. $h(x) = [(2x + 1)^{10} + 1]^{10}$

26. $f(t) = \left[\left(1 + \dfrac{1}{t}\right)^{-1} + 1\right]^{-1}$

27. $F(x) = 2x(2x + 1)^2(2x + 3)^3$

28. $f(x) = x^2 \sin(5x)$

29. $g(t) = \cos(t^2 + 3t)\sin(5t - 7)$

30. $g(t) = \cos(\tfrac{1}{t})e^{5t^2}$

31. $a(t) = 7t^3 e^{\tan t^2}$

32. $y = \sqrt{\sin(\cos^2 x)}$

33. $k(x) = \cos(x\sin x^3)$

34. If $k(x) = f(g(x))$ with $f(2) = -4$, $g(2) = 2$, $f'(2) = 3$, and $g'(2) = 5$. Find $k'(2)$.

35. Suppose $r(x) = f(g(h(x)))$, where $h(1) = 2$, $g(2) = 3$, $h'(1) = 3$, $g'(2) = 5$, and $f'(3) = 6$. Find $r'(1)$.

36. If f and g are functions whose graphs are shown, evaluate the expressions.
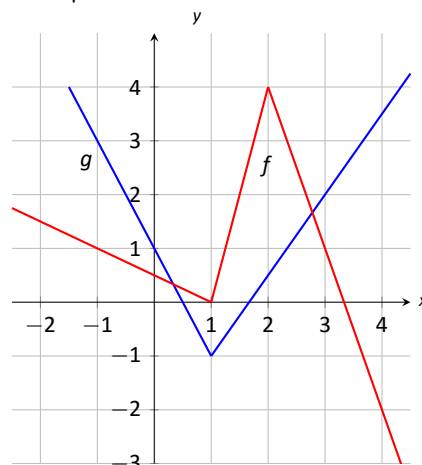
(a) $(f \circ g)'(-1)$ (b) $(g \circ f)'(0)$
(c) $(g \circ g)'(-1)$ (d) $(f \circ f)'(4)$

37.

x	$f(x)$	$f'(x)$	$g(x)$	$g'(x)$
1	4	5	4	5
4	0	7	1	$\frac{1}{2}$
6	6	4	6	3

Use the given table of values for f, g, f', and g' to find

(a) $(f \circ g)'(6)$

(b) $(g \circ f)'(1)$

(c) $(g \circ g)'(6)$

(d) $(f \circ f)'(1)$

In Exercises 38–41, find the equations of tangent line to the graph of the function at the given point. Note: the functions here are the same as in Exercises 7 – 10.

38. $f(x) = (4x^3 - x)^{10}$ at $x = 0$

39. $f(t) = (3t - 2)^5$ at $t = 1$

40. $g(\theta) = (\sin\theta + \cos\theta)^3$ at $\theta = \pi/2$

41. $h(t) = e^{3t^2 + t - 1}$ at $t = -1$

42. Compute $\dfrac{d}{dx}\left(\ln(kx)\right)$ two ways:

(a) Using the Chain Rule, and

(b) by first using the logarithm rule $\ln(ab) = \ln a + \ln b$, then taking the derivative.

43. Compute $\dfrac{d}{dx}\left(\ln(x^k)\right)$ two ways:

    (a) Using the Chain Rule, and

    (b) by first using the logarithm rule $\ln(a^p) = p\ln a$, then taking the derivative.

44. Use the Chain Rule to prove the following:

    (a) The derivative of an even function is an odd function.

    (b) The derivative of an odd function is an even function.

45. Use the Chain Rule and Product Rule to give an alternative proof of the Quotient Rule. (Hint: write $f(x)/g(x)$ as $f(x) \cdot [g(x)]^{-1}$).

46. Use the Chain Rule to express the second derivative of $f(g(x))$ in terms of first and second derivatives of f and g.

Review

47. The "wind chill factor" is a measurement of how cold it "feels" during cold, windy weather. Let $W(w)$ be the wind chill factor, in degrees Fahrenheit, when it is $25°$F outside with a wind of w mph.

    (a) What are the units of $W'(w)$?

    (b) What would you expect the sign of $W'(10)$ to be?

48. Find the derivatives of $f(x) = x^2 e^x \cot x$

2.6 Implicit Differentiation

In the previous sections we learned to find the derivative, $\frac{dy}{dx}$, or y', when y is given *explicitly* as a function of x. That is, if we know $y = f(x)$ for some function f, we can find y'. For example, given $y = 3x^2 - 7$, we can easily find $y' = 6x$. (Here we explicitly state how x and y are related. Knowing x, we can directly find y.)

Sometimes the relationship between y and x is not explicit; rather, it is *implicit*. For instance, we might know that $x^2 - y = 4$. This equality defines a relationship between x and y; if we know x, we could figure out y. Can we still find y'? In this case, sure; we solve for y to get $y = x^2 - 4$ (hence we now know y explicitly) and then differentiate to get $y' = 2x$.

Sometimes the *implicit* relationship between x and y is complicated. Suppose we are given $\sin(y) + y^3 = 6 - x^3$. A graph of this equation is given in Figure 2.19. In this case there is absolutely no way to solve for y in terms of elementary functions. The surprising thing is, however, that we can still find y' via a process known as **implicit differentiation**.

Implicit differentiation is a technique based on the Chain Rule that is used to find a derivative when the relationship between the variables is given implicitly rather than explicitly (solved for one variable in terms of the other).

We begin by reviewing the Chain Rule. Let f and g be functions of x. Then

$$\frac{d}{dx}\Big(f(g(x))\Big) = f'(g(x)) \cdot g'(x).$$

Suppose now that $y = g(x)$. We can rewrite the above as

$$\frac{d}{dx}\Big(f(y)\Big) = f'(y) \cdot y', \quad \text{or} \quad \frac{d}{dx}\Big(f(y)\Big) = f'(y) \cdot \frac{dy}{dx}. \tag{2.1}$$

These equations look strange; the key concept to learn here is that we can find y' even if we don't exactly know how y and x relate.
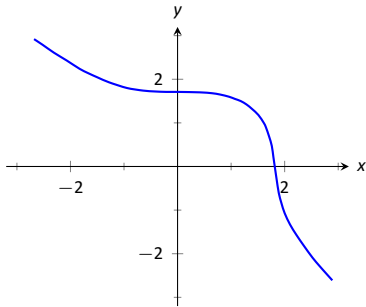
Figure 2.19: A graph of the equation $\sin(y) + y^3 = 6 - x^3$.

Watch the video:
Showing explicit and implicit differentiation give same result at
`https://youtu.be/2CsQ_11S2_Y`

We demonstrate this process in the following example.

Notes:

Example 1   **Using Implicit Differentiation**

Find y' given that $\sin(y) + y^3 = 6 - x^3$.

Solution   We start by taking the derivative of both sides (thus maintaining the equality). We have:

$$\frac{d}{dx}\left(\sin(y) + y^3\right) = \frac{d}{dx}\left(6 - x^3\right).$$

The right hand side is easy; it returns $-3x^2$.

The left hand side requires more consideration. We take the derivative term–by–term. Using the technique derived from Equation (2.1) above, we can see that

$$\frac{d}{dx}\left(\sin y\right) = \cos y \cdot y'.$$

We apply the same process to the y^3 term.

$$\frac{d}{dx}\left(y^3\right) = \frac{d}{dx}\left((y)^3\right) = 3(y)^2 \cdot y'.$$

Putting this together with the right hand side, we have

$$\cos(y)y' + 3y^2y' = -3x^2.$$

Now solve for y'.

$$\cos(y)y' + 3y^2y' = -3x^2.$$
$$\left(\cos y + 3y^2\right)y' = -3x^2$$
$$y' = \frac{-3x^2}{\cos y + 3y^2}$$

This equation for y' probably seems unusual for it contains both x and y terms. How is it to be used? We'll address that next.

Implicit functions are generally harder to deal with than explicit functions. With an explicit function, given an x value, we have an explicit formula for computing the corresponding y value. With an implicit function, one often has to find x and y values *at the same time* that satisfy the equation. It is much easier to demonstrate that a given point satisfies the equation than to actually find such a point.

For instance, we can affirm easily that the point $(\sqrt[3]{6}, 0)$ lies on the graph of the equation $\sin y + y^3 = 6 - x^3$. Plugging in 0 for y, we see the left hand side is 0. Setting $x = \sqrt[3]{6}$, we see the right hand side is also 0; the equation is satisfied. The following example finds an equation of the tangent line to this equation at this point.

Notes:

Example 2   **Using Implicit Differentiation to find a tangent line**
Find the equation of the line tangent to the implicitly defined curve $\sin y + y^3 = 6 - x^3$ at the point $(\sqrt[3]{6}, 0)$.

SOLUTION   In Example 1 we found that

$$y' = \frac{-3x^2}{\cos y + 3y^2}.$$

We find the slope of the tangent line at the point $(\sqrt[3]{6}, 0)$ by substituting $\sqrt[3]{6}$ for x and 0 for y. Thus at the point $(\sqrt[3]{6}, 0)$, we have the slope as

$$y' = \frac{-3(\sqrt[3]{6})^2}{\cos 0 + 3 \cdot 0^2} = \frac{-3\sqrt[3]{36}}{1} \approx -9.91.$$

Therefore an equation of the tangent line to the implicitly defined curve $\sin y + y^3 = 6 - x^3$ at the point $(\sqrt[3]{6}, 0)$ is

$$y = -3\sqrt[3]{36}(x - \sqrt[3]{6}) + 0 \approx -9.91x + 18.$$

The curve and this tangent line are shown in Figure 2.20.

This suggests a general method for implicit differentiation. For the steps below assume y is a function of x.

1. Take the derivative of each term in the equation. Treat the x terms like normal. When taking the derivatives of y terms, the usual rules apply except that, because of the Chain Rule, we need to multiply each term by y'.

2. Get all the y' terms on one side of the equal sign and put the remaining terms on the other side.

3. Factor out y'; solve for y' by dividing.

Practical Note: When working by hand, it may be beneficial to use the symbol $\frac{dy}{dx}$ instead of y', as the latter can be easily confused for y or y^1.

Example 3   **Using Implicit Differentiation**
Given the implicitly defined function $y^3 + x^2y^4 = 1 + 2x$, find y'.

Figure 2.20: The equation $\sin y + y^3 = 6 - x^3$ and its tangent line at the point $(\sqrt[3]{6}, 0)$.

Notes:

SOLUTION   We will take the implicit derivatives term by term. Using the Chain Rule the derivative of y^3 is $3y^2y'$.

The second term, x^2y^4 is a little more work. It requires the Product Rule as it is the product of two functions of x: x^2 and y^4. We see that $\dfrac{d}{dx}(x^2y^4)$ is

$$x^2 \cdot \frac{d}{dx}(y^4) + \frac{d}{dx}(x^2) \cdot y^4$$
$$x^2 \cdot (4y^3y') + 2x \cdot y^4$$

The first part of this expression requires a y' because we are taking the derivative of a y term. The second part does not require it because we are taking the derivative of x^2.

The derivative of the right hand side of the equation is found to be 2. In all, we get:

$$3y^2y' + 4x^2y^3y' + 2xy^4 = 2.$$

Move terms around so that the left side consists only of the y' terms and the right side consists of all the other terms:

$$3y^2y' + 4x^2y^3y' = 2 - 2xy^4.$$

Factor out y' from the left side and solve to get

$$y' = \frac{2 - 2xy^4}{3y^2 + 4x^2y^3}.$$

To confirm the validity of our work, let's find the equation of a tangent line to this curve at a point. It is easy to confirm that the point $(0, 1)$ lies on the graph of this curve. At this point, $y' = 2/3$. So the equation of the tangent line is $y = 2/3(x - 0) + 1$. The equation and its tangent line are graphed in Figure 2.21.

Notice how our curve looks much different than other functions we have worked with up to this point. Such curves are important in many areas of mathematics, so developing tools to deal with them is also important.

Example 4   **Using Implicit Differentiation**
Given the implicitly defined curve $\sin(x^2y^2) + y^3 = x + y$, find y'.

SOLUTION   Differentiating term by term, we find the most difficulty in
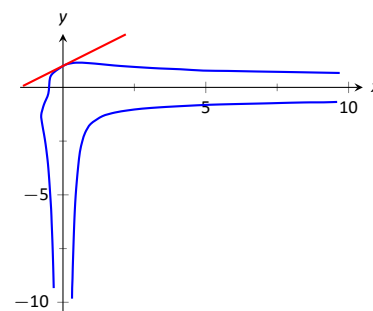
Figure 2.21: A graph of the equation $y^3 + x^2y^4 = 1 + 2x$ along with its tangent line at the point $(0, 1)$.

Notes:

the first term. It requires both the Chain and Product Rules.

$$\frac{d}{dx}\left(\sin(x^2y^2)\right) = \cos(x^2y^2) \cdot \frac{d}{dx}(x^2y^2)$$

$$= \cos(x^2y^2) \cdot \left(x^2\frac{d}{dx}(y^2) + \frac{d}{dx}(x^2) \cdot y^2\right)$$

$$= \cos(x^2y^2) \cdot (x^2 \cdot 2yy' + 2xy^2)$$

$$= 2(x^2yy' + xy^2)\cos(x^2y^2).$$

We leave the derivatives of the other terms to the reader. After taking the derivatives of both sides, we have

$$2(x^2yy' + xy^2)\cos(x^2y^2) + 3y^2y' = 1 + y'.$$

We now have to be careful to properly solve for y', particularly because of the product on the left. It is best to multiply out the product. Doing this, we get

$$2x^2y\cos(x^2y^2)y' + 2xy^2\cos(x^2y^2) + 3y^2y' = 1 + y'.$$

From here we can safely move around terms to get the following:

$$2x^2y\cos(x^2y^2)y' + 3y^2y' - y' = 1 - 2xy^2\cos(x^2y^2).$$

Then we can solve for y' to get

$$y' = \frac{1 - 2xy^2\cos(x^2y^2)}{2x^2y\cos(x^2y^2) + 3y^2 - 1}.$$

A graph of this implicit equation is given in Figure 2.22(a). It is easy to verify that the points $(0,0)$, $(0,1)$ and $(0,-1)$ all lie on the graph. We can find the slopes of the tangent lines at each of these points using our formula for y'.

At $(0,0)$, the slope is -1.

At $(0,1)$, the slope is $1/2$.

At $(0,-1)$, the slope is also $1/2$.

The tangent lines have been added to the graph of the function in Figure 2.22(b).

Quite a few "famous" curves have equations that are given implicitly. We can use implicit differentiation to find the slope at various points on those curves. We investigate two such curves in the next examples.

(a)

(b)
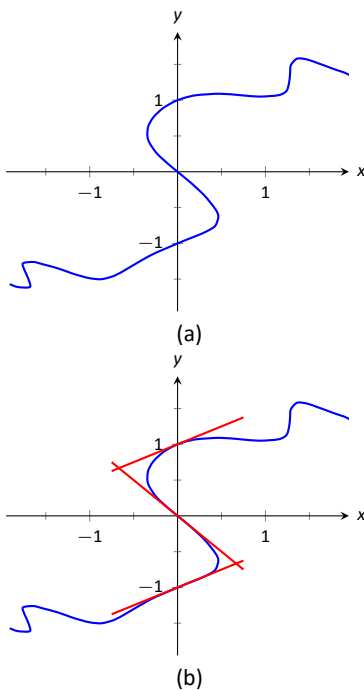
Figure 2.22: A graph of the equation $\sin(x^2y^2) + y^3 = x + y$ and certain tangent lines.

Notes:

Example 5   **Finding slopes of tangent lines to a circle**

Find the slope of the tangent line to the circle $x^2+y^2 = 1$ at the point $(1/2, \sqrt{3}/2)$.

> **SOLUTION**   Taking derivatives, we get $2x+2yy' = 0$. Solving for y' gives:
>
> $$y' = \frac{-x}{y}.$$
>
> This is a clever formula. Recall that the slope of the line through the origin and the point (x, y) on the circle will be y/x. We have found that the slope of the tangent line to the circle at that point is the opposite reciprocal of y/x, namely, $-x/y$. Hence these two lines are always perpendicular.
>
> At the point $(1/2, \sqrt{3}/2)$, we have the tangent line's slope as
>
> $$y' = \frac{-1/2}{\sqrt{3}/2} = \frac{-1}{\sqrt{3}} \approx -0.577.$$
>
> A graph of the circle and its tangent line at $(1/2, \sqrt{3}/2)$ is given in Figure 2.23, along with a thin dashed line from the origin that is perpendicular to the tangent line. (It turns out that all normal lines to a circle pass through the center of the circle.)
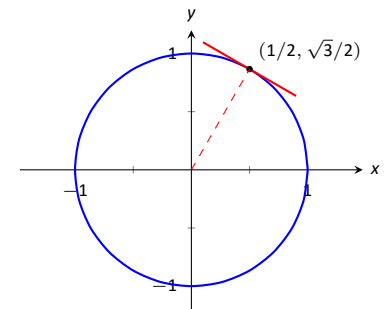
Figure 2.23: The unit circle with its tangent line at $(1/2, \sqrt{3}/2)$.

This section has shown how to find the derivatives of implicitly defined functions, whose graphs include a wide variety of interesting and unusual shapes. Implicit differentiation can also be used to further our understanding of "regular" differentiation.

Implicit Differentiation and the Second Derivative

We can use implicit differentiation to find higher order derivatives. In theory, this is simple: first find $\frac{dy}{dx}$, then take its derivative with respect to x. In practice, it is not hard, but it often requires a bit of algebra. We demonstrate this in an example.

Example 6   **Finding the second derivative**

Given $x^2 + y^2 = 1$, find $\dfrac{d^2y}{dx^2} = y''$.

> **SOLUTION**   We found that $y' = \frac{dy}{dx} = -x/y$ in Example 5. To find y'', we

Notes:

apply implicit differentiation to y'.

$$y'' = \frac{d}{dx}(y')$$

$$= \frac{d}{dx}\left(-\frac{x}{y}\right) \qquad \text{now use the Quotient Rule}$$

$$= -\frac{y(1) - x(y')}{y^2} \qquad \text{replace } y' \text{ with } -x/y$$

$$= -\frac{y - x(-x/y)}{y^2}$$

$$= -\frac{y^2 + x^2}{y^3}, \qquad \text{since we were given } x^2 + y^2 = 1$$

$$= -\frac{1}{y^3}.$$

We can see that $y'' > 0$ when $y < 0$ and $y'' < 0$ when $y > 0$. In Section 3.4, we will see how this relates to the shape of the graph.

Implicit differentiation proves to be useful as it allows us to find the instantaneous rates of change of a variety of functions.

In this chapter we have defined the derivative, given rules to facilitate its computation, and given the derivatives of a number of standard functions. We restate the most important of these in the following theorem, intended to be a reference for further work.

Notes:

Theorem 21   Glossary of Derivatives of Elementary Functions

Let u and v be differentiable functions, and let c and n be real numbers, $n \neq 0$.

1. $\frac{d}{dx}(cu) = cu'$

2. $\frac{d}{dx}(u \pm v) = u' \pm v'$

3. $\frac{d}{dx}(u \cdot v) = uv' + u'v$

4. $\frac{d}{dx}\left(\frac{u}{v}\right) = \frac{u'v - uv'}{v^2}$

5. $\frac{d}{dx}(u(v)) = u'(v)v'$

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

7. $\frac{d}{dx}(c) = 0$

8. $\frac{d}{dx}(x) = 1$

9. $\frac{d}{dx}(\ln x) = \frac{1}{x}$

10. $\frac{d}{dx}(e^x) = e^x$

11. $\frac{d}{dx}(\sin x) = \cos x$

12. $\frac{d}{dx}(\cos x) = -\sin x$

13. $\frac{d}{dx}(\tan x) = \sec^2 x$

14. $\frac{d}{dx}(\cot x) = -\csc^2 x$

15. $\frac{d}{dx}(\sec x) = \sec x \tan x$

16. $\frac{d}{dx}(\csc x) = -\csc x \cot x$

Notes:

Exercises 2.6

Terms and Concepts

1. In your own words, explain the difference between implicit functions and explicit functions.

2. Implicit differentiation is based on what other differentiation rule?

3. T/F: Implicit differentiation can be used to find the derivative of $y = \sqrt{x}$.

4. T/F: Implicit differentiation can be used to find the derivative of $y = x^{3/4}$.

Problems

In Exercises 5–21, find $\dfrac{dy}{dx}$ using implicit differentiation.

5. $x^4 + y^2 + y = 7$

6. $x^{2/5} + y^{2/5} = 1$

7. $\cos x + \sin y = 1$

8. $\dfrac{x}{y} = 10$

9. $\dfrac{y}{x} = 10$

10. $x^2 \tan y = 50$

11. $(3x^2 + 2y^3)^4 = 2$

12. $(y^2 + 2y - x)^2 = 200$

13. $\dfrac{x^2 + y}{x + y^2} = 17$

14. $\dfrac{\sin(x) + y}{\cos(y) + x} = 1$

15. $\ln(x^2 + y^2) = e$

16. $\ln(x^2 + xy + y^2) = 1$

17. $xe^x = ye^y$

18. $y \sin(x^3) = x \sin(y^3)$

19. $\sqrt{xy} = 1 + x^2 y$

20. $x^2 y - y^2 x = 1$

21. $x^2 + y^2 + xy = 7$

22. Show that $\dfrac{dy}{dx}$ is the same for each of the following implicitly defined functions.

    (a) $xy = 1$

    (b) $x^2 y^2 = 1$

    (c) $\sin(xy) = 1$

    (d) $\ln(xy) = 1$

In Exercises 23–28, find the equation of the tangent line to the graph of the implicitly defined function at the indicated points. As a visual aid, each function is graphed.

23. $x^{2/5} + y^{2/5} = 1$

    (a) At $(1, 0)$.

    (b) At $(0.1, 0.281)$ (which does not *exactly* lie on the curve, but is very close).

24. $x^4 + y^4 = 1$

    (a) At $(1, 0)$.
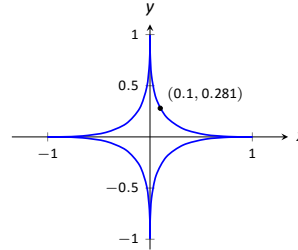
    (b) At $(\sqrt{0.6}, \sqrt{0.8})$.

    (c) At $(0, 1)$.

25. $(x^2 + y^2 - 4)^3 = 108y^2$

    (a) At $(0, 4)$.
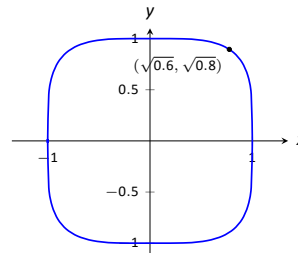
    (b) At $(2, -\sqrt[4]{108})$.

26. $(x^2 + y^2 + x)^2 = x^2 + y^2$

    (a) At $(0, 1)$.
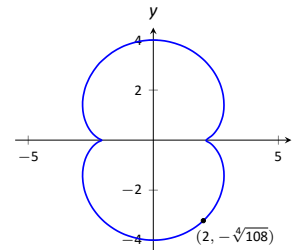
    (b) At $\left(-\dfrac{3}{4}, \dfrac{3\sqrt{3}}{4}\right)$.
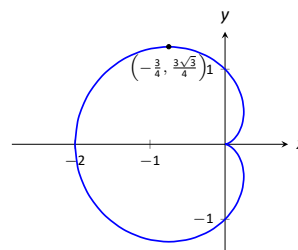
27. $(x-2)^2 + (y-3)^2 = 9$

   (a) At $\left(\dfrac{7}{2}, \dfrac{6+3\sqrt{3}}{2}\right)$.

   (b) At $\left(\dfrac{4+3\sqrt{3}}{2}, \dfrac{3}{2}\right)$.
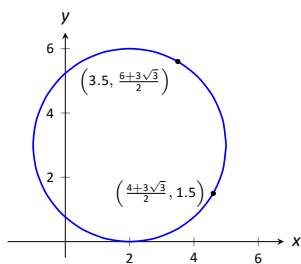
28. $x^2 + 2xy - y^2 + x = 2$

   (a) At $(-2, 0)$.

   (b) At $(1, 2)$.
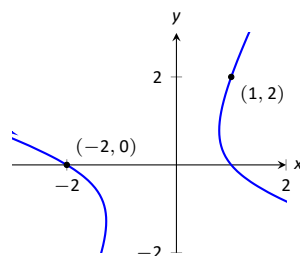
In Exercises 29–32, an implicitly defined function is given. Find $\dfrac{d^2y}{dx^2}$. Note: these are the same problems used in Exercises 5 – 8.

29. $x^4 + y^2 + y = 7$

30. $x^{2/5} + y^{2/5} = 1$

31. $\cos x + \sin y = 1$

32. $\dfrac{x}{y} = 10$

3: THE GRAPHICAL BEHAVIOR OF FUNCTIONS

Our study of limits led to continuous functions, which is a certain class of functions that behave in a particularly nice way. Limits then gave us an even nicer class of functions, functions that are differentiable.

This chapter explores many of the ways we can take advantage of the information that continuous and differentiable functions provide.

3.1 Extreme Values

Given any quantity described by a function, we are often interested in the largest and/or smallest values that quantity attains. For instance, if a function describes the speed of an object, it seems reasonable to want to know the fastest/slowest the object traveled. If a function describes the value of a stock, we might want to know the highest/lowest values the stock attained over the past year. We call such values *extreme values.*

Note: The extreme values of a function are "y" values, values the function attains, not the input values.

> **Definition 14   Extreme Values**
>
> Let f be defined on an interval I containing c.
>
> 1. $f(c)$ is the **minimum** (also, **absolute minimum**) of f on I if $f(c) \leq f(x)$ for all x in I.
>
> 2. $f(c)$ is the **maximum** (also, **absolute maximum**) of f on I if $f(c) \geq f(x)$ for all x in I.
>
> The maximum and minimum values are the **extreme values**, or **extrema**, of f on I.

Consider Figure 3.1. The function displayed in (a) has a maximum, but no minimum, as the interval over which the function is defined is open. In (b), the function has a minimum, but no maximum; there is a discontinuity in the "natural" place for the maximum to occur. Finally, the function shown in (c) has both a maximum and a minimum; note that the function is continuous and the interval on which it is defined is closed.

It is possible for discontinuous functions defined on an open interval to have both a maximum and minimum value, but we have just seen examples where they did not. On the other hand, continuous functions on a closed interval *always* have a maximum and minimum value.

(a)

(b)

(c)

Figure 3.1: Graphs of functions with and without extreme values.

> **Theorem 22   The Extreme Value Theorem**
> Let f be a continuous function defined on a closed interval I. Then f has both a maximum and minimum value on I.

This theorem states that f has extreme values, but it does not offer any advice about how/where to find these values. The process can seem to be fairly easy, as the next example illustrates. After the example, we will draw on lessons learned to form a more general and powerful method for finding extreme values.

> Watch the video:
> Finding Critical Numbers — Example 2 at
> `https://youtu.be/3-6bdDXzl9M`

Example 1   Approximating extreme values
Consider $f(x) = 2x^3 - 9x^2$ on $I = [-1, 5]$, as graphed in Figure 3.2. Approximate the extreme values of f.

SOLUTION   The graph is drawn in such a way to draw attention to certain points. It certainly seems that the smallest y value is -27, found when $x = 3$. It also seems that the largest y value is 25, found at the endpoint of I, $x = 5$. We use the word *seems*, for by the graph alone we cannot be sure the smallest value is not less than -27. Since the problem asks for an approximation, we approximate the extreme values to be 25 and -27.

Figure 3.2: A graph of $f(x) = 2x^3 - 9x^2$ as in Example 1.

Notice how the minimum value came at "the bottom of a valley," and the maximum value came at an endpoint. Also note that while 0 is not an extreme value, it would be if we narrowed our interval to $[-1, 4]$. The idea that the point $(0, 0)$ is the location of an extreme value for some interval is important, leading us to a definition.

Notes:

Definition 15   Relative Minimum and Relative Maximum

Let f be defined on an interval I containing c.

1. If there is an open interval containing c such that $f(c)$ is the minimum value, then $f(c)$ is a **relative minimum** of f. We also say that f has a relative minimum at $(c, f(c))$.

2. If there is an open interval containing c such that $f(c)$ is the maximum value, then $f(c)$ is a **relative maximum** of f. We also say that f has a relative maximum at $(c, f(c))$.

The relative maximum and minimum values comprise the **relative extrema** of f.

Note: The terms *local minimum*, *local maximum*, and *local extrema* are often used as synonyms for *relative minimum*, *relative maximum*, and *relative extrema*.

We briefly practice using these definitions.

Example 2   Approximating relative extrema
Consider $f(x) = (3x^4 - 4x^3 - 12x^2 + 5)/5$, as shown in Figure 3.3. Approximate the relative extrema of f. At each of these points, evaluate f'.

SOLUTION   We still do not have the tools to exactly find the relative extrema, but the graph does allow us to make reasonable approximations. It seems f has relative minima at $x = -1$ and $x = 2$, with values of $f(-1) = 0$ and $f(2) = -5.4$. It also seems that f has a relative maximum at the point $(0, 1)$.

We approximate the relative minima to be 0 and -5.4; we approximate the relative maximum to be 1.

It is straightforward to evaluate $f'(x) = \frac{1}{5}(12x^3 - 12x^2 - 24x)$ at $x = 0, 1$ and 2. In each case, $f'(x) = 0$.

Figure 3.3: A graph of $f(x) = (3x^4 - 4x^3 - 12x^2 + 5)/5$ as in Example 2.

Example 3   Approximating relative extrema
Approximate the relative extrema of $f(x) = (x - 1)^{2/3} + 2$, shown in Figure 3.4. At each of these points, evaluate f'.

SOLUTION   The figure implies that f does not have any relative maxima, but has a relative minimum at $(1, 2)$. In fact, the graph suggests that not only is this point a relative minimum, $y = f(1) = 2$ is *the* minimum value of the function.

We compute $f'(x) = \frac{2}{3}(x - 1)^{-1/3}$. When $x = 1$, f' is undefined.

Figure 3.4: A graph of $f(x) = (x-1)^{2/3}+2$ as in Example 3.

Notes:

What can we learn from the previous two examples? We were able to visually approximate relative extrema, and at each such point, the derivative was either 0 or it was not defined. This observation holds for all functions, leading to a definition and a theorem.

Definition 16   Critical Numbers and Critical Points

Let f be defined at c. The value c is a **critical number** (or **critical value**) of f if $f'(c) = 0$ or $f'(c)$ is not defined.

If c is a critical number of f, then the point $(c, f(c))$ is a **critical point** of f.

Theorem 23   Fermat's Theorem

Let a function f have a relative extrema at the point $(c, f(c))$. Then c is a critical number of f.

It isn't too hard to see why this should be true. If f' is defined at a relative extreme, then the tangent line must be horizontal. Otherwise, we'd be able to move along the graph in the direction given by the tangent line to get a more extreme value.

Be careful to understand that this theorem states "All relative extrema occur at critical points." It does not say "All critical numbers produce relative extrema." For instance, consider $f(x) = x^3$. Since $f'(x) = 3x^2$, it is straightforward to determine that $x = 0$ is a critical number of f. However, f has no relative extrema, as illustrated in Figure 3.5.
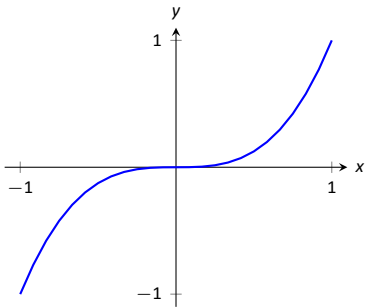
Figure 3.5: A graph of $f(x) = x^3$ which has a critical value of $x = 0$, but no relative extrema.

Theorem 22 states that a continuous function on a closed interval will have absolute extrema, that is, both an absolute maximum and an absolute minimum. These extrema occur either at the endpoints or at critical values in the interval. We combine these concepts to offer a strategy for finding extrema.

Notes:

> **Key Idea 4   Finding Extrema on a Closed Interval**
> Let f be a continuous function defined on a closed interval $[a, b]$. To find the maximum and minimum values of f on $[a, b]$:
>
> 1. Evaluate f at the endpoints a and b of the interval.
>
> 2. Find the critical numbers of f in (a, b).
>
> 3. Evaluate f at each critical number.
>
> 4. The absolute maximum of f is the largest of these values, and the absolute minimum of f is the least of these values.

We practice these ideas in the next examples.

Example 4   Finding extreme values
Find the extreme values of $f(x) = 2x^3 + 3x^2 - 12x$ on $[0, 3]$, graphed in Figure 3.6(a).

SOLUTION   We follow the steps outlined in Key Idea 4. We first evaluate f at the endpoints:

$$f(0) = 2(0)^3 + 3(0)^2 - 12(0) = 0 \quad \text{and} \quad f(3) = 2(3)^3 + 3(3)^2 - 12(3) = 45.$$

Next, we find the critical values of f on $[0, 3]$. We see that $f'(x) = 6x^2 + 6x - 12 = 6(x + 2)(x - 1)$; therefore the critical values of f are $x = -2$ and $x = 1$. Since $x = -2$ does not lie in the interval $[0, 3]$, we ignore it. Evaluating f at the only critical number in our interval gives: $f(1) = 2(1)^3 + 3(1)^2 - 12(1) = -7$.

The table in Figure 3.6(b) gives f evaluated at the "important" x values in $[0, 3]$. We can easily see the maximum and minimum values of f: the maximum value is 45 and the minimum value is -7.

Note that all this was done without the aid of a graph; this work followed an analytic algorithm and did not depend on any visualization. Figure 3.6(a) shows f and we can confirm our answer, but it is important to understand that these answers can be found without graphical assistance.

We practice again.

Example 5   Finding extreme values
Find the maximum and minimum values of f on $[-4, 2]$, where

$$f(x) = \begin{cases} (x - 1)^2 & x \le 0 \\ x + 1 & x > 0 \end{cases}.$$

(a)

x	$f(x)$
0	0
1	-7
3	45

(b)
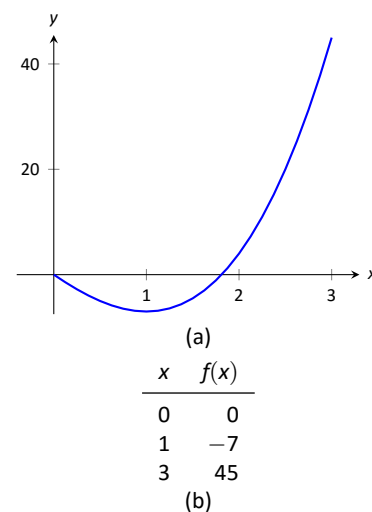
Figure 3.6: A graph and table of extreme values of $f(x) = 2x^3 + 3x^2 - 12x$ on $[0, 3]$ as in Example 4.

Notes:

x	$f(x)$
-4	25
0	1
2	3

(a)
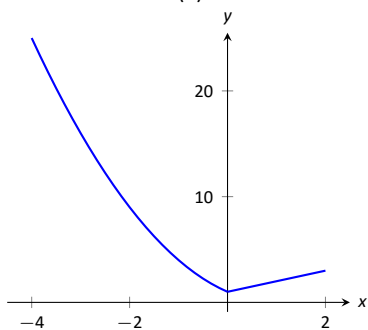
(b)

Figure 3.7: A table of extreme values and graph of $f(x)$ on $[-4, 2]$ as in Example 5.

x	$f(x)$
-2	-0.65
$-\sqrt{\pi}$	-1
0	1
$\sqrt{\pi}$	-1
2	-0.65

(a)
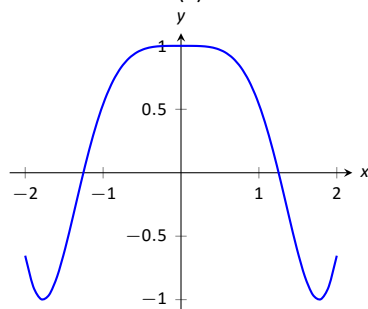
(b)

Figure 3.8: A table of extreme values and graph of $f(x) = \cos(x^2)$ on $[-2, 2]$ in Example 6.

SOLUTION   Here f is piecewise–defined, but we can still apply Key Idea 4 because it is continuous. Evaluating f at the endpoints gives:

$$f(-4) = (-4 - 1)^2 = (-5)^2 = 25 \quad \text{and} \quad f(2) = 2 + 1 = 3.$$

We now find the critical numbers of f. We have to define f' in a piecewise manner; it is

$$f'(x) = \begin{cases} 2(x - 1) & x < 0 \\ 1 & x > 0 \end{cases}.$$

Note that while f is defined for all of $[-4, 2]$, f' is not, as the derivative of f does not exist when $x = 0$. (From the left, the derivative approaches -2; from the right the derivative is 1.) Thus one critical number of f is $x = 0$.

We now set $f'(x) = 0$. When $x > 0$, $f'(x)$ is never 0. When $x < 0$, $f'(x)$ is also never 0. (We may be tempted to say that $f'(x) = 0$ when $x = 1$. However, this is nonsensical, for we only consider $f'(x) = 2(x - 1)$ when $x < 0$, so we will ignore a solution that says $x = 1$.)

So we have three important x values to consider: $x = -4, 2$ and 0. We have already evaluated the first two, and $f(0) = (0 - 1)^2 = (-1)^2 = 1$. Collecting these values into Figure 3.7(a), we see that the absolute minimum of f is 1 and the absolute maximum of f is 25. Our answer is confirmed by the graph of f in Figure 3.7(b).

Example 6   **Finding extreme values**
Find the extrema of $f(x) = \cos(x^2)$ on $[-2, 2]$.

SOLUTION   We again use Key Idea 4. Evaluating f at the endpoints of the interval gives: $f(-2) = f(2) = \cos(4) \approx -0.6536$. We now find the critical values of f.

Applying the Chain Rule, we find $f'(x) = -2x \sin(x^2)$. Set $f'(x) = 0$ and solve for x to find the critical values of f.

We have $f'(x) = 0$ when $x = 0$ and when $\sin(x^2) = 0$. In general, $\sin t = 0$ when $t = \ldots - 2\pi, -\pi, 0, \pi, \ldots$ Thus $\sin(x^2) = 0$ when $x^2 = 0, \pi, 2\pi, \ldots$ (x^2 is always positive so we ignore $-\pi$, etc.) So $\sin(x^2) = 0$ when $x = 0, \pm\sqrt{\pi}, \pm\sqrt{2\pi}, \ldots$. The only values to fall in the given interval of $[-2, 2]$ are 0 and $\pm\sqrt{\pi}$, approximately ± 1.77.

We again construct a table of important values in Figure 3.8(a). In this example we have 5 values to consider: $x = 0, \pm 2, \pm\sqrt{\pi}$.

From the table it is clear that the maximum value of f on $[-2, 2]$ is 1; the minimum value is -1. The graph in Figure 3.8(b) confirms our results.

We consider one more example.

Notes:

Example 7   Finding extreme values

Find the extreme values of $f(x) = \sqrt{1 - x^2}$.

SOLUTION   A closed interval is not given, so we find the extreme values of f on its domain. This f is defined whenever $1 - x^2 \geq 0$; thus the domain of f is $[-1, 1]$. Evaluating f at either endpoint returns 0.

Using the Chain Rule, we find $f'(x) = \dfrac{-x}{\sqrt{1 - x^2}}$. The critical points of f are found when $f'(x) = 0$ or when f' is undefined. It is straightforward to find that $f'(x) = 0$ when $x = 0$, and f' is undefined when $x = \pm 1$, the endpoints of the interval. The table of important values is given in Figure 3.9(b). The maximum value is 1, and the minimum value is 0.

We have seen that continuous functions on closed intervals always have a maximum and minimum value, and we have also developed a technique to find these values. In the next section, we further our study of the information we can glean from "nice" functions with the Mean Value Theorem. On a closed interval, we can find the *average rate of change* of a function (as we did at the beginning of Chapter 2). We will see that differentiable functions always have a point at which their *instantaneous* rate of change is same as the *average* rate of change. This is surprisingly useful, as we'll see.

(a)

x	$f(x)$
-1	0
0	1
1	0

(b)
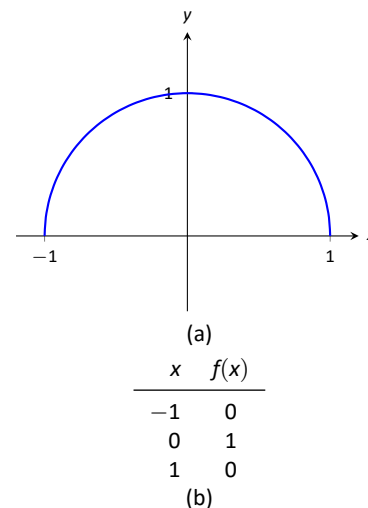
Figure 3.9: A graph and table of extrema of $f(x) = \sqrt{1 - x^2}$ on $[-1, 1]$ as in Example 7.

Note: We implicitly found the derivative of $x^2 + y^2 = 1$, the unit circle, in Example 2.6.5 as $\frac{dy}{dx} = -x/y$. In Example 7, half of the unit circle is given as $y = f(x) = \sqrt{1 - x^2}$. We found $f'(x) = \frac{-x}{\sqrt{1-x^2}}$. Recognize that the denominator of this fraction is y; that is, we again found $f'(x) = \frac{dy}{dx} = -x/y$.

Notes:

Exercises 3.1

Terms and Concepts

1. Describe what an "extreme value" of a function is in your own words.
2. Sketch the graph of a function f on $(-1, 1)$ that has both a maximum and minimum value.
3. Describe the difference between absolute and relative maxima in your own words.
4. Sketch the graph of a function f where f has a relative maximum at $x = 1$ and $f'(1)$ is undefined.
5. T/F: If c is a critical value of a function f, then f has either a relative maximum or relative minimum at $x = c$.

Problems

In Exercises 6–7, identify each of the marked points as being an absolute maximum or minimum, a relative maximum or minimum, or none of the above. (A point could be more than one.)

6.

7.

8.
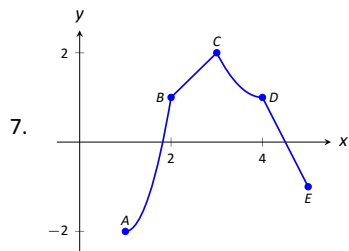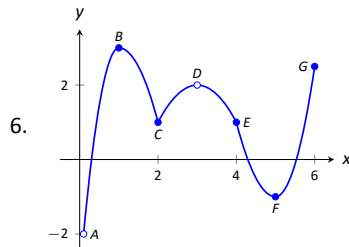
    (a) Sketch the graph of a function that has a local minimum at 3 and is differentiable at 3.

    (b) Sketch the graph of a function that has a local minimum at 3 and is continuous but not differentiable at 3.

    (c) Sketch the graph of a function that has a local minimum at 3 and is not continuous at 3.

In Exercises 9–15, evaluate $f'(x)$ at the points indicated in the graph.

9. $f(x) = \dfrac{2}{x^2 + 1}$
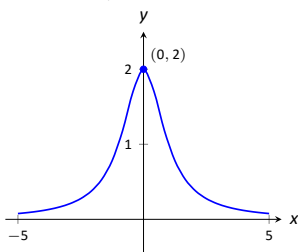
10. $f(x) = x^2\sqrt{6 - x^2}$
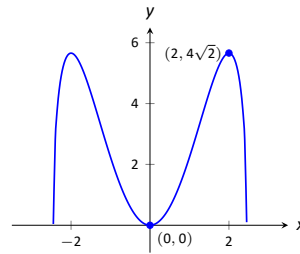
11. $f(x) = \sin x$
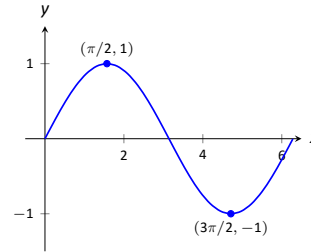
12. $f(x) = x^2\sqrt{4 - x}$
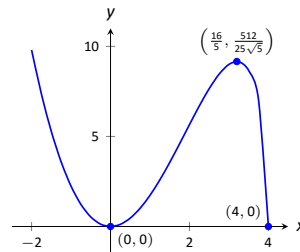
13. $f(x) = \begin{cases} x^2 & x \le 0 \\ x^5 & x > 0 \end{cases}$
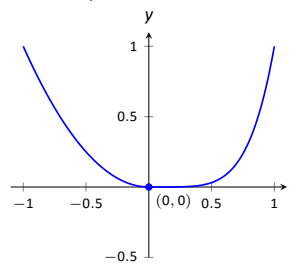
14. $f(x) = \begin{cases} x^2 & x \le 0 \\ x & x > 0 \end{cases}$
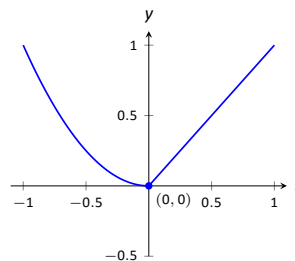
15. $f(x) = \dfrac{(x-2)^{2/3}}{x} + 1$
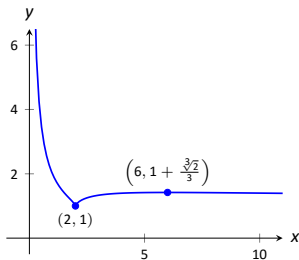
In Exercises 16–25, find the extreme values of the function on the given interval.

16. $f(x) = x^2 + x + 4$ on $[-1, 2]$.

17. $f(x) = x^3 - \dfrac{9}{2}x^2 - 30x + 3$ on $[0, 6]$.

18. $f(x) = 3\sin x$ on $[\pi/4, 2\pi/3]$.

19. $f(x) = x^2\sqrt{4 - x^2}$ on $[-2, 2]$.

20. $f(x) = x + \dfrac{3}{x}$ on $[1, 5]$.

21. $f(x) = \dfrac{x^2}{x^2 + 5}$ on $[-3, 5]$.

22. $f(x) = e^x \cos x$ on $[0, \pi]$.

23. $f(x) = e^x \sin x$ on $[0, \pi]$.

24. $f(x) = \dfrac{\ln x}{x}$ on $[1, 4]$.

25. $f(x) = x^{2/3} - x$ on $[0, 2]$.

26. Show that 4 is a critical number of $f(x) = (x - 4)^3 + 7$ but f does not have a relative extreme value at 4.

27. A cubic function is a polynomial of degree 3; that is, it has the form $ax^3 + bx^2 + cx + d$, where $a \neq 0$.

    (a) Show that a cubic function can have 2, 1, or 0 critical numbers. Give examples and sketches to illustrate the 3 possibilities.

    (b) How many local extreme values can a cubic function have?

28. Suppose that a and b are positive numbers. Find the extreme values of $f(x) = x^a(1 - x)^b$ on $[0, 1]$.

Review

29. Find $\dfrac{dy}{dx}$, where $x^2y - y^2x = 1$.

30. Find the equation of the line tangent to the graph of $x^2 + y^2 + xy = 7$ at the point $(1, 2)$.

31. Let $f(x) = x^3 + x$.

Evaluate $\lim\limits_{s \to 0} \dfrac{f(x + s) - f(x)}{s}$.

3.2   The Mean Value Theorem

We motivate this section with the following question: Suppose you leave your house and drive to your friend's house in a city 100 miles away, completing the trip in two hours. Is there necessarily a moment during the trip when you are going 50 miles per hour?

In answering this question, it is clear that the *average* speed for the entire trip is 50 mph (i.e. 100 miles in 2 hours), but the question is whether or not your *instantaneous* speed is ever exactly 50 mph. More simply, does your speedometer ever read exactly 50 mph?. Figure 3.10 shows a graphical interpretation of this question. The answer, under some very reasonable assumptions, is "yes."

Let's now see why this situation is in a calculus text by translating it into mathematical symbols.

First assume that the function $y = f(t)$ gives the distance (in miles) traveled from your home at time t (in hours) where $0 \leq t \leq 2$. In particular, this gives $f(0) = 0$ and $f(2) = 100$. The slope of the secant line (average velocity) connecting the starting and ending points $(0, f(0))$ and $(2, f(2))$ is therefore

$$\frac{\varDelta f}{\varDelta t} = \frac{f(2) - f(0)}{2 - 0} = \frac{100 - 0}{2} = 50 \, \text{mph}.$$

The slope at any point on the graph itself (instantaneous velocity) is given by the derivative $f'(t)$. So, since the answer to the question above is "yes," this means that at some time during the trip, the derivative takes on the value of 50 mph. Symbolically,

$$f'(c) = \frac{f(2) - f(0)}{2 - 0} = 50$$

for some time $0 \leq c \leq 2$.

How about more generally? Given any function $y = f(x)$ and a range $a \leq x \leq b$ does the value of the derivative at some point between a and b have to match the slope of the secant line connecting the points $(a, f(a))$ and $(b, f(b))$? Or equivalently, does the equation $f'(c) = \frac{f(b) - f(a)}{b - a}$ have to hold for some $a < c < b$?

Let's look at two functions in an example.

Example 1   **Comparing average and instantaneous rates of change**
Consider functions

$$f_1(x) = \frac{1}{x^2} \quad \text{and} \quad f_2(x) = |x|$$

with $a = -1$ and $b = 1$ as shown in Figure 3.11(a) and (b), respectively. Both functions have a value of 1 at a and b. Therefore the slope of the secant line
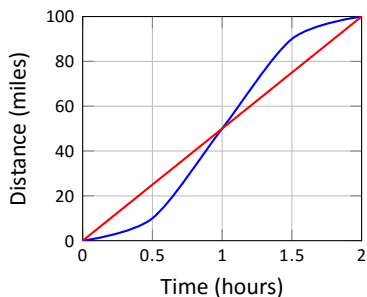
Figure 3.10: Distance traveled as a function of time. Must there be a tangent line parallel to the average slope?
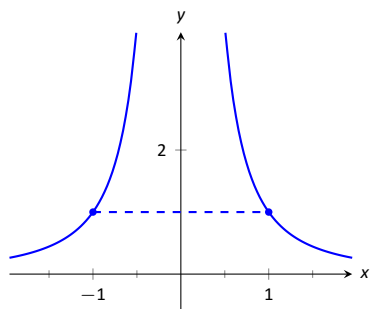
(a)
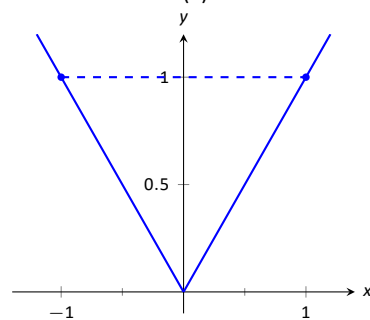
(b)

Figure 3.11: A graph of $f_1(x) = 1/x^2$ and $f_2(x) = |x|$ in Example 1.

Notes:

connecting the end points is 0 in each case. But if you look at the plots of each, you can see that there are no points on either graph where the tangent lines have slope zero. Therefore we have found that there is no c in $[-1, 1]$ such that

$$f'(c) = \frac{f(1) - f(-1)}{1 - (-1)} = 0.$$

So what went "wrong"? It may not be surprising to find that the discontinuity of f_1 and the corner of f_2 play a role. If our functions had been continuous and differentiable, would we have been able to find that special value c? This is our motivation for the following theorem.

> **Theorem 24   The Mean Value Theorem of Differentiation**
>
> Let $y = f(x)$ be a continuous function on the closed interval $[a, b]$ and differentiable on the open interval (a, b). There exists a value c, $a < c < b$, such that
> $$f'(c) = \frac{f(b) - f(a)}{b - a}.$$
> That is, there is a value c in (a, b) where the instantaneous rate of change of f at c is equal to the average rate of change of f on $[a, b]$.
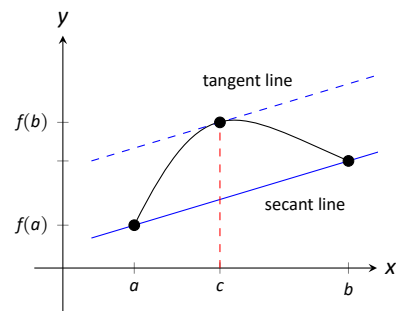
Figure 3.12: A graph of illustrating the Mean Value Theorem of Differentiation

Note that the reasons that the functions in Example 1 fail are indeed that f_1 has a discontinuity on the interval $[-1, 1]$ and f_2 is not differentiable at the origin.

We will give a proof of the Mean Value Theorem below. To do so, we use Rolle's Theorem, stated here.

> **Theorem 25   Rolle's Theorem**
>
> Let f be continuous on $[a, b]$ and differentiable on (a, b), where $f(a) = f(b)$. There is some c in (a, b) such that $f'(c) = 0$.

Consider Figure 3.13 where the graph of a function f is given, where $f(a) = f(b)$. It should make intuitive sense that if f is differentiable (and hence, continuous) that there would be a value c in (a, b) where $f'(c) = 0$; that is, there would be a relative maximum or minimum of f in (a, b). Rolle's Theorem guarantees at least one; there may be more.
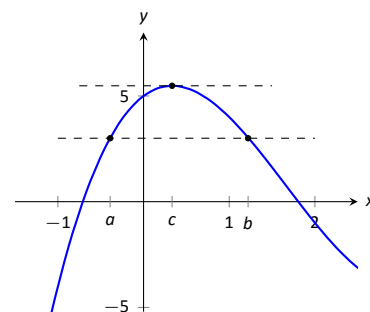
Figure 3.13: A graph of $f(x) = x^3 - 5x^2 + 3x + 5$, where $f(a) = f(b)$. Note the existence of c, where $a < c < b$, where $f'(c) = 0$.

Notes:

Rolle's Theorem is really just a special case of the Mean Value Theorem. If $f(a) = f(b)$, then the *average* rate of change on (a, b) is 0, and the theorem guarantees some c where $f'(c) = 0$. We will prove Rolle's Theorem, then use it to prove the Mean Value Theorem.

Proof of Rolle's Theorem

Let f be differentiable on (a, b) where $f(a) = f(b)$. We consider two cases.

Case 1: Consider the case when f is constant on $[a, b]$; that is, $f(x) = f(a) = f(b)$ for all x in $[a, b]$. Then $f'(x) = 0$ for all x in $[a, b]$, showing there is at least one value c in (a, b) where $f'(c) = 0$.

Case 2: Now assume that f is not constant on $[a, b]$. The Extreme Value Theorem guarantees that f has a maximal and minimal value on $[a, b]$, found either at the endpoints or at a critical value in (a, b). Since $f(a) = f(b)$ and f is not constant, it is clear that the maximum and minimum cannot *both* be found at the endpoints. Assume, without loss of generality, that the maximum of f is not found at the endpoints. Therefore there is a c in (a, b) such that $f(c)$ is the maximum value of f. By Theorem 23, c must be a critical number of f; since f is differentiable, we have that $f'(c) = 0$, completing the proof of the theorem.   □

Example 2   Exactly One Root

Show that $f(x) = 8x^7 + x^3 + 3x + 2$ has exactly one real root.

SOLUTION   We'll do this in two steps. The first step is to use the Intermediate Value Theorem to show that there is at least one root. The second step is to use Rolle's Theorem to show that there is at most one root. (Because f is a polynomial, it is continuous and differentiable, so both of these theorems apply.)

We can apply the Intermediate Value Theorem on the interval $[-1, 0]$. Since $f(-1) = -10 < 0 < f(0) = 2$, the Intermediate Value Theorem tells us that there is at least one place in $[-1, 0]$ where $f(x) = 0$. This means that there is at least one root, but there may be more in the interval (and there may be more outside the interval where we haven't even looked).

We will now use Rolle's Theorem to show that f has at most one root. Suppose for this paragraph that f had two (or more) roots. Then by Rolle's Theorem, there is some c in between the roots so that $0 = f'(c) = 56x^6 + 3x^2 + 3$. But this cannot happen, since f' is always at least 3.

Therefore, f has at most one root. Combining this with "there is at least one root", we see that f has exactly one root. (Notice that because both the Intermediate Value Theorem and Rolle's Theorem are existential theorems, we don't know what the root is, only that it must exist.)

Notes:

We will now use Rolle's Theorem to prove the Mean Value Theorem.

Proof of the Mean Value Theorem

Define the function

$$g(x) = f(x) - \frac{f(b) - f(a)}{b - a} x.$$

We know g is differentiable on (a, b) and continuous on $[a, b]$ since f is. We also see that

$$g(b) - g(a) = f(b) - \frac{f(b) - f(a)}{b - a} b - f(a) + \frac{f(b) - f(a)}{b - a} a$$

$$= \left(f(b) - f(a) \right) - \frac{f(b) - f(a)}{b - a} (b - a) = 0$$

which shows that $g(a) = g(b)$. We can then apply Rolle's theorem to guarantee the existence of $c \in (a, b)$ such that $g'(c) = 0$. But note that

$$0 = g'(c) = f'(c) - \frac{f(b) - f(a)}{b - a} \; ;$$

hence

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

which is what we sought to prove. □

Going back to the very beginning of the section, we see that the only assumption we would need about our distance function $f(t)$ is that it be continuous and differentiable for t from 0 to 2 hours (both reasonable assumptions). By the Mean Value Theorem, we are guaranteed a time during the trip where our instantaneous speed is 50 mph. This fact is used in practice. Some law enforcement agencies monitor traffic speeds while in aircraft. They do not measure speed with radar, but rather by timing individual cars as they pass over lines painted on the highway whose distances apart are known. The officer is able to measure the *average* speed of a car between the painted lines; if that average speed is greater than the posted speed limit, the officer is assured that the driver exceeded the speed limit at some time.

Note that the Mean Value Theorem is an *existence* theorem. It states that a special value *c exists*, but it does not give any indication about how to find it. It turns out that when we need the Mean Value Theorem, existence is all we need.

Notes:

Watch the video:
The Mean Value Theorem at
`https://youtu.be/xYOrYLq3fE0`

Example 3   **Using the Mean Value Theorem**

Consider $f(x) = x^3 + 5x + 5$ on $[-3, 3]$. Find c in $[-3, 3]$ that satisfies the Mean Value Theorem.

SOLUTION   The average rate of change of f on $[-3, 3]$ is:

$$\frac{f(3) - f(-3)}{3 - (-3)} = \frac{84}{6} = 14.$$

We want to find c such that $f'(c) = 14$. We find $f'(x) = 3x^2 + 5$. We set this equal to 14 and solve for x.

$$f'(x) = 14$$
$$3x^2 + 5 = 14$$
$$x^2 = 3$$
$$x = \pm\sqrt{3} \approx \pm 1.732$$

We have found 2 values c in $[-3, 3]$ where the instantaneous rate of change is equal to the average rate of change; the Mean Value Theorem guaranteed at least one. In Figure 3.14 f is graphed with a dashed line representing the average rate of change; the lines tangent to f at $x = \pm\sqrt{3}$ are also given. Note how these lines are parallel (i.e., have the same slope) as the dashed line.

While the Mean Value Theorem has practical use (for instance, the speed monitoring application mentioned before), it is mostly used to advance other theory. We will use it in the next section to relate the shape of a graph to its derivative.
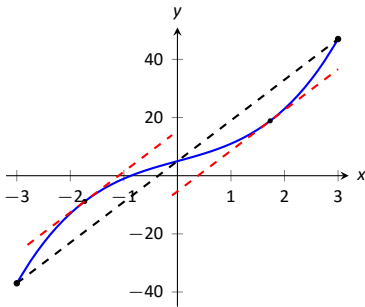
Figure 3.14: Demonstrating the Mean Value Theorem in Example 3.

Notes:

Exercises 3.2

Terms and Concepts

1. Explain in your own words what the Mean Value Theorem states.
2. Explain in your own words what Rolle's Theorem states.

Problems

In Exercises 3–10, a function $f(x)$ and interval $[a, b]$ are given. Check if Rolle's Theorem can be applied to f on $[a, b]$; if so, find c in $[a, b]$ such that $f'(c) = 0$.

3. $f(x) = 6$ on $[-1, 1]$.

4. $f(x) = 6x$ on $[-1, 1]$.

5. $f(x) = x^2 + x - 6$ on $[-3, 2]$.

6. $f(x) = x^2 + x - 2$ on $[-3, 2]$.

7. $f(x) = x^2 + x$ on $[-2, 2]$.

8. $f(x) = \sin x$ on $[\pi/6, 5\pi/6]$.

9. $f(x) = \cos x$ on $[0, \pi]$.

10. $f(x) = \dfrac{1}{x^2 - 2x + 1}$ on $[0, 2]$.

In Exercises 11–19, a function $f(x)$ and interval $[a, b]$ are given. Check if the Mean Value Theorem can be applied to f on $[a, b]$; if so, find a value c in $[a, b]$ guaranteed by the Mean Value Theorem.

11. $f(x) = x^2 + 3x - 1$ on $[-2, 2]$.

12. $f(x) = 5x^2 - 6x + 8$ on $[0, 5]$.

13. $f(x) = \sqrt{9 - x^2}$ on $[0, 3]$.

14. $f(x) = \sqrt{25 - x}$ on $[0, 9]$.

15. $f(x) = \dfrac{x^2 - 9}{x^2 - 1}$ on $[0, 2]$.

16. $f(x) = \ln x$ on $[1, 5]$.

17. $f(x) = \tan x$ on $[-\pi/4, \pi/4]$.

18. $f(x) = x^3 - 2x^2 + x + 1$ on $[-2, 2]$.

19. $f(x) = 2x^3 - 5x^2 + 6x + 1$ on $[-5, 2]$.

20. Suppose that f is continuous on $[1, 4]$ and differentiable on $(1, 4)$. If $f(1) = 10$ and $f'(x) \geq 2$ for $1 \leq x \leq 4$, how small can $f(4)$ possibly be?

21. Does there exist a function f such that $f(0) = -1, f(2) = 4$, and $f'(x) \leq 2$ for all x?

22. Show that the equation $1 + 2x + x^3 + 4x^5 = 0$ has exactly one real root.

23. Show that a polynomial of degree 3 has at most 3 real roots.

24.
   (a) Suppose that f is differentiable everywhere and has 2 roots. Show that f' has at least one real root.

   (b) Suppose that f is twice differentiable everywhere and has 3 roots. Show that f'' has at least one real root.

25. Let p, q, and r be constants, and define $f(x) = px^2 + qx + r$. Show that the Mean Value Theorem applied to f for the interval $[a, b]$ is always satisfied at the midpoint of the interval.

Review

26. Find the extreme values of $f(x) = x^2 - 3x + 9$ on $[-2, 5]$.

27. Describe the critical points of $f(x) = \cos x$.

28. Describe the critical points of $f(x) = \tan x$.

3.3 Increasing and Decreasing Functions

Our study of "nice" functions f in this chapter has so far focused on individual points: points where f is maximal/minimal, points where $f'(x) = 0$ or f' does not exist, and points c where $f'(c)$ is the average rate of change of f on some interval.

In this section we begin to study how functions behave *between* special points; we begin studying in more detail the shape of their graphs.

We start with an intuitive concept. Given the graph in Figure 3.15, where would you say the function is *increasing*? *Decreasing*? Even though we have not defined these terms mathematically, one likely answered that f is increasing when $x > 1$ and decreasing when $x < 1$. We formally define these terms here.
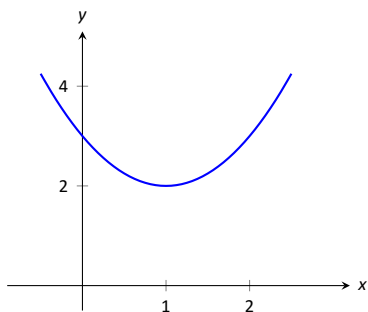
Figure 3.15: A graph of a function f used to illustrate the concepts of *increasing* and *decreasing*.

Definition 17   Increasing and Decreasing Functions

Let f be a function defined on an interval I.

1. f is **increasing** on I if for every $a < b$ in I, $f(a) \leq f(b)$.

2. f is **decreasing** on I if for every $a < b$ in I, $f(a) \geq f(b)$.

A function is **strictly increasing** when $a < b$ in I implies $f(a) < f(b)$, with a similar definition holding for **strictly decreasing**.

Informally, a function is increasing if as x gets larger (i.e., looking left to right) $f(x)$ gets larger.

Our interest lies in finding intervals in the domain of f on which f is either increasing or decreasing. Such information should seem useful. For instance, if f describes the speed of an object, we might want to know when the speed was increasing or decreasing (i.e., when the object was accelerating vs. decelerating). If f describes the population of a city, we should be interested in when the population is growing or declining.

To find such intervals, we again consider secant lines. Let f be an increasing, differentiable function on an open interval I, such as the one shown in Figure 3.16, and let $a < b$ be given in I. The secant line on the graph of f from $x = a$ to $x = b$ is drawn; it has a slope of $(f(b) - f(a))/(b - a)$. But note:
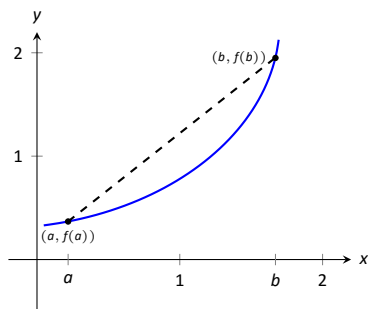
Figure 3.16: Examining the secant line of an increasing function.

$$\frac{f(b) - f(a)}{b - a} \Rightarrow \frac{\text{numerator} > 0}{\text{denominator} > 0} \Rightarrow \frac{\text{slope of the}}{\text{secant line} > 0} \Rightarrow \begin{array}{c} \text{Average rate of} \\ \text{change of } f \text{ on} \\ [a, b] \text{ is} > 0. \end{array}$$

We have shown mathematically what may have already been obvious: when f is increasing, its secant lines will have a positive slope. Now recall the Mean

---

Notes:

Value Theorem guarantees that there is a number c, where $a < c < b$, such that

$$f'(c) = \frac{f(b) - f(a)}{b - a} > 0.$$

By considering all such secant lines in I, we strongly imply that $f'(x) \geq 0$ on I. A similar statement can be made for decreasing functions.

Our above logic can be summarized as "If f is increasing, then f' is probably positive." Theorem 26 below turns this around by stating "If f' is postive, then f is increasing." This leads us to a method for finding when functions are increasing and decreasing.

Theorem 26   Test For Increasing/Decreasing Functions

Let f be a continuous function on $[a, b]$ and differentiable on (a, b).

1. If $f'(c) > 0$ for all c in (a, b), then f is increasing on $[a, b]$.

2. If $f'(c) < 0$ for all c in (a, b), then f is decreasing on $[a, b]$.

3. If $f'(c) = 0$ for all c in (a, b), then f is constant on $[a, b]$.

Note: Theorem 26 (parts 1 and 2) also holds if $f'(c) = 0$ for a finite number of values of c in I.

Let a and b be in I where $f'(a) > 0$ and $f'(b) < 0$. If f' is continuous, then we can use the Intermediate Value Theorem. Even if f' isn't continuous, Darboux's Theorem (which is too advanced to prove here) shows that f' has the Intermediate Value property. Either way, there must be some value c between a and b where $f'(c) = 0$. This leads us to the following method for finding intervals on which a function is increasing or decreasing.

Key Idea 5   Finding Intervals on which f is Increasing or Decreasing

Let f be a differentiable function on an interval I. To find intervals on which f is increasing and decreasing:

1. Find the critical values of f. That is, find all c in I where $f'(c) = 0$ or f' is not defined.

2. Use the critical values to divide I into subintervals.

3. Pick any point p in each subinterval, and find the sign of $f'(p)$.

   (a) If $f'(p) > 0$, then f is increasing on that subinterval.

   (b) If $f'(p) < 0$, then f is decreasing on that subinterval.

Notes:

Watch the video:
Finding Intervals of Increase/Decrease Local Max/Mins at
`https://youtu.be/-W4dOqFzyQY`

We demonstrate using Key Idea 5 in the following example.

Example 1 **Finding intervals of increasing/decreasing**
Let $f(x) = x^3 + x^2 - x + 1$. Find intervals on which f is increasing or decreasing.

SOLUTION Using Key Idea 5, we first find the critical values of f. We have $f'(x) = 3x^2 + 2x - 1 = (3x - 1)(x + 1)$, so $f'(x) = 0$ when $x = -1$ and when $x = 1/3$. We see that f' is never undefined.

Since an interval was not specified for us to consider, we consider the entire domain of f which is $(-\infty, \infty)$. We thus break the whole real line into three subintervals based on the two critical values we just found: $(-\infty, -1)$, $(-1, 1/3)$ and $(1/3, \infty)$. This is shown in the following sign chart.
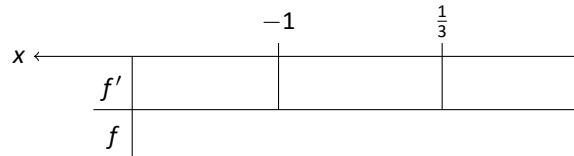
We now pick a value p in each subinterval and find the sign of $f'(p)$. All we care about is the sign, so we do not actually have to fully compute $f'(p)$; pick "nice" values that make this simple.

Subinterval 1, $(-\infty, -1)$: We (arbitrarily) pick $p = -2$. We can compute $f'(-2)$ directly: $f'(-2) = 3(-2)^2 + 2(-2) - 1 = 7 > 0$. We conclude that f is increasing on $(-\infty, -1)$.

Note we can arrive at the same conclusion without computation. For instance, we could choose $p = -100$. The first term in $f'(-100)$, i.e., $3(-100)^2$ is clearly positive and very large. The other terms are small in comparison, so we know $f'(-100) > 0$. All we need is the sign.

Subinterval 2, $(-1, 1/3)$: We pick $p = 0$ since that value seems easy to deal with and note that $f'(0) = -1 < 0$. We conclude f is decreasing on $(-1, 1/3)$.

Notes:

Subinterval 3, $(1/3, \infty)$: Pick an arbitrarily large value for $p > 1/3$ and note that $f'(p) = 3p^2 + 2p - 1 > 0$. We conclude that f is increasing on $(1/3, \infty)$ and use all of our information to complete our sign chart.

We can verify our calculations by considering Figure 3.17, where f is graphed. The graph also presents f'; note how $f' > 0$ when f is increasing and $f' < 0$ when f is decreasing.

One is justified in wondering why so much work is done when the graph seems to make the intervals very clear. We give three reasons why the above work is worthwhile.

First, the points at which f switches from increasing to decreasing are not precisely known given a graph. The graph shows us something significant happens near $x = -1$ and $x = 0.3$, but we cannot determine exactly where from the graph.

One could argue that just finding critical values is important; once we know the significant points are $x = -1$ and $x = 1/3$, the graph shows the increasing/decreasing traits just fine. That is true. However, the technique prescribed here helps reinforce the relationship between increasing/decreasing and the sign of f'. Once mastery of this concept (and several others) is obtained, one finds that either (a) just the critical points are computed and the graph shows all else that is desired, or (b) a graph is never produced, because determining increasing/decreasing using f' is straightforward and the graph is unnecessary. So our second reason why the above work is worthwhile is this: once mastery of a subject is gained, one has *options* for finding needed information. We are working to develop mastery.

Finally, our third reason: many problems we face "in the real world" are very complex. Solutions are tractable only through the use of computers to do many calculations for us. Computers do not solve problems "on their own," however; they need to be taught (i.e., *programmed*) to do the right things. It would be beneficial to give a function to a computer and have it return maximum and minimum values, intervals on which the function is increasing and decreasing, the locations of relative maxima, etc. The work that we are doing here is easily programmable. It is hard to teach a computer to "look at the graph and see if it is going up or down." It is easy to teach a computer to "determine if a number is greater than or less than 0."
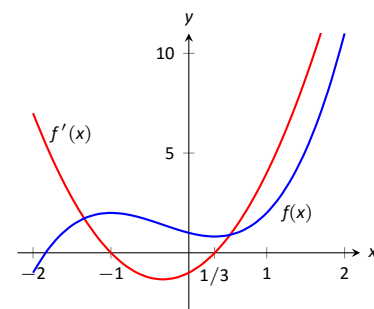
Figure 3.17: A graph of $f(x)$ in Example 1, showing where f is increasing and decreasing.

Notes:

In Section 3.1 we learned the definition of relative maxima and minima and found that they occur at critical points. We are now learning that functions can switch from increasing to decreasing (and vice–versa) at critical points. This new understanding of increasing and decreasing creates a great method of determining whether a critical point corresponds to a maximum, minimum, or neither. Imagine a function increasing until a critical point at $x = c$, after which it decreases. A quick sketch helps confirm that $f(c)$ must be a relative maximum. A similar statement can be made for relative minimums, see Figure 3.18. We formalize this concept in a theorem.
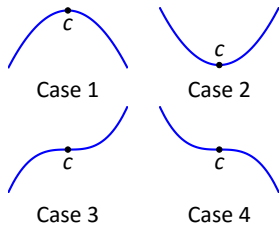
Figure 3.18: The four cases of Theorem 27

Theorem 27 First Derivative Test

Let f be differentiable on I and let c be a critical number in I.

1. If the sign of f' switches from positive to negative at c, then $f(c)$ is a relative maximum of f.

2. If the sign of f' switches from negative to positive at c, then $f(c)$ is a relative minimum of f.

3. If the sign of f' is positive before and after c, then $f(c)$ is not a relative extrema of f.

4. If the sign of f' is negative before and after c, then $f(c)$ is not a relative extrema of f.

Example 2 Using the First Derivative Test

Find the intervals on which f is increasing and decreasing, and use the First Derivative Test to determine the relative extrema of f, where

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

SOLUTION We start by noting the domain of f: $(-\infty, 1) \cup (1, \infty)$. Key Idea 5 describes how to find intervals where f is increasing and decreasing *when the domain of f is an interval.* Since the domain of f in this example is the union of two intervals, we apply the techniques of Key Idea 5 to both intervals of the domain of f.

Since f is not defined at $x = 1$, the increasing/decreasing nature of f could switch at this value. We do not formally consider $x = 1$ to be a critical value of f, but we will include it in our list of critical values that we find next.

Notes:

Using the Quotient Rule, we find

$$f'(x) = \frac{x^2 - 2x - 3}{(x-1)^2}.$$

We need to find the critical values of f; we want to know when $f'(x) = 0$ and when f' is not defined. That latter is straightforward: when the denominator of $f'(x)$ is 0, f' is undefined. That occurs when $x = 1$, which we've already recognized as an important value.

$f'(x) = 0$ when the numerator of $f'(x)$ is 0. That occurs when $x^2 - 2x - 3 = (x-3)(x+1) = 0$; i.e., when $x = -1, 3$.

We have found that f has two critical numbers, $x = -1, 3$, and at $x = 1$ something important might also happen. These three numbers divide the real number line into 4 subintervals:

$$(-\infty, -1), \quad (-1, 1), \quad (1, 3) \quad \text{and} \quad (3, \infty).$$

Pick a number p from each subinterval and test the sign of f' at p to determine whether f is increasing or decreasing on that interval. Again, we do well to avoid complicated computations; notice that the denominator of f' is *always* positive so we can ignore it during our work.

Interval 1, $(-\infty, -1)$: Choosing a very small number (i.e., a negative number with a large magnitude) p returns $p^2 - 2p - 3$ in the numerator of f'; that will be positive. Hence f is increasing on $(-\infty, -1)$.

Interval 2, $(-1, 1)$: Choosing 0 seems simple: $f'(0) = -3 < 0$. We conclude f is decreasing on $(-1, 1)$.

Interval 3, $(1, 3)$: Choosing 2 seems simple: $f'(2) = -3 < 0$. Again, f is decreasing.

Interval 4, $(3, \infty)$: Choosing a very large number p from this subinterval will give a positive numerator and (of course) a positive denominator. So f is increasing on $(3, \infty)$.

In summary, f is increasing on $(-\infty, -1)$ and $(3, \infty)$ and is decreasing on $(-1, 1)$ and $(1, 3)$. Since at $x = -1$, the sign of f' switched from positive to negative, Theorem 27 states that $f(-1)$ is a relative maximum of f. At $x = 3$, the sign of f' switched from negative to positive, meaning $f(3)$ is a relative minimum. At $x = 1$, f is not defined, so there is no relative extrema at $x = 1$.

Notes:

Note: In our sign charts, we will use "U" to indicate that something is undefined.
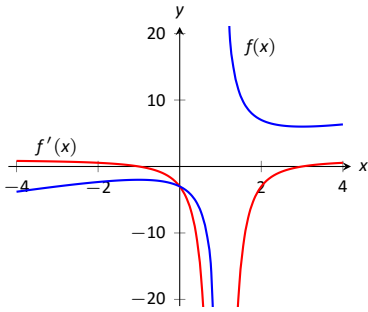
Figure 3.19: A graph of $f(x)$ in Example 2, showing where f is increasing and decreasing.

This is summarized in the number line shown above. Also, Figure 3.19 shows a graph of f, confirming our calculations. This figure also shows f', again demonstrating that f is increasing when $f' > 0$ and decreasing when $f' < 0$.

One is often tempted to think that functions always alternate "increasing, decreasing, increasing, decreasing,..." around critical values. Our previous example demonstrated that this is not always the case. While $x = 1$ was not technically a critical value, it was an important value we needed to consider. We found that f was decreasing on "both sides of $x = 1$."

Example 3   **Using the First Derivative Test**
Find the intervals on which $f(x) = x^{8/3} - 4x^{2/3}$ is increasing and decreasing and identify the relative extrema.

SOLUTION   We again start with taking derivatives. Since we know we want to solve $f'(x) = 0$, we will do some algebra after taking derivatives.

$$f(x) = x^{\frac{8}{3}} - 4x^{\frac{2}{3}}$$
$$f'(x) = \frac{8}{3}x^{\frac{5}{3}} - \frac{8}{3}x^{-\frac{1}{3}}$$
$$= \frac{8}{3}x^{-\frac{1}{3}}\left(x^{\frac{6}{3}} - 1\right)$$
$$= \frac{8}{3}x^{-\frac{1}{3}}(x^2 - 1)$$
$$= \frac{8}{3}x^{-\frac{1}{3}}(x - 1)(x + 1).$$

This derivation of f' shows that $f'(x) = 0$ when $x = \pm 1$ and f' is not defined when $x = 0$. Thus we have 3 critical values, breaking the number line into 4 subintervals: $(-\infty, -1)$, $(-1, 0)$, $(0, 1)$, and (1∞).

Interval 1, $(\infty, -1)$: We choose $p = -2$; we can easily verify that $f'(-2) < 0$. So f is decreasing on $(-\infty, -1)$.

Interval 2, $(-1, 0)$: Choose $p = -1/2$. Once more we practice finding the sign of $f'(p)$ without computing an actual value. We have $f'(p) = (8/3)p^{-1/3}(p -$

Notes:

$1)(p + 1)$; find the sign of each of the three terms.

$$f'(p) = \frac{8}{3} \cdot \underbrace{p^{-\frac{1}{3}}}_{<0} \cdot \underbrace{(p-1)}_{<0} \underbrace{(p+1)}_{>0}.$$

We have a "negative × negative × positive" giving a positive number; f is increasing on $(-1, 0)$.

Interval 3, $(0, 1)$: We do a similar sign analysis as before, using p in $(0, 1)$.

$$f'(p) = \frac{8}{3} \cdot \underbrace{p^{-\frac{1}{3}}}_{>0} \cdot \underbrace{(p-1)}_{<0} \underbrace{(p+1)}_{>0}.$$

We have 2 positive factors and one negative factor; $f'(p) < 0$ and so f is decreasing on $(0, 1)$.

Interval 4, $(1, \infty)$: Similar work to that done for the other three intervals shows that $f'(x) > 0$ on $(1, \infty)$, so f is increasing on this interval. We can now put all this information into a chart.

	-1		0		1	
x						
f'	$-$		$+$		$-$	$+$
f	decr	min	incr	max	decr	min	incr

We conclude by stating that f is increasing on $(-1, 0)$ and $(1, \infty)$ and decreasing on $(-\infty, -1)$ and $(0, 1)$. The sign of f' changes from negative to positive around $x = -1$ and $x = 1$, meaning by Theorem 27 that $f(-1)$ and $f(1)$ are relative minima of f. As the sign of f' changes from positive to negative at $x = 0$, we have a relative maximum at $f(0)$. Figure 3.20 shows a graph of f, confirming our result. We also graph f', highlighting once more that f is increasing when $f' > 0$ and is decreasing when $f' < 0$.

We examine one more example.

Example 4   **Using the First Derivative Test with Trigonometry**

Find the intervals on which $f(\theta) = \cos\theta - \cos^2\theta$ is increasing and decreasing and find the relative extrema on the interval $[0, 2\pi]$.

SOLUTION   We see that $f'(\theta) = -\sin\theta + 2\cos\theta\sin\theta = \sin\theta(2\cos\theta - 1)$. Therefore, $f'(\theta) = 0$ when $\theta = 0, \frac{\pi}{3}, \pi, \frac{5\pi}{3}, 2\pi$. This breaks our number line into four intervals.
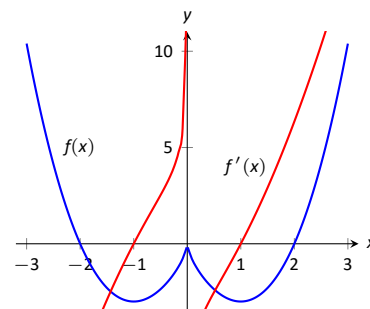
Figure 3.20: A graph of $f(x)$ in Example 3, showing where f is increasing and decreasing.

Notes:

Interval 1, $(0, \frac{\pi}{3})$: We choose $\theta = \frac{\pi}{6}$, and see that $\sin\theta$ and $2\cos\theta - 1$ are both positive. Therefore, $f' > 0$.

Interval 2, $(\frac{\pi}{3}, \pi)$: When $\theta = \frac{\pi}{2}$, $\sin\theta$ is positive, but $2\cos\theta - 1$ is negative. Therefore, $f' < 0$.

Interval 3, $(\pi, \frac{5\pi}{3})$: When $\theta = \frac{3\pi}{2}$, $\sin\theta$ and $2\cos\theta - 1$ are both negative. Therefore, $f' > 0$.

Interval 4, $(\frac{5\pi}{3}, 2\pi)$: When $\theta = \frac{5\pi}{6}$, $\sin\theta < 0$ and $2\cos\theta - 1 > 0$ so that $f' < 0$. We summarize this information in a chart.

	0		$\frac{\pi}{3}$		π		$\frac{5\pi}{3}$		2π
x									
f'		$+$		$-$		$+$		$-$	
f		incr	max	decr	min	incr	max	decr	

This means that f is increasing on $(0, \frac{\pi}{3})$ and $(\pi, \frac{5\pi}{3})$ and decreasing on $(\frac{\pi}{3}, \pi)$ and $(\frac{5\pi}{3}, 2\pi)$, so that the relative maxima are $f(\frac{\pi}{3})$ and $f(\frac{5\pi}{3})$ and the relative minimum is $f(\pi)$. (The values $f(0)$ and $f(2\pi)$ would also be relative minima, but relative extrema are not allowed to occur at the endpoints of an interval.)

We have seen how the first derivative of a function helps determine when the function is going "up" or "down." In the next section, we will see how the second derivative helps determine how the graph of a function curves.

Notes:

Exercises 3.3

Terms and Concepts

1. In your own words describe what it means for a function to be increasing.

2. What does a decreasing function "look like"?

3. Sketch a graph of a function on $[0, 2]$ that is increasing but not strictly increasing.

4. Give an example of a function describing a situation where it is "bad" to be increasing and "good" to be decreasing.

5. A function f has derivative $f'(x) = (\sin x + 2)e^{x^2+1}$, where $f'(x) > 1$ for all x. Is f increasing, decreasing, or can we not tell from the given information?

Problems

6. Given the graph of f, identify the intervals of increasing and decreasing as well as the x coordinates of the relative extrema.

7. Given the graph of f, identify the intervals of increasing and decreasing as well as the x coordinates of the relative extrema.
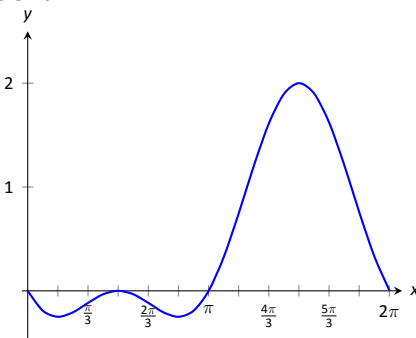
8. Given the graph of f', identify the intervals of increasing and decreasing as well as the x coordinates of the relative extrema.

9. Given the graph of f', identify the intervals of increasing and decreasing as well as the x coordinates of the relative extrema.
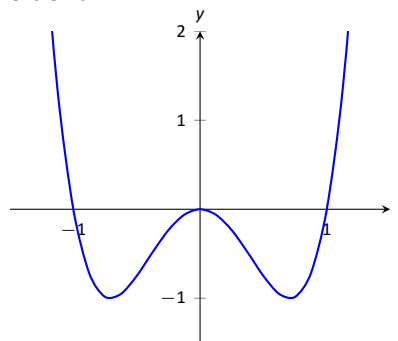
In Exercises 10–17, a function $f(x)$ is given.

(a) Compute $f'(x)$.

(b) Graph f and f' on the same axes (using technology is permitted) and verify Theorem 26.

10. $f(x) = 2x + 3$

11. $f(x) = x^2 - 3x + 5$

12. $f(x) = \cos x$

13. $f(x) = \tan x$

14. $f(x) = x^3 - 5x^2 + 7x - 1$

15. $f(x) = 2x^3 - x^2 + x - 1$

16. $f(x) = x^4 - 5x^2 + 4$

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

In Exercises 18–38, a function $f(x)$ is given.

(a) Give the domain of f.

(b) Find the critical numbers of f.

(c) Create a number line to determine the intervals on which f is increasing and decreasing.

(d) Use the First Derivative Test to determine whether each critical point is a relative maximum, minimum, or neither.

18. $f(x) = x^2 + 2x - 3$

19. $f(x) = x^3 + 3x^2 + 3$

20. $f(x) = 2x^3 + x^2 - x + 3$

21. $f(x) = x^3 - 3x^2 + 3x - 1$

22. $f(x) = \dfrac{1}{x^2 - 2x + 2}$

23. $f(x) = \dfrac{x^2 - 4}{x^2 - 1}$

24. $f(x) = \dfrac{x}{x^2 - 2x - 8}$

25. $f(x) = \dfrac{(x - 2)^{2/3}}{x}$

26. $f(x) = \sin x \cos x$ on $(-\pi, \pi)$.

27. $f(x) = x^5 - 5x$

28. $f(x) = x - 2 \sin x$ on $0 \le x \le 3\pi$

29. $f(x) = \cos^2 x - 2 \sin x$ on $0 \le x \le 2\pi$

30. $f(x) = x\sqrt{x - 3}$

31. $f(x) = (x^2 - 1)^3$

32. $f(x) = x^{1/3}(x + 4)$

33. $f(\theta) = 2 \cos \theta + \cos^2 \theta$ on $[0, 2\pi]$

34. $f(x) = 2\sqrt{x} - 4x^2$

35. $f(x) = 5x^{2/3} - 2x^{5/3}$

36. $f(x) = \frac{1}{2}x^4 - 4x^2 + 3$

37. $f(x) = \sin^3 x$ on $[0, 2\pi]$

38. $f(x) = (x + 1)^5 - 5x - 2$

Review

39. Consider $f(x) = x^2 - 3x + 5$ on $[-1, 2]$; find c guaranteed by the Mean Value Theorem.

40. Consider $f(x) = \sin x$ on $[-\pi/2, \pi/2]$; find c guaranteed by the Mean Value Theorem.

3.4 Concavity and the Second Derivative

Our study of "nice" functions continues. The previous section showed how the first derivative of a function, f', can relay important information about f. We now apply the same technique to f' itself, and learn what this tells us about f.

The key to studying f' is to consider its derivative, namely f'', which is the second derivative of f. When $f'' > 0$, f' is increasing. When $f'' < 0$, f' is decreasing. As with f, f' has relative maxima and minima where $f'' = 0$ or is undefined.

This section explores how knowing information about f'' gives information about f.

Concavity

We begin with a definition, then explore its meaning.

> **Definition 18   Concave Up and Concave Down**
> Let f be differentiable on an interval I. The graph of f is **concave up** on I if f' is increasing. The graph of f is **concave down** on I if f' is decreasing. If f' is constant then the graph of f is said to have **no concavity**.

Geometrically, a function is concave up if its graph lies above its tangent lines, so that it curves upward. A function is concave down if its graph lies below its tangent lines, so that it curves downward.

The graph of a function f is concave up when f' is increasing. That means as one looks at a concave up graph from left to right, the slopes of the tangent lines will be increasing. Consider Figure 3.21(a), where a concave up graph is shown along with some tangent lines. Notice how the tangent line on the left is steep, downward, corresponding to a small value of f'. On the right, the tangent line is steep, upward, corresponding to a large value of f'.

If a function is decreasing and concave up, then its rate of decrease is slowing; it is "leveling off." If the function is increasing and concave up, then the *rate* of increase is increasing. The function is increasing at a faster and faster rate.

Now consider a function which is concave down. We essentially repeat the above paragraphs with slight variation.

The graph of a function f is concave down when f' is decreasing. That means as one looks at a concave down graph from left to right, the slopes of the tangent lines will be decreasing. Consider Figure 3.21(b), where a concave down graph is shown along with some tangent lines. Notice how the tangent line on the left

Note: We often state that "f is concave up" instead of "the graph of f is concave up" for simplicity.

(a)

(b)
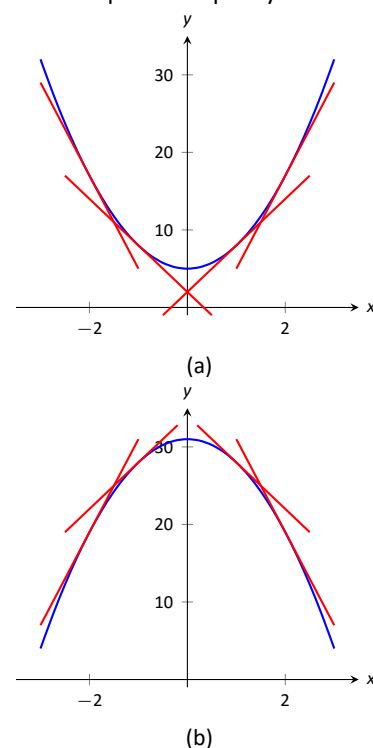
Figure 3.21: A function f with a graph that is (a) concave up and (b) concave down. Notice how the slopes of the tangent lines, when looking from left to right, are (a) increasing and (b) decreasing.

Notes:

is steep, upward, corresponding to a large value of f'. On the right, the tangent line is steep, downward, corresponding to a small value of f'.

If a function is increasing and concave down, then its rate of increase is slowing; it is "leveling off." If the function is decreasing and concave down, then the *rate* of decrease is decreasing. The function is decreasing at a faster and faster rate.

Our definition of concave up and concave down is given in terms of when the first derivative is increasing or decreasing. We can apply the results of the previous section and to find intervals on which a graph is concave up or down. That is, we recognize that f' is increasing when $f'' > 0$, etc.

Note: A mnemonic for remembering what concave up/down means is: "Concave up is like a cup; concave down is like a frown." It is admittedly terrible, but it works.

> **Theorem 28   Test for Concavity**
> Let f be twice differentiable on an interval I. The graph of f is concave up if $f'' > 0$ on I, and is concave down if $f'' < 0$ on I.

If knowing where a graph is concave up/down is important, it makes sense that the places where the graph changes from one to the other is also important. This leads us to a definition.

> **Definition 19   Point of Inflection**
> A **point of inflection** is a point on the graph of f at which the concavity of f changes.

Figure 3.22 shows a graph of a function with inflection points labeled.

If the concavity of f changes at a point $(c, f(c))$, then f' is changing from increasing to decreasing (or, decreasing to increasing) at $x = c$. That means that the sign of f'' is changing from positive to negative (or, negative to positive) at $x = c$. This leads to the following theorem.

> **Theorem 29   Points of Inflection**
> If $(c, f(c))$ is a point of inflection on the graph of f, then either $f'' = 0$ or f'' is not defined at c.

Figure 3.22: A graph of a function with its inflection points marked. The intervals where concave up/down are also indicated.

We have identified the concepts of concavity and points of inflection. It is now time to practice using these concepts; given a function, we should be able to find its points of inflection and identify intervals on which it is concave up or down. We do so in the following examples.

Notes:

Watch the video:
Finding Local Maximums/Minimums — Second Derivative Test at
`https://youtu.be/QtXCIxB6kW8`

Example 1 Finding intervals of concave up/down, inflection points

Let $f(x) = x^3 - 3x + 1$. Find the inflection points of f and the intervals on which it is concave up/down.

SOLUTION We start by finding $f'(x) = 3x^2 - 3$ and $f''(x) = 6x$. To find the inflection points, we use Theorem 29 and find where $f''(x) = 0$ or where f'' is undefined. We find f'' is always defined, and is 0 only when $x = 0$. So the point $(0, 1)$ is the only possible point of inflection.

This possible inflection point divides the real line into two intervals, $(-\infty, 0)$ and $(0, \infty)$. We use a process similar to the one used in the previous section to determine increasing/decreasing. Pick any $c < 0$; $f''(c) < 0$ so f is concave down on $(-\infty, 0)$. Pick any $c > 0$; $f''(c) > 0$ so f is concave up on $(0, \infty)$. Since the concavity changes at $x = 0$, the point $(0, 1)$ is an inflection point.

The number line in Figure 3.23(a) illustrates the process of determining concavity (to save space, we will abbreviate "concave down", "concave up", and "inflection point" to "CD", "CU", and "IP", respectively). Figure 3.23(b) shows a graph of f and f'', confirming our results. Notice how f is concave down precisely when $f''(x) < 0$ and concave up when $f''(x) > 0$.

Example 2 Finding intervals of concave up/down, inflection points

Let $f(x) = x/(x^2 - 1)$. Find the inflection points of f and the intervals on which it is concave up/down.

SOLUTION The first thing we see is that f itself is not defined at $x = \pm 1$, having a domain of $(-\infty, -1) \cup (-1, 1) \cup (1, \infty)$. Since the domain of f is the union of three intervals, it makes sense that the concavity of f could switch across intervals. We cannot say that f has points of inflection at $x = \pm 1$ as they are not part of the domain, but we must still consider these x-values to be important and will include them in our number line.

We need to find f' and f''. Using the Quotient Rule and simplifying, we find

$$f'(x) = \frac{-(1 + x^2)}{(x^2 - 1)^2} \quad \text{and} \quad f''(x) = \frac{2x(x^2 + 3)}{(x^2 - 1)^3}.$$

Figure 3.23: A number line determining the concavity of f and a graph of f used in Example 1.

Notes:

To find the possible points of inflection, we seek to find where $f''(x) = 0$ and where f'' is not defined. Solving $f''(x) = 0$ reduces to solving $2x(x^2 + 3) = 0$; we find $x = 0$. We find that f'' is not defined when $x = \pm 1$, for then the denominator of f'' is 0 (of course, f is not defined at these points either).

The important x-values at which concavity might switch are $x = -1$, $x = 0$ and $x = 1$, which split the number line into four intervals as shown in our sign chart below. We determine the concavity on each. Keep in mind that all we are concerned with is the *sign* of f'' on the interval.

Interval 1, $(-\infty, -1)$: Select a number c in this interval with a large magnitude (for instance, $c = -100$). The denominator of $f''(x)$ will be positive. In the numerator, the $(c^2 + 3)$ will be positive and the $2c$ term will be negative. Thus the numerator is negative and $f''(c)$ is negative. We conclude f is concave down on $(-\infty, -1)$.

Interval 2, $(-1, 0)$: For any number c in this interval, the term $2c$ in the numerator will be negative, the term $(c^2 + 3)$ in the numerator will be positive, and the term $(c^2 - 1)^3$ in the denominator will be negative. Thus $f''(c) > 0$ and f is concave up on this interval.

Interval 3, $(0, 1)$: Any number c in this interval will be positive and "small." Thus the numerator is positive while the denominator is negative. Thus $f''(c) < 0$ and f is concave down on this interval.

Interval 4, $(1, \infty)$: Choose a large value for c. It is evident that $f''(c) > 0$, so we conclude that f is concave up on $(1, \infty)$.

We conclude that f is concave up on $(-1, 0)$ and $(1, \infty)$ and concave down on $(-\infty, -1)$ and $(0, 1)$. There is only one point of inflection, $(0, 0)$, as f is not defined at $x = \pm 1$. Our work is confirmed by the graph of f in Figure 3.24. Notice how f is concave up whenever f'' is positive, and concave down when f'' is negative.
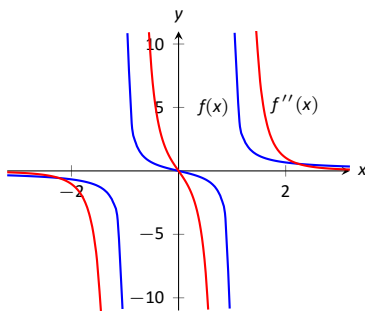
Figure 3.24: A graph of $f(x)$ and $f''(x)$ in Example 2.

Notes:

Recall that relative maxima and minima of f are found at critical points of f; that is, they are found when $f'(x) = 0$ or when f' is undefined. Likewise, the relative maxima and minima of f' are found when $f''(x) = 0$ or when f'' is undefined; note that these are the inflection points of f.

What does a "relative maximum of f'" *mean*? The derivative measures the rate of change of f; maximizing f' means finding where f is increasing the most – where f has the steepest tangent line. A similar statement can be made for minimizing f'; it corresponds to where f has the steepest negatively–sloped tangent line.

We utilize this concept in the next example.

Example 3   Understanding inflection points

The sales of a certain product over a three-year span are modeled by $S(t) = t^4 - 8t^2 + 20$, where t is the time in years, shown in Figure 3.25. Over the first two years, sales are decreasing. Find the point at which sales are decreasing at their greatest rate.

SOLUTION   We want to maximize the rate of decrease, which is to say, we want to find where S' has a minimum. To do this, we find where S'' is 0. We find $S'(t) = 4t^3 - 16t$ and $S''(t) = 12t^2 - 16$. Setting $S''(t) = 0$ and solving, we get $t = 2/\sqrt{3} \approx 1.16$ (we ignore the negative value of t since it does not lie in the domain of our function S).

Figure 3.25: A graph of $S(t)$ in Example 3, modeling the sale of a product over time.

0		$\frac{2}{\sqrt{3}} \approx 1.2$	
t			
f''	$-$		$+$
f	CD	IP	CU

This is both the inflection point and the point of maximum decrease. This is the point at which things first start looking up for the company. After the inflection point, it will still take some time before sales start to increase, but at least sales are not decreasing quite as quickly as they had been.

A graph of $S(t)$ and $S'(t)$ is given in Figure 3.26. When $S'(t) < 0$, sales are decreasing; note how at $t \approx 1.16$, $S'(t)$ is minimized. That is, sales are decreasing at the fastest rate at $t \approx 1.16$. On the interval of $(1.16, 2)$, S is decreasing but concave up, so the decline in sales is "leveling off."

Figure 3.26: A graph of $S(t)$ in Example 3 along with $S'(t)$.

Not every critical point corresponds to a relative extrema; $f(x) = x^3$ has a critical point at $(0, 0)$ but no relative maximum or minimum. Likewise, just because $f''(x) = 0$ we cannot conclude concavity changes at that point. We were

Notes:
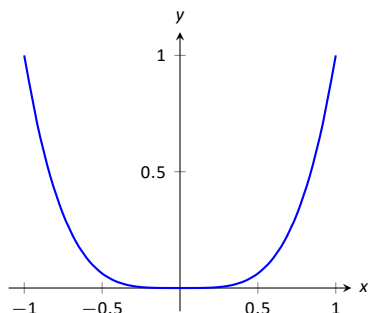
Figure 3.27: A graph of $f(x) = x^4$. Clearly f is always concave up, despite the fact that $f''(x) = 0$ when $x = 0$. It this example, the *possible* point of inflection $(0, 0)$ is not a point of inflection.
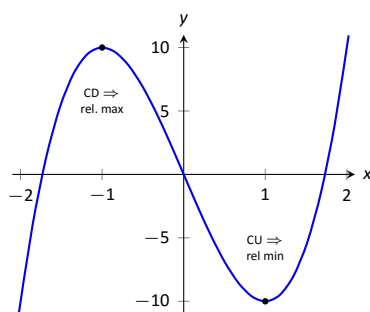
Figure 3.28: Demonstrating the fact that relative maxima occur when the graph is concave down and relative minima occur when the graph is concave up.
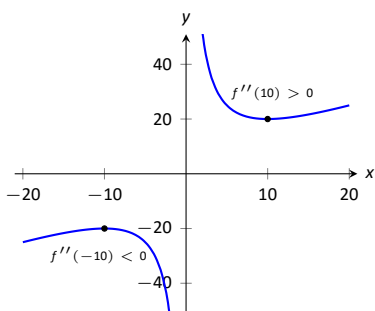
Figure 3.29: A graph of $f(x)$ in Example 4. The second derivative is evaluated at each critical point. When the graph is concave up, the critical point represents a local minimum; when the graph is concave down, the critical point represents a local maximum.

careful before to use terminology "*possible* point of inflection" since we needed to check to see if the concavity changed. The canonical example of $f''(x) = 0$ *without* concavity changing is $f(x) = x^4$. At $x = 0$, $f''(x) = 0$ but f is always concave up, as shown in Figure 3.27.

The Second Derivative Test

The first derivative of a function gave us a test to find if a critical value corresponded to a relative maximum, minimum, or neither. The second derivative gives us another way to test if a critical point is a local maximum or minimum. The following theorem officially states something that is intuitive: if a critical value occurs in a region where a function f is concave up, then that critical value must correspond to a relative minimum of f, etc. See Figure 3.28 for a visualization of this.

Theorem 30   The Second Derivative Test

Let c be a critical value of f where $f''(c)$ is defined.

1. If $f''(c) > 0$, then f has a local minimum at $(c, f(c))$.

2. If $f''(c) < 0$, then f has a local maximum at $(c, f(c))$.

Note that if $f''(c) = 0$, then the Second Derivative Test is inconclusive. The Second Derivative Test relates to the First Derivative Test in the following way. If $f''(c) > 0$, then the graph is concave up at a critical point c and f' itself is growing. Since $f'(c) = 0$ and f' is growing at c, then it must go from negative to positive at c. This means the function goes from decreasing to increasing, indicating a local minimum at c.

Example 4   Using the Second Derivative Test
Let $f(x) = 100/x + x$. Find the critical points of f and use the Second Derivative Test to label them as relative maxima or minima.

Solution   We find $f'(x) = -100/x^2 + 1$ and $f''(x) = 200/x^3$. We set $f'(x) = 0$ and solve for x to find the critical values (note that f' is not defined at $x = 0$, but neither is f so this is not a critical value.) We find the critical values are $x = \pm 10$. Evaluating f'' at $x = 10$ gives $0.1 > 0$, so there is a local minimum at $x = 10$. Evaluating $f''(-10) = -0.1 < 0$, determining a relative maximum at $x = -10$. These results are confirmed in Figure 3.29.

Notes:

We have been learning how the first and second derivatives of a function relate information about the graph of that function. We have found intervals of increasing and decreasing, intervals where the graph is concave up and down, along with the locations of relative extrema and inflection points. In Chapter 1 we saw how limits explained asymptotic behavior. In the next section we combine all of this information to produce accurate sketches of functions.

Notes:

Exercises 3.4

Terms and Concepts

1. Sketch a graph of a function $f(x)$ that is concave up on $(0, 1)$ and is concave down on $(1, 2)$.

2. Sketch a graph of a function $f(x)$ that is:

   (a) Increasing, concave up on $(0, 1)$,

   (b) increasing, concave down on $(1, 2)$,

   (c) decreasing, concave down on $(2, 3)$ and

   (d) increasing, concave down on $(3, 4)$.

3. Is is possible for a function to be increasing and concave down on $(0, \infty)$ with a horizontal asymptote of $y = 1$? If so, give a sketch of such a function.

4. Is is possible for a function to be increasing and concave up on $(0, \infty)$ with a horizontal asymptote of $y = 1$? If so, give a sketch of such a function.

Problems

5. Given the graph of f'', identify the concavity of f and its inflection points.
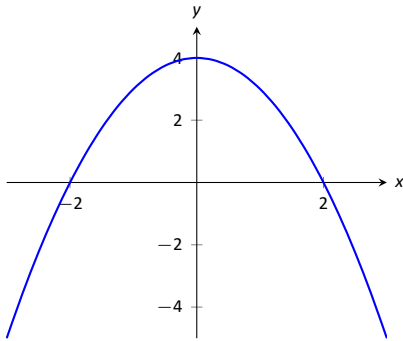
6. Given the graph of f', identify the concavity of f and its inflection points.
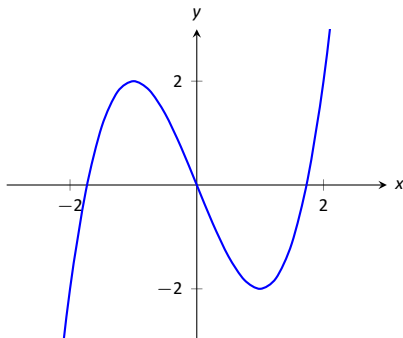
7. Given the graph of f, identify the concavity of f and its inflection points.
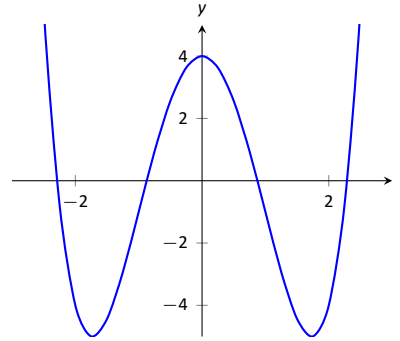
In Exercises 8–18, a function $f(x)$ is given.

   (a) Compute $f''(x)$.

   (b) Graph f and f'' on the same axes (using technology is permitted) and verify Theorem 28.

8. $f(x) = -7x + 3$

9. $f(x) = -4x^2 + 3x - 8$

10. $f(x) = 4x^2 + 3x - 8$

11. $f(x) = x^3 - 3x^2 + x - 1$

12. $f(x) = -x^3 + x^2 - 2x + 5$

13. $f(x) = \cos x$

14. $f(x) = \sin x$

15. $f(x) = \tan x$

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

17. $f(x) = \dfrac{1}{x}$

18. $f(x) = \dfrac{1}{x^2}$

In Exercises 19–37, a function $f(x)$ is given.

   (a) Find the possible points of inflection of f.

   (b) Create a number line to determine the intervals on which f is concave up or concave down.

   (c) Find the critical points of f and use the Second Derivative Test, when possible, to determine the relative extrema.

   (d) Find the x values where $f'(x)$ has a relative maximum or minimum.

19. $f(x) = x^2 - 2x + 1$

20. $f(x) = -x^2 - 5x + 7$

21. $f(x) = x^3 - x + 1$

22. $f(x) = 2x^3 - 3x^2 + 9x + 5$

23. $f(x) = \dfrac{x^4}{4} + \dfrac{x^3}{3} - 2x + 3$

24. $f(x) = -3x^4 + 8x^3 + 6x^2 - 24x + 2$

25. $f(x) = x^4 - 4x^3 + 6x^2 - 4x + 1$

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

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

28. $f(x) = \sin x + \cos x$ on $(-\pi, \pi)$

29. $f(x) = x^2 e^x$

30. $f(x) = x^2 \ln x$

31. $f(x) = e^{-x^2}$

32. $f(x) = x\sqrt{x + 3}$

33. $f(x) = \cos^2 x - 2 \sin x$ on $(0, 2\pi)$

34. $f(x) = x^3 - 3x^2 - 9x + 4$

35. $f(x) = x^4 - 2x^2 + 3$

36. $f(x) = 1 + 3x^2 - 2x^3$

37. $f(x) = \sqrt{x} - \sqrt[4]{x}$

3.5   Curve Sketching

We have been learning how we can understand the behavior of a function based on its first and second derivatives. While we have been treating the properties of a function separately (increasing and decreasing, concave up and concave down, etc.), we combine them here to produce an accurate graph of the function without plotting lots of extraneous points.

Why bother? Graphing utilities are very accessible, whether on a computer, a hand–held calculator, or a smartphone. These resources are usually very fast and accurate. We will see that our method is not particularly fast – it will require time (but it is not *hard*). So again: why bother?

We are attempting to understand the behavior of a function f based on the information given by its derivatives. While all of a function's derivatives relay information about it, it turns out that "most" of the behavior we care about is explained by f' and f''. Understanding the interactions between the graph of f and f' and f'' is important and is illustrated in Figure 3.30. To gain this understanding, one might argue that all that is needed is to look at lots of graphs. This is true to a point, but is somewhat similar to stating that one understands how an engine works after looking only at pictures. It is true that the basic ideas will be conveyed, but "hands–on" access increases understanding.

The following Key Idea summarizes what we have learned so far that is applicable to sketching graphs of functions and gives a framework for putting that information together. It is followed by several examples.
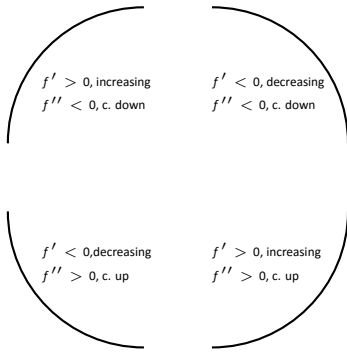
Figure 3.30: Demonstrating the 4 ways that concavity interacts with increasing/decreasing, along with the relationships with the first and second derivatives.

Key Idea 6   Curve Sketching

To produce an accurate sketch a given function f, consider the following steps.

1. Find the domain of f. Generally, we assume that the domain is the entire real line then find restrictions, such as where a denominator is 0 or where negatives appear under the radical.

2. Find the location of any vertical asymptotes of f (usually done in conjunction with the previous step).

3. Find the x and y-intercepts of f, and any symmetry.

(continued)

Notes:

Key Idea 6   Curve Sketching — Continued

4. Consider the limits $\lim\limits_{x\to-\infty} f(x)$ and $\lim\limits_{x\to\infty} f(x)$ to determine the end behavior of the function.

5. Find the critical values of f.

6. Find the possible points of inflection of f.

7. Create a number line that includes all critical points, possible points of inflection, and locations of vertical asymptotes. For each interval created, determine whether f is increasing or decreasing, concave up or down.

8. Evaluate f at each critical point and possible point of inflection. Plot these points on a set of axes. Connect these points with curves exhibiting the proper concavity. Sketch asymptotes and x and y-intercepts where applicable.

Watch the video:
Summary of Curve Sketching — Example 2, Part 1 of 4 at
`https://youtu.be/DMYUsv8ZaoY`

Example 1   Curve sketching
Use Key Idea 6 to sketch $f(x) = 3x^3 - 10x^2 + 4x + 10$.

   SOLUTION   We follow the steps outlined in the Key Idea.

1. The domain of f is the entire real line; there are no values x for which $f(x)$ is not defined.

2. There are no vertical asymptotes.

3. We see that $f(0) = 10$, and f does not appear to factor easily (so we skip finding the roots). It has no symmetry.

Notes:

4. We determine the end behavior using limits as x approaches $\pm$infinity.

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

We do not have any horizontal asymptotes.

5. Find the critical values of f. We compute $f'(x) = 9x^2 - 20x + 4 = (9x - 2)(x - 2)$, so that $x = \frac{2}{9}, 2$.

6. Find the possible points of inflection of f. We see $f''(x) = 18x - 20$, so that

$$f''(x) = 0 \Rightarrow x = 10/9 \approx 1.111.$$

7. We place the values $x = \frac{2}{9}, \frac{10}{9}, 2$ on a number line. We mark each subinterval as increasing or decreasing, concave up or down, using the techniques used in Sections 3.3 and 3.4.

		$\frac{2}{9} \approx 0.2$		$\frac{10}{9} \approx 1.1$		2	
f'	$+$		$-$		$-$		$+$
f''	$-$		$-$		$+$		$+$
f	incr CD	max	decr CD	IP	decr CU	min	incr CU

8. We plot the appropriate points on axes as shown in Figure 3.31(a) and connect the points with the proper concavity. Our curve crosses the y axis at $y = 10$ and crosses the x axis near $x = -0.75$. In Figure 3.31(b) we show a graph of f drawn with a computer program, verifying the accuracy of our sketch.

Example 2   Curve sketching

Sketch $f(x) = \dfrac{x^2 - x - 2}{x^2 - x - 6}$.

SOLUTION   We again follow the steps outlined in Key Idea 6.

1. In determining the domain, we assume it is all real numbers and look for restrictions. We find that at $x = -2$ and $x = 3$, $f(x)$ is not defined. So the domain of f is $D = \{\text{real numbers } x \mid x \neq -2, 3\}$.

(a)

(b)

Figure 3.31: Sketching f in Example 1.

Notes:

2. The vertical asymptotes of f are at $x = -2$ and $x = 3$, the places where f is undefined. We see that $\lim\limits_{x \to -2^-} f(x) = \infty$, $\lim\limits_{x \to -2^+} f(x) = -\infty$, $\lim\limits_{x \to 3^-} f(x) = -\infty$, and $\lim\limits_{x \to 3^+} f(x) = \infty$.

3. We see that $f(0) = \frac{1}{3}$ and that $f(x) = 0$ when $0 = x^2 - x - 2 = (x-2)(x+1)$ so that $x = -1, 2$. There is no symmetry.

4. There is a horizontal asymptote of $y = 1$, as $\lim\limits_{x \to -\infty} f(x) = 1$ and $\lim\limits_{x \to \infty} f(x) = 1$.

5. To find the critical values of f, we first find $f'(x)$. Using the Quotient Rule, we find
$$f'(x) = \frac{-8x + 4}{(x^2 + x - 6)^2} = \frac{-8x + 4}{(x-3)^2(x+2)^2}.$$

$f'(x) = 0$ when $x = 1/2$, and f' is undefined when $x = -2, 3$. Since f' is undefined only when f is, these are not critical values. The only critical value is $x = 1/2$.

6. To find the possible points of inflection, we find $f''(x)$, again employing the Quotient Rule:
$$f''(x) = \frac{24x^2 - 24x + 56}{(x-3)^3(x+2)^3}.$$

We find that $f''(x)$ is never 0 (setting the numerator equal to 0 and solving for x, we find the only roots to this quadratic are imaginary) and f'' is undefined when $x = -2, 3$. Thus concavity will possibly only change at $x = -2$ and $x = 3$ (although these are not inflection points, since f is not defined there).

7. We place the values $x = 1/2$, $x = -2$ and $x = 3$ on a number line. We mark in each interval whether f is increasing or decreasing, concave up or down. We see that f has a relative maximum at $x = 1/2$; concavity changes only at the vertical asymptotes.

x		-2		$\frac{1}{2}$		3	
f'	$+$		$+$		$-$		$-$
f''	$+$		$-$		$-$		$+$
f	incr CU	U	incr CD	max	decr CD	U	decr CU
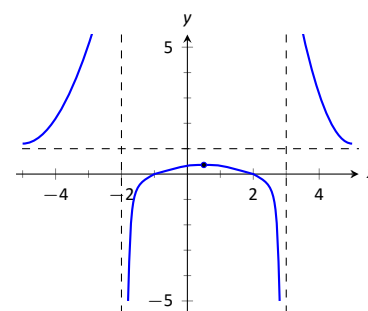
(a)
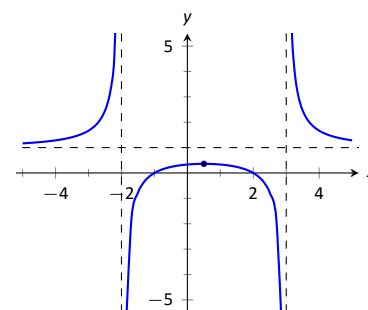
(b)

Figure 3.32: Sketching f in Example 2.

Notes:

8. In Figure 3.32(a), we plot the points from the number line on a set of axes and connect the points with the appropriate concavity. We also show f crossing the x axis at $x = -1$ and $x = 2$. Figure 3.32(b) shows a computer generated graph of f, which verifies the accuracy of our sketch.

Example 3   **Curve sketching**

Sketch $f(x) = \dfrac{5(x-2)(x+1)}{x^2 + 2x + 4}$.

    SOLUTION   We again follow Key Idea 6.

1. We assume that the domain of f is all real numbers and consider restrictions. The only restrictions come when the denominator is 0, but this never occurs. Therefore the domain of f is all real numbers, $\mathbb{R}$.

2. There are no vertical asymptotes.

3. We see that $f(0) = -\dfrac{5}{2}$ and that $f(x) = 0$ when $x = -1, 2$.

4. We have a horizontal asymptote of $y = 5$, as

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

5. We find the critical values of f by setting $f'(x) = 0$ and solving for x. We find

$$f'(x) = \frac{15x(x+4)}{(x^2 + 2x + 4)^2} \quad \Rightarrow \quad f'(x) = 0 \text{ when } x = -4, 0.$$

6. We find the possible points of inflection by solving $f''(x) = 0$ for x. We find

$$f''(x) = -\frac{30x^3 + 180x^2 - 240}{(x^2 + 2x + 4)^3}.$$

The cubic in the numerator does not factor very "nicely." We instead approximate the roots at $x = -5.759$, $x = -1.305$ and $x = 1.064$.

7. We place the critical points and possible inflection points on a number line and mark each interval as increasing/decreasing, concave up/down appropriately.

Notes:

x ←		−5.8		−4		−1.3		0		1.1	
f′	+		+		−		−		+		+
f″	+		−		−		+		+		−
f	incr CU	IP	incr CD	max	decr CD	IP	decr CU	min	incr CU	IP	incr CD

8. In Figure 3.33(a) we plot the significant points from the number line as well as the two roots of f, x = −1 and x = 2, and connect the points with the appropriate concavity. Figure 3.33(b) shows a computer generated graph of f, affirming our results (but the top left was slightly off).

In each of our examples, we found a few, significant points on the graph of f that corresponded to changes in increasing/decreasing or concavity. We connected these points with curves, and finished by showing a very accurate, computer generated graph.

Why are computer graphics so good? It is not because computers are "smarter" than we are. Rather, it is largely because computers are much faster at computing than we are. In general, computers graph functions much like most students do when first learning to draw graphs: they plot equally spaced points, then connect the dots using lines. By using lots of points, the connecting lines are short and the graph looks smooth.

This does a fine job of graphing in most cases (in fact, this is the method used for many graphs in this text). However, in regions where the graph is very "curvy," this can generate noticeable sharp edges on the graph unless a large number of points are used. High quality computer algebra systems, such as *Mathematica*, use special algorithms to plot lots of points only where the graph is "curvy."

In Figure 3.34, a graph of y = sin x is given, generated by *Mathematica*. The small points represent each of the places *Mathematica* sampled the function. Notice how at the "bends" of sin x, lots of points are used; where sin x is relatively straight, fewer points are used. (Many points are also used at the endpoints to ensure the "end behavior" is accurate.)

How does *Mathematica* know where the graph is "curvy"? Calculus. When we study *curvature* in a later chapter, we will see how the first and second derivatives of a function work together to provide a measurement of "curviness." *Mathematica* employs algorithms to determine regions of "high curvature" and plots extra points there.

Again, the goal of this section is not "How to graph a function when there is no computer to help." Rather, the goal is "Understand that the shape of the graph of a function is largely determined by understanding the behavior of the
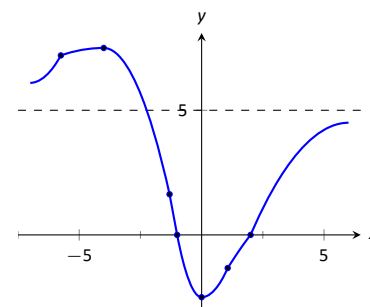
(a)
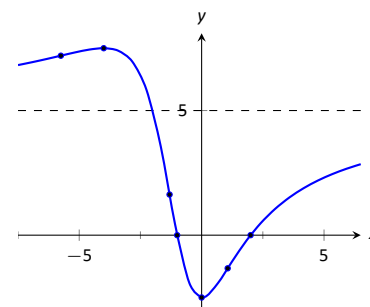
(b)

Figure 3.33: Sketching f in Example 3.

Notes:
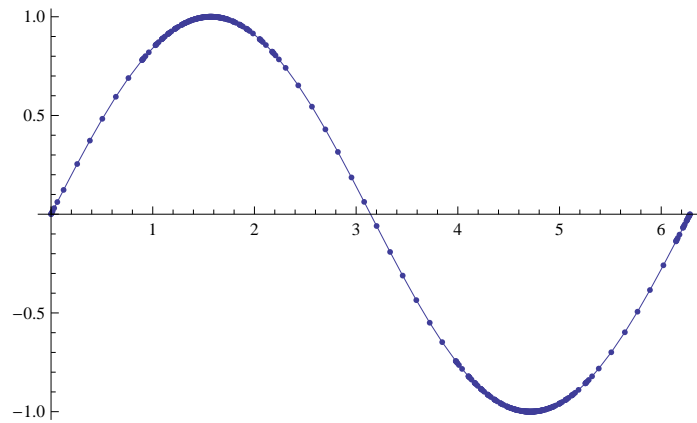
Figure 3.34: A graph of $y = \sin x$ generated by *Mathematica*.

function at a few key places." In Example 3, we were able to accurately sketch a complicated graph using only 5 points and knowledge of asymptotes!

There are many applications of our understanding of derivatives beyond curve sketching. The next chapter explores some of these applications, demonstrating just a few kinds of problems that can be solved with a basic knowledge of differentiation.

Notes:

Exercises 3.5

Terms and Concepts

1. Why is sketching curves by hand beneficial even though technology is readily available?
2. T/F: When sketching graphs of functions, it is useful to find the critical points.
3. T/F: When sketching graphs of functions, it is useful to find the possible points of inflection.
4. T/F: When sketching graphs of functions, it is useful to find the horizontal and vertical asymptotes.

Problems

5. Given the graph of f, identify the concavity of f, its inflection points, its regions of increasing and decreasing, and its relative extrema.
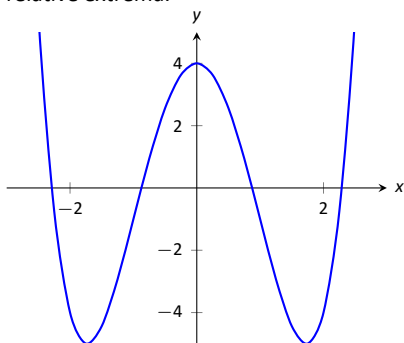
6. Given the graph of f', identify the concavity of f, its inflection points, its regions of increasing and decreasing, and its relative extrema.
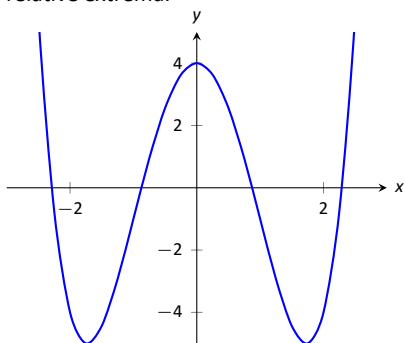
In Exercises 7–12, practice using Key Idea 6 by applying the principles to the given functions with familiar graphs.

7. $f(x) = 2x + 4$
8. $f(x) = -x^2 + 1$
9. $f(x) = \sin x$
10. $f(x) = e^x$
11. $f(x) = \dfrac{1}{x}$
12. $f(x) = \dfrac{1}{x^2}$

In Exercises 13–46, sketch a graph of the given function using Key Idea 6. Show all work; check your answer with technology.

13. $f(x) = x^3 - 2x^2 + 4x + 1$
14. $f(x) = -x^3 + 5x^2 - 3x + 2$
15. $f(x) = x^3 + 3x^2 + 3x + 1$
16. $f(x) = x^3 - x^2 - x + 1$
17. $f(x) = (x - 2)\ln(x - 2)$
18. $f(x) = (x - 2)^2 \ln(x - 2)$
19. $f(x) = \dfrac{x^2 - 4}{x^2}$
20. $f(x) = \dfrac{x^2 - 4x + 3}{x^2 - 6x + 8}$
21. $f(x) = x + \sin x$ on $[0, 2\pi]$.
22. $f(x) = \dfrac{x^2 - 2x + 1}{x^2 - 6x + 8}$
23. $f(x) = x\sqrt{x + 1}$
24. $f(x) = x^2 e^x$
25. $f(x) = \sin x \cos x$ on $[-\pi, \pi]$
26. $f(x) = (x - 3)^{2/3} + 2$
27. $f(x) = \dfrac{(x - 1)^{2/3}}{x}$
28. $f(x) = \sqrt{\dfrac{x}{x - 5}}$
29. $f(x) = \sec x - 2\cos x$ on $[0, 2\pi]$.
30. $f(x) = x\sqrt{2 - x^2}$
31. $f(x) = \dfrac{x}{\sqrt{x^2 - 1}}$
32. $f(x) = x^{5/3} - 5x^{2/3}$
33. $f(x) = \dfrac{\sin x}{2 + \cos x}$ on $[0, 2\pi]$.
34. $f(x) = \dfrac{x}{x^2 + 3}$
35. $f(x) = \dfrac{4x^2 - 4x + 1}{4x^2 - 12x + 9}$
    Hint: $f(x)$ can be simplified in a variety of ways. Use whichever simplification works best for your current task.
36. $y = \sqrt{x^2 + x} - x$
37. $y = x + \cos x$
38. $y = x \tan x$ on $\left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$
39. $y = \sin x + \sqrt{3}\cos x$ on $[-2\pi, 2\pi]$
40. $y = \csc x - 2\sin x$ on $(0, \pi)$
41. $f(x) = \dfrac{3}{x^2 + 4}$
42. $f(x) = \dfrac{3}{x^2 - 4}$
43. $f(x) = \dfrac{x}{x^2 + 4}$
44. $f(x) = \dfrac{x}{x^2 - 4}$
45. $f(x) = \dfrac{x}{x - 4}$
46. $f(x) = \dfrac{(x - 1)^2}{x^2 + 1}$

In Exercises 47–52, sketch the graph of a function that satisfies all of the given conditions.

47. $f'(0) = f'(2) = f'(4) = 0$,
$f'(x) > 0$ if $x < 0$ or $2 < x < 4$,
$f'(x) < 0$ if $0 < x < 2$ or $x > 4$,
$f''(x) > 0$ if $1 < x < 3$, $f''(x) < 0$ if $x < 1$ or $x > 4$

48. $f'(5) = 0$, $f'(x) < 0$ when $x < 5$,
$f'(x) > 0$ if $x > 5$, $f''(2) = 0$, $f''(8) = 0$,
$f''(x) < 0$ if $x < 2$ or $x > 8$,
$f''(x) > 0$ if $2 < x < 8$

49. $f(0) = 0$, $f'(-2) = f'(1) = f'(9) = 0$,
$\lim\limits_{x \to \infty} f(x) = 0$, $\lim\limits_{x \to 6} f(x) = -\infty$,
$f'(x) < 0$ on $(-\infty, -2)$, $(1, 6)$, $(9, \infty)$,
$f'(x) > 0$ on $(-2, 1)$, $(6, 9)$,
$f''(x) > 0$ on $(-\infty, 0)$, $(12, \infty)$,
$f''(x) < 0$ on $(0, 6)$, $(6, 12)$

50. f is odd, $f'(x) < 0$ on $(0, 2)$,
$f'(x) > 0$ on $(2, \infty)$, $f''(x) > 0$ on $(0, 3)$,
$f''(x) < 0$ on $(3, \infty)$, $\lim\limits_{x \to \infty} f(x) = -2$

51. concave up on $(-\infty, -1)$, $(1, \infty)$;
concave down on $(-1, 1)$;
increasing on $(-\infty, 0)$; and
decreasing on $(0, \infty)$.

52. $\lim\limits_{x \to -\infty} f(x) = 1$, $\lim\limits_{x \to 3^-} f(x) = \infty$,
$\lim\limits_{x \to \infty} f(x) = -1$, $\lim\limits_{x \to 3^+} f(x) = -\infty$;
$f'(x) > 0$ on $(-\infty, -2)$, $(-1, 0)$, $(2, 3)$, $(3, \infty)$;
$f'(x) < 0$ on $(-2, -1)$, $(0, 2)$;
$f''(x) > 0$ on $(-\infty, -3)$, $(1, 3)$; and
$f''(x) < 0$ on $(-3, -1)$, $(-1, 1)$, $(3, \infty)$.

In Exercises 53–55, a function with the parameters a and b are given. Describe the critical points and possible points of inflection of f in terms of a and b.

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

54. $f(x) = \sin(ax + b)$

55. $f(x) = (x - a)(x - b)$

56. Given $x^2 + y^2 = 1$, use implicit differentiation to find $\dfrac{dy}{dx}$ and $\dfrac{d^2y}{dx^2}$. Use this information to justify the sketch of the unit circle.

186

4: Applications of the Derivative

In Chapter 3, we learned how the first and second derivatives of a function influence its graph. In this chapter we explore other applications of the derivative.

4.1 Related Rates

When two quantities are related by an equation, knowing the value of one quantity can determine the value of the other. For instance, the circumference and radius of a circle are related by $C = 2\pi r$; knowing that $C = 6\pi$ in determines the radius must be 3 in.

The topic of **related rates** takes this one step further: knowing the *rate* at which one quantity is changing can determine the *rate* at which the other changes.

We demonstrate the concepts of related rates through examples.

Note: This section relies heavily on implicit differentiation, so referring back to Section 2.6 may help.

Example 1   Understanding related rates
The radius of a circle is growing at a rate of 5 in/hr. At what rate is the circumference growing?

SOLUTION   The circumference and radius of a circle are related by $C = 2\pi r$. We are given information about how the length of r changes with respect to time; that is, we are told $\frac{dr}{dt} = 5$ in/hr. We want to know how the length of C changes with respect to time, i.e., we want to know $\frac{dC}{dt}$.

Implicitly differentiate both sides of $C = 2\pi r$ with respect to t:

$$C = 2\pi r$$
$$\frac{d}{dt}(C) = \frac{d}{dt}(2\pi r)$$
$$\frac{dC}{dt} = 2\pi \frac{dr}{dt}.$$

As we know $\frac{dr}{dt} = 5$ in/hr, we know

$$\frac{dC}{dt} = 2\pi 5 = 10\pi \text{ in/hr.}$$

Before we look at another example, we'll state a few ideas on approaching these problems.

Key Idea 7   Solving Related Rates Problems

1. Understand the problem. Clearly identify what quantities are given and what are to be found. Make a sketch if helpful.

2. Assign mathematical notation to all quantities, including those that are functions of time.

3. Create an equation relevant to the context of the problem, using the information given.

4. Substitute constant quantities and if necessary, use the given information to eliminate other variables.

5. Use the Chain Rule to differentiate both sides of the equation.

6. Substitute the known quantities, and solve for the unknown rate.

The important thing to remember is that you must differentiate before you substitute varying values. Otherwise, you'll substitute a constant for what should be a variable, and it's derivative will be zero. Consider another example.

Example 2   Finding related rates
Water streams out of a faucet at a rate of 2 in^3/s onto a flat surface at a constant rate, forming a circular puddle that is $1/8$ in deep.

1. At what rate is the area of the puddle growing?

2. At what rate is the radius of the circle growing?

   SOLUTION

1. We can answer this question two ways: using "common sense" or related rates. The common sense method states that the volume of the puddle is growing by 2 in^3/s, where

   $$\text{volume of puddle} = \text{area of circle} \times \text{depth.}$$

Notes:

Since the depth is constant at $1/8$ in, the area must be growing by $16\text{in}^2/\text{s}$.

This approach reveals the underlying related–rates principle. Let V and A represent the Volume and Area of the puddle. We know $V = A \times \frac{1}{8}$. Take the derivative of both sides with respect to t, employing implicit differentiation.

$$V = \frac{1}{8}A$$

$$\frac{d}{dt}(V) = \frac{d}{dt}\left(\frac{1}{8}A\right)$$

$$\frac{dV}{dt} = \frac{1}{8}\frac{dA}{dt}$$

As $\frac{dV}{dt} = 2$, we know $2 = \frac{1}{8}\frac{dA}{dt}$, and hence $\frac{dA}{dt} = 16$. Thus the area is growing by 16 in^2/s.

2. To start, we need an equation that relates what we know to the radius. We know that $V = \pi r^2 h = \frac{\pi}{8}r^2$. Implicitly derive both sides with respect to t:

$$V = \frac{\pi}{8}r^2$$

$$\frac{d}{dt}(V) = \frac{d}{dt}\left(\frac{\pi}{8}r^2\right)$$

$$\frac{dV}{dt} = \frac{\pi}{8}\cdot 2r\frac{dr}{dt} = \frac{\pi}{4}r\frac{dr}{dt}$$

Solving for $\frac{dr}{dt}$, we have $\frac{dr}{dt} = \frac{\frac{dV}{dt}}{\frac{\pi}{4}r} = \frac{2}{\frac{\pi}{4}r} = \frac{8}{\pi r}$.

Note how our answer is not a number, but rather a function of r. In other words, *the rate at which the radius is growing depends on how big the circle already is.* If the circle is very large, adding 2 in^3 of water will not make the circle much bigger at all. If the circle is dime–sized, adding the same amount of water will make a radical change in the radius of the circle.

In some ways, our problem was (intentionally) ill–posed. We need to specify a current radius in order to know a rate of change. When the puddle has a radius of 10 in, the radius is growing at a rate of

$$\frac{dr}{dt} = \frac{8}{10\pi} = \frac{4}{5\pi} \approx 0.25 \text{ in/s}.$$

Notes:

B = 1/2

N

Car

E

A = 1/2

C

Officer
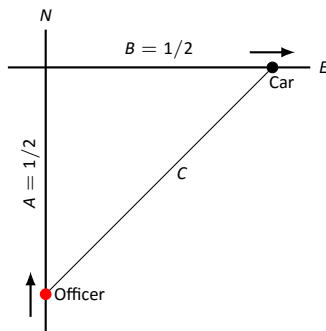
Figure 4.1: A sketch of a police car (at bottom) attempting to measure the speed of a car (at right) in Example 3.

Note: Example 3 is both interesting and impractical. It highlights the difficulty in using radar in a non–linear fashion, and explains why "in real life" the police officer would follow the other driver to determine their speed, and not pull out pencil and paper.

The principles here are important, though. Many automated vehicles make judgments about other moving objects based on perceived distances, radar–like measurements and the concepts of related rates.

Example 3   Studying related rates

Radar guns measure the rate of distance change between the gun and the object it is measuring. For instance, a reading of "55 mph" means the object is moving away from the gun at a rate of 55 miles per hour, whereas a measurement of "−25 mph" would mean that the object is approaching the gun at a rate of 25 miles per hour.

If the radar gun is moving (say, attached to a police car) then radar readouts are only immediately understandable if the gun and the object are moving along the same line. If a police officer is traveling 60 mph and gets a readout of 15 mph, he knows that the car ahead of him is moving away at a rate of 15 miles an hour, meaning the car is traveling 75 mph. (This straight–line principle is one reason officers park on the side of the highway and try to shoot straight back down the road. It gives the most accurate reading.)

Suppose an officer is driving due north at 30 mph and sees a car moving due east, as shown in Figure 4.1. Using his radar gun, he measures a reading of 20 mph. By using landmarks, he believes both he and the other car are about 1/2 mile from the intersection of their two roads.

If the speed limit on the other road is 55 mph, is the other driver speeding?

SOLUTION   Using the diagram in Figure 4.1, let's label what we know about the situation. As both the police officer and other driver are 1/2 mile from the intersection, we have $A = 1/2$, $B = 1/2$, and through the Pythagorean Theorem, $C = 1/\sqrt{2} \approx 0.707$.

We know the police officer is traveling at 30 mph; that is, $\frac{dA}{dt} = -30$. The reason this rate of change is negative is that A is getting smaller; the distance between the officer and the intersection is shrinking. The radar measurement is $\frac{dC}{dt} = 20$. We want to find $\frac{dB}{dt}$.

We need an equation that relates B to A and/or C. The Pythagorean Theorem is a good choice: $A^2 + B^2 = C^2$. Differentiate both sides with respect to t:

$$A^2 + B^2 = C^2$$
$$\frac{d}{dt}\left(A^2 + B^2\right) = \frac{d}{dt}\left(C^2\right)$$
$$2A\frac{dA}{dt} + 2B\frac{dB}{dt} = 2C\frac{dC}{dt}$$

We have values for everything except $\frac{dB}{dt}$. Solving for this we have

$$\frac{dB}{dt} = \frac{C\frac{dC}{dt} - A\frac{dA}{dt}}{B} \approx 58.28 \text{ mph.}$$

The other driver appears to be speeding slightly.

Notes:

Example 4   **Studying related rates**

A camera is placed on a tripod 10 ft from the side of a road. The camera is to turn to track a car that is to drive by at 100 mph for a promotional video. The video's planners want to know what kind of motor the tripod should be equipped with in order to properly track the car as it passes by. Figure 4.2 shows the proposed setup.

How fast must the camera be able to turn to track the car?

    SOLUTION   We seek information about how fast the camera is to *turn*; therefore, we need an equation that will relate an angle θ to the position of the camera and the speed and position of the car.

Figure 4.2 suggests we use a trigonometric equation. Letting *x* represent the distance the car is from the point on the road directly in front of the camera, we have

$$\tan\theta = \frac{x}{10}. \tag{4.1}$$

As the car is moving at 100 mph, we have $\frac{dx}{dt} = -100$ mph (as in the last example, since *x* is getting smaller as the car travels, $\frac{dx}{dt}$ is negative). We need to convert the measurements so they use the same units; rewrite -100 mph in terms of ft/s:

$$\frac{dx}{dt} = -100\frac{m}{hr} = -100\frac{m}{hr} \cdot 5280\frac{ft}{m} \cdot \frac{1}{3600}\frac{hr}{s} = -146.\overline{6}\ ft/s.$$

Now take the derivative of both sides of Equation (4.1) using implicit differentiation:

$$\tan\theta = \frac{x}{10}$$
$$\frac{d}{dt}\left(\tan\theta\right) = \frac{d}{dt}\left(\frac{x}{10}\right)$$
$$\sec^2\theta\,\frac{d\theta}{dt} = \frac{1}{10}\frac{dx}{dt}$$
$$\frac{d\theta}{dt} = \frac{\cos^2\theta}{10}\frac{dx}{dt} \tag{4.2}$$

We want to know the fastest the camera has to turn. Common sense tells us this is when the car is directly in front of the camera (i.e., when $\theta = 0$). Our mathematics bears this out. In Equation (4.2) we see this is when $\cos^2\theta$ is largest; this is when $\cos\theta = 1$, or when $\theta = 0$.

With $\frac{dx}{dt} \approx -146.67$ ft/s, we have

$$\frac{d\theta}{dt} = -\frac{1\text{ rad}}{10\text{ ft}}146.67\text{ ft/s} = -14.667\text{ radians/s}.$$

We find that $\frac{d\theta}{dt}$ is negative; this matches our diagram in Figure 4.2 for θ is getting smaller as the car approaches the camera.
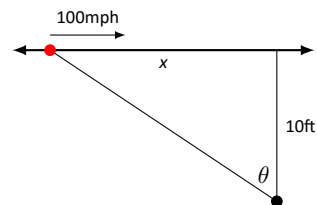
Figure 4.2: Tracking a speeding car (at left) with a rotating camera.

Notes:

What is the practical meaning of -14.667 radians/s? Recall that 1 circular revolution goes through 2π radians, thus 14.667 rad/s means $14.667/(2\pi) \approx$ 2.33 revolutions per second. The negative sign indicates the camera is rotating in a clockwise fashion.

We introduced the derivative as a function that gives the slopes of tangent lines of functions. This section stresses the "rate of change" aspect of the derivative to find a relationship between the rates of change of two related quantities. In the next section we use Extreme Value concepts to *optimize* quantities.

Notes:

Exercises 4.1

Terms and Concepts

1. T/F: Implicit differentiation is often used when solving "related rates" type problems.

2. T/F: A study of related rates is part of the standard police officer training.
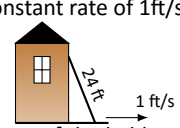
Problems

3. The area of a square is increasing at a rate of 42 ft^2/min. How fast is the side length increasing when the length is 7 ft?

4. Water flows onto a flat surface at a rate of 5cm^3/s forming a circular puddle 10mm deep. How fast is the radius growing when the radius is:

   (a) 1 cm?

   (b) 10 cm?

   (c) 100 cm?

5. A circular balloon is inflated with air flowing at a rate of 10cm^3/s. How fast is the radius of the balloon increasing when the radius is:

   (a) 1 cm?

   (b) 10 cm?

   (c) 100 cm?

6. Consider the traffic situation introduced in Example 3. How fast is the "other car" traveling if the officer and the other car are each 1/2 mile from the intersection, the other car is traveling *due west*, the officer is traveling north at 50mph, and the radar reading is −80mph?

7. Consider the traffic situation introduced in Example 3. Calculate how fast the "other car" is traveling in each of the following situations.

   (a) The officer is traveling due north at 50mph and is 1/2 mile from the intersection, while the other car is 1 mile from the intersection traveling west and the radar reading is −80mph?

   (b) The officer is traveling due north at 50mph and is 1 mile from the intersection, while the other car is 1/2 mile from the intersection traveling west and the radar reading is −80mph?
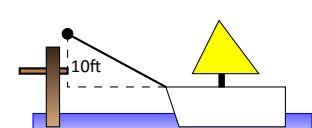
8. An F-22 aircraft is flying at 500mph with an elevation of 10,000ft on a straight–line path that will take it directly over an anti–aircraft gun.
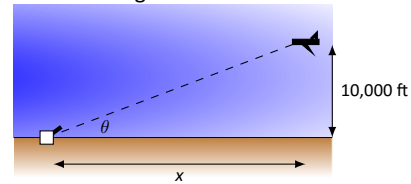
   How fast must the gun be able to turn to accurately track the aircraft when the plane is:

   (a) 1 mile away?

   (b) 1/5 mile away?

   (c) Directly overhead?

9. An F-22 aircraft is flying at 500mph with an elevation of 100ft on a straight–line path that will take it directly over an anti–aircraft gun as in Exercise 8 (note the lower elevation here).
   How fast must the gun be able to turn to accurately track the aircraft when the plane is:

   (a) 1000 feet away?

   (b) 100 feet away?

   (c) Directly overhead?

10. A 24ft. ladder is leaning against a house while the base is pulled away at a constant rate of 1ft/s.

   At what rate is the top of the ladder sliding down the side of the house when the base is:

   (a) 1 foot from the house?

   (b) 10 feet from the house?

   (c) 23 feet from the house?

   (d) 24 feet from the house?

11. A boat is being pulled into a dock at a constant rate of 30ft/min by a winch located 10ft above the deck of the boat.

   At what rate is the boat approaching the dock when the boat is:

   (a) 50 feet out?

   (b) 15 feet out?

   (c) 1 foot from the dock?

   (d) What happens when the length of rope pulling in the boat is less than 10 feet long?

12. An inverted cylindrical cone, 20ft deep and 10ft across at the top, is being filled with water at a rate of $10\text{ft}^3/\text{min}$. At what rate is the water rising in the tank when the depth of the water is:

    (a) 1 foot?

    (b) 10 feet?

    (c) 19 feet?

How long will the tank take to fill when starting at empty?

13. A rope, attached to a weight, goes up through a pulley at the ceiling and back down to a worker. The man holds the rope at the same height as the connection point between rope and weight.
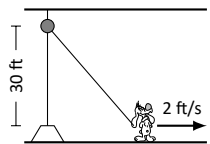
Suppose the man stands directly next to the weight (i.e., a total rope length of 60 ft) and begins to walk away at a rate of 2ft/s. How fast is the weight rising when the man has walked:

    (a) 10 feet?

    (b) 40 feet?

How far must the man walk to raise the weight all the way to the pulley?

14. Consider the situation described in Exercise 13. Suppose the man starts 40ft from the weight and begins to walk away at a rate of 2ft/s.

    (a) How long is the rope?

    (b) How fast is the weight rising after the man has walked 10 feet?

    (c) How fast is the weight rising after the man has walked 40 feet?

    (d) How far must the man walk to raise the weight all the way to the pulley?

15. A hot air balloon lifts off from ground rising vertically. From 100 feet away, a 5' woman tracks the path of the balloon. When her sightline with the balloon makes a $45°$ angle with the horizontal, she notes the angle is increasing at about $5°/\text{min}$.

    (a) What is the elevation of the balloon?

    (b) How fast is it rising?

16. A company that produces landscaping materials is dumping sand into a conical pile. The sand is being poured at a rate of $5\text{ft}^3/\text{sec}$; the physical properties of the sand, in conjunction with gravity, ensure that the cone's height is roughly 2/3 the length of the diameter of the circular base. How fast is the cone rising when it has a height of 30 feet?

4.2 Optimization

In Section 3.1 we learned about extreme values – the largest and smallest values a function attains on an interval. We motivated our interest in such values by discussing how it made sense to want to know the highest/lowest values of a stock, or the fastest/slowest an object was moving. In this section we apply the concepts of extreme values to solve "word problems," i.e., problems stated in terms of situations that require us to create the appropriate mathematical framework in which to solve the problem.

Watch the video:
Optimization Problem #7 — Minimizing the Area of Two Squares With Total Perimeter of Fixed Length at `https://youtu.be/BbTwa4Dbmmo`

We start with a classic example which is followed by a discussion of the topic of optimization.

Example 1   Optimization: perimeter and area
A man has 100 feet of fencing, a large yard, and a small dog. He wants to create a rectangular enclosure for his dog with the fencing that provides the maximal area. What dimensions provide the maximal area?

SOLUTION   One can likely guess the correct answer — that is great. We will proceed to show how calculus can provide this answer in a context that proves this answer is correct.

It helps to make a sketch of the situation. Our enclosure is sketched twice in Figure 4.3, either with green grass and nice fence boards or as a simple rectangle. Either way, drawing a rectangle forces us to realize that we need to know the dimensions of this rectangle so we can create an area function – after all, we are trying to maximize the area.

We let x and y denote the lengths of the sides of the rectangle. Clearly,

$$\text{Area} = xy.$$

We do not yet know how to handle functions with 2 variables; we need to reduce this down to a single variable. We know more about the situation: the man has 100 feet of fencing. By knowing the perimeter of the rectangle must be 100, we can create another equation:

$$\text{Perimeter} = 100 = 2x + 2y.$$
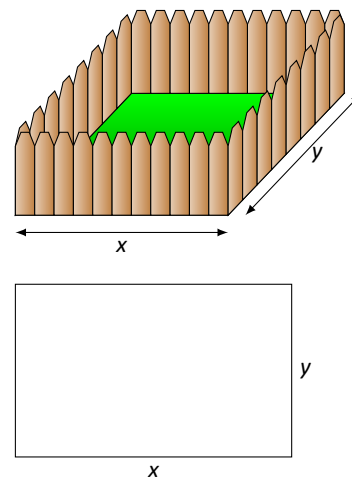
Figure 4.3: A sketch of the enclosure in Example 1.

Notes:

We now have 2 equations and 2 unknowns. In the latter equation, we solve for y:

$$y = 50 - x.$$

Now substitute this expression for y in the area equation:

$$\text{Area} = A(x) = x(50 - x).$$

Note we now have an equation of one variable; we can truly call the Area a function of x.

This function only makes sense when $0 \leq x \leq 50$, otherwise we get negative values of area. So we find the extreme values of $A(x)$ on the interval $[0, 50]$.

To find the critical points, we take the derivative of $A(x)$ and set it equal to 0, then solve for x.

$$\begin{aligned} A(x) &= x(50 - x) \\ &= 50x - x^2 \\ A'(x) &= 50 - 2x \end{aligned}$$

We solve $50 - 2x = 0$ to find $x = 25$; this is the only critical point. We evaluate $A(x)$ at the endpoints of our interval and at this critical point to find the extreme values; in this case, all we care about is the maximum.

Clearly $A(0) = 0$ and $A(50) = 0$, whereas $A(25) = 625$ ft^2. This is the maximum. Since we earlier found $y = 50 - x$, we find that y is also 25. Thus the dimensions of the rectangular enclosure with perimeter of 100 ft. with maximum area is a square, with sides of length 25 ft.

This example is very simplistic and a bit contrived. (After all, most people create a design then buy fencing to meet their needs, and not buy fencing and plan later.) But it models well the necessary process: create equations that describe a situation, reduce an equation to a single variable, then find the needed extreme value.

"In real life," problems are much more complex. The equations are often *not* reducible to a single variable (hence multi–variable calculus is needed) and the equations themselves may be difficult to form. Understanding the principles here will provide a good foundation for the mathematics you will likely encounter later.

We outline here the basic process of solving these optimization problems.

Notes:

Key Idea 8　　**Solving Optimization Problems**

1. Understand the problem. Clearly identify what quantity is to be maximized or minimized. Make a sketch if helpful.

2. Create equations relevant to the context of the problem, using the information given. (One of these should describe the quantity to be optimized. We'll call this the *fundamental equation.*)

3. If the fundamental equation defines the quantity to be optimized as a function of more than one variable, reduce it to a single variable function using substitutions derived from the other equations.

4. Identify the domain of this function, keeping in mind the context of the problem.

5. Find the extreme values of this function on the determined domain.

6. Identify the values of all relevant quantities of the problem.

We will use Key Idea 8 in a variety of examples.

Example 2　　Optimization: perimeter and area

Here is another classic calculus problem: A woman has a 100 feet of fencing, a small dog, and a large yard that contains a stream (that is mostly straight). She wants to create a rectangular enclosure with maximal area that uses the stream as one side. (Apparently her dog won't swim away.) What dimensions provide the maximal area?

SOLUTION　　We will follow the steps outlined by Key Idea 8.

1. We are maximizing *area*. A sketch of the region will help; Figure 4.4 gives two sketches of the proposed enclosed area. A key feature of the sketches is to acknowledge that one side is not fenced.

2. We want to maximize the area; as in the example before,

$$\text{Area} = xy.$$

This is our fundamental equation. This defines area as a function of two variables, so we need another equation to reduce it to one variable.
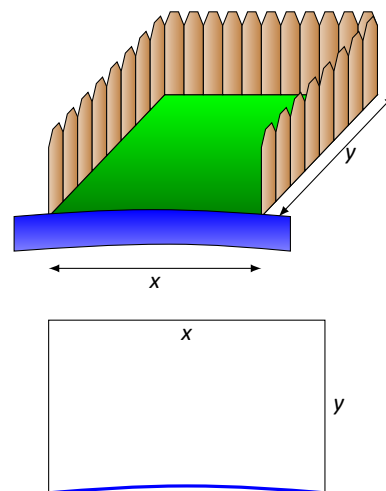
Figure 4.4: A sketch of the enclosure in Example 2.

Notes:

We again appeal to the perimeter; here the perimeter is

$$\text{Perimeter} = 100 = x + 2y.$$

Note how this is different than in our previous example.

3. We now reduce the fundamental equation to a single variable. In the perimeter equation, solve for y: $y = 50 - x/2$. We can now write Area as

$$\text{Area} = A(x) = x(50 - x/2) = 50x - \frac{1}{2}x^2.$$

Area is now defined as a function of one variable.

4. We want the area to be nonnegative. Since $A(x) = x(50 - x/2)$, we want $x \geq 0$ and $50 - x/2 \geq 0$. The latter inequality implies that $x \leq 100$, so $0 \leq x \leq 100$.

5. We now find the extreme values. At the endpoints, the minimum is found, giving an area of 0.

   Find the critical points. We have $A'(x) = 50 - x$; setting this equal to 0 and solving for x returns $x = 50$. This gives an area of

$$A(50) = 50(25) = 1250.$$

6. We earlier set $y = 50 - x/2$; thus $y = 25$. Thus our rectangle will have two sides of length 25 and one side of length 50, with a total area of 1250 ft^2.

Keep in mind as we do these problems that we are practicing a *process*; that is, we are learning to turn a situation into a system of equations. These equations allow us to write a certain quantity as a function of one variable, which we then optimize.
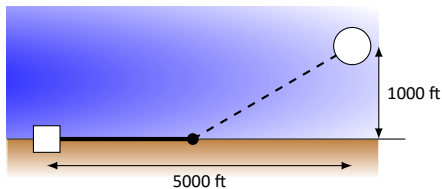
Figure 4.5: Running a power line from the power station to an offshore facility with minimal cost in Example 3.

Example 3   Optimization: minimizing cost
A power line needs to be run from a power station located on the beach to an offshore facility. Figure 4.5 shows the distances between the power station to the facility.

It costs $50/ft. to run a power line along the land, and $130/ft. to run a power line under water. How much of the power line should be run along the land to minimize the overall cost? What is the minimal cost?

Notes:

SOLUTION   We will follow the strategy of Key Idea 8 implicitly, without specifically numbering steps.

There are two immediate solutions that we could consider, each of which we will reject through "common sense." First, we could minimize the distance by directly connecting the two locations with a straight line. However, this requires that all the wire be laid underwater, the most costly option. Second, we could minimize the underwater length by running a wire all 5000 ft. along the beach, directly across from the offshore facility. This has the undesired effect of having the longest distance of all, probably ensuring a non–minimal cost.

The optimal solution likely has the line being run along the ground for a while, then underwater, as the figure implies. We need to label our unknown distances – the distance run along the ground and the distance run underwater. Recognizing that the underwater distance can be measured as the hypotenuse of a right triangle, we choose to label the distances as shown in Figure 4.6.

By choosing x as we did, we make the expression under the square root simple. We now create the cost function.

Figure 4.6: Labeling unknown distances in Example 3.

$$
\begin{aligned}
\text{Cost} \ &= \ \text{land cost} \ &+& \ \text{water cost} \\
&= \ \$50 \times \text{land distance} \ &+& \ \$130 \times \text{water distance} \\
&= \ 50(5000 - x) \ &+& \ 130\sqrt{x^2 + 1000^2}.
\end{aligned}
$$

So we have $c(x) = 50(5000 - x) + 130\sqrt{x^2 + 1000^2}$. This function only makes sense on the interval $[0, 5000]$. While we are fairly certain the endpoints will not give a minimal cost, we still evaluate $c(x)$ at each to verify.

$$c(0) = 380,000 \qquad c(5000) \approx 662,873.$$

We now find the critical values of $c(x)$. We compute $c'(x)$ as

$$c'(x) = -50 + \frac{130x}{\sqrt{x^2 + 1000^2}}.$$

Recognize that this is never undefined. Setting $c'(x) = 0$ and solving for x,

Notes:

we have:

$$-50 + \frac{130x}{\sqrt{x^2 + 1000^2}} = 0$$

$$\frac{130x}{\sqrt{x^2 + 1000^2}} = 50$$

$$\frac{130^2 x^2}{x^2 + 1000^2} = 50^2$$

$$130^2 x^2 = 50^2(x^2 + 1000^2)$$

$$130^2 x^2 - 50^2 x^2 = 50^2 \cdot 1000^2$$

$$(130^2 - 50^2)x^2 = 50,000^2$$

$$x^2 = \frac{50,000^2}{130^2 - 50^2}$$

$$x = \frac{50,000}{\sqrt{130^2 - 50^2}}$$

$$x = \frac{50,000}{120} = \frac{1250}{3} \approx 416.67.$$

Evaluating $c(x)$ at $x = 416.67$ gives a cost of about \$370,000. The distance the power line is laid along land is $5000 - 416.67 = 4583.33$ ft., and the underwater distance is $\sqrt{416.67^2 + 1000^2} \approx 1083$ ft.
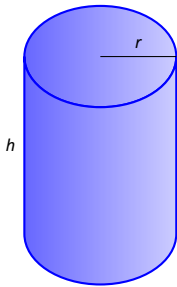
Figure 4.7: A sketch for Example 4.

Example 4   Optimization: Minimizing Surface Area
Design a closed cylindrical can of volume 8 ft^3 so that it uses the least amount of metal. In other words, minimize the surface area of the can.

SOLUTION   Following the strategy of Key Idea 8, we make a sketch in Figure 4.7 and identify the quantity to be minimized as the surface area of the cylinder. The formula for the surface area is our fundamental equation since it relates all of our relevant quantities.

$$A = \underbrace{\pi r^2}_{\text{Top}} + \underbrace{\pi r^2}_{\text{Bottom}} + \underbrace{2\pi rh}_{\text{Side}} = 2\pi r^2 + 2\pi rh$$

Our surface area is now defined in terms of two variables. To reduce this to a single variable we use the volume of a can, $V = \pi r^2 h$. Since the can must have $V = 8$ ft^3, we set $\pi r^2 h = 8$. Thus $h = \dfrac{8}{\pi r^2}$ and

$$A(r) = 2\pi r^2 + 2\pi r \frac{8}{\pi r^2} = 2\pi r^2 + \frac{16}{r}$$

Notes:

Next we find the critical values of $A(r)$. We compute

$$A'(r) = 4\pi r - \frac{16}{r^2} = \frac{4\pi r^3 - 16}{r^2}$$

and find that $A'(r) = 0$ when $r^3 = \dfrac{4}{\pi}$, that is, $r = \left(\dfrac{4}{\pi}\right)^{1/3} \approx 1.08$ ft.

Looking back at $A(r)$, we notice that r is not restricted to a closed interval. The radius can take on any positive value making the interval of optimization $(0, \infty)$. Since we do not have endpoints to test in $A(r)$ we consider what happens to $A(r)$ as r approaches the endpoints of $(0, \infty)$. We see that

$$A(r) \to \infty \text{ as } r \to \infty \qquad \text{(because of the } r^2 \text{ term) and}$$

$$A(r) \to \infty \text{ as } r \to 0 \qquad \text{(because of the } \frac{16}{r} \text{ term)}$$

Thus, the surface area must be minimized at the critical value we found. Finally, we determine the height of the cylinder.

$$h = \frac{8}{\pi r^2} = \frac{8}{\pi} r^{-2} = 2\left(\frac{4}{\pi}\right)\left(\frac{4}{\pi}\right)^{-2/3} = 2\left(\frac{4}{\pi}\right)^{1/3} \approx 2.17 \text{ ft.}$$

Notice that the height is twice the length of the radius. This means that the surface area is minimized when the can is as tall as it is wide.

In the exercises you will see a variety of situations that require you to combine problem–solving skills with calculus. Focus on the *process*; learn how to form equations from situations that can be manipulated into what you need. Eschew memorizing how to do "this kind of problem" as opposed to "that kind of problem." Learning a process will benefit one far longer than memorizing a specific technique.

The next section introduces another application of the derivative: *differentials*. Given $y = f(x)$, they offer a method of approximating the change in y after x changes by a small amount.

Notes:

Exercises 4.2

Terms and Concepts

1. T/F: An "optimization problem" is essentially an "extreme values" problem in a "story problem" setting.
2. T/F: This section teaches one to find the extreme values of function that have more than one variable.

Problems

3. Find the maximum product of two numbers (not necessarily integers) that have a sum of 100.
4. Find the minimum sum of two positive numbers whose product is 500.
5. Find the maximum sum of two positive numbers whose product is 500.
6. Find the maximum sum of two numbers, each of which is in $[0, 300]$ whose product is 500.
7. Find the maximal area of a right triangle with hypotenuse of length 1.
8. A rancher has 1000 feet of fencing in which to construct adjacent, equally sized rectangular pens. What dimensions should these pens have to maximize the enclosed area?
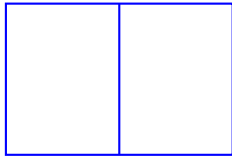
9. A standard soda can is roughly cylindrical and holds 355cm^3 of liquid. What dimensions should the cylinder be to minimize the material needed to produce the can? Based on your dimensions, determine whether or not the standard can is produced to minimize the material costs.
10. Find the dimensions of a cylindrical can with a volume of 206in^3 that minimizes the surface area.
    The "#10 can"is a standard sized can used by the restaurant industry that holds about 206in^3 with a diameter of 6 2/16in and height of 7in. Does it seem these dimensions where chosen with minimization in mind?
11. The United States Postal Service charges more for boxes whose combined length and girth exceeds 108" (the "length" of a package is the length of its longest side; the girth is the perimeter of the cross section, i.e., $2w + 2h$).
    What is the maximum volume of a package with a square cross section ($w = h$) that does not exceed the 108" standard?
12. The strength S of a wooden beam is directly proportional to its cross sectional width w and the square of its height h; that is, $S = kwh^2$ for some constant k.
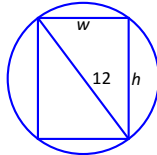
    Given a circular log with diameter of 12 inches, what sized beam can be cut from the log with maximum strength?

13. A power line is to be run to an offshore facility in the manner described in Example 3. The offshore facility is 2 miles at sea and 5 miles along the shoreline from the power plant. It costs $50,000 per mile to lay a power line underground and $80,000 to run the line underwater.
    How much of the power line should be run underground to minimize the overall costs?
14. A power line is to be run to an offshore facility in the manner described in Example 3. The offshore facility is 5 miles at sea and 2 miles along the shoreline from the power plant. It costs $50,000 per mile to lay a power line underground and $80,000 to run the line underwater.
    How much of the power line should be run underground to minimize the overall costs?
15. A woman throws a stick into a lake for her dog to fetch; the stick is 20 feet down the shore line and 15 feet into the water from there. The dog may jump directly into the water and swim, or run along the shore line to get closer to the stick before swimming. The dog runs about 22ft/s and swims about 1.5ft/s.
    How far along the shore should the dog run to minimize the time it takes to get to the stick? (Hint: the figure from Example 3 can be useful.)
16. A woman throws a stick into a lake for her dog to fetch; the stick is 15 feet down the shore line and 30 feet into the water from there. The dog may jump directly into the water and swim, or run along the shore line to get closer to the stick before swimming. The dog runs about 22ft/s and swims about 1.5ft/s.
    How far along the shore should the dog run to minimize the time it takes to get to the stick? *(Google "calculus dog" to learn more about a dog's ability to minimize times.)*
17. What are the dimensions of the rectangle with largest area that can be drawn inside the unit circle?
18. Four squares are going to be cut from a larger square piece of paper of side length 10 inches. After the paper is folded into a topless box, what is the largest volume the box could have?
19. The material to make the sides of a box costs 2 ¢/in^2. Making the bottom costs 4 ¢/in^2, while the top costs 1 ¢/in^2. What are the dimensions of the least expensive box with a square base and a volume of 10 in^3?
20. A box needs to have a surface area of 12 in^2 and be twice as long as it is wide. What is the largest volume the box can have?

4.3 Differentials

In Section 2.2 we explored the meaning and use of the derivative. This section starts by revisiting some of those ideas.

Recall that the derivative of a function f can be used to find the slopes of lines tangent to the graph of f. At $x = c$, the tangent line to the graph of f has equation

$$y = f'(c)(x - c) + f(c).$$

The tangent line can be used to find good approximations of $f(x)$ for values of x near c. (This tangent line is also called the *linearization* of f at c.)

For instance, we can approximate $\sin 1.1$ using the tangent line to the graph of $f(x) = \sin x$ at $x = \pi/3 \approx 1.05$. Recall that $\sin(\pi/3) = \sqrt{3}/2 \approx 0.866$, and $\cos(\pi/3) = 1/2$. Thus the tangent line to $f(x) = \sin x$ at $x = \pi/3$ is:

$$\ell(x) = \frac{1}{2}(x - \pi/3) + 0.866.$$

In Figure 4.8(a), we see a graph of $f(x) = \sin x$ graphed along with its tangent line at $x = \pi/3$. The small rectangle shows the region that is displayed in Figure 4.8(b). In this figure, we see how we are approximating $\sin 1.1$ with the tangent line, evaluated at 1.1. Together, the two figures show how close these values are.

Using this line to approximate $\sin 1.1$, we have:

$$\ell(1.1) = \frac{1}{2}(1.1 - \pi/3) + 0.866$$

$$= \frac{1}{2}(0.053) + 0.866 = 0.8925.$$

(We leave it to the reader to see how good of an approximation this is.)

We now generalize this concept. Given $f(x)$ and an x–value c, the tangent line is $\ell(x) = f'(c)(x - c) + f(c)$. Clearly, $f(c) = \ell(c)$. Let Δx be a small number, representing a small change in x value. We assert that:

$$f(c + \Delta x) \approx \ell(c + \Delta x),$$

since the tangent line to a function approximates well the values of that function near $x = c$.

As the x value changes from c to $c + \Delta x$, the y value of f changes from $f(c)$ to $f(c + \Delta x)$. We call this change of y value Δy. That is:

$$\Delta y = f(c + \Delta x) - f(c).$$

(a)

(b)
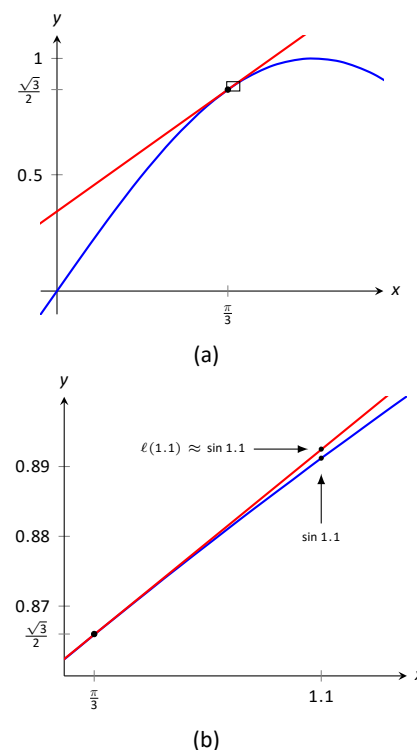
Figure 4.8: Graphing $f(x) = \sin x$ and its tangent line at $x = \pi/3$ in order to estimate $\sin 1.1$.

Notes:

Replacing $f(c + \Delta x)$ with its tangent line approximation, we have

$$\Delta y \approx \ell(c + \Delta x) - f(c)$$
$$= f'(c)\big((c + \Delta x) - c\big) + f(c) - f(c)$$
$$= f'(c)\Delta x \tag{4.3}$$

This final equation is important; we'll come back to it in Key Idea 9.

We introduce two new variables, dx and dy in the context of a formal definition.

Definition 20   Differentials of x and y.

Let $y = f(x)$ be differentiable. The **differential of x**, denoted dx, is any nonzero real number (usually taken to be a small number). The **differential of y**, denoted dy, is

$$dy = f'(x)dx.$$

We can solve for $f'(x)$ in the above equation: $f'(x) = dy/dx$. This states that the derivative of f with respect to x is the differential of y divided by the differential of x; this is **not** the alternate notation for the derivative, $\frac{dy}{dx}$. This latter notation was chosen because of the fraction–like qualities of the derivative, but again, it is one symbol and not a fraction.

It is helpful to organize our new concepts and notations in one place and an accompanying figure.
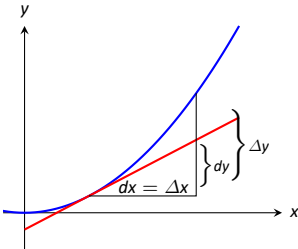
Figure 4.9: The distances involved in Key Idea 9.

Key Idea 9   Differential Notation

Let $y = f(x)$ be a differentiable function.

1. Δx represents a small, nonzero change in x value.

2. dx represents a small, nonzero change in x value (i.e., $\Delta x = dx$).

3. Δy is the change in y value as x changes by Δx; hence

    $$\Delta y = f(x + \Delta x) - f(x).$$

4. $dy = f'(x)dx$ which, by Equation (4.3), is an *approximation* of the change in y value as x changes by Δx; $dy \approx \Delta y$.

Notes:

What is the value of differentials? Like many mathematical concepts, differentials provide both practical and theoretical benefits. We explore both here.

Watch the video:
Differentials 2 at
`https://youtu.be/AvM8-LUdg84`

Example 1   **Finding and using differentials**
Consider $f(x) = x^2$. Knowing $f(3) = 9$, approximate $f(3.1)$.

SOLUTION   The x value is changing from $x = 3$ to $x = 3.1$; therefore, we see that $dx = 0.1$. If we know how much the y value changes from $f(3)$ to $f(3.1)$ (i.e., if we know Δy), we will know exactly what $f(3.1)$ is (since we already know $f(3)$). We can approximate Δy with dy.

$$\Delta y \approx dy$$
$$= f'(3)dx$$
$$= 2 \cdot 3 \cdot 0.1 = 0.6.$$

We expect the y value to change by about 0.6, so we approximate $f(3.1) \approx 9.6$.

We leave it to the reader to verify this, but the preceding discussion links the differential to the tangent line of $f(x)$ at $x = 3$. One can verify that the tangent line, evaluated at $x = 3.1$, also gives $y = 9.6$.

Of course, it is easy to compute the actual answer (by hand or with a calculator): $3.1^2 = 9.61$. (Before we get too cynical and say "Then why bother?", note our approximation is *really* good!)

So why bother?

In "most" real life situations, we do not know the function that describes a particular behavior. Instead, we can only take measurements of how things change – measurements of the derivative.

Imagine water flowing down a winding channel. It is easy to measure the speed and direction (i.e., the *velocity*) of water at any location. It is very hard to create a function that describes the overall flow, hence it is hard to predict where a floating object placed at the beginning of the channel will end up. However, we can *approximate* the path of an object using differentials. Over small

Notes:

intervals, the path taken by a floating object is essentially linear. Differentials allow us to approximate the true path by piecing together lots of short, linear paths. This technique is called Euler's Method, studied in introductory Differential Equations courses.

We use differentials once more to approximate the value of a function. Even though calculators are very accessible, it is neat to see how these techniques can sometimes be used to easily compute something that looks rather hard.

Example 2   Using differentials to approximate a function value
Approximate $\sqrt{4.5}$.

SOLUTION   We expect $\sqrt{4.5} \approx 2$, yet we can do better. Let $f(x) = \sqrt{x}$, and let $c = 4$. Thus $f(4) = 2$. We can compute $f'(x) = 1/(2\sqrt{x})$, so $f'(4) = 1/4$.

We approximate the difference between $f(4.5)$ and $f(4)$ using differentials, with $dx = 0.5$:

$$f(4.5) - f(4) = \Delta y \approx dy = f'(4) \cdot dx = 1/4 \cdot 1/2 = 1/8 = 0.125.$$

The approximate change in f from $x = 4$ to $x = 4.5$ is 0.125, so we approximate $\sqrt{4.5} \approx 2.125$.

Differentials are important when we discuss *integration*. When we study that topic, we will use notation such as

$$\int f(x)\, dx$$

quite often. While we don't discuss here what all of that notation means, note the existence of the differential dx. Proper handling of *integrals* comes with proper handling of differentials.

In light of that, we practice finding differentials in general.

Example 3   Finding differentials
In each of the following, find the differential dy.

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

SOLUTION

1. $y = \sin x$:   As $f(x) = \sin x$, $f'(x) = \cos x$. Thus

$$dy = \cos(x)dx.$$

Notes:

2. $y = e^x(x^2 + 2)$:   Let $f(x) = e^x(x^2 + 2)$. We need $f'(x)$, requiring the Product Rule.

   We have $f'(x) = e^x(x^2 + 2) + 2xe^x$, so

   $$dy = \left(e^x(x^2 + 2) + 2xe^x\right)dx.$$

3. $y = \sqrt{x^2 + 3x - 1}$:   Let $f(x) = \sqrt{x^2 + 3x - 1}$; we need $f'(x)$, requiring the Chain Rule.

   We have $f'(x) = \frac{1}{2}(x^2 + 3x - 1)^{-\frac{1}{2}}(2x + 3) = \frac{2x + 3}{2\sqrt{x^2 + 3x - 1}}$. Thus

   $$dy = \frac{(2x + 3)dx}{2\sqrt{x^2 + 3x - 1}}.$$

Finding the differential dy of $y = f(x)$ is really no harder than finding the derivative of f; we just *multiply $f'(x)$* by dx. It is important to remember that we are not simply adding the symbol "dx" at the end.

We have seen a practical use of differentials as they offer a good method of making certain approximations. Another use is *error propagation.* Suppose a length is measured to be x, although the actual value is $x + \Delta x$ (where we hope Δx is small). This measurement of x may be used to compute some other value; we can think of this as $f(x)$ for some function f. As the true length is $x + \Delta x$, one really should have computed $f(x + \Delta x)$. The difference between $f(x)$ and $f(x + \Delta x)$ is the propagated error.

How close are $f(x)$ and $f(x + \Delta x)$? This is a difference in "y" values;

$$f(x + \Delta x) - f(x) = \Delta y \approx dy.$$

We can approximate the propagated error using differentials.

Example 4   **Using differentials to approximate propagated error**
A steel ball bearing is to be manufactured with a diameter of 2cm. The manufacturing process has a tolerance of ± 0.1mm in the diameter. Given that the density of steel is about 7.85g/cm^3, estimate the propagated error in the mass of the ball bearing.

SOLUTION   The mass of a ball bearing is found using the equation "mass = volume $\times$ density." In this situation the mass function is a product of the radius of the ball bearing, hence it is $m = 7.85\frac{4}{3}\pi r^3$. The differential of the mass is

$$dm = 31.4\pi r^2 dr.$$

Notes:

The radius is to be 1cm; the manufacturing tolerance in the radius is ± 0.05mm, or ± 0.005cm. The propagated error is approximately:

$$\Delta m \approx dm$$
$$= 31.4\pi(1)^2(\pm 0.005)$$
$$= \pm 0.493\text{g}$$

Is this error significant? It certainly depends on the application, but we can get an idea by computing the *relative error*. The ratio between amount of error to the total mass is

$$\frac{dm}{m} = \pm\frac{0.493}{7.85\frac{4}{3}\pi}$$
$$= \pm\frac{0.493}{32.88}$$
$$= \pm 0.015,$$

or $\pm 1.5\%$.

We leave it to the reader to confirm this, but if the diameter of the ball was supposed to be 10cm, the same manufacturing tolerance would give a propagated error in mass of ± 12.33g, which corresponds to a *percent error* of $\pm 0.188\%$. While the amount of error is much greater ($12.33 > 0.493$), the percent error is much lower.

Notes:

Exercises 4.3

Terms and Concepts

1. T/F: Given a differentiable function $y = f(x)$, we are generally free to choose a value for dx, which then determines the value of dy.

2. T/F: The symbols "dx" and "Δx" represent the same concept.

3. T/F: The symbols "dy" and "Δy" represent the same concept.

4. T/F: Differentials are important in the study of integration.

5. How are differentials and tangent lines related?

Problems

In Exercises 6–16, use differentials to approximate the given value by hand.

6. 2.05^2

7. 5.93^2

8. 5.1^3

9. 6.8^3

10. $\sqrt{16.5}$

11. $\sqrt{24}$

12. $\sqrt[3]{63}$

13. $\sqrt[3]{8.5}$

14. $\sin 3$

15. $\cos 1.5$

16. $e^{0.1}$

In Exercises 17–29, compute the differential dy.

17. $y = x^2 + 3x - 5$

18. $y = x^7 - x^5$

19. $y = \dfrac{1}{4x^2}$

20. $y = (2x + \sin x)^2$

21. $y = x^2 e^{3x}$

22. $y = \dfrac{4}{x^4}$

23. $y = \dfrac{2x}{\tan x + 1}$

24. $y = \ln(5x)$

25. $y = e^x \sin x$

26. $y = \cos(\sin x)$

27. $y = \dfrac{x+1}{x+2}$

28. $y = 3^x \ln x$

29. $y = x \ln x - x$

30. Show that the linearization of $f(x) = (1 + x)^k$ at $x = 0$ is $L(x) = 1 + kx$.

In Exercises 31–36, use the linear approximation $(1 + x)^k \approx 1 + kx$ to find an approximation for the function $f(x)$ for values of x near zero.

31. $f(x) = (1 - x)^6$

32. $f(x) = \dfrac{2}{1 - x}$

33. $f(x) = \dfrac{1}{\sqrt{1 + x}}$

34. $f(x) = \sqrt{2 + x^2}$

35. $f(x) = (4 + 3x)^{1/3}$

36. $f(x) = \sqrt[3]{\left(1 - \dfrac{1}{2 + x}\right)^2}$

37. A set of plastic spheres are to be made with a diameter of 1cm. If the manufacturing process is accurate to 1mm, what is the propagated error in volume of the spheres?

38. The distance, in feet, a stone drops in t seconds is given by $d(t) = 16t^2$. The depth of a hole is to be approximated by dropping a rock and listening for it to hit the bottom. What is the propagated error if the time measurement is accurate to $2/10^{ths}$ of a second and the measured time is:

    (a) 2 seconds?

    (b) 5 seconds?

39. What is the propagated error in the measurement of the cross sectional area of a circular log if the diameter is measured at $15''$, accurate to $1/4''$?

40. A wall is to be painted that is exactly $8'$ high and is measured to be $10'$, $7''$ long. Find the propagated error in the measurement of the wall's surface area if the measurement of the length is accurate to $1/2''$.

Exercises 41–45 explore some issues related to surveying in which distances are approximated using other measured distances and measured angles. *(Hint: Convert all angles to radians before computing.)*

41. The length l of a long wall is to be approximated. The angle θ, as shown in the diagram (not to scale), is measured to be $85.2°$, accurate to $1°$. Assume that the triangle formed is a right triangle.
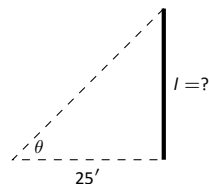
    (a) What is the measured length l of the wall?

    (b) What is the propagated error?

    (c) What is the percent error?

42. Answer the questions of Exercise 41, but with a measured angle of $71.5°$, accurate to $1°$, measured from a point $100'$ from the wall.

43. The length *l* of a long wall is to be calculated by measuring the angle θ shown in the diagram (not to scale). Assume the formed triangle is an isosceles triangle. The measured angle is $143°$, accurate to $1°$.
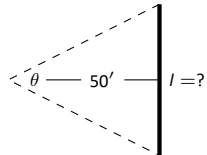
(a) What is the measured length of the wall?

(b) What is the propagated error?

(c) What is the percent error?

44. The length of the walls in Exercises 41 – 43 are essentially the same. Which setup gives the most accurate result?

45. Consider the setup in Exercise 43. This time, assume the angle measurement of $143°$ is exact but the measured $50'$ from the wall is accurate to $6''$. What is the approximate percent error?

4.4 Newton's Method

Solving equations is one of the most important things we do in mathematics, yet we are surprisingly limited in what we can solve analytically. For instance, equations as simple as $x^5 + x + 1 = 0$ or $\cos x = x$ cannot be solved by algebraic methods in terms of familiar functions. Fortunately, there are methods that can give us *approximate* solutions to equations like these. These methods can usually give an approximation correct to as many decimal places as we like. In Section 1.6 we learned about the Bisection Method. This section focuses on another technique (which generally works faster), called Newton's Method.

Newton's Method is built around tangent lines. The main idea is that if x is sufficiently close to a root of $f(x)$, then the tangent line to the graph at $(x, f(x))$ will cross the x-axis at a point closer to the root than x.

We start Newton's Method with an initial guess about roughly where the root is. Call this x_0. (See Figure 4.10(a).) Draw the tangent line to the graph at $(x_0, f(x_0))$ and see where it meets the x-axis. Call this point x_1. Then repeat the process – draw the tangent line to the graph at $(x_1, f(x_1))$ and see where it meets the x-axis. (See Figure 4.10(b).) Call this point x_2. Repeat the process again to get x_3, x_4, etc. This sequence of points will often converge rather quickly to a root of f.

We can use this *geometric* process to create an *algebraic* process. Let's look at how we found x_1. We started with the tangent line to the graph at $(x_0, f(x_0))$. The slope of this tangent line is $f'(x_0)$ and the equation of the line is

$$y = f'(x_0)(x - x_0) + f(x_0).$$

This line crosses the x-axis when $y = 0$, and the x–value where it crosses is what we called x_1. So let $y = 0$ and replace x with x_1, giving the equation:

$$0 = f'(x_0)(x_1 - x_0) + f(x_0).$$

Now solve for x_1:

$$x_1 = x_0 - \frac{f(x_0)}{f'(x_0)}.$$

Since we repeat the same geometric process to find x_2 from x_1, we have

$$x_2 = x_1 - \frac{f(x_1)}{f'(x_1)}.$$

In general, given an approximation x_n, we can find the next approximation, x_{n+1} as follows:

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}.$$

We summarize this process as follows.

(a)

(b)

(c)
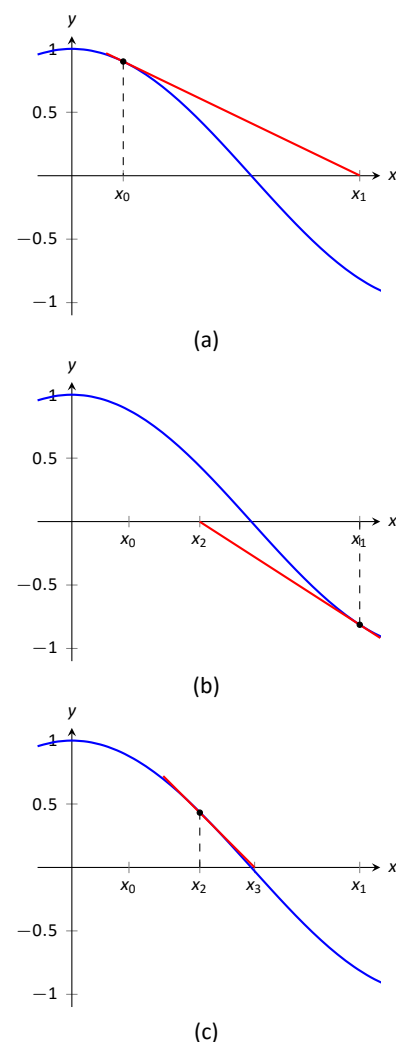
Figure 4.10: Demonstrating the geometric concept behind Newton's Method.

Notes:

> **Key Idea 10   Newton's Method**
>
> Let f be a differentiable function on an interval I with a root in I. To approximate the value of the root, accurate to d decimal places:
>
> 1. Choose a value x_0 as an initial approximation of the root. (This is often done by looking at a graph of f.)
>
> 2. Create successive approximations iteratively; given an approximation x_n, compute the next approximation x_{n+1} as
>
> $$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}.$$
>
> 3. Stop the iterations when successive approximations do not differ in the first d places after the decimal point.

Note: Newton's Method is not infallible. The sequence of approximate values may not converge, or it may converge so slowly that one is "tricked" into thinking a certain approximation is better than it actually is. These issues will be discussed at the end of the section.

Watch the video:
Newton's Method at
`https://youtu.be/1uN8cBGVpfs`

Let's practice Newton's Method with a concrete example.

Example 1   Using Newton's Method
Approximate the real root of $x^3 - x^2 - 1 = 0$, accurate to the first 3 places after the decimal, using Newton's Method and an initial approximation of $x_0 = 1$.

SOLUTION   To begin, we compute $f'(x) = 3x^2 - 2x$. Then we apply the Newton's Method algorithm, outlined in Key Idea 10.

$$x_1 = 1 - \frac{f(1)}{f'(1)} = 1 - \frac{1^3 - 1^2 - 1}{3 \cdot 1^2 - 2 \cdot 1} = 2,$$

$$x_2 = 2 - \frac{f(2)}{f'(2)} = 2 - \frac{2^3 - 2^2 - 1}{3 \cdot 2^2 - 2 \cdot 2} = 1.625,$$

Notes:

$$x_3 = 1.625 - \frac{f(1.625)}{f'(1.625)} = 1.625 - \frac{1.625^3 - 1.625^2 - 1}{3 \cdot 1.625^2 - 2 \cdot 1.625} \approx 1.48579,$$

$$x_4 = 1.48579 - \frac{f(1.48579)}{f'(1.48579)} \approx 1.46596,$$

$$x_5 = 1.46596 - \frac{f(1.46596)}{f'(1.46596)} \approx 1.46557$$

We performed 5 iterations of Newton's Method to find a root accurate to the first 3 places after the decimal; our final approximation is 1.465. The exact value of the root, to six decimal places, is 1.465571; It turns out that our x_5 is accurate to more than just 3 decimal places.

A graph of $f(x)$ is given in Figure 4.11. We can see from the graph that our initial approximation of $x_0 = 1$ was not particularly accurate; a closer guess would have been $x_0 = 1.5$. Our choice was based on ease of initial calculation, and shows that Newton's Method can be robust enough that we do not have to make a very accurate initial approximation.

We can automate this process on a calculator that has an `Ans` key that returns the result of the previous calculation. Start by pressing 1 and then `Enter`. (We have just entered our initial guess, $x_0 = 1$.) Now compute

$$\texttt{Ans} - \frac{f(\texttt{Ans})}{f'(\texttt{Ans})}$$

by entering the following and repeatedly press the `Enter` key:

```
Ans-(Ans^3-Ans^2-1)/(3*Ans^2-2*Ans)
```

Each time we press the `Enter` key, we are finding the successive approximations, x_1, x_2, ..., and each one is getting closer to the root. In fact, once we get past around x_7 or so, the approximations don't appear to be changing. They actually are changing, but the change is far enough to the right of the decimal point that it doesn't show up on the calculator's display. When this happens, we can be pretty confident that we have found an accurate approximation.

We can use a similar approach in most spreadsheet programs, which intelligently copy formulas. Start by entering 1 in cell A1. Then in cell A2, enter:

```
A1-(A1^3-A1^2-1)/(3*A1^2-2*A1)
```

Copy this cell, and paste it into A3. The spreadsheet will automatically change A1 to A2, giving you the next approximation. Continue pasting this into A4, A5, and so on. Each time we paste the formula, we are finding the successive approximations, and each one is getting closer to the root.

Using a calculator in this manner makes the calculations simple; many iterations can be computed very quickly.
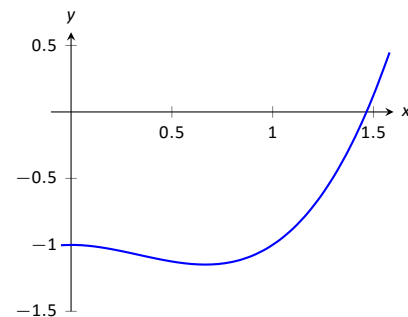
Figure 4.11: A graph of $f(x) = x^3 - x^2 - 1$ in Example 1.

Notes:
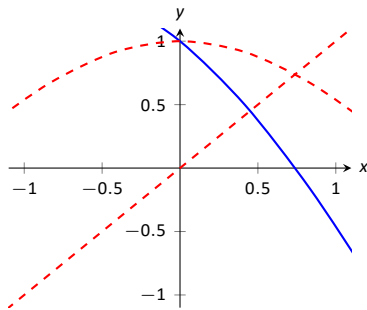
Example 2 **Using Newton's Method to find where functions intersect**

Use Newton's Method to approximate a solution to $\cos x = x$, accurate to 5 places after the decimal.

SOLUTION Newton's Method provides a method of solving $f(x) = 0$; it is not (directly) a method for solving equations like $f(x) = g(x)$. However, this is not a problem; we can rewrite the latter equation as $f(x) - g(x) = 0$ and then use Newton's Method.

So we rewrite $\cos x = x$ as $\cos x - x = 0$. Written this way, we are finding a root of $f(x) = \cos x - x$. We compute $f'(x) = -\sin x - 1$. Next we need a starting value, x_0. Consider Figure 4.12, where $f(x) = \cos x - x$ is graphed. It seems that $x_0 = 0.75$ is pretty close to the root, so we will use that as our x_0. (The figure also shows the graphs of $y = \cos x$ and $y = x$, drawn with dashed lines. Note how they intersect at the same x value as when $f(x) = 0$.)

We now compute x_1, x_2, etc. The formula for x_1 is

$$x_1 = 0.75 - \frac{\cos(0.75) - 0.75}{-\sin(0.75) - 1} \approx 0.7391111388.$$

Apply Newton's Method again to find x_2:

$$x_2 = 0.7391111388 - \frac{\cos(0.7391111388) - 0.7391111388}{-\sin(0.7391111388) - 1} \approx 0.7390851334.$$

We can continue this way, but it is really best to automate this process. On a calculator with an Ans key, we would start by pressing 0.75, then Enter, inputting our initial approximation. We then enter:

Ans - (cos(Ans)-Ans)/(-sin(Ans)-1).

(In a spreadsheet, we would enter A1-(cos(A1)-A1)/(-sin(A1)-1) in A2.)

Repeatedly pressing the Enter key gives successive approximations. We quickly find:

$$x_3 = 0.7390851332$$
$$x_4 = 0.7390851332.$$

Our approximations x_2 and x_3 did not differ for at least the first 5 places after the decimal, so we could have stopped. However, using our calculator in the manner described is easy, so finding x_4 was not hard. It is interesting to see how we found an approximation, accurate to as many decimal places as our calculator displays, in just 4 iterations.

Figure 4.12: A graph of $f(x) = \cos x - x$ used to find an initial approximation of its root.

Notes:

If you know how to program, you can translate the following pseudocode into your favorite language to perform the computation in this problem.

```
x = .75
while true
    oldx = x
    x = x - (cos(x)-x)/(-sin(x)-1)
    print x
    if abs(x-oldx) < .0000000001
        break
```

This code calculates x_1, x_2, etc., storing each result in the variable x. The previous approximation is stored in the variable oldx. We continue looping until the difference between two successive approximations, abs(x-oldx), is less than some small tolerance, in this case, .0000000001.

Figure 4.13: A graph of $f(x) = x^3 - x^2 - 1$, showing why an initial approximation of $x_0 = 0$ with Newton's Method fails.

Convergence of Newton's Method

What should one use for the initial guess, x_0? Generally, the closer to the actual root the initial guess is, the better. However, some initial guesses should be avoided. For instance, consider Example 1 where we sought the root to $f(x) = x^3 - x^2 - 1$. Choosing $x_0 = 0$ would have been a particularly poor choice. Consider Figure 4.13, where $f(x)$ is graphed along with its tangent line at $x = 0$. Since $f'(0) = 0$, the tangent line is horizontal and does not intersect the x–axis. Graphically, we see that Newton's Method fails.

We can also see analytically that it fails. Since

$$x_1 = 0 - \frac{f(0)}{f'(0)}$$

and $f'(0) = 0$, we see that x_1 is not well defined.

This problem can also occur if, for instance, it turns out that $f'(x_5) = 0$. Adjusting the initial approximation x_0 by a very small amount will likely fix the problem.

It is also possible for Newton's Method to not converge while each successive approximation is well defined. Consider $f(x) = x^{1/3}$, as shown in Figure 4.14. It is clear that the root is $x = 0$, but let's approximate this with $x_0 = 0.1$. Figure 4.14(a) shows graphically the calculation of x_1; notice how it is farther from the root than x_0. Figures 4.14(b) and (c) show the calculation of x_2 and x_3, which are even farther away; our successive approximations are getting worse. (It turns out that in this particular example, each successive approximation is twice as far from the true answer as the previous approximation.)

There is no "fix" to this problem; Newton's Method simply will not work and another method must be used.

Notes:

(a)

(b)

(c)

Figure 4.14: Newton's Method fails to find a root of $f(x) = x^{1/3}$, regardless of the choice of x_0.

While Newton's Method does not always work, it does work "most of the time," and it is generally very fast. Once the approximations get close to the root, Newton's Method can as much as double the number of correct decimal places with each successive approximation. A course in Numerical Analysis will introduce the reader to more iterative root finding methods, as well as give greater detail about the strengths and weaknesses of Newton's Method.

We first learned of the derivative in the context of instantaneous rates of change and slopes of tangent lines. We furthered our understanding of the power of the derivative by studying how it relates to the graph of a function (leading to ideas of increasing/decreasing and concavity). This chapter has put the derivative to yet more uses:

- Related Rates (furthering our use of the derivative to find instantaneous rates of change)

- Optimization (applied extreme values), and

- Differentials (useful for various approximations and for something called integration).

- Equation solving (Newton's Method)

In the next chapters, we will consider the "reverse" problem to computing the derivative: given a function f, can we find a function whose derivative is f? Be able to do so opens up an incredible world of mathematics and applications.

Notes:

Exercises 4.4

Terms and Concepts

1. T/F: Given a function $f(x)$, Newton's Method produces an exact solution to $f(x) = 0$.

2. T/F: In order to get a solution to $f(x) = 0$ accurate to d places after the decimal, at least $d + 1$ iterations of Newtons' Method must be used.

Problems

In Exercises 3–7, the roots of $f(x)$ are known or are easily found. Use 5 iterations of Newton's Method with the given initial approximation to approximate the root. Compare it to the known value of the root.

3. $f(x) = \cos x, x_0 = 1.5$

4. $f(x) = \sin x, x_0 = 1$

5. $f(x) = x^2 + x - 2, x_0 = 0$

6. $f(x) = x^2 - 2, x_0 = 1.5$

7. $f(x) = \ln x, x_0 = 2$

In Exercises 8–11, use Newton's Method to approximate all roots of the given functions accurate to 3 places after the decimal. If an interval is given, find only the roots that lie in that interval. Use technology to obtain good initial approximations.

8. $f(x) = x^3 + 5x^2 - x - 1$

9. $f(x) = x^4 + 2x^3 - 7x^2 - x + 5$

10. $f(x) = x^{17} - 2x^{13} - 10x^8 + 10$ on $(-2, 2)$

11. $f(x) = x^2 \cos x + (x - 1) \sin x$ on $(-3, 3)$

In Exercises 12–15, use Newton's Method to approximate when the given functions are equal, accurate to 3 places after the decimal. Use technology to obtain good initial approximations.

12. $f(x) = x^2, g(x) = \cos x$

13. $f(x) = x^2 - 1, g(x) = \sin x$

14. $f(x) = e^{x^2}, g(x) = \cos x$

15. $f(x) = x, g(x) = \tan x$ on $[-6, 6]$

16. Why does Newton's Method fail in finding a root of $f(x) = x^3 - 3x^2 + x + 3$ when $x_0 = 1$?

17. Why does Newton's Method fail in finding a root of $f(x) = -17x^4 + 130x^3 - 301x^2 + 156x + 156$ when $x_0 = 1$?

In Exercises 18–21, use Newton's Method to approximate the given value.

18. $\sqrt{16.5}$.

19. $\sqrt{24}$.

20. $\sqrt[3]{63}$.

21. $\sqrt[3]{8.5}$.

22. Show graphically what happens when Newton's Method is used at different x_0 for the function shown. (a) $x_0 = 0$ (b) $x_0 = 1$ (c) $x_0 = 3$ (d) $x_0 = 4$ (e) $x_0 = 5$
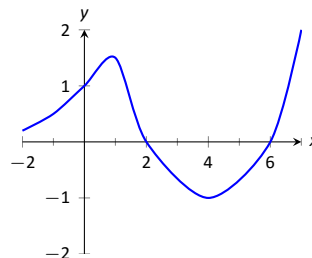
5: INTEGRATION

We have spent considerable time considering the derivatives of a function and their applications. In the following chapters, we are going to starting thinking in "the other direction." That is, given a function $f(x)$, we are going to consider functions $F(x)$ such that $F'(x) = f(x)$. There are numerous reasons this will prove to be useful: these functions will help us compute areas, volumes, mass, force, pressure, work, and much more.

5.1 Antiderivatives and Indefinite Integration

Given a function $y = f(x)$, a *differential equation* is one that incorporates y, x, and the derivatives of y. For instance, a simple differential equation is:

$$y' = 2x.$$

Solving a differential equation amounts to finding a function y that satisfies the given equation. Take a moment and consider that equation; can you find a function y such that $y' = 2x$?

Can you find another?

And yet another?

Hopefully one was able to come up with at least one solution: $y = x^2$. "Finding another" may have seemed impossible until one realizes that a function like $y = x^2 + 1$ also has a derivative of $2x$. Once that discovery is made, finding "yet another" is not difficult; the function $y = x^2 + 123{,}456{,}789$ also has a derivative of $2x$. The differential equation $y' = 2x$ has many solutions. This leads us to some definitions.

Definition 21   Antiderivatives

Let a function $f(x)$ be given. An **antiderivative** of $f(x)$ is a function $F(x)$ such that $F'(x) = f(x)$.

We refer to *an* antiderivative of f, as opposed to *the* antiderivative of f, since antiderivatives are not unique. We often use upper-case letters to denote antiderivatives.

> **Theorem 31   Antiderivative Forms**
> Let $F(x)$ and $G(x)$ be antiderivatives of $f(x)$ on an interval. Then there exists a constant C such that
>
> $$G(x) = F(x) + C.$$

Proof
Suppose that a and b are two distinct points in the interval. Then by applying the Mean Value Theorem to the function $G(x) - F(x)$, there is a point c between a and b so that

$$(G(b) - F(b)) - (G(a) - F(a)) = (G'(c) - F'(c))(b - a) = (f(c) - f(c))(b - a) = 0.$$

Because this holds for any a and b in the interval, $G(b) - F(b)$ is constant for all possible b. $\qquad\qquad\qquad\square$

Given a function f and one of its antiderivatives F, we know *all* antiderivatives of f have the form $F(x) + C$ for some constant C.

> **Definition 22   Indefinite Integrals**
> The set of all antiderivatives of $f(x)$ is the **indefinite integral of f**, denoted by
>
> $$\int f(x)\, dx.$$

Using Definitions 21 and 22, we can say that on an interval

$$\int f(x)\, dx = F(x) + C.$$

Let's analyze this indefinite integral notation.
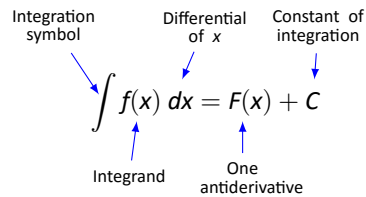
Notes:

The integration symbol, $\int$, is in reality an "elongated S," representing "take the sum." We will later see how *sums* and *antiderivatives* are related.

The function we want to find an antiderivative of is called the *integrand*. It contains the differential of the variable we are integrating with respect to. The $\int$ symbol and the differential dx are not "bookends" with a function sandwiched in between; rather, the symbol $\int$ means "find all antiderivatives of what follows," and the function $f(x)$ and dx are multiplied together; the dx does not "just sit there."

Let's practice using this notation.

Example 1   **Evaluating indefinite integrals**

Evaluate $\int \sin x\, dx$.

SOLUTION   We are asked to find all functions $F(x)$ such that $F'(x) = \sin x$. Some thought will lead us to one solution: $F(x) = -\cos x$, because $\frac{d}{dx}(-\cos x) = \sin x$.

The indefinite integral of $\sin x$ is thus $-\cos x$, plus a constant of integration. So:

$$\int \sin x\, dx = -\cos x + C.$$

A commonly asked question is "What happened to the *dx*?" The unenlightened response is "Don't worry about it. It just goes away." A full understanding includes the following.

This process of *antidifferentiation* is really solving a *differential* question. The integral

$$\int \sin x\, dx$$

presents us with a differential, $dy = \sin x\, dx$. It is asking: "What is y?" We found lots of solutions, all of the form $y = -\cos x + C$.

Letting $dy = \sin x\, dx$, rewrite

$$\int \sin x\, dx \quad \text{as} \quad \int dy.$$

This is asking: "What functions have a differential of the form *dy*?" The answer is "Functions of the form $y + C$, where C is a constant." What is y? We have lots of choices, all differing by a constant; the simplest choice is $y = -\cos x$.

Understanding all of this is more important later as we try to find antiderivatives of more complicated functions. In this section, we will simply explore the rules of indefinite integration, and one can succeed for now with answering "What happened to the *dx*?" with "It went away."

Let's practice once more before stating integration rules.

Note: Recall from Definition 20 that dx is any nonzero real number and $dy = f'(x)dx$.

Notes:

Example 2   **Evaluating indefinite integrals**

Evaluate $\int (3x^2 + 4x + 5) \, dx$.

SOLUTION   We seek a function $F(x)$ whose derivative is $3x^2 + 4x + 5$. When taking derivatives, we can consider functions term–by–term, so we can likely do that here.

What functions have a derivative of $3x^2$? Some thought will lead us to a cubic, specifically $x^3 + C_1$, where C_1 is a constant.

What functions have a derivative of $4x$? Here the x term is raised to the first power, so we likely seek a quadratic. Some thought should lead us to $2x^2 + C_2$, where C_2 is a constant.

Finally, what functions have a derivative of 5? Functions of the form $5x + C_3$, where C_3 is a constant.

Our answer appears to be

$$\int (3x^2 + 4x + 5) \, dx = x^3 + C_1 + 2x^2 + C_2 + 5x + C_3.$$

We do not need three separate constants of integration; combine them as one constant, giving the final answer of

$$\int (3x^2 + 4x + 5) \, dx = x^3 + 2x^2 + 5x + C.$$

It is easy to verify our answer; take the derivative of $x^3 + 2x^2 + 5x + C$ and see we indeed get $3x^2 + 4x + 5$.

This final step of "verifying our answer" is important both practically and theoretically. In general, taking derivatives is easier than finding antiderivatives so checking our work is easy and vital as we learn.

We also see that taking the derivative of our answer returns the function in the integrand. Thus we can say that:

$$\frac{d}{dx} \left(\int f(x) \, dx \right) = f(x).$$

Differentiation "undoes" the work done by antidifferentiation.

Theorem 21 gave a list of the derivatives of common functions we had learned at that point. We restate part of that list here to stress the relationship between derivatives and antiderivatives. This list will also be useful as a glossary of common antiderivatives as we learn.

Notes:

Theorem 32   Derivatives and Antiderivatives

Common Differentiation Rules   Common Indefinite Integral Rules

1. $\frac{d}{dx}\left(cf(x)\right) = c \cdot f'(x)$

1. $\int c \cdot f(x)\, dx = c \cdot \int f(x)\, dx$

2. $\frac{d}{dx}\left(f(x) \pm g(x)\right) = f'(x) \pm g'(x)$

2. $\int \left(f(x) \pm g(x)\right)\, dx = \int f(x)\, dx \pm \int g(x)\, dx$

3. $\frac{d}{dx}\left(C\right) = 0$

3. $\int 0\, dx = C$

We highlight a few important points from Theorem 32:

- Rule #1 states $\int c \cdot f(x)\, dx = c \cdot \int f(x)\, dx$. This is the Constant Multiple Rule: we can temporarily ignore constants when finding antiderivatives, just as we did when computing derivatives (i.e., $\frac{d}{dx}\left(3x^2\right)$ is just as easy to compute as $\frac{d}{dx}\left(x^2\right)$). An example:

$$\int 5\cos x\, dx = 5 \cdot \int \cos x\, dx = 5 \cdot (\sin x + C) = 5\sin x + C.$$

In the last step we can consider the constant as also being multiplied by 5, but "5 times a constant" is still a constant, so we just write "C".

- Rule #2 is the Sum/Difference Rule: we can split integrals apart when the integrand contains terms that are added/subtracted, as we did in Example 2. So:

$$\begin{aligned}
\int (3x^2 + 4x + 5)\, dx &= \int 3x^2\, dx + \int 4x\, dx + \int 5\, dx \\
&= 3\int x^2\, dx + 4\int x\, dx + \int 5\, dx \\
&= 3 \cdot \frac{1}{3}x^3 + 4 \cdot \frac{1}{2}x^2 + 5x + C \\
&= x^3 + 2x^2 + 5x + C
\end{aligned}$$

In practice we generally do not write out all these steps, but we demonstrate them here for completeness.

Notes:

<div style="border:1px solid green">

Theorem 33 **Derivatives and Antiderivatives**

Common Derivatives   Common Indefinite Integrals

4. $\frac{d}{dx}(x^n) = n \cdot x^{n-1}$ 4. $\int x^n\, dx = \frac{x^{n+1}}{n+1} + C$ $(n \neq -1)$

5. $\frac{d}{dx}(\ln|x|) = \frac{1}{x}$ 5. $\int \frac{1}{x}\, dx = \ln|x| + C$

6. $\frac{d}{dx}(e^x) = e^x$ 6. $\int e^x\, dx = e^x + C$

7. $\frac{d}{dx}(\sin x) = \cos x$ 7. $\int \cos x\, dx = \quad \sin x + C$

8. $\frac{d}{dx}(\cos x) = -\sin x$ 8. $\int \sin x\, dx = -\cos x + C$

9. $\frac{d}{dx}(\tan x) = \sec^2 x$ 9. $\int \sec^2 x\, dx = \quad \tan x + C$

10. $\frac{d}{dx}(\cot x) = -\csc^2 x$ 10. $\int \csc^2 x\, dx = -\cot x + C$

11. $\frac{d}{dx}(\sec x) = \sec x \tan x$ 11. $\int \sec x \tan x\, dx = \quad \sec x + C$

12. $\frac{d}{dx}(\csc x) = -\csc x \cot x$ 12. $\int \csc x \cot x\, dx = -\csc x + C$

</div>

- Rule #4 is the Power Rule of indefinite integration. There are two important things to keep in mind:

  1. Notice the restriction that $n \neq -1$. This is important: $\int \frac{1}{x}\, dx \neq$ "$\frac{1}{0}x^0 + C$"; rather, see Rule #5.

  2. We are presenting antidifferentiation as the "inverse operation" of differentiation. Here is a useful quote to remember:

     "Inverse operations do the opposite things in the opposite order."

     When taking a derivative using the Power Rule, we **first** *multiply* by the power, then **second** *subtract* 1 from the power. To find the antiderivative, do the opposite things in the opposite order: **first** *add* one to the power, then **second** *divide* by the power.

- Note that Rule #5 incorporates the absolute value of *x*. The exercises will work the reader through why this is the case; for now, know the absolute value is important and cannot be ignored.

Notes:

Initial Value Problems

In Section 2.3 we saw that the derivative of a position function gave a velocity function, and the derivative of a velocity function describes acceleration. We can now go "the other way:" the antiderivative of an acceleration function gives a velocity function, etc. While there is just one derivative of a given function, there are infinite antiderivatives. Therefore we cannot ask "What is *the* velocity of an object whose acceleration is -32ft/s^2?", since there is more than one answer.

> Watch the video:
> Antiderivatives: Acceleration, Velocity, Position Functions — A Word Problem at
> `https://youtu.be/brNADtx8Qu8`

We can find *the* answer if we provide more information with the question, as done in the following example. Often the additional information comes in the form of an *initial value*, a value of the function that one knows beforehand.

Example 3   Solving initial value problems
The acceleration due to gravity of a falling object is -32 ft/s^2. At time $t = 3$, a falling object had a velocity of -10 ft/s. Find the equation of the object's velocity.

SOLUTION   We want to know a velocity function, $v(t)$. We know two things:

- The acceleration, i.e., $v'(t) = -32$, and

- the velocity at a specific time, i.e., $v(3) = -10$.

Using the first piece of information, we know that $v(t)$ is an antiderivative of $v'(t) = -32$. So we begin by finding the indefinite integral of -32:

$$v(t) = \int (-32)\, dt = -32t + C.$$

Now we use the fact that $v(3) = -10$ to find C:

$$v(t) = -32t + C$$
$$v(3) = -10$$
$$-32(3) + C = -10$$
$$C = 86$$

Notes:

Thus $v(t) = -32t + 86$. We can use this equation to understand the motion of the object: when $t = 0$, the object had a velocity of $v(0) = 86$ ft/s. Since the velocity is positive, the object was moving upward.

When did the object begin moving down? Immediately after $v(t) = 0$:

$$-32t + 86 = 0 \quad \Rightarrow \quad t = \frac{43}{16} \approx 2.69\text{s}.$$

Recognize that we are able to determine quite a bit about the path of the object knowing just its acceleration and its velocity at a single point in time.

Example 4   **Solving initial value problems**

Find $f(t)$, given that $f''(t) = \cos t$, $f'(0) = 3$ and $f(0) = 5$.

    SOLUTION   We start by finding $f'(t)$, which is an antiderivative of $f''(t)$:

$$f'(t) = \int f''(t)\, dt = \int \cos t\, dt = \sin t + C.$$

So $f'(t) = \sin t + C$ for the correct value of C. We are given that $f'(0) = 3$, so:

$$f'(0) = 3 \quad \Rightarrow \quad \sin 0 + C = 3 \quad \Rightarrow \quad C = 3.$$

Using the initial value, we have found $f'(t) = \sin t + 3$.

We now find $f(t)$ by integrating again.

$$f(t) = \int f'(t)\, dt = \int (\sin t + 3)\, dt = -\cos t + 3t + C.$$

We are given that $f(0) = 5$, so

$$-\cos 0 + 3(0) + C = 5$$
$$-1 + C = 5$$
$$C = 6$$

Thus $f(t) = -\cos t + 3t + 6$.

This section introduced antiderivatives and the indefinite integral. We found they are needed when finding a function given information about its derivative(s). For instance, we found a position function given a velocity function.

In the next section, we will see how position and velocity are unexpectedly related by the areas of certain regions on a graph of the velocity function. Then, in Section 5.4, we will see how areas and antiderivatives are closely tied together.

Notes:

Exercises 5.1

Terms and Concepts

1. Define the term "antiderivative" in your own words.

2. Is it more accurate to refer to "the" antiderivative of $f(x)$ or "an" antiderivative of $f(x)$?

3. Use your own words to define the indefinite integral of $f(x)$.

4. Fill in the blanks: "Inverse operations do the _____ things in the _____ order."

5. What is an "initial value problem"?

6. The derivative of a position function is a _____ function.

7. The antiderivative of an acceleration function is a _____ function.

Problems

In Exercises 8–40, evaluate the given indefinite integral.

8. $\int 3x^3 \, dx$

9. $\int x^8 \, dx$

10. $\int (10x^2 - 2) \, dx$

11. $\int dt$

12. $\int \frac{1}{3t^2} \, dt$

13. $\int \frac{3}{t^2} \, dt$

14. $\int \frac{1}{\sqrt{x}} \, dx$

15. $\int \sec^2 \theta \, d\theta$

16. $\int \sin \theta \, d\theta$

17. $\int (\sec x \tan x + \csc x \cot x) \, dx$

18. $\int \frac{e^t}{2} \, dt$

19. $\int (2t + 3)^2 \, dt$

20. $\int (t^2 + 3)(t^3 - 2t) \, dt$

21. $\int x^2 x^3 \, dx$

22. $\int e^\pi \, dx$

23. $\int \frac{3}{x^4} \, dx$

24. $\int \frac{4x^5 - 7}{x^3} \, dx$

25. $\int \sqrt{x^7} \, dx$

26. $\int \frac{x^3 - 7x}{\sqrt{x}} \, dx$

27. $\int 5 - \frac{2}{3}x^2 + \frac{3}{4}x^3 \, dx$

28. $\int u^6 - 2u^5 - u^3 + \frac{2}{7} \, du$

29. $\int (u + 4)(2u + 1) \, du$

30. $\int \sqrt{t}(t^2 + 3t + 2) \, dt$

31. $\int \frac{1 + \sqrt{x} + x}{\sqrt{x}} \, dx$

32. $\int \sin^2 x + \cos^2 x \, dx$

33. $\int 2 + \tan^2 \theta \, d\theta$

34. $\int \sec t(\sec t + \tan t) \, dt$

35. $\int \frac{1 - \sin^2 t}{\sin^2 t} \, dt$

36. $\int \frac{\sin 2x}{\sin x} \, dx$

37. $\int \frac{4 + 6u}{\sqrt{u}} \, du$

38. $\int \frac{\sin \theta + \sin \theta \tan^2 \theta}{\sec^2} \, d\theta$

39. $\int \frac{2 + t}{\sqrt[3]{t^2}} \, dt$

40. $\int \sqrt[4]{x^5} + \sqrt[5]{x^4} \, dx$

41. This problem investigates why Theorem 32 states that $\int \frac{1}{x} \, dx = \ln |x| + C$.

    (a) What is the domain of $y = \ln x$?

    (b) Find $\frac{d}{dx} \left(\ln x \right)$.

    (c) What is the domain of $y = \ln(-x)$?

    (d) Find $\frac{d}{dx} \left(\ln(-x) \right)$.

    (e) You should find that $1/x$ has two types of antiderivatives, depending on whether $x > 0$ or $x < 0$. In one expression, give a formula for $\int \frac{1}{x} \, dx$ that takes these different domains into account, and explain your answer.

In Exercises 42–52, find $f(x)$ described by the given initial value problem.

42. $f'(x) = \sin x$ and $f(0) = 2$

43. $f'(x) = 5e^x$ and $f(0) = 10$

44. $f'(x) = 4x^3 - 3x^2$ and $f(-1) = 9$

45. $f'(x) = \sec^2 x$ and $f(\pi/4) = 5$

46. $f''(x) = 5$ and $f'(0) = 7, f(0) = 3$

47. $f''(x) = 7x$ and $f'(1) = -1, f(1) = 10$

48. $f''(x) = 5e^x$ and $f'(0) = 3, f(0) = 5$

49. $f''(\theta) = \sin\theta$ and $f'(\pi) = 2, f(\pi) = 4$

50. $f''(x) = 0$ and $f'(1) = 3, f(1) = 1$

51. $f'(x) = \dfrac{-2}{x^3}$ and $f(1) = 2$

52. $f'(x) = \dfrac{1}{\sqrt{x}}$ and $f(4) = 0$

53. An object is moving so that its velocity at time t is given by $v(t) = 3\sqrt{t}$. If the object was at the origin at time $t = 0$, find it's position $s(t)$ at time t.

54. A nickel dropped from the top of the North Dakota State Capital Building has acceleration $a(t) = -32$ ft/sec^2 (ignoring air resistance), initial velocity $v(0) = 0$, and initial height $s(0) = 241.67$ ft. How long will it take the nickel to hit the ground?

55. Given the graph of f below, sketch the graph of the antiderivative F of f that passes through the origin. What do the graphs of the other antiderivatives of f look like?
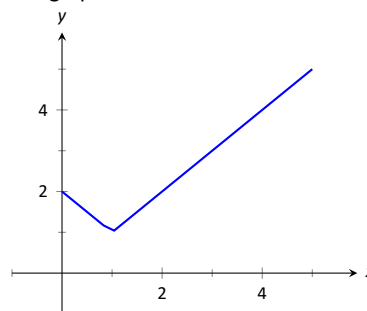
56. Given the graph of f below, sketch the graph of the antiderivative F of f that passes through the origin. What do the graphs of the other antiderivatives of f look like?
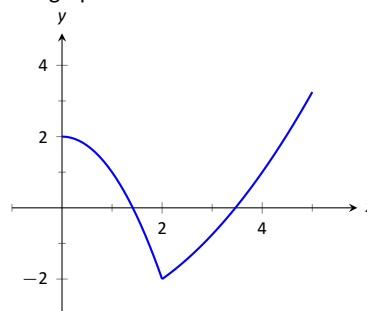
Review

57. Use information gained from the first and second derivatives to sketch $f(x) = \dfrac{1}{e^x + 1}$.

58. Given $y = x^2 e^x \cos x$, find dy.

5.2 The Definite Integral

We start with an easy problem. An object travels in a straight line at a constant velocity of 5 ft/s for 10 seconds. How far away from its starting point is the object?

We approach this problem with the familiar "Distance = Rate × Time" equation. In this case, Distance = 5ft/s × 10s = 50 feet.

It is interesting to note that this solution of 50 feet can be represented graphically. Consider Figure 5.1, where the constant velocity of 5ft/s is graphed on the axes. Shading the area under the line from $t = 0$ to $t = 10$ gives a rectangle with an area of 50 square units; when one considers the units of the axes, we can say this area represents 50 ft.

Now consider a slightly harder situation (and not particularly realistic): an object travels in a straight line with a constant velocity of 5ft/s for 10 seconds, then instantly reverses course at a rate of 2ft/s for 4 seconds. (Since the object is traveling in the opposite direction when reversing course, we say the velocity is a constant -2ft/s.) How far away from the starting point is the object – what is its *displacement*?

Here we use "Distance = Rate$_1$ × Time$_1$ + Rate$_2$ × Time$_2$," which is

$$\text{Distance } = 5 \cdot 10 + (-2) \cdot 4 = 42 \text{ ft.}$$

Hence the object is 42 feet from its starting location.

We can again depict this situation graphically. In Figure 5.2 we have the velocities graphed as straight lines on $[0, 10]$ and $[10, 14]$, respectively. The displacement of the object is

$$\text{"Area above the } t\text{–axis} \quad - \quad \text{Area below the } t\text{–axis,"}$$

which is easy to calculate as $50 - 8 = 42$ feet.

Now consider a more difficult problem.

Example 1   Finding position using velocity

The velocity of an object moving straight up/down under the acceleration of gravity is given as $v(t) = -32t + 48$, where time t is given in seconds and velocity is in ft/s. When $t = 0$, the object had a height of 0 ft.

1. What was the initial velocity of the object?

2. What was the maximum height of the object?

3. What was the height of the object at time $t = 2$?

y (ft/s)
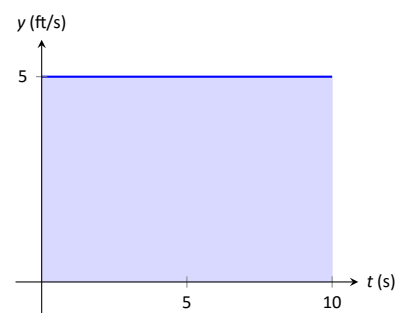
Figure 5.1: The area under a constant velocity function corresponds to distance traveled.
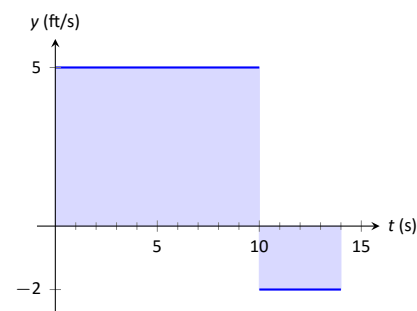
Figure 5.2: The total displacement is the area above the t–axis minus the area below the t–axis.

Notes:

SOLUTION   It is straightforward to find the initial velocity; at time $t = 0$, $v(0) = -32 \cdot 0 + 48 = 48$ ft/s.

To answer questions about the height of the object, we need to find the object's position function $s(t)$. This is an initial value problem, which we studied in the previous section. We are told the initial height is 0, i.e., $s(0) = 0$. We know $s'(t) = v(t) = -32t + 48$. To find s, we find the indefinite integral of $v(t)$:

$$s(t) = \int v(t)\, dt = \int (-32t + 48)\, dt = -16t^2 + 48t + C.$$

Since $s(0) = 0$, we conclude that $C = 0$ and $s(t) = -16t^2 + 48t$.

To find the maximum height of the object, we need to find the maximum of s. Recalling our work finding extreme values, we find the critical points of s by setting its derivative equal to 0 and solving for t:

$$s'(t) = -32t + 48 = 0 \quad \Rightarrow \quad t = 48/32 = 1.5\text{s}.$$

(Notice how we ended up just finding when the velocity was 0ft/s.) The first derivative test shows this is a maximum, so the maximum height of the object is found at

$$s(1.5) = -16(1.5)^2 + 48(1.5) = 36\text{ft}.$$

The height at time $t = 2$ is now straightforward to compute: it is $s(2) = 32$ft.

While we have answered all three questions, let's look at them again graphically, using the concepts of area that we explored earlier.

Figure 5.3 shows a graph of $v(t)$ on axes from $t = 0$ to $t = 3$. It is again straightforward to find $v(0)$. How can we use the graph to find the maximum height of the object?

Recall how in our previous work that the displacement of the object (in this case, its height) was found as the area under the velocity curve, as shaded in the figure. Moreover, the area between the curve and the t–axis that is below the t–axis counted as "negative" area. That is, it represents the object coming back toward its starting position. So to find the maximum distance from the starting point – the maximum height – we find the area under the velocity line that is above the t–axis, i.e., from $t = 0$ to $t = 1.5$. This region is a triangle; its area is

$$\text{Area } = \frac{1}{2}\text{Base} \times \text{Height} = \frac{1}{2} \times 1.5\text{s} \times 48\text{ft/s} = 36\text{ft},$$

which matches our previous calculation of the maximum height.

Finally, we find the total *signed* area under the velocity function from $t = 0$ to $t = 2$ to find the $s(2)$, the height at $t = 2$, which is a displacement, the distance from the current position to the starting position. That is,
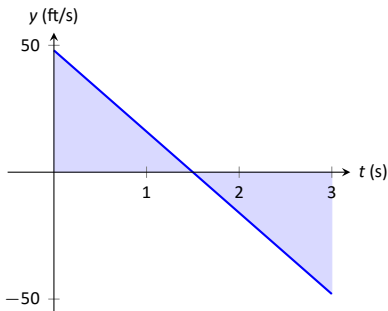
Figure 5.3: A graph of $v(t) = -32t + 48$; the shaded areas help determine displacement.

Notes:

Displacement = Area above the t–axis $-$ Area below t–axis.

The regions are triangles, and we find

$$\text{Displacement} = \frac{1}{2}(1.5\text{s})(48\text{ft/s}) - \frac{1}{2}(.5\text{s})(16\text{ft/s}) = 32\text{ft}.$$

This also matches our previous calculation of the height at $t = 2$.

Notice how we answered each question in this example in two ways. Our first method was to manipulate equations using our understanding of antiderivatives and derivatives. Our second method was geometric: we answered questions looking at a graph and finding the areas of certain regions of this graph.

The above example does not *prove* a relationship between area under a velocity function and displacement, but it does indicate that there may be a relationship. Section 5.4 will fully establish fact that the area under a velocity function is displacement.

Given a graph of a continuous function $y = f(x)$, we will find that there is great use in computing the area between the curve $y = f(x)$ and the x-axis. Because of this, we need to define some terms. The **total signed area** from $x = a$ to $x = b$ under a continuous function f is

(area under f and above the x–axis on $[a, b]$) $-$
(area above f and under the x–axis on $[a, b]$).

Definition 23   The Definite Integral
Let $y = f(x)$ be continuous on a closed interval $[a, b]$. The **definite integral of f on $[a, b]$** is the total signed area of f on $[a, b]$, denoted

$$\int_a^b f(x)\ dx,$$

where a and b are the **bounds of integration.**

By our definition, the definite integral gives the "signed area under f." We usually drop the word "signed" when talking about the definite integral, and simply say the definite integral gives "the area under f" or, more commonly, "the area under the curve."

The previous section introduced the indefinite integral, which is related to antiderivatives. We have now defined the definite integral, which relates to areas under a curve. The two are very much related, as we'll see when we learn

Notes:

the Fundamental Theorem of Calculus in Section 5.4. Recall that earlier we said that the "$\int$" symbol was an "elongated S" that represented finding a "sum." In the context of the definite integral, this notation makes a bit more sense, as we are adding up areas under the function f.

We practice using this notation.

Example 2   Evaluating definite integrals

Consider the function f given in Figure 5.4. Find:

1. $\displaystyle\int_0^3 f(x)\ dx$

2. $\displaystyle\int_3^5 f(x)\ dx$

3. $\displaystyle\int_0^5 f(x)\ dx$

4. $\displaystyle\int_0^3 5f(x)\ dx$

5. $\displaystyle\int_1^1 f(x)\ dx$

SOLUTION

1. $\int_0^3 f(x)\ dx$ is the area under f on the interval $[0, 3]$. This region is a triangle, so the area is $\int_0^3 f(x)\ dx = \frac{1}{2}(3)(1) = 1.5$.

2. $\int_3^5 f(x)\ dx$ represents the area of the triangle found under the x–axis on $[3, 5]$. The area is $\frac{1}{2}(2)(1) = 1$; since it is found *under* the x–axis, this is "negative area." Therefore $\int_3^5 f(x)\ dx = -1$.

3. $\int_0^5 f(x)\ dx$ is the total signed area under f on $[0, 5]$. This is $1.5 + (-1) = 0.5$.

4. $\int_0^3 5f(x)\ dx$ is the area under $5f$ on $[0, 3]$. This is sketched in Figure 5.5. Again, the region is a triangle, with height 5 times that of the height of the original triangle. Thus the area is $\int_0^3 5f(x)\ dx = 15/2 = 7.5$.

5. $\int_1^1 f(x)\ dx$ is the area under f on the "interval" $[1, 1]$. This describes a line segment, not a region; it has no width. Therefore the area is 0.

This example illustrates some of the properties of the definite integral, given in Theorem 34.

So far, when we have computed a definite integral $\int_a^b f(x)\ dx$, we have required that $a \leq b$. In practice, it is sometimes convenient to be able to compute $\int_a^b f(x)\ dx$ for $a > b$. To do so, we introduce the convention that for any a and b, $\int_a^b f(x)\ dx = -\int_b^a f(x)\ dx$. It will be clear why this makes sense after we introduce Riemann sums.

Figure 5.4: A graph of $f(x)$ in Example 2.
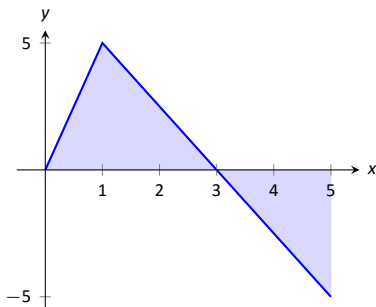
Figure 5.5: A graph of $5f$ in Example 2. (Yes, it looks just like the graph of f in Figure 5.4, just with a different y-scale.)

Notes:

Theorem 34   Properties of the Definite Integral

Let f and g be continuous on a closed interval I that contains the values a, b, and c, and let k be a constant. The following hold:

1. $\displaystyle\int_a^a f(x)\,dx = 0$

2. $\displaystyle\int_a^b f(x)\,dx = -\int_b^a f(x)\,dx$

3. $\displaystyle\int_a^b f(x)\,dx + \int_b^c f(x)\,dx = \int_a^c f(x)\,dx$

4. $\displaystyle\int_a^b (f(x) \pm g(x))\,dx = \int_a^b f(x)\,dx \pm \int_a^b g(x)\,dx$

5. $\displaystyle\int_a^b k \cdot f(x)\,dx = k \cdot \int_a^b f(x)\,dx$

We will justify these properties after introducing Riemann sums. For now, we note that properties 1 and 5 are illustrated in Example 2 and property 2 is our convention from above. To see why property 3 makes sense geometrically, consider the figure below:
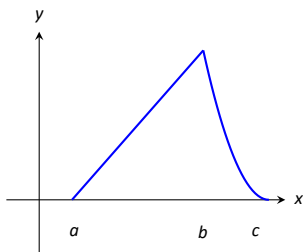
Property 3 says that the total area under this curve should be the sum of the area under the curve from a to b and the area under the curve from b to c.

What if the picture were like the following?

Notes:
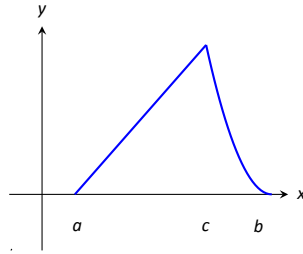
Then we have

$$\int_a^b f(x)\,dx = \int_a^c f(x)\,dx + \int_c^b f(x)\,dx$$

and we can apply property 2.

$$\int_a^c f(x)\,dx = \int_a^b f(x)\,dx - \int_c^b f(x)\,dx, \qquad \text{so property 2 yields}$$

$$\int_a^c f(x)\,dx = \int_a^b f(x)\,dx + \int_b^c f(x)\,dx$$

Example 3   **Evaluating definite integrals using Theorem 34.**
Consider the graph of a function $f(x)$ shown in Figure 5.6. Answer the following:
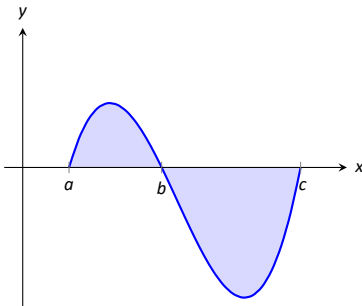
Figure 5.6: A graph of a function in Example 3.

1. Which value is greater: $\displaystyle\int_a^b f(x)\,dx$ or $\displaystyle\int_b^c f(x)\,dx$?

2. Is $\displaystyle\int_a^c f(x)\,dx$ greater or less than 0?

3. Which value is greater: $\displaystyle\int_a^b f(x)\,dx$ or $\displaystyle\int_c^b f(x)\,dx$?

SOLUTION

1. $\int_a^b f(x)\,dx$ has a positive value (since the area is above the x–axis) whereas $\int_b^c f(x)\,dx$ has a negative value. Hence $\int_a^b f(x)\,dx$ is bigger.

2. $\int_a^c f(x)\,dx$ is the total signed area under f between $x = a$ and $x = c$. Since the region below the x–axis looks to be larger than the region above, we conclude that the definite integral has a value less than 0.

3. Note how the second integral has the bounds "reversed." Therefore $\int_c^b f(x)\,dx$ represents a positive number, greater than the area described by the first definite integral. Hence $\int_c^b f(x)\,dx$ is greater.

Notes:

The area definition of the definite integral allows us to use geometry to compute the definite integral of some simple functions.

Example 4   Evaluating definite integrals using geometry
Evaluate the following definite integrals:

1. $\displaystyle\int_{-2}^{5} (2x - 4)\, dx$ 2. $\displaystyle\int_{-3}^{3} \sqrt{9 - x^2}\, dx.$

SOLUTION

1. It is useful to sketch the function in the integrand, as shown in Figure 5.7(a). We see we need to compute the areas of two regions, which we have labeled R_1 and R_2. Both are triangles, so the area computation is straightforward:

$$R_1 : \frac{1}{2}(4)(8) = 16 \qquad R_2 : \frac{1}{2}(3)6 = 9.$$

Region R_1 lies under the x–axis, hence it is counted as negative area (we can think of the triangle's height as being "-8"), so

$$\int_{-2}^{5} (2x - 4)\, dx = -16 + 9 = -7.$$

2. Recognize that the integrand of this definite integral describes a half circle, as sketched in Figure 5.7(b), with radius 3. Thus the area is:

$$\int_{-3}^{3} \sqrt{9 - x^2}\, dx = \frac{1}{2}\pi r^2 = \frac{9}{2}\pi.$$

Example 5   Understanding motion given velocity
Consider the graph of a velocity function of an object moving in a straight line, given in Figure 5.8, where the numbers in the given regions gives the area of that region. Assume that the definite integral of a velocity function gives displacement. Find the maximum speed of the object and its maximum displacement from its starting position.

    SOLUTION   Since the graph gives velocity, finding the maximum speed is simple: it looks to be 15ft/s.

At time $t = 0$, the displacement is 0; the object is at its starting position. At time $t = a$, the object has moved backward 11 feet. Between times $t = a$ and $t = b$, the object moves forward 38 feet, bringing it into a position 27 feet forward of its starting position. From $t = b$ to $t = c$ the object is moving

Notes:

(a)

(b)
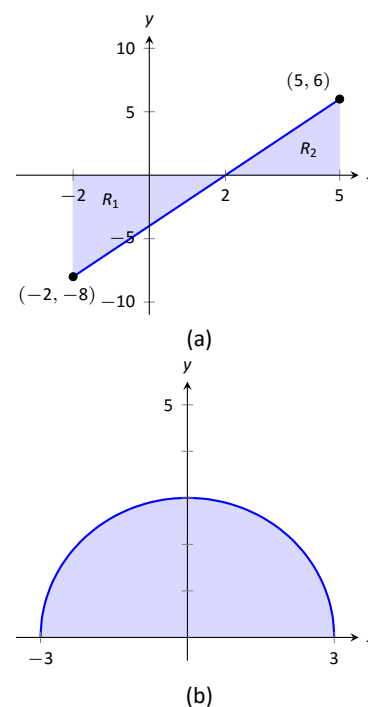
Figure 5.7: A graph of $f(x) = 2x - 4$ in (a) and $f(x) = \sqrt{9 - x^2}$ in (b), from Example 4.
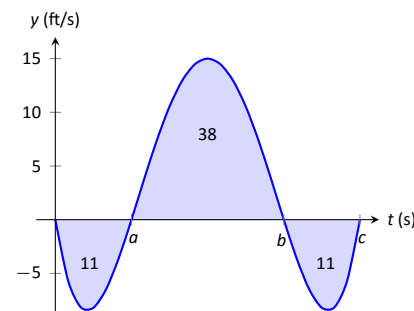
Figure 5.8: A graph of a velocity in Example 5.

Note: The *displacement* of the object is different from the distance traveled since the object moves backwards and forwards at different times in this example. The displacement measures how far the object is from where it started, without regard for how far it actually traveled to get there.

backwards again, hence its maximum displacement is 27 feet from its starting position.

We can also analyze the displacement by drawing the path of the particle's location as time varies, as in Figure 5.9.
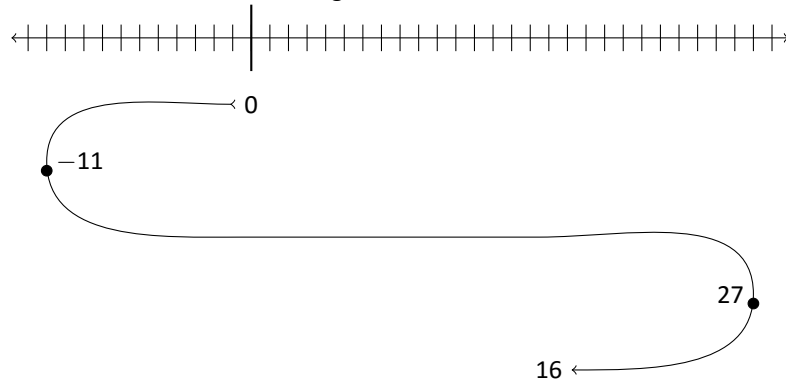
Figure 5.9: Number line for Example 5.

The object starts at the origin, and moves to the left with a negative velocity 11 units. It then reverses direction and moves to the right with a positive velocity 38 units, arriving at 27. Finally, it reverses direction again and moves to the left with a negative velocity 11 units, ending at 26.

Watch the video:
Definite Integral as Area 2 — Breaking Up the Region
at
`https://youtu.be/Z7uyQjcFSy4`

In our examples, we have either found the areas of regions that have nice geometric shapes (such as rectangles, triangles and circles) or the areas were given to us. Consider Figure 5.10, where a region below $y = x^2$ is shaded. What is its area? The function $y = x^2$ is relatively simple, yet the shape it defines has an area that is not simple to find geometrically.

In the next section we will explore how to find the areas of such regions.

Figure 5.10: What is the area below $y = x^2$ on $[0, 3]$? The region is not a usual geometric shape.

Notes:

Exercises 5.2

Terms and Concepts

1. What is "total signed area"?

2. What is "displacement"?

3. What is $\int_3^3 \sin x \, dx$?

4. Give a single definite integral that has the same value as
$$\int_0^1 (2x+3)\, dx + \int_1^2 (2x+3)\, dx.$$

Problems

In Exercises 5–9, a graph of a function $f(x)$ is given. Using the geometry of the graph, evaluate the definite integrals.

5.
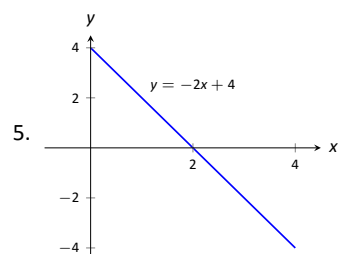
(a) $\int_0^1 (-2x+4)\, dx$ (d) $\int_1^3 (-2x+4)\, dx$

(b) $\int_0^2 (-2x+4)\, dx$ (e) $\int_2^4 (-2x+4)\, dx$

(c) $\int_0^3 (-2x+4)\, dx$ (f) $\int_0^1 (-6x+12)\, dx$

6.
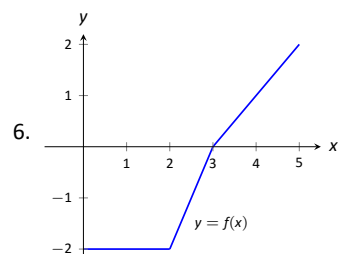
(a) $\int_0^2 f(x)\, dx$ (d) $\int_2^5 f(x)\, dx$

(b) $\int_0^3 f(x)\, dx$ (e) $\int_5^3 f(x)\, dx$

(c) $\int_0^5 f(x)\, dx$ (f) $\int_0^3 -2f(x)\, dx$

7.
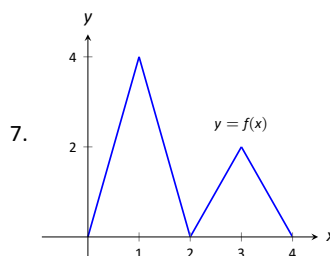
(a) $\int_0^2 f(x)\, dx$ (d) $\int_0^1 4x\, dx$

(b) $\int_2^4 f(x)\, dx$ (e) $\int_2^3 (2x-4)\, dx$

(c) $\int_2^4 2f(x)\, dx$ (f) $\int_2^3 (4x-8)\, dx$

8.
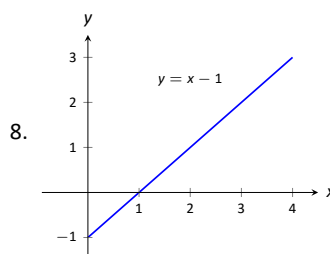
(a) $\int_0^1 (x-1)\, dx$ (d) $\int_2^3 (x-1)\, dx$

(b) $\int_0^2 (x-1)\, dx$ (e) $\int_1^4 (x-1)\, dx$

(c) $\int_0^3 (x-1)\, dx$ (f) $\int_1^4 \big((x-1)+1\big)\, dx$

9.
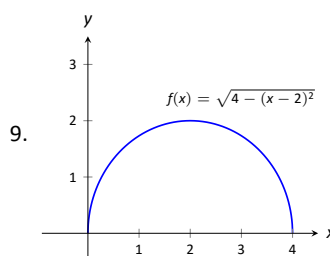
(a) $\int_0^2 f(x)\, dx$ (c) $\int_0^4 f(x)\, dx$

(b) $\int_2^4 f(x)\, dx$ (d) $\int_0^4 5f(x)\, dx$

In Exercises 10–13, a graph of a function $f(x)$ is given; the numbers inside the shaded regions give the area of that region. Evaluate the definite integrals using this area information.
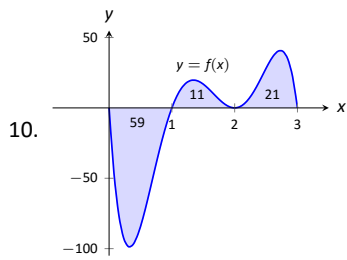
10.

(a) $\int_0^1 f(x)\,dx$

(b) $\int_0^2 f(x)\,dx$

(c) $\int_0^3 f(x)\,dx$

(d) $\int_1^2 -3f(x)\,dx$

(e) $\int_0^2 |f(x)|\,dx$

(f) $\int_0^3 |f(x)|\,dx$

11.

(a) $\int_0^2 f(x)\,dx$

(b) $\int_2^4 f(x)\,dx$

(c) $\int_0^4 f(x)\,dx$

(d) $\int_0^1 f(x)\,dx$

(e) $\int_0^2 |f(x)|\,dx$

(f) $\int_0^4 |f(x)|\,dx$

12.

(a) $\int_{-2}^{-1} f(x)\,dx$

(b) $\int_1^2 f(x)\,dx$

(c) $\int_{-1}^1 f(x)\,dx$

(d) $\int_0^1 f(x)\,dx$

(e) $\int_0^2 |f(x)|\,dx$

(f) $\int_0^1 |f(x)|\,dx$

13.
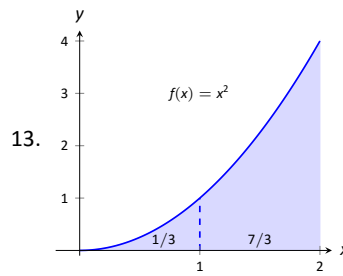
(a) $\int_0^2 5x^2\,dx$

(b) $\int_0^2 (x^2 + 3)\,dx$

(c) $\int_1^3 (x-1)^2\,dx$

(d) $\int_2^4 \left((x-2)^2 + 5\right) dx$

In Exercises 14–15, a graph of the velocity function of an object moving in a straight line is given. Answer the questions based on that graph.
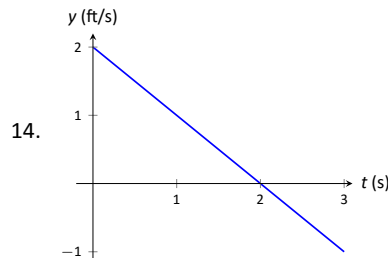
14.

(a) What is the object's maximum velocity?

(b) What is the object's maximum displacement?

(c) What is the object's total displacement on $[0, 3]$?

15.
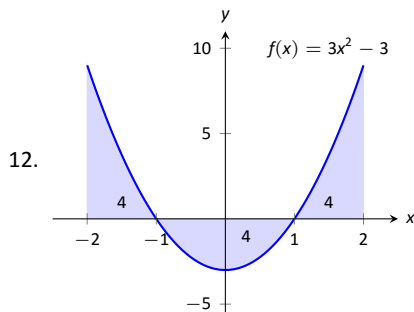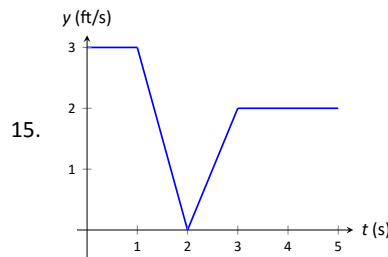
(a) What is the object's maximum velocity?

(b) What is the object's maximum displacement?

(c) What is the object's total displacement on $[0, 5]$?

16. An object is thrown straight up with a velocity, in ft/s, given by $v(t) = -32t + 64$, where t is in seconds, from a height of 48 feet.

(a) What is the object's maximum velocity?

(b) What is the object's maximum displacement?

(c) When does the maximum displacement occur?

(d) When will the object reach a height of 0? (Hint: find when the displacement is -48ft.)

17. An object is thrown straight up with a velocity, in ft/s, given by $v(t) = -32t + 96$, where t is in seconds, from a height of 64 feet.

   (a) What is the object's initial velocity?

   (b) When is the object's displacement 0?

   (c) How long does it take for the object to return to its initial height?

   (d) When will the object reach a height of 210 feet?

In Exercises 18–21, let

- $\int_0^2 f(x)\, dx = 5,$

- $\int_0^3 f(x)\, dx = 7,$

- $\int_0^2 g(x)\, dx = -3,$ and

- $\int_2^3 g(x)\, dx = 5.$

Use these values to evaluate the given definite integrals.

18. $\int_0^2 \big(f(x) + g(x)\big)\, dx$

19. $\int_0^3 \big(f(x) - g(x)\big)\, dx$

20. $\int_2^3 \big(3f(x) + 2g(x)\big)\, dx$

21. Find values for a and b such that
$$\int_0^3 \big(af(x) + bg(x)\big)\, dx = 0$$

In Exercises 22–25, let

- $\int_0^3 s(t)\, dt = 10,$

- $\int_3^5 s(t)\, dt = 8,$

- $\int_3^5 r(t)\, dt = -1,$ and

- $\int_0^5 r(t)\, dt = 11.$

Use these values to evaluate the given definite integrals.

22. $\int_0^3 \big(s(t) + r(t)\big)\, dt$

23. $\int_5^0 \big(s(t) - r(t)\big)\, dt$

24. $\int_3^3 \big(\pi s(t) - 7r(t)\big)\, dt$

25. Find values for a and b such that
$$\int_0^5 \big(ar(t) + bs(t)\big)\, dt = 0$$

Review

In Exercises 26–29, evaluate the given indefinite integral.

26. $\int \big(x^3 - 2x^2 + 7x - 9\big)\, dx$

27. $\int \big(\sin x - \cos x + \sec^2 x\big)\, dx$

28. $\int \big(\sqrt[3]{t} + \frac{1}{t^2} + 2^t\big)\, dt$

29. $\int \big(\frac{1}{x} - \csc x \cot x\big)\, dx$

5.3 Riemann Sums

In the previous section we defined the definite integral of a function on $[a, b]$ to be the signed area between the curve and the x–axis. Some areas were simple to compute; we ended the section with a region whose area was not simple to compute. In this section we develop a technique to find such areas.

A fundamental calculus technique is to first answer a given problem with an approximation, then refine that approximation to make it better, then use limits in the refining process to find the exact answer. That is exactly what we will do here.

Consider the region given in Figure 5.11, which is the area under $y = 4x - x^2$ on $[0, 4]$. What is the signed area of this region – i.e., what is $\int_0^4 (4x - x^2)\, dx$?

We start by approximating. We can surround the region with a rectangle with height and width of 4 and find the area is approximately 16 square units. This is obviously an *over–approximation*; we are including area in the rectangle that is not under the parabola.

We have an approximation of the area, using one rectangle. How can we refine our approximation to make it better? The key to this section is this answer: *use more rectangles.*

Let's use 4 rectangles of equal width of 1. This *partitions* the interval $[0, 4]$ into 4 *subintervals*, $[0, 1]$, $[1, 2]$, $[2, 3]$ and $[3, 4]$. On each subinterval we will draw a rectangle.

There are three common ways to determine the height of these rectangles: the **Left Hand Rule**, the **Right Hand Rule**, and the **Midpoint Rule**. The **Left Hand Rule** says to evaluate the function at the left–hand endpoint of the subinterval and make the rectangle that height. In Figure 5.12, the rectangle drawn on the interval $[2, 3]$ has height determined by the Left Hand Rule; it has a height of $f(2)$. (The rectangle is labeled "LHR.")

The **Right Hand Rule** says the opposite: on each subinterval, evaluate the function at the right endpoint and make the rectangle that height. In the figure, the rectangle drawn on $[0, 1]$ is drawn using $f(1)$ as its height; this rectangle is labeled "RHR.".

The **Midpoint Rule** says that on each subinterval, evaluate the function at the midpoint and make the rectangle that height. The rectangle drawn on $[1, 2]$ was made using the Midpoint Rule, with a height of $f(1.5)$. That rectangle is labeled "MPR."

These are the three most common rules for determining the heights of approximating rectangles, but one is not forced to use one of these three methods. The rectangle on $[3, 4]$ has a height of approximately $f(3.53)$, very close to the Midpoint Rule. It was chosen so that the area of the rectangle is *exactly* the area of the region under f on $[3, 4]$. (Later you'll be able to figure how to do this, too.)
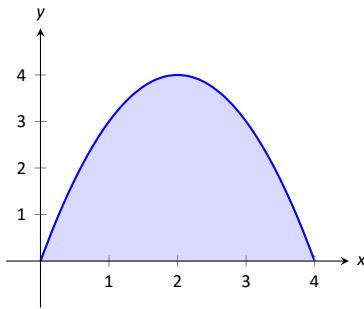
Figure 5.11: A graph of $f(x) = 4x - x^2$. What is the area of the shaded region?
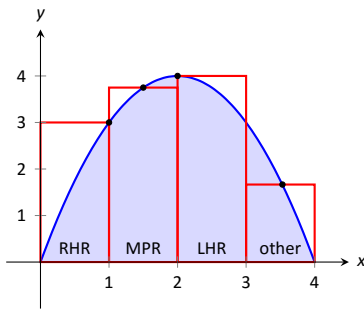
Figure 5.12: Approximating $\int_0^4 (4x - x^2)\, dx$ using rectangles. The heights of the rectangles are determined using different rules.

Notes:

The following example will approximate the value of $\int_0^4 (4x - x^2)\, dx$ using these rules.

Example 1   Using the Left Hand, Right Hand and Midpoint Rules

Approximate the value of $\int_0^4 (4x - x^2)\, dx$ using the Left Hand Rule, the Right Hand Rule, and the Midpoint Rule, using 4 equally spaced subintervals.

SOLUTION   We break the interval $[0, 4]$ into four subintervals as before. In Figure 5.13 we first see 4 rectangles drawn on $f(x) = 4x - x^2$ using the Left Hand Rule. (The areas of the rectangles are given in each figure.)

Note how in the first subinterval, $[0, 1]$, the rectangle has height $f(0) = 0$. We add up the areas of each rectangle (height× width) for our Left Hand Rule approximation:

$$f(0) \cdot 1 + f(1) \cdot 1 + f(2) \cdot 1 + f(3) \cdot 1 =$$
$$0 + 3 + 4 + 3 = 10.$$

Figure 5.13 next shows 4 rectangles drawn under f using the Right Hand Rule; note how the $[3, 4]$ subinterval has a rectangle of height 0.

These rectangle seem to be the mirror image of those found with the Left Hand Rule. (This is because of the symmetry of our shaded region.) Our approximation gives the same answer as before, though calculated a different way:

$$f(1) \cdot 1 + f(2) \cdot 1 + f(3) \cdot 1 + f(4) \cdot 1 =$$
$$3 + 4 + 3 + 0 = 10.$$

Figure 5.13 last shows 4 rectangles drawn under f using the Midpoint Rule. This gives an approximation of $\int_0^4 (4x - x^2)\, dx$ as:

$$f(0.5) \cdot 1 + f(1.5) \cdot 1 + f(2.5) \cdot 1 + f(3.5) \cdot 1 =$$
$$1.75 + 3.75 + 3.75 + 1.75 = 11.$$

Our three methods provide two approximations of $\int_0^4 (4x - x^2)\, dx$: 10 and 11.

Summation Notation

It is hard to tell at this moment which is a better approximation: 10 or 11? We can continue to refine our approximation by using more rectangles. The notation

Left Hand Rule

Right Hand Rule

Midpoint Rule
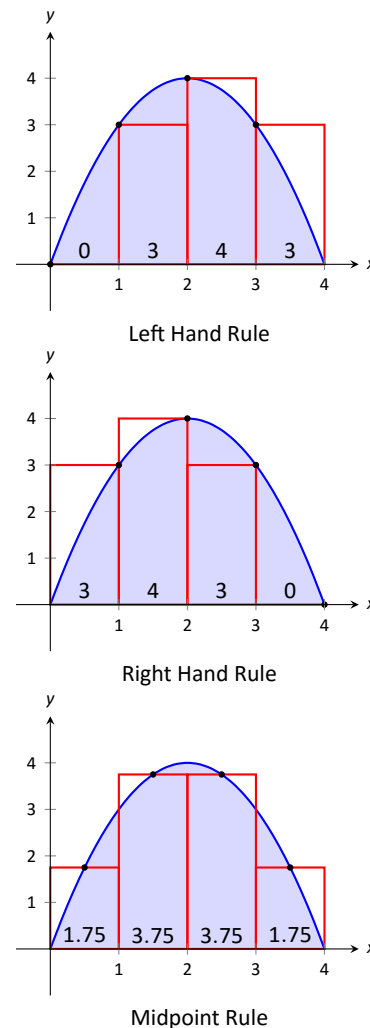
Figure 5.13: Approximating $\int_0^4 (4x - x^2)\, dx$ in Example 1.

Notes:

can become unwieldy, though, as we add up longer and longer lists of numbers. We introduce **summation notation** to ameliorate this problem.

Suppose we wish to add up a list of numbers a_1, a_2, a_3, …, a_9. Instead of writing

$$a_1 + a_2 + a_3 + a_4 + a_5 + a_6 + a_7 + a_8 + a_9,$$

we use summation notation and write

$$\sum_{i=1}^{9} a_i$$

Lets analyze this notation.
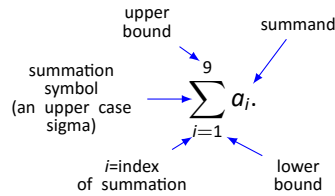
Figure 5.14: Understanding summation notation.

The upper case sigma, $\sum$, represents the term "sum." The index of summation in this example is i; any symbol can be used. By convention, the index takes on only the integer values between (and including) the lower and upper bounds.

Let's practice using this notation.

Example 2   **Using summation notation**
Let the numbers $\{a_i\}$ be defined as $a_i = 2i - 1$ for integers i, where $i \geq 1$. So $a_1 = 1$, $a_2 = 3$, $a_3 = 5$, etc. (The output is the positive odd integers). Evaluate the following summations:

1. $\displaystyle\sum_{i=1}^{6} a_i$
2. $\displaystyle\sum_{i=3}^{7}(3a_i - 4)$
3. $\displaystyle\sum_{i=1}^{4}(a_i)^2$

SOLUTION

1.
$$\sum_{i=1}^{6} a_i = a_1 + a_2 + a_3 + a_4 + a_5 + a_6$$
$$= 1 + 3 + 5 + 7 + 9 + 11$$
$$= 36.$$

Notes:

2. Note the starting value is different than 1:

$$\sum_{i=3}^{7}(3a_i - 4) = (3a_3 - 4) + (3a_4 - 4) + (3a_5 - 4) + (3a_6 - 4) + (3a_7 - 4)$$

$$= 11 + 17 + 23 + 29 + 35$$

$$= 115.$$

3.

$$\sum_{i=1}^{4}(a_i)^2 = (a_1)^2 + (a_2)^2 + (a_3)^2 + (a_4)^2$$

$$= 1^2 + 3^2 + 5^2 + 7^2$$

$$= 84$$

It might seem odd to stress a new, concise way of writing summations only to write each term out as we add them up. It is. The following theorem gives some of the properties of summations that allow us to work with them without writing individual terms. Examples will follow.

Theorem 35   Properties of Summations

1. $\displaystyle\sum_{i=1}^{n} c = c \cdot n$, where c is a constant.

5. $\displaystyle\sum_{i=1}^{n} i = \frac{n(n+1)}{2}$

2. $\displaystyle\sum_{i=m}^{n}(a_i \pm b_i) = \sum_{i=m}^{n} a_i \pm \sum_{i=m}^{n} b_i$

6. $\displaystyle\sum_{i=1}^{n} i^2 = \frac{n(n+1)(2n+1)}{6}$

3. $\displaystyle\sum_{i=m}^{n} c \cdot a_i = c \cdot \sum_{i=m}^{n} a_i$

7. $\displaystyle\sum_{i=1}^{n} i^3 = \left(\frac{n(n+1)}{2}\right)^2$

4. $\displaystyle\sum_{i=m}^{j} a_i + \sum_{i=j+1}^{n} a_i = \sum_{i=m}^{n} a_i$

Note: In practice we will sometimes need variations on formulas 5, 6, and 7 above. For example, we note that

$$\sum_{i=0}^{n} i = 0 + 1 + 2 + \cdots + n = 0 + \sum_{i=1}^{n} i = 0 + \frac{n(n+1)}{2} = \frac{n(n+1)}{2},$$

Notes:

so we see that

$$\sum_{i=0}^{n} i = \frac{n(n+1)}{2}.$$

Similarly, we find that

$$\sum_{i=0}^{n} i^2 = \frac{n(n+1)(2n+1)}{6}, \quad \text{and}$$

$$\sum_{i=0}^{n} i^3 = \left(\frac{n(n+1)}{2} \right)^2$$

Example 3   **Evaluating summations using Theorem 35**
Revisit Example 2 and, using Theorem 35, evaluate

$$\sum_{i=1}^{6} a_i = \sum_{i=1}^{6} (2i - 1).$$

SOLUTION

$$\sum_{i=1}^{6} (2i - 1) = \sum_{i=1}^{6} 2i - \sum_{i=1}^{6} (1) \qquad \text{(Theorem 35(2))}$$

$$= \left(2 \sum_{i=1}^{6} i \right) - \sum_{i=1}^{6} (1) \qquad \text{(Theorem 35(3))}$$

$$= 2 \left(\frac{6(6+1)}{2} \right) - 6 \qquad \text{(Theorem 35(1,5))}$$

$$= 2(21) - 6 = 36$$

We obtained the same answer without writing out all six terms. When dealing with small sizes of n, it may be faster to write the terms out by hand. However, Theorem 35 is incredibly important when dealing with large sums as we'll soon see.

Riemann Sums

Consider again $\int_{0}^{4} (4x - x^2)\, dx$. We will approximate this definite integral using 16 equally spaced subintervals and the Right Hand Rule in Example 4. Before doing so, it will pay to do some careful preparation.
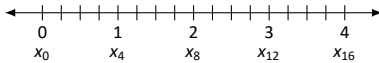
Figure 5.15: Dividing $[0, 4]$ into 16 equally spaced subintervals.

Notes:

Figure 5.15 shows a number line of $[0, 4]$ subdivided into 16 equally spaced subintervals. We denote 0 as x_0; we have marked the values of x_4, x_8, x_{12}, and x_{16}. We could mark them all, but the figure would get crowded. While it is easy to figure that $x_9 = 2.25$, in general, we want a method of determining the value of x_i without consulting the figure. Consider:

$$x_i = x_0 + i\Delta x$$

number of subintervals between x_0 and x_i

starting value

subinterval size

So $x_9 = x_0 + 9(4/16) = 9/4 = 2.25$.

If we had partitioned $[0, 4]$ into 100 equally spaced subintervals, each subinterval would have length $\Delta x = 4/100 = 0.04$. We could compute x_{31} as

$$x_{31} = x_0 + 31(4/100) = 124/100 = 1.24.$$

(That was far faster than creating a sketch first.)

Given any subdivision of $[0, 4]$, the first subinterval is $[x_0, x_1]$; the second is $[x_1, x_2]$; the ith subinterval is $[x_{i-1}, x_i]$.

When using the Left Hand Rule, the height of the ith rectangle will be $f(x_{i-1})$.

When using the Right Hand Rule, the height of the ith rectangle will be $f(x_i)$.

When using the Midpoint Rule, the height of the ith rectangle will be $f\left(\frac{x_{i-1}+x_i}{2}\right)$.

Thus approximating $\int_0^4 (4x - x^2)\, dx$ with 16 equally spaced subintervals can be expressed as follows, where $\Delta x = 4/16 = 1/4$:

Left Hand Rule: $\displaystyle\sum_{i=1}^{16} f(x_{i-1})\Delta x$

Right Hand Rule: $\displaystyle\sum_{i=1}^{16} f(x_i)\Delta x$

Midpoint Rule: $\displaystyle\sum_{i=1}^{16} f\left(\frac{x_{i-1}+x_i}{2}\right)\Delta x$

Notes:

Watch the video:
Calculating a Definite Integral Using Riemann Sums
— Part 1 at
`https://youtu.be/gFpHHTxsDkI`

We use these formulas in the next two examples. The following example lets us practice using the Left Hand Rule and the summation formulas introduced in Theorem 35.

Example 4   Approximating definite integrals using sums
Approximate $\displaystyle\int_0^4 (4x - x^2)\, dx$ using the Right Hand Rule and summation formulas with 16 and 1000 equally spaced intervals.

SOLUTION   Using the formula derived before, using 16 equally spaced intervals and the Right Hand Rule, we can approximate the definite integral as

$$\sum_{i=1}^{16} f(x_i)\Delta x.$$

We have $\Delta x = 4/16 = 0.25$, $x_i = 0 + i\Delta x = i\Delta x$, and $f(x_i) = f(i\Delta x) = 4i\Delta x - i^2\Delta x^2$. Using the summation formulas, we see:

$$
\begin{aligned}
\int_0^4 (4x - x^2)\, dx &\approx \sum_{i=1}^{16} f(x_i)\Delta x \\
&= \sum_{i=1}^{16} f(i\Delta x)\Delta x \\
&= \sum_{i=1}^{16} (4i\Delta x - i^2(\Delta x)^2)\Delta x && \text{(from above)} \\
&= \sum_{i=1}^{16} (4i(\Delta x)^2 - i^2(\Delta x)^3) \\
&= \sum_{i=1}^{16} 4i(\Delta x)^2 - \sum_{i=1}^{16} i^2(\Delta x)^3 && \text{(Theorem 35(2))}
\end{aligned}
$$

Notes:

$$= 4(\Delta x)^2 \sum_{i=1}^{16} i - (\Delta x)^3 \sum_{i=1}^{16} i^2 \qquad \text{(*)} \quad \text{(Theorem 35(3))}$$

$$= 4\left(\frac{1}{4}\right)^2 \left(\frac{(16)(17)}{2}\right) - \left(\frac{1}{4}\right)^3 \left(\frac{(16)(17)(33)}{6}\right) \quad \text{(Theorem 35(5,6))}$$

$$= 34 - \frac{187}{8} = \frac{85}{8} = 10.625$$

We were able to sum up the areas of 16 rectangles with very little computation. In Figure 5.16 the function and the 16 rectangles are graphed. While some rectangles over-approximate the area, others under-approximate the area by about the same amount. Thus our approximate area of 10.625 is likely a fairly good approximation.

Notice Equation (*); by changing the 16's to 1000's and changing the value of Δx to $4/1000 = 0.004$, we can use the equation to sum up the areas of 1000 rectangles. We do so here, skipping from the original summand to the equivalent of Equation (*) to save space.

$$\int_0^4 (4x - x^2)\, dx \approx \sum_{i=1}^{1000} f(x_i)\Delta x$$

$$= 4(\Delta x)^2 \sum_{i=1}^{1000} i - (\Delta x)^3 \sum_{i=1}^{1000} i^2$$

$$= 4(.004)^2 \left(\frac{(1000)(1001)}{2}\right) - (0.004)^3 \left(\frac{(1000)(1001)(2001)}{6}\right)$$

$$= 10.666656$$

Using many, many rectangles, we likely have a good approximation of $\int_0^4 (4x - x^2)\, dx$. That is,

$$\int_0^4 (4x - x^2)\, dx \approx 10.666656.$$

Figure 5.16: Approximating $\int_0^4 (4x - x^2)\, dx$ with the Right Hand Rule and 16 evenly spaced subintervals.

Before the above example, we stated what the summations for the Left Hand, Right Hand and Midpoint Rules looked like. Each had the same basic structure, which was:

1. each rectangle has the same width, which we referred to as Δx, and

2. each rectangle's height is determined by evaluating f at a particular point in each subinterval. For instance, the Left Hand Rule states that each rectangle's height is determined by evaluating f at the left hand endpoint of the subinterval the rectangle lives on.

Notes:

One could partition an interval $[a, b]$ with subintervals that did not have the same size. We refer to the length of the first subinterval as Δx_1, the length of the second subinterval as Δx_2, and so on, giving the length of the ith subinterval as Δx_i. Also, one could determine each rectangle's height by evaluating f at *any* point in the ith subinterval. We refer to the point picked in the first subinterval as c_1, the point picked in the second subinterval as c_2, and so on, with c_i representing the point picked in the ith subinterval. Thus the height of the ith subinterval would be $f(c_i)$, and the area of the ith rectangle would be $f(c_i)\Delta x_i$.

Summations of rectangles with area $f(c_i)\Delta x_i$ are named after mathematician Georg Friedrich Bernhard Riemann, as given in the following definition.

Definition 24   Riemann Sum

Let f be defined on the closed interval $[a, b]$ and let Δx be a partition of $[a, b]$, with

$$a = x_0 < x_1 < \ldots < x_{n-1} < x_n = b.$$

Let Δx_i denote the length of the ith subinterval $[x_{i-1}, x_i]$ and let c_i denote any value in the ith subinterval. The sum

$$\sum_{i=1}^{n} f(c_i)\Delta x_i$$

is a **Riemann sum** of f on $[a, b]$.
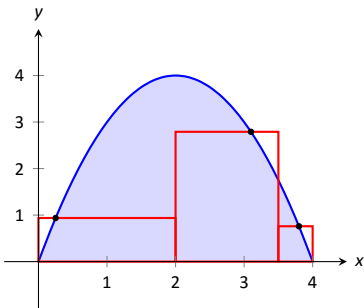
Figure 5.17: An example of a general Riemann sum to approximate $\int_0^4 (4x - x^2)\, dx$.

Figure 5.17 shows the approximating rectangles of a Riemann sum of $\int_0^4 (4x - x^2)\, dx$. While the rectangles in this example do not approximate well the shaded area, they demonstrate that the subinterval widths may vary and the heights of the rectangles can be determined without following a particular rule.

Usually, Riemann sums are calculated using one of the three methods we have introduced. The uniformity of construction makes computations easier. We have $\Delta x_i = \Delta x = \dfrac{b - a}{n}$ and the ith term of the partition is $x_i = a + i\Delta x$. Then the Left Hand Rule uses $c_i = x_{i-1}$, the Right Hand Rule uses $c_i = x_i$, and the Midpoint Rule uses $c_i = \dfrac{x_{i-1} + x_i}{2}$.

Let's do another example.

Example 5   Approximating definite integrals with sums

Approximate $\displaystyle\int_{-2}^{3} (5x + 2)\, dx$ using the Midpoint Rule and 10 equally spaced intervals.

Notes:

SOLUTION   We see that

$$\Delta x = \frac{3 - (-2)}{10} = \frac{1}{2} \quad \text{and} \quad x_i = (-2) + \frac{1}{2}i = \frac{i}{2} - 2.$$

As we are using the Midpoint Rule, we will also need x_{i-1} and $\dfrac{x_{i-1} + x_i}{2}$. Since $x_i = \frac{i}{2} - \frac{5}{2}$, $x_{i-1} = \frac{i-1}{2} - 2 = \frac{i}{2} - \frac{5}{2}$. This gives

$$\frac{x_{i-1} + x_i}{2} = \frac{(\frac{i}{2} - \frac{5}{2}) + (\frac{i}{2} - 2)}{2} = \frac{i - \frac{9}{2}}{2} = \frac{i}{2} - \frac{9}{4}.$$

We now construct the Riemann sum and compute its value using summation formulas.

$$\int_{-2}^{3} (5x + 2)\, dx \approx \sum_{i=1}^{10} f\left(\frac{x_{i-1} + x_i}{2}\right) \Delta x$$

$$= \sum_{i=1}^{10} f\left(\frac{i}{2} - \frac{9}{4}\right) \Delta x$$

$$= \sum_{i=1}^{10} \left(5\left(\frac{i}{2} - \frac{9}{4}\right) + 2\right)\left(\frac{1}{2}\right)$$

$$= \sum_{i=1}^{10} \left(\frac{5i}{4} - \frac{37}{8}\right)$$

$$= \left(\frac{5}{4}\sum_{i=1}^{10}(i) - \sum_{i=1}^{10}\left(\frac{37}{8}\right)\right)$$

$$= \left(\frac{5}{4} \cdot \frac{(10)(11)}{2} - 10 \cdot \frac{37}{8}\right)$$

$$= \frac{45}{2} = 22.5$$

Note the graph of $f(x) = 5x + 2$ in Figure 5.18. The regions whose area is computed by the definite integral are triangles, meaning we can find the exact answer without summation techniques. We find that the exact answer is indeed 22.5. One of the strengths of the Midpoint Rule is that often each rectangle includes area that should not be counted, but misses other area that should. When Δx is small, these two amounts are about equal and these errors almost "subtract each other out." In this example, since our function is a line, these errors are exactly equal and they do subtract each other out, giving us the exact answer.

Note too that when the function is negative, the rectangles have a "negative" height. When we compute the area of the rectangle, we use $f(c_i)\Delta x$; when f is negative, the area is counted as negative.
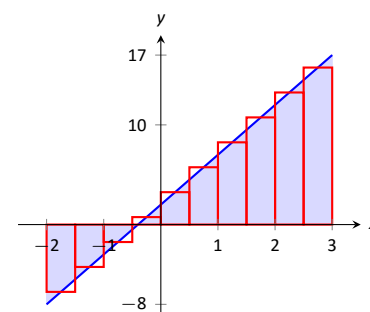
Figure 5.18: Approximating $\int_{-2}^{3}(5x + 2)\, dx$ using the Midpoint Rule and 10 evenly spaced subintervals in Example 5.

Notes:

Notice in the previous example that while we used 10 equally spaced intervals, the number "10" didn't play a big role in the calculations until the very end. Mathematicians love to abstract ideas; let's approximate the area of another region using n subintervals, where we do not specify a value of n until the very end.

Example 6   Approximating definite integrals with a sum formula

Revisit $\int_0^4 (4x - x^2)\, dx$ yet again. Approximate this definite integral using the Right Hand Rule with n equally spaced subintervals.

SOLUTION   We see that $\Delta x = \dfrac{4 - 0}{n} = \dfrac{4}{n}$. We also find $x_i = 0 + \Delta x i = \dfrac{4i}{n}$.

We construct the Right Hand Rule Riemann sum as follows. Be sure to follow each step carefully. If you get stuck, and do not understand how one line proceeds to the next, you may skip to the result and consider how this result is used. You should come back, though, and work through each step for full understanding.

$$\int_0^4 (4x - x^2)\, dx \approx \sum_{i=1}^{n} f(x_i)\Delta x$$

$$= \sum_{i=1}^{n} f\left(\frac{4i}{n}\right)\Delta x$$

$$= \sum_{i=1}^{n} \left[4\frac{4i}{n} - \left(\frac{4i}{n}\right)^2\right]\frac{4}{n}$$

$$= \sum_{i=1}^{n} \left(\frac{64}{n^2}\right)i - \sum_{i=1}^{n}\left(\frac{64}{n^3}\right)i^2$$

$$= \left(\frac{64}{n^2}\right)\sum_{i=1}^{n}i - \left(\frac{64}{n^3}\right)\sum_{i=1}^{n}i^2$$

$$= \left(\frac{64}{n^2}\right)\cdot\frac{n(n+1)}{2} - \left(\frac{64}{n^3}\right)\frac{n(n+1)(2n+1)}{6}$$

$$= \frac{32(n+1)}{n} - \frac{32(n+1)(2n+1)}{3n^2} \qquad \text{(now simplify)}$$

$$= \frac{32}{3}\left(1 - \frac{1}{n^2}\right)$$

The result is an amazing, easy to use formula. To approximate the definite integral with 10 equally spaced subintervals and the Right Hand Rule, set $n = 10$

Notes:

and compute

$$\int_0^4 (4x - x^2)\, dx \approx \frac{32}{3}\left(1 - \frac{1}{10^2}\right) = 10.56.$$

Recall how earlier we approximated the definite integral with 4 subintervals; with $n = 4$, the formula gives 10, our answer as before.

It is now easy to approximate the integral with 1,000,000 subintervals! Hand-held calculators will round off the answer a bit prematurely giving an answer of 10.66666667. (The actual answer is 10.666666666656.)

We now take an important leap. Up to this point, our mathematics has been limited to geometry and algebra (finding areas and manipulating expressions). Now we apply *calculus*. For any *finite n*, we know that

$$\int_0^4 (4x - x^2)\, dx \approx \frac{32}{3}\left(1 - \frac{1}{n^2}\right).$$

Both common sense and high–level mathematics tell us that as n gets large, the approximation gets better. In fact, if we take the *limit* as $n \to \infty$, we get the *exact area* described by $\int_0^4 (4x - x^2)\, dx$. That is,

$$\begin{aligned}
\int_0^4 (4x - x^2)\, dx &= \lim_{n \to \infty} \frac{32}{3}\left(1 - \frac{1}{n^2}\right) \\
&= \frac{32}{3}(1 - 0) \\
&= \frac{32}{3} = 10.\overline{6}
\end{aligned}$$

This is a fantastic result. By considering n equally–spaced subintervals, we obtained a formula for an approximation of the definite integral that involved our variable n. As n grows large – without bound – the error shrinks to zero and we obtain the exact area.

This section started with a fundamental calculus technique: make an approximation, refine the approximation to make it better, then use limits in the refining process to get an exact answer. That is precisely what we just did.

Let's practice this again.

Example 7   **Approximating definite integrals with a sum formula**

Find a formula that approximates $\int_{-1}^5 x^3\, dx$ using the Right Hand Rule and n equally spaced subintervals, then take the limit as $n \to \infty$ to find the exact area.

Notes:
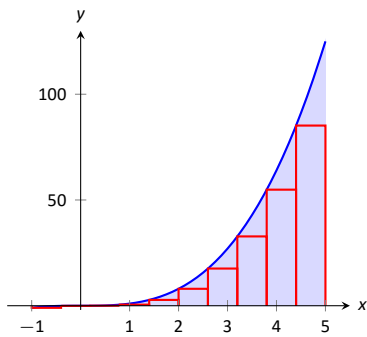
Figure 5.19: Approximating $\int_{-1}^{5} x^3\, dx$ using the Left Hand Rule and 10 evenly spaced subintervals.

SOLUTION   We see that $\Delta x = \frac{5-(-1)}{n} = \frac{6}{n}$ and $x_i = (-1) + i\Delta x = -1 + \frac{6i}{n}$.

The Riemann sum corresponding to the Right Hand Rule is (followed by simplifications):

$$\int_{-1}^{5} x^3\, dx \approx \sum_{i=1}^{n} f(x_i)\Delta x$$

$$= \sum_{i=1}^{n} f(-1 + i\Delta x)\Delta x$$

$$= \sum_{i=1}^{n} \left(-1 + i\frac{6}{n}\right)^3 \frac{6}{n}$$

$$= \sum_{i=1}^{n} \frac{1296 i^3}{n^4} - \frac{648 i^2}{n^3} + \frac{108 i}{n^2} - \frac{6}{n}$$

$$= \frac{1296}{n^4} \sum_{i=1}^{n} i^3 - \frac{648}{n^3} \sum_{i=1}^{n} i^2 + \frac{108}{n^2} \sum_{i=1}^{n} i - \sum_{i=1}^{n} \frac{6}{n}$$

$$= \frac{1296}{n^4} \left(\frac{n(n+1)}{2}\right)^2 - \frac{648}{n^3} \frac{n(n+1)(2n+1)}{6} + \frac{108}{n^2} \frac{n(n+1)}{2} - 6$$

$$= 156 + \frac{378}{n} + \frac{216}{n^2} \qquad \text{(after a sizable amount of algebra)}$$

Once again, we have found a compact formula for approximating the definite integral with n equally spaced subintervals and the Right Hand Rule. Using 10 subintervals, we have an approximation of 195.96 (these rectangles are shown in Figure 5.19). Using $n = 100$ gives an approximation of 159.802.

Now find the exact answer using a limit:

$$\int_{-1}^{5} x^3\, dx = \lim_{n\to\infty} \left(156 - \frac{378}{n} + \frac{216}{n^2}\right) = 156.$$

Limits of Riemann Sums

We have used limits to find the exact value of certain definite integrals. Will this always work? We will show, given not–very–restrictive conditions, that yes, it will always work.

The previous two examples demonstrated how an expression such as

$$\sum_{i=1}^{n} f(x_i)\Delta x$$

Notes:

can be rewritten as an expression explicitly involving n, such as $\dfrac{32}{3}\left(1 - \dfrac{1}{n^2}\right)$.

Viewed in this manner, we can think of the summation as a function of n. An n value is given (where n is a positive integer), and the sum of areas of n equally spaced rectangles is returned, using the Left Hand, Right Hand, or Midpoint Rules.

Given a definite integral $\displaystyle\int_a^b f(x)\,dx$, let:

- $S_L(n) = \displaystyle\sum_{i=1}^{n} f(x_{i-1})\Delta x$, the sum of equally spaced rectangles formed using the Left Hand Rule,

- $S_R(n) = \displaystyle\sum_{i=1}^{n} f(x_i)\Delta x$, the sum of equally spaced rectangles formed using the Right Hand Rule, and

- $S_M(n) = \displaystyle\sum_{i=1}^{n} f\left(\dfrac{x_{i-1} + x_i}{2}\right)\Delta x$, the sum of equally spaced rectangles formed using the Midpoint Rule.

Recall the definition of a limit as $n \to \infty$: $\displaystyle\lim_{n\to\infty} S_L(n) = K$ if, given any $\varepsilon > 0$, there exists $N > 0$ such that

$$|S_L(n) - K| < \varepsilon \quad \text{when} \quad n \geq N.$$

The following theorem states that we can use any of our three rules to find the exact value of a definite integral $\displaystyle\int_a^b f(x)\,dx$. It also goes two steps further. The theorem states that the height of each rectangle doesn't have to be determined following a specific rule, but could be $f(c_i)$, where c_i is any point in the ith subinterval, as discussed before Riemann Sums where defined in Definition 24.

The theorem goes on to state that the rectangles do not need to be of the same width. Using the notation of Definition 24, let Δx_i denote the length of the ith subinterval in a partition of $[a, b]$. Now let $\|\Delta x\|$ represent the length of the largest subinterval in the partition: that is, $\|\Delta x\|$ is the largest of all the Δx_i's (this is sometimes called the *size* of the partition). If $\|\Delta x\|$ is small, then $[a, b]$ must be partitioned into many subintervals, since all subintervals must have small lengths. "Taking the limit as $\|\Delta x\|$ goes to zero" implies that the number n of subintervals in the partition is growing to infinity, as the largest subinterval length is becoming arbitrarily small. We then interpret the expression

$$\lim_{\|\Delta x\|\to 0} \sum_{i=1}^{n} f(c_i)\Delta x_i$$

Notes:

as "the limit of the sum of rectangles, where the width of each rectangle can be different but getting small, and the height of each rectangle is not necessarily determined by a particular rule." The theorem states that this Riemann Sum also gives the value of the definite integral of f over $[a, b]$.

Theorem 36   Definite Integrals and the Limit of Riemann Sums

Let f be continuous on the closed interval $[a, b]$ and let $S_L(n)$, $S_R(n)$ and $S_M(n)$ be defined as before. Then:

1. $\displaystyle\lim_{n\to\infty} S_L(n) = \lim_{n\to\infty} S_R(n) = \lim_{n\to\infty} S_M(n) = \lim_{n\to\infty} \sum_{i=1}^{n} f(c_i)\Delta x,$

2. $\displaystyle\lim_{n\to\infty} \sum_{i=1}^{n} f(c_i)\Delta x = \int_{a}^{b} f(x)\, dx,$ and

3. $\displaystyle\lim_{\|\Delta x\|\to 0} \sum_{i=1}^{n} f(c_i)\Delta x_i = \int_{a}^{b} f(x)\, dx.$

Now that we have more tools to work with, we can justify the remaining properties in Theorem 34.

Proof

1. To see why this property holds note that for any Riemann sum we have $\Delta x = 0$, from which we see that:

$$\int_{a}^{b} f(x)\, dx = \lim_{n\to\infty} \sum_{i=1}^{n} f(c_i)\Delta x \quad \text{(by Theorem 36(2))}$$
$$= \lim_{n\to\infty} 0$$
$$= 0$$

2. Applying Theorem 36(2), we have:

$$\int_{a}^{b} f(x)\, dx = \lim_{n\to\infty} \sum_{i=1}^{n} f(c_i)\Delta x.$$

When we compute $\displaystyle\int_{b}^{a} f(x)\, dx$, we can use the same partitions and the same points c_i, so the heights $f(c_i)$ will remain the same. Since we want

Notes:

to start at $x = b$ and finish at $x = a$, we use $\widetilde{\Delta x} = \frac{a-b}{n} = -\Delta x$. We now have:

$$\int_b^a f(x)\,dx = \lim_{n\to\infty} \sum_{i=1}^n f(c_i)\widetilde{\Delta x} \qquad \text{(Theorem 36(2))}$$

$$= \lim_{n\to\infty} \sum_{i=1}^n f(c_i)(-\Delta x)$$

$$= \lim_{n\to\infty} -\left(\sum_{i=1}^n f(c_i)\Delta x \right) \qquad \text{(using Theorem 35(3))}$$

$$= -\lim_{n\to\infty} \sum_{i=1}^n f(c_i)\Delta x$$

$$= -\int_a^b f(x)\,dx \qquad \text{(Theorem 36(2))}$$

3. This property was justified previously.

4. To see why this property holds, we again use Theorems 35 and 36. In this case we have:

$$\int_a^b (f(x) + g(x))\,dx = \lim_{n\to\infty} (f(c_i) + g(c_i))\Delta x$$

$$= \lim_{n\to\infty} \sum_{i=1}^n (f(c_i)\Delta x + g(c_i)\Delta x)$$

$$= \lim_{n\to\infty} \left(\sum_{i=1}^n f(c_i)\Delta x + \sum_{i=0}^{n-1} g(c_i)\Delta x \right)$$

$$= \lim_{n\to\infty} \sum_{i=1}^n f(c_i)\Delta x + \lim_{n\to\infty} \sum_{i=1}^n g(c_i)\Delta x$$

$$= \int_a^b f(x)\,dx + \int_a^b g(x)\,dx$$

5. The justification of this property is left as an exercise. □

Notes:

> **Theorem 37   Further Properties of the Definite Integral**
> Let f be continuous on the interval $[a, b]$ and let k, m, and M be constants. The following hold:
>
> 1. $\displaystyle\int_a^b k\,dx = k(b-a).$
>
> 2. If $m \le f(x)$ for all x in $[a, b]$, then $m(b-a) \le \displaystyle\int_a^b f(x)\,dx.$
>
> 3. If $f(x) \le M$ for all x in $[a, b]$, then $\displaystyle\int_a^b f(x)\,dx \le M(b-a).$

Proof
Before justifying these properties, note that for any subdivision of $[a, b]$ we have:

$$\sum_{i=1}^n \Delta x = n\frac{b-a}{n} = b - a.$$

To see why (a) holds, let k be a constant. We apply Theorem 36 to see that:

$$\int_a^b k\,dx = \lim_{n\to\infty} \sum_{i=1}^n k\Delta x$$

$$= \lim_{n\to\infty} k\left(\sum_{i=1}^n \Delta x\right) \qquad \text{(using Theorem 35)}$$

$$= k\left(\lim_{n\to\infty} \sum_{i=1}^n \Delta x\right)$$

$$= k\left(\lim_{n\to\infty} (b-a)\right)$$

$$= k(b-a)$$

We can now use this property to see why (b) holds. Let f and m be as given.

Notes:

Then we have:

$$m(b - a) = \int_a^b m\,dx$$

$$= \lim_{n \to \infty} \sum_{i=1}^{n} m\,\Delta x \qquad \text{(Theorem 36)}$$

$$\leq \lim_{n \to \infty} \sum_{i=1}^{n} f(c_i)\,\Delta x$$

$$= \int_a^b f(x)\,dx \qquad \text{(Theorem 36)}$$

Justifying property (c) is similar and is left as an exercise.   □

We summarize what we have learned over the past few sections here.

- Knowing the "area under the curve" can be useful. One common example is: the area under a velocity curve is displacement.

- We have defined the definite integral, $\int_a^b f(x)\,dx$, to be the signed area under f on the interval $[a, b]$.

- While we can approximate a definite integral many ways, we have focused on using rectangles whose heights can be determined using: the Left Hand Rule, the Right Hand Rule and the Midpoint Rule.

- Sums of rectangles of this type are called Riemann sums.

- The exact value of the definite integral can be computed using the limit of a Riemann sum. We generally use one of the above methods as it makes the algebra simpler.

We first learned of derivatives through limits then learned rules that made the process simpler. We know of a way to evaluate a definite integral using limits; in the next section we will see how the Fundamental Theorem of Calculus makes the process simpler. The key feature of this theorem is its connection between the indefinite integral and the definite integral.

Notes:

Exercises 5.3

Terms and Concepts

1. A fundamental calculus technique is to use _____ to refine approximations to get an exact answer.

2. What is the upper bound in the summation $\sum_{i=7}^{14}(48i-201)$?

3. This section approximates definite integrals using what geometric shape?

4. T/F: A sum using the Right Hand Rule is an example of a Riemann Sum.

Problems

In Exercises 5–11, write out each term of the summation and compute the sum.

5. $\sum_{i=2}^{4} i^2$

6. $\sum_{i=-1}^{3}(4i-2)$

7. $\sum_{i=-2}^{2} \sin(\pi i/2)$

8. $\sum_{i=1}^{5} \frac{1}{i}$

9. $\sum_{i=1}^{6}(-1)^i i$

10. $\sum_{i=1}^{4}\left(\frac{1}{i} - \frac{1}{i+1}\right)$

11. $\sum_{i=0}^{5}(-1)^i \cos(\pi i)$

In Exercises 12–15, write each sum in summation notation.

12. $3 + 6 + 9 + 12 + 15$

13. $-1 + 0 + 3 + 8 + 15 + 24 + 35 + 48 + 63$

14. $\frac{1}{2} + \frac{2}{3} + \frac{3}{4} + \frac{4}{5}$

15. $1 - e + e^2 - e^3 + e^4$

In Exercises 16–22, evaluate the summation using Theorem 35.

16. $\sum_{i=1}^{25} i$

17. $\sum_{i=1}^{10}(3i^2 - 2i)$

18. $\sum_{i=1}^{15}(2i^3 - 10)$

19. $\sum_{i=1}^{10}(-4i^3 + 10i^2 - 7i + 11)$

20. $\sum_{i=1}^{10}(i^3 - 3i^2 + 2i + 7)$

21. $1 + 2 + 3 + \ldots + 99 + 100$

22. $1 + 4 + 9 + \ldots + 361 + 400$

Theorem 35 states

$$\sum_{i=1}^{n} a_i = \sum_{i=1}^{k} a_i + \sum_{i=k+1}^{n} a_i \text{ , so}$$

$$\sum_{i=k+1}^{n} a_i = \sum_{i=1}^{n} a_i - \sum_{i=1}^{k} a_i \text{ .}$$

Use this fact, along with other parts of Theorem 35, to evaluate the summations given in Exercises 23–26.

23. $\sum_{i=11}^{20} i$

24. $\sum_{i=16}^{25} i^3$

25. $\sum_{i=7}^{12} 4$

26. $\sum_{i=5}^{10} 4i^3$

In Exercises 27–30, express the limit as a definite integral.

27. $\lim_{n\to\infty} \frac{\pi}{n} \sum_{i=1}^{n} \frac{\sin \frac{\pi i}{n}}{1 + \frac{\pi i}{n}}$

28. $\lim_{n\to\infty} \sum_{i=1}^{n} \frac{3}{n}\left(2 + \frac{3i}{n}\right)\sqrt{1 + \left(2 + \frac{3i}{n}\right)^3}$

29. $\lim_{n\to\infty} \frac{5}{n} \sum_{i=1}^{n}\left(5\left(2 + \frac{5i}{n}\right)^3 - 4\left(2 + \frac{5i}{n}\right) + 7\right)$

30. $\lim_{n\to\infty} \frac{2}{n} \sum_{i=1}^{n} \frac{1 + \frac{2i}{n}}{\left(1 + \frac{2i}{n}\right)^2 + 4}$

In Exercises 31–33, express the definite integral as a limit of a sum.

31. $\int_{2}^{5} 4 - 2x\, dx$

32. $\int_{-2}^{0} x^2 + 3x\, dx$

33. $\int_{-\pi/2}^{\pi/2} \frac{\sin^3 x}{2 + \cos x}\, dx$

In Exercises 34–39, a definite integral $\int_{a}^{b} f(x)\, dx$ is given.

(a) Graph $f(x)$ on $[a, b]$.

(b) Add to the sketch rectangles using the provided rule.

(c) Approximate $\int_{a}^{b} f(x)\, dx$ by summing the areas of the rectangles.

34. $\int_{-3}^{3} x^2\, dx$, with 6 rectangles using the Left Hand Rule.

35. $\int_{0}^{2} (5 - x^2)\, dx$, with 4 rectangles using the Midpoint Rule.

36. $\int_{0}^{\pi} \sin x\, dx$, with 6 rectangles using the Right Hand Rule.

37. $\int_{1}^{3} \sqrt{10 - x^2}\, dx$ with 4 rectangles using the Right Hand Rule.

38. $\int_{1}^{2} \ln x\, dx$, with 3 rectangles using the Midpoint Rule.

39. $\int_{1}^{9} \frac{1}{x}\, dx$, with 4 rectangles using the Right Hand Rule.

In Exercises 40–45, a definite integral $\int_{a}^{b} f(x)\, dx$ is given. As demonstrated in Examples 6 and 7, do the following.

(a) Find a formula to approximate $\int_{a}^{b} f(x)\, dx$ using n subintervals and the provided rule.

(b) Evaluate the formula using $n = 10$, 100 and $1{,}000$.

(c) Find the limit of the formula, as $n \to \infty$, to find the exact value of $\int_{a}^{b} f(x)\, dx$.

40. $\int_{0}^{1} x^3\, dx$, using the Right Hand Rule.

41. $\int_{-1}^{1} 3x^2\, dx$, using the Left Hand Rule.

42. $\int_{-1}^{3} (3x - 1)\, dx$, using the Midpoint Rule.

43. $\int_{1}^{4} (2x^2 - 3)\, dx$, using the Left Hand Rule.

44. $\int_{-10}^{10} (5 - x)\, dx$, using the Right Hand Rule.

45. $\int_{0}^{1} (x^3 - x^2)\, dx$, using the Right Hand Rule.

46. Use six rectangles to approximate the area under the given graph of f from $x = 0$ to $x = 12$, using:

(a) The Left Hand Rule,

(b) The Right Hand Rule,

(c) The Midpoint Rule.
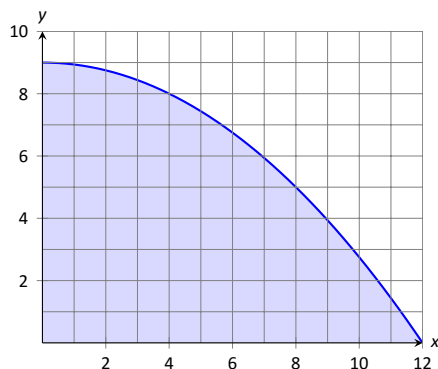
47. A car accelerates from 0 to 40 mph in 30 seconds. The speedometer reading at each 5 second interval during this time is given in the table below. Estimate how far the car travels during this 30 second period using the velocities at:

(a) The beginning of each time interval.

(b) The end of each time interval.

t (sec)	0	5	10	15	20	25	30
v (mph)	0	6	14	23	30	36	40

48. Use Theorems 35 and 36 to justify the remaining property in Theorem 34:

$$\int_{a}^{b} k \cdot f(x)\, dx = k \int_{a}^{b} f(x)\, dx$$

49. Use Theorems 35 and 36 to justify the remaining property in Theorem 37: If $f(x) \leq M$ for all x in $[a, b]$, then

$$\int_{a}^{b} f(x)\, dx \leq M(b - a).$$

Review

In Exercises 50–54, find an antiderivative of the given function.

50. $f(x) = 5 \sec^2 x$

51. $f(x) = \dfrac{7}{x}$

52. $g(t) = 4t^5 - 5t^3 + 8$

53. $g(t) = \cos t + \sin t$

54. $f(x) = \dfrac{1}{\sqrt{x}}$

5.4   The Fundamental Theorem of Calculus

In this section we will find connections between differential calculus (derivatives and antiderivatives) and integral calculus (definite integrals). These connections between the major ideas of calculus are important enough to be called the Fundamental Theorem of Calculus. These connections will also explain why we use the term indefinite integral for the set of all antiderivatives, and why we use such similar notations for antiderivatives and definite integrals.

Let $f(t)$ be a continuous function defined on $[a, b]$. The definite integral $\int_a^b f(x)\,dx$ is the "area under f" on $[a, b]$. We can turn this concept into a function by letting the upper (or lower) bound vary.

Let $F(x) = \int_a^x f(t)\,dt$. It computes the area under f on $[a, x]$ as illustrated in Figure 5.20. We can study this function using our knowledge of the definite integral. For instance, $F(a) = 0$ since $\int_a^a f(t)\,dt = 0$.

The first part of the Fundamental Theorem of Calculus tells us how to find derivatives of these kinds of functions.
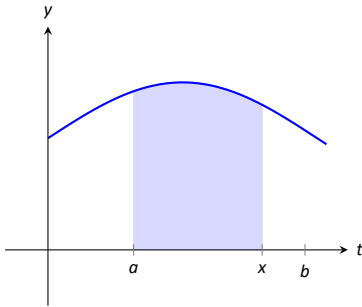
Figure 5.20: The area of the shaded region is $F(x) = \int_a^x f(t)\,dt$.

Theorem 38   The Fundamental Theorem of Calculus, Part 1

Let f be continuous on $[a, b]$ and let $F(x) = \int_a^x f(t)\,dt$. Then F is a differentiable function on (a, b), and

$$F'(x) = f(x).$$

Proof

In order to see why this is true, we must compute $\lim\limits_{h \to 0} \dfrac{F(x + h) - F(x)}{h}$. Suppose x and $x + h$ are in $[a, b]$. Theorem 34 implies that

$$\int_a^{x+h} f(t)\,dt = \int_a^x f(t)\,dt + \int_x^{x+h} f(t)\,dt,$$

which we can rewrite as

$$\int_x^{x+h} f(t)\,dt = \int_a^{x+h} f(t)\,dt - \int_a^x f(t)\,dt.$$

This allows us to simplify the denominator of the difference quotient in our limit as follows:

$$F(x + h) - F(x) = \int_a^{x+h} f(t)\,dt - \int_a^x f(t)\,dt \quad \text{(by the definition of } F\text{)}$$

$$= \int_x^{x+h} f(t)\,dt,$$

Notes:

so we see that

$$\lim_{h \to 0} \frac{F(x+h) - F(x)}{h} = \lim_{h \to 0} \frac{1}{h} \int_{x}^{x+h} f(t)\, dt.$$

Assume for the moment that $h > 0$. Since x and $x + h$ are both in $[a, b]$ and f is continuous on $[a, b]$, f is also continuous on $[x, x + h]$. Applying the Extreme Value Theorem (Theorem 22), we know that f must have an absolute minimum value $f(u) = m$ and an absolute maximum value $f(v) = M$ on this interval. In other words, $m \leq f(t) \leq M$ whenever $x \leq t \leq x + h$. Using the Comparison Properties of Integrals, we can now say that

$$\int_{x}^{x+h} m\, dt \leq \int_{x}^{x+h} f(t)\, dt \leq \int_{x}^{x+h} M\, dt.$$

Computing the outer integrals, this becomes

$$m(x + h - x) \leq \int_{x}^{x+h} f(t)\, dt \leq M(x + h - x), \quad \text{or}$$

$$mh \leq \int_{x}^{x+h} f(t)\, dt \leq Mh.$$

Since $h > 0$, we may divide by h to obtain

$$f(u) = m \leq \frac{1}{h} \int_{x}^{x+h} f(t)\, dt \leq M = f(v).$$

Now suppose that $h < 0$. Preceding as before, we know that f has an absolute minimum value $f(u) = m$ and an absolute maximum value $f(v) = M$ on the interval $[x + h, x]$. We know that $m \leq f(t) \leq M$ whenever $x + h \leq t \leq x$, so we have

$$\int_{x+h}^{x} m\, dt \leq \int_{x+h}^{x} f(t)\, dt \leq \int_{x+h}^{x} M\, dt.$$

Once again we compute to obtain

$$-mh \leq \int_{x+h}^{x} f(t)\, dt \leq -Mh.$$

Since $-h > 0$, we can divide by $-h$ to obtain:

$$m \leq -\frac{1}{h} \int_{x+h}^{x} f(t)\, dt \leq M$$

$$f(u) = m \leq \frac{1}{h} \int_{x}^{x+h} f(t)\, dt \leq M = f(v) \quad \text{(using Theorem 34(2))}$$

Notes:

We are now ready to compute the desired limit,

$$\lim_{h \to 0} \frac{F(x+h) - F(x)}{h} = \lim_{h \to 0} \frac{1}{h} \int_x^{x+h} f(t)\, dt.$$

Whether $h > 0$ or $h < 0$, we know that

$$f(u) \le \frac{1}{h} \int_x^{x+h} f(t)\, dt \le f(v),$$

where u and v are both between x and $x + h$. Note that

$$\lim_{h \to 0} (x + h) = x \quad \text{and} \quad \lim_{h \to 0} x = x,$$

so the Squeeze Theorem (Theorem 5) says that

$$\lim_{h \to 0} u = x \quad \text{and} \quad \lim_{h \to 0} v = x.$$

Since f is continuous at x, we know that

$$\lim_{h \to 0} f(u) = f(x) \quad \text{and} \quad \lim_{h \to 0} f(v) = f(x).$$

Finally, we know that

$$f(u) \le \frac{1}{h} \int_x^{x+h} f(t)\, dt \le f(v),$$

so applying the Squeeze Theorem again tells us that

$$\lim_{h \to 0} \frac{1}{h} \int_x^{x+h} f(t)\, dt = f(x).$$

Therefore $F'(x) = f(x)$ as desired. $\qquad\square$

Watch the video:
Fundamental Theorem of Calculus Part 1 at
`https://youtu.be/PGmVvIglZx8`

Initially this seems simple, as demonstrated in the following example.

Notes:

Example 1   **Using the Fundamental Theorem of Calculus, Part 1**

Let $F(x) = \displaystyle\int_{-5}^{x} (t^2 + \sin t)\, dt$. What is $F'(x)$?

SOLUTION   Using the Fundamental Theorem of Calculus, we have
$F'(x) = x^2 + \sin x$.

This simple example reveals something incredible: $F(x)$ is an antiderivative of $x^2 + \sin x$. Therefore, $F(x) = \frac{1}{3}x^3 - \cos x + C$ for some value of C. (We can find C, but generally we do not care. We know that $F(-5) = 0$, which allows us to compute C. In this case, $C = \cos(-5) + \frac{125}{3}$.)

We have done more than found a complicated way of computing an antiderivative. Consider a function f defined on an open interval containing a, b and c. Suppose we want to compute $\int_a^b f(t)\, dt$. First, let $F(x) = \int_c^x f(t)\, dt$. Using the properties of the definite integral found in Theorem 34, we know

$$\int_a^b f(t)\, dt = \int_a^c f(t)\, dt + \int_c^b f(t)\, dt$$

$$= -\int_c^a f(t)\, dt + \int_c^b f(t)\, dt$$

$$= -F(a) + F(b)$$

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

We now see how indefinite integrals and definite integrals are related: we can evaluate a definite integral using antiderivatives. Furthermore, Theorem 31 told us that any other antiderivative G differs from F by a constant: $G(x) = F(x) + C$. This means that $G(b) - G(a) = (F(b) + C) - (F(a) + C) = F(b) - F(a)$, and the formula we've just found holds for any antiderivative. Consequently, it does not matter what value of C we use, and we might as well let $C = 0$. This proves the second part of the Fundamental Theorem of Calculus.

Theorem 39   The Fundamental Theorem of Calculus, Part 2

Let f be continuous on $[a, b]$ and let F be *any* antiderivative of f. Then

$$\int_a^b f(x)\, dx = F(b) - F(a).$$

Notes:

Watch the video:
The Fundamental Theorem of Calculus. Part 2 at
`https://youtu.be/nHnZVFeQvNQ`

Example 2   **Using the Fundamental Theorem of Calculus, Part 2**
We spent a great deal of time in the previous section studying $\int_0^4 (4x - x^2)\, dx$.
Using the Fundamental Theorem of Calculus, evaluate this definite integral.

   SOLUTION   We need an antiderivative of $f(x) = 4x - x^2$. All antideriva-
tives of f have the form $F(x) = 2x^2 - \frac{1}{3}x^3 + C$; for simplicity, choose $C = 0$.
   The Fundamental Theorem of Calculus states

$$\int_0^4 (4x - x^2)\, dx = F(4) - F(0) = \left(2(4)^2 - \frac{1}{3}4^3\right) - \left(0 - 0\right) = 32 - \frac{64}{3} = 32/3.$$

This is the same answer we obtained using limits in the previous section, just
with much less work.

Notation: A special notation is often used in the process of evaluating defi-
nite integrals using the Fundamental Theorem of Calculus. Instead of explicitly
writing $F(b) - F(a)$, the notation $F(x)\Big|_a^b$ is used. Thus the solution to Example 2
would be written as:

$$\int_0^4 (4x - x^2)\, dx = \left(2x^2 - \frac{1}{3}x^3\right)\Big|_0^4 = \left(2(4)^2 - \frac{1}{3}4^3\right) - \left(0 - 0\right) = 32/3.$$

Example 3   **Using the Fundamental Theorem of Calculus, Part 2**
Evaluate the following definite integrals.

1. $\int_{-2}^{2} x^3\, dx$ 2. $\int_0^{\pi} \sin x\, dx$ 3. $\int_0^5 e^t\, dt$ 4. $\int_4^9 \sqrt{u}\, du$ 5. $\int_1^5 2\, dx$

   SOLUTION

   1. $\int_{-2}^{2} x^3\, dx = \frac{1}{4}x^4 \Big|_{-2}^{2} = \left(\frac{1}{4}2^4\right) - \left(\frac{1}{4}(-2)^4\right) = 0.$

Notes:

2. $\displaystyle\int_0^\pi \sin x\,dx = -\cos x\Big|_0^\pi = -\cos\pi - \left(-\cos 0\right) = 1 + 1 = 2.$

   (This is interesting; it says that the area under one "hump" of a sine curve is 2.)

3. $\displaystyle\int_0^5 e^t\,dt = e^t\Big|_0^5 = e^5 - e^0 = e^5 - 1 \approx 147.41.$

4. $\displaystyle\int_4^9 \sqrt{u}\,du = \int_4^9 u^{\frac{1}{2}}\,du = \frac{2}{3}u^{\frac{3}{2}}\Big|_4^9 = \frac{2}{3}\left(9^{\frac{3}{2}} - 4^{\frac{3}{2}}\right) = \frac{2}{3}(27 - 8) = \frac{38}{3}.$

5. $\displaystyle\int_1^5 2\,dx = 2x\Big|_1^5 = 2(5) - 2 = 2(5 - 1) = 8.$

   This integral is interesting; the integrand is a constant function, hence we are finding the area of a rectangle with width $(5 - 1) = 4$ and height 2. Notice how the evaluation of the definite integral led to $2(4) = 8$.

   In general, if c is a constant, then $\int_a^b c\,dx = c(b - a)$.

The Fundamental Theorem of Calculus and the Chain Rule

Part 1 of the Fundamental Theorem of Calculus (FTC) states that given $F(x) = \displaystyle\int_a^x f(t)\,dt$, $F'(x) = f(x)$. Using other notation, $\dfrac{d}{dx}\left(F(x)\right) = f(x)$. While we have just practiced evaluating definite integrals, sometimes finding antiderivatives is impossible and we need to rely on other techniques to approximate the value of a definite integral. Functions written as $F(x) = \int_a^x f(t)\,dt$ are useful in such situations.

It may be of further use to compose such a function with another. As an example, we may compose $F(x)$ with $g(x)$ to get

$$F\big(g(x)\big) = \int_a^{g(x)} f(t)\,dt.$$

What is the derivative of such a function? The Chain Rule can be employed to state

$$\frac{d}{dx}\left(F\big(g(x)\big)\right) = F'\big(g(x)\big)g'(x) = f\big(g(x)\big)g'(x).$$

An example will help us understand this.

Example 4 **The FTC, Part 1, and the Chain Rule**

Find the derivative of $F(x) = \displaystyle\int_2^{x^2} \ln t\,dt.$

Notes:

SOLUTION   We can view $F(x)$ as being the function $G(x) = \displaystyle\int_2^x \ln t\, dt$ composed with $g(x) = x^2$; that is, $F(x) = G\big(g(x)\big)$. The Fundamental Theorem of Calculus states that $G'(x) = \ln x$. The Chain Rule gives us

$$
\begin{aligned}
F'(x) &= G'\big(g(x)\big)g'(x) \\
&= \ln(g(x))g'(x) \\
&= \ln(x^2)2x \\
&= 2x\ln x^2
\end{aligned}
$$

Normally, the steps defining $G(x)$ and $g(x)$ are skipped.

Practice this once more.

Example 5   **The FTC, Part 1, and the Chain Rule**

Find the derivative of $F(x) = \displaystyle\int_{\cos x}^5 t^3\, dt$.

SOLUTION   Note that $F(x) = -\displaystyle\int_5^{\cos x} t^3\, dt$. Viewed this way, the derivative of F is straightforward:

$$
F'(x) = \sin x \cos^3 x.
$$

Understanding Motion with the Fundamental Theorem of Calculus

We established, starting with Key Idea 3, that the derivative of a position function is a velocity function, and the derivative of a velocity function is an acceleration function. Now consider definite integrals of velocity and acceleration functions. Specifically, if $v(t)$ is a velocity function, what does $\displaystyle\int_a^b v(t)\, dt$ mean?

The Fundamental Theorem of Calculus states that

$$
\int_a^b v(t)\, dt = V(b) - V(a),
$$

where $V(t)$ is any antiderivative of $v(t)$. Since $v(t)$ is a velocity function, $V(t)$ must be a position function, and $V(b) - V(a)$ measures a change in position, or **displacement**.

Notes:

How would we measure total distance traveled? We have to consider the intervals when $v(t) \geq 0$ and when $v(t) \leq 0$. Therefore,

$$\text{total distance traveled} = \int_a^b |v(t)|\ dt.$$

Example 6   **Finding displacement and total distance traveled**

A ball is thrown straight up with velocity given by $v(t) = -32t + 20$ft/s, where t is measured in seconds. Find, and interpret,

$$1.\ \int_0^1 v(t)\ dt \qquad \text{and} \qquad 2.\ \int_0^1 |v(t)|\ dt.$$

SOLUTION

1. Using the Fundamental Theorem of Calculus, we have

$$\int_0^1 v(t)\ dt = \int_0^1 (-32t + 20)\ dt$$
$$= -16t^2 + 20t \Big|_0^1$$
$$= 4 \text{ ft.}$$

    Thus if a ball is thrown straight up into the air with velocity $v(t) = -32t + 20$, the height of the ball, 1 second later, will be 4 feet above the initial height. We will see in part 2. that the *distance traveled* is much farther. It has gone up to its peak and is falling down, but the difference between its height at $t = 0$ and $t = 1$ is 4 ft.

2. Here we are trying to find the total distance traveled by the ball. We must first consider where $v(t) > 0$ and $v(t) < 0$.

$$v(t) = -32t + 20 = 0$$
$$-32t = -20$$
$$t = \frac{5}{8}$$

Notes:
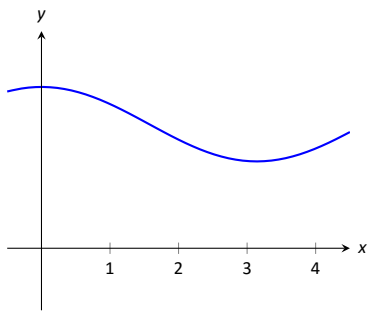
(a)
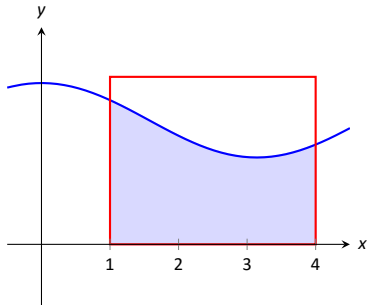
(b)
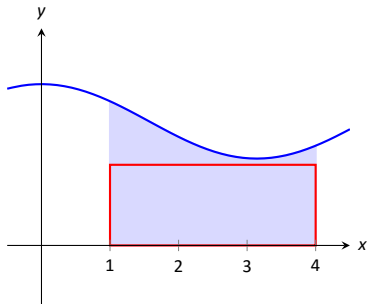
(c)
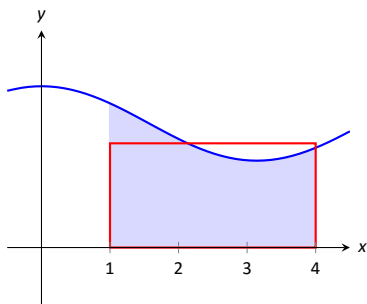
(d)

Figure 5.21: A graph of a function f to introduce the Mean Value Theorem and differently sized rectangles giving upper and lower bounds on $\int_1^4 f(x)\,dx$; the last rectangle matches the area exactly.

This means $v(t) > 0$ for $t < \frac{5}{8}$ and $v(t) < 0$ for $t > \frac{5}{8}$ so we have

$$\int_0^1 |v(t)|\ dt = \int_0^{5/8} v(t)\,dt + \int_{5/8}^1 -v(t)\,dt$$

$$= \int_0^{5/8} -32t + 20\,dt + \int_{5/8}^1 32t - 20\,dt$$

$$= \frac{34}{4} = 8.5 \text{ ft.}$$

Integrating a rate of change function gives total change. Velocity is the rate of position change; integrating velocity gives the total change of position, i.e., displacement.

Integrating a speed function gives a similar, though different, result. Speed is also the rate of position change, but does not account for direction. So integrating a speed function gives total change of position, without the possibility of "negative position change." Hence the integral of a speed function gives *distance traveled.*

As acceleration is the rate of velocity change, integrating an acceleration function gives total change in velocity. We do not have a simple term for this analogous to displacement. If $a(t) = 5\text{miles/h}^2$ and t is measured in hours, then

$$\int_0^3 a(t)\,dt = 15$$

means the velocity has increased by 15m/h from $t = 0$ to $t = 3$.

The Mean Value Theorem and Average Value

Consider the graph of a function f in Figure 5.21(a) and the area defined by $\int_1^4 f(x)\,dx$. Three rectangles are then drawn; in (b), the height of the rectangle is greater than f on $[1, 4]$, hence the area of this rectangle is is greater than $\int_1^4 f(x)\,dx$.

In (c), the height of the rectangle is smaller than f on $[1, 4]$, hence the area of this rectangle is less than $\int_1^4 f(x)\,dx$.

Finally, in (d) the height of the rectangle is such that the area of the rectangle is *exactly* that of $\int_1^4 f(x)\,dx$. Since rectangles that are "too big", as in (b), and rectangles that are "too little," as in (c), give areas greater/lesser than $\int_1^4 f(x)\,dx$, it makes sense that there is a rectangle, whose top intersects $f(x)$ somewhere on $[1, 4]$, whose area is *exactly* that of the definite integral.

We state this idea formally in a theorem.

Notes:

> **Theorem 40** **The Mean Value Theorem of Integration**
> Let f be continuous on $[a, b]$. There exists a value c in $[a, b]$ such that
>
> $$\int_a^b f(x)\, dx = f(c)(b - a).$$

This is an *existential* statement; c exists, but we do not provide a method of finding it. Theorem 40 is directly connected to the Mean Value Theorem of Differentiation, given as Theorem 24.

Proof
If $a = b$, then $\int_a^a f(x)\, dx = 0 = f(a)(a - a)$. Otherwise, we define the following for x in $[a, b]$:

$$F(x) = \int_a^x f(t)\, dt.$$

Applying Theorem 38 we know that F is differentiable on $[a, b]$ and that $F'(x) = f(x)$ for any x in $[a, b]$. We may now apply the Mean Value Theorem for Differentiation (Theorem 24) to see that there is a value c in (a, b) such that

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

Note that $F'(c) = f(c)$ and that $F(b) - F(a) = \int_a^b f(x)\, dx$ by Theorem 39. Therefore we can rewrite our equation as:

$$f(c) = \frac{\int_a^b f(x)\, dx}{b - a}, \text{ or}$$

$$f(c)(b - a) = \int_a^b f(x)\, dx. \qquad \square$$

We demonstrate the principles involved in this version of the Mean Value Theorem in the following example.

Example 7 **Using the Mean Value Theorem**

Consider $\int_0^\pi \sin x\, dx$. Find a value c guaranteed by the Mean Value Theorem.

SOLUTION We first need to evaluate $\int_0^\pi \sin x\, dx$. (This was previously done in Example 3.)

$$\int_0^\pi \sin x\, dx = -\cos x \Big|_0^\pi = 2.$$
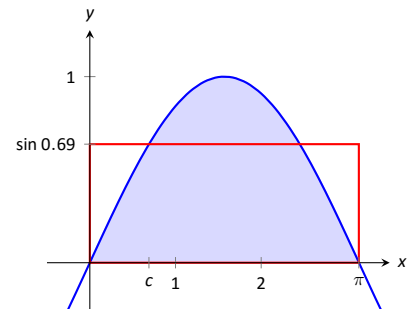
Figure 5.22: A graph of $y = \sin x$ on $[0, \pi]$ and the rectangle guaranteed by the Mean Value Theorem.

Notes:

Thus we seek a value c in $[0, \pi]$ such that $\pi \sin c = 2$.

$$\pi \sin c = 2 \quad \Rightarrow \quad \sin c = \frac{2}{\pi} \quad \Rightarrow \quad c = \sin^{-1}\left(\frac{2}{\pi}\right) \approx 0.69.$$

In Figure 5.22 $\sin x$ is sketched along with a rectangle with height $\sin(0.69)$. The area of the rectangle is the same as the area under $\sin x$ on $[0, \pi]$.

Let f be a function on $[a, b]$ with c such that $f(c)(b-a) = \int_a^b f(x)\, dx$. Consider $\int_a^b \big(f(x) - f(c)\big)\, dx$:

$$\int_a^b \big(f(x) - f(c)\big)\, dx = \int_a^b f(x) - \int_a^b f(c)\, dx$$
$$= f(c)(b - a) - f(c)(b - a)$$
$$= 0.$$

When $f(x)$ is shifted by $-f(c)$, the amount of area under f above the x–axis on $[a, b]$ is the same as the amount of area below the x–axis above f; see Figure 5.23 for an illustration of this. In this sense, we can say that $f(c)$ is the *average value* of f on $[a, b]$.

The value $f(c)$ is the average value in another sense. First, recognize that the Mean Value Theorem can be rewritten as

$$f(c) = \frac{1}{b - a} \int_a^b f(x)\, dx,$$

for some value of c in $[a, b]$. Next, partition the interval $[a, b]$ into n equally spaced subintervals, $a = x_1 < x_2 < \ldots < x_{n+1} = b$ and choose any c_i in $[x_i, x_{i+1}]$. The average of the numbers $f(c_1), f(c_2), \ldots, f(c_n)$ is:

$$\frac{1}{n}\Big(f(c_1) + f(c_2) + \ldots + f(c_n)\Big) = \frac{1}{n}\sum_{i=1}^{n} f(c_i).$$

Figure 5.23: On top, a graph of $y = f(x)$ and the rectangle guaranteed by the Mean Value Theorem. Below, $y = f(x)$ is shifted down by $f(c)$; the resulting "area under the curve" is 0.

Notes:

Multiply this last expression by 1 in the form of $\frac{(b-a)}{(b-a)}$:

$$\frac{1}{n}\sum_{i=1}^{n}f(c_i) = \sum_{i=1}^{n}f(c_i)\frac{1}{n}$$

$$= \sum_{i=1}^{n}f(c_i)\frac{1}{n}\frac{(b-a)}{(b-a)}$$

$$= \frac{1}{b-a}\sum_{i=1}^{n}f(c_i)\frac{b-a}{n}$$

$$= \frac{1}{b-a}\sum_{i=1}^{n}f(c_i)\Delta x \quad \text{(where } \Delta x = (b-a)/n)$$

Now take the limit as $n \to \infty$:

$$\lim_{n\to\infty}\frac{1}{b-a}\sum_{i=1}^{n}f(c_i)\Delta x \quad = \quad \frac{1}{b-a}\int_{a}^{b}f(x)\,dx \quad = \quad f(c).$$

This tells us this: when we evaluate f at n (somewhat) equally spaced points in $[a, b]$, the average value of these samples is $f(c)$ as $n \to \infty$.

This leads us to a definition.

Definition 25   **The Average Value of f on $[a, b]$**

Let f be continuous on $[a, b]$. The **average value of f on $[a, b]$** is $f(c)$, where c is a value in $[a, b]$ guaranteed by the Mean Value Theorem. I.e.,

$$\text{Average Value of } f \text{ on } [a, b] = \frac{1}{b-a}\int_{a}^{b}f(x)\,dx.$$

An application of this definition is given in the following example.

Example 8   **Finding the average value of a function**

An object moves back and forth along a straight line with a velocity given by $v(t) = (t-1)^2$ on $[0, 3]$, where t is measured in seconds and $v(t)$ is measured in ft/s.

What is the average velocity of the object?

SOLUTION   By our definition, the average velocity is:

$$\frac{1}{3-0}\int_{0}^{3}(t-1)^2\,dt = \frac{1}{3}\int_{0}^{3}\left(t^2 - 2t + 1\right)dt = \frac{1}{3}\left(\frac{1}{3}t^3 - t^2 + t\right)\Big|_{0}^{3} = 1 \text{ ft/s.}$$

Notes:

We can understand the above example through a simpler situation. Suppose you drove 100 miles in 2 hours. What was your average speed? The answer is simple: displacement/time = 100 miles/2 hours = 50 mph.

What was the displacement of the object in Example 8? We calculate this by integrating its velocity function: $\int_0^3 (t-1)^2\, dt = 3$ ft. Its final position was 3 feet from its initial position after 3 seconds: its average velocity was 1 ft/s.

This section has laid the groundwork for a lot of great mathematics to follow. The most important lesson is this: definite integrals can be evaluated using antiderivatives. Since the previous section established that definite integrals are the limit of Riemann sums, we can later create Riemann sums to approximate values other than "area under the curve," convert the sums to definite integrals, then evaluate these using the Fundamental Theorem of Calculus. This will allow us to compute the work done by a variable force, the volume of certain solids, the arc length of curves, and more.

The downside is this: generally speaking, computing antiderivatives is much more difficult than computing derivatives. Much of our time in Calculus II will be devoted to techniques of finding antiderivatives so that a wide variety of definite integrals can be evaluated.

Notes:

Exercises 5.4

Terms and Concepts

1. How are definite and indefinite integrals related?

2. What constant of integration is most commonly used when evaluating definite integrals?

3. T/F: If f is a continuous function, then $F(x) = \int_a^x f(t)\, dt$ is also a continuous function.

4. The definite integral can be used to find "the area under a curve." Give two other uses for definite integrals.

Problems

In Exercises 5–34, use the Fundamental Theorem of Calculus Part 2 to evaluate the definite integral.

5. $\int_1^3 (3x^2 - 2x + 1)\, dx$

6. $\int_0^4 (x - 1)^2\, dx$

7. $\int_{-1}^1 (x^3 - x^5)\, dx$

8. $\int_{\pi/2}^{\pi} \cos x\, dx$

9. $\int_0^{\pi/4} \sec^2 x\, dx$

10. $\int_1^e \frac{1}{x}\, dx$

11. $\int_{-2}^{-1} (4 - 2x^3)\, dx$

12. $\int_0^{\pi} (2 \cos x - 2 \sin x)\, dx$

13. $\int_1^3 e^x\, dx$

14. $\int_0^4 \sqrt{t}\, dt$

15. $\int_9^{25} \frac{1}{\sqrt{t}}\, dt$

16. $\int_1^8 \sqrt[3]{x}\, dx$

17. $\int_1^2 \frac{1}{x}\, dx$

18. $\int_1^2 \frac{1}{x^2}\, dx$

19. $\int_0^1 x^3\, dx$

20. $\int_0^1 x^{100}\, dx$

21. $\int_{-10}^{-5} 3\, dx$

22. $\int_{\pi/6}^{\pi/3} \csc x \cot x\, dx$

23. $\int_0^2 \left| x^2 - 1 \right|\, dx$

24. $\int_0^3 \left| 1 - 2x \right|\, dx$

25. $\int_{-1}^2 (u + 4)(2u + 1)\, du$

26. $\int_1^9 \frac{1 + \sqrt{x} + x}{\sqrt{x}}\, dx$

27. $\int_{\pi/7}^{\pi} \sin^2 x + \cos^2 x\, dx$

28. $\int_{-\pi/4}^{\pi/4} 2 + \tan^2 \theta\, d\theta$

29. $\int_0^{\pi/4} \sec t(\sec t + \tan t)\, dt$

30. $\int_{\pi/6}^{\pi/2} \frac{\sin 2x}{\sin x}\, dx$

31. $\int_1^4 \frac{4 + 6u}{\sqrt{u}}\, du$

32. $\int_0^{\pi/3} \frac{\sin \theta + \sin \theta \tan^2 \theta}{\sec^2}\, d\theta$

33. $\int_1^8 \frac{2 + t}{\sqrt[3]{t^2}}\, dt$

34. $\int_0^1 \sqrt[4]{x^5} + \sqrt[5]{x^4}\, dx$

35. Explain why:

    (a) $\int_{-1}^1 x^n\, dx = 0$, when n is a positive, odd integer, and

    (b) $\int_{-1}^1 x^n\, dx = 2\int_0^1 x^n\, dx$ when n is a positive, even integer.

In Exercises 36–39, find a value c guaranteed by the Mean Value Theorem.

36. $\int_0^2 x^2\, dx$

37. $\int_{-2}^2 x^2\, dx$

38. $\int_0^1 e^x\, dx$

39. $\int_0^{16} \sqrt{x}\, dx$

In Exercises 40–45, find the average value of the function on the given interval.

40. $f(x) = \sin x$ on $[0, \pi/2]$

41. $y = \sin x$ on $[0, \pi]$

42. $y = x$ on $[0, 4]$

43. $y = x^2$ on $[0, 4]$

44. $y = x^3$ on $[0, 4]$

45. $g(t) = 1/t$ on $[1, e]$

In Exercises 46–49, a velocity function of an object moving along a straight line is given. Find (a) the displacement of the object over the given time interval and (b) the total distance traveled by the object over the given time interval.

46. $v(t) = -32t + 20$ft/s on $[0, 5]$

47. $v(t) = -32t + 200$ft/s on $[0, 10]$

48. $v(t) = \cos t$ ft/s on $[0, 3\pi/2]$

49. $v(t) = \sqrt[4]{t}$ ft/s on $[0, 16]$

In Exercises 50–53, an acceleration function of an object moving along a straight line is given. Find the change of the object's velocity over the given time interval.

50. $a(t) = -32$ft/s^2 on $[0, 2]$

51. $a(t) = 10$ft/s^2 on $[0, 5]$

52. $a(t) = t$ ft/s^2 on $[0, 2]$

53. $a(t) = \cos t$ ft/s^2 on $[0, \pi]$

In Exercises 54–61, use the Fundamental Theorem of Calculus Part 1 to find $F'(x)$.

54. $F(x) = \int_2^{x^3+x} \dfrac{1}{t}\, dt$

55. $F(x) = \int_{x^3}^0 t^3\, dt$

56. $F(x) = \int_x^{x^2} (t+2)\, dt$

57. $F(x) = \int_{\ln x}^{e^x} \sin t\, dt$

58. $F(x) = \int_1^x \dfrac{\ln t + 4}{t^2 + 7}\, dt$

59. $F(x) = \int_2^{\sin x} \cos^3 t + 3\tan^3 t\, dt$

60. $F(x) = \int_{5x^3}^4 \dfrac{\sqrt{\cos t + 5}}{t^2 + e^t}\, dt$

61. $F(x) = \int_{\tan^2 x}^{10} \ln t + e^{t^2-7}\, dt$

62. Let $g(x) = \int_0^x f(t)\, dt$ where f is the function whose graph is shown below.

    (a) Evaluate $g(x)$ for $x = 0, 1, 2, 3, 4, 5, 6$.

    (b) Estimate $g(7)$.

    (c) Where does g have a minimum value? a maximum value?

    (d) Sketch the graph of g.
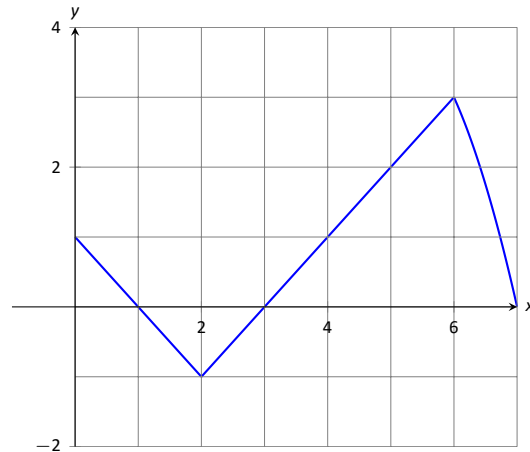
5.5   Substitution

We motivate this section with an example. Let $f(x) = (x^2 + 3x - 5)^{10}$. We can compute $f'(x)$ using the Chain Rule. It is:

$$f'(x) = 10(x^2 + 3x - 5)^9 \cdot (2x + 3) = (20x + 30)(x^2 + 3x - 5)^9.$$

Now consider this: What is $\int (20x + 30)(x^2 + 3x - 5)^9 \, dx$? We have the answer in front of us;

$$\int (20x + 30)(x^2 + 3x - 5)^9 \, dx = (x^2 + 3x - 5)^{10} + C.$$

How would we have evaluated this indefinite integral without starting with $f(x)$ as we did?

   This section explores *integration by substitution.* It allows us to "undo the Chain Rule." Substitution allows us to evaluate the above integral without knowing the original function first.

   The underlying principle is to rewrite a "complicated" integral of the form $\int f(x) \, dx$ as a not–so–complicated integral $\int h(u) \, du$. We'll formally establish later how this is done. First, consider again our introductory indefinite integral, $\int (20x + 30)(x^2 + 3x - 5)^9 \, dx$. Arguably the most "complicated" part of the integrand is $(x^2 + 3x - 5)^9$. We wish to make this simpler; we do so through a substitution. Let $u = x^2 + 3x - 5$. Thus

$$(x^2 + 3x - 5)^9 = u^9.$$

We have established u as a function of x, so now consider the differential of u:

$$du = (2x + 3)dx.$$

Keep in mind that $(2x+3)$ and dx are multiplied; the dx is not "just sitting there."

   Return to the original integral and do some substitutions through algebra:

$$\begin{aligned}
\int (20x + 30)(x^2 + 3x - 5)^9 \, dx &= \int 10(2x + 3)(x^2 + 3x - 5)^9 \, dx \\
&= \int 10\underbrace{(x^2 + 3x - 5)}_{u}{}^9 \underbrace{(2x + 3) \, dx}_{du} \\
&= \int 10u^9 \, du \\
&= u^{10} + C \quad \text{(replace u with $x^2 + 3x - 5$)} \\
&= (x^2 + 3x - 5)^{10} + C
\end{aligned}$$

Note: Recall from Section 4.3 that the differential of x, denoted dx, is any nonzero real number. If u is a function of x, then the differential of u, denoted du, is defined by $du = u'(x) \, dx$.

Notes:

One might well look at this and think "I (sort of) followed how that worked, but I could never come up with that on my own," but the process is learnable. This section contains numerous examples through which the reader will gain understanding and mathematical maturity enabling them to regard substitution as a natural tool when evaluating integrals.

We stated before that integration by substitution "undoes" the Chain Rule. Specifically, let $F(x)$ and $g(x)$ be differentiable functions and consider the derivative of their composition:

$$\frac{d}{dx}\Big(F\big(g(x)\big)\Big) = F'(g(x))g'(x).$$

Thus

$$\int F'(g(x))g'(x)\,dx = F(g(x)) + C.$$

Integration by substitution works by recognizing the "inside" function $g(x)$ and replacing it with a variable. By setting $u = g(x)$, we can rewrite the derivative as

$$\frac{d}{dx}\Big(F(u)\Big) = F'(u)u'.$$

Since $du = g'(x)dx$, we can rewrite the above integral as

$$\int F'(g(x))g'(x)\,dx = \int F'(u)du = F(u) + C = F(g(x)) + C.$$

This concept is important so we restate it in the context of a theorem.

Theorem 41   Integration by Substitution

Let F and g be differentiable functions, where the range of g is an interval I contained in the domain of F. Then

$$\int F'(g(x))g'(x)\,dx = F(g(x)) + C.$$

If $u = g(x)$, then $du = g'(x)dx$ and

$$\int F'(g(x))g'(x)\,dx = \int F'(u)\,du = F(u) + C = F(g(x)) + C.$$

The point of substitution is to make the integration step easy. Indeed, the step $\int F'(u)\,du = F(u) + C$ looks easy, as the antiderivative of the derivative of F is just F, plus a constant. The "work" involved is making the proper substitution. There is not a step–by–step process to memorize; rather, experience will be your guide. To gain experience, we now embark on many examples.

Notes:

Watch the video:
Integration by U-Substitution (Indefinite Integral) at
`https://youtu.be/li1SMPsqNuw`

Example 1 **Integrating by substitution**

Evaluate $\int x \sin(x^2 + 5) \, dx$.

SOLUTION Knowing that substitution is related to the Chain Rule, we choose to let u be the "inside" function of $\sin(x^2 + 5)$. (This is not *always* a good choice, but it is often the best place to start.)

Let $u = x^2 + 5$, hence $du = 2x \, dx$. The integrand has an $x \, dx$ term, but not a $2x \, dx$ term. (Recall that multiplication is commutative, so the x does not physically have to be next to dx for there to be an $x \, dx$ term.) We can divide both sides of the du expression by 2:

$$du = 2x \, dx \quad \Rightarrow \quad \frac{1}{2} du = x \, dx.$$

We can now substitute.

$$\int x \sin(x^2 + 5) \, dx = \int \sin(\underbrace{x^2 + 5}_{u}) \underbrace{x \, dx}_{\frac{1}{2} du}$$

$$= \int \frac{1}{2} \sin u \, du$$

$$= -\frac{1}{2} \cos u + C \quad \text{(now replace } u \text{ with } x^2 + 5\text{)}$$

$$= -\frac{1}{2} \cos(x^2 + 5) + C.$$

Thus $\int x \sin(x^2 + 5) \, dx = -\frac{1}{2} \cos(x^2 + 5) + C$. We can check our work by evaluating the derivative of the right hand side.

Example 2 **Integrating by substitution**

Evaluate $\int \cos(5x) \, dx$.

Notes:

SOLUTION   Again let u replace the "inside" function. Letting $u = 5x$, we have $du = 5dx$. Since our integrand does not have a $5dx$ term, we can divide the previous equation by 5 to obtain $\frac{1}{5}du = dx$. We can now substitute.

$$\int \cos(5x)\, dx = \int \cos(\underbrace{5x}_{u})\, \underbrace{dx}_{\frac{1}{5}du}$$

$$= \int \frac{1}{5} \cos u\, du$$

$$= \frac{1}{5} \sin u + C$$

$$= \frac{1}{5} \sin(5x) + C.$$

We can again check our work through differentiation.

The previous example exhibited a common, and simple, type of substitution. The "inside" function was a linear function (in this case, $y = 5x$). When the inside function is linear, the resulting integration is very predictable, so that we can say

$$\int F'(ax + b)\, dx = \frac{1}{a}F(ax + b) + C.$$

For example, $\int \sin(7x - 4)\, dx = -\frac{1}{7}\cos(7x - 4) + C$. Our next example can use this idea, but we will only employ it after going through all of the steps.

Example 3   **Integrating by substituting a linear function**

Evaluate $\displaystyle\int \frac{7}{-3x + 1}\, dx$.

SOLUTION   View this as a composition of functions $f(g(x))$, where $f(x) = 7/x$ and $g(x) = -3x + 1$. Employing our understanding of substitution, we let $u = -3x + 1$, the inside function. Thus $du = -3dx$. The integrand lacks a -3; hence divide the previous equation by -3 to obtain $-du/3 = dx$. We can now evaluate the integral through substitution.

$$\int \frac{7}{-3x + 1}\, dx = \int \left(\frac{7}{u}\right)\left(\frac{du}{-3}\right)$$

$$= \frac{-7}{3} \int \frac{du}{u}$$

$$= \frac{-7}{3} \ln|u| + C$$

$$= -\frac{7}{3} \ln|-3x + 1| + C.$$

Notes:

Not all integrals that benefit from substitution have a clear "inside" function. Several of the following examples will demonstrate ways in which this occurs.

Example 4 **Integrating by substitution**

Evaluate $\int \sin x \cos x \, dx$.

SOLUTION There is not a composition of function here to exploit; rather, just a product of functions. Do not be afraid to experiment; when given an integral to evaluate, it is often beneficial to think "If I let u be *this*, then du must be *that* ..." and see if this helps simplify the integral at all.

In this example, let's set $u = \sin x$. Then $du = \cos x \, dx$, which we have as part of the integrand. The substitution becomes very straightforward:

$$\int \sin x \cos x \, dx = \int u \, du$$
$$= \frac{1}{2} u^2 + C$$
$$= \frac{1}{2} \sin^2 x + C.$$

One would do well to ask "What would happen if we let $u = \cos x$?" The result is just as easy to find, yet looks very different. The challenge to the reader is to evaluate the integral letting $u = \cos x$ and discover why the answer is the same, yet looks different.

Our examples so far have required "basic substitution." The next example demonstrates how substitutions can be made that often strike the new learner as being "nonstandard."

Example 5 **Integrating by substitution**

Evaluate $\int x\sqrt{x + 3} \, dx$.

SOLUTION Recognizing the composition of functions, set $u = x + 3$. Then $du = dx$, giving what seems initially to be a simple substitution. But at this stage, we have:

$$\int x\sqrt{x + 3} \, dx = \int x\sqrt{u} \, du.$$

We cannot evaluate an integral that has both an x and an u in it. We need to convert the x to an expression involving just u.

Notes:

Since we set $u = x+3$, we can also state that $u-3 = x$. Thus we can replace x in the integrand with $u - 3$. It will also be helpful to rewrite $\sqrt{u}$ as $u^{\frac{1}{2}}$.

$$\int x\sqrt{x+3}\, dx = \int (u-3)u^{\frac{1}{2}}\, du$$
$$= \int \left(u^{\frac{3}{2}} - 3u^{\frac{1}{2}} \right) du$$
$$= \frac{2}{5}u^{\frac{5}{2}} - 2u^{\frac{3}{2}} + C$$
$$= \frac{2}{5}(x+3)^{\frac{5}{2}} - 2(x+3)^{\frac{3}{2}} + C.$$

Checking your work is always a good idea. In this particular case, some algebra will be needed to make one's answer match the integrand in the original problem.

Example 6   **Integrating by substitution**

Evaluate $\displaystyle\int \frac{1}{x \ln x}\, dx$.

SOLUTION   This is another example where there does not seem to be an obvious composition of functions. The line of thinking used in Example 5 is useful here: choose something for u and consider what this implies du must be. If u can be chosen such that du also appears in the integrand, then we have chosen well.

Choosing $u = 1/x$ makes $du = -1/x^2\, dx$; that does not seem helpful. However, setting $u = \ln x$ makes $du = 1/x\, dx$, which is part of the integrand. Thus:

$$\int \frac{1}{x \ln x}\, dx = \int \underbrace{\frac{1}{\ln x}}_{1/u} \underbrace{\frac{1}{x}\, dx}_{du}$$
$$= \int \frac{1}{u}\, du$$
$$= \ln|u| + C$$
$$= \ln|\ln x| + C.$$

The final answer is interesting; the natural log of the natural log. Take the derivative to confirm this answer is indeed correct.

Notes:

Integrals Involving Trigonometric Functions

Section 8.2 delves deeper into integrals of a variety of trigonometric functions; here we use substitution to establish a foundation that we will build upon.

The next three examples will help fill in some missing pieces of our antiderivative knowledge. We know the antiderivatives of the sine and cosine functions; what about the other standard functions tangent, cotangent, secant and cosecant? We discover these next.

Example 7 **Integration by substitution: antiderivatives of** $\tan x$

Evaluate $\displaystyle\int \tan x \, dx$.

SOLUTION The previous paragraph established that we did not know the antiderivatives of tangent, hence we must assume that we have learned something in this section that can help us evaluate this indefinite integral.

Rewrite $\tan x$ as $\sin x / \cos x$. While the presence of a composition of functions may not be immediately obvious, recognize that $\cos x$ is "inside" the $1/x$ function. Therefore, we see if setting $u = \cos x$ returns usable results. We have that $du = -\sin x \, dx$, hence $-du = \sin x \, dx$. We can integrate:

$$
\begin{aligned}
\int \tan x \, dx &= \int \frac{\sin x}{\cos x} \, dx \\
&= \int \underbrace{\frac{1}{\cos x}}_{u} \underbrace{\sin x \, dx}_{-du} \\
&= \int \frac{-1}{u} \, du \\
&= -\ln|u| + C \\
&= -\ln|\cos x| + C.
\end{aligned}
$$

Some texts prefer to bring the -1 inside the logarithm as a power of $\cos x$, as in:

$$
\begin{aligned}
-\ln|\cos x| + C &= \ln\left|(\cos x)^{-1}\right| + C \\
&= \ln\left|\frac{1}{\cos x}\right| + C \\
&= \ln|\sec x| + C.
\end{aligned}
$$

Thus the result they give is $\int \tan x \, dx = \ln|\sec x| + C$. These two answers are equivalent.

Notes:

Example 8   **Integrating by substitution: antiderivatives of** $\sec x$

Evaluate $\int \sec x\, dx$.

SOLUTION   This example employs a wonderful trick: multiply the integrand by "1" so that we see how to integrate more clearly. In this case, we write "1" as

$$1 = \frac{\sec x + \tan x}{\sec x + \tan x}.$$

This may seem like it came out of left field, but it works beautifully. Consider:

$$\int \sec x\, dx = \int \sec x \cdot \frac{\sec x + \tan x}{\sec x + \tan x}\, dx$$

$$= \int \frac{\sec^2 x + \sec x \tan x}{\sec x + \tan x}\, dx.$$

Now let $u = \sec x + \tan x$; this means $du = (\sec x \tan x + \sec^2 x)\, dx$, which is our numerator. Thus:

$$= \int \frac{du}{u}$$

$$= \ln |u| + C$$

$$= \ln |\sec x + \tan x| + C.$$

We can use similar techniques to those used in Examples 7 and 8 to find antiderivatives of $\cot x$ and $\csc x$ (which the reader can explore in the exercises.) We summarize our results here.

Theorem 42   **Antiderivatives of Trigonometric Functions**

1. $\int \sin x\, dx = -\cos x + C$ 4. $\int \csc x\, dx = -\ln |\csc x + \cot x| + C$

2. $\int \cos x\, dx = \sin x + C$ 5. $\int \sec x\, dx = \ln |\sec x + \tan x| + C$

3. $\int \tan x\, dx = \ln |\sec x| + C$ 6. $\int \cot x\, dx = \ln |\sin x| + C$

Notes:

Simplifying the Integrand

It is common to be reluctant to manipulate the integrand of an integral; at first, our grasp of integration is tenuous and one may think that working with the integrand will improperly change the results. Integration by substitution works using a different logic: as long as *equality* is maintained, the integrand can be manipulated so that its *form* is easier to deal with. The next example demonstrates a common way in which using algebra first makes the integration easier to perform.

Example 9   **Integration by alternate methods**

Evaluate $\int \dfrac{x^2 + 2x + 3}{\sqrt{x}} \, dx$ with, and without, substitution.

SOLUTION   We already know how to integrate this particular example. Rewrite $\sqrt{x}$ as $x^{\frac{1}{2}}$ and simplify the fraction:

$$\frac{x^2 + 2x + 3}{x^{1/2}} = x^{\frac{3}{2}} + 2x^{\frac{1}{2}} + 3x^{-\frac{1}{2}}.$$

We can now integrate using the Power Rule:

$$\int \frac{x^2 + 2x + 3}{x^{1/2}} \, dx = \int \left(x^{\frac{3}{2}} + 2x^{\frac{1}{2}} + 3x^{-\frac{1}{2}} \right) \, dx$$

$$= \frac{2}{5} x^{\frac{5}{2}} + \frac{4}{3} x^{\frac{3}{2}} + 6x^{\frac{1}{2}} + C$$

This is a perfectly fine approach. We demonstrate how this can also be solved using substitution as its implementation is rather clever.

Let $u = \sqrt{x} = x^{\frac{1}{2}}$; therefore

$$du = \frac{1}{2} x^{-\frac{1}{2}} dx = \frac{1}{2\sqrt{x}} \, dx \quad \Rightarrow \quad 2du = \frac{1}{\sqrt{x}} \, dx.$$

This gives us $\int \dfrac{x^2 + 2x + 3}{\sqrt{x}} \, dx = \int (x^2 + 2x + 3) \cdot 2 \, du$. What are we to do with the other x terms? Since $u = x^{\frac{1}{2}}$, we have $u^2 = x$ and $u^4 = x^2$. We can then replace x^2 and x with appropriate powers of u. We thus have

$$\int \frac{x^2 + 2x + 3}{\sqrt{x}} \, dx = \int (x^2 + 2x + 3) \cdot 2 \, du$$

$$= \int 2(u^4 + 2u^2 + 3) \, du$$

$$= \frac{2}{5} u^5 + \frac{4}{3} u^3 + 6u + C$$

$$= \frac{2}{5} x^{\frac{5}{2}} + \frac{4}{3} x^{\frac{3}{2}} + 6x^{\frac{1}{2}} + C,$$

Notes:

which is obviously the same answer we obtained before. In this situation, substitution is arguably more work than our other method. The fantastic thing is that it works. It demonstrates how flexible integration is.

Substitution and Definite Integration

So far this section has focused on learning a new technique for finding antiderivatives. In practice, we will frequently be interested in finding definite integrals. We can use this antiderivative to evaluate the definite integral, but there is a more efficient method.

At its heart, (using the notation of Theorem 41) substitution converts integrals of the form $\int F'(g(x))g'(x)\, dx$ into an integral of the form $\int F'(u)\, du$ with the substitution of $u = g(x)$. The following theorem states how the bounds of a definite integral can be changed as the substitution is performed.

> **Theorem 43   Substitution with Definite Integrals**
> Let F and g be differentiable functions, where the range of g is an interval I that is contained in the domain of F. Then
> $$\int_a^b F'\big(g(x)\big)g'(x)\, dx = \int_{g(a)}^{g(b)} F'(u)\, du.$$

In effect, Theorem 43 states that once you convert to integrating with respect to u, you do not need to switch back to evaluating with respect to x. A few examples will help one understand.

Example 10   Definite integrals and substitution: changing the bounds
Evaluate $\displaystyle\int_0^2 \cos(3x - 1)\, dx$ using Theorem 43.

SOLUTION   Observing the composition of functions, let $u = 3x - 1$, hence $du = 3dx$. As $3dx$ does not appear in the integrand, divide the latter equation by 3 to get $du/3 = dx$.

By setting $u = 3x-1$, we are implicitly stating that $g(x) = 3x-1$. Theorem 43 states that the new lower bound is $g(0) = -1$; the new upper bound is $g(2) = 5$. We now evaluate the definite integral:

Notes:

$$\int_1^2 \cos(3x - 1)\, dx = \int_{-1}^5 \cos u \frac{du}{3}$$

$$= \left. \frac{1}{3} \sin u \right|_{-1}^5$$

$$= \frac{1}{3} \big(\sin 5 - \sin(-1) \big).$$

(a)

(b)

Figure 5.24: Graphing the areas defined by the definite integrals of Example 10.

Notice how once we converted the integral to be in terms of u, we never went back to using x.

The graphs in Figure 5.24 tell more of the story. In (a) the area defined by the original integrand is shaded, whereas in (b) the area defined by the new integrand is shaded. In this particular situation, the areas look very similar; the new region is "shorter" but "wider," giving the same area.

Example 11 **Definite integrals and substitution: changing the bounds**

Evaluate $\displaystyle\int_0^{\pi/2} \sin x \cos x\, dx$ using Theorem 43.

Solution We saw the corresponding indefinite integral in Example 4. In that example we set $u = \sin x$ but stated that we could have let $u = \cos x$. For variety, we do the latter here.

Let $u = g(x) = \cos x$, giving $du = -\sin x\, dx$ and hence $\sin x\, dx = -du$. The new upper bound is $g(\pi/2) = 0$; the new lower bound is $g(0) = 1$. Note how the lower bound is actually larger than the upper bound now. We have

$$\int_0^{\pi/2} \sin x \cos x\, dx = \int_1^0 -u\, du \qquad \text{(switch bounds \& change sign)}$$

$$= \int_0^1 u\, du$$

$$= \left. \frac{1}{2} u^2 \right|_0^1 = \frac{1}{2}.$$

In Figure 5.25 we have again graphed the two regions defined by our definite integrals. Unlike the previous example, they bear no resemblance to each other. However, Theorem 43 guarantees that they have the same area.
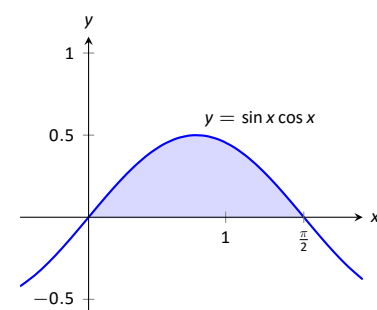
(a)
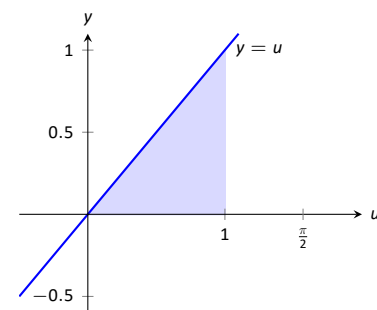
(b)

Figure 5.25: Graphing the areas defined by the definite integrals of Example 11.

Notes:

285

Example 12   **Definite integrals and substitution: changing the bounds**

Evaluate $\displaystyle\int_0^2 xe^{x^2+1}\,dx$ using Theorem 43.

SOLUTION   We note the composition of functions and let $u = x^2 + 1$, hence $du = 2x\,dx$. We divide the differential by 2 to get $\frac{du}{2} = x\,dx$.

Setting $g(x) = u = x^2 + 1$, we find that the new lower bound is $g(0) = 1$; the new upper bound is $g(2) = 5$. We now evaluate:

$$\int_0^2 xe^{x^2+1}\,dx = \int_1^5 e^u \frac{du}{2}$$
$$= \frac{1}{2}e^u \Big|_1^5$$
$$= \frac{1}{2}(e^5 - e^1)$$
$$= \frac{e}{2}(e^4 - 1).$$

Notes:

Exercises 5.5

Terms and Concepts

1. Substitution "undoes" what derivative rule?

2. T/F: One can sometimes use algebra to rewrite the integrand of an integral to make it easier to evaluate.

Problems

In Exercises 3–51, evaluate the indefinite integral.

3. $\int 3x^2 \left(x^3 - 5\right)^7 \, dx$

4. $\int (2x - 5) \left(x^2 - 5x + 7\right)^3 \, dx$

5. $\int x \left(x^2 + 1\right)^8 \, dx$

6. $\int (12x + 14) \left(3x^2 + 7x - 1\right)^5 \, dx$

7. $\int \dfrac{1}{2x + 7} \, dx$

8. $\int \dfrac{1}{\sqrt{2x + 3}} \, dx$

9. $\int \dfrac{x}{\sqrt{x + 3}} \, dx$

10. $\int \dfrac{x^3 - x}{\sqrt{x}} \, dx$

11. $\int \dfrac{e^{\sqrt{x}}}{\sqrt{x}} \, dx$

12. $\int \dfrac{x^4}{\sqrt{x^5 + 1}} \, dx$

13. $\int \dfrac{\frac{1}{x} + 1}{x^2} \, dx$

14. $\int \dfrac{\ln(x)}{x} \, dx$

15. $\int \sin^2(x) \cos(x) \, dx$

16. $\int \cos(3 - 6x) \, dx$

17. $\int \sec^2(4 - x) \, dx$

18. $\int \sec(2x) \, dx$

19. $\int \tan^2(x) \sec^2(x) \, dx$

20. $\int x \cos \left(x^2\right) \, dx$

21. $\int \cot x \, dx$. Do not just refer to Theorem 42 for the answer; justify it through Substitution.

22. $\int \csc x \, dx$. Do not just refer to Theorem 42 for the answer; justify it through Substitution.

23. $\int e^{3x - 1} \, dx$

24. $\int e^{x^3} x^2 \, dx$

25. $\int e^{x^2 - 2x + 1}(x - 1) \, dx$

26. $\int \dfrac{e^x + 1}{e^x} \, dx$

27. $\int \dfrac{e^x - e^{-x}}{e^{2x}} \, dx$

28. $\int \dfrac{\ln x}{x} \, dx$

29. $\int \dfrac{\left(\ln x\right)^2}{x} \, dx$

30. $\int \dfrac{\ln \left(x^3\right)}{x} \, dx$

31. $\int \dfrac{1}{x \ln \left(x^2\right)} \, dx$

32. $\int \dfrac{x^2}{\left(x^3 + 3\right)^2} \, dx$

33. $\int \left(3x^2 + 2x\right) \left(5x^3 + 5x^2 + 2\right)^8 \, dx$

34. $\int \dfrac{x}{\sqrt{1 - x^2}} \, dx$

35. $\int x^2 \csc^2 \left(x^3 + 1\right) \, dx$

36. $\int \sin(x) \sqrt{\cos(x)} \, dx$

37. $\int \dfrac{1}{x - 5} \, dx$

38. $\int \dfrac{7}{3x + 2} \, dx$

39. $\int \dfrac{2x + 7}{x^2 + 7x + 3} \, dx$

40. $\int \dfrac{9(2x + 3)}{3x^2 + 9x + 7} \, dx$

41. $\int \dfrac{3x - 3}{\sqrt{x^2 - 2x - 6}} \, dx$

42. $\int \dfrac{x - 3}{\sqrt{x^2 - 6x + 8}} \, dx$

43. $\int \dfrac{\cos \sqrt{x}}{\sqrt{x}} \, dx$

44. $\int \sec^2 \theta \tan \theta \, d\theta$

45. $\int x \sqrt{2x + 3} \, dx$

46. $\int \dfrac{x^3}{\left(x^2 + 1\right)^3} \, dx$

47. $\int \dfrac{2x^5}{x^2 + 1} \, dx$

48. $\int 3x^8 \left(x^3 + 2\right)^8 \, dx$

49. $\displaystyle\int \sin\left(\frac{x}{3}\right) dx$

50. $\displaystyle\int \sin^5\left(\frac{x}{4}\right) \cos\left(\frac{x}{4}\right) dx$

51. $\displaystyle\int x^{1/2} \cos(x^{3/2} + 1)\, dx$

In Exercises 52–62, evaluate the definite integral.

52. $\displaystyle\int_1^3 \frac{1}{x-5}\, dx$

53. $\displaystyle\int_2^6 x\sqrt{x-2}\, dx$

54. $\displaystyle\int_{-\pi/2}^{\pi/2} \sin^2 x \cos x\, dx$

55. $\displaystyle\int_0^1 2x(1-x^2)^4\, dx$

56. $\displaystyle\int_{-2}^{-1} (x+1)e^{x^2+2x+1}\, dx$

57. $\displaystyle\int_0^{\pi/4} e^{\tan x} \sec^2 x\, dx$

58. $\displaystyle\int_{-1}^1 \frac{x}{1+x^2}\, dx$

59. $\displaystyle\int_1^{\ln 3} \frac{e^x}{1+e^x}\, dx$

60. $\displaystyle\int_0^1 \frac{2x^2+1}{(2x^3+3x+2)^3}\, dx$

61. $\displaystyle\int_{-1}^2 \frac{x}{\sqrt{x+2}}\, dx$

62. $\displaystyle\int_0^{\frac{\pi}{4}} \cos^5(2x) \sin(2x)\, dx$

6: Applications of Integration

We begin this chapter with a reminder of a few key concepts from Chapter 5. Let f be a continuous function on $[a, b]$ which is partitioned into n equally spaced subintervals as

$$a = x_0 < x_1 < \cdots < x_{n-1} < x_n = b.$$

Let $\Delta x = (b - a)/n$ denote the length of the subintervals, and let c_i be any x-value in the ith subinterval. Definition 24 states that the sum

$$\sum_{i=1}^{n} f(c_i)\,\Delta x$$

is a *Riemann Sum.* Riemann Sums are often used to approximate some quantity (area, volume, work, pressure, etc.). The *approximation* becomes *exact* by taking the limit

$$\lim_{n \to \infty} \sum_{i=1}^{n} f(c_i)\,\Delta x.$$

Theorem 36 connects limits of Riemann Sums to definite integrals:

$$\lim_{n \to \infty} \sum_{i=1}^{n} f(c_i)\,\Delta x = \int_a^b f(x)\,dx.$$

Finally, the Fundamental Theorem of Calculus states how definite integrals can be evaluated using antiderivatives.

This chapter employs the following technique to a variety of applications. Suppose the value Q of a quantity is to be calculated. We first approximate the value of Q using a Riemann Sum, then find the exact value via a definite integral. We spell out this technique in the following Key Idea.
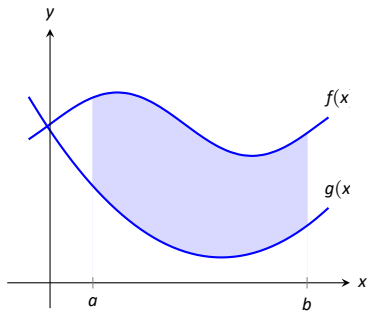
(a)
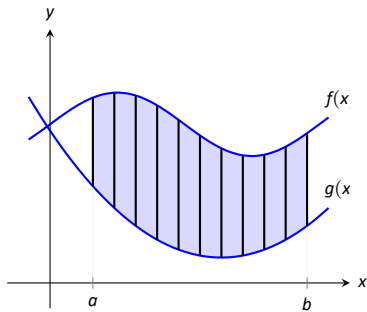
(b)
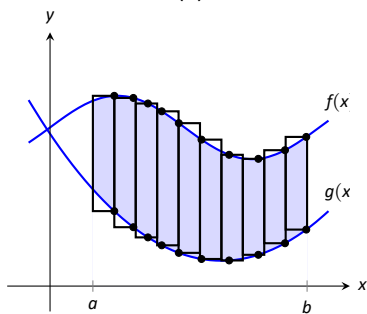
(c)

Figure 6.1: Subdividing a region into vertical slices and approximating the areas with rectangles.

> **Key Idea 11   Application of Definite Integrals Strategy**
>
> Let a quantity be given whose value Q is to be computed.
>
> 1. Divide the quantity into n smaller "subquantities" of value Q_i.
>
> 2. Identify a variable x and function $f(x)$ such that each subquantity can be approximated with the product $f(c_i)\Delta x$, where Δx represents a small change in x. Thus $Q_i \approx f(c_i)\Delta x$.
>
> 3. Recognize that $Q \approx \sum_{i=1}^{n} Q_i = \sum_{i=1}^{n} f(c_i)\Delta x$, which is a Riemann Sum.
>
> 4. Taking the appropriate limit gives $Q = \int_a^b f(x)\, dx$

This Key Idea will make more sense after we have had a chance to use it several times. We begin with Area Between Curves.

6.1   Area Between Curves

We are often interested in knowing the area of a region. Forget momentarily that we addressed this already in Section 5.4 and approach it instead using the technique described in Key Idea 11.

Let Q be the area of a region bounded by continuous functions f and g. If we break the region into many subregions, we have an obvious equation:

Total Area = sum of the areas of the subregions.

The issue to address next is how to systematically break a region into subregions. A graph will help. Consider Figure 6.1 (a) where a region between two curves is shaded. While there are many ways to break this into subregions, one particularly efficient way is to "slice" it vertically, as shown in Figure 6.1 (b), into n equally spaced slices.

We now approximate the area of a slice. Again, we have many options, but using a rectangle seems simplest. Picking any x-value c_i in the ith slice, we set the height of the rectangle to be $f(c_i) - g(c_i)$, the difference of the corresponding y-values. The width of the rectangle is a small difference in x-values, which we represent with Δx. Figure 6.1 (c) shows sample points c_i chosen in each subinterval and appropriate rectangles drawn. Each slice has an area approximately equal to $\big(f(c_i) - g(c_i)\big)\Delta x$; hence, the total area is approximately the Riemann

---

Notes:

Sum

$$Q \approx \sum_{i=1}^{n} \big(f(c_i) - g(c_i) \big) \Delta x.$$

Taking the limit as $n \to \infty$ gives the exact area as $\int_a^b \big(f(x) - g(x) \big) \, dx$.

Theorem 44   Area Between Curves

Let $f(x)$ and $g(x)$ be continuous functions defined on $[a, b]$ where $f(x) \geq g(x)$ for all x in $[a, b]$. The area of the region bounded by the curves $y = f(x)$, $y = g(x)$ and the lines $x = a$ and $x = b$ is

$$\int_a^b \big(f(x) - g(x) \big) \, dx.$$

Often, we do not know which function is greater (or they switch within the domain of integration). In that case, we can say that the area is $\int_a^b |f(x) - g(x)| \, dx$, which may involve dividing the domain of integration into pieces.

Watch the video:
Finding Areas Between Curves at
`https://youtu.be/DRFyNHdVgUA`

Example 1   Finding area enclosed by curves
Find the area of the region bounded by $f(x) = \sin x + 2$, $g(x) = \frac{1}{2} \cos(2x) - 1$, $x = 0$ and $x = 4\pi$, as shown in Figure 6.2.

SOLUTION   The graph verifies that the upper boundary of the region is given by f and the lower bound is given by g. Therefore the area of the region is the value of the integral

$$\int_0^{4\pi} \big(f(x) - g(x) \big) \, dx = \int_0^{4\pi} \left(\sin x + 2 - \left(\frac{1}{2} \cos(2x) - 1 \right) \right) dx$$

$$= -\cos x - \frac{1}{4} \sin(2x) + 3x \Big|_0^{4\pi}$$

$$= 12\pi \text{ units}^2.$$
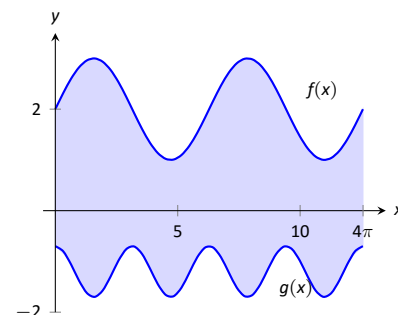
Figure 6.2: Graphing an enclosed region in Example 1.

Notes:
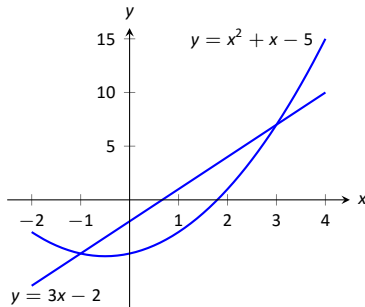
Figure 6.3: Sketching the region enclosed by $y = x^2 + x - 5$ and $y = 3x - 2$ in Example 2.

Example 2   Finding area between curves
Find the area of the region enclosed by $y = x^2 + x - 5$ and $y = 3x - 2$.

SOLUTION   It will help to sketch these two functions, as done in Figure 6.3. The region whose area we seek is completely bounded by these two functions; they seem to intersect at $x = -1$ and $x = 3$. To check, set $x^2 + x - 5 = 3x - 2$ and solve for x:

$$x^2 + x - 5 = 3x - 2$$
$$(x^2 + x - 5) - (3x - 2) = 0$$
$$x^2 - 2x - 3 = 0$$
$$(x - 3)(x + 1) = 0$$
$$x = -1,\ 3.$$

Following Theorem 44, the area is

$$\int_{-1}^{3} \left(3x - 2 - (x^2 + x - 5)\right)\, dx = \int_{-1}^{3} (-x^2 + 2x + 3)\, dx$$
$$= \left(-\frac{1}{3}x^3 + x^2 + 3x\right)\Big|_{-1}^{3}$$
$$= -\frac{1}{3}(27) + 9 + 9 - \left(\frac{1}{3} + 1 - 3\right)$$
$$= 10\frac{2}{3} = 10.\overline{6}$$

Figure 6.4: Graphing a region enclosed by two functions in Example 3.

Example 3   Finding total area enclosed by curves
Find the total area of the region enclosed by the functions $f(x) = -2x + 5$ and $g(x) = x^3 - 7x^2 + 12x - 3$ as shown in Figure 6.4.

SOLUTION   A quick calculation shows that $f = g$ at $x = 1, 2$ and 4. One can proceed thoughtlessly by computing $\int_{1}^{4} \left(f(x) - g(x)\right)\, dx$, but this ignores the fact that on $[1, 2]$, $g(x) > f(x)$. (In fact, the thoughtless integration returns $-9/4$, hardly the expected value of an *area*.) Thus we compute the total area by breaking the interval $[1, 4]$ into two subintervals, $[1, 2]$ and $[2, 4]$ and using the

Notes:

proper integrand in each.

$$\text{Total Area} = \int_1^2 \big(g(x) - f(x)\big)\, dx + \int_2^4 \big(f(x) - g(x)\big)\, dx$$

$$= \int_1^2 \big(x^3 - 7x^2 + 14x - 8\big)\, dx + \int_2^4 \big(-x^3 + 7x^2 - 14x + 8\big)\, dx$$

$$= \frac{5}{12} + \frac{8}{3}$$

$$= \frac{37}{12}\ \text{units}^2.$$

The previous example makes note that we are expecting area to be *positive*. When first learning about the definite integral, we interpreted it as "signed area under the curve," allowing for "negative area." That doesn't apply here; area is to be positive.

The previous example also demonstrates that we often have to break a given region into subregions before applying Theorem 44. The following example shows another situation where this is applicable, along with an alternate view of applying the Theorem.

Example 4   **Finding area: integrating with respect to y**
Find the area of the region enclosed by the functions $y = \sqrt{x} + 2$, $y = -(x - 1)^2 + 3$ and $y = 2$, as shown in Figure 6.5.

    SOLUTION   We give two approaches to this problem. In the first approach, we notice that the region's "top" is defined by two different curves. On $[0, 1]$, the top function is $y = \sqrt{x} + 2$; on $[1, 2]$, the top function is $y = -(x - 1)^2 + 3$. Thus we compute the area as the sum of two integrals:

$$\text{Total Area} = \int_0^1 \Big(\big(\sqrt{x} + 2\big) - 2\Big)\, dx + \int_1^2 \Big(\big(-(x-1)^2 + 3\big) - 2\Big)\, dx$$

$$= 2/3 + 2/3$$

$$= 4/3.$$

The second approach is clever and very useful in certain situations. We are used to viewing curves as functions of x; we input an x-value and a y-value is returned. Some curves can also be described as functions of y: input a y-value and an x-value is returned. We can rewrite the equations describing the boundary by solving for x:

$$y = \sqrt{x} + 2 \quad \Rightarrow \quad x = (y - 2)^2$$

$$y = -(x - 1)^2 + 3 \quad \Rightarrow \quad x = \sqrt{3 - y} + 1.$$
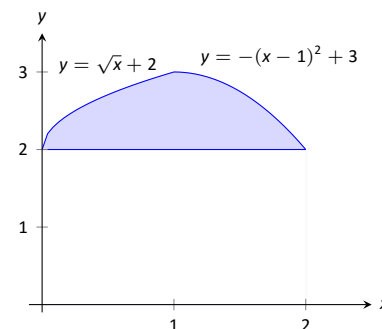
Figure 6.5: Graphing a region for Example 4.

Notes:
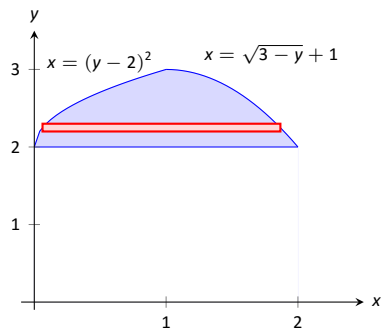
Figure 6.6: The region used in Example 4 with boundaries relabeled as functions of y.

Figure 6.6 shows the region with the boundaries relabeled. A horizontal rectangle is also pictured. The width of the rectangle is a small change in y: Δy. The height of the rectangle is a difference in x-values. The "top" x-value is the largest value, i.e., the rightmost. The "bottom" x-value is the smaller, i.e., the leftmost. Therefore the height of the rectangle is

$$\left(\sqrt{3-y}+1\right)-(y-2)^2.$$

The area is found by integrating the above function with respect to y with the appropriate bounds. We determine these by considering the y-values the region occupies. It is bounded below by $y=2$, and bounded above by $y=3$. That is, both the "top" and "bottom" functions exist on the y interval $[2,3]$. Thus

$$\text{Total Area} = \int_2^3 \left(\sqrt{3-y}+1-(y-2)^2\right)\,dy$$
$$= \left(-\frac{2}{3}(3-y)^{3/2}+y-\frac{1}{3}(y-2)^3\right)\Big|_2^3$$
$$= 4/3.$$

The important thing to notice is that by integrating with respect to y instead of x, we only had to do one integral and did not need to find the point at which to switch from one integration to another.

This calculus–based technique of finding area can be useful even with shapes that we normally think of as "easy." Example 5 computes the area of a triangle. While the formula "$\frac{1}{2} \times \text{base} \times \text{height}$" is well known, in arbitrary triangles it can be nontrivial to compute the height. Calculus makes the problem simple.

Example 5 Finding the area of a triangle
Compute the area of the regions bounded by the lines $y=3-x$, $y=x+1$ and $y=5x-15$, as shown in Figure 6.7.

SOLUTION Recognize that there are two "bottom" functions to this region, causing us to use two definite integrals.

$$\text{Total Area} = \int_1^3 \left((x+1)-(3-x)\right)\,dx + \int_3^4 \left((x+1)-(5x-15)\right)\,dx$$
$$= 4+2$$
$$= 6.$$

We can also approach this by converting each function into a function of y. This also requires 2 integrals, so there isn't really any advantage to doing so. We do it here for demonstration purposes.
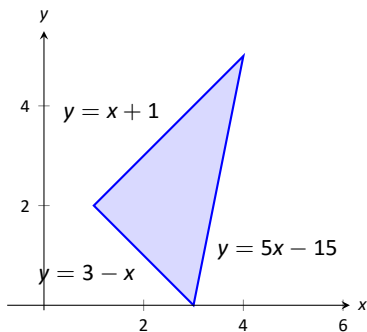
Figure 6.7: Graphing a triangular region in Example 5.

Notes:

The "top" function is always $x = \frac{y}{5} + 3$ while there are two "bottom" functions: $x = 3 - y$ and $x = y - 1$. Being mindful of the proper integration bounds, we have

$$\text{Total Area} = \int_0^2 \left(\left(\frac{y}{5} + 3\right) - (3 - y)\right)\, dy + \int_2^5 \left(\left(\frac{y}{5} + 3\right) - (y - 1)\right)\, dy$$
$$= \frac{12}{5} + \frac{18}{5}$$
$$= 6.$$

Of course, the final answer is the same (and we see that integrating with respect to x was probably easier, since it avoided fractions).

In the next section we apply Key Idea 11 to finding the volumes of certain solids.

Notes:

Exercises 6.1

Terms and Concepts

1. T/F: The area between curves is always positive.

2. T/F: Calculus can be used to find the area of basic geometric shapes.

3. In your own words, describe how to find the total area enclosed by $y = f(x)$ and $y = g(x)$.

Problems

In Exercises 4–9, find the area of the shaded region in the given graph.

4.

5.

6.

7.

8.

9.
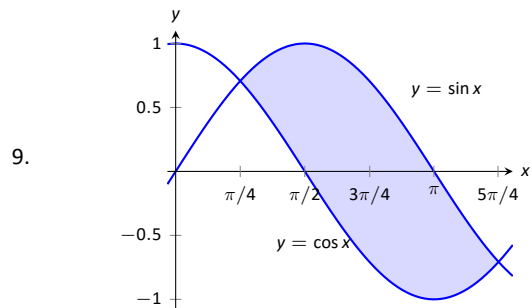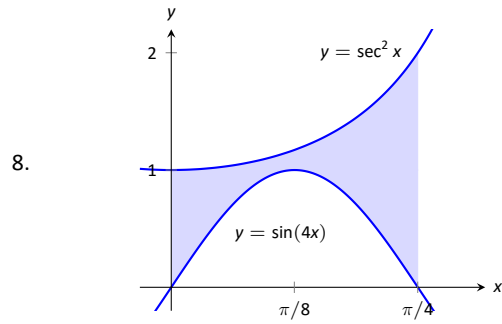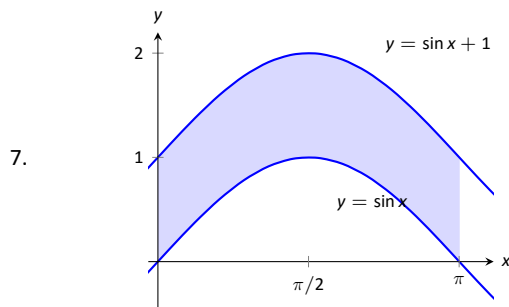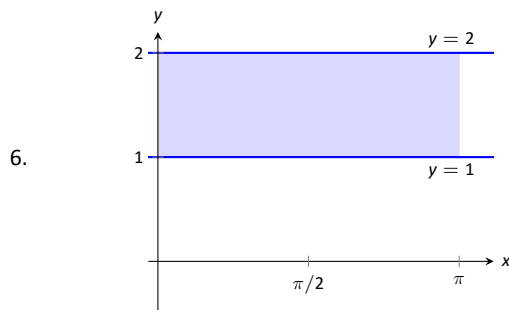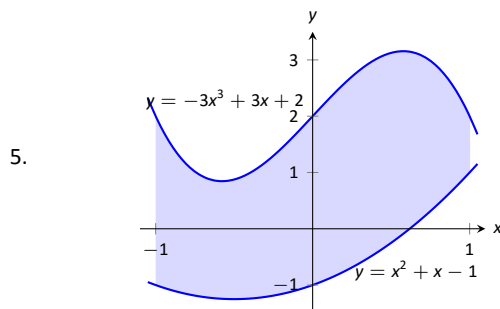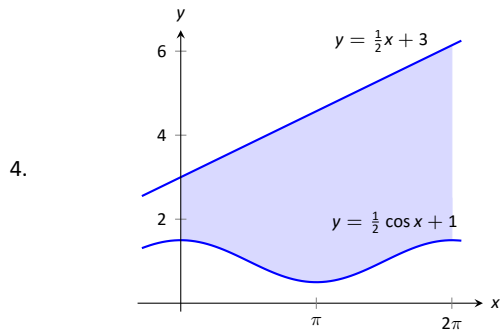
In Exercises 10–19, find the area of the region bounded by the given curves.

10. $f(x) = 2x^2 + 5x - 3$, $g(x) = x^2 + 4x - 1$

11. $f(x) = x^2 - 3x + 2$, $g(x) = -3x + 3$

12. $f(x) = \sin x$, $g(x) = 2x/\pi$

13. $f(x) = x^3 - 4x^2 + x - 1$, $g(x) = -x^2 + 2x - 4$

14. $f(x) = x$, $g(x) = \sqrt{x}$

15. $f(x) = -x^3 + 5x^2 + 2x + 1$, $g(x) = 3x^2 + x + 3$

16. $x = 2y^2$, $x + y = 1$

17. $x = y^2 - 1$, $x = 1 - y^2$

18. $4x + y^2 = 12$, $x = y$

19. $x = y^2 - 4y$, $x = 2y - y^2$

20. The functions $f(x) = \cos(2x)$ and $g(x) = \sin x$ intersect infinitely many times, forming an infinite number of repeated, enclosed regions. Find the areas of these regions.

In Exercises 21–25, find the area of the enclosed region in two ways:

1. by treating the boundaries as functions of x, and

2. by treating the boundaries as functions of y.

21.

22.

$y = \sqrt{x}$
$y = -2x + 3$
$y = -\frac{1}{2}x$

23.

$y = x + 2$
$y = x^2$

24.

$x = -\frac{1}{2}y + 1$
$x = \frac{1}{2}y^2$

25.

$y = x^{1/3}$
$y = \sqrt{x - 1/2}$
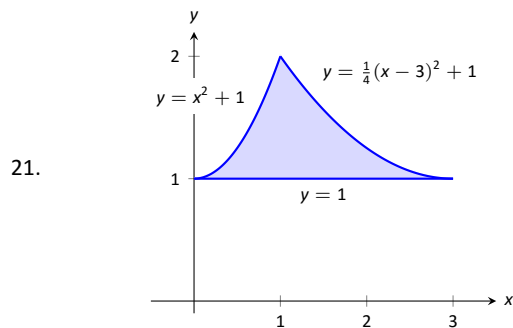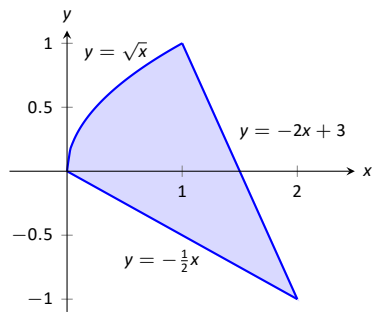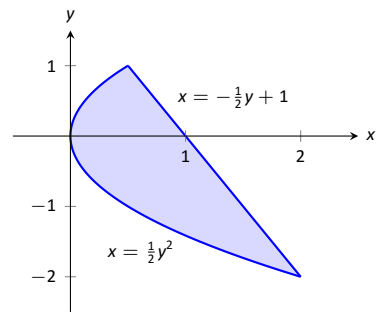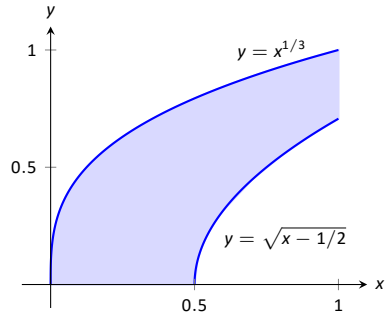
In Exercises 26–29, find the area triangle formed by the given three points.
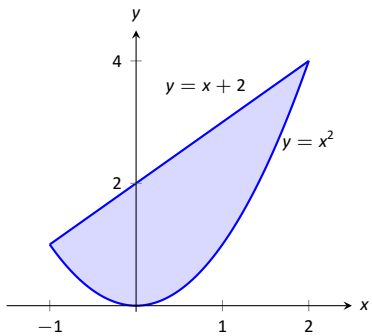
26. $(1, 1)$, $(2, 3)$, and $(3, 3)$

27. $(-1, 1)$, $(1, 3)$, and $(2, -1)$

28. $(1, 1)$, $(3, 3)$, and $(3, 3)$

29. $(0, 0)$, $(2, 5)$, and $(5, 2)$

6.2 Volume by Cross-Sectional Area; Disk and Washer Methods

The volume of a general right cylinder, as shown in Figure 6.8, is

$$\text{Area of the base} \times \text{height}.$$

We can use this fact as the building block in finding volumes of a variety of shapes.

Given an arbitrary solid, we can *approximate* its volume by cutting it into n thin slices. When the slices are thin, each slice can be approximated well by a general right cylinder. Thus the volume of each slice is approximately its cross-sectional area × thickness.

By orienting a solid along the x-axis, we can let $A(x_i)$ represent the cross-sectional area of the ith slice, and let Δx represent the thickness of the slices (the thickness is a small change in x). The total volume of the solid is approximately:

$$\text{Volume} \approx \sum_{i=1}^{n} \Big[\text{Area} \ \times \ \text{thickness} \Big]$$

$$= \sum_{i=1}^{n} A(x_i)\,\Delta x.$$

Recognize that this is a Riemann Sum. By taking a limit (as the thickness of the slices goes to 0) we can find the volume exactly.

$$\text{Volume} = \lim_{n \to \infty} \sum_{i=1}^{n} A(x_i)\,\Delta x$$

with $\Delta x = \frac{b-a}{n}$ and $x_i = a + i\Delta x$. We recognize this as a definite integral.

Theorem 45   Volume By Cross-Sectional Area

The volume V of a solid, oriented along the x-axis with cross-sectional area $A(x)$ from $x = a$ to $x = b$, is

$$V = \int_{a}^{b} A(x)\,dx.$$

Example 1   Finding the volume of a solid

Find the volume of a pyramid with a square base of side length 10 in and a height of 5 in.

base area $= A$
Volume $= A \cdot h$

Figure 6.8: The volume of a general right cylinder

Notes:

SOLUTION   There are many ways to "orient" the pyramid along the *x*-axis; Figure 6.9(a) gives one such way, with the pointed top of the pyramid at the origin and the *x*-axis going through the center of the base.

Each cross section of the pyramid is a square. To determine its area $A(x)$, we need to determine the side lengths of the square.

When $x = 5$, the square has side length 10; when $x = 0$, the square has side length 0. Since the edges of the pyramid are lines, it is easy to figure that each cross-sectional square has side length 2*x*, giving $A(x) = (2x)^2 = 4x^2$.

If one were to cut a slice out of the pyramid at $x = 3$, as shown in Figure 6.9(b), one would have a shape with square bottom and top with sloped sides. If the slice were thin, both the bottom and top squares would have sides lengths of about 6, and thus the cross–sectional area of the bottom and top would be about 36in^2. Letting Δx represent the thickness of the slice, the volume of this slice would then be about $36\Delta x\text{in}^3$.

Cutting the pyramid into *n* slices divides the total volume into *n* equally–spaced smaller pieces, each with volume $(2x_i)^2\Delta x$, where x_i is the approximate location of the slice along the *x*-axis and Δx represents the thickness of each slice. One can approximate total volume of the pyramid by summing up the volumes of these slices:

$$\text{Volume} \approx \sum_{i=1}^{n}(2x_i)^2\Delta x.$$

Taking the limit as $n \to \infty$ gives the actual volume of the pyramid; recognizing this sum as a Riemann Sum allows us to find the exact answer using a definite integral, matching the definite integral given by Theorem 45.

We have

$$V = \lim_{n\to\infty}\sum_{i=1}^{n}(2x_i)^2\Delta x$$

$$= \int_0^5 4x^2\,dx$$

$$= \frac{4}{3}x^3\Big|_0^5$$

$$= \frac{500}{3}\text{ in}^3.$$

We can check our work by consulting the general equation for the volume of a pyramid (see the back cover under "Volume of A General Cone"):

$$\tfrac{1}{3} \times \text{area of base} \times \text{height}.$$

Certainly, using this formula from geometry is faster than our new method, but the calculus–based method can be applied to much more than just cones.

(a)

(b)
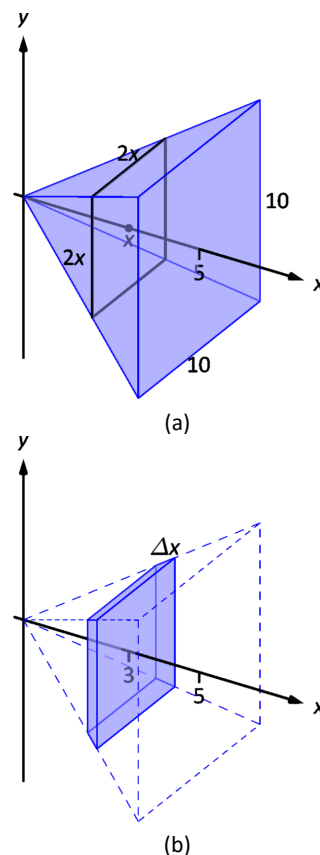
Figure 6.9: Orienting a pyramid along the *x*-axis (top) and cutting a slice in the pyramid (bottom) in Example 1.

Notes:

An important special case of Theorem 45 is when the solid is a **solid of revolution**, that is, when the solid is formed by rotating a shape around an axis.

Start with a function $y = f(x)$ from $x = a$ to $x = b$. Revolving this curve about a horizontal axis encloses a three-dimensional solid whose cross sections are disks (thin circles), perpendicular to the axis of rotation. Let $R(x)$ represent the radius of the cross-sectional disk at x; the area of this disk is $\pi[R(x)]^2$. Applying Theorem 45 gives the Disk Method.

Key Idea 12   The Disk Method

Let a solid be enclosed by revolving the curve $y = f(x)$ from $x = a$ to $x = b$ around a horizontal axis, and let $R(x)$ be the radius of the cross-sectional disk at x. The volume of the solid is

$$V = \pi \int_a^b [R(x)]^2 \, dx.$$

Watch the video:
Longer Version — Volumes using Disks/Washers at
`https://youtu.be/nZq0Kc067Z8`

Notes:

Example 2   **Finding volume using the Disk Method**

Find the volume of the solid formed by revolving about the x-axis the region bounded by the curves $y = 1/x$, $x = 1$, $x = 2$ and the x-axis.

SOLUTION   A sketch can help us understand this problem. In Figure 6.10(a) we have sketched the region we will be rotating. In Figure 6.10(b), the curve $y = 1/x$ is sketched along with the sample slice, a disk, at x with radius $R(x) = 1/x$. In Figure 6.10(c) the whole solid is pictured, along with the sample slice.

The volume of the sample slice shown in part (b) of the figure is approximately $\pi R(x_i)^2 \Delta x$, where $R(x_i)$ is the radius of the disk shown and Δx is the thickness of that slice. The radius $R(x_i)$ is the distance from the x-axis to the curve, hence $R(x_i) = 1/x_i$.

Slicing the solid into n equally–spaced slices, we can approximate the total volume by adding up the approximate volume of each slice:

$$\text{Approximate volume} = \sum_{i=1}^{n} \pi \left(\frac{1}{x_i} \right)^2 \Delta x.$$

Taking the limit of the above sum as $n \to \infty$ gives the actual volume; recognizing this sum as a Riemann sum allows us to evaluate the limit with a definite integral, which matches the formula given in Key Idea 12:

$$V = \lim_{n \to \infty} \sum_{i=1}^{n} \pi \left(\frac{1}{x_i} \right)^2 \Delta x$$

$$= \pi \int_{1}^{2} \left(\frac{1}{x} \right)^2 dx$$

$$= \pi \int_{1}^{2} \frac{1}{x^2} dx$$

$$= \pi \left[-\frac{1}{x} \right] \Big|_{1}^{2}$$

$$= \pi \left[-\frac{1}{2} - (-1) \right]$$

$$= \frac{\pi}{2} \text{ units}^3.$$

(a)

(b)

(c)
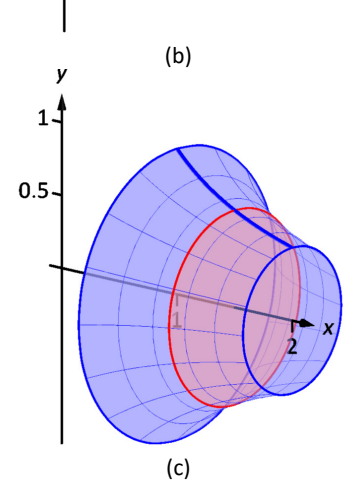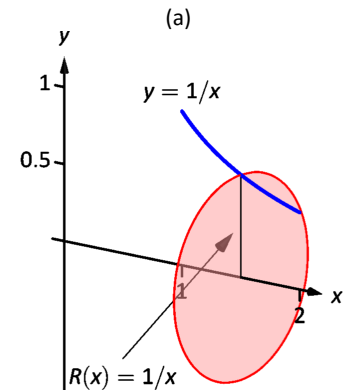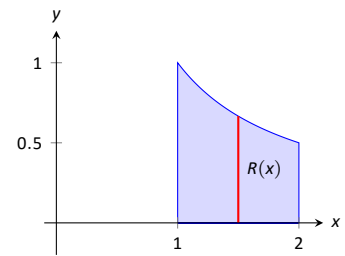
Figure 6.10: Sketching a solid in Example 2.

Notes:

(a)

(b)

(c)

Figure 6.11: Sketching a solid in Example 3.

While Key Idea 12 is given in terms of functions of x, the principle involved can be applied to functions of y when the axis of rotation is vertical, not horizontal. We demonstrate this in the next example.

Example 3   **Finding volume using the Disk Method**
Find the volume of the solid formed by revolving about the y-axis the region bounded by the curves $y = 1/x$, $y = 1$, $y = 0.5$, and the y-axis.

SOLUTION   Since the axis of rotation is vertical, our perpendicular cross sections have thickness Δy and radius $x = R(y)$. We need to convert the function into a function of y. Since $y = 1/x$ defines the curve, we rewrite it as $x = 1/y$.

Thus we are rotating about the y-axis the region bounded by the curves $x = 1/y$, $y = 1/2$, $y = 1$, and the y-axis to form a solid. The region of revolution is sketched in Figure 6.11(a), the curve and sample sample disk are sketched in Figure 6.11(b), and a full sketch of the solid is in Figure 6.11(b). We integrate to find the volume:

$$V = \pi \int_{1/2}^{1} \frac{1}{y^2} \, dy$$

$$= -\frac{\pi}{y} \Big|_{1/2}^{1}$$

$$= \pi \text{ units}^3.$$

We can also compute the volume of solids of revolution that have a hole in the center. The general principle is simple: compute the volume of the solid irrespective of the hole, then subtract the volume of the hole. If the outside radius of the solid is $R(x)$ and the inside radius (defining the hole) is $r(x)$, then the volume is

$$V = \pi \int_{a}^{b} [R(x)]^2 \, dx - \pi \int_{a}^{b} [r(x)]^2 \, dx = \pi \int_{a}^{b} \left([R(x)]^2 - [r(x)]^2 \right) \, dx.$$

One can generate a solid of revolution with a hole in the middle by revolving a region about an axis. Consider Figure 6.12(a), where a region is sketched along with a dashed, horizontal axis of rotation. By rotating the region about the axis, a solid is formed. Each cross section of this solid will be a washer (a disk with a hole in the center) as sketched in Figure 6.12(b). The outside of the washer has radius $R(x)$, whereas the inside has radius $r(x)$. The entire solid is sketched in Figure 6.12(c). This leads us to the Washer Method.

Notes:

Figure 6.12: Establishing the Washer Method.
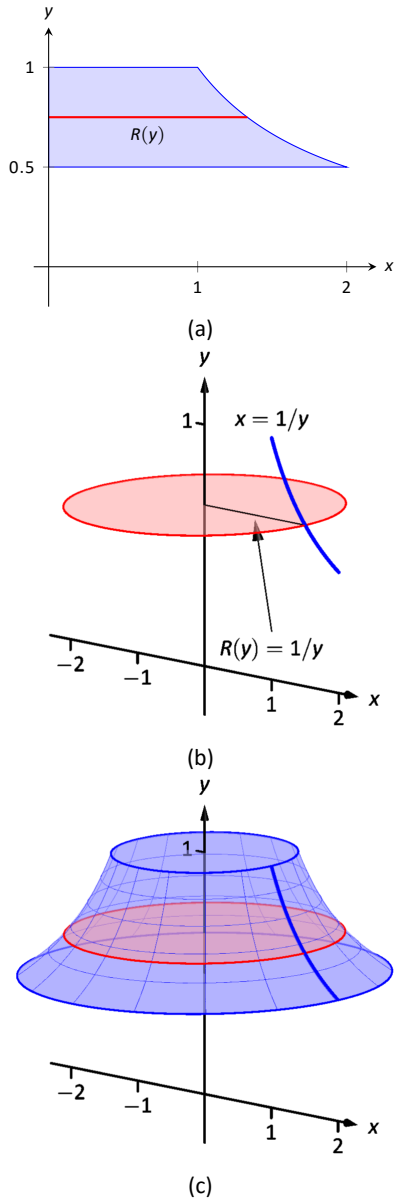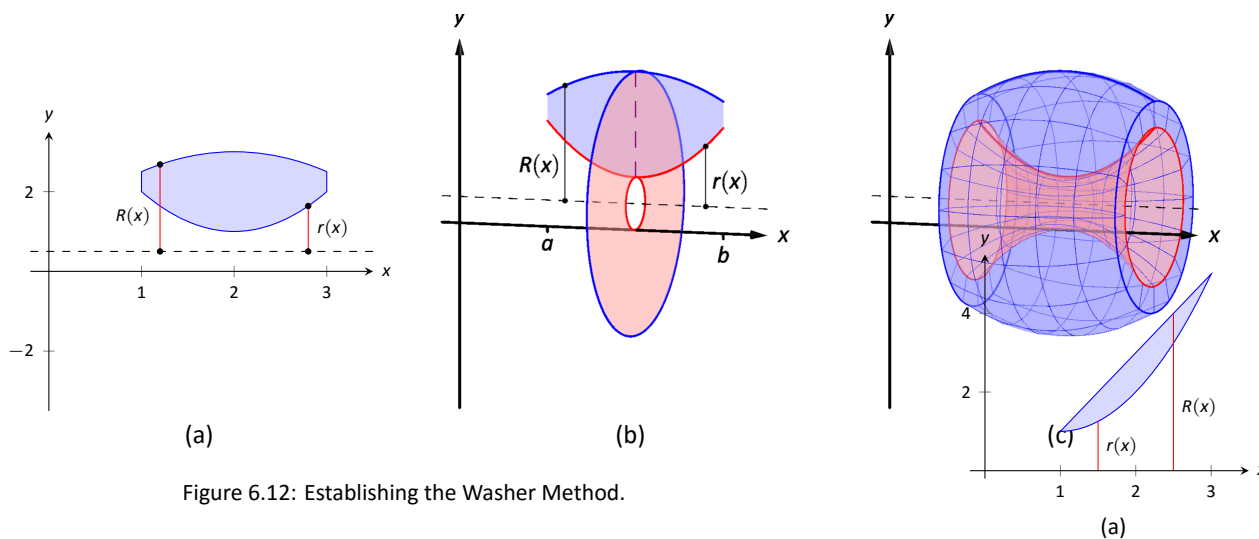
Key Idea 13   The Washer Method
Let a region bounded by $y = f(x)$, $y = g(x)$, $x = a$ and $x = b$ be rotated about a horizontal axis that does not intersect the region, forming a solid. Each cross section at x will be a washer with outside radius $R(x)$ and inside radius $r(x)$. The volume of the solid is

$$V = \pi \int_a^b \left([R(x)]^2 - [r(x)]^2 \right) dx.$$

Even though we introduced it first, the Disk Method is just a special case of the Washer Method with an inside radius of $r(x) = 0$.

Example 4   Finding volume with the Washer Method
Find the volume of the solid formed by rotating the region bounded by $y = x^2 - 2x + 2$ and $y = 2x - 1$ about the x-axis.

SOLUTION   A sketch of the region will help, as given in Figure 6.13(a). Rotating about the x-axis will produce cross sections in the shape of washers, as shown in Figure 6.13(b); the complete solid is shown in part (c). The outside radius of this washer is $R(x) = 2x + 1$; the inside radius is $r(x) = x^2 - 2x + 2$. As the region is bounded from $x = 1$ to $x = 3$, we integrate as follows to compute
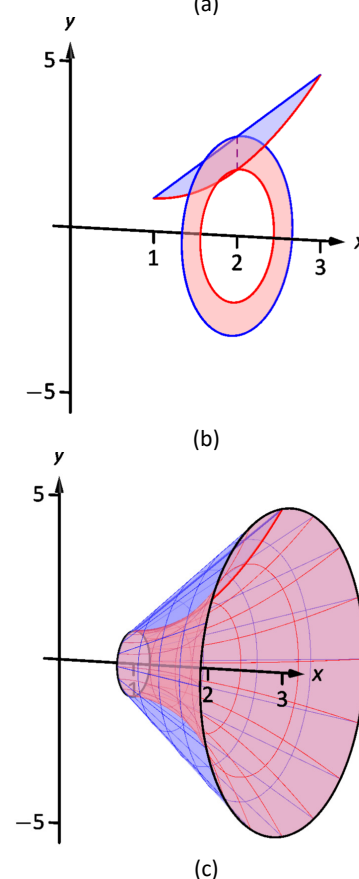
Figure 6.13: Sketching the region, a sample slice, and solid in Example 4.

Notes:

the volume.

$$V = \pi \int_1^3 \left((2x-1)^2 - (x^2 - 2x + 2)^2 \right) dx$$

$$= \pi \int_1^3 \left(-x^4 + 4x^3 - 4x^2 + 4x - 3 \right) dx$$

$$= \pi \left[-\frac{1}{5}x^5 + x^4 - \frac{4}{3}x^3 + 2x^2 - 3x \right]\Big|_1^3$$

$$= \frac{104}{15}\pi \text{ units}^3.$$

When rotating about a vertical axis, the outside and inside radius functions must be functions of y.

Example 5   **Finding volume with the Washer Method**
Find the volume of the solid formed by rotating the region bounded by $y = x^2$ and $x = y^2$ about the y-axis.

SOLUTION   In Figure 6.14 we have a sketch of the region (a), a sample slice (b), and the solid (c). Rotating about the y-axis will produce cross sections in the shape of washers, as shown in Figure (not yet created); the complete solid is shown in part (c). Since the axis of rotation is vertical, each radius is a function of y. The outside radius of this washer is $R(y) = \sqrt{y}$ and the inside radius is $r(y) = y^2$. As the region is bounded from $y = 0$ to $y = 1$, we integrate as follows to compute the volume.

$$V = \pi \int_0^1 \left((\sqrt{y})^2 - (y^2)^2 \right) \, dy$$

$$= \pi \int_0^1 y - y^4 \, dy$$

$$= \pi \left[\frac{1}{2}y^2 - \frac{1}{5}y^5 \right]\Big|_0^1$$

$$= \frac{3\pi}{10} \text{units}^3.$$

Notes:

(a)

(b)

(c)
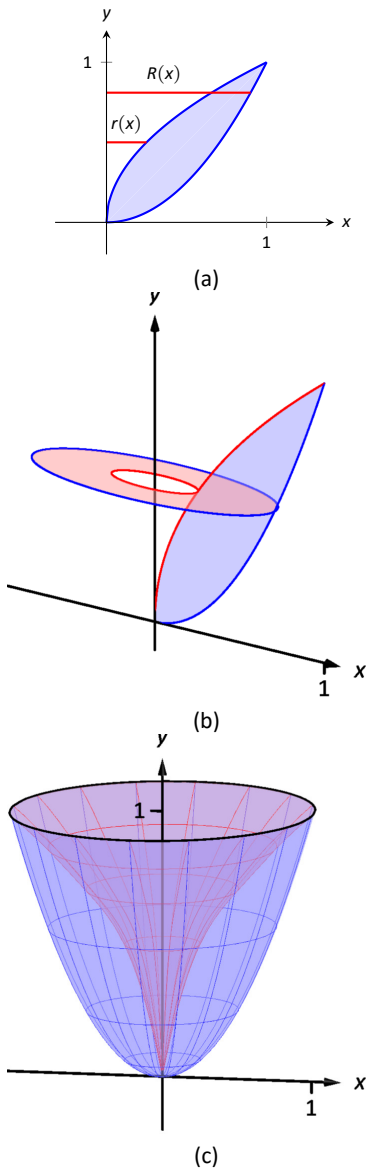
Figure 6.14: Sketching the region, a sample slice, and the solid in Example 5.

Example 6   **Finding volume with the Washer Method**

Find the volume of the solid formed by rotating the triangular region with vertices at $(1, 1)$, $(2, 1)$ and $(2, 3)$ about the y-axis.

     Solution   The triangular region is sketched in Figure 6.15(a); the sample slice is sketched in (b) and the full solid is drawn in (c). They help us establish the outside and inside radii. Since the axis of rotation is vertical, each radius is a function of y.

     The outside radius $R(y)$ is formed by the line connecting $(2, 1)$ and $(2, 3)$; it is a constant function, as regardless of the y-value the distance from the line to the axis of rotation is 2. Thus $R(y) = 2$.

     The inside radius is formed by the line connecting $(1, 1)$ and $(2, 3)$. The equation of this line is $y = 2x - 1$, but we need to refer to it as a function of y. Solving for x gives $r(y) = \frac{1}{2}(y + 1)$.

     We integrate over the y-bounds of $y = 1$ to $y = 3$. Thus the volume is

$$
\begin{aligned}
V &= \pi \int_1^3 \left(2^2 - \left(\frac{1}{2}(y + 1) \right)^2 \right) dy \\
&= \pi \int_1^3 \left(-\frac{1}{4}y^2 - \frac{1}{2}y + \frac{15}{4} \right) dy \\
&= \pi \left[-\frac{1}{12}y^3 - \frac{1}{4}y^2 + \frac{15}{4}y \right]\Big|_1^3 \\
&= \frac{10}{3}\pi \text{ units}^3.
\end{aligned}
$$

     In the previous examples, the axis of rotation has either been the x or y axis. We will now consider a problem where the axis of rotation is some other horizontal line.

(a)

(b)

(c)
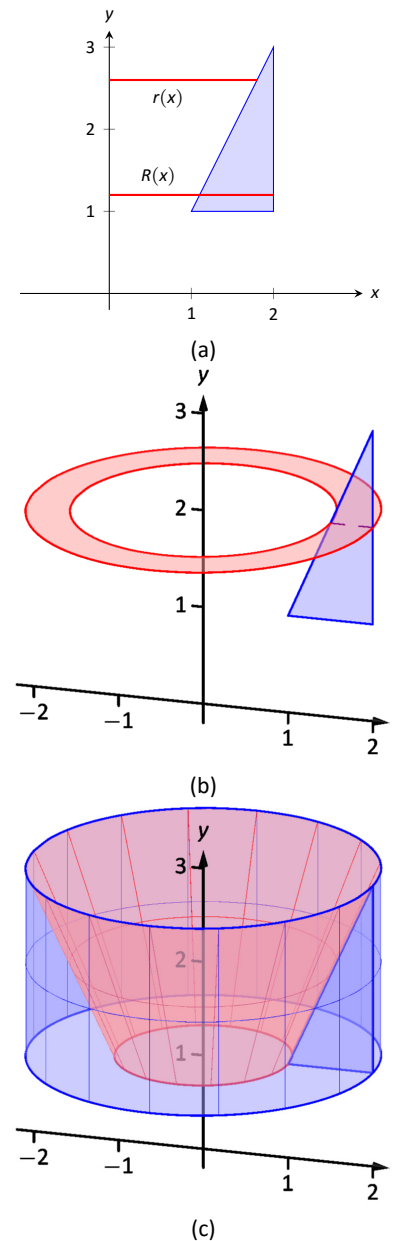
Figure 6.15: Sketching the region, a sample slice, and the solid in Example 6.

Notes:

(a)

(b)

(c)
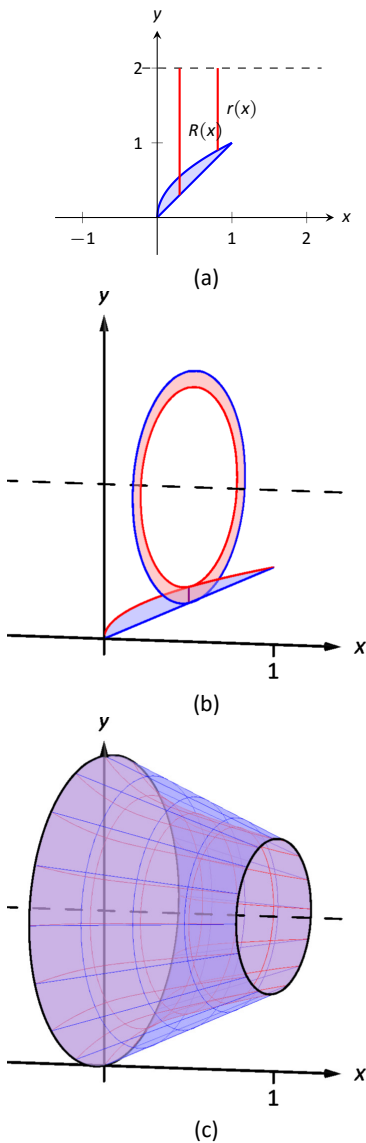
Figure 6.16: Sketching the solid in Example 7.

Example 7   **Finding volume with the Washer Method**

Find the volume of the solid formed by rotating the region bounded by $y = \sqrt{x}$ and $y = x$ about $y = 2$.

SOLUTION   Figure 6.16 shows the region we are rotating (a), a sample slice (b) and the full solid (c). The axis of rotation is horizontal so the radii must be functions of x. The radii is the distance from the axis of rotation to the curve so the outside radius of this washer is $R(x) = 2 - x$ and the inside radius is $r(x) = 2 - \sqrt{x}$. The region is bounded from $x = 0$ to $x = 1$, thus the volume is

$$V = \pi \int_0^1 \left((2-x)^2 - (2-\sqrt{x})^2 \right) \, dx$$

$$= \pi \int_0^1 (4 - 4x + x^2) - (4 - 4\sqrt{x} + x) \, dx$$

$$= \pi \int_0^1 x^2 - 5x + 4\sqrt{x} \, dx$$

$$= \pi \left[\frac{1}{3}x^3 - \frac{5}{2}x^2 + \frac{8}{3}x^{3/2} \right]\Bigg|_0^1$$

$$= \frac{\pi}{2} \text{units}^3.$$

This section introduced a new application of the definite integral. Our default view of the definite integral is that it gives "the area under the curve." However, we can establish definite integrals that represent other quantities; in this section, we computed volume.

The ultimate goal of this section is not to compute volumes of solids. That can be useful, but what is more useful is the understanding of this basic principle of integral calculus, outlined in Key Idea 11: to find the exact value of some quantity,

- we start with an approximation (in this section, slice the solid and approximate the volume of each slice),

- then make the approximation better by refining our original approximation (i.e., use more slices),

- then use limits to establish a definite integral which gives the exact value.

We practice this principle in the next section where we find volumes by slicing solids in a different way.

Notes:

Exercises 6.2

Terms and Concepts

1. T/F: A solid of revolution is formed by revolving a shape around an axis.

2. In your own words, explain how the Disk and Washer Methods are related.

3. Explain the how the units of volume are found in the integral of Theorem 45: if $A(x)$ has units of in^2, how does $\int A(x)\,dx$ have units of in^3?

Problems

In Exercises 4–7, a region of the Cartesian plane is shaded. Use the Disk/Washer Method to find the volume of the solid of revolution formed by revolving the region about the *x*-axis.

7.

4.

5.

6.
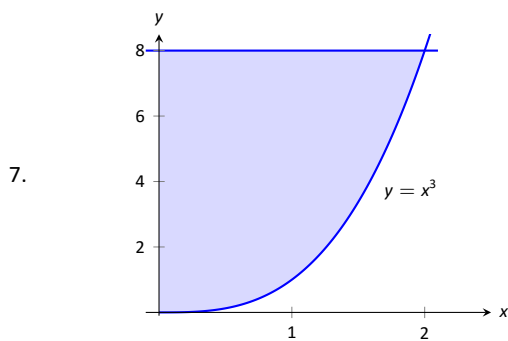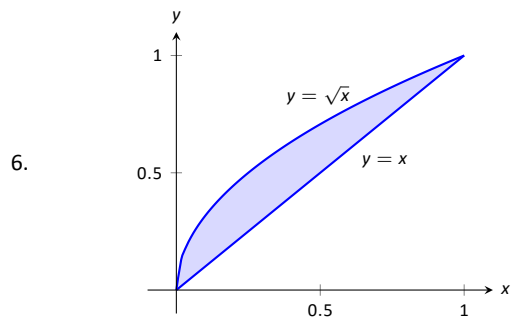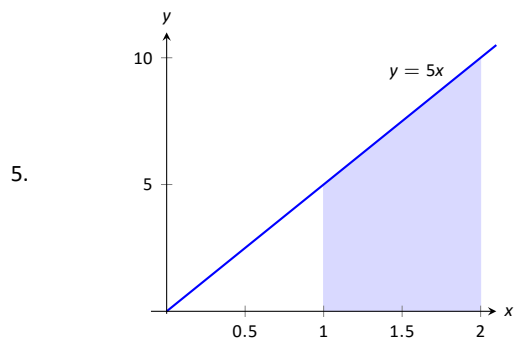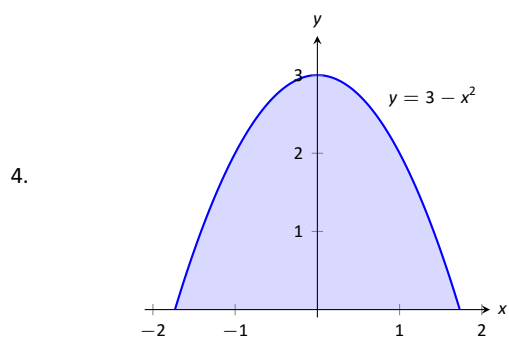
In Exercises 8–11, a region of the Cartesian plane is shaded. Use the Disk/Washer Method to find the volume of the solid of revolution formed by revolving the region about the *y*-axis.

8.

9.

10.

11.
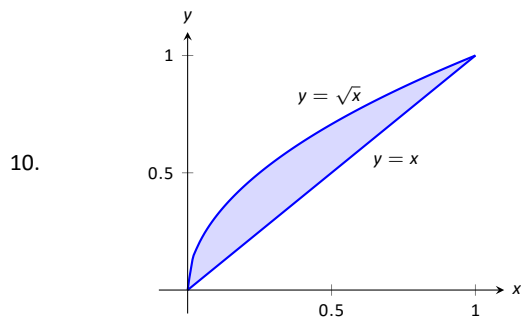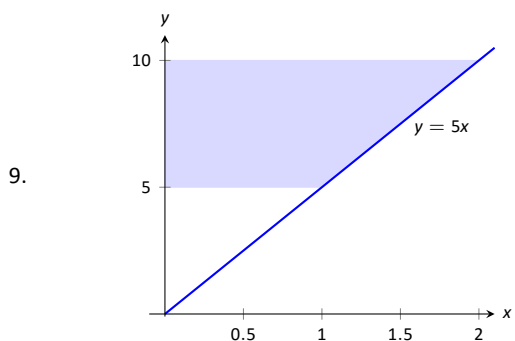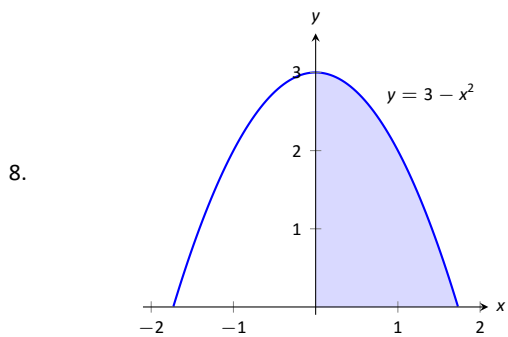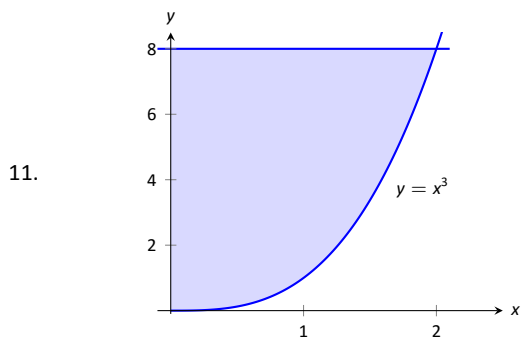
In Exercises 12–17, a region of the Cartesian plane is described. Use the Disk/Washer Method to find the volume of the solid of revolution formed by rotating the region about each of the given axes.

12. Region bounded by: $y = \sqrt{x}$, $y = 0$ and $x = 1$.
    Rotate about:

    (a) the x-axis   (c) the y-axis

    (b) $y = 1$ (d) $x = 1$

13. Region bounded by: $y = 4 - x^2$ and $y = 0$.
    Rotate about:

    (a) the x-axis   (c) $y = -1$

    (b) $y = 4$ (d) $x = 2$

14. The triangle with vertices $(1, 1)$, $(1, 2)$ and $(2, 1)$.
    Rotate about:

    (a) the x-axis   (c) the y-axis

    (b) $y = 2$ (d) $x = 1$

15. Region bounded by $y = x^2 - 2x + 2$ and $y = 2x - 1$.
    Rotate about:

    (a) the x-axis   (c) $y = 5$

    (b) $y = 1$

16. Region bounded by $y = 2x$, $y = x$ and $x = 2$.
    Rotate about:

    (a) the x-axis   (c) the y-axis

    (b) $y = 4$ (d) $x = 2$

17. Region bounded by $y = \cos x$, $x = 0$, $x = \dfrac{\pi}{4}$ and the x-axis.
    Rotate about:

    (a) the x-axis   (c) $y = -1$

    (b) $y = 1$

In Exercises 18–21, a solid is described. Orient the solid along the x-axis such that a cross-sectional area function $A(x)$ can be obtained, then apply Theorem 45 to find the volume of the solid.

18. A right circular cone with height of 10 and base radius of 5.
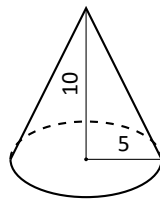
19. A skew right circular cone with height of 10 and base radius of 5. (Hint: all cross-sections are circles.)
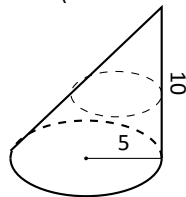
20. A right triangular cone with height of 10 and whose base is a right, isosceles triangle with side length 4.
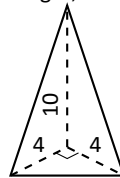
21. A solid with length 10 with a rectangular base and triangular top, wherein one end is a square with side length 5 and the other end is a triangle with base and height of 5.
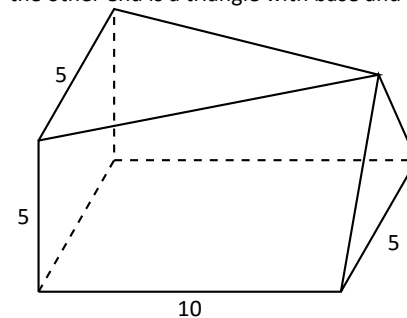
6.3 The Shell Method

Often a given problem can be solved in more than one way. A particular method may be chosen out of convenience, personal preference, or perhaps necessity. Ultimately, it is good to have options.

The previous section introduced the Disk and Washer Methods, which computed the volume of solids of revolution by integrating the cross–sectional area of the solid. This section develops another method of computing volume, the **Shell Method.** Instead of slicing the solid perpendicular to the axis of rotation creating cross-sections, we now slice it parallel to the axis of rotation, creating "shells."

Consider Figure 6.17, where the region shown in (a) is rotated around the y-axis forming the solid shown in (c). A small slice of the region is drawn in (a), *parallel* to the axis of rotation. When the region is rotated, this thin slice forms a **cylindrical shell**, as pictured in part (b) of the figure. The previous section approximated a solid with lots of thin disks (or washers); we now approximate a solid with many thin cylindrical shells.

To compute the volume of one shell, first consider the paper label on a soup can with radius r and height h. What is the area of this label? A simple way of determining this is to cut the label and lay it out flat, forming a rectangle with height h and length $2\pi r$. Thus the area is $A = 2\pi rh$; see Figure 6.18(a).

Do a similar process with a cylindrical shell, with height h, thickness Δx, and approximate radius r. Cutting the shell and laying it flat forms a rectangular solid with length $2\pi r$, height h and depth Δx. Thus the volume is $V \approx 2\pi rh\Delta x$; see Figure 6.18(b). (We say "approximately" since our radius was an approximation.)

By breaking the solid into n cylindrical shells, we can approximate the volume of the solid as

$$V \approx \sum_{i=1}^{n} 2\pi r_i h_i \Delta x,$$

where r_i, h_i and Δx are the radius, height and thickness of the ith shell, respectively.

This is a Riemann Sum. Taking a limit as the thickness of the shells approaches 0 leads to a definite integral.

$$V = \lim_{n \to \infty} \sum_{i=1}^{n} 2\pi r_i h_i \Delta x$$

$$= 2\pi \int_{a}^{b} r(x)h(x)\, dx$$

(a)

(b)

(c)
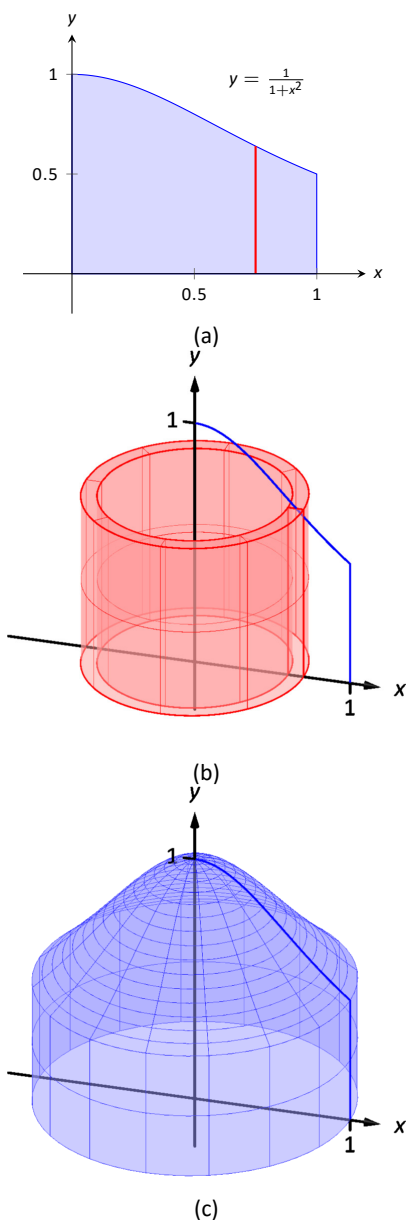
Figure 6.17: Introducing the Shell Method.

Notes:
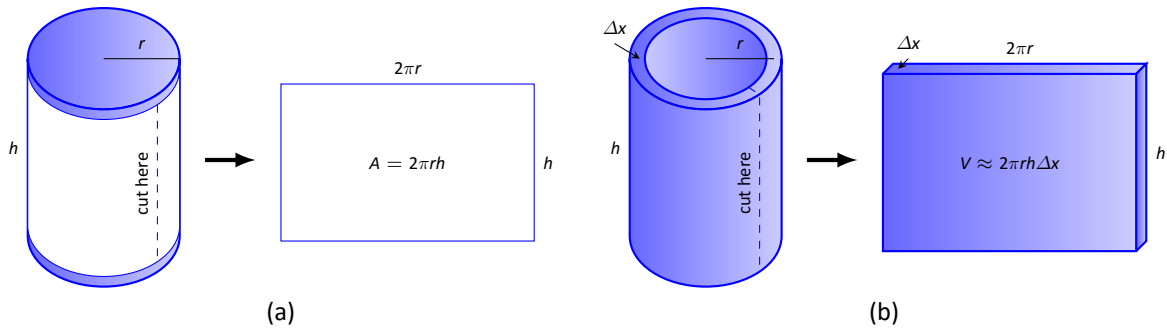
Figure 6.18: Determining the volume of a thin cylindrical shell.

Key Idea 14   The Shell Method

Let a solid be formed by revolving a region R, bounded by $x = a$ and $x = b$, around a vertical axis. Let $r(x)$ represent the distance from the axis of rotation to x (i.e., the radius of a sample shell) and let $h(x)$ represent the height of the solid at x (i.e., the height of the shell). The volume of the solid is

$$V = 2\pi \int_a^b r(x)h(x)\,dx.$$

Special Cases:

1. When the region R is bounded above by $y = f(x)$ and below by $y = g(x)$, then $h(x) = f(x) - g(x)$.

2. When the axis of rotation is the y-axis (i.e., $x = 0$) then $r(x) = x$.

Watch the video:
Volumes of Revolution — Cylindrical Shells at
`https://youtu.be/V6nTsxumjgU`

Let's practice using the Shell Method.

Example 1   Finding volume using the Shell Method
Find the volume of the solid formed by rotating the region bounded by $y = 0$, $y = 1/(1 + x^2)$, $x = 0$ and $x = 1$ about the y-axis.

Notes:

SOLUTION This is the region used to introduce the Shell Method in Figure 6.17, but is sketched again in Figure 6.19 for closer reference. A line is drawn in the region parallel to the axis of rotation representing a shell that will be carved out as the region is rotated about the y-axis.

The distance this line is from the axis of rotation determines $r(x)$; as the distance from x to the y-axis is x, we have $r(x) = x$. The height of this line determines $h(x)$; the top of the line is at $y = 1/(1 + x^2)$, whereas the bottom of the line is at $y = 0$. Thus $h(x) = 1/(1 + x^2) - 0 = 1/(1 + x^2)$. The region is bounded from $x = 0$ to $x = 1$, so the volume is

$$V = 2\pi \int_0^1 \frac{x}{1 + x^2} \, dx.$$

Figure 6.19: Graphing a region in Example 1.

This requires substitution. Let $u = 1 + x^2$, so $du = 2x \, dx$. We also change the bounds: $u(0) = 1$ and $u(1) = 2$. Thus we have:

$$= \pi \int_1^2 \frac{1}{u} \, du$$

$$= \pi \ln u \Big|_1^2$$

$$= \pi \ln 2 \text{ units}^3.$$

Note: in order to find this volume using the Disk Method, two integrals would be needed to account for the regions above and below $y = 1/2$.

With the Shell Method, nothing special needs to be accounted for to compute the volume of a solid that has a hole in the middle, as demonstrated next.

Example 2 **Finding volume using the Shell Method**
Find the volume of the solid formed by rotating the triangular region determined by the points $(0, 1)$, $(1, 1)$ and $(1, 3)$ about the line $x = 3$.

SOLUTION The region is sketched in Figure 6.20(a) along with a line within the region parallel to the axis of rotation. In part (b) of the figure, we see a sample shell, and in part (c) the whole solid is shown.

The height of the sample shell is the distance from $y = 1$ to $y = 2x + 1$, the line that connects the points $(0, 1)$ and $(1, 3)$. Thus $h(x) = 2x + 1 - 1 = 2x$. The radius of the sample shell is the distance from x to $x = 3$; that is, it is $r(x) = 3 - x$.

(a)

(b)

(c)

Figure 6.20: Graphing a region in Example 2.

Notes:

The *x*-bounds of the region are $x = 0$ to $x = 1$, giving

$$V = 2\pi \int_0^1 (3 - x)(2x)\, dx$$

$$= 2\pi \int_0^1 (6x - 2x^2)\, dx$$

$$= 2\pi \left(3x^2 - \frac{2}{3}x^3\right)\Big|_0^1$$

$$= \frac{14}{3}\pi \text{ units}^3.$$

When revolving a region around a horizontal axis, we must consider the radius and height functions in terms of *y*, not *x*.

Example 3   Finding volume using the Shell Method
Find the volume of the solid formed by rotating the region given in Example 2 about the *x*-axis.

SOLUTION   The region is sketched in Figure 6.21(a). In part (b) of the figure the sample shell is drawn, and the solid is sketched in (c). (Note that the triangular region looks "short and wide" here, whereas in the previous example the same region looked "tall and narrow." This is because the bounds on the graphs are different.)

The height of the sample shell is an *x*-distance, between $x = \frac{1}{2}y - \frac{1}{2}$ and $x = 1$. Thus $h(y) = 1 - (\frac{1}{2}y - \frac{1}{2}) = -\frac{1}{2}y + \frac{3}{2}$. The radius is the distance from *y* to the *x*-axis, so $r(y) = y$. The *y* bounds of the region are $y = 1$ and $y = 3$, leading to the integral

$$V = 2\pi \int_1^3 \left[y\left(-\frac{1}{2}y + \frac{3}{2}\right)\right] dy$$

$$= 2\pi \int_1^3 \left[-\frac{1}{2}y^2 + \frac{3}{2}y\right] dy$$

$$= 2\pi \left[-\frac{1}{6}y^3 + \frac{3}{4}y^2\right]\Big|_1^3$$

$$= 2\pi \left[\frac{9}{4} - \frac{7}{12}\right]$$

$$= \frac{10}{3}\pi \text{ units}^3.$$

(a)

(b)

(c)

Figure 6.21: Graphing a region in Example 3.

Notes:

The following example shows how there are times when it does not matter which method you choose to evaluate the volume of a solid. In Example 6.2.7 we found the volume of the solid formed by rotating the region bounded by $y = \sqrt{x}$ and $y = x$ about $y = 2$. We will now demonstrate how to find the volume with the shell method. Note that your answer should be the same whichever method you choose.

Example 4   **Using the shell method instead of the washer method**
Find the volume of the solid formed by rotating the region bounded by $y = \sqrt{x}$ and $y = x$ about $y = 2$ using the Shell Method.

SOLUTION   Since our shells are parallel to the axis of rotation, we must consider the radius and height functions in terms of y. The radius of a sample shell will be $r(y) = 2 - y$ and the height of a sample shell will be $h(y) = y = y^2$. The y bounds for the region will be $y = 0$ to $y = 1$ resulting in the integral

$$V = 2\pi \int_0^1 (2 - y)(y - y^2)\, dy$$

$$= 2\pi \int_0^1 y^3 - 3y^2 + 2y\, dy$$

$$= \frac{\pi}{2} \text{units}^3.$$

At the beginning of this section it was stated that "it is good to have options." The next example finds the volume of a solid rather easily with the Shell Method, but using the Washer Method would be quite a chore.

Example 5   **Finding volume using the Shell Method**
Find the volume of the solid formed by rotating the region bounded by $y = 3x - x^2$ and $y = x$ about the y-axis.

SOLUTION   The region, a sample shell, and the resulting solid are shown in Figure 6.22. The radius of a sample shell is $r(x) = x$; the height of a sample shell is $h(x) = (3x - x^2) - x = 2x - x^2$. The x bounds on the region are $x = 0$

Notes:

(a)

(b)

(c)
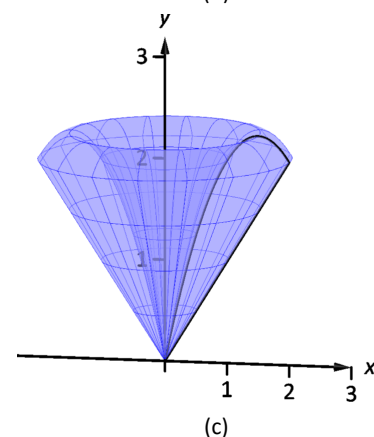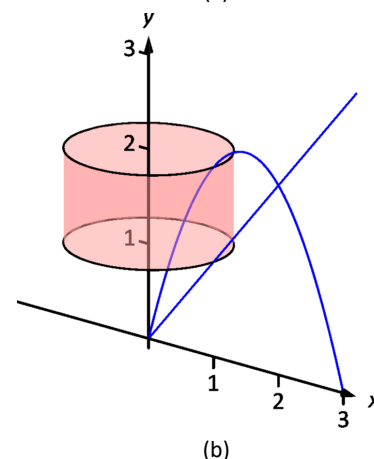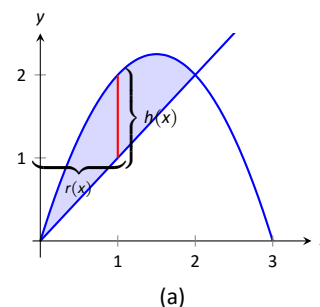
Figure 6.22: Graphing a region in Example 5.

to $x = 2$ leading to the integral

$$V = 2\pi \int_0^2 x(2x - x^2)\, dx$$

$$= 2\pi \int_0^2 2x^2 - x^3\, dx$$

$$= 2\pi \left[\frac{2}{3}x^3 - \frac{1}{4}x^4 \right]\Big|_0^2$$

$$= \frac{4\pi}{3}$$

Note that in order to use the Washer Method, we would need to solve $y = 3x - x^2$ for x, requiring us to complete the square. We must evaluate two integrals as we have two different sample slices. The volume can be computed as

$$V = 2\pi \int_0^2 \left(y - \left(\frac{3}{2} - \sqrt{\frac{9}{4} - y} \right) \right)^2 dy$$

$$+ 2\pi \int_2^{9/4} \left(\left(\frac{3}{2} + \sqrt{\frac{9}{4} - y} \right) - \left(\frac{3}{2} - \sqrt{\frac{9}{4} - y} \right) \right)^2 dy$$

$$= 2\pi \int_0^2 \left(y - \frac{3}{2} + \sqrt{\frac{9}{4} - y} \right)^2 dy + 2\pi \int_2^{9/4} \left(2\sqrt{\frac{9}{4} - y} \right)^2 dy$$

While this integral is not impossible to solve, using the Shell Method gave us a significantly easier way to compute the volume.

As in the previous section, the real goal of this section is not to be able to compute volumes of certain solids. Rather, it is to be able to solve a problem by first approximating, then using limits to refine the approximation to give the exact value. In this section, we approximate the volume of a solid by cutting it into thin cylindrical shells. By summing up the volumes of each shell, we get an approximation of the volume. By taking a limit as the number of equally spaced shells goes to infinity, our summation can be evaluated as a definite integral, giving the exact value.

Notes:

Exercises 6.3

Terms and Concepts

1. T/F: A solid of revolution is formed by revolving a shape around an axis.

2. T/F: The Shell Method can only be used when the Washer Method fails.

3. T/F: The Shell Method works by integrating cross–sectional areas of a solid.

4. T/F: When finding the volume of a solid of revolution that was revolved around a vertical axis, the Shell Method integrates with respect to x.

Problems

In Exercises 5–8, a region of the Cartesian plane is shaded. Use the Shell Method to find the volume of the solid of revolution formed by revolving the region about the y-axis.

5.

$y = 3 - x^2$

6.

$y = 5x$

7.

$y = x^3$

8.

$y = \sqrt{x}$

$y = x$
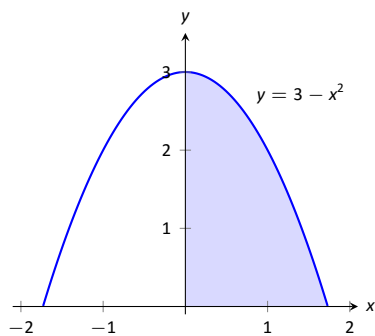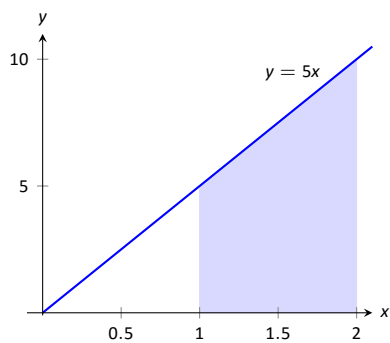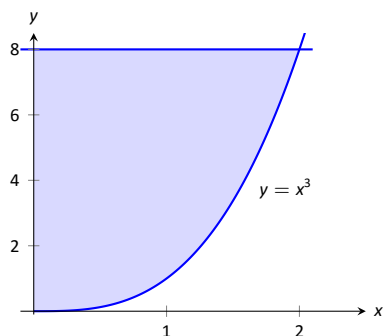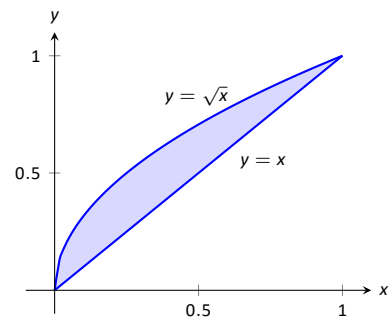
In Exercises 9–12, a region of the Cartesian plane is shaded. Use the Shell Method to find the volume of the solid of revolution formed by revolving the region about the x-axis.

9.

$y = 3 - x^2$

10.

$y = 5x$

11.

$y = x^3$

12.
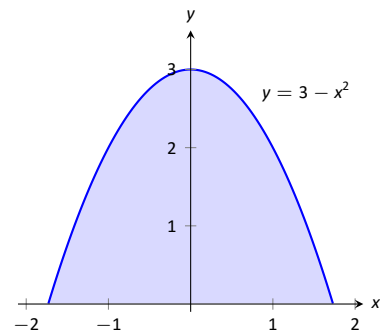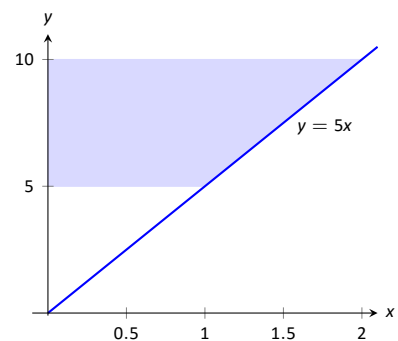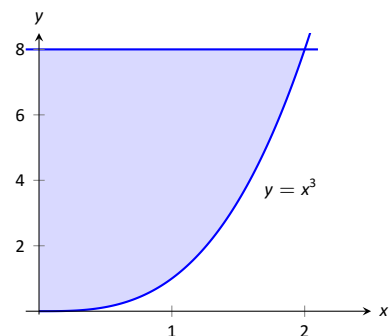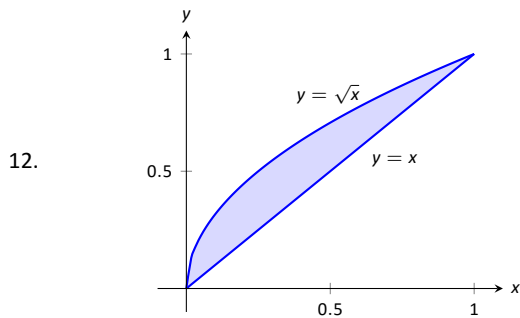
In Exercises 13–17, a region of the Cartesian plane is described. Use the Shell Method to find the volume of the solid of revolution formed by rotating the region about each of the given axes.

13. Region bounded by: $y = \sqrt{x}$, $y = 0$ and $x = 1$.
    Rotate about:

    (a) the y-axis   (c) the x-axis

    (b) $x = 1$ (d) $y = 1$

14. Region bounded by: $y = 4 - x^2$ and $y = 0$.
    Rotate about:

    (a) $x = 2$ (c) the x-axis

    (b) $x = -2$ (d) $y = 4$

15. The triangle with vertices $(1, 1)$, $(1, 2)$ and $(2, 1)$.
    Rotate about:

    (a) the y-axis   (c) the x-axis

    (b) $x = 1$ (d) $y = 2$

16. Region bounded by $y = x^2 - 2x + 2$ and $y = 2x - 1$.
    Rotate about:

    (a) the y-axis   (c) $x = -1$

    (b) $x = 1$

17. Region bounded by $y = 2x$, $y = x$ and $x = 2$.
    Rotate about:

    (a) the y-axis   (c) the x-axis

    (b) $x = 2$ (d) $y = 4$

In Exercises 18–20, use your choice of the Washer or Shell Method to find the indicated volume.

18. Region bounded by $y = x^4$, $y = 0$, and $x = 1$.
    Rotate about:
    (a) the y-axis   (b) the x-axis

19. Region bounded by $y = x^3 + 1$, $x = 0$, and $y = 2$.
    Rotate about:
    (a) the y-axis   (b) $y = 2$

20. Region bounded by $y = 4x^2$ and $4x + y = 8$.
    Rotate about

    (a) the x-axis   (c) $y = 16$

    (b) $x = 1$

6.4 Work

Work is the scientific term used to describe the action of a force which moves an object. When a constant force F is applied to move an object a distance d, the amount of work performed is $W = F \cdot d$.

The SI unit of force is the Newton, (kg·m/s^2), and the SI unit of distance is a meter (m). The fundamental unit of work is one Newton–meter, or a Joule (J). That is, applying a force of one Newton for one meter performs one Joule of work. In Imperial units (as used in the United States), force is measured in pounds (lb) and distance is measured in feet (ft), hence work is measured in ft–lb.

When force is constant, the measurement of work is straightforward. For instance, lifting a 200 lb object 5 ft performs $200 \cdot 5 = 1000$ ft–lb of work.

What if the force applied is variable? For instance, imagine a climber pulling a 200 ft rope up a vertical face. The rope becomes lighter as more is pulled in, requiring less force and hence the climber performs less work.

In general, let $F(x)$ be a force function on an interval $[a, b]$. We want to measure the amount of work done applying the force F from $x = a$ to $x = b$. We can approximate the amount of work being done by partitioning $[a, b]$ into subintervals $a = x_0 < x_1 < \cdots < x_n = b$ and assuming that F is constant on each subinterval. Let c_i be a value in the ith subinterval $[x_i, x_{i+1}]$. Then the work done on this interval is approximately $W_i \approx F(c_i) \cdot (x_{i+1} - x_i) = F(c_i)\Delta x_i$, a constant force $\times$ the distance over which it is applied. The total work is

$$W = \sum_{i=1}^{n} W_i \approx \sum_{i=1}^{n} F(c_i)\Delta x_i.$$

This, of course, is a Riemann sum. Taking a limit as the subinterval lengths go to zero give an exact value of work which can be evaluated through a definite integral.

Note: *Mass* and *weight* are closely related, yet different, concepts. The mass m of an object is a quantitative measure of that object's resistance to acceleration. The weight w of an object is a measurement of the force applied to the object by the acceleration of gravity g.

Since the two measurements are proportional, $w = m \cdot g$, they are often used interchangeably in everyday conversation. When computing work, one must be careful to note which is being referred to. When mass is given, it must be multiplied by the acceleration of gravity to reference the related force.

Key Idea 15 Work

Let $F(x)$ be a continuous function on $[a, b]$ describing the amount of force being applied to an object in the direction of travel from distance $x = a$ to distance $x = b$. The total work W done on $[a, b]$ is

$$W = \int_a^b F(x)\, dx.$$

Notes:

Watch the video:
Finding Work using Calculus — The Cable/Rope Problem at
`https://youtu.be/2pbInn9PkHQ`

Example 1   Computing work performed: applying variable force
A 60 m climbing rope is hanging over the side of a tall cliff. How much work is performed in pulling the rope up to the top, where the rope has a mass of 66 g/m?

Solution   We need to create a force function $F(x)$ on the interval $[0, 60]$. To do so, we must first decide what x is measuring: it is the length of the rope still hanging or is it the amount of rope pulled in? As long as we are consistent, either approach is fine. We adopt for this example the convention that x is the amount of rope pulled in. This seems to match intuition better; pulling up the first 10 meters of rope involves $x = 0$ to $x = 10$ instead of $x = 60$ to $x = 50$.

As x is the amount of rope pulled in, the amount of rope still hanging is $60-x$. This length of rope has a mass of 66 g/m, or 0.066 kg/m. The the mass of the rope still hanging is $0.066(60 - x)$ kg; multiplying this mass by the acceleration of gravity, 9.8 m/s^2, gives our variable force function

$$F(x) = (9.8)(0.066)(60 - x) = 0.6468(60 - x).$$

Thus the total work performed in pulling up the rope is

$$W = \int_0^{60} 0.6468(60 - x)\, dx = 1,164.24 \text{ J}.$$

By comparison, consider the work done in lifting the entire rope 60 meters. The rope weights $60 \times 0.066 \times 9.8 = 38.808$ N, so the work applying this force for 60 meters is $60 \times 38.808 = 2,328.48$ J. This is exactly twice the work calculated before (and we leave it to the reader to understand why.)

Example 2   Computing work performed: applying variable force
Consider again pulling a 60 m rope up a cliff face, where the rope has a mass of 66 g/m. At what point is exactly half the work performed?

Notes:

Solution   From Example 1 we know the total work performed is 1, 164.24 J. We want to find a height h such that the work in pulling the rope from a height of $x = 0$ to a height of $x = h$ is 582.12, half the total work. Thus we want to solve for h in the equation

$$\int_0^h 0.6468(60 - x) \, dx = 582.12.$$

We see that

$$\int_0^h 0.6468(60 - x) \, dx = 582.12$$

$$\left(38.808x - 0.3234x^2\right)\Big|_0^h = 582.12$$

$$38.808h - 0.3234h^2 = 582.12$$

$$-0.3234h^2 + 38.808h - 582.12 = 0 \qquad \text{(Apply the Quadratic Formula)}$$

$$h \approx 17.57 \text{ and } 102.43$$

As the rope is only 60m long, the only sensible answer is $h = 17.57$. Thus about half the work is done pulling up the first 17.57m the other half of the work is done pulling up the remaining 42.43m.

Note: In Example 2, we find that half of the work performed in pulling up a 60 m rope is done in the last 42.43 m. Why is it not coincidental that $60/\sqrt{2} = 42.43$?

Example 3   **Computing work performed: applying variable force**

A box of 100 lb of sand is being pulled up at a uniform rate a distance of 50 ft over 1 minute. The sand is leaking from the box at a rate of 1 lb/s. The box itself weighs 5 lb and is pulled by a rope weighing .2 lb/ft.

1. How much work is done lifting just the rope?

2. How much work is done lifting just the box and sand?

3. What is the total amount of work performed?

Solution

1. We start by forming the force function $F_r(x)$ for the rope (where the subscript denotes we are considering the rope). As in the previous example, let x denote the amount of rope, in feet, pulled in. (This is the same as saying x denotes the height of the box.) The weight of the rope with x feet pulled in is $F_r(x) = 0.2(50 - x) = 10 - 0.2x$. (Note that we do not have to include the acceleration of gravity here, for the *weight* of the rope per foot is given, not its *mass* per meter as before.) The work performed lifting the rope is

$$W_r = \int_0^{50} (10 - 0.2x) \, dx = 250 \text{ ft–lb.}$$

Notes:

2. The sand is leaving the box at a rate of 1 lb/s. As the vertical trip is to take one minute, we know that 60 lb will have left when the box reaches its final height of 50 ft. Again letting x represent the height of the box, we have two points on the line that describes the weight of the sand: when $x = 0$, the sand weight is 100 lb, producing the point $(0, 100)$; when $x = 50$, the sand in the box weighs 40 lb, producing the point $(50, 40)$. The slope of this line is $\frac{100-40}{0-50} = -1.2$, giving the equation of the weight of the sand at height x as $w(x) = -1.2x + 100$. The box itself weighs a constant 5 lb, so the total force function is $F_b(x) = -1.2x + 105$. Integrating from $x = 0$ to $x = 50$ gives the work performed in lifting box and sand:

$$W_b = \int_0^{50} (-1.2x + 105)\, dx = 3750 \text{ ft–lb.}$$

3. The total work is the sum of W_r and W_b: $250 + 3750 = 4000$ ft–lb. We can also arrive at this via integration:

$$\begin{aligned} W &= \int_0^{50} (F_r(x) + F_b(x))\, dx \\ &= \int_0^{50} (10 - 0.2x - 1.2x + 105)\, dx \\ &= \int_0^{50} (-1.4x + 115)\, dx \\ &= 4000 \text{ ft–lb.} \end{aligned}$$

Hooke's Law and Springs

Hooke's Law states that the force required to compress or stretch a spring x units from its natural length is proportional to x; that is, this force is $F(x) = kx$ for some constant k. For example, if a force of 1 N stretches a given spring 2 cm, then a force of 5 N will stretch the spring 10 cm. Converting the distances to meters, we have that stretching this spring 0.02 m requires a force of $F(0.02) = k(0.02) = 1$ N, hence $k = 1/0.02 = 50$ N/m.

Example 4   **Computing work performed: stretching a spring**
A force of 20 lb stretches a spring from a natural length of 7 inches to a length of 12 inches. How much work was performed in stretching the spring to this length?

Notes:

Solution In many ways, we are not at all concerned with the actual length of the spring, only with the amount of its change. Hence, we do not care that 20 lb of force stretches the spring to a length of 12 inches, but rather that a force of 20 lb stretches the spring by 5 in. This is illustrated in Figure 6.23; we only measure the change in the spring's length, not the overall length of the spring.
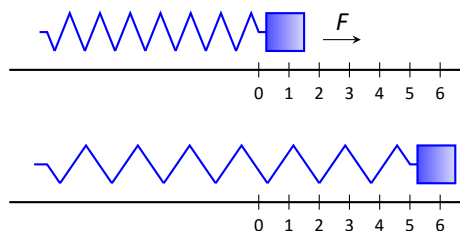
Figure 6.23: Illustrating the important aspects of stretching a spring in computing work in Example 4.

Converting the units of length to feet, we have

$$F(5/12) = (5/12)k = 20 \text{ lb}.$$

Thus $k = 48$ lb/ft and $F(x) = 48x$.

We compute the total work performed by integrating $F(x)$ from $x = 0$ to $x = 5/12$:

$$W = \int_0^{5/12} 48x \, dx$$
$$= 24x^2 \Big|_0^{5/12}$$
$$= 25/6 \text{ ft–lb}.$$

Pumping Fluids

Another useful example of the application of integration to compute work comes in the pumping of fluids, often illustrated in the context of emptying a storage tank by pumping the fluid out the top. This situation is different than our previous examples for the forces involved are constant. After all, the force required to move one cubic foot of water (about 62.4 lb) is the same regardless of its location in the tank. What is variable is the distance that cubic foot of water has

Fluid	lb/ft^3	kg/m^3
Gasoline	45.93	737.22
Methanol	49.3	791.3
Fuel Oil	55.46	890.13
Water	62.4	1000
Milk, whole	63.6	1020
Milk, nonfat	65.4	1050
Concrete	150	2400
Iodine	307	4927
Mercury	844	13546

Figure 6.24: Weight and Mass densities

Notes:

to travel; water closer to the top travels less distance than water at the bottom, producing less work.

We demonstrate how to compute the total work done in pumping a fluid out of the top of a tank in the next two examples.

Example 5   Computing work performed: pumping fluids
A cylindrical storage tank with a radius of 10 ft and a height of 30 ft is filled with water, which weighs approximately 62.4 lb/ft^3. Compute the amount of work performed by pumping the water up to a point 5 feet above the top of the tank.

SOLUTION   We will refer often to Figure 6.25 which illustrates the salient aspects of this problem.

We start as we often do: we partition an interval into subintervals. We orient our tank vertically since this makes intuitive sense with the base of the tank at $y = 0$. Hence the top of the water is at $y = 30$, meaning we are interested in subdividing the y-interval $[0, 30]$ into n subintervals as

$$0 = y_0 < y_1 < \cdots < y_n = 30.$$

Consider the work W_i of pumping only the water residing in the i th subinterval, illustrated in Figure 6.25. The force required to move this water is equal to its weight which we calculate as volume × density. The volume of water in this subinterval is $V_i = 10^2 \pi \Delta y_i$; its density is 62.4 lb/ft^3. Thus the required force is $6240\pi \Delta y_i$ lb.

We approximate the distance the force is applied by using any y-value contained in the i th subinterval; for simplicity, we arbitrarily use y_i for now (it will not matter later on). The water will be pumped to a point 5 feet above the top of the tank, that is, to the height of $y = 35$ ft. Thus the distance the water at height y_i travels is $35 - y_i$ ft.

In all, the approximate work W_i peformed in moving the water in the i th subinterval to a point 5 feet above the tank is

$$W_i \approx 6240\pi \Delta y_i (35 - y_i).$$

To approximate the total work performed in pumping out all the water from the tank, we sum all the work W_i performed in pumping the water from each of the n subintervals of $[0, 30]$:

$$W \approx \sum_{i=1}^{n} W_i = \sum_{i=1}^{n} 6240\pi \Delta y_i (35 - y_i).$$
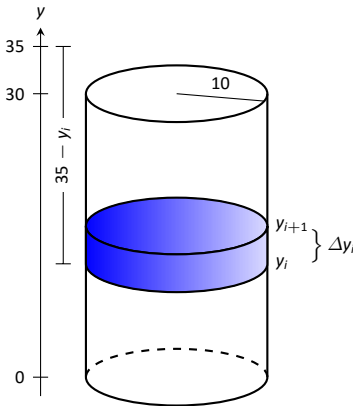
Figure 6.25: Illustrating a water tank in order to compute the work required to empty it in Example 5.

Notes:

This is a Riemann sum. Taking the limit as the subinterval length goes to 0 gives

$$W = \int_0^{30} 6240\pi(35 - y)\, dy$$
$$= 6240\pi\left(35y - \frac{y^2}{2}\right)\Big|_0^{30}$$
$$= 11,762,123 \text{ ft–lb}$$
$$\approx 1.176 \times 10^7 \text{ ft–lb}.$$

We can "streamline" the above process a bit as we may now recognize what the important features of the problem are. Figure 6.26 shows the tank from Example 5 without the ith subinterval identified. Instead, we just draw a sample slice. This helps establish the height a small amount of water must travel along with the force required to move it (where the force is volume $\times$ density).

We demonstrate the concepts again in the next examples.

Example 6   **Computing work performed: pumping fluids**
A conical water tank has its top at ground level and its base 10 feet below ground. The radius of the cone at ground level is 2 ft. It is filled with water weighing 62.4 lb/ft^3 and is to be emptied by pumping the water to a spigot 3 feet above ground level. Find the total amount of work performed in emptying the tank.

SOLUTION   The conical tank is sketched in Figure 6.27. We can orient the tank in a variety of ways; we could let $y = 0$ represent the base of the tank and $y = 10$ represent the top of the tank, but we choose to keep the convention of the wording given in the problem and let $y = 0$ represent ground level and hence $y = -10$ represents the bottom of the tank. The actual "height" of the water does not matter; rather, we are concerned with the distance the water travels.

The figure also sketches a cross–sectional circle. The radius of this circle is variable, depending on y. When $y = -10$, the circle has radius 0; when $y = 0$, the circle has radius 2. These two points, $(-10, 0)$ and $(0, 2)$, allow us to find the equation of the line that gives the radius of the cross–sectional circle, which is $r(y) = y/5 + 2$. Hence the volume of water at this height is $V(y) = \pi(y/5 + 2)^2 dy$, where dy represents a very small height of the slice. The force required to move the water at height y is $F(y) = 62.4 \times V(y)$.

The distance the water at height y travels is given by $h(y) = 3 - y$. Thus the
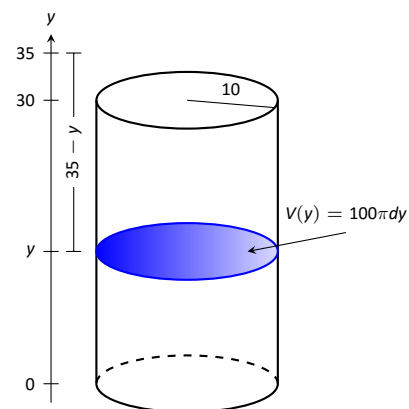
Figure 6.26: A simplified illustration for computing work.
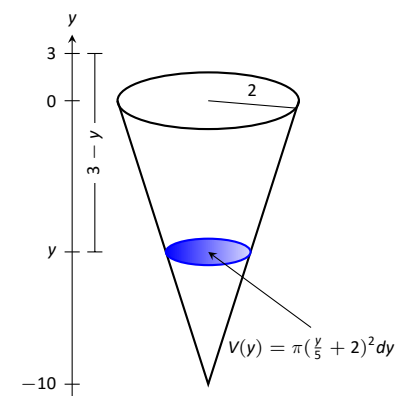
Figure 6.27: A graph of the conical water tank in Example 6.

Notes:

total work done in pumping the water from the tank is

$$W = \int_{-10}^{0} 62.4\pi(y/5+2)^2(3-y)\,dy$$

$$= 62.4\pi \int_{-10}^{0} \left(-\frac{1}{25}y^3 - \frac{17}{25}y^2 - \frac{8}{5}y + 12\right)\,dy$$

$$= 62.2\pi \cdot \frac{220}{3}\ \text{ft–lb.}$$

Example 7   **Computing work performed: pumping fluids**
A rectangular swimming pool is 20 ft wide and has a 3 ft "shallow end" and a 6 ft "deep end." It is to have its water pumped out to a point 2 ft above the current top of the water. The cross–sectional dimensions of the water in the pool are given in Figure 6.28(a). (Note that the "20 ft wide" is into the picture; the pool is 25 ft long.) Compute the amount of work performed in draining the pool.

SOLUTION   For the purposes of this problem we choose to set $y = 0$ to represent the bottom of the pool, meaning the top of the water is at $y = 6$. Figure 6.28(b) shows the pool oriented with this y-axis, along with 2 sample slices as the pool must be split into two different regions.

The top region lies in the y-interval of $[3, 6]$, where the length of the sample slice is 25 ft as shown. As the pool is 20 ft wide, this sample slice of water has a volume of $V(y) = 20 \cdot 25 \cdot dy$. The water is to be pumped to a height of $y = 8$, so the height function is $h(y) = 8 - y$. The work done in pumping this top region of water is

$$W_t = 62.4 \int_{3}^{6} 500(8-y)\,dy = 327,600\ \text{ft–lb.}$$

The bottom region lies in the y-interval of $[0, 3]$; we need to compute the length of the sample slice in this interval.

One end of the sample slice is at $x = 0$ and the other is along the line segment joining the points $(10, 0)$ and $(15, 3)$. The equation of this line is $y = 3(x-10)/5$; as we will be integrating with respect to y, we rewrite this equation as $x = 5y/3 + 10$. So the length of the sample slice is a difference of x-values: $x = 0$ and $x = 5y/3 + 10$, giving a length of $x = 5y/3 + 10$.

Again, as the pool is 20 ft wide, this slice of water has a volume of $V(y) = 20 \cdot (5y/3 + 10) \cdot dy$; the height function is the same as before at $h(y) = 8 - y$. The work performed in emptying this part of the pool is

$$W_b = 62.4 \int_{0}^{3} 20(5y/3 + 10)(8-y)\,dy = 299,520\ \text{ft–lb.}$$

25 ft

3 ft.

6 ft.

10 ft.

10 ft.

(a)

(15, 3)

(10, 0)

(b)
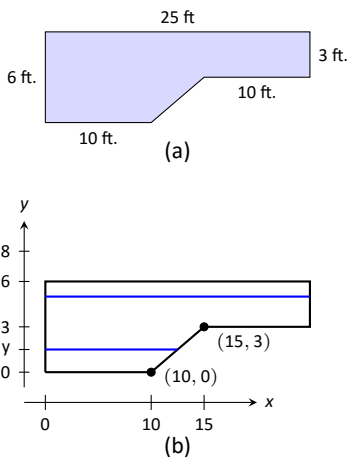
Figure 6.28: The cross–section of a swimming pool filled with water in Example 7 and two sample slices.

Notes:

The total work in emptying the pool is

$$W = W_b + W_t = 327,600 + 299,520 = 627,120 \text{ ft–lb.}$$

Notice how the emptying of the bottom of the pool performs almost as much work as emptying the top. The top portion travels a shorter distance but has more water. In the end, this extra water produces more work.

The next section introduces one final application of the definite integral, the calculation of fluid force on a plate.

Notes:

Exercises 6.4

Terms and Concepts

1. What are the typical units of work?

2. If a man has a mass of 80 kg on Earth, will his mass on the moon be bigger, smaller, or the same?

3. If a woman weighs 130 lb on Earth, will her weight on the moon be bigger, smaller, or the same?

Problems

4. A 100 ft rope, weighing 0.1 lb/ft, hangs over the edge of a tall building.

    (a) How much work is done pulling the entire rope to the top of the building?

    (b) How much rope is pulled in when half of the total work is done?

5. A 50 m rope, with a mass of 0.2 kg/m, hangs over the edge of a tall building.

    (a) How much work is done pulling the entire rope to the top of the building?

    (b) How much work is done pulling in the first 20 m?

6. A rope of length ℓ ft hangs over the edge of tall cliff. (Assume the cliff is taller than the length of the rope.) The rope weighs d lb/ft.

    (a) How much work is done pulling the entire rope to the top of the cliff?

    (b) What percentage of the total work is done pulling in the first half of the rope?

    (c) How much rope is pulled in when half of the total work is done?

7. A 20 m rope with a mass of 0.5 kg/m hangs over the edge of a 10 m building. How much work is done pulling the rope to the top?

8. A crane lifts a 2,000 lb load vertically 30 ft with a 1" cable weighing 1.68 lb/ft.

    (a) How much work is done lifting the cable alone?

    (b) How much work is done lifting the load alone?

    (c) Could one conclude that the work done lifting the cable is negligible compared to the work done lifting the load?

9. A 100 lb bag of sand is lifted uniformly 120 ft in one minute. Sand leaks from the bag at a rate of 1/4 lb/s. What is the total work done in lifting the bag?

10. A box weighing 2 lb lifts 10 lb of sand vertically 50 ft. A crack in the box allows the sand to leak out such that 9 lb of sand is in the box at the end of the trip. Assume the sand leaked out at a uniform rate. What is the total work done in lifting the box and sand?

11. A force of 1000 lb compresses a spring 3 in. How much work is performed in compressing the spring?

12. A force of 2 N stretches a spring 5 cm. How much work is performed in stretching the spring?

13. A force of 50 lb compresses a spring from a natural length of 18 in to 12 in. How much work is performed in compressing the spring?

14. A force of 20 lb stretches a spring from a natural length of 6 in to 8 in. How much work is performed in stretching the spring?

15. A force of 7 N stretches a spring from a natural length of 11 cm to 21 cm. How much work is performed in stretching the spring from a length of 16 cm to 21 cm?

16. A force of f N stretches a spring d m from its natural length. How much work is performed in stretching the spring?

17. A 20 lb weight is attached to a spring. The weight rests on the spring, compressing the spring from a natural length of 1 ft to 6 in.
How much work is done in lifting the box 1.5 ft (i.e, the spring will be stretched 1 ft beyond its natural length)?

18. A 20 lb weight is attached to a spring. The weight rests on the spring, compressing the spring from a natural length of 1 ft to 6 in.
How much work is done in lifting the box 6 in (i.e, bringing the spring back to its natural length)?

19. A 5 m tall cylindrical tank with radius of 2 m is filled with 3 m of gasoline, with a mass of 737.22 kg/m^3. Compute the total work performed in pumping all the gasoline to the top of the tank.

20. A 6 ft cylindrical tank with a radius of 3 ft is filled with water, which weighs 62.4 lb/ft^3. The water is to be pumped to a point 2 ft above the top of the tank.

    (a) How much work is performed in pumping all the water from the tank?

    (b) How much work is performed in pumping 3 ft of water from the tank?

    (c) At what point is 1/2 of the total work done?

21. A gasoline tanker is filled with gasoline which weighs 45.93 lb/ft^3. The dispensing valve at the base is jammed shut, forcing the operator to empty the tank via pumping the gas to a point 1 ft above the top of the tank. Assume the tank is a perfect cylinder, 20 ft long with a diameter of 7.5 ft. How much work is performed in pumping all the gasoline from the tank?

22. A fuel oil storage tank is 10 ft deep with trapezoidal sides, 5 ft at the top and 2 ft at the bottom, and is 15 ft wide (see diagram below). Given that fuel oil weighs 55.46 lb/ft^3, find the work performed in pumping all the oil from the tank to a point 3 ft above the top of the tank.
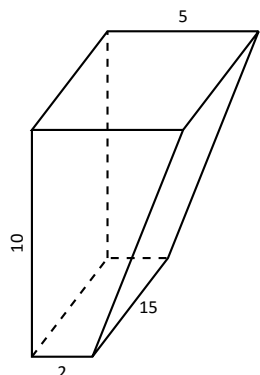
23. A conical water tank is 5 m deep with a top radius of 3 m. (This is similar to Example 6.) The tank is filled with pure water, with a mass of 1000 kg/m^3.

    (a) Find the work performed in pumping all the water to the top of the tank.

    (b) Find the work performed in pumping the top 2.5 m of water to the top of the tank.

    (c) Find the work performed in pumping the top half of the water, by volume, to the top of the tank.

24. A water tank has the shape of a truncated cone, with dimensions given below, and is filled with water which weighs 62.4 lb/ft^3. Find the work performed in pumping all water to a point 1 ft above the top of the tank.
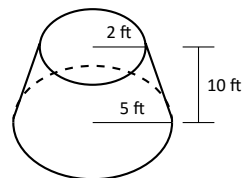
25. A water tank has the shape of an inverted pyramid, with dimensions given below, and is filled with water with a mass of 1000 kg/m^3. Find the work performed in pumping all water to a point 5 m above the top of the tank.
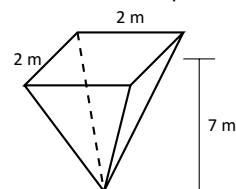
26. A water tank has the shape of an truncated, inverted pyramid, with dimensions given below, and is filled with water with a mass of 1000 kg/m^3. Find the work performed in pumping all water to a point 1 m above the top of the tank.
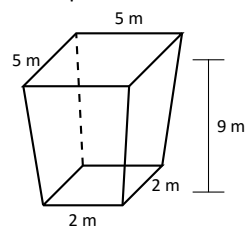
6.5   Fluid Forces

In the unfortunate situation of a car driving into a body of water, the conventional wisdom is that the water pressure on the doors will quickly be so great that they will be effectively unopenable. (Survival techniques suggest immediately opening the door, rolling down or breaking the window, or waiting until the water fills up the interior at which point the pressure is equalized and the door will open. See Mythbusters episode #72 to watch Adam Savage test these options.)

How can this be true? How much force does it take to open the door of a submerged car? In this section we will find the answer to this question by examining the forces exerted by fluids.

We start with **pressure**, which is related to **force** by the following equations:

$$\text{Pressure} = \frac{\text{Force}}{\text{Area}} \quad \Leftrightarrow \quad \text{Force} = \text{Pressure} \times \text{Area}.$$

In the context of fluids, we have the following definition.

Definition 26   Fluid Pressure

Let w be the weight–density of a fluid. The **pressure** p exerted on an object at depth d in the fluid is $p = w \cdot d$.

We use this definition to find the **force** exerted on a horizontal sheet by considering the sheet's area.

Watch the video:
Fluid Force on a Vertical Plane at
`http://www.ltcconline.net/greenl/`
`courses/106/videos/fluidSemiCircle/`
`fluidSemiCircle1.htm`

Example 1   Computing fluid force

1. A cylindrical storage tank has a radius of 2 ft and holds 10 ft of a fluid with a weight–density of 50 lb/ft^3. (See Figure 6.29.) What is the force exerted on the base of the cylinder by the fluid?
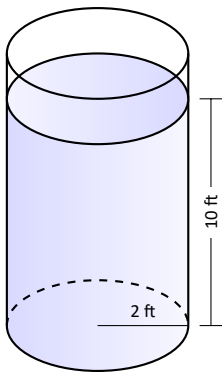
Figure 6.29: A cylindrical tank in Example 1.

Notes:

2. A rectangular tank whose base is a 5 ft square has a circular hatch at the bottom with a radius of 2 ft. The tank holds 10 ft of a fluid with a weight–density of 50 lb/ft^3. (See Figure 6.30.) What is the force exerted on the hatch by the fluid?

SOLUTION

1. Using Definition 26, we calculate that the pressure exerted on the cylinder's base is $w \cdot d = 50$ lb/ft$^3 \times 10$ ft $= 500$ lb/ft^2. The area of the base is $\pi \cdot 2^2 = 4\pi$ ft^2. So the force exerted by the fluid is

$$F = 500 \times 4\pi = 6283 \text{ lb}.$$

Note that we effectively just computed the *weight* of the fluid in the tank.

2. The dimensions of the tank in this problem are irrelevant. All we are concerned with are the dimensions of the hatch and the depth of the fluid. Since the dimensions of the hatch are the same as the base of the tank in the previous part of this example, as is the depth, we see that the fluid force is the same. That is, $F = 6283$ lb.

A key concept to understand here is that we are effectively measuring the weight of a 10 ft column of water above the hatch. The size of the tank holding the fluid does not matter.
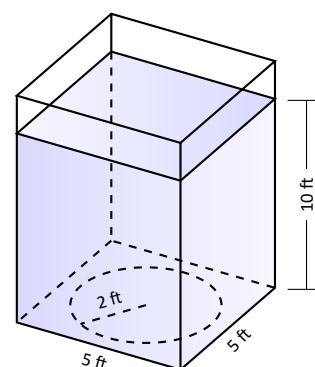
Figure 6.30: A rectangular tank in Example 1.

The previous example demonstrates that computing the force exerted on a horizontally oriented plate is relatively easy to compute. What about a vertically oriented plate? For instance, suppose we have a circular porthole located on the side of a submarine. How do we compute the fluid force exerted on it?

Pascal's Principle states that the pressure exerted by a fluid at a depth is equal in all directions. Thus the pressure on any portion of a plate that is 1 ft below the surface of water is the same no matter how the plate is oriented. (Thus a hollow cube submerged at a great depth will not simply be "crushed" from above, but the sides will also crumple in. The fluid will exert force on *all* sides of the cube.)

So consider a vertically oriented plate as shown in Figure 6.31 submerged in a fluid with weight–density w. What is the total fluid force exerted on this plate? We find this force by first approximating the force on small horizontal strips.

Let the top of the plate be at depth b and let the bottom be at depth a. (For now we assume that surface of the fluid is at depth 0, so if the bottom of the plate is 3 ft under the surface, we have $a = -3$. We will come back to this later.) We partition the interval $[a, b]$ into n subintervals

$$a = y_0 < y_1 < \ldots < y_n = b,$$
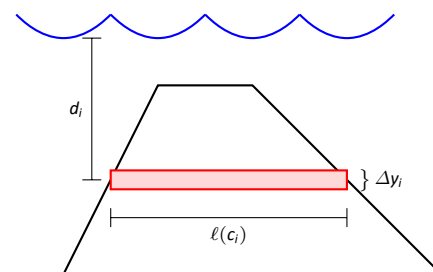
Figure 6.31: A thin, vertically oriented plate submerged in a fluid with weight–density w.

Notes:

with the ith subinterval having length Δy_i. The force F_i exerted on the plate in the ith subinterval is $F_i = $ Pressure $\times$ Area.

The pressure is depth $\times w$. We approximate the depth of this thin strip by choosing any value d_i in $[y_i, y_{i+1}]$; the depth is approximately $-d_i$. (Our convention has d_i being a negative number, so $-d_i$ is positive.) For convenience, we let d_i be an endpoint of the subinterval; we let $d_i = y_i$.

The area of the thin strip is approximately length $\times$ width. The width is Δy_i. The length is a function of some y-value c_i in the ith subinterval. We state the length is $\ell(c_i)$. Thus

$$F_i = \text{Pressure} \times \text{Area}$$
$$= -y_i \cdot w \times \ell(c_i) \cdot \Delta y_i.$$

To approximate the total force, we add up the approximate forces on each of the n thin strips:

$$F = \sum_{i=1}^{n} F_i \approx \sum_{i=1}^{n} -w \cdot y_i \cdot \ell(c_i) \cdot \Delta y_i.$$

This is, of course, another Riemann Sum. We can find the exact force by taking a limit as the subinterval lengths go to 0; we evaluate this limit with a definite integral.

Key Idea 16   Fluid Force on a Vertically Oriented Plate

Let a vertically oriented plate be submerged in a fluid with weight–density w where the top of the plate is at $y = b$ and the bottom is at $y = a$. Let $\ell(y)$ be the length of the plate at y.

1. If $y = 0$ corresponds to the surface of the fluid, then the force exerted on the plate by the fluid is

$$F = \int_a^b w \cdot (-y) \cdot \ell(y)\, dy.$$

2. In general, let $d(y)$ represent the distance between the surface of the fluid and the plate at y. Then the force exerted on the plate by the fluid is

$$F = \int_a^b w \cdot d(y) \cdot \ell(y)\, dy.$$

Notes:

Example 2   **Finding fluid force**

Consider a thin plate in the shape of an isosceles triangle as shown in Figure 6.32 submerged in water with a weight–density of 62.4 lb/ft^3. If the bottom of the plate is 10 ft below the surface of the water, what is the total fluid force exerted on this plate?

   SOLUTION   We approach this problem in two different ways to illustrate the different ways Key Idea 16 can be implemented. First we will let $y = 0$ represent the surface of the water, then we will consider an alternate convention.

1. We let $y = 0$ represent the surface of the water; therefore the bottom of the plate is at $y = -10$. We center the triangle on the y-axis as shown in Figure 6.33. The depth of the plate at y is $-y$ as indicated by the Key Idea. We now consider the length of the plate at y.

   We need to find equations of the left and right edges of the plate. The right hand side is a line that connects the points $(0, -10)$ and $(2, -6)$: that line has equation $x = \frac{1}{2}(y + 10)$. (Find the equation in the familiar $y = mx + b$ format and solve for x.) Likewise, the left hand side is described by the line $x = -\frac{1}{2}(y + 10)$. The total length is the distance between these two lines: $\ell(y) = \frac{1}{2}(y + 10) - (-\frac{1}{2}(y + 10)) = y + 10$.

   The total fluid force is then:

   $$F = \int_{-10}^{-6} 62.4(-y)(y + 10)\, dy$$

   $$= 62.4 \cdot \frac{176}{3} \approx 3660.8 \text{ lb}.$$

2. Sometimes it seems easier to orient the thin plate nearer the origin. For instance, consider the convention that the bottom of the triangular plate is at $(0, 0)$, as shown in Figure 6.34. The equations of the left and right hand sides are easy to find. They are $y = 2x$ and $y = -2x$, respectively, which we rewrite as $x = y/2$ and $x = -y/2$. Thus the length function is $\ell(y) = y/2 - (-y/2) = y$.

   As the surface of the water is 10 ft above the base of the plate, we have that the surface of the water is at $y = 10$. Thus the depth function is the distance between $y = 10$ and y; $d(y) = 10 - y$. We compute the total fluid force as:

   $$F = \int_0^4 62.4(10 - y)(y)\, dy$$

   $$\approx 3660.8 \text{ lb}.$$

The correct answer is, of course, independent of the placement of the plate in the coordinate plane as long as we are consistent.

Notes:
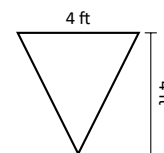
Figure 6.32: A thin plate in the shape of an isosceles triangle in Example 2.
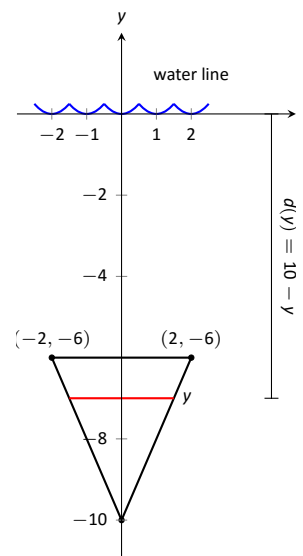
Figure 6.33: Sketching the triangular plate in Example 2 with the convention that the water level is at $y = 0$.
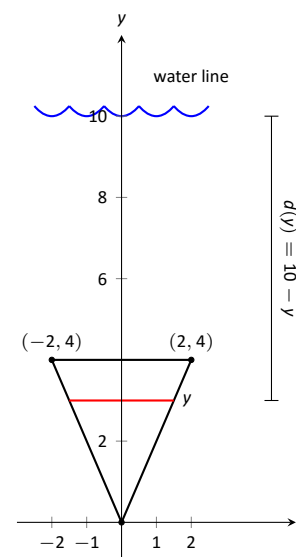
Figure 6.34: Sketching the triangular plate in Example 2 with the convention that the base of the triangle is at $(0, 0)$.
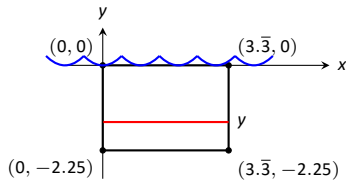
Figure 6.35:   Sketching a submerged car door in Example 3.

Example 3   **Finding fluid force**
Find the total fluid force on a car door submerged up to the bottom of its window in water, where the car door is a rectangle 40" long and 27" high (based on the dimensions of a 2005 Fiat Grande Punto.)

SOLUTION   The car door, as a rectangle, is drawn in Figure 6.35. Its length is $10/3$ ft and its height is 2.25 ft. We adopt the convention that the top of the door is at the surface of the water, both of which are at $y = 0$. Using the weight–density of water of 62.4 lb/ft^3, we have the total force as

$$F = \int_{-2.25}^{0} 62.4(-y)10/3 \; dy$$

$$= \int_{-2.25}^{0} -208y \; dy$$

$$= -104y^2 \Big|_{-2.25}^{0}$$

$$= 526.5 \text{ lb.}$$

Most adults would find it very difficult to apply over 500 lb of force to a car door while seated inside, making the door effectively impossible to open. This is counter–intuitive as most assume that the door would be relatively easy to open. The truth is that it is not, hence the survival tips mentioned at the beginning of this section.
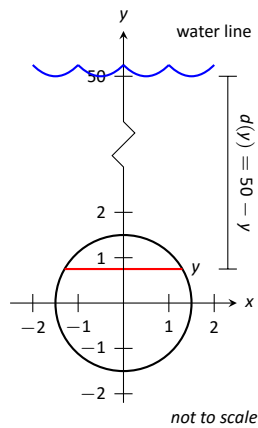
Figure 6.36: Measuring the fluid force on an underwater porthole in Example 4.

Example 4   **Finding fluid force**
An underwater observation tower is being built with circular viewing portholes enabling visitors to see underwater life. Each vertically oriented porthole is to have a 3 ft diameter whose center is to be located 50 ft underwater. Find the total fluid force exerted on each porthole. Also, compute the fluid force on a horizontally oriented porthole that is under 50 ft of water.

SOLUTION   We place the center of the porthole at the origin, meaning the surface of the water is at $y = 50$ and the depth function will be $d(y) = 50-y$; see Figure 6.36
   The equation of a circle with a radius of 1.5 is $x^2 + y^2 = 2.25$; solving for x we have $x = \pm\sqrt{2.25 - y^2}$, where the positive square root corresponds to the right side of the circle and the negative square root corresponds to the left side of the circle. Thus the length function at depth y is $\ell(y) = 2\sqrt{2.25 - y^2}$.

Notes:

Integrating on $[-1.5, 1.5]$ we have:

$$F = 62.4 \int_{-1.5}^{1.5} 2(50 - y)\sqrt{2.25 - y^2}\, dy$$

$$= 62.4 \int_{-1.5}^{1.5} \left(100\sqrt{2.25 - y^2} - 2y\sqrt{2.25 - y^2}\right) dy$$

$$= 6240 \int_{-1.5}^{1.5} \left(\sqrt{2.25 - y^2}\right) dy - 62.4 \int_{-1.5}^{1.5} \left(2y\sqrt{2.25 - y^2}\right) dy.$$

The second integral above can be evaluated using Substitution. Let $u = 2.25 - y^2$ with $du = -2y\, dy$. The new bounds are: $u(-1.5) = 0$ and $u(1.5) = 0$; the new integral will integrate from $u = 0$ to $u = 0$, hence the integral is 0.

The first integral above finds the area of half a circle of radius 1.5, thus the first integral evaluates to $6240 \cdot \pi \cdot 1.5^2/2 = 22,054$. Thus the total fluid force on a vertically oriented porthole is $22,054$ lb.

Finding the force on a horizontally oriented porthole is more straightforward:

$$F = \text{Pressure} \times \text{Area} = 62.4 \cdot 50 \times \pi \cdot 1.5^2 = 22,054 \text{ lb}.$$

That these two forces are equal is not coincidental; it turns out that the fluid force applied to a vertically oriented circle whose center is at depth d is the same as force applied to a horizontally oriented circle at depth d.

We end this chapter with a reminder of the true skills meant to be developed here. We are not truly concerned with an ability to find fluid forces or the volumes of solids of revolution. Work done by a variable force is important, though measuring the work done in pulling a rope up a cliff is probably not.

What we are actually concerned with is the ability to solve certain problems by first approximating the solution, then refining the approximation, then recognizing if/when this refining process results in a definite integral through a limit. Knowing the formulas found inside the special boxes within this chapter is beneficial as it helps solve problems found in the exercises, and other mathematical skills are strengthened by properly applying these formulas. However, more importantly, understand how each of these formulas was constructed. Each is the result of a summation of approximations; each summation was a Riemann sum, allowing us to take a limit and find the exact answer through a definite integral.

Notes:

Exercises 6.5

Terms and Concepts

1. State in your own words Pascal's Principle.

2. State in your own words how pressure is different from force.

Problems

In Exercises 3–12, find the fluid force exerted on the given plate, submerged in water with a weight density of 62.4 lb/ft^3.

3.

1 ft

2 ft

2 ft

4.

1 ft

2 ft

1 ft

5.

5 ft

6 ft

4 ft

6.

5 ft

4 ft

6 ft

7.

5 ft

2 ft

8.

5 ft

4 ft

9.

5 ft

2 ft

4 ft

10.

5 ft

4 ft

2 ft

11.

1 ft

2 ft

2 ft

12.

1 ft

2 ft

2 ft

2 ft
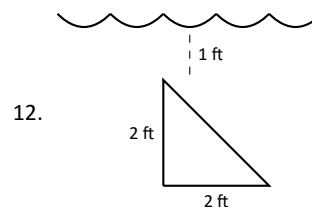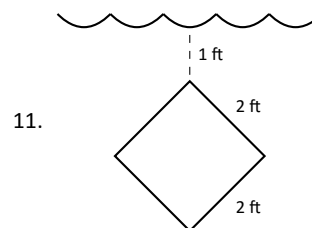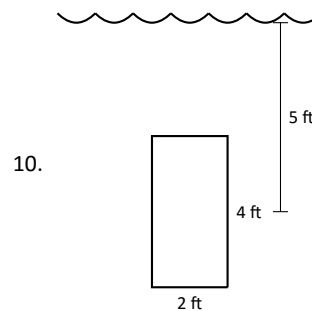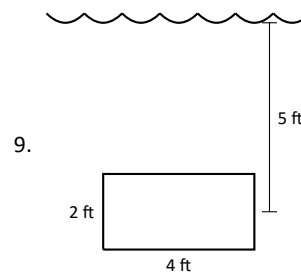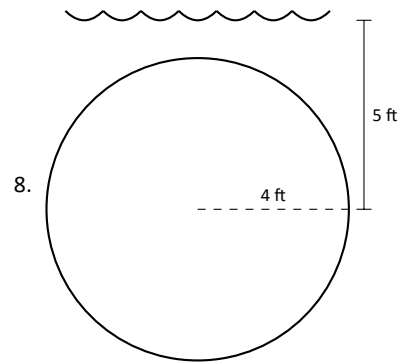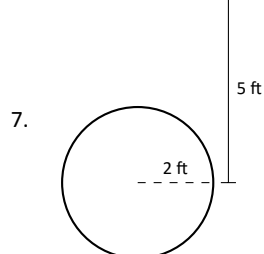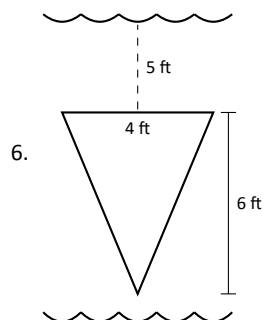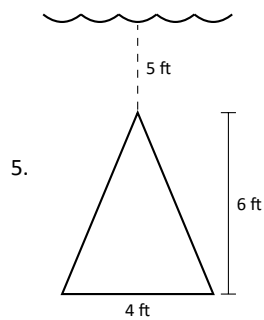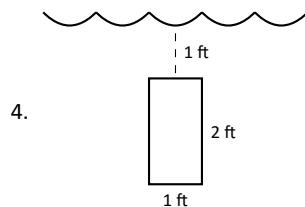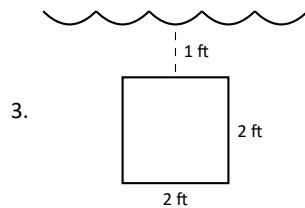
In Exercises 13–18, the side of a container is pictured. Find the fluid force exerted on this plate when the container is full of:

1. water, with a weight density of 62.4 lb/ft^3, and

2. concrete, with a weight density of 150 lb/ft^3.

13.

5 ft

3 ft

14.

4 ft

4 ft

$y = x^2$

15.

$y = 4 - x^2$

4 ft

4 ft

16.

2 ft

$y = -\sqrt{1 - x^2}$

17.

$y = \sqrt{1 - x^2}$

2 ft

18.

6 ft

$y = -\sqrt{9 - x^2}$
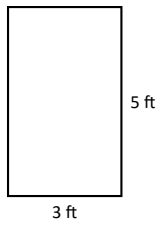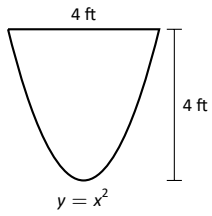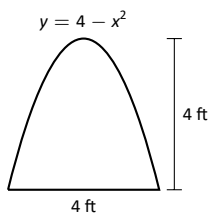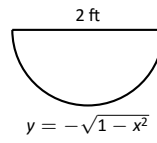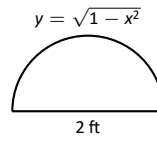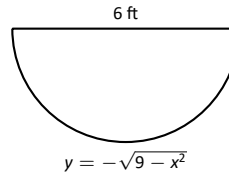
19. How deep must the center of a vertically oriented circular plate with a radius of 1 ft be submerged in water, with a weight density of 62.4 lb/ft^3, for the fluid force on the plate to reach 1,000 lb?

20. How deep must the center of a vertically oriented square plate with a side length of 2 ft be submerged in water, with a weight density of 62.4 lb/ft^3, for the fluid force on the plate to reach 1,000 lb?

Notes:

Calculus II

7: Inverse Functions and L'Hôpital's Rule

This chapter completes our differentiation toolkit. The first and most important tool will be how to differentiate inverse functions. We'll be able to use this to differentiate exponential and logarithmic functions, which we stated in Theorem 14 but did not prove.

7.1   Inverse Functions

We say that two functions f and g are *inverses* if $g(f(x)) = x$ for all x in the domain of f and $f(g(x)) = x$ for all x in the domain of g. A function can only have an inverse if it is one-to-one, i.e. if we never have $f(x_1) = f(x_2)$ for different elements x_1 and x_2 of the domain. This is equivalent to saying that the graph of the function passes the horizontal line test. The inverse of f is denoted f^{-1}, which should not be confused with the function $1/f(x)$.

Key Idea 17   Inverse Functions

For a one-to–one function f,

- The domain of f^{-1} is the range of f; the range of f^{-1} is the domain of f.

- $f^{-1}(f(x)) = x$ for all x in the domain of f.

- $f(f^{-1}(x)) = x$ for all x in the domain of f^{-1}.

- The graph of $y = f^{-1}(x)$ is the reflection across $y = x$ of the graph of $y = f(x)$.

- $y = f^{-1}(x)$ if and only if $f(y) = x$ and y is in the domain of f.

Watch the video:
Finding the Inverse of a Function or Showing One Does not Exist, Ex 3 at
`https://youtu.be/BmjbDINGZGg`
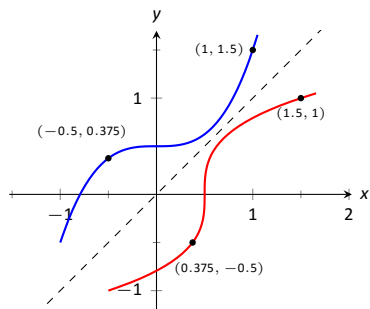
Figure 7.1: A function f along with its inverse f^{-1}. (Note how it does not matter which function we refer to as f; the other is f^{-1}.)

To determine whether or not f and g are inverses for each other, we check to see whether or not $g(f(x)) = x$ for all x in the domain of f, and $f(g(x)) = x$ for all x in the domain of g.

Example 1   **Verifying Inverses**

Determine whether or not the following pairs of functions are inverses:

1. $f(x) = 3x + 1$; $g(x) = \dfrac{x - 1}{3}$

2. $f(x) = x^3 + 1$; $g(x) = x^{1/3} - 1$

SOLUTION

1. To check the composition we plug $f(x)$ in for x in the definition of g as follows:

$$g(f(x)) = \frac{f(x) - 1}{3} = \frac{(3x + 1) - 1}{3} = \frac{3x}{3} = x$$

So $g(f(x)) = x$ for all x in the domain of f. Likewise, you can check that $f(g(x)) = x$ for all x in the domain of g, so f and g are inverses.

2. If we try to proceed as before, we find that:

$$g(f(x)) = (f(x))^{1/3} - 1 = (x^3 + 1)^{1/3} - 1$$

This doesn't seem to be the same as the identity function x. To verify this, we find a number a in the domain of f and show that $g(f(a)) \neq a$ for that value. Let's try $x = 1$. Since $f(1) = 1^3 + 1 = 2$, we find that $g(f(1)) = g(2) = 2^{1/3} - 1 \approx 0.26$. Since $g(f(1)) \neq 1$, these functions are not inverses.

Functions that are not one-to-one.
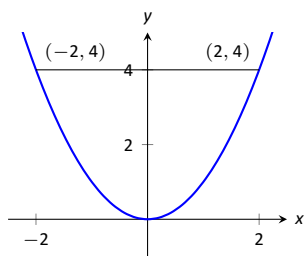
Figure 7.2: The function $f(x) = x^2$ is not one-to-one.

Unfortunately, not every function we would like to find an inverse for is one-to-one. For example, the function $f(x) = x^2$ is not one-to-one because $f(-2) = f(2) = 4$. If f^{-1} is an inverse for f, then $f^{-1}(f(-2)) = -2$ implies that $f^{-1}(4) = -2$. On the other hand, $f^{-1}(f(2)) = 2$, so $f^{-1}(4) = 2$. We cannot have it both ways if f^{-1} is a function, so no such inverse exists. We can find a partial solution to this dilemma by restricting the domain of f. There are many possible choices, but traditionally we restrict the domain to the interval $[0, \infty)$. The function $f^{-1}(x) = \sqrt{x}$ is now an inverse for this restricted version of f.

Notes:

The inverse sine function

We consider the function $f(x) = \sin x$, which is not one-to-one. A piece of the graph of f is in Figure 7.3(a). In order to find an appropriate restriction of the domain of f, we look for consecutive critical points where f takes on its minimum and maximum values. In this case, we use the interval $[-\pi/2, \pi/2]$. We define the inverse of f on this restricted range by $y = \sin^{-1} x$ if and only if $\sin y = x$ and $-\pi/2 \leq y \leq \pi/2$. The graph is a reflection of the graph of g across the line $y = x$, as seen in Figure 7.3(b).

(a)   (b)
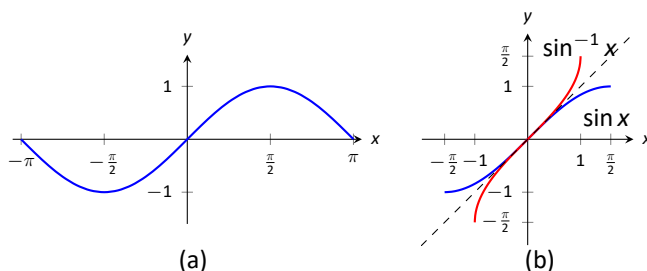
Figure 7.3: (a) A portion of $y = \sin x$. (b) A one–to–one portion of $y = \sin x$ along with $y = \sin^{-1} x$.

The inverse tangent function

Next we consider the function $f(x) = \tan x$, which is also not one-to-one. A piece of the graph of f is given in Figure 7.4(a). In order to find an interval on which the function is one-to-one and on which the function takes on all values in the range, we use an interval between consecutive vertical asymptotes. Traditionally, the interval $(-\pi/2, \pi/2)$ is chosen. Note that we choose the open interval in this case because the function f is not defined at the endpoints. So we define $y = \tan^{-1} x$ if and only if $\tan y = x$ and $-\pi/2 < y < \pi/2$. The graph of $y = \tan^{-1} x$ is shown in Figure 7.4(b). Also note that the vertical asymptotes of the original function are reflected to become horizontal asymptotes of the inverse function.

Notes:
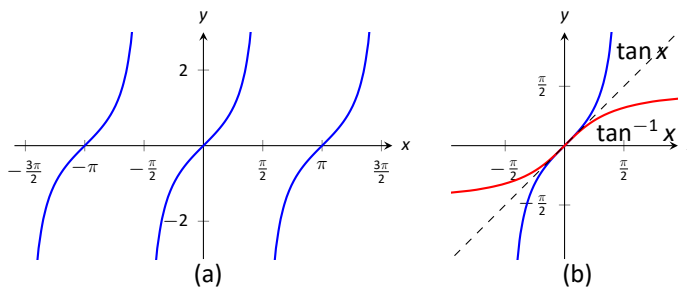
Figure 7.4: (a) A portion of $y = \tan x$. (b) A one–to–one portion of $y = \tan x$ along with $y = \tan^{-1}$.

The other inverse trigonometric functions are defined in a similar fashion. The resulting domains and ranges are summarized in Figure 7.5.

Function	Restricted Domain	Range	Inverse Function	Domain	Range
$\sin x$	$[-\pi/2, \pi/2]$	$[-1, 1]$	$\sin^{-1} x$	$[-1, 1]$	$[-\pi/2, \pi/2]$
$\cos x$	$[0, \pi]$	$[-1, 1]$	$\cos^{-1} x$	$[-1, 1]$	$[0, \pi]$
$\tan x$	$(-\pi/2, \pi/2)$	$(-\infty, \infty)$	$\tan^{-1} x$	$(-\infty, \infty)$	$(-\pi/2, \pi/2)$
$\csc x$	$[-\pi/2, 0) \cup (0, \pi/2]$	$(-\infty, -1] \cup [1, \infty)$	$\csc^{-1} x$	$(-\infty, -1] \cup [1, \infty)$	$[-\pi/2, 0) \cup (0, \pi/2]$
$\sec x$	$[0, \pi/2) \cup (\pi/2, \pi]$	$(-\infty, -1] \cup [1, \infty)$	$\sec^{-1} x$	$(-\infty, -1] \cup [1, \infty)$	$[0, \pi/2) \cup (\pi/2, \pi]$
$\cot x$	$(0, \pi)$	$(-\infty, \infty)$	$\cot^{-1} x$	$(-\infty, \infty)$	$(0, \pi)$

Figure 7.5: Domains and ranges of the trigonometric and inverse trigonometric functions.

Sometimes, arcsin is used instead of $\sin^{-1}$. Similar "arc" functions are used for the other inverse trigonometric functions as well.

Example 2   Evaluating Inverse Trigonometric Functions

Find exact values for the following:

1. $\tan^{-1}(1)$

2. $\cos(\sin^{-1}(\sqrt{3}/2))$

3. $\sin^{-1}(\sin(7\pi/6))$

4. $\tan(\cos^{-1}(11/15))$

SOLUTION

1. $\tan^{-1}(1) = \pi/4$

2. $\cos(\sin^{-1}(\sqrt{3}/2)) = \cos(\pi/3) = 1/2$

3. Since $7\pi/6$ is not in the range of the inverse sine function, we should be careful with this one.

$$\sin^{-1}(\sin(7\pi/6)) = \sin^{-1}(-1/2) = -\pi/6.$$

4. Since we don't know the value of $\cos^{-1}(11/15)$, we let θ stand for this value. We know that θ is an angle between 0 and π and that $\cos(\theta) =$

Notes:

11/15. In Figure 7.6, we use this information to construct a right triangle with angle θ, where the adjacent side over the hypotenuse must equal 11/15. Applying the Pythagorean Theorem we find that

$$y = \sqrt{15^2 - 11^2} = \sqrt{104} = 2\sqrt{26}.$$

Finally, we have:

$$\tan(\cos^{-1}(11/15)) = \tan(\theta) = \frac{2\sqrt{26}}{11}.$$
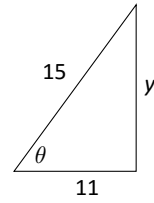
Figure 7.6: A right triangle for the situation in Example 2 (4).

Notes:

Exercises 7.1

Terms and Concepts

1. T/F: Every function has an inverse.

2. In your own words explain what it means for a function to be "one to one."

3. If $(1, 10)$ lies on the graph of $y = f(x)$, what can be said about the graph of $y = f^{-1}(x)$?

Problems

In Exercises 4–8, verify that the given functions are inverses.

4. $f(x) = 2x + 6$ and $g(x) = \frac{1}{2}x - 3$

5. $f(x) = x^2 + 6x + 11$, $x \geq 3$ and $g(x) = \sqrt{x - 2} - 3$, $x \geq 2$

6. $f(x) = x^2 + 6x + 11$, $x \leq 3$ and $g(x) = -\sqrt{x - 2} - 3$, $x \geq 2$.

7. $f(x) = \dfrac{3}{x - 5}$, $x \neq 5$ and $g(x) = \dfrac{3 + 5x}{x}$, $x \neq 0$

8. $f(x) = \dfrac{x + 1}{x - 1}$, $x \neq 1$ and $g(x) = f(x)$

In Exercises 9–12, find a restriction of the domain of the given function on which the function will have an inverse.

9. $f(x) = \sqrt{16 - x^2}$

10. $g(x) = \sqrt{x^2 - 16}$

11. $r(t) = t^2 - 6t + 9$

12. $f(x) = \dfrac{1 - \sqrt{x}}{1 + \sqrt{x}}$

In Exercises 13–19, find the exact value.

13. $\tan^{-1}(0)$

14. $\tan^{-1}(\tan(\pi/7))$

15. $\cos(\cos^{-1}(-1/5))$

16. $\sin^{-1}(\sin(8\pi/3))$

17. $\sin(\tan^{-1}(1))$

18. $\cos(\tan^{-1}(3/7))$

19. $\sec(\sin^{-1}(-3/5))$

7.2 Derivatives of Inverse Functions

In this section we will figure out how to differentiate the inverse of a function. To do so, we recall that if f and g are inverses, then $f(g(x)) = x$ for all x in the domain of f. Differentiating and simplifying yields:

$$f(g(x)) = x$$
$$f'(g(x))g'(x) = 1$$
$$g'(x) = \frac{1}{f'(g(x))} \quad \text{assuming } f'(x) \text{ is nonzero}$$

Note that the derivation above assumes that the function g is differentiable. It is possible to prove that g must be differentiable if f' is nonzero, but the proof is beyond the scope of this text. However, assuming this fact we have shown the following:

Theorem 46   **Derivatives of Inverse Functions**

Let f be differentiable and one-to-one on an open interval I, where $f'(x) \neq 0$ for all x in I, let J be the range of f on I, let g be the inverse function of f, and let $f(a) = b$ for some a in I. Then g is a differentiable function on J, and in particular,

$$\left(f^{-1}\right)'(b) = g'(b) = \frac{1}{f'(a)}$$

$$\left(f^{-1}\right)'(x) = g'(x) = \frac{1}{f'(g(x))}$$

The results of Theorem 46 are not trivial; the notation may seem confusing at first. Careful consideration, along with examples, should earn understanding.

Watch the video:
Derivative of an Inverse Function, Ex 2 at
`https://youtu.be/RKfGMXOpn2k`

In the next example we apply Theorem 46 to the arcsine function.

Example 1   **Finding the derivative of an inverse trigonometric function**
Let $y = \sin^{-1} x$. Find y' using Theorem 46.

Notes:

SOLUTION   Adopting our previously defined notation, let $g(x) = \sin^{-1} x$ and $f(x) = \sin x$. Thus $f'(x) = \cos x$. Applying Theorem 46, we have

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

$$= \frac{1}{\cos(\sin^{-1} x)}.$$
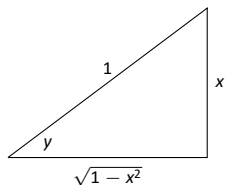
Figure 7.7: A right triangle defined by $y = \sin^{-1}(x/1)$ with the length of the third leg found using the Pythagorean Theorem.

This last expression is not immediately illuminating. Drawing a figure will help, as shown in Figure 7.7. Recall that the sine function can be viewed as taking in an angle and returning a ratio of sides of a right triangle, specifically, the ratio "opposite over hypotenuse." This means that the arcsine function takes as input a ratio of sides and returns an angle. The equation $y = \sin^{-1} x$ can be rewritten as $y = \sin^{-1}(x/1)$; that is, consider a right triangle where the hypotenuse has length 1 and the side opposite of the angle with measure y has length x. This means the final side has length $\sqrt{1 - x^2}$, using the Pythagorean Theorem.

Therefore $\cos(\sin^{-1} x) = \cos y = \sqrt{1 - x^2}/1 = \sqrt{1 - x^2}$, resulting in

$$\frac{d}{dx}\left(\sin^{-1} x\right) = g'(x) = \frac{1}{\sqrt{1 - x^2}}.$$

Remember that the input x of the arcsine function is a ratio of a side of a right triangle to its hypotenuse; the absolute value of this ratio will be less than 1. Therefore $1 - x^2$ will be positive.

In order to make $y = \sin x$ one-to-one, we restrict its domain to $[-\pi/2, \pi/2]$; on this domain, the range is $[-1, 1]$. Therefore the domain of $y = \sin^{-1} x$ is $[-1, 1]$ and the range is $[-\pi/2, \pi/2]$. When $x = \pm 1$, note how the derivative of the arcsine function is undefined; this corresponds to the fact that as $x \to \pm 1$, the tangent lines to arcsine approach vertical lines with undefined slopes.

In Figure 7.8 we see $f(x) = \sin x$ and $f^{-1}(x) = \sin^{-1} x$ graphed on their respective domains. The line tangent to $\sin x$ at the point $(\pi/3, \sqrt{3}/2)$ has slope $\cos \pi/3 = 1/2$. The slope of the corresponding point on $\sin^{-1} x$, the point $(\sqrt{3}/2, \pi/3)$, is

$$\frac{1}{\sqrt{1 - (\sqrt{3}/2)^2}} = \frac{1}{\sqrt{1 - 3/4}} = \frac{1}{\sqrt{1/4}} = \frac{1}{1/2} = 2,$$

verifying Theorem 46 yet again: at corresponding points, a function and its inverse have reciprocal slopes.

Using similar techniques, we can find the derivatives of all the inverse trigonometric functions after first restricting their domains according to Figure 7.5 to allow them to be invertible.
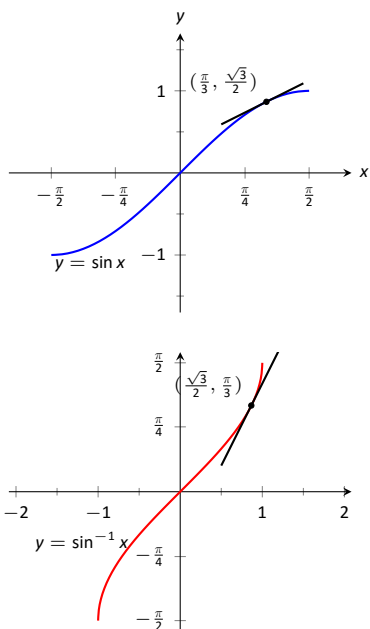
Figure 7.8: Graphs of $y = \sin x$ and $y = \sin^{-1} x$ along with corresponding tangent lines.

Notes:

> **Theorem 47   Derivatives of Inverse Trigonometric Functions**
> The inverse trigonometric functions are differentiable on all open sets contained in their domains (as listed in Figure 7.5) and their derivatives are as follows:
>
> 1. $\dfrac{d}{dx}\left(\sin^{-1}x\right) = \dfrac{1}{\sqrt{1-x^2}}$ 4. $\dfrac{d}{dx}\left(\cos^{-1}x\right) = -\dfrac{1}{\sqrt{1-x^2}}$
>
> 2. $\dfrac{d}{dx}\left(\sec^{-1}x\right) = \dfrac{1}{|x|\sqrt{x^2-1}}$ 5. $\dfrac{d}{dx}\left(\csc^{-1}x\right) = -\dfrac{1}{|x|\sqrt{x^2-1}}$
>
> 3. $\dfrac{d}{dx}\left(\tan^{-1}x\right) = \dfrac{1}{1+x^2}$ 6. $\dfrac{d}{dx}\left(\cot^{-1}x\right) = -\dfrac{1}{1+x^2}$

Note how the last three derivatives are merely the negatives of the first three, respectively. Because of this, the first three are used almost exclusively throughout this text.

Example 2   Finding derivatives of inverse functions
Find the derivatives of the following functions:

1. $f(x) = \cos^{-1}(x^2)$ 2. $g(x) = \dfrac{\sin^{-1}x}{\sqrt{1-x^2}}$ 3. $f(x) = \sin^{-1}(\cos x)$

Solution

1. We use Theorem 47 and the Chain Rule to find:

$$f'(x) = -\frac{1}{\sqrt{1-(x^2)^2}}(2x) = -\frac{2x}{\sqrt{1-x^4}}$$

2. We use Theorem 47 and the Quotient Rule to compute:

$$g'(x) = \frac{\left(\frac{1}{\sqrt{1-x^2}}\right)\sqrt{1-x^2} - (\sin^{-1}x)\left(\frac{1}{2\sqrt{1-x^2}}(-2x)\right)}{\left(\sqrt{1-x^2}\right)^2}$$

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

Notes:

3. We apply Theorem 47 and the Chain Rule again to compute:

$$f'(x) = \frac{1}{\sqrt{1 - \cos^2 x}}(-\sin x)$$
$$= \frac{-\sin x}{\sqrt{\sin^2 x}}$$
$$= \frac{-\sin x}{\sin x}$$
$$= -1.$$

Theorem 47 allows us to integrate some functions that we could not integrate before. For example,

$$\int \frac{dx}{\sqrt{1 - x^2}} = \sin^{-1} x + C.$$

Combining these formulas with u-substitution yields the following:

Theorem 48   Integrals Involving Inverse Trigonometric Functions
Let $a > 0$.

1. $\int \dfrac{1}{a^2 + x^2}\, dx = \dfrac{1}{a} \tan^{-1}\left(\dfrac{x}{a}\right) + C$

2. $\int \dfrac{1}{\sqrt{a^2 - x^2}}\, dx = \sin^{-1}\left(\dfrac{x}{a}\right) + C$

3. $\int \dfrac{1}{x\sqrt{x^2 - a^2}}\, dx = \dfrac{1}{a} \sec^{-1}\left(\dfrac{|x|}{a}\right) + C$

We will look at the second part of this theorem. The other parts are similar and are left as exercises.

First we note that the integrand involves the number a^2, but does not explicitly involve a. We make the assumption that $a > 0$ in order to simplify what follows. We can rewrite the integral as follows:

$$\int \frac{dx}{\sqrt{a^2 - x^2}} = \int \frac{dx}{\sqrt{a^2(1 - (x/a)^2)}} = \int \frac{dx}{a\sqrt{1 - (x/a)^2}}$$

Notes:

We next use the substitution $u = x/a$ and $du = dx/a$ to find:

$$\int \frac{dx}{a\sqrt{1-(x/a)^2}} = \int \frac{a}{a\sqrt{1-u^2}}\,du$$

$$= \int \frac{du}{\sqrt{1-u^2}}$$

$$= \sin^{-1}u + C$$

$$= \sin^{-1}(x/a) + C$$

We conclude this section with several examples.

Example 3   **Finding antiderivatives involving inverse functions**
Find the following integrals.

$$1. \quad \int \frac{dx}{100 + x^2} \qquad 2. \quad \int \frac{\sin^{-1}x}{\sqrt{1-x^2}}\,dx \qquad 3. \int \frac{dx}{x^2 + 2x + 5}$$

SOLUTION

1. $\displaystyle \int \frac{dx}{100 + x^2} = \int \frac{dx}{10^2 + x^2} = \frac{1}{10}\tan^{-1}(x/10) + C$

2. We use the substitution $u = \sin^{-1}x$ and $du = \frac{dx}{\sqrt{1-x^2}}$ to find:

$$\int \frac{\sin^{-1}x}{\sqrt{1-x^2}} = \int u\,du = \frac{1}{2}u^2 + C = \frac{1}{2}\left(\sin^{-1}x\right)^2 + C$$

3. This does not immediately look like one of the forms in Theorem 48, but we can complete the square in the denominator to see that

$$\int \frac{dx}{x^2 + 2x + 5} = \int \frac{dx}{(x^2 + 2x + 1) + 4} = \int \frac{dx}{4 + (x+1)^2}$$

We now use the substitution $u = x + 1$ and $du = dx$ to find:

$$\int \frac{dx}{4 + (x+1)^2} = \int \frac{du}{4 + u^2} = \frac{1}{2}\tan^{-1}(u/2) + C = \frac{1}{2}\tan^{-1}\left(\frac{x+1}{2}\right) + C.$$

Notes:

Exercises 7.2

Terms and Concepts

1. If $(1, 10)$ lies on the graph of $y = f(x)$ and $f'(1) = 5$, what can be said about $y = f^{-1}(x)$?

Problems

In Exercises 2–6, an invertible function $f(x)$ is given along with a point that lies on its graph. Using Theorem 46, evaluate $\left(f^{-1}\right)'(x)$ at the indicated value.

2. $f(x) = 5x + 10$
   Point$= (2, 20)$
   Evaluate $\left(f^{-1}\right)'(20)$

3. $f(x) = x^2 - 2x + 4,\ x \geq 1$
   Point$= (3, 7)$
   Evaluate $\left(f^{-1}\right)'(7)$

4. $f(x) = \sin 2x,\ -\pi/4 \leq x \leq \pi/4$
   Point$= (\pi/6, \sqrt{3}/2)$
   Evaluate $\left(f^{-1}\right)'(\sqrt{3}/2)$

5. $f(x) = x^3 - 6x^2 + 15x - 2$
   Point$= (1, 8)$
   Evaluate $\left(f^{-1}\right)'(8)$

6. $f(x) = \dfrac{1}{1 + x^2},\ x \geq 0$
   Point$= (1, 1/2)$
   Evaluate $\left(f^{-1}\right)'(1/2)$

In Exercises 7–15, compute the derivative of the given function.

7. $h(t) = \sin^{-1}(2t)$

8. $f(t) = \sec^{-1}(2t)$

9. $g(x) = \tan^{-1}(2x)$

10. $f(x) = x \sin^{-1} x$

11. $g(t) = \sin t \cos^{-1} t$

12. $h(x) = \dfrac{\sin^{-1} x}{\cos^{-1} x}$

13. $g(x) = \tan^{-1}(\sqrt{x})$

14. $f(x) = \sec^{-1}(1/x)$

15. $f(x) = \sin(\sin^{-1} x)$

In Exercises 16–18, compute the derivative of the given function in two ways:

    (a) By simplifying first, then taking the derivative, and

    (b) by using the Chain Rule first then simplifying.

Verify that the two answers are the same.

16. $f(x) = \sin(\sin^{-1} x)$

17. $f(x) = \tan^{-1}(\tan x)$

18. $f(x) = \sin(\cos^{-1} x)$

In Exercises 19–20, find the equation of the line tangent to the graph of f at the indicated x value.

19. $f(x) = \sin^{-1} x$ at $x = \frac{\sqrt{2}}{2}$

20. $f(x) = \cos^{-1}(2x)$ at $x = \frac{\sqrt{3}}{4}$

21. A regulation hockey goal is 6 feet wide. If a player is skating towards the end line on a line perpendicular to the end line and 10 feet from the imaginary line joining the center of one goal to the center of the other, the angle between the player and the goal first increases and then begins to decrease. In order to maximize this angle, how far from the end line should the player be when they shoot the puck?
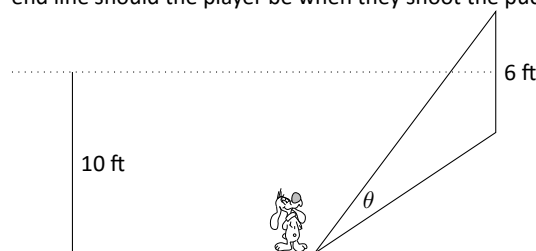
In Exercises 22–28, compute the indicated integral.

22. $\displaystyle \int_{1/\sqrt{2}}^{1/2} \frac{2}{\sqrt{1 - x^2}}\, dx$

23. $\displaystyle \int_{0}^{\sqrt{3}} \frac{4}{9 + x^2}\, dx$

24. $\displaystyle \int \frac{\sin^{-1} r}{\sqrt{1 - r^2}}\, dr$

25. $\displaystyle \int \frac{x^3}{4 + x^8}\, dx$

26. $\displaystyle \int \frac{e^t}{\sqrt{10 - e^{2t}}}\, dt$

27. $\displaystyle \int \frac{1}{\sqrt{3 - x^2 + 2x}}\, dx$

28. $\displaystyle \int \frac{1}{\sqrt{x}(1 + x)}\, dx$

7.3   Exponential and Logarithmic Functions

In this section we will define general exponential and logarithmic functions and find their derivatives.

General exponential functions

Consider first the function $f(x) = 2^x$. If x is rational, then we know how to compute 2^x. What do we mean by 2^π though? We compute this by first looking at 2^r for rational numbers r that are very close to π, then finding a limit. In our case we might compute 2^3, $2^{3.1}$, $2^{3.14}$, etc. We then define 2^π to be the limit of these numbers. Note that this is actually a different kind of limit than we have dealt with before since we only consider rational number close to π, not all real numbers close to π. We will see one way to make this more precise in Chapter 9. Graphically, we can plot the values of 2^x for x rational and get something like the dotted curve in Figure 7.9. In order to define the remaining values, we are "connecting the dots" in a way that makes the function continuous.

It follows from continuity and the properties of limits that exponential functions will satisfy the familiar properties of exponents (see Section 2.0). This implies that

$$\left(\frac{1}{2}\right)^x = (2^{-1})^x = 2^{-x},$$

so the graph of $g(x) = (1/2)^x$ is the reflection of f across the y-axis, as in Figure 7.10.

We can go through the same process as above for any base $a > 0$, though we are not usually interested in the constant function 1^x.
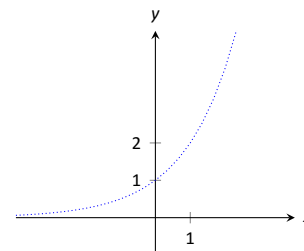
Figure 7.9: The function 2^x for rational values of x.
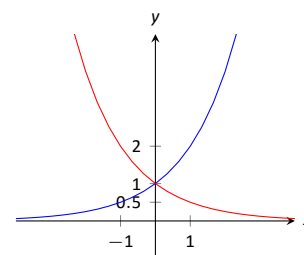
Figure 7.10: The functions 2^x and 2^{-x}.

Key Idea 18   Properties of Exponential Functions
For $a > 0$ and $a \neq 1$ the exponential function $f(x) = a^x$ satisfies:

1. $a^0 = 1$

2. $\lim_{x \to \infty} a^x = \begin{cases} \infty & a > 1 \\ 0 & a < 1 \end{cases}$

3. $a^x > 0$ for all x

4. $\lim_{x \to -\infty} a^x = \begin{cases} 0 & a > 1 \\ \infty & a < 1 \end{cases}$

Notes:

Derivatives of exponential functions

Suppose $f(x) = a^x$ for some $a > 0$. We can use the rules of exponents to find the derivative of f:

$$
\begin{aligned}
f'(x) &= \lim_{h \to 0} \frac{f(x+h) - f(x)}{h} \\
&= \lim_{h \to 0} \frac{a^{x+h} - a^x}{h} \\
&= \lim_{h \to 0} \frac{a^x a^h - a^x}{h} \\
&= \lim_{h \to 0} \frac{a^x(a^h - 1)}{h} \\
&= a^x \lim_{h \to 0} \frac{a^h - 1}{h} \qquad \text{(since } a^x \text{ does not depend on } h\text{)}
\end{aligned}
$$

So we know that $f'(x) = a^x \lim_{h \to 0} \dfrac{a^h - 1}{h}$, but can we say anything about that remaining limit? First we note that

$$
f'(0) = \lim_{h \to 0} \frac{a^{0+h} - a^0}{h} = \lim_{h \to 0} \frac{a^h - 1}{h},
$$

so we have $f'(x) = a^x f'(0)$. The actual value of the limit $\lim_{h \to 0} \dfrac{a^h - 1}{h}$ depends on the base a, but it can be proved that it does exist. We will figure out just what this limit is later, but for now we note that the easiest differentiation formulas come from using a base a that makes $\lim_{h \to 0} \dfrac{a^h - 1}{h} = 1$. This base is the number $e \approx 2.71828$ and the exponential function e^x is called the natural exponential function. This leads to the following result.

> **Theorem 49   Derivative of Exponential Functions**
> For any base $a > 0$, the exponential function $f(x) = a^x$ has derivative $f'(x) = a^x f'(0)$. The natural exponential function $g(x) = e^x$ has derivative $g'(x) = e^x$.

Notes:

Watch the video:
Derivatives of Exponential Functions at
`https://youtu.be/U3PyUcEd7IU`

General logarithmic functions

Before reviewing general logarithmic functions, we'll first remind ourselves of the laws of logarithms.

Key Idea 19   Properties of Logarithms

For $a, x, y > 0$ and $a \neq 1$, we have

1. $\log_a(xy) = \log_a x + \log_a y$

2. $\log_a \dfrac{x}{y} = \log_a x - \log_a y$

3. $\log_x y = \dfrac{\log_a y}{\log_a x}$, when $x \neq 1$

4. $\log_a x^y = y \log_a x$

5. $\log_a 1 = 0$

6. $\log_a a = 1$

Let us consider the function $f(x) = a^x$ where $a \neq 1$. We know that $f'(x) = f'(0)a^x$, where $f'(0)$ is a constant that depends on the base a. Since $a^x > 0$ for all x, this implies that $f'(x)$ is either always positive or always negative, depending on the sign of $f'(0)$. This in turn implies that f is strictly monotonic, so f is one-to-one. We can now say that f has an inverse. We call this inverse the logarithm with base a, denoted $f^{-1}(x) = \log_a x$. When $a = e$, this is the natural logarithm function $\ln x$. So we can say that $y = \log_a x$ if and only if $a^y = x$. Since the range of the exponential function is the set of positive real numbers, the domain of the logarithm function is also the set of positive real numbers. Reflecting the graph of $y = a^x$ across the line $y = x$ we find that (for $a > 1$) the graph of the logarithm looks like Figure 7.11.
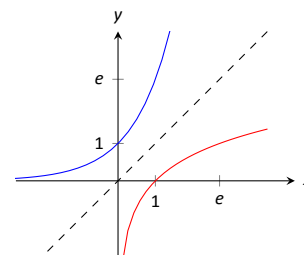
Figure 7.11: The functions $y = a^x$ and $y = \log_a x$ for $a > 1$.

Notes:

353

Key Idea 20   Properties of Logarithmic Functions

For $a > 0$ and $a \neq 1$ the logarithmic function $f(x) = \log_a x$ satisfies:

1. The domain of $f(x) = log_a x$ is $(0, \infty)$ and the range is $(-\infty, \infty)$.

2. $y = \log_a x$ if and only if $a^y = x$.

3. $\displaystyle\lim_{x \to \infty} \log_a x = \begin{cases} \infty & \text{if } a > 1 \\ -\infty & \text{if } a < 1 \end{cases}$

4. $\displaystyle\lim_{x \to 0^+} \log_a x = \begin{cases} -\infty & \text{if } a > 1 \\ \infty & \text{if } a < 1 \end{cases}$

Using the inverse of the natural exponential function, we can determine what the value of $f'(0)$ is in the formula $(a^x)' = f'(0)a^x$. To do so, we note that $a = e^{\ln a}$ since the exponential and logarithm functions are inverses. Hence we can write:

$$a^x = \left(e^{\ln a}\right)^x = e^{x \ln a}$$

Now since $\ln a$ is a constant, we can use the Chain Rule to see that:

$$\frac{d}{dx}a^x = \frac{d}{dx}e^{x \ln a} = e^{x \ln a}(\ln a) = a^x \ln a$$

Comparing this to our previous result, we can restate our theorem:

Theorem 50   Derivative of Exponential Functions

For any base $a > 0$, the exponential function $f(x) = a^x$ has derivative $f'(x) = a^x \ln a$. The natural exponential function $g(x) = e^x$ has derivative $g'(x) = e^x$.

Change of base

In the previous computation, we found it convenient to rewrite the general exponential function in terms of the natural exponential function. A related formula allows us to rewrite the general logarithmic function in terms of the natural

Notes:

logarithm. To see how this works, suppose that $y = \log_a x$, then we have:

$$a^y = x$$
$$\ln(a^y) = \ln x$$
$$y \ln a = \ln x$$
$$y = \frac{\ln x}{\ln a}$$
$$\log_a x = \frac{\ln x}{\ln a}.$$

This change of base formula allows us to use facts about the natural logarithm to derive facts about the general logarithm.

Derivatives of logarithmic functions

Since the natural logarithm function is the inverse of the natural exponential function, we can use the formula $(f^{-1}(x))' = \dfrac{1}{f'(f^{-1}(x))}$ to find the derivative of $y = \ln x$. We know that $\dfrac{d}{dx} e^x = e^x$, so we get:

$$\frac{d}{dx} \ln x = \frac{1}{e^y} = \frac{1}{e^{\ln x}} = \frac{1}{x}.$$

Now we can apply the change of base formula to find the derivative of a general logarithmic function:

$$\frac{d}{dx} \log_a x = \frac{d}{dx} \left(\frac{\ln x}{\ln a} \right) = \frac{1}{\ln a} \left(\frac{d}{dx} \ln x \right) = \frac{1}{x \ln a}.$$

Example 1   Finding Derivatives of Logs and Exponentials
Find derivatives of the following functions.

1. $f(x) = x 3^{4x-7}$ 2. $g(x) = 2^{x^2}$ 3. $h(x) = \dfrac{x}{\log_5 x}$

   SOLUTION

1. We apply both the Product and Chain Rules:
$$f'(x) = 3^{4x-7} + x\left(3^{4x-7} \ln 3\right)(4) = (1 + 4x \ln 3)3^{4x-7}$$

2. We apply the Chain Rule:
$$g'(x) = 2^{x^2} \ln 2 (2x) = 2^{x^2+1} x \ln 2.$$

3. Applying the Quotient Rule:
$$h'(x) = \frac{\log_5 x - x\left(\frac{1}{x \ln 5}\right)}{(\log_5 x)^2} = \frac{(\log_5 x)(\ln 5) - 1}{(\log_5 x)^2 \ln 5}$$

Notes:

355

Example 2   **The Derivative of the Natural Log**

Find the derivative of the function $y = \ln |x|$.

 Solution We can rewrite our function as

$$y = \begin{cases} \ln x & \text{if } x > 0 \\ \ln(-x) & \text{if } x < 0 \end{cases}$$

Applying the Chain Rule, we see that $\frac{dy}{dx} = \frac{1}{x}$ for $x > 0$, and $\frac{dy}{dx} = \frac{-1}{-x} = \frac{1}{x}$ for $x < 0$. Hence we have

$$\frac{d}{dx} \ln |x| = \frac{1}{x} \quad \text{for } x \neq 0.$$

Antiderivatives

Combining these new results, we arrive at the following theorem:

Theorem 51 **Derivatives and Antiderivatives of Exponentials and Logarithms**

Given a base $a > 0$ and $a \neq 1$, the following hold:

1. $\dfrac{d}{dx} e^x = e^x$

2. $\dfrac{d}{dx} a^x = a^x \ln a$

3. $\dfrac{d}{dx} \ln x = \dfrac{1}{x}$

4. $\dfrac{d}{dx} \log_a x = \dfrac{1}{x \ln a}$

5. $\displaystyle\int e^x \, dx = e^x + C$

6. $\displaystyle\int a^x \, dx = \dfrac{a^x}{\ln a} + C$

7. $\displaystyle\int \dfrac{dx}{x} = \ln |x| + C$

Example 3 **Finding Antiderivatives**

Find the following antiderivatives.

$$1. \int 3^x \, dx \qquad 2. \int x^2 e^{x^3} \, dx \qquad 3. \int \frac{x \, dx}{x^2 + 1}$$

Notes:

SOLUTION

1. Applying our theorem,

$$\int 3^x \, dx = \frac{3^x}{\ln 3} + C$$

2. We use the substitution $u = x^3$, $du = 3x^2 \, dx$:

$$\int x^2 e^{x^3} \, dx = \frac{1}{3} \int e^u \, du$$

$$= \frac{1}{3} e^u + C$$

$$= \frac{1}{3} e^{x^3} + C$$

3. Using the substitution $u = x^2 + 1$, $du = 2x \, dx$:

$$\int \frac{x \, dx}{x^2 + 1} = \frac{1}{2} \int \frac{du}{u}$$

$$= \frac{1}{2} \ln |u| + C$$

$$= \frac{1}{2} \ln \left| x^2 + 1 \right| + C$$

$$= \frac{1}{2} \ln(x^2 + 1) + C$$

Note that we do not yet have an antiderivative for the function $f(x) = \ln x$. We remedy this in Section 8.1 with Example 8.1.5.

Logarithmic Differentiation

Consider the function $y = x^x$; it is graphed in Figure 7.12. It is well–defined for $x > 0$ and we might be interested in finding equations of lines tangent and normal to its graph. How do we take its derivative?

The function is not a power function: it has a "power" of x, not a constant. It is not an exponential function: it has a "base" of x, not a constant.

A differentiation technique known as *logarithmic differentiation* becomes useful here. The basic principle is this: take the natural log of both sides of an equation $y = f(x)$, then use implicit differentiation to find y'. We demonstrate this in the following example.
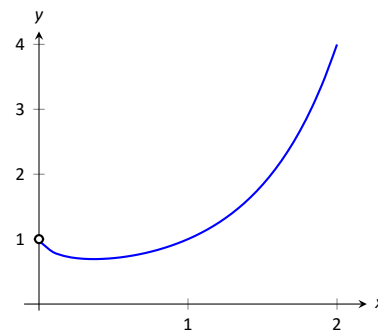
Figure 7.12: A plot of $y = x^x$.

Notes:

Example 4   **Using Logarithmic Differentiation**

Given $y = x^x$, use logarithmic differentiation to find y'.

SOLUTION   As suggested above, we start by taking the natural log of both sides then applying implicit differentiation.

$$y = x^x$$
$$\ln(y) = \ln(x^x) \qquad\qquad \text{(apply logarithm rule)}$$
$$\ln(y) = x \ln x \qquad\qquad \text{(now use implicit differentiation)}$$
$$\frac{d}{dx}\Big(\ln(y)\Big) = \frac{d}{dx}\Big(x \ln x\Big)$$
$$\frac{y'}{y} = \ln x + x \cdot \frac{1}{x}$$
$$\frac{y'}{y} = \ln x + 1$$
$$y' = y\big(\ln x + 1\big) \qquad\qquad \text{(substitute } y = x^x)$$
$$y' = x^x\big(\ln x + 1\big).$$

To "test" our answer, let's use it to find the equation of the tangent line at $x = 1.5$. The point on the graph our tangent line must pass through is $(1.5, 1.5^{1.5}) \approx (1.5, 1.837)$. Using the equation for y', we find the slope as

$$y' = 1.5^{1.5}\big(\ln 1.5 + 1\big) \approx 1.837(1.405) \approx 2.582.$$

Thus the equation of the tangent line is $y = 1.6833(x - 1.5) + 1.837$. Figure 7.13 graphs $y = x^x$ along with this tangent line.
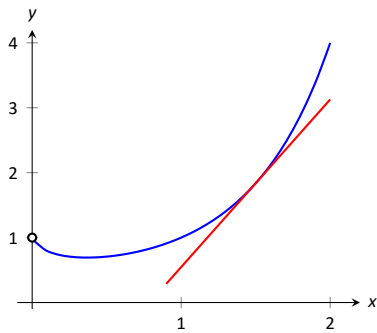
Figure 7.13: A graph of $y = x^x$ and its tangent line at $x = 1.5$.

Notes:

Exercises 7.3

Problems

In Exercises 1–4, find the domain of the function.

1. $f(x) = e^{x^2+1}$

2. $f(t) = \ln(1 - x^2)$

3. $g(x) = \ln(x^2)$

4. $f(x) = \dfrac{2}{\log_3(x^2 + 1)}$

In Exercises 5–13, find the derivative of the function.

5. $f(t) = e^{t^3-1}$

6. $g(r) = x^2 \log_2 x$

7. $f(x) = \dfrac{\log_5 x}{5^x}$

8. $f(x) = 4^{x^5}$

9. $f(x) = 7^{\log_7 x}$

10. $g(x) = e^{x^2} \sin(x - \ln x)$

11. $h(r) = \tan^{-1}(3^r)$

12. $h(x) = \log_{10}\left(\dfrac{x^2 + 1}{x^4}\right)$

13. $f(t) = \ln te^t$

14. Find the two values of n so that the function $y = e^{nx}$ satisfies the differential equation $y'' + y' - 6y = 0$.

In Exercises 15–23, evaluate the integral.

15. $\displaystyle\int_0^2 5^x \, dx$

16. $\displaystyle\int_1^3 \dfrac{\log_3 x}{x} \, dx$

17. $\displaystyle\int x3^{x^2-1} \, dx$

18. $\displaystyle\int \dfrac{\cos(\ln x)}{x} \, dx$

19. $\displaystyle\int e^x \sin(e^x) \cos(e^x) \, dx$

20. $\displaystyle\int_1^8 \log_2 x \, dx$

21. $\displaystyle\int_0^5 \dfrac{3^x}{3^x + 2} \, dx$

22. $\displaystyle\int \dfrac{1 - t}{1 + t^2} \, dt$

23. $\displaystyle\int \dfrac{1}{(1 + x^2)\tan^{-1} x} \, dx$

24. Let $f(x) = x^2$ and $g(x) = 2^x$.

   (a) Since $f(2) = 2^2 = 4$ and $g(2) = 2^2 = 4$, $f(2) = g(2)$. Find a positive number $c > 2$ so that $f(c) = g(c)$.

   (b) Explain how you can be sure that there is at least one negative number a so that $f(a) = g(a)$.

   (c) Use the Bisection Method to estimate the number a accurate to within .05.

   (d) Assume you were to graph $f(x)$ and $g(x)$ on the same graph with unit length equal to 1 inch along both coordinate axes. Approximately how high is the graph of f when $x = 18$? The graph of g?

In Exercises 25–32, use logarithmic differentiation to find $\dfrac{dy}{dx}$, then find the equation of the tangent line at the indicated x–value.

25. $y = (1 + x)^{1/x}$, $x = 1$

26. $y = (2x)^{x^2}$, $x = 1$

27. $y = \dfrac{x^x}{x + 1}$, $x = 1$

28. $y = x^{\sin(x)+2}$, $x = \pi/2$

29. $y = \dfrac{x + 1}{x + 2}$, $x = 1$

30. $y = \dfrac{(x + 1)(x + 2)}{(x + 3)(x + 4)}$, $x = 0$

31. $y = x^{e^x}, x = 1$

32. $y = (\cot x)^{\cos x}, x = \pi$

7.4   Hyperbolic Functions

The **hyperbolic functions** are functions that have many applications to mathematics, physics, and engineering. Among many other applications, they are used to describe the formation of satellite rings around planets, to describe the shape of a rope hanging from two points, and have application to the theory of special relativity. This section defines the hyperbolic functions and describes many of their properties, especially their usefulness to calculus.

These functions are sometimes referred to as the "hyperbolic trigonometric functions" as there are many connections between them and the standard trigonometric functions. Figure 7.14 demonstrates one such connection. Just as cosine and sine are used to define points on the circle defined by $x^2 + y^2 = 1$, the functions **hyperbolic cosine** and **hyperbolic sine** are used to define points on the hyperbola $x^2 - y^2 = 1$.

We begin with their definitions.
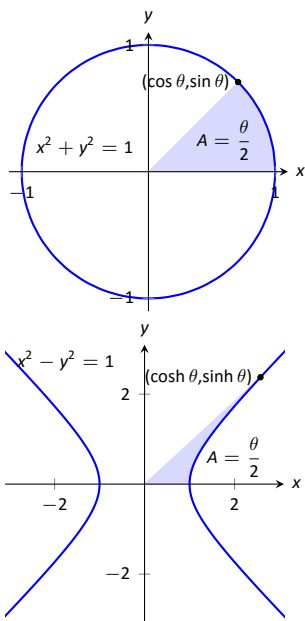
Figure 7.14: Using trigonometric functions to define points on a circle and hyperbolic functions to define points on a hyperbola.

Definition 27   Hyperbolic Functions

1. $\cosh x = \dfrac{e^x + e^{-x}}{2}$ 4. $\operatorname{sech} x = \dfrac{1}{\cosh x}$

2. $\sinh x = \dfrac{e^x - e^{-x}}{2}$ 5. $\operatorname{csch} x = \dfrac{1}{\sinh x}$

3. $\tanh x = \dfrac{\sinh x}{\cosh x}$ 6. $\coth x = \dfrac{\cosh x}{\sinh x}$

The hyperbolic functions are graphed in Figure 7.15. In the graphs of $\cosh x$ and $\sinh x$, graphs of $e^x/2$ and $e^{-x}/2$ are included with dashed lines. As x gets "large," $\cosh x$ and $\sinh x$ each act like $e^x/2$; when x is a large negative number, $\cosh x$ acts like $e^{-x}/2$ whereas $\sinh x$ acts like $-e^{-x}/2$.

Notice the domains of $\tanh x$ and $\operatorname{sech} x$ are $(-\infty, \infty)$, whereas both $\coth x$ and $\operatorname{csch} x$ have vertical asymptotes at $x = 0$. Also note the ranges of these functions, especially $\tanh x$: as $x \to \infty$, both $\sinh x$ and $\cosh x$ approach $e^{-x}/2$, hence $\tanh x$ approaches 1.

Pronunciation Note:
"cosh" rhymes with "gosh,"
"sinh" rhymes with "pinch," and
"tanh" rhymes with "ranch,"

Notes:

$f(x) = \cosh x$

$f(x) = \sinh x$

$f(x) = \coth x$

$f(x) = \tanh x$

$f(x) = \operatorname{sech} x$

$f(x) = \operatorname{csch} x$
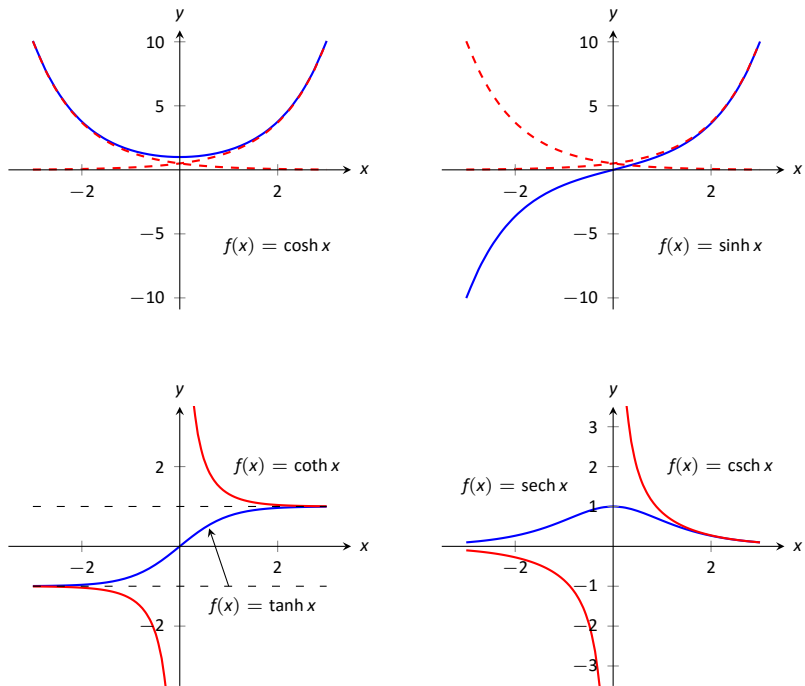
Figure 7.15: Graphs of the hyperbolic functions.

Watch the video:
Hyperbolic Functions — The Basics at
`https://youtu.be/G1C1Z5aTZSQ`

The following example explores some of the properties of these functions that bear remarkable resemblance to the properties of their trigonometric counterparts.

Example 1   **Exploring properties of hyperbolic functions**
Use Definition 27 to rewrite the following expressions.

Notes:

1. $\cosh^2 x - \sinh^2 x$

2. $\tanh^2 x + \text{sech}^2 x$

3. $2 \cosh x \sinh x$

4. $\frac{d}{dx}\left(\cosh x\right)$

5. $\frac{d}{dx}\left(\sinh x\right)$

6. $\frac{d}{dx}\left(\tanh x\right)$

SOLUTION

1.
$$\cosh^2 x - \sinh^2 x = \left(\frac{e^x + e^{-x}}{2}\right)^2 - \left(\frac{e^x - e^{-x}}{2}\right)^2$$
$$= \frac{e^{2x} + 2e^x e^{-x} + e^{-2x}}{4} - \frac{e^{2x} - 2e^x e^{-x} + e^{-2x}}{4}$$
$$= \frac{4}{4} = 1.$$

So $\cosh^2 x - \sinh^2 x = 1$.

2.
$$\tanh^2 x + \text{sech}^2 x = \frac{\sinh^2 x}{\cosh^2 x} + \frac{1}{\cosh^2 x}$$
$$= \frac{\sinh^2 x + 1}{\cosh^2 x} \qquad \text{Now use identity from \#1.}$$
$$= \frac{\cosh^2 x}{\cosh^2 x} = 1.$$

So $\tanh^2 x + \text{sech}^2 x = 1$.

3.
$$2 \cosh x \sinh x = 2 \left(\frac{e^x + e^{-x}}{2}\right)\left(\frac{e^x - e^{-x}}{2}\right)$$
$$= 2 \cdot \frac{e^{2x} - e^{-2x}}{4}$$
$$= \frac{e^{2x} - e^{-2x}}{2} = \sinh(2x).$$

Thus $2 \cosh x \sinh x = \sinh(2x)$.

4.
$$\frac{d}{dx}\left(\cosh x\right) = \frac{d}{dx}\left(\frac{e^x + e^{-x}}{2}\right)$$
$$= \frac{e^x - e^{-x}}{2}$$
$$= \sinh x.$$

So $\frac{d}{dx}\left(\cosh x\right) = \sinh x$.

Notes:

362

5.
$$\frac{d}{dx}\left(\sinh x\right) = \frac{d}{dx}\left(\frac{e^x - e^{-x}}{2}\right)$$
$$= \frac{e^x + e^{-x}}{2}$$
$$= \cosh x.$$

So $\frac{d}{dx}\left(\sinh x\right) = \cosh x$.

6.
$$\frac{d}{dx}\left(\tanh x\right) = \frac{d}{dx}\left(\frac{\sinh x}{\cosh x}\right)$$
$$= \frac{\cosh x \cosh x - \sinh x \sinh x}{\cosh^2 x}$$
$$= \frac{1}{\cosh^2 x}$$
$$= \operatorname{sech}^2 x.$$

So $\frac{d}{dx}\left(\tanh x\right) = \operatorname{sech}^2 x$.

The following Key Idea summarizes many of the important identities relating to hyperbolic functions. Each can be verified by referring back to Definition 27.

Key Idea 21   Useful Hyperbolic Function Properties

Basic Identities

1. $\cosh^2 x - \sinh^2 x = 1$
2. $\tanh^2 x + \operatorname{sech}^2 x = 1$
3. $\coth^2 x - \operatorname{csch}^2 x = 1$
4. $\cosh 2x = \cosh^2 x + \sinh^2 x$
5. $\sinh 2x = 2 \sinh x \cosh x$
6. $\cosh^2 x = \dfrac{\cosh 2x + 1}{2}$
7. $\sinh^2 x = \dfrac{\cosh 2x - 1}{2}$

Derivatives

1. $\frac{d}{dx}\left(\cosh x\right) = \sinh x$
2. $\frac{d}{dx}\left(\sinh x\right) = \cosh x$
3. $\frac{d}{dx}\left(\tanh x\right) = \operatorname{sech}^2 x$
4. $\frac{d}{dx}\left(\operatorname{sech} x\right) = -\operatorname{sech} x \tanh x$
5. $\frac{d}{dx}\left(\operatorname{csch} x\right) = -\operatorname{csch} x \coth x$
6. $\frac{d}{dx}\left(\coth x\right) = -\operatorname{csch}^2 x$

Integrals

1. $\displaystyle\int \cosh x \, dx = \sinh x + C$
2. $\displaystyle\int \sinh x \, dx = \cosh x + C$
3. $\displaystyle\int \tanh x \, dx = \ln(\cosh x) + C$
4. $\displaystyle\int \coth x \, dx = \ln|\sinh x| + C$

We practice using Key Idea 21.

Notes:

Example 2   **Derivatives and integrals of hyperbolic functions**
Evaluate the following derivatives and integrals.

1. $\dfrac{d}{dx}\left(\cosh 2x\right)$

2. $\displaystyle\int \operatorname{sech}^2(7t-3)\, dt$

3. $\displaystyle\int_0^{\ln 2} \cosh x \, dx$

SOLUTION

1. Using the Chain Rule directly, we have $\frac{d}{dx}\left(\cosh 2x\right) = 2\sinh 2x$.

   Just to demonstrate that it works, let's also use the Basic Identity found in Key Idea 21: $\cosh 2x = \cosh^2 x + \sinh^2 x$.

   $$\frac{d}{dx}\left(\cosh 2x\right) = \frac{d}{dx}\left(\cosh^2 x + \sinh^2 x\right) = 2\cosh x \sinh x + 2\sinh x \cosh x$$
   $$= 4\cosh x \sinh x.$$

   Using another Basic Identity, we can see that $4\cosh x \sinh x = 2\sinh 2x$. We get the same answer either way.

2. We employ substitution, with $u = 7t - 3$ and $du = 7dt$. Applying Key Idea 21 we have:

   $$\int \operatorname{sech}^2(7t-3)\, dt = \frac{1}{7}\tanh(7t-3) + C.$$

3.
   $$\int_0^{\ln 2} \cosh x \, dx = \sinh x \Big|_0^{\ln 2} = \sinh(\ln 2) - \sinh 0 = \sinh(\ln 2).$$

   We can simplify this last expression as $\sinh x$ is based on exponentials:

   $$\sinh(\ln 2) = \frac{e^{\ln 2} - e^{-\ln 2}}{2} = \frac{2 - 1/2}{2} = \frac{3}{4}.$$

Inverse Hyperbolic Functions

Just as the inverse trigonometric functions are useful in certain integrations, the inverse hyperbolic functions are useful with others. Figure 7.16 shows the restrictions on the domains to make each function one-to-one and the resulting domains and ranges of their inverse functions. Their graphs are shown in Figure 7.17.

Notes:

Because the hyperbolic functions are defined in terms of exponential functions, their inverses can be expressed in terms of logarithms as shown in Key Idea 22. It is often more convenient to refer to $\sinh^{-1} x$ than to $\ln\left(x + \sqrt{x^2 + 1}\right)$, especially when one is working on theory and does not need to compute actual values. On the other hand, when computations are needed, technology is often helpful but many hand-held calculators lack a *convenient* $\sinh^{-1} x$ button. (Often it can be accessed under a menu system, but not conveniently.) In such a situation, the logarithmic representation is useful. The reader is not encouraged to memorize these, but rather know they exist and know how to use them when needed.

Function	Domain	Range	Function	Domain	Range
$\cosh x$	$[0, \infty)$	$[1, \infty)$	$\cosh^{-1} x$	$[1, \infty)$	$[0, \infty)$
$\sinh x$	$(-\infty, \infty)$	$(-\infty, \infty)$	$\sinh^{-1} x$	$(-\infty, \infty)$	$(-\infty, \infty)$
$\tanh x$	$(-\infty, \infty)$	$(-1, 1)$	$\tanh^{-1} x$	$(-1, 1)$	$(-\infty, \infty)$
$\text{sech}\, x$	$[0, \infty)$	$(0, 1]$	$\text{sech}^{-1} x$	$(0, 1]$	$[0, \infty)$
$\text{csch}\, x$	$(-\infty, 0) \cup (0, \infty)$	$(-\infty, 0) \cup (0, \infty)$	$\text{csch}^{-1} x$	$(-\infty, 0) \cup (0, \infty)$	$(-\infty, 0) \cup (0, \infty)$
$\coth x$	$(-\infty, 0) \cup (0, \infty)$	$(-\infty, -1) \cup (1, \infty)$	$\coth^{-1} x$	$(-\infty, -1) \cup (1, \infty)$	$(-\infty, 0) \cup (0, \infty)$

Figure 7.16: Domains and ranges of the hyperbolic and inverse hyperbolic functions.
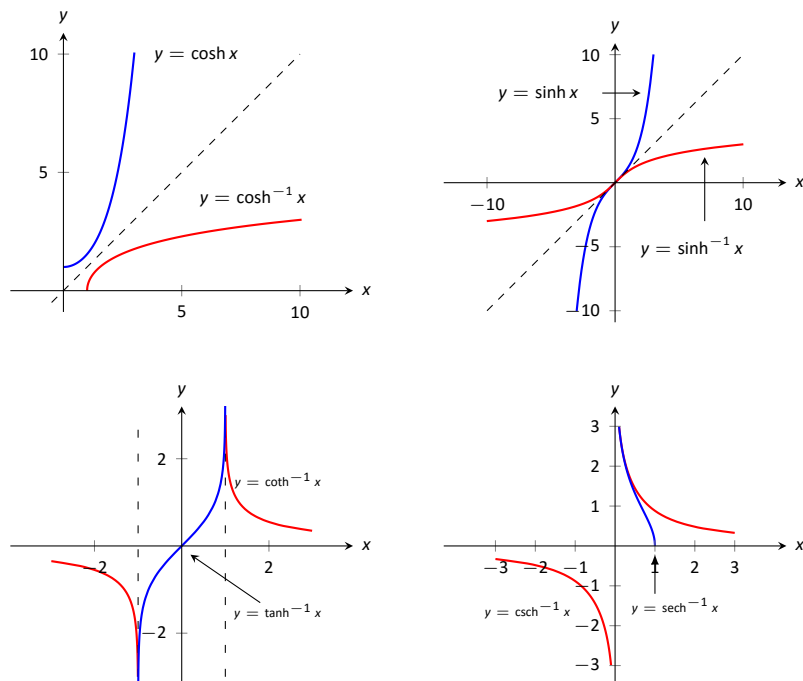
Figure 7.17: Graphs of the hyperbolic functions and their inverses.

Now let's consider the inverses of the hyperbolic functions. We begin with the function $f(x) = \sinh x$. Since $f'(x) = \cosh x > 0$ for all real x, f is increasing

---

Notes:

and must be one-to-one. We proceed as in Section 7.1:

$$y = \frac{e^x - e^{-x}}{2}$$
$$2y = e^x - e^{-x} \qquad \text{(now multiply by } e^x)$$
$$2ye^x = e^{2x} - 1 \qquad \text{(a quadratic form)}$$
$$\left(e^x\right)^2 - 2ye^x - 1 = 0 \qquad \text{(use the quadratic formula)}$$
$$e^x = \frac{2y \pm \sqrt{4y^2 - 4}}{2}$$
$$e^x = y \pm \sqrt{y^2 + 1} \qquad \text{(use the fact that } e^x > 0)$$
$$e^x = y + \sqrt{y^2 + 1}$$
$$x = \ln(y + \sqrt{y^2 + 1})$$

Finally, interchange the variable to find that

$$\sinh^{-1} x = \ln(x + \sqrt{x^2 + 1}).$$

In a similar manner we find that the inverses of the other hyperbolic functions are given by:

Key Idea 22   Logarithmic definitions of Inverse Hyperbolic Functions

1. $\cosh^{-1} x = \ln\left(x + \sqrt{x^2 - 1}\right); \ x \geq 1$ 4. $\sinh^{-1} x = \ln\left(x + \sqrt{x^2 + 1}\right)$

2. $\tanh^{-1} x = \dfrac{1}{2} \ln\left(\dfrac{1 + x}{1 - x}\right); \ |x| < 1$ 5. $\coth^{-1} x = \dfrac{1}{2} \ln\left(\dfrac{x + 1}{x - 1}\right); \ |x| > 1$

3. $\operatorname{sech}^{-1} x = \ln\left(\dfrac{1 + \sqrt{1 - x^2}}{x}\right); \ 0 < x \leq 1$ 6. $\operatorname{csch}^{-1} x = \ln\left(\dfrac{1}{x} + \dfrac{\sqrt{1 + x^2}}{|x|}\right); \ x \neq 0$

The following Key Ideas give the derivatives and integrals relating to the inverse hyperbolic functions. In Key Idea 24, both the inverse hyperbolic and logarithmic function representations of the antiderivative are given, based on Key Idea 22. Again, these latter functions are often more useful than the former.

Notes:

Key Idea 23   Derivatives Involving Inverse Hyperbolic Functions

1. $\dfrac{d}{dx}\left(\cosh^{-1}x\right)=\dfrac{1}{\sqrt{x^2-1}}$; $x>1$

2. $\dfrac{d}{dx}\left(\sinh^{-1}x\right)=\dfrac{1}{\sqrt{x^2+1}}$

3. $\dfrac{d}{dx}\left(\tanh^{-1}x\right)=\dfrac{1}{1-x^2}$; $|x|<1$

4. $\dfrac{d}{dx}\left(\operatorname{sech}^{-1}x\right)=\dfrac{-1}{x\sqrt{1-x^2}}$; $0<x<1$

5. $\dfrac{d}{dx}\left(\operatorname{csch}^{-1}x\right)=\dfrac{-1}{|x|\sqrt{1+x^2}}$; $x\neq0$

6. $\dfrac{d}{dx}\left(\coth^{-1}x\right)=\dfrac{1}{1-x^2}$; $|x|>1$

Key Idea 24   Integrals Involving Inverse Hyperbolic Functions

1. $\displaystyle\int\dfrac{1}{\sqrt{x^2-a^2}}\,dx = \cosh^{-1}\left(\dfrac{x}{a}\right)+C$; $0<a<x$ $= \ln\left|x+\sqrt{x^2-a^2}\right|+C$

2. $\displaystyle\int\dfrac{1}{\sqrt{x^2+a^2}}\,dx = \sinh^{-1}\left(\dfrac{x}{a}\right)+C$; $a>0$ $= \ln\left|x+\sqrt{x^2+a^2}\right|+C$

3. $\displaystyle\int\dfrac{1}{a^2-x^2}\,dx = \begin{cases}\frac{1}{a}\tanh^{-1}\left(\frac{x}{a}\right)+C & |x|<|a| \\ \frac{1}{a}\coth^{-1}\left(\frac{x}{a}\right)+C & |a|<|x|\end{cases}$ $= \dfrac{1}{2a}\ln\left|\dfrac{a+x}{a-x}\right|+C$

4. $\displaystyle\int\dfrac{1}{x\sqrt{a^2-x^2}}\,dx = -\dfrac{1}{a}\operatorname{sech}^{-1}\left(\dfrac{x}{a}\right)+C$; $0<x<a$ $= \dfrac{1}{a}\ln\left(\dfrac{x}{a+\sqrt{a^2-x^2}}\right)+C$

5. $\displaystyle\int\dfrac{1}{x\sqrt{x^2+a^2}}\,dx = -\dfrac{1}{a}\operatorname{csch}^{-1}\left|\dfrac{x}{a}\right|+C$; $x\neq0$, $a>0$ $= \dfrac{1}{a}\ln\left|\dfrac{x}{a+\sqrt{a^2+x^2}}\right|+C$

We practice using the derivative and integral formulas in the following example.

Example 3   Derivatives and integrals involving inverse hyperbolic functions

Evaluate the following.

Notes:

1. $\dfrac{d}{dx}\left[\cosh^{-1}\left(\dfrac{3x-2}{5}\right)\right]$

3. $\displaystyle\int \dfrac{1}{\sqrt{9x^2+10}}\,dx$

2. $\displaystyle\int \dfrac{1}{x^2-1}\,dx$

SOLUTION

1. Applying Key Idea 23 with the Chain Rule gives:

$$\dfrac{d}{dx}\left[\cosh^{-1}\left(\dfrac{3x-2}{5}\right)\right] = \dfrac{1}{\sqrt{\left(\frac{3x-2}{5}\right)^2-1}}\cdot\dfrac{3}{5}.$$

2. Multiplying the numerator and denominator by (-1) gives: $\displaystyle\int \dfrac{1}{x^2-1}\,dx =$

$\displaystyle\int \dfrac{-1}{1-x^2}\,dx$. The second integral can be solved with a direct application of item #3 from Key Idea 24, with $a = 1$. Thus

$$\int \dfrac{1}{x^2-1}\,dx = -\int \dfrac{1}{1-x^2}\,dx$$

$$= \begin{cases} -\tanh^{-1}(x) + C & x^2 < 1 \\ -\coth^{-1}(x) + C & 1 < x^2 \end{cases}$$

$$= -\dfrac{1}{2}\ln\left|\dfrac{x+1}{x-1}\right| + C$$

$$= \dfrac{1}{2}\ln\left|\dfrac{x-1}{x+1}\right| + C. \qquad\qquad (7.1)$$

3. This requires a substitution, then item #2 of Key Idea 24 can be applied. Let $u = 3x$, hence $du = 3dx$. We have

$$\int \dfrac{1}{\sqrt{9x^2+10}}\,dx = \dfrac{1}{3}\int \dfrac{1}{\sqrt{u^2+10}}\,du.$$

Note $a^2 = 10$, hence $a = \sqrt{10}$. Now apply the integral rule.

$$= \dfrac{1}{3}\sinh^{-1}\left(\dfrac{3x}{\sqrt{10}}\right) + C$$

$$= \dfrac{1}{3}\ln\left|3x + \sqrt{9x^2+10}\right| + C.$$

Notes:

This section covers a lot of ground. New functions were introduced, along with some of their fundamental identities, their derivatives and antiderivatives, their inverses, and the derivatives and antiderivatives of these inverses. Four Key Ideas were presented, each including quite a bit of information.

Do not view this section as containing a source of information to be memorized, but rather as a reference for future problem solving. Key Idea 24 contains perhaps the most useful information. Know the integration forms it helps evaluate and understand how to use the inverse hyperbolic answer and the logarithmic answer.

The next section takes a brief break from demonstrating new integration techniques. It instead demonstrates a technique of evaluating limits that return indeterminate forms. This technique will be useful in Section 8.6, where limits will arise in the evaluation of certain definite integrals.

Notes:

Exercises 7.4

Terms and Concepts

1. In Key Idea 21, the equation $\int \tanh x \, dx = \ln(\cosh x) + C$ is given. Why is "$\ln|\cosh x|$" not used — i.e., why are absolute values not necessary?

2. The hyperbolic functions are used to define points on the right hand portion of the hyperbola $x^2 - y^2 = 1$, as shown in Figure 7.14. How can we use the hyperbolic functions to define points on the left hand portion of the hyperbola?

Problems

In Exercises 3–10, verify the given identity using Definition 27, as done in Example 1.

3. $\coth^2 x - \csch^2 x = 1$

4. $\cosh 2x = \cosh^2 x + \sinh^2 x$

5. $\cosh^2 x = \dfrac{\cosh 2x + 1}{2}$

6. $\sinh^2 x = \dfrac{\cosh 2x - 1}{2}$

7. $\dfrac{d}{dx}[\sech x] = -\sech x \tanh x$

8. $\dfrac{d}{dx}[\coth x] = -\csch^2 x$

9. $\int \tanh x \, dx = \ln(\cosh x) + C$

10. $\int \coth x \, dx = \ln|\sinh x| + C$

In Exercises 11–21, find the derivative of the given function.

11. $f(x) = \cosh 2x$

12. $f(x) = \tanh(x^2)$

13. $f(x) = \ln(\sinh x)$

14. $f(x) = \sinh x \cosh x$

15. $f(x) = x \sinh x - \cosh x$

16. $f(x) = \sech^{-1}(x^2)$

17. $f(x) = \sinh^{-1}(3x)$

18. $f(x) = \cosh^{-1}(2x^2)$

19. $f(x) = \tanh^{-1}(x + 5)$

20. $f(x) = \tanh^{-1}(\cos x)$

21. $f(x) = \cosh^{-1}(\sec x)$

In Exercises 22–26, find the equation of the line tangent to the function at the given x-value.

22. $f(x) = \sinh x$ at $x = 0$

23. $f(x) = \cosh x$ at $x = \ln 2$

24. $f(x) = \sech^2 x$ at $x = \ln 3$

25. $f(x) = \sinh^{-1} x$ at $x = 0$

26. $f(x) = \cosh^{-1} x$ at $x = \sqrt{2}$

In Exercises 27–34, evaluate the given indefinite integral.

27. $\int \tanh(2x) \, dx$

28. $\int \cosh(3x - 7) \, dx$

29. $\int \sinh x \cosh x \, dx$

30. $\int \dfrac{1}{9 - x^2} \, dx$

31. $\int \dfrac{2x}{\sqrt{x^4 - 4}} \, dx$

32. $\int \dfrac{\sqrt{x}}{\sqrt{1 + x^3}} \, dx$

33. $\int \dfrac{e^x}{e^{2x} + 1} \, dx$

34. $\int \sech x \, dx$ (Hint: multiply by $\frac{\cosh x}{\cosh x}$; set $u = \sinh x$.)

In Exercises 35–36, evaluate the given definite integral.

35. $\displaystyle\int_{-1}^{1} \sinh x \, dx$

36. $\displaystyle\int_{-\ln 2}^{\ln 2} \cosh x \, dx$

7.5 L'Hôpital's Rule

This section is concerned with a technique for evaluating certain limits that will be useful in later chapters.

Our treatment of limits exposed us to "0/0", an indeterminate form. If $\lim\limits_{x \to c} f(x) = 0$ and $\lim\limits_{x \to c} g(x) = 0$, we do not conclude that $\lim\limits_{x \to c} f(x)/g(x)$ is 0/0; rather, we use 0/0 as notation to describe the fact that both the numerator and denominator approach 0. The expression 0/0 has no numeric value; other work must be done to evaluate the limit.

Other indeterminate forms exist; they are: ∞/∞, $0 \cdot \infty$, $\infty - \infty$, 0^0, 1^∞ and ∞^0. Just as "0/0" does not mean "divide 0 by 0," the expression "∞/∞" does not mean "divide infinity by infinity." Instead, it means "a quantity is growing without bound and is being divided by another quantity that is growing without bound." We cannot determine from such a statement what value, if any, results in the limit. Likewise, "$0 \cdot \infty$" does not mean "multiply zero by infinity." Instead, it means "one quantity is shrinking to zero, and is being multiplied by a quantity that is growing without bound." We cannot determine from such a description what the result of such a limit will be.

This section introduces L'Hôpital's Rule, a method of resolving limits that produce the indeterminate forms 0/0 and ∞/∞. We'll also show how algebraic manipulation can be used to convert other indeterminate expressions into one of these two forms so that our new rule can be applied.

Theorem 52   L'Hôpital's Rule, Part 1

Let f and g be differentiable functions on an open interval I containing a.

1. If $\lim\limits_{x \to a} f(x) = 0$, $\lim\limits_{x \to a} g(x) = 0$, and $g'(x) \neq 0$ except possibly at $x = a$, then

$$\lim\limits_{x \to a} \frac{f(x)}{g(x)} = \lim\limits_{x \to a} \frac{f'(x)}{g'(x)}.$$

2. If $\lim\limits_{x \to a} f(x) = \pm\infty$ and $\lim\limits_{x \to a} g(x) = \pm\infty$, then

$$\lim\limits_{x \to a} \frac{f(x)}{g(x)} = \lim\limits_{x \to a} \frac{f'(x)}{g'(x)}.$$

A similar statement holds if we just look at the one sided limits $\lim\limits_{x \to a^-}$ and $\lim\limits_{x \to a^+}$.

Notes:

> **Theorem 53   L'Hôpital's Rule, Part 2**
>
> Let f and g be differentiable functions on the open interval (c, ∞) for some value c and $g'(x) \neq 0$ on (c, ∞).
>
> 1. If $\lim\limits_{x \to \infty} f(x) = 0$ and $\lim\limits_{x \to \infty} g(x) = 0$, then
>
> $$\lim_{x \to \infty} \frac{f(x)}{g(x)} = \lim_{x \to \infty} \frac{f'(x)}{g'(x)}.$$
>
> 2. If $\lim\limits_{x \to \infty} f(x) = \pm\infty$ and $\lim\limits_{x \to \infty} g(x) = \pm\infty$, then
>
> $$\lim_{x \to \infty} \frac{f(x)}{g(x)} = \lim_{x \to \infty} \frac{f'(x)}{g'(x)}.$$
>
> Similar statements can be made where x approaches $-\infty$.

We demonstrate the use of L'Hôpital's Rule in the following examples; we will often use "LHR" as an abbreviation of "L'Hôpital's Rule."

Example 1   Using L'Hôpital's Rule

Evaluate the following limits, using L'Hôpital's Rule as needed.

1. $\lim\limits_{x \to 0} \dfrac{\sin x}{x}$

2. $\lim\limits_{x \to 1} \dfrac{\sqrt{x+3} - 2}{1 - x}$

3. $\lim\limits_{x \to 0} \dfrac{x^2}{1 - \cos x}$

4. $\lim\limits_{x \to -3} \dfrac{x^3 + 27}{x^2 + 9}$

5. $\lim\limits_{x \to \infty} \dfrac{3x^2 - 100x + 2}{4x^2 + 5x - 1000}$

6. $\lim\limits_{x \to \infty} \dfrac{e^x}{x^3}$

SOLUTION

1. This has the indeterminate form $0/0$. We proved this limit is 1 in Example 1.3.4 using the Squeeze Theorem. Here we use L'Hôpital's Rule to show its power.

$$\lim_{x \to 0} \frac{\sin x}{x} \overset{\text{by LHR}}{=} \lim_{x \to 0} \frac{\cos x}{1} = 1.$$

While this seems easier than using the Squeeze Theorem to find this limit, we note that applying L'Hôpital's Rule here requires us to know the derivative of $\sin x$. We originally encountered this limit when we were trying to find that derivative.

Notes:

2. This has the indeterminate form $0/0$.

$$\lim_{x\to 1}\frac{\sqrt{x+3}-2}{1-x} \overset{\text{by LHR}}{=} \lim_{x\to 1}\frac{\frac{1}{2}(x+3)^{-1/2}}{-1} = -\frac{1}{4}.$$

3. This has the indeterminate form $0/0$.

$$\lim_{x\to 0}\frac{x^2}{1-\cos x} \overset{\text{by LHR}}{=} \lim_{x\to 0}\frac{2x}{\sin x}.$$

This latter limit also evaluates to the $0/0$ indeterminate form. To evaluate it, we apply L'Hôpital's Rule again.

$$\lim_{x\to 0}\frac{2x}{\sin x} \overset{\text{by LHR}}{=} \frac{2}{\cos x} = 2.$$

Thus $\displaystyle\lim_{x\to 0}\frac{x^2}{1-\cos x} = 2$.

4.
$$\lim_{x\to -3}\frac{x^3+27}{x^2+9} = \frac{0}{18} = 0$$

We cannot use L'Hôpital's Rule in this case because the original limit does not return an indeterminate form, so L'Hôpital's Rule does not apply. In fact, the inappropriate use of L'Hôpital's Rule here would result in the incorrect limit $\frac{3}{2}$.

5. We can evaluate this limit already using Key Idea 2; the answer is 3/4. We apply L'Hôpital's Rule to demonstrate its applicability.

$$\lim_{x\to\infty}\frac{3x^2-100x+2}{4x^2+5x-1000} \overset{\text{by LHR}}{=} \lim_{x\to\infty}\frac{6x-100}{8x+5} \overset{\text{by LHR}}{=} \lim_{x\to\infty}\frac{6}{8} = \frac{3}{4}.$$

6. $\displaystyle\lim_{x\to\infty}\frac{e^x}{x^3} \overset{\text{by LHR}}{=} \lim_{x\to\infty}\frac{e^x}{3x^2} \overset{\text{by LHR}}{=} \lim_{x\to\infty}\frac{e^x}{6x} \overset{\text{by LHR}}{=} \lim_{x\to\infty}\frac{e^x}{6} = \infty.$

Recall that this means that the limit does not exist; as x approaches ∞, the expression e^x/x^3 grows without bound. We can infer from this that e^x grows "faster" than x^3; as x gets large, e^x is far larger than x^3. (This has important implications in computing when considering efficiency of algorithms.)

Notes:

Indeterminate Forms $0 \cdot \infty$ and $\infty - \infty$

L'Hôpital's Rule can only be applied to ratios of functions. When faced with an indeterminate form such as $0 \cdot \infty$ or $\infty - \infty$, we can sometimes apply algebra to rewrite the limit so that L'Hôpital's Rule can be applied. We demonstrate the general idea in the next example.

> Watch the video:
> L'Hôpital's Rule — Indeterminate Powers at
> `https://youtu.be/kEnwac_9lyg`

Example 2   **Applying L'Hôpital's Rule to other indeterminate forms**
Evaluate the following limits.

1. $\displaystyle\lim_{x \to 0^+} x \cdot e^{1/x}$

2. $\displaystyle\lim_{x \to 0^-} x \cdot e^{1/x}$

3. $\displaystyle\lim_{x \to \infty} \left(\ln(x+1) - \ln x \right)$

4. $\displaystyle\lim_{x \to \infty} \left(x^2 - e^x \right)$

SOLUTION

1. As $x \to 0^+$, note that $x \to 0$ and $e^{1/x} \to \infty$. Thus we have the indeterminate form $0 \cdot \infty$. We rewrite the expression $x \cdot e^{1/x}$ as $\dfrac{e^{1/x}}{1/x}$; now, as $x \to 0^+$, we get the indeterminate form ∞/∞ to which L'Hôpital's Rule can be applied.

$$\lim_{x \to 0^+} x \cdot e^{1/x} = \lim_{x \to 0^+} \frac{e^{1/x}}{1/x} \overset{\text{by LHR}}{=} \lim_{x \to 0^+} \frac{(-1/x^2)e^{1/x}}{-1/x^2} = \lim_{x \to 0^+} e^{1/x} = \infty.$$

   Interpretation: $e^{1/x}$ grows "faster" than x shrinks to zero, meaning their product grows without bound.

2. As $x \to 0^-$, note that $x \to 0$ and $e^{1/x} \to e^{-\infty} \to 0$. The the limit evaluates to $0 \cdot 0$ which is not an indeterminate form. We conclude then that

$$\lim_{x \to 0^-} x \cdot e^{1/x} = 0.$$

Notes:

3. This limit initially evaluates to the indeterminate form $\infty - \infty$. By applying a logarithmic rule, we can rewrite the limit as

$$\lim_{x \to \infty} \left(\ln(x+1) - \ln x \right) = \lim_{x \to \infty} \ln \left(\frac{x+1}{x} \right).$$

As $x \to \infty$, the argument of the natural logarithm approaches ∞ / ∞, to which we can apply L'Hôpital's Rule.

$$\lim_{x \to \infty} \frac{x+1}{x} \overset{\text{by LHR}}{=} \lim_{x \to \infty} \frac{1}{1} = 1.$$

Since $x \to \infty$ implies $\dfrac{x+1}{x} \to 1$, it follows that

$$x \to \infty \quad \text{implies} \quad \ln \left(\frac{x+1}{x} \right) \to \ln 1 = 0.$$

Thus

$$\lim_{x \to \infty} \left(\ln(x+1) - \ln x \right) = \lim_{x \to \infty} \ln \left(\frac{x+1}{x} \right) = 0.$$

Interpretation: since this limit evaluates to 0, it means that for large x, there is essentially no difference between $\ln(x+1)$ and $\ln x$; their difference is essentially 0.

4. The limit $\lim_{x \to \infty} \left(x^2 - e^x \right)$ initially returns the indeterminate form $\infty - \infty$.

We can rewrite the expression by factoring out x^2; $x^2 - e^x = x^2 \left(1 - \dfrac{e^x}{x^2} \right)$.

We need to evaluate how e^x / x^2 behaves as $x \to \infty$:

$$\lim_{x \to \infty} \frac{e^x}{x^2} \overset{\text{by LHR}}{=} \lim_{x \to \infty} \frac{e^x}{2x} \overset{\text{by LHR}}{=} \lim_{x \to \infty} \frac{e^x}{2} = \infty.$$

Thus $\lim_{x \to \infty} x^2 (1 - e^x / x^2)$ evaluates to $\infty \cdot (-\infty)$, which is not an indeterminate form; rather, $\infty \cdot (-\infty)$ evaluates to $-\infty$. We conclude that $\lim_{x \to \infty} \left(x^2 - e^x \right) = -\infty$.

Interpretation: as x gets large, the difference between x^2 and e^x grows very large.

Notes:

Indeterminate Forms 0^0, 1^∞ and ∞^0

When faced with a limit that returns one of the indeterminate forms 0^0, 1^∞, or ∞^0, it is often useful to use the natural logarithm to convert to an indeterminate form we already know how to find the limit of, then use the natural exponential function find the original limit. This is possible because the natural logarithm and natural exponential functions are inverses and because they are both continuous. The following Key Idea expresses the concept, which is followed by an example that demonstrates its use.

Key Idea 25   **Evaluating Limits Involving Indeterminate Forms 0^0, 1^∞ and ∞^0**

If $\lim\limits_{x\to c} \ln\big(f(x)\big) = L$,   then $\lim\limits_{x\to c} f(x) = \lim\limits_{x\to c} e^{\ln(f(x))} = e^L$.

Example 3   **Using L'Hôpital's Rule with indeterminate forms involving exponents**

Evaluate the following limits.

$$1. \quad \lim_{x\to\infty} \left(1 + \frac{1}{x}\right)^x \qquad\qquad 2. \quad \lim_{x\to 0^+} x^x.$$

SOLUTION

1. This is equivalent to a special limit given in Theorem 6; these limits have important applications in mathematics and finance. Note that the exponent approaches ∞ while the base approaches 1, leading to the indeterminate form 1^∞. Let $f(x) = (1 + 1/x)^x$; the problem asks to evaluate $\lim\limits_{x\to\infty} f(x)$. Let's first evaluate $\lim\limits_{x\to\infty} \ln\big(f(x)\big)$.

$$\lim_{x\to\infty} \ln\big(f(x)\big) = \lim_{x\to\infty} \ln\left(1 + \frac{1}{x}\right)^x$$
$$= \lim_{x\to\infty} x\ln\left(1 + \frac{1}{x}\right)$$
$$= \lim_{x\to\infty} \frac{\ln\left(1 + \frac{1}{x}\right)}{1/x}$$

Notes:

This produces the indeterminate form 0/0, so we apply L'Hôpital's Rule.

$$= \lim_{x\to\infty} \frac{\frac{1}{1+1/x}\cdot(-1/x^2)}{(-1/x^2)}$$

$$= \lim_{x\to\infty} \frac{1}{1+1/x}$$

$$= 1.$$

Thus $\lim_{x\to\infty} \ln\big(f(x)\big) = 1$. We return to the original limit and apply Key Idea 25.

$$\lim_{x\to\infty} \left(1+\frac{1}{x}\right)^x = \lim_{x\to\infty} f(x) = \lim_{x\to\infty} e^{\ln(f(x))} = e^1 = e.$$

This is another way to determine the value of the number e.

2. This limit leads to the indeterminate form 0^0. Let $f(x) = x^x$ and consider first $\lim_{x\to 0^+} \ln\big(f(x)\big)$.

$$\lim_{x\to 0^+} \ln\big(f(x)\big) = \lim_{x\to 0^+} \ln\big(x^x\big)$$

$$= \lim_{x\to 0^+} x\ln x$$

This produces the indeterminate form $0(-\infty)$, so we rewrite it in order to apply L'Hôpital's Rule.

$$= \lim_{x\to 0^+} \frac{\ln x}{1/x}.$$

This produces the indeterminate form $-\infty/\infty$ so we apply L'Hôpital's Rule.

$$= \lim_{x\to 0^+} \frac{1/x}{-1/x^2}$$

$$= \lim_{x\to 0^+} -x$$

$$= 0.$$

Thus $\lim_{x\to 0^+} \ln\big(f(x)\big) = 0$. We return to the original limit and apply Key Idea 25.

$$\lim_{x\to 0^+} x^x = \lim_{x\to 0^+} f(x) = \lim_{x\to 0^+} e^{\ln(f(x))} = e^0 = 1.$$

This result is supported by the graph of $f(x) = x^x$ given in Figure 7.18.
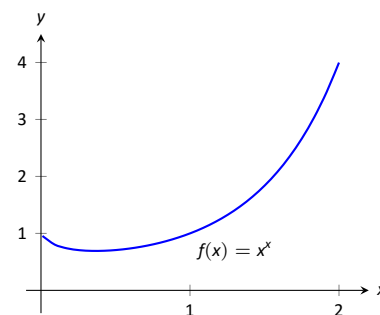
Figure 7.18: A graph of $f(x) = x^x$ supporting the fact that as $x \to 0^+$, $f(x) \to 1$.

Notes:

Exercises 7.5

Terms and Concepts

1. List the different indeterminate forms described in this section.

2. T/F: l'Hôpital's Rule provides a faster method of computing derivatives.

3. T/F: l'Hôpital's Rule states that $\dfrac{d}{dx}\left[\dfrac{f(x)}{g(x)}\right] = \dfrac{f'(x)}{g'(x)}$.

4. Explain what the indeterminate form "1^{∞}" means.

5. Fill in the blanks: The Quotient Rule is applied to $\dfrac{f(x)}{g(x)}$ when taking _____; l'Hôpital's Rule is applied when taking certain _____.

6. Create (but do not evaluate!) a limit that returns "∞^0".

7. Create a function $f(x)$ such that $\lim\limits_{x\to 1} f(x)$ returns "0^0".

Problems

In Exercises 8–52, evaluate the given limit.

8. $\lim\limits_{x\to 1} \dfrac{x^2 + x - 2}{x - 1}$

9. $\lim\limits_{x\to 2} \dfrac{x^2 + x - 6}{x^2 - 7x + 10}$

10. $\lim\limits_{x\to \pi} \dfrac{\sin x}{x - \pi}$

11. $\lim\limits_{x\to \pi/4} \dfrac{\sin x - \cos x}{\cos(2x)}$

12. $\lim\limits_{x\to 0} \dfrac{\sin(5x)}{x}$

13. $\lim\limits_{x\to 0} \dfrac{\sin(2x)}{x + 2}$

14. $\lim\limits_{x\to 0} \dfrac{\sin(2x)}{\sin(3x)}$

15. $\lim\limits_{x\to 0} \dfrac{\sin(ax)}{\sin(bx)}$

16. $\lim\limits_{x\to 0^+} \dfrac{e^x - 1}{x^2}$

17. $\lim\limits_{x\to 0^+} \dfrac{e^x - x - 1}{x^2}$

18. $\lim\limits_{x\to 0^+} \dfrac{x - \sin x}{x^3 - x^2}$

19. $\lim\limits_{x\to \infty} \dfrac{x^4}{e^x}$

20. $\lim\limits_{x\to \infty} \dfrac{\sqrt{x}}{e^x}$

21. $\lim\limits_{x\to \infty} \dfrac{e^x}{\sqrt{x}}$

22. $\lim\limits_{x\to \infty} \dfrac{e^x}{2^x}$

23. $\lim\limits_{x\to \infty} \dfrac{e^x}{3^x}$

24. $\lim\limits_{x\to 3} \dfrac{x^3 - 5x^2 + 3x + 9}{x^3 - 7x^2 + 15x - 9}$

25. $\lim\limits_{x\to -2} \dfrac{x^3 + 4x^2 + 4x}{x^3 + 7x^2 + 16x + 12}$

26. $\lim\limits_{x\to \infty} \dfrac{\ln x}{x}$

27. $\lim\limits_{x\to \infty} \dfrac{\ln(x^2)}{x}$

28. $\lim\limits_{x\to \infty} \dfrac{\left(\ln x\right)^2}{x}$

29. $\lim\limits_{x\to 0^+} x \ln x$

30. $\lim\limits_{x\to 0^+} \sqrt{x} \ln x$

31. $\lim\limits_{x\to 0^+} x e^{1/x}$

32. $\lim\limits_{x\to \infty} \left(x^3 - x^2\right)$

33. $\lim\limits_{x\to \infty} \left(\sqrt{x} - \ln x\right)$

34. $\lim\limits_{x\to -\infty} x e^x$

35. $\lim\limits_{x\to 0^+} \dfrac{1}{x^2} e^{-1/x}$

36. $\lim\limits_{x\to 0^+} (1 + x)^{1/x}$

37. $\lim\limits_{x\to 0^+} (2x)^x$

38. $\lim\limits_{x\to 0^+} (2/x)^x$

39. $\lim\limits_{x\to 0^+} (\sin x)^x$

40. $\lim\limits_{x\to 1^+} (1 - x)^{1-x}$

41. $\lim\limits_{x\to \infty} (x)^{1/x}$

42. $\lim\limits_{x\to \infty} (1/x)^x$

43. $\lim\limits_{x\to 1^+} (\ln x)^{1-x}$

44. $\lim\limits_{x\to \infty} (1 + x)^{1/x}$

45. $\lim\limits_{x\to \infty} (1 + x^2)^{1/x}$

46. $\lim\limits_{x\to \pi/2} \tan x \cos x$

47. $\lim\limits_{x\to \pi/2} \tan x \sin(2x)$

48. $\lim\limits_{x\to 1^+} \left(\dfrac{1}{\ln x} - \dfrac{1}{x - 1}\right)$

49. $\lim\limits_{x\to 3^+} \left(\dfrac{5}{x^2 - 9} - \dfrac{x}{x - 3}\right)$

50. $\lim\limits_{x\to \infty} x \tan(1/x)$

51. $\lim\limits_{x\to \infty} \dfrac{(\ln x)^3}{x}$

52. $\lim\limits_{x\to 1} \dfrac{x^2 + x - 2}{\ln x}$

8: Techniques of Integration

Chapter 5 introduced the antiderivative and connected it to signed areas under a curve through the Fundamental Theorem of Calculus. The chapter after explored more applications of definite integrals than just area. As evaluating definite integrals will become even important, we will want to find antiderivatives of a variety of functions.

This chapter is devoted to exploring techniques of antidifferentiation. While not every function has an antiderivative in terms of elementary functions, we can still find antiderivatives of a wide variety of functions.

8.1 Integration by Parts

Here's a simple integral that we can't yet evaluate:

$$\int x \cos x \, dx.$$

It's a simple matter to take the derivative of the integrand using the Product Rule, but there is no Product Rule for integrals. However, this section introduces *Integration by Parts*, a method of integration that is based on the Product Rule for derivatives. It will enable us to evaluate this integral.

The Product Rule says that if u and v are functions of x, then $(uv)' = u'v + uv'$. For simplicity, we've written u for $u(x)$ and v for $v(x)$. Suppose we integrate both sides with respect to x. This gives

$$\int (uv)' \, dx = \int (u'v + uv') \, dx.$$

By the Fundamental Theorem of Calculus, the left side integrates to uv. The right side can be broken up into two integrals, and we have

$$uv = \int u'v \, dx + \int uv' \, dx.$$

Solving for the second integral we have

$$\int uv' \, dx = uv - \int u'v \, dx.$$

Using differential notation, we can write

$$u' = \frac{du}{dx}$$
$$v' = \frac{dv}{dx} \qquad \Rightarrow \qquad \begin{aligned} du &= u'dx \\ dv &= v'dx. \end{aligned}$$

Thus, the equation above can be written as follows:

$$\int u \, dv = uv - \int v \, du.$$

This is the Integration by Parts formula. For reference purposes, we state this in a theorem.

Theorem 54   Integration by Parts

Let u and v be differentiable functions of x on an interval I containing a and b. Then

$$\int u \, dv = uv - \int v \, du,$$

and applying FTC part 2 we have

$$\int_{x=a}^{x=b} u \, dv = uv \Big|_a^b - \int_{x=a}^{x=b} v \, du.$$

Watch the video:
Integration by Parts — Definite Integral at
`https://youtu.be/zGGI4PkHzhI`

Let's try an example to understand our new technique.

Example 1   Integrating using Integration by Parts

Evaluate $\displaystyle\int x \cos x \, dx$.

SOLUTION   The key to Integration by Parts is to identify part of the integrand as "u" and part as "dv." Regular practice will help one make good identifications, and later we will introduce some principles that help. For now, let $u = x$ and $dv = \cos x \, dx$.

It is generally useful to make a small table of these values.

$$\begin{array}{ll} u = x & dv = \cos x \, dx \\ du = \, ? & v = \, ? \end{array} \quad \Rightarrow \quad \begin{array}{ll} u = x & dv = \cos x \, dx \\ du = dx & v = \sin x \end{array}$$

Notes:

Right now we only know u and dv as shown on the left; on the right we fill in the rest of what we need. If $u = x$, then $du = dx$. Since $dv = \cos x\, dx$, v is an antiderivative of $\cos x$, so $v = \sin x$.

Now substitute all of this into the Integration by Parts formula, giving

$$\int x \cos x\, dx = x \sin x - \int \sin x\, dx.$$

We can then integrate $\sin x$ to get $-\cos x + C$ and overall our answer is

$$\int x \cos x\, dx = x \sin x + \cos x + C.$$

We have two important notes here: (1) notice how the antiderivative contains the product, $x \sin x$. This product is what makes integration by parts necessary. And (2) antidifferentiating dv does result in $v + C$. The intermediate $+C$s are all added together and represented by one $+C$ in the final answer.

The example above demonstrates how Integration by Parts works in general. We try to identify u and dv in the integral we are given, and the key is that we usually want to choose u and dv so that du is simpler than u and v is hopefully not too much more complicated than dv. This will mean that the integral on the right side of the Integration by Parts formula, $\int v\, du$ will be simpler to integrate than the original integral $\int u\, dv$.

In the example above, we chose $u = x$ and $dv = \cos x\, dx$. Then $du = dx$ was simpler than u and $v = \sin x$ is no more complicated than dv. Therefore, instead of integrating $x \cos x\, dx$, we could integrate $\sin x\, dx$, which we knew how to do.

If we had chosen $u = \cos x$ and $dv = x\, dx$, so that $du = -\sin x\, dx$ and $v = \frac{1}{2}x^2$, then

$$\int x \cos x\, dx = \frac{1}{2}x^2 \cos x - \left(-\frac{1}{2}\right) \int x^2 \sin x\, dx.$$

We then need to integrate $x^2 \sin x$, which is more complicated than our original integral, making this an unproductive choice.

We now consider another example.

Example 2   **Integrating using Integration by Parts**

Evaluate $\displaystyle\int xe^x\, dx$.

SOLUTION   Notice that x becomes simpler when differentiated and e^x is unchanged by differentiation or integration. This suggests that we should let $u = x$ and $dv = e^x\, dx$:

$$
\begin{array}{ll}
u = x & dv = e^x\, dx \\
du = ? & v = ?
\end{array}
\qquad \Rightarrow \qquad
\begin{array}{ll}
u = x & dv = e^x\, dx \\
du = dx & v = e^x
\end{array}
$$

Notes:

The Integration by Parts formula gives

$$\int xe^x \, dx = xe^x - \int e^x \, dx.$$

The integral on the right is simple; our final answer is

$$\int xe^x \, dx = xe^x - e^x + C.$$

Note again how the antiderivatives contain a product term.

Example 3   **Integrating using Integration by Parts**

Evaluate $\int x^2 \cos x \, dx$.

   SOLUTION   Let $u = x^2$ instead of the trigonometric function, hence $dv = \cos x \, dx$. Then $du = 2x \, dx$ and $v = \sin x$ as shown below.

$$\begin{array}{ll} u = x^2 & dv = \cos x \, dx \\ du = \, ? & v = \, ? \end{array} \quad \Rightarrow \quad \begin{array}{ll} u = x^2 & dv = \cos x \, dx \\ du = 2x \, dx & v = \sin x \end{array}$$

   The Integration by Parts formula gives

$$\int x^2 \cos x \, dx = x^2 \sin x - \int 2x \sin x \, dx.$$

At this point, the integral on the right is indeed simpler than the one we started with, but to evaluate it, we need to do Integration by Parts again. Here we choose $u = 2x$ and $dv = \sin x$ and fill in the rest below.

$$\begin{array}{ll} u = 2x & dv = \sin x \, dx \\ du = \, ? & v = \, ? \end{array} \quad \Rightarrow \quad \begin{array}{ll} u = 2x & dv = \sin x \, dx \\ du = 2 \, dx & v = -\cos x \end{array}$$

   This means that

$$\int x^2 \cos x \, dx = x^2 \sin x - \left(-2x \cos x - \int -2 \cos x \, dx \right).$$

The integral all the way on the right is now something we can evaluate. It evaluates to $-2 \sin x$. Then going through and simplifying, being careful to keep all the signs straight, our answer is

$$\int x^2 \cos x \, dx = x^2 \sin x + 2x \cos x - 2 \sin x + C.$$

Notes:

Example 4   **Integrating using Integration by Parts**

Evaluate $\int e^x \cos x \, dx$.

SOLUTION   This is a classic problem. In this particular example, one can let u be either $\cos x$ or e^x; we choose $u = e^x$ and hence $dv = \cos x \, dx$. Then $du = e^x \, dx$ and $v = \sin x$ as shown below.

$$
\begin{array}{ll}
u = e^x & dv = \cos x \, dx \\
du = \,? & v = \,?
\end{array}
\quad \Rightarrow \quad
\begin{array}{ll}
u = e^x & dv = \cos x \, dx \\
du = e^x \, dx & v = \sin x
\end{array}
$$

Notice that du is no simpler than u, going against our general rule (but bear with us). The Integration by Parts formula yields

$$\int e^x \cos x \, dx = e^x \sin x - \int e^x \sin x \, dx.$$

The integral on the right is not much different than the one we started with, so it seems like we have gotten nowhere. Let's keep working and apply Integration by Parts to the new integral. So what should we use for u and dv this time? We may feel like letting the trigonometric function be dv and the exponential be u was a bad choice last time since we still can't integrate the new integral. However, if we let $u = \sin x$ and $dv = e^x \, dx$ this time we will reverse what we just did, taking us back to the beginning. So, we let $u = e^x$ and $dv = \sin x \, dx$. This leads us to the following:

$$
\begin{array}{ll}
u = e^x & dv = \sin x \, dx \\
du = \,? & v = \,?
\end{array}
\quad \Rightarrow \quad
\begin{array}{ll}
u = e^x & dv = \sin x \, dx \\
du = e^x \, dx & v = -\cos x
\end{array}
$$

The Integration by Parts formula then gives:

$$\int e^x \cos x \, dx = e^x \sin x - \left(-e^x \cos x - \int -e^x \cos x \, dx \right)$$

$$= e^x \sin x + e^x \cos x - \int e^x \cos x \, dx.$$

It seems we are back right where we started, as the right hand side contains $\int e^x \cos x \, dx$. But this is actually a good thing.

Add $\int e^x \cos x \, dx$ to both sides. This gives

$$2 \int e^x \cos x \, dx = e^x \sin x + e^x \cos x$$

Notes:

Now divide both sides by 2:

$$\int e^x \cos x\, dx = \frac{1}{2}\left(e^x \sin x + e^x \cos x\right).$$

Simplifying a little and adding the constant of integration, our answer is thus

$$\int e^x \cos x\, dx = \frac{1}{2}e^x \left(\sin x + \cos x\right) + C.$$

Example 5   **Integrating using Integration by Parts: antiderivative of** $\ln x$

Evaluate $\int \ln x\, dx$.

SOLUTION   One may have noticed that we have rules for integrating the familiar trigonometric functions and e^x, but we have not yet given a rule for integrating $\ln x$. That is because $\ln x$ can't easily be integrated with any of the rules we have learned up to this point. But we can find its antiderivative by a clever application of Integration by Parts. Set $u = \ln x$ and $dv = dx$. This is a good strategy to learn as it can help in other situations. This determines $du = (1/x)\, dx$ and $v = x$ as shown below.

$$\begin{array}{cc} u = \ln x & dv = dx \\ du = ? & v = ? \end{array} \quad \Rightarrow \quad \begin{array}{cc} u = \ln x & v = x \\ du = 1/x\, dx & dv = dx \end{array}$$

Putting this all together in the Integration by Parts formula, things work out very nicely:

$$\int \ln x\, dx = x \ln x - \int x\, \frac{1}{x}\, dx$$

$$= x \ln x - \int 1\, dx$$

$$= x \ln x - x + C.$$

Example 6   **Using Integration by Parts: antiderivative of** $\tan^{-1} x$

Evaluate $\int \tan^{-1} x\, dx$.

Notes:

SOLUTION The same strategy we used above works here. Let $u = \tan^{-1} x$ and $dv = dx$. Then $du = 1/(1 + x^2)\, dx$ and $v = x$. The Integration by Parts formula gives

$$\int \tan^{-1} x \, dx = x \tan^{-1} x - \int \frac{x}{1 + x^2} \, dx.$$

The integral on the right can be solved by substitution. Taking $t = 1 + x^2$, we get $dt = 2x\, dx$. The integral then becomes

$$\int \tan^{-1} x \, dx = x \tan^{-1} x - \frac{1}{2} \int \frac{1}{t} \, dt.$$

The integral on the right evaluates to $\ln|t| + C$, which becomes $\ln(1 + x^2) + C$. Therefore, the answer is

$$\int \tan^{-1} x \, dx = x \tan^{-1} x - \ln(1 + x^2) + C.$$

Since $1 + x^2 > 0$, we do not need to include the absolute value in the $\ln(1 + x^2)$ term.

Substitution Before Integration

When taking derivatives, it was common to employ multiple rules (such as using both the Quotient and the Chain Rules). It should then come as no surprise that some integrals are best evaluated by combining integration techniques. In particular, here we illustrate making an "unusual" substitution first before using Integration by Parts.

Example 7 **Integration by Parts after substitution**

Evaluate $\displaystyle\int \cos(\ln x) \, dx$.

SOLUTION The integrand contains a composition of functions, leading us to think Substitution would be beneficial. Letting $u = \ln x$, we have $du = 1/x\, dx$. This seems problematic, as we do not have a $1/x$ in the integrand. But consider:

$$du = \frac{1}{x} \, dx \Rightarrow x \cdot du = dx.$$

Since $u = \ln x$, we can use inverse functions to solve for $x = e^u$. Therefore we have that

$$dx = x \cdot du$$
$$= e^u \, du.$$

Notes:

We can thus replace $\ln x$ with u and dx with $e^u \, du$. Thus we rewrite our integral as

$$\int \cos(\ln x)\, dx = \int e^u \cos u \, du.$$

We evaluated this integral in Example 4. Using the result there, we have:

$$
\begin{aligned}
\int \cos(\ln x)\, dx &= \int e^u \cos u \, du \\
&= \frac{1}{2} e^u \big(\sin u + \cos u \big) + C \\
&= \frac{1}{2} e^{\ln x} \big(\sin(\ln x) + \cos(\ln x) \big) + C \\
&= \frac{1}{2} x \big(\sin(\ln x) + \cos(\ln x) \big) + C.
\end{aligned}
$$

Definite Integrals and Integration By Parts

So far we have focused only on evaluating indefinite integrals. Of course, we can use Integration by Parts to evaluate definite integrals as well, as Theorem 54 states. We do so in the next example.

Example 8   **Definite integration using Integration by Parts**

Evaluate $\displaystyle\int_1^2 x^2 \ln x \, dx$.

SOLUTION   To simplify the integral we let $u = \ln x$ and $dv = x^2 \, dx$. We then get $du = (1/x)\, dx$ and $v = x^3/3$ as shown below.

$$
\begin{array}{cc}
u = \ln x & dv = x^2 \, dx \\
du = \, ? & v = \, ?
\end{array}
\quad \Rightarrow \quad
\begin{array}{cc}
u = \ln x & v = x^3/3 \\
du = 1/x \, dx & dv = x^2 \, dx
\end{array}
$$

This may seem counterintuitive since the power on the algebraic factor has

---

Notes:

increased ($v = x^3/3$), but as we see this is a wise choice:

$$\int_1^2 x^2 \ln x \, dx = \frac{x^3}{3} \ln x \bigg|_1^2 - \int_1^2 \frac{x^3}{3} \frac{1}{x} \, dx$$

$$= \frac{x^3}{3} \ln x \bigg|_1^2 - \int_1^2 \frac{x^2}{3} \, dx$$

$$= \frac{x^3}{3} \ln x \bigg|_1^2 - \frac{x^3}{9} \bigg|_1^2$$

$$= \left(\frac{x^3}{3} \ln x - \frac{x^3}{9} \right) \bigg|_1^2$$

$$= \left(\frac{8}{3} \ln 2 - \frac{8}{9} \right) - \left(\frac{1}{3} \ln 1 - \frac{1}{9} \right)$$

$$= \frac{8}{3} \ln 2 - \frac{7}{9}.$$

In general, Integration by Parts is useful for integrating certain products of functions, like $\int xe^x \, dx$ or $\int x^3 \sin x \, dx$. It is also useful for integrals involving logarithms and inverse trigonometric functions.

As stated before, integration is generally more difficult than differentiation. We are developing tools for handling a large array of integrals, and experience will tell us when one tool is preferable/necessary over another. For instance, consider the three similar–looking integrals

$$\int xe^x \, dx, \qquad \int xe^{x^2} \, dx \qquad \text{and} \qquad \int xe^{x^3} \, dx.$$

While the first is calculated easily with Integration by Parts, the second is best approached with Substitution. Taking things one step further, the third integral has no answer in terms of elementary functions, so none of the methods we learn in calculus will get us the exact answer. We will learn how to approximate this integral in Chapter 9

Integration by Parts is a very useful method, second only to substitution. In the following sections of this chapter, we continue to learn other integration techniques. The next section focuses on handling integrals containing trigonometric functions.

Notes:

Exercises 8.1

Terms and Concepts

1. T/F: Integration by Parts is useful in evaluating integrands that contain products of functions.

2. T/F: Integration by Parts can be thought of as the "opposite of the Chain Rule."

Problems

In Exercises 3–36, evaluate the given indefinite integral.

3. $\int x \sin x \, dx$

4. $\int x e^{-x} \, dx$

5. $\int x^2 \sin x \, dx$

6. $\int x^3 \sin x \, dx$

7. $\int x e^{x^2} \, dx$

8. $\int x^3 e^x \, dx$

9. $\int x e^{-2x} \, dx$

10. $\int e^x \sin x \, dx$

11. $\int e^{2x} \cos x \, dx$

12. $\int e^{2x} \sin(3x) \, dx$

13. $\int e^{5x} \cos(5x) \, dx$

14. $\int \sin x \cos x \, dx$

15. $\int \sin^{-1} x \, dx$

16. $\int \tan^{-1}(2x) \, dx$

17. $\int x \tan^{-1} x \, dx$

18. $\int \sin^{-1} x \, dx$

19. $\int x \ln x \, dx$

20. $\int (x - 2) \ln x \, dx$

21. $\int x \ln(x - 1) \, dx$

22. $\int x \ln(x^2) \, dx$

23. $\int x^2 \ln x \, dx$

24. $\int (\ln x)^2 \, dx$

25. $\int (\ln(x + 1))^2 \, dx$

26. $\int x \sec^2 x \, dx$

27. $\int x \csc^2 x \, dx$

28. $\int x \sqrt{x - 2} \, dx$

29. $\int x \sqrt{x^2 - 2} \, dx$

30. $\int \sec x \tan x \, dx$

31. $\int x \sec x \tan x \, dx$

32. $\int x \csc x \cot x \, dx$

33. $\int x \cosh x \, dx$

34. $\int x \sinh x \, dx$

35. $\int \sinh^{-1} x \, dx$

36. $\int \tanh^{-1} x \, dx$

In Exercises 37–42, evaluate the indefinite integral after first making a substitution.

37. $\int \sin(\ln x) \, dx$

38. $\int \sin(\sqrt{x}) \, dx$

39. $\int \ln(\sqrt{x}) \, dx$

40. $\int e^{\sqrt{x}} \, dx$

41. $\int e^{\ln x} \, dx$

42. $\int x^3 e^{x^2} \, dx$

In Exercises 43–51, evaluate the definite integral. Note: the corresponding indefinite integrals appear in Exercises 3 – 11.

43. $\int_0^\pi x \sin x \, dx$

44. $\int_{-1}^1 x e^{-x} \, dx$

45. $\int_{-\pi/4}^{\pi/4} x^2 \sin x \, dx$

46. $\int_{-\pi/2}^{\pi/2} x^3 \sin x \, dx$

47. $\int_0^{\sqrt{\ln 2}} x e^{x^2} \, dx$

48. $\displaystyle\int_0^1 x^3 e^x \, dx$

49. $\displaystyle\int_1^2 xe^{-2x} \, dx$

50. $\displaystyle\int_0^\pi e^x \sin x \, dx$

51. $\displaystyle\int_{-\pi/2}^{\pi/2} e^{2x} \cos x \, dx$

8.2   Trigonometric Integrals

Trigonometric functions are useful for describing periodic behavior. This section describes several techniques for finding antiderivatives of certain combinations of trigonometric functions.

Integrals of the form $\displaystyle\int \sin^m x \cos^n x \, dx$

In learning the technique of Substitution, we saw the integral $\int \sin x \cos x \, dx$ in Example 5.5.4. The integration was not difficult, and one could easily evaluate the indefinite integral by letting $u = \sin x$ or by letting $u = \cos x$. This integral is easy since the power of both sine and cosine is 1.

We generalize this integral and consider integrals of the form $\int \sin^m x \cos^n x \, dx$, where m, n are nonnegative integers. Our strategy for evaluating these integrals is to use the identity $\cos^2 x + \sin^2 x = 1$ to convert high powers of one trigonometric function into the other, leaving a single sine or cosine term in the integrand. We summarize the general technique in the following Key Idea.

Watch the video:
Trigonometric Integrals — Part 2 of 6 at
`https://youtu.be/zyg9k1je7Fg`

Notes:

Key Idea 26   Integrals Involving Powers of Sine and Cosine

Consider $\int \sin^m x \cos^n x \, dx$, where m, n are nonnegative integers.

1. If m is odd, then $m = 2k + 1$ for some integer k. Rewrite

$$\sin^m x = \sin^{2k+1} x = \sin^{2k} x \sin x = (\sin^2 x)^k \sin x = (1 - \cos^2 x)^k \sin x.$$

Then

$$\int \sin^m x \cos^n x \, dx = \int (1 - \cos^2 x)^k \sin x \cos^n x \, dx = -\int (1 - u^2)^k u^n \, du,$$

where $u = \cos x$ and $du = -\sin x \, dx$.

2. If n is odd, then using substitutions similar to that outlined above we have

$$\int \sin^m x \cos^n x \, dx = \int u^m (1 - u^2)^k \, du,$$

where $u = \sin x$ and $du = \cos x \, dx$.

3. If both m and n are even, use the half–angle identities

$$\cos^2 x = \frac{1 + \cos(2x)}{2} \quad \text{and} \quad \sin^2 x = \frac{1 - \cos(2x)}{2}$$

to reduce the degree of the integrand. Expand the result and apply the principles of this Key Idea again.

We practice applying Key Idea 26 in the next examples.

Example 1   Integrating powers of sine and cosine

Evaluate $\int \sin^5 x \cos^8 x \, dx$.

SOLUTION   The power of the sine factor is odd, so we rewrite $\sin^5 x$ as

$$\sin^5 x = \sin^4 x \sin x = (\sin^2 x)^2 \sin x = (1 - \cos^2 x)^2 \sin x.$$

Our integral is now $\int (1 - \cos^2 x)^2 \cos^8 x \sin x \, dx$. Let $u = \cos x$, hence $du =$

Notes:

$-\sin x\, dx$. Making the substitution and expanding the integrand gives

$$\int (1 - \cos^2 x)^2 \cos^8 x \sin x\, dx = -\int (1 - u^2)^2 u^8\, du$$

$$= -\int \left(1 - 2u^2 + u^4\right) u^8\, du$$

$$= -\int \left(u^8 - 2u^{10} + u^{12}\right)\, du$$

$$= -\frac{1}{9}u^9 + \frac{2}{11}u^{11} - \frac{1}{13}u^{13} + C$$

$$= -\frac{1}{9}\cos^9 x + \frac{2}{11}\cos^{11} x - \frac{1}{13}\cos^{13} x + C.$$

Example 2   **Integrating powers of sine and cosine**

Evaluate $\displaystyle\int \sin^5 x \cos^9 x\, dx$.

SOLUTION   The powers of both the sine and cosine factors are odd, therefore we can apply the techniques of Key Idea 26 to either power. We choose to work with the power of the sine factor since that has a smaller exponent. We rewrite $\sin^5 x$ as

$$\sin^5 x = \sin^4 x \sin x$$

$$= (1 - \cos^2 x)^2 \sin x$$

$$= (1 - 2\cos^2 x + \cos^4 x) \sin x.$$

This lets us rewrite the integral as

$$\int \sin^5 x \cos^9 x\, dx = \int \left(1 - 2\cos^2 x + \cos^4 x\right) \sin x \cos^9 x\, dx.$$

Substituting and integrating with $u = \cos x$ and $du = -\sin x\, dx$, we have

$$\int \left(1 - 2\cos^2 x + \cos^4 x\right) \sin x \cos^9 x\, dx$$

$$= -\int \left(1 - 2u^2 + u^4\right) u^9\, du$$

$$= -\int u^9 - 2u^{11} + u^{13}\, du$$

$$= -\frac{1}{10}u^{10} + \frac{1}{6}u^{12} - \frac{1}{14}u^{14} + C$$

$$= -\frac{1}{10}\cos^{10} x + \frac{1}{6}\cos^{12} x - \frac{1}{14}\cos^{14} x + C.$$

Notes:

Instead, another approach would be to rewrite $\cos^9 x$ as

$$\cos^9 x = \cos^8 x \cos x$$
$$= (\cos^2 x)^4 \cos x$$
$$= (1 - \sin^2 x)^4 \cos x$$
$$= (1 - 4\sin^2 x + 6\sin^4 x - 4\sin^6 x + \sin^8 x)\cos x.$$

We rewrite the integral as

$$\int \sin^5 x \cos^9 x \, dx = \int \sin^5 x \left(1 - 4\sin^2 x + 6\sin^4 x - 4\sin^6 x + \sin^8 x\right)\cos x \, dx.$$

Now substitute and integrate, using $u = \sin x$ and $du = \cos x \, dx$.

$$\int \sin^5 x \left(1 - 4\sin^2 x + 6\sin^4 x - 4\sin^6 x + \sin^8 x\right)\cos x \, dx$$
$$= \int u^5 (1 - 4u^2 + 6u^4 - 4u^6 + u^8) \, du$$
$$= \int \left(u^5 - 4u^7 + 6u^9 - 4u^{11} + u^{13}\right) du$$
$$= \frac{1}{6}u^6 - \frac{1}{2}u^8 + \frac{3}{5}u^{10} - \frac{1}{3}u^{12} + \frac{1}{14}u^{14} + C$$
$$= \frac{1}{6}\sin^6 x - \frac{1}{2}\sin^8 x + \frac{3}{5}\sin^{10} x - \frac{1}{3}\sin^{12} x + \frac{1}{14}\sin^{14} x + C.$$

Technology Note: The work we are doing here can be a bit tedious, but the skills developed (problem solving, algebraic manipulation, etc.) are important. Nowadays problems of this sort are often solved using a computer algebra system. The powerful program *Mathematica*® integrates $\int \sin^5 x \cos^9 x \, dx$ as

$$f(x) = -\frac{45\cos(2x)}{16384} - \frac{5\cos(4x)}{8192} + \frac{19\cos(6x)}{49152} + \frac{\cos(8x)}{4096} - \frac{\cos(10x)}{81920} - \frac{\cos(12x)}{24576} - \frac{\cos(14x)}{114688},$$

which clearly has a different form than our second answer in Example 2, which is

$$g(x) = \frac{1}{6}\sin^6 x - \frac{1}{2}\sin^8 x + \frac{3}{5}\sin^{10} x - \frac{1}{3}\sin^{12} x + \frac{1}{14}\sin^{14} x.$$

Figure 8.1 shows a graph of f and g; they are clearly not equal, but they differ *only by a constant*: $g(x) = f(x) + C$ for some constant C. So we have two different antiderivatives of the same function, meaning both answers are correct.
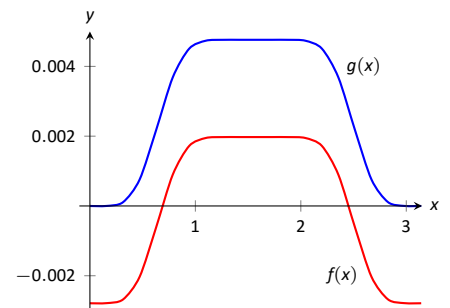
Notes:

Figure 8.1: A plot of $f(x)$ and $g(x)$ from Example 2 and the Technology Note.

Example 3   **Integrating powers of sine and cosine**

Evaluate $\int \sin^2 x \, dx$.

SOLUTION   The power of sine is even so we employ a half-angle identity, algebra and a u- substitution as follows:

$$\int \sin^2 x \, dx = \int \frac{1 - \cos(2x)}{2} \, dx$$

$$= \frac{1}{2} \int 1 - \cos(2x) \, dx$$

$$= \frac{1}{2} \left(x - \frac{1}{2} \sin(2x) \right) + C$$

$$= \frac{1}{2}x - \frac{1}{4} \sin(2x) + C.$$

Example 4   **Integrating powers of sine and cosine**

Evaluate $\int \cos^4 x \sin^2 x \, dx$.

SOLUTION   The powers of sine and cosine are both even, so we employ the half–angle formulas and algebra as follows.

$$\int \cos^4 x \sin^2 x \, dx = \int \left(\frac{1 + \cos(2x)}{2} \right)^2 \left(\frac{1 - \cos(2x)}{2} \right) \, dx$$

$$= \int \frac{1 + 2\cos(2x) + \cos^2(2x)}{4} \cdot \frac{1 - \cos(2x)}{2} \, dx$$

$$= \int \frac{1}{8} \left(1 + \cos(2x) - \cos^2(2x) - \cos^3(2x) \right) \, dx$$

The $\cos(2x)$ term is easy to integrate. The $\cos^2(2x)$ term is another trigonometric integral with an even power, requiring the half–angle formula again. The $\cos^3(2x)$ term is a cosine function with an odd power, requiring a substitution as done before. We integrate each in turn below.

$$\int \cos(2x) \, dx = \frac{1}{2} \sin(2x) + C.$$

$$\int \cos^2(2x) \, dx = \int \frac{1 + \cos(4x)}{2} \, dx = \frac{1}{2} \left(x + \frac{1}{4} \sin(4x) \right) + C.$$

Notes:

Finally, we rewrite $\cos^3(2x)$ as

$$\cos^3(2x) = \cos^2(2x)\cos(2x) = \left(1 - \sin^2(2x)\right)\cos(2x).$$

Letting $u = \sin(2x)$, we have $du = 2\cos(2x)\,dx$, hence

$$\int \cos^3(2x)\,dx = \int \left(1 - \sin^2(2x)\right)\cos(2x)\,dx$$

$$= \int \frac{1}{2}(1 - u^2)\,du$$

$$= \frac{1}{2}\left(u - \frac{1}{3}u^3\right) + C$$

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

Putting all the pieces together, we have

$$\int \cos^4 x \sin^2 x\,dx = \int \frac{1}{8}\left(1 + \cos(2x) - \cos^2(2x) - \cos^3(2x)\right)dx$$

$$= \frac{1}{8}\left[x + \frac{1}{2}\sin(2x) - \frac{1}{2}\left(x + \frac{1}{4}\sin(4x)\right) - \frac{1}{2}\left(\sin(2x) - \frac{1}{3}\sin^3(2x)\right)\right] + C$$

$$= \frac{1}{8}\left[\frac{1}{2}x - \frac{1}{8}\sin(4x) + \frac{1}{6}\sin^3(2x)\right] + C.$$

The process above was a bit long and tedious, but being able to work a problem such as this from start to finish is important.

Integrals of the form $\displaystyle\int \tan^m x \sec^n x\,dx$.

When evaluating integrals of the form $\int \sin^m x \cos^n x\,dx$, the Pythagorean Theorem allowed us to convert even powers of sine into even powers of cosine, and vise–versa. If, for instance, the power of sine was odd, we pulled out one $\sin x$ and converted the remaining even power of $\sin x$ into a function using powers of $\cos x$, leading to an easy substitution.

The same basic strategy applies to integrals of the form $\int \tan^m x \sec^n x\,dx$, albeit a bit more nuanced. The following three facts will prove useful:

- $\frac{d}{dx}(\tan x) = \sec^2 x$,

- $\frac{d}{dx}(\sec x) = \sec x \tan x$, and

- $1 + \tan^2 x = \sec^2 x$ (the Pythagorean Theorem).

Notes:

If the integrand can be manipulated to separate a $\sec^2 x$ term with the remaining secant power even, or if a $\sec x \tan x$ term can be separated with the remaining $\tan x$ power even, the Pythagorean Theorem can be employed, leading to a simple substitution. This strategy is outlined in the following Key Idea.

Key Idea 27   Integrals Involving Powers of Tangent and Secant

Consider $\displaystyle\int \tan^m x \sec^n x\, dx$, where m, n are nonnegative integers.

1. If n is even, then $n = 2k$ for some integer k. Rewrite $\sec^n x$ as

$$\sec^n x = \sec^{2k} x = \sec^{2k-2} x \sec^2 x = (1 + \tan^2 x)^{k-1} \sec^2 x.$$

Then

$$\int \tan^m x \sec^n x\, dx = \int \tan^m x(1 + \tan^2 x)^{k-1} \sec^2 x\, dx = \int u^m(1 + u^2)^{k-1}\, du,$$

where $u = \tan x$ and $du = \sec^2 x\, dx$.

2. If m is odd, then $m = 2k + 1$ for some integer k. Rewrite $\tan^m x \sec^n x$ as

$$\tan^m x \sec^n x = \tan^{2k+1} x \sec^n x = \tan^{2k} x \sec^{n-1} x \sec x \tan x = (\sec^2 x - 1)^k \sec^{n-1} x \sec x \tan x.$$

Then

$$\int \tan^m x \sec^n x\, dx = \int (\sec^2 x - 1)^k \sec^{n-1} x \sec x \tan x\, dx = \int (u^2 - 1)^k u^{n-1}\, du,$$

where $u = \sec x$ and $du = \sec x \tan x\, dx$.

3. If n is odd and m is even, then $m = 2k$ for some integer k. Convert $\tan^m x$ to $(\sec^2 x - 1)^k$. Expand the new integrand and use Integration By Parts, with $dv = \sec^2 x\, dx$.

4. If m is even and $n = 0$, rewrite $\tan^m x$ as

$$\tan^m x = \tan^{m-2} x \tan^2 x = \tan^{m-2} x(\sec^2 x - 1) = \tan^{m-2} \sec^2 x - \tan^{m-2} x.$$

So

$$\int \tan^m x\, dx = \underbrace{\int \tan^{m-2} \sec^2 x\, dx}_{\text{apply rule \#1}} - \underbrace{\int \tan^{m-2} x\, dx}_{\text{apply rule \#4 again}}.$$

The techniques described in items 1 and 2 of Key Idea 27 are relatively straight-

---

Notes:

forward, but the techniques in items 3 and 4 can be rather tedious. A few examples will help with these methods.

Example 5   **Integrating powers of tangent and secant**

Evaluate $\int \tan^2 x \sec^6 x \, dx$.

SOLUTION   Since the power of secant is even, we use rule #1 from Key Idea 27 and pull out a $\sec^2 x$ in the integrand. We convert the remaining powers of secant into powers of tangent.

$$\int \tan^2 x \sec^6 x \, dx = \int \tan^2 x \sec^4 x \sec^2 x \, dx$$

$$= \int \tan^2 x \left(1 + \tan^2 x\right)^2 \sec^2 x \, dx$$

Now substitute, with $u = \tan x$, with $du = \sec^2 x \, dx$.

$$= \int u^2 \left(1 + u^2\right)^2 \, du$$

We leave the integration and subsequent substitution to the reader. The final answer is

$$= \frac{1}{3} \tan^3 x + \frac{2}{5} \tan^5 x + \frac{1}{7} \tan^7 x + C.$$

We derived integrals for tangent and secant in Section 5.5 and will regularly use them when evaluating integrals of the form $\tan^m x \sec^n x \, dx$. As a reminder:

$$\int \tan x \, dx = \ln |\sec x| + C$$

$$\int \sec x \, dx = \ln |\sec x + \tan x| + C$$

Example 6   **Integrating powers of tangent and secant**

Evaluate $\int \sec^3 x \, dx$.

SOLUTION   We apply rule #3 from Key Idea 27 as the power of secant is odd and the power of tangent is even (0 is an even number). We use Integration by Parts; the rule suggests letting $dv = \sec^2 x \, dx$, meaning that $u = \sec x$.

Notes:

$$u = \sec x \quad dv = \sec^2 x\, dx$$
$$du = ? \qquad v = ?$$

$\Rightarrow$

$$u = \sec x \qquad\qquad dv = \sec^2 x\, dx$$
$$du = \sec x \tan x\, dx \qquad v = \tan x$$

Figure 8.2: Setting up Integration by Parts.

Employing Integration by Parts, we have

$$\int \sec^3 x\, dx = \int \underbrace{\sec x}_{u} \cdot \underbrace{\sec^2 x\, dx}_{dv}$$

$$= \sec x \tan x - \int \sec x \tan^2 x\, dx.$$

This new integral also requires applying rule #3 of Key Idea 27:

$$= \sec x \tan x - \int \sec x \big(\sec^2 x - 1 \big)\, dx$$

$$= \sec x \tan x - \int \sec^3 x\, dx + \int \sec x\, dx$$

$$= \sec x \tan x - \int \sec^3 x\, dx + \ln |\sec x + \tan x|\, dx$$

Note: Remember that in Example 5.5.8, we found that $\int \sec x\, dx = \ln |\sec x + \tan x| + C$

In previous applications of Integration by Parts, we have seen where the original integral has reappeared in our work. We resolve this by adding $\int \sec^3 x\, dx$ to both sides, giving:

$$2 \int \sec^3 x\, dx = \sec x \tan x + \ln |\sec x + \tan x|$$

$$\int \sec^3 x\, dx = \frac{1}{2} \Big(\sec x \tan x + \ln |\sec x + \tan x| \Big) + C.$$

We give one more example.

Example 7   **Integrating powers of tangent and secant**

Evaluate $\int \tan^6 x\, dx$.

SOLUTION   We employ rule #4 of Key Idea 27.

$$\int \tan^6 x\, dx = \int \tan^4 x \tan^2 x\, dx$$

$$= \int \tan^4 x \big(\sec^2 x - 1 \big)\, dx$$

$$= \int \tan^4 x \sec^2 x\, dx - \int \tan^4 x\, dx$$

Notes:

We integrate the first integral with substitution, $u = \tan x$ and $du = \sec^2 x\, dx$; and the second by employing rule #4 again.

$$= \int u^4\, du - \int \tan^2 x \tan^2 x\, dx$$

$$= \frac{1}{5}\tan^5 x - \int \tan^2 x\left(\sec^2 x - 1\right)\, dx$$

$$= \frac{1}{5}\tan^5 x - \int \tan^2 x \sec^2 x\, dx + \int \tan^2 x\, dx$$

Again, use substitution for the first integral and rule #4 for the second.

$$= \frac{1}{5}\tan^5 x - \frac{1}{3}\tan^3 x + \int \left(\sec^2 x - 1\right)\, dx$$

$$= \frac{1}{5}\tan^5 x - \frac{1}{3}\tan^3 x + \tan x - x + C.$$

Integrals of the form $\displaystyle\int \cot^m x \csc^n x\, dx$

Not surprisingly, evaluating integrals of the form $\int \cot^m x \csc^n x\, dx$ is similar to evaluating $\int \tan^m x \sec^n x\, dx$. The guidelines from Key Idea 27 and the following three facts will be useful:

$$\frac{d}{dx}(\cot x) = -\csc^2 x$$

$$\frac{d}{dx}(\csc x) = -\csc x \cot x, \qquad \text{and}$$

$$\csc^2 x = \cot^2 x + 1$$

Example 8   **Integrating powers of cotangent and cosecant**

Evaluate $\displaystyle\int \cot^2 x \csc^4 x\, dx$

SOLUTION   Since the power of cosecant is even we will let $u = \cot x$ and save a $\csc^2 x$ for the resulting $du = -\csc^2 x\, dx$.

$$\int \cot^2 x \csc^4 x\, dx = \int \cot^2 x \csc^2 x \csc^2 x\, dx$$

$$= \int \cot^2 x(1 + \cot^2 x)\csc^2 x\, dx$$

$$= -\int u^2(1 + u^2)\, du.$$

Notes:

The integration and substitution required to finish this example are similar to that of previous examples in this section. The result is

$$-\frac{1}{3}\cot^3 x - \frac{1}{5}\cot^5 x + C.$$

Integrals of the form $\int \sin(mx)\sin(nx)\,dx$, $\int \cos(mx)\cos(nx)\,dx$, and $\int \sin(mx)\cos(nx)\,dx$.

Functions that contain products of sines and cosines of differing periods are important in many applications including the analysis of sound waves. Integrals of the form

$$\int \sin(mx)\sin(nx)\,dx, \quad \int \cos(mx)\cos(nx)\,dx \quad \text{and} \quad \int \sin(mx)\cos(nx)\,dx$$

are best approached by first applying the Product to Sum Formulas of Trigonometry found in the back cover of this text, namely

$$\sin(mx)\sin(nx) = \frac{1}{2}\Big[\cos\big((m-n)x\big) - \cos\big((m+n)x\big)\Big]$$
$$\cos(mx)\cos(nx) = \frac{1}{2}\Big[\cos\big((m-n)x\big) + \cos\big((m+n)x\big)\Big]$$
$$\sin(mx)\cos(nx) = \frac{1}{2}\Big[\sin\big((m-n)x\big) + \sin\big((m+n)x\big)\Big]$$

Example 9   **Integrating products of** $\sin(mx)$ **and** $\cos(nx)$

Evaluate $\int \sin(5x)\cos(2x)\,dx$.

SOLUTION   The application of the formula and subsequent integration are straightforward:

$$\int \sin(5x)\cos(2x)\,dx = \int \frac{1}{2}\Big[\sin(3x) + \sin(7x)\Big]\,dx$$
$$= -\frac{1}{6}\cos(3x) - \frac{1}{14}\cos(7x) + C.$$

Notes:

Integrating other combinations of trigonometric functions

Combinations of trigonometric functions that we have not discussed in this chapter are evaluated by applying algebra, trigonometric identities and other integration strategies to create an equivalent integrand that we can evaluate. To evaluate "crazy" combinations, those not readily manipulated into a familiar form, one should use integral tables. A table of "common crazy" combinations can be found at the end of this text.

These latter examples were admittedly long, with repeated applications of the same rule. Try to not be overwhelmed by the length of the problem, but rather admire how robust this solution method is. A trigonometric function of a high power can be systematically reduced to trigonometric functions of lower powers until all antiderivatives can be computed.

The next section introduces an integration technique known as Trigonometric Substitution, a clever combination of Substitution and the Pythagorean Theorem.

Notes:

Exercises 8.2

Terms and Concepts

1. T/F: $\int \sin^2 x \cos^2 x \, dx$ cannot be evaluated using the techniques described in this section since both powers of $\sin x$ and $\cos x$ are even.

2. T/F: $\int \sin^3 x \cos^3 x \, dx$ cannot be evaluated using the techniques described in this section since both powers of $\sin x$ and $\cos x$ are odd.

3. T/F: This section addresses how to evaluate indefinite integrals such as $\int \sin^5 x \tan^3 x \, dx$.

Problems

In Exercises 4–30, evaluate the indefinite integral.

4. $\int \cos^2 x \, dx$

5. $\int \cos^4 x \, dx$

6. $\int \sin^3 x \cos^2 x \, dx$

7. $\int \sin^3 x \cos^3 x \, dx$

8. $\int \sin^6 x \cos^5 x \, dx$

9. $\int \cos^2 x \tan^3 x \, dx$

10. $\int \sin^2 x \cos^2 x \, dx$

11. $\int \sin^3 x \sqrt{\cos x} \, dx$

12. $\int \sin(x) \cos(2x) \, dx$

13. $\int \sin(3x) \sin(7x) \, dx$

14. $\int \sin(\pi x) \sin(2\pi x) \, dx$

15. $\int \cos(x) \cos(2x) \, dx$

16. $\int \cos\left(\frac{\pi}{2}x\right) \cos(\pi x) \, dx$

17. $\int \tan^2 x \, dx$

18. $\int \tan^2 x \sec^4 x \, dx$

19. $\int \tan^3 x \sec^4 x \, dx$

20. $\int \tan^3 x \sec^2 x \, dx$

21. $\int \tan^3 x \sec^3 x \, dx$

22. $\int \tan^5 x \sec^5 x \, dx$

23. $\int \tan^4 x \, dx$

24. $\int \sec^5 x \, dx$

25. $\int \tan^2 x \sec x \, dx$

26. $\int \tan^2 x \sec^3 x \, dx$

27. $\int \csc x \, dx$

28. $\int \cot^3 x \csc^3 x \, dx$

29. $\int \cot^3 x \, dx$

30. $\int \cot^6 x \csc^4 x \, dx$

In Exercises 31–39, evaluate the definite integral. Note: the corresponding indefinite integrals appear in the previous set.

31. $\int_0^\pi \sin x \cos^4 x \, dx$

32. $\int_{-\pi}^\pi \sin^3 x \cos x \, dx$

33. $\int_{-\pi/2}^{\pi/2} \sin^2 x \cos^7 x \, dx$

34. $\int_0^{\pi/2} \sin(5x) \cos(3x) \, dx$

35. $\int_{-\pi/2}^{\pi/2} \cos(x) \cos(2x) \, dx$

36. $\int_0^{\pi/4} \tan^4 x \sec^2 x \, dx$

37. $\int_{-\pi/4}^{\pi/4} \tan^2 x \sec^4 x \, dx$

38. $\int_{\frac{\pi}{6}}^{\frac{\pi}{2}} \cot^2 x \, dx$

39. $\int_{\frac{\pi}{4}}^{\frac{\pi}{2}} \cot^3 x \, dx$

8.3 Trigonometric Substitution

In Section 5.2 we defined the definite integral as the "signed area under the curve." In that section we had not yet learned the Fundamental Theorem of Calculus, so we evaluated special definite integrals which described nice, geometric shapes. For instance, we were able to evaluate

$$\int_{-3}^{3} \sqrt{9 - x^2} \, dx = \frac{9\pi}{2} \tag{8.1}$$

as we recognized that $f(x) = \sqrt{9 - x^2}$ described the upper half of a circle with radius 3.

We have since learned a number of integration techniques, including Substitution and Integration by Parts, yet we are still unable to evaluate the above integral without resorting to a geometric interpretation. This section introduces Trigonometric Substitution, a method of integration that fills this gap in our integration skill. This technique works on the same principle as Substitution as found in Section 5.5, though it can feel "backward." In Section 5.5, we set $u = f(x)$, for some function f, and replaced $f(x)$ with u. In this section, we will set $x = f(\theta)$, where f is a trigonometric function, then replace x with $f(\theta)$.

> Watch the video:
> Trigonometric Substitution — Example 3 / Part 1 at
> https://youtu.be/yW6OduOYHLO

We start by demonstrating this method in evaluating the integral in Equation (8.1). After the example, we will generalize the method and give more examples.

Example 1 Using Trigonometric Substitution

Evaluate $\displaystyle\int_{-3}^{3} \sqrt{9 - x^2} \, dx$.

SOLUTION We begin by noting that $9 \sin^2 \theta + 9 \cos^2 \theta = 9$, and hence $9 \cos^2 \theta = 9 - 9 \sin^2 \theta$. If we let $x = 3 \sin \theta$, then $9 - x^2 = 9 - 9 \sin^2 \theta = 9 \cos^2 \theta$.

Setting $x = 3 \sin \theta$ gives $dx = 3 \cos \theta \, d\theta$. We are almost ready to substitute. We also change our bounds of integration. The bound $x = -3$ corresponds to $\theta = -\pi/2$ (for when $\theta = -\pi/2$, $x = 3 \sin \theta = -3$). Likewise, the bound of

Notes:

$x = 3$ is replaced by the bound $\theta = \pi/2$. Thus

$$\int_{-3}^{3} \sqrt{9 - x^2}\, dx = \int_{-\pi/2}^{\pi/2} \sqrt{9 - 9 \sin^2 \theta}\,(3 \cos \theta)\, d\theta$$

$$= \int_{-\pi/2}^{\pi/2} 3\sqrt{9 \cos^2 \theta} \cos \theta\, d\theta$$

$$= \int_{-\pi/2}^{\pi/2} 3 \left|3 \cos \theta\right| \cos \theta\, d\theta.$$

On $[-\pi/2, \pi/2]$, $\cos \theta$ is always positive, so we can drop the absolute value bars, then employ a half–angle formula:

$$= \int_{-\pi/2}^{\pi/2} 9 \cos^2 \theta\, d\theta$$

$$= \int_{-\pi/2}^{\pi/2} \frac{9}{2} \left(1 + \cos(2\theta)\right) d\theta$$

$$= \frac{9}{2} \left(\theta + \frac{1}{2} \sin(2\theta)\right)\Big|_{-\pi/2}^{\pi/2} = \frac{9}{2}\pi.$$

This matches our answer from before.

We now describe in detail Trigonometric Substitution. This method excels when dealing with integrands that contain $\sqrt{a^2 - x^2}$, $\sqrt{x^2 - a^2}$ and $\sqrt{x^2 + a^2}$. The following Key Idea outlines the procedure for each case, followed by more examples.

Notes:

> **Key Idea 28   Trigonometric Substitution**
>
> (a) For integrands containing $\sqrt{a^2 - x^2}$:
>
>   Let $x = a\sin\theta$,   for $-\pi/2 \le \theta \le \pi/2$ and $a > 0$.
>
>   On this interval, $\cos\theta \ge 0$, so $\sqrt{a^2 - x^2} = a\cos\theta$
>
> (b) For integrands containing $\sqrt{x^2 + a^2}$:
>
>   Let $x = a\tan\theta$,   for $-\pi/2 < \theta < \pi/2$ and $a > 0$.
>
>   On this interval, $\sec\theta > 0$, so $\sqrt{x^2 + a^2} = a\sec\theta$
>
> (c) For integrands containing $\sqrt{x^2 - a^2}$:
>
>   Let $x = a\sec\theta$,   restricting our work to where $x \ge a > 0$,
>   so $x/a \ge 1$, and $0 \le \theta < \pi/2$.
>
>   On this interval, $\tan\theta \ge 0$, so $\sqrt{x^2 - a^2} = a\tan\theta$

Example 2   Using Trigonometric Substitution

Evaluate $\displaystyle\int \frac{1}{\sqrt{5 + x^2}}\, dx$.

SOLUTION   Using Key Idea 28(b), we recognize $a = \sqrt{5}$ and set $x = \sqrt{5}\tan\theta$. This makes $dx = \sqrt{5}\sec^2\theta\, d\theta$. We will use the fact that $\sqrt{5 + x^2} = \sqrt{5 + 5\tan^2\theta} = \sqrt{5\sec^2\theta} = \sqrt{5}\sec\theta$. Substituting, we have:

$$\int \frac{1}{\sqrt{5 + x^2}}\, dx = \int \frac{1}{\sqrt{5 + 5\tan^2\theta}}\sqrt{5}\sec^2\theta\, d\theta$$

$$= \int \frac{\sqrt{5}\sec^2\theta}{\sqrt{5}\sec\theta}\, d\theta$$

$$= \int \sec\theta\, d\theta$$

$$= \ln|\sec\theta + \tan\theta| + C.$$

While the integration steps are over, we are not yet done. The original problem was stated in terms of x, whereas our answer is given in terms of θ. We must convert back to x.

The lengths of the sides of the reference triangle in Figure 8.3 are determined by the Pythagorean Theorem. With $x = \sqrt{5}\tan\theta$, we have

$$\tan\theta = \frac{x}{\sqrt{5}} \quad \text{and} \quad \sec\theta = \frac{\sqrt{x^2 + 5}}{\sqrt{5}}.$$
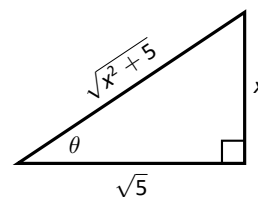
Figure 8.3: A reference triangle for Example 2

Notes:

This gives

$$\int \frac{1}{\sqrt{5+x^2}}\, dx = \ln \left| \sec\theta + \tan\theta \right| + C$$

$$= \ln \left| \frac{\sqrt{x^2+5}}{\sqrt{5}} + \frac{x}{\sqrt{5}} \right| + C.$$

We can leave this answer as is, or we can use a logarithmic identity to simplify it. Note:

$$\ln \left| \frac{\sqrt{x^2+5}}{\sqrt{5}} + \frac{x}{\sqrt{5}} \right| + C = \ln \left| \frac{1}{\sqrt{5}} \left(\sqrt{x^2+5} + x \right) \right| + C$$

$$= \ln \left| \frac{1}{\sqrt{5}} \right| + \ln \left| \sqrt{x^2+5} + x \right| + C$$

$$= \ln \left| \sqrt{x^2+5} + x \right| + C,$$

where the $\ln\left(1/\sqrt{5}\right)$ term is absorbed into the constant C. (In Section 7.4 we learned another way of approaching this problem.)

Example 3　　　　**Using Trigonometric Substitution**

Evaluate $\displaystyle\int \sqrt{4x^2-1}\, dx$.

Solution　　　　We start by rewriting the integrand so that it looks like $\sqrt{x^2-a^2}$ for some value of a:

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

$$= 2\sqrt{x^2 - \left(\frac{1}{2}\right)^2}.$$

So we have $a = 1/2$, and following Key Idea 28(c), we set $x = \frac{1}{2}\sec\theta$, and hence

---

Notes:

$dx = \frac{1}{2} \sec\theta \tan\theta \, d\theta$. We now rewrite the integral with these substitutions:

$$\int \sqrt{4x^2 - 1} \, dx = \int 2\sqrt{x^2 - \left(\frac{1}{2}\right)^2} \, dx$$

$$= \int 2\sqrt{\frac{1}{4}\sec^2\theta - \frac{1}{4}} \left(\frac{1}{2}\sec\theta\tan\theta\right) d\theta$$

$$= \int \sqrt{\frac{1}{4}(\sec^2\theta - 1)} \left(\sec\theta\tan\theta\right) d\theta$$

$$= \int \sqrt{\frac{1}{4}\tan^2\theta} \left(\sec\theta\tan\theta\right) d\theta$$

$$= \int \frac{1}{2}\tan^2\theta \sec\theta \, d\theta$$

$$= \frac{1}{2} \int \left(\sec^2\theta - 1\right) \sec\theta \, d\theta$$

$$= \frac{1}{2} \int \left(\sec^3\theta - \sec\theta\right) d\theta.$$

We integrated $\sec^3\theta$ in Example 8.2.6, finding its antiderivatives to be

$$\int \sec^3\theta \, d\theta = \frac{1}{2}\left(\sec\theta\tan\theta + \ln|\sec\theta + \tan\theta|\right) + C.$$

Thus

$$\int \sqrt{4x^2 - 1} \, dx$$

$$= \frac{1}{2} \int \left(\sec^3\theta - \sec\theta\right) d\theta$$

$$= \frac{1}{2}\left(\frac{1}{2}\left(\sec\theta\tan\theta + \ln|\sec\theta + \tan\theta|\right) - \ln|\sec\theta + \tan\theta|\right) + C$$

$$= \frac{1}{4}\left(\sec\theta\tan\theta - \ln|\sec\theta + \tan\theta|\right) + C.$$

We are not yet done. Our original integral is given in terms of x, whereas our final answer, as given, is in terms of θ. We need to rewrite our answer in terms of x. With $a = 1/2$, and $x = \frac{1}{2}\sec\theta$, we use the Pythagorean Theorem to determine the lengths of the sides of the reference triangle in Figure 8.4.

$$\tan\theta = \frac{\sqrt{x^2 - \frac{1}{4}}}{\frac{1}{2}} = 2\sqrt{x^2 - \frac{1}{4}} \qquad \text{and} \qquad \sec\theta = 2x.$$
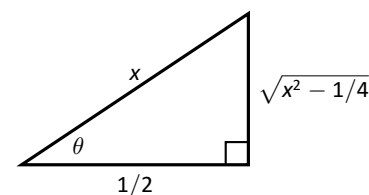
Figure 8.4: A reference triangle for Example 3

Notes:

Therefore,

$$\int \sqrt{4x^2 - 1}\, dx = \frac{1}{4}\left(\sec\theta \tan\theta - \ln|\sec\theta + \tan\theta| \right) + C$$

$$= \frac{1}{4}\left(2x \cdot 2\sqrt{x^2 - \frac{1}{4}} - \ln\left|2x + 2\sqrt{x^2 - \frac{1}{4}}\right| \right) + C$$

$$= \frac{1}{4}\left(4x\sqrt{x^2 - \frac{1}{4}} - \ln\left|2x + 2\sqrt{x^2 - \frac{1}{4}}\right| \right) + C$$

$$= \frac{1}{4}\left(2x\sqrt{4x^2 - 1} - \ln\left|2x + \sqrt{4x^2 - 1}\right| \right) + C.$$

Example 4   **Using Trigonometric Substitution**

Evaluate $\displaystyle\int \frac{\sqrt{4 - x^2}}{x^2}\, dx$.

SOLUTION   We use Key Idea 28(a) with $a = 2$, $x = 2\sin\theta$, $dx = 2\cos\theta$ and hence $\sqrt{4 - x^2} = 2\cos\theta$. This gives

$$\int \frac{\sqrt{4 - x^2}}{x^2}\, dx = \int \frac{2\cos\theta}{4\sin^2\theta}(2\cos\theta)\, d\theta$$

$$= \int \cot^2\theta\, d\theta$$

$$= \int (\csc^2\theta - 1)\, d\theta$$

$$= -\cot\theta - \theta + C.$$

We need to rewrite our answer in terms of x. Using the Pythagorean Theorem we determine the lengths of the sides of the reference triangle in Figure 8.5. We have $\cot\theta = \sqrt{4 - x^2}/x$ and $\theta = \sin^{-1}(x/2)$. Thus

$$\int \frac{\sqrt{4 - x^2}}{x^2}\, dx = -\frac{\sqrt{4 - x^2}}{x} - \sin^{-1}\left(\frac{x}{2}\right) + C.$$
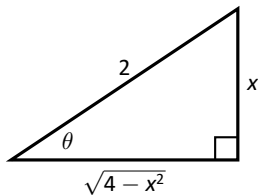
Figure 8.5: A reference triangle for Example 4

Trigonometric Substitution can be applied in many situations, even those not of the form $\sqrt{a^2 - x^2}$, $\sqrt{x^2 - a^2}$ or $\sqrt{x^2 + a^2}$. In the following example, we apply it to an integral we already know how to handle.

Example 5   **Using Trigonometric Substitution**

Evaluate $\displaystyle\int \frac{1}{x^2 + 1}\, dx$.

Notes:

SOLUTION We know the answer already as $\tan^{-1} x + C$. We apply Trigonometric Substitution here to show that we get the same answer without inherently relying on knowledge of the derivative of the arctangent function.

Using Key Idea 28(b), let $x = \tan\theta$, $dx = \sec^2\theta\,d\theta$ and note that $x^2 + 1 = \tan^2\theta + 1 = \sec^2\theta$. Thus

$$\int \frac{1}{x^2 + 1}\,dx = \int \frac{1}{\sec^2\theta} \sec^2\theta\,d\theta$$

$$= \int 1\,d\theta$$

$$= \theta + C.$$

Since $x = \tan\theta$, $\theta = \tan^{-1} x$, and we conclude that $\displaystyle\int \frac{1}{x^2 + 1}\,dx = \tan^{-1} x + C.$

The next example is similar to the previous one in that it does not involve a square–root. It shows how several techniques and identities can be combined to obtain a solution.

Example 6 **Using Trigonometric Substitution**

Evaluate $\displaystyle\int \frac{1}{(x^2 + 6x + 10)^2}\,dx.$

SOLUTION We start by completing the square, then make the substitution $u = x + 3$, followed by the trigonometric substitution of $u = \tan\theta$:

$$\int \frac{1}{(x^2 + 6x + 10)^2}\,dx = \int \frac{1}{\left((x+3)^2 + 1\right)^2}\,dx = \int \frac{1}{(u^2 + 1)^2}\,du.$$

Now make the substitution $u = \tan\theta$, $du = \sec^2\theta\,d\theta$:

$$= \int \frac{1}{(\tan^2\theta + 1)^2} \sec^2\theta\,d\theta$$

$$= \int \frac{1}{(\sec^2\theta)^2} \sec^2\theta\,d\theta$$

$$= \int \cos^2\theta\,d\theta.$$

Applying a half–angle formula, we have

$$= \int \left(\frac{1}{2} + \frac{1}{2}\cos(2\theta)\right)\,d\theta$$

$$= \frac{1}{2}\theta + \frac{1}{4}\sin(2\theta) + C. \tag{8.2}$$

Note: Remember the sine and cosine double angle identities:

$$\sin 2\theta = 2\sin\theta\cos\theta$$
$$\cos 2\theta = \cos^2\theta - \sin^2\theta$$
$$= 2\cos^2\theta - 1$$
$$= 1 - 2\sin^2\theta$$

They are often needed for writing your final answer in terms of x.

Notes:

We need to return to the variable x. As $u = \tan\theta$, $\theta = \tan^{-1} u$. Using the identity $\sin(2\theta) = 2\sin\theta\cos\theta$ and using the reference triangle found in Key Idea 28(b), we have

$$\frac{1}{4}\sin(2\theta) = \frac{1}{2}\frac{u}{\sqrt{u^2+1}}\cdot\frac{1}{\sqrt{u^2+1}} = \frac{1}{2}\frac{u}{u^2+1}.$$

Finally, we return to x with the substitution $u = x+3$. We start with the expression in Equation (8.2):

$$\frac{1}{2}\theta + \frac{1}{4}\sin(2\theta) + C = \frac{1}{2}\tan^{-1}u + \frac{1}{2}\frac{u}{u^2+1} + C$$

$$= \frac{1}{2}\tan^{-1}(x+3) + \frac{x+3}{2(x^2+6x+10)} + C.$$

Stating our final result in one line,

$$\int\frac{1}{(x^2+6x+10)^2}\,dx = \frac{1}{2}\tan^{-1}(x+3) + \frac{x+3}{2(x^2+6x+10)} + C.$$

Our last example returns us to definite integrals, as seen in our first example. Given a definite integral that can be evaluated using Trigonometric Substitution, we could first evaluate the corresponding indefinite integral (by changing from an integral in terms of x to one in terms of θ, then converting back to x) and then evaluate using the original bounds. It is much more straightforward, though, to change the bounds as we substitute.

Example 7   Definite integration and Trigonometric Substitution

Evaluate $\displaystyle\int_0^5\frac{x^2}{\sqrt{x^2+25}}\,dx$.

SOLUTION   Using Key Idea 28(b), we set $x = 5\tan\theta$, $dx = 5\sec^2\theta\,d\theta$, and note that $\sqrt{x^2+25} = 5\sec\theta$. As we substitute, we change the bounds of integration.

The lower bound of the original integral is $x = 0$. As $x = 5\tan\theta$, we solve for θ and find $\theta = \tan^{-1}(x/5)$. Thus the new lower bound is $\theta = \tan^{-1}(0) = 0$. The original upper bound is $x = 5$, thus the new upper bound is $\theta = \tan^{-1}(5/5) = \pi/4$.

Thus we have

$$\int_0^5\frac{x^2}{\sqrt{x^2+25}}\,dx = \int_0^{\pi/4}\frac{25\tan^2\theta}{5\sec\theta}5\sec^2\theta\,d\theta$$

$$= 25\int_0^{\pi/4}\tan^2\theta\sec\theta\,d\theta.$$

Notes:

We encountered this indefinite integral in Example 3 where we found

$$\int \tan^2 \theta \sec \theta \, d\theta = \frac{1}{2}\left(\sec \theta \tan \theta - \ln|\sec \theta + \tan \theta| \right).$$

So

$$25 \int_0^{\pi/4} \tan^2 \theta \sec \theta \, d\theta = \frac{25}{2} \left(\sec \theta \tan \theta - \ln|\sec \theta + \tan \theta| \right)\Big|_0^{\pi/4}$$

$$= \frac{25}{2}\left(\sqrt{2} - \ln(\sqrt{2} + 1) \right).$$

The next section introduces Partial Fraction Decomposition, which is an algebraic technique that turns "complicated" fractions into sums of "simpler" fractions, making integration easier.

Notes:

Exercises 8.3

Terms and Concepts

1. Trigonometric Substitution works on the same principles as Integration by Substitution, though it can feel "_____".

2. If one uses Trigonometric Substitution on an integrand containing $\sqrt{25 - x^2}$, then one should set $x =$_____.

3. Consider the Pythagorean Identity $\sin^2 \theta + \cos^2 \theta = 1$.

    (a) What identity is obtained when both sides are divided by $\cos^2 \theta$?

    (b) Use the new identity to simplify $9 \tan^2 \theta + 9$.

4. Why does Key Idea 28(a) state that $\sqrt{a^2 - x^2} = a \cos \theta$, and not $|a \cos \theta|$?

Problems

In Exercises 5–12, apply Trigonometric Substitution to evaluate the indefinite integrals.

5. $\displaystyle\int \sqrt{x^2 + 1}\, dx$

6. $\displaystyle\int \sqrt{x^2 - 1}\, dx$

7. $\displaystyle\int \sqrt{4x^2 + 1}\, dx$

8. $\displaystyle\int \sqrt{1 - 9x^2}\, dx$

9. $\displaystyle\int \sqrt{16x^2 - 1}\, dx$

10. $\displaystyle\int \frac{8}{\sqrt{x^2 + 2}}\, dx$

11. $\displaystyle\int \frac{3}{\sqrt{7 - x^2}}\, dx$

12. $\displaystyle\int \frac{5}{\sqrt{x^2 - 8}}\, dx$

In Exercises 13–20, evaluate the indefinite integrals. Some may be evaluated without Trigonometric Substitution.

13. $\displaystyle\int \frac{\sqrt{x^2 - 11}}{x}\, dx$

14. $\displaystyle\int \frac{x}{\sqrt{x^2 - 3}}\, dx$

15. $\displaystyle\int \frac{x}{(x^2 + 9)^{3/2}}\, dx$

16. $\displaystyle\int \frac{5x^2}{\sqrt{x^2 - 10}}\, dx$

17. $\displaystyle\int \frac{1}{(x^2 + 4x + 13)^2}\, dx$

18. $\displaystyle\int x^2(1 - x^2)^{-3/2}\, dx$

19. $\displaystyle\int \frac{\sqrt{5 - x^2}}{7x^2}\, dx$

20. $\displaystyle\int \frac{x^2}{\sqrt{x^2 + 3}}\, dx$

In Exercises 21–26, evaluate the definite integrals by making the proper trigonometric substitution *and* changing the bounds of integration.

21. $\displaystyle\int_{-1}^{1} \sqrt{1 - x^2}\, dx$

22. $\displaystyle\int_{4}^{8} \sqrt{x^2 - 16}\, dx$

23. $\displaystyle\int_{0}^{2} \sqrt{x^2 + 4}\, dx$

24. $\displaystyle\int_{-1}^{1} \frac{1}{(x^2 + 1)^2}\, dx$

25. $\displaystyle\int_{-1}^{1} \sqrt{9 - x^2}\, dx$

26. $\displaystyle\int_{-1}^{1} x^2\sqrt{1 - x^2}\, dx$

8.4 Partial Fraction Decomposition

In this section we investigate the antiderivatives of rational functions. Recall that rational functions are functions of the form $f(x) = \frac{p(x)}{q(x)}$, where $p(x)$ and $q(x)$ are polynomials and $q(x) \neq 0$. Such functions arise in many contexts, one of which is the solving of certain fundamental differential equations.

We begin with an example that demonstrates the motivation behind this section. Consider the integral $\int \frac{1}{x^2 - 1} \, dx$. We do not have a simple formula for this (if the denominator were $x^2 + 1$, we would recognize the antiderivative as being the arctangent function). It can be solved using Trigonometric Substitution, but note how the integral is easy to evaluate once we realize:

$$\frac{1}{x^2 - 1} = \frac{1/2}{x - 1} - \frac{1/2}{x + 1}.$$

Thus

$$\int \frac{1}{x^2 - 1} \, dx = \int \frac{1/2}{x - 1} \, dx - \int \frac{1/2}{x + 1} \, dx$$

$$= \frac{1}{2} \ln|x - 1| - \frac{1}{2} \ln|x + 1| + C.$$

This section teaches how to *decompose*

$$\frac{1}{x^2 - 1} \quad \text{into} \quad \frac{1/2}{x - 1} - \frac{1/2}{x + 1}.$$

We start with a rational function $f(x) = \frac{p(x)}{q(x)}$, where p and q do not have any common factors. We first consider the degree of p and q.

- If the $\deg(p) \geq \deg(q)$ then we use polynomial long division to divide q into p to determine a remainder $r(x)$ where $\deg(r) < \deg(q)$. We then write $f(x) = s(x) + \frac{r(x)}{q(x)}$ and apply partial fraction decomposition to $\frac{r(x)}{q(x)}$.

- If the $\deg(p) < \deg(q)$ we can apply partial fraction decomposition to $\frac{p(x)}{q(x)}$ without additional work.

Partial fraction decomposition is based on an algebraic theorem that guarantees that any polynomial, and hence q, can use real numbers to factor into the product of linear and irreducible quadratics factors. The following Key Idea states how to decompose a rational function into a sum of rational functions whose denominators are all of lower degree than q.

An *irreducible quadratic* is one that cannot factor into linear terms with real coefficients.

Notes:

Key Idea 29   **Partial Fraction Decomposition**

Let $\dfrac{p(x)}{q(x)}$ be a rational function, where $\deg(p) < \deg(q)$.

1. **Factor q(x)** : Write $q(x)$ as the product of its linear and irreducible quadratic factors of the form $(ax + b)^m$ and $(ax^2 + bx + c)^n$ where m and n are the highest powers of each factor that divide q.

   - **Linear Terms:** For each linear factor of $q(x)$ the decomposition of $\dfrac{p(x)}{q(x)}$ will contain the following terms:

   $$\frac{A_1}{(ax + b)} + \frac{A_2}{(ax + b)^2} + \cdots \frac{A_m}{(ax + b)^m}$$

   - **Irreducible Quadratic Terms:** For each irreducible quadratic factor of $q(x)$ the decomposition of $\dfrac{p(x)}{q(x)}$ will contain the following terms:

   $$\frac{B_1 x + C_1}{(ax^2 + bx + c)} + \frac{B_2 x + C_2}{(ax^2 + bx + c)^2} + \cdots \frac{B_n x + C_n}{(ax^2 + bx + c)^n}$$

2. **Finding the Coefficients A_i, B_i, and C_i:**

   - Set $\dfrac{p(x)}{q(x)}$ equal to the sum of its linear and irreducible quadratic terms.

   $$\frac{p(x)}{q(x)} = \frac{A_1}{(ax + b)} + \cdots \frac{A_m}{(ax + b)^m} + \frac{B_1 x + C_1}{(ax^2 + bx + c)} + \cdots \frac{B_n x + C_n}{(ax^2 + bx + c)^n}$$

   - Multiply this equation by the factored form of $q(x)$ and simplify to clear the denominators.

   - Solve for the coefficients A_i, B_i, and C_i by

     (a) multiplying out the remaining terms and collecting like powers of x, equating the resulting coefficients and solving the resulting system of linear equations, **or**

     (b) substituting in values for x that eliminate terms so the simplified equation can be solved for a coefficient.

Notes:

Watch the video:
Integration Using method of Partial Fractions at
`https://youtu.be/6qVgHWxdlZ0`

The following examples will demonstrate how to put this Key Idea into practice. In Example 1, we focus on the setting up the decomposition of a rational function.

Example 1 **Decomposing into partial fractions**

Decompose $f(x) = \dfrac{1}{(x+5)(x-2)^3(x^2+x+2)(x^2+x+7)^2}$ without solving for the resulting coefficients.

SOLUTION The denominator is already factored, as both x^2+x+2 and x^2+x+7 are irreducible quadratics. We need to decompose $f(x)$ properly. Since $(x+5)$ is a linear factor that divides the denominator, there will be a

$$\frac{A}{x+5}$$

term in the decomposition.

As $(x-2)^3$ divides the denominator, we will have the following terms in the decomposition:

$$\frac{B}{x-2}, \quad \frac{C}{(x-2)^2} \quad \text{and} \quad \frac{D}{(x-2)^3}.$$

The x^2+x+2 term in the denominator results in a $\dfrac{Ex+F}{x^2+x+2}$ term.

Finally, the $(x^2+x+7)^2$ term results in the terms

$$\frac{Gx+H}{x^2+x+7} \quad \text{and} \quad \frac{Ix+J}{(x^2+x+7)^2}.$$

All together, we have

$$\frac{1}{(x+5)(x-2)^3(x^2+x+2)(x^2+x+7)^2} =$$
$$\frac{A}{x+5}+\frac{B}{x-2}+\frac{C}{(x-2)^2}+\frac{D}{(x-2)^3}+\frac{Ex+F}{x^2+x+2}+\frac{Gx+H}{x^2+x+7}+\frac{Ix+J}{(x^2+x+7)^2}$$

Solving for the coefficients A, B, ..., J would be a bit tedious but not "hard." In the next example we demonstrate solving for the coefficients using both methods given in Key Idea 29.

Notes:

Example 2   **Decomposing into partial fractions**

Perform the partial fraction decomposition of $\dfrac{1}{x^2 - 1}$.

SOLUTION   The denominator can be written as the product of two linear factors: $x^2 - 1 = (x - 1)(x + 1)$. Thus

$$\frac{1}{x^2 - 1} = \frac{A}{x - 1} + \frac{B}{x + 1}. \tag{8.3}$$

Using the method described in Key Idea 29 2(a) to solve for A and B, first multiply through by $x^2 - 1 = (x - 1)(x + 1)$:

$$\begin{aligned}
1 &= \frac{A(x - 1)(x + 1)}{x - 1} + \frac{B(x - 1)(x + 1)}{x + 1} \\
&= A(x + 1) + B(x - 1) \tag{8.4}\\
&= Ax + A + Bx - B \\
&= (A + B)x + (A - B) \qquad \text{collect like terms.}
\end{aligned}$$

The next step is key. For clarity's sake, rewrite the equality we have as

$$0x + 1 = (A + B)x + (A - B).$$

On the left, the coefficient of the x term is 0; on the right, it is $(A + B)$. Since both sides are equal for all values of x, we must have that $0 = A + B$. Likewise, on the left, we have a constant term of 1; on the right, the constant term is $(A - B)$. Therefore we have $1 = A - B$.

We have two linear equations with two unknowns. This one is easy to solve by hand, leading to

$$\begin{array}{ccc}
A + B = 0 & & A = 1/2 \\
A - B = 1 & \Rightarrow & B = -1/2.
\end{array}$$

Thus

$$\frac{1}{x^2 - 1} = \frac{1/2}{x - 1} - \frac{1/2}{x + 1}.$$

Before solving for A and B using the method described in Key Idea 29 2(b), we note that Equations (8.3) and (8.4) are not equivalent. Only the second equation holds for all values of x, including $x = -1$ and $x = 1$, by continuity of polynomials. Thus, we can choose values for x that eliminate terms in the polynomial to solve for A and B.

$$1 = A(x + 1) + B(x - 1).$$

Notes:

If we choose $x = -1$,

$$1 = A(0) + B(-2)$$

$$B = -\frac{1}{2}.$$

Next choose $x = 1$:

$$1 = A(2) + B(0)$$

$$A = \frac{1}{2}.$$

Resulting in the same decomposition as above.

In Example 3, we solve for the decomposition coefficients using the system of linear equations (method 2a). The margin note explains how to solve using substitution (method 2b).

Example 3 **Integrating using partial fractions**

Use partial fraction decomposition to integrate $\displaystyle\int \frac{1}{(x-1)(x+2)^2}\, dx.$

SOLUTION We decompose the integrand as follows, as described by Key Idea 29:

$$\frac{1}{(x-1)(x+2)^2} = \frac{A}{x-1} + \frac{B}{x+2} + \frac{C}{(x+2)^2}. \tag{8.5}$$

To solve for A, B and C, we multiply both sides by $(x-1)(x+2)^2$ and collect like terms:

$$1 = A(x+2)^2 + B(x-1)(x+2) + C(x-1) \tag{8.6}$$
$$= Ax^2 + 4Ax + 4A + Bx^2 + Bx - 2B + Cx - C$$
$$= (A+B)x^2 + (4A+B+C)x + (4A-2B-C)$$

We have

$$0x^2 + 0x + 1 = (A+B)x^2 + (4A+B+C)x + (4A-2B-C)$$

leading to the equations

$$A+B = 0, \quad 4A+B+C = 0 \quad \text{and} \quad 4A-2B-C = 1.$$

These three equations of three unknowns lead to a unique solution:

$$A = 1/9, \quad B = -1/9 \quad \text{and} \quad C = -1/3.$$

Note: Equations (8.5) and (8.6) are not equivalent for $x = 1$ and $x = -2$. However, due to the continuity of polynomials we can let $x = 1$ to simplify the right hand side to $A(1+2)^2 = 9A$. Since the left hand side is still 1, we have $1 = 9A$, so that $A = 1/9$.

Likewise, when $x = -2$; this leads to the equation $1 = -3C$. Thus $C = -1/3$.

Knowing A and C, we can find the value of B by choosing yet another value of x, such as $x = 0$, and solving for B.

Notes:

Thus

$$\int \frac{1}{(x-1)(x+2)^2} \, dx = \int \frac{1/9}{x-1} \, dx + \int \frac{-1/9}{x+2} \, dx + \int \frac{-1/3}{(x+2)^2} \, dx.$$

Each can be integrated with a simple substitution with $u = x-1$ or $u = x+2$. The end result is

$$\int \frac{1}{(x-1)(x+2)^2} \, dx = \frac{1}{9} \ln|x-1| - \frac{1}{9} \ln|x+2| + \frac{1}{3(x+2)} + C.$$

Example 4   **Integrating using partial fractions**

Use partial fraction decomposition to integrate $\displaystyle \int \frac{x^3}{(x-5)(x+3)} \, dx$.

Solution   Key Idea 29 presumes that the degree of the numerator is less than the degree of the denominator. Since this is not the case here, we begin by using polynomial division to reduce the degree of the numerator. We omit the steps, but encourage the reader to verify that

$$\frac{x^3}{(x-5)(x+3)} = x + 2 + \frac{19x+30}{(x-5)(x+3)}.$$

Using Key Idea 29, we can rewrite the new rational function as:

$$\frac{19x+30}{(x-5)(x+3)} = \frac{A}{x-5} + \frac{B}{x+3}$$

for appropriate values of A and B. Clearing denominators, we have

$$19x + 30 = A(x+3) + B(x-5).$$

As in the previous examples we choose values of x to eliminate terms in the polynomial. If we choose $x = -3$,

$$19(-3) + 30 = A(0) + B(-8)$$

$$B = \frac{27}{8}.$$

Next choose $x = 5$:

$$19(5) + 30 = A(8) + B(0)$$

$$A = \frac{125}{8}.$$

Notes:

We can now integrate:

$$\int \frac{x^3}{(x-5)(x+3)}\,dx = \int \left(x + 2 + \frac{125/8}{x-5} + \frac{27/8}{x+3}\right)\,dx$$

$$= \frac{x^2}{2} + 2x + \frac{125}{8}\ln|x-5| + \frac{27}{8}\ln|x+3| + C.$$

Before the next example we remind the reader of a rational integrand evaluated by trigonometric substitution:

$$\int \frac{1}{x^2 + a^2}\,dx = \frac{1}{a}\tan^{-1}\left(\frac{x}{a}\right) + C.$$

Example 5   **Integrating using partial fractions**

Use partial fraction decomposition to evaluate $\displaystyle\int \frac{7x^2 + 31x + 54}{(x+1)(x^2 + 6x + 11)}\,dx$.

SOLUTION   The degree of the numerator is less than the degree of the denominator so we begin by applying Key Idea 29. We have:

$$\frac{7x^2 + 31x + 54}{(x+1)(x^2 + 6x + 11)} = \frac{A}{x+1} + \frac{Bx + C}{x^2 + 6x + 11}.$$

Now clear the denominators.

$$7x^2 + 31x + 54 = A(x^2 + 6x + 11) + (Bx + C)(x + 1).$$

Again, we choose values of x to eliminate terms in the polynomial. If we choose $x = -1$,

$$30 = 6A + (-B + C)(0)$$
$$A = 5.$$

Although none of the other terms can be zeroed out, we continue by letting $A = 5$ and substituting helpful values of x. Choosing $x = 0$, we notice

$$54 = 55 + C$$
$$C = -1.$$

Finally, choose $x = 1$ (any value other than -1 and 0 can be used, 1 is easy to work with)

$$92 = 90 + (B - 1)(2)$$
$$B = 2.$$

Notes:

Thus

$$\int \frac{7x^2 + 31x + 54}{(x+1)(x^2 + 6x + 11)}\, dx = \int \left(\frac{5}{x+1} + \frac{2x-1}{x^2 + 6x + 11} \right)\, dx.$$

The first term of this new integrand is easy to evaluate; it leads to a $5 \ln |x+1|$ term. The second term is not hard, but takes several steps and uses substitution techniques.

The integrand $\dfrac{2x-1}{x^2 + 6x + 11}$ has a quadratic in the denominator and a linear term in the numerator. This leads us to try substitution. Let $u = x^2 + 6x + 11$, so $du = (2x + 6)\, dx$. The numerator is $2x - 1$, not $2x + 6$, but we can get a $2x + 6$ term in the numerator by adding 0 in the form of "$7 - 7$."

$$\frac{2x-1}{x^2 + 6x + 11} = \frac{2x - 1 + 7 - 7}{x^2 + 6x + 11}$$

$$= \frac{2x+6}{x^2 + 6x + 11} - \frac{7}{x^2 + 6x + 11}.$$

We can now integrate the first term with substitution, leading to a $\ln \left| x^2 + 6x + 11 \right|$ term. The final term can be integrated using arctangent. First, complete the square in the denominator:

$$\frac{7}{x^2 + 6x + 11} = \frac{7}{(x+3)^2 + 2}.$$

An antiderivative of the latter term can be found using Key Idea 28 and substitution:

$$\int \frac{7}{x^2 + 6x + 11}\, dx = \frac{7}{\sqrt{2}} \tan^{-1} \left(\frac{x+3}{\sqrt{2}} \right) + C.$$

Let's start at the beginning and put all of the steps together.

$$\int \frac{7x^2 + 31x + 54}{(x+1)(x^2 + 6x + 11)}\, dx$$

$$= \int \left(\frac{5}{x+1} + \frac{2x-1}{x^2 + 6x + 11} \right)\, dx$$

$$= \int \frac{5}{x+1}\, dx + \int \frac{2x+6}{x^2 + 6x + 11}\, dx - \int \frac{7}{x^2 + 6x + 11}\, dx$$

$$= 5 \ln |x+1| + \ln \left| x^2 + 6x + 11 \right| - \frac{7}{\sqrt{2}} \tan^{-1} \left(\frac{x+3}{\sqrt{2}} \right) + C.$$

As with many other problems in calculus, it is important to remember that one is not expected to "see" the final answer immediately after seeing the problem. Rather, given the initial problem, we break it down into smaller problems that are easier to solve. The final answer is a combination of the answers of the smaller problems.

Notes:

Partial Fraction Decomposition is an important tool when dealing with rational functions. Note that at its heart, it is a technique of algebra, not calculus, as we are rewriting a fraction in a new form. Regardless, it is very useful in the realm of calculus as it lets us evaluate a certain set of "complicated" integrals. The next section will require the reader to determine an appropriate method for evaluating a variety of integrals.

Notes:

Exercises 8.4

Terms and Concepts

1. Fill in the blank: Partial Fraction Decomposition is a method of rewriting _____ functions.

2. T/F: It is sometimes necessary to use polynomial division before using Partial Fraction Decomposition.

3. Decompose $\dfrac{1}{x^2 - 3x}$ without solving for the coefficients, as done in Example 1.

4. Decompose $\dfrac{7 - x}{x^2 - 9}$ without solving for the coefficients, as done in Example 1.

5. Decompose $\dfrac{x - 3}{x^2 - 7}$ without solving for the coefficients, as done in Example 1.

6. Decompose $\dfrac{2x + 5}{x^3 + 7x}$ without solving for the coefficients, as done in Example 1.

Problems

In Exercises 7–32, evaluate the indefinite integral.

7. $\displaystyle\int \dfrac{7x + 7}{x^2 + 3x - 10}\, dx$

8. $\displaystyle\int \dfrac{7x - 2}{x^2 + x}\, dx$

9. $\displaystyle\int \dfrac{-4}{3x^2 - 12}\, dx$

10. $\displaystyle\int \dfrac{x + 7}{(x + 5)^2}\, dx$

11. $\displaystyle\int \dfrac{-3x - 20}{(x + 8)^2}\, dx$

12. $\displaystyle\int \dfrac{9x^2 + 11x + 7}{x(x + 1)^2}\, dx$

13. $\displaystyle\int \dfrac{-12x^2 - x + 33}{(x - 1)(x + 3)(3 - 2x)}\, dx$

14. $\displaystyle\int \dfrac{94x^2 - 10x}{(7x + 3)(5x - 1)(3x - 1)}\, dx$

15. $\displaystyle\int \dfrac{x^2 + x + 1}{x^2 + x - 2}\, dx$

16. $\displaystyle\int \dfrac{x^3}{x^2 - x - 20}\, dx$

17. $\displaystyle\int \dfrac{2x^2 - 4x + 6}{x^2 - 2x + 3}\, dx$

18. $\displaystyle\int \dfrac{1}{x^3 + 2x^2 + 3x}\, dx$

19. $\displaystyle\int \dfrac{dx}{x^4 - x^2}$

20. $\displaystyle\int \dfrac{x^2 + x + 5}{x^2 + 4x + 10}\, dx$

21. $\displaystyle\int \dfrac{12x^2 + 21x + 3}{(x + 1)(3x^2 + 5x - 1)}\, dx$

22. $\displaystyle\int \dfrac{6x^2 + 8x - 4}{(x - 3)(x^2 + 6x + 10)}\, dx$

23. $\displaystyle\int \dfrac{1 - x + 2x^2 - x^3}{x(x^2 + 1)^2}\, dx$

24. $\displaystyle\int \dfrac{2x^2 + x + 1}{(x + 1)(x^2 + 9)}\, dx$

25. $\displaystyle\int \dfrac{x^2 - 20x - 69}{(x - 7)(x^2 + 2x + 17)}\, dx$

26. $\displaystyle\int \dfrac{x^3 + x^2 + 2x + 1}{(x^2 + 1)(x^2 + 2)}\, dx$

27. $\displaystyle\int \dfrac{x}{x^4 + 4x^2 + 3}\, dx$

28. $\displaystyle\int \dfrac{x - 3}{(x^2 + 2x + 4)^2}\, dx$

29. $\displaystyle\int \dfrac{9x^2 - 60x + 33}{(x - 9)(x^2 - 2x + 11)}\, dx$

30. $\displaystyle\int \dfrac{6x^2 + 45x + 121}{(x + 2)(x^2 + 10x + 27)}\, dx$

31. $\displaystyle\int \dfrac{1}{x^4 - 16}\, dx$

32. $\displaystyle\int \dfrac{1}{x^2 + x}\, dx$

In Exercises 33–36, evaluate the definite integral.

33. $\displaystyle\int_1^2 \dfrac{8x + 21}{(x + 2)(x + 3)}\, dx$

34. $\displaystyle\int_0^5 \dfrac{14x + 6}{(3x + 2)(x + 4)}\, dx$

35. $\displaystyle\int_{-1}^1 \dfrac{x^2 + 5x - 5}{(x - 10)(x^2 + 4x + 5)}\, dx$

36. $\displaystyle\int_0^1 \dfrac{x}{(x + 1)(x^2 + 2x + 1)}\, dx$

8.5   Integration Strategies

We've now seen a fair number of different integration techniques and so we should probably pause at this point to talk a little bit about a strategy to use for determining the correct technique to use when faced with an integral.

There are a couple of points that need to be made about this strategy. First, it isn't a hard and fast set of rules for determining the method that should be used. It is really nothing more than a general set of guidelines that will help us to identify techniques that may work. Some integrals can be done in more than one way and so depending on the path you take through the strategy you may end up with a different technique than someone else who also went through this strategy.

Second, while the strategy is presented as a way to identify the technique that could be used on an integral keep in mind that, for many integrals, it can also automatically exclude certain techniques as well. When going through the strategy keep two lists in mind. The first list is integration techniques that simply won't work and the second list is techniques that look like they might work. After going through the strategy, if the second list has only one entry then that is the technique to use. If on the other hand, there is more than one possible technique to use we will have to decide on which is liable to be the best for us to use. Unfortunately there is no way to teach which technique is the best as that usually depends upon the person and which technique they find to be the easiest.

Third, don't forget that many integrals can be evaluated in multiple ways and so more than one technique may be used on it. This has already been mentioned in each of the previous points, but is important enough to warrant a separate mention. Sometimes one technique will be significantly easier than the others and so don't just stop at the first technique that appears to work. Always identify all possible techniques and then go back and determine which you feel will be the easiest for you to use.

Next, it's entirely possible that you will need to use more than one method to completely evaluate an integral. For instance a substitution may lead to using integration by parts or partial fractions integral.

Notes:

> **Key Idea 30** **Guidelines for Choosing an Integration Strategy**
>
> 1. Simplify the integrand, if possible
>
> 2. See if a "simple" substitution will work
>
> 3. Identify the type of integral
>
> 4. Relate the integral to an integral we already know how to do
>
> 5. Try multiple techniques
>
> 6. Try again

Let's expand on the ideas of the previous Key Idea.

1. **Simplify the integrand, if possible.** This step is very important in the integration process. Many integrals can be taken from very difficult to very easy with a little simplification or manipulation. Don't forget basic trigonometric and algebraic identities as these can often be used to simplify the integral.

   We used this idea when we were looking at integrals involving trigonometric functions. For example consider the following integral.

   $$\int \cos^2 x \, dx$$

   the integral can't be done as it is, however by recalling the identity,

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

   the integral becomes very easy to do.

   Note that this example also shows that simplifications does not necessarily mean that we'll write the integrand in a "simpler" form. It ony means that we'll write the integrand into a form that we can deal with and this is often longer and/or "messier" than the original integral.

2. **See if a "simple" substitution will work.** Look to see if a simple substitution can be used instead of the often more complicated methods from Calculus II. For example consider both of the following integrals.

   $$\int \frac{x}{x^2 - 1} \, dx \qquad \int x\sqrt{x^2 - 1} \, dx$$

Notes:

The first integral can be done with the method of partial fractions and the second could be done with a trigonometric substitution.

However, both could also be evaluated using the substitution $u = x^2 - 1$ and the work involved in the substitution would be significantly less than the work involved in either partial fractions or trigonometric substitution.

So, always look for quick, simple substitutions before moving on to the more complicated Calculus II techniques.

3. **Identify the type of integral.** Note that any integral may fall into more than one of these types. Because of this fact it's usually best to go all the way through the list and identify all possible types since one may be easier than the other and it's entirely possible that the easier type is listed lower in the list.

   (a) Is the integrand a rational expression (i.e. is the integrand a polynomial divided by a polynomial)? If so then partial fractions (Section 8.4) may work on the integral.

   (b) Is the integrand a polynomial times a trigonometric function, exponential, or logarithm? If so, then integration by parts (Section 8.1) may work.

   (c) Is the integrand a product of sines and cosines, secants and tangents, or cosecants and cotangents? If so, then the topics from Section 8.2 may work. Likewise, don't forget that some quotients involving these functions can also be done using these techniques.

   (d) Does the integrand involve $\sqrt{b^2x^2 + a^2}$, $\sqrt{b^2x^2 - a^2}$, or $\sqrt{a^2 - b^2x^2}$? If so, then a trigonometric substitution (Section 8.3) might work nicely.

   (e) Does the integrand have roots other than those listed above in it? If so then the substitution $u = \sqrt[n]{g(x)}$ might work.

   (f) Does the integrand have a quadratic in it? If so then completing the square on the quadratic might put it into a form that we can deal with.

4. **Relate the integral to an integral we already know how to do.** In other words, can we use a substitution or manipulation to write the integrand into a form that does fit into the forms we've looked at previously in this chapter. A typical example is the following integral.

$$\int \cos x \sqrt{1 + \sin^2 x}\, dx$$

This integral doesn't obviously fit into any of the forms we looked at in this chapter. However, with the substitution $u = \sin x$ we can reduce the

Notes:

integral to the form

$$\int \sqrt{1 + u^2}\, dx$$

which is a trigonometric substitution problem.

5. **Try multiple techniques.** In this step we need to ask ourselves if it is possible that we'll need to use multiple techniques. The example in the previous part is a good example. Using a substitution didn't allow us to actually do the integral. All it did was put the integral into a form that we could use a different technique on.

   Don't ever get locked into the idea that an integral will only required one step to completely evaluate it. Many will require more than one step.

6. **Try again.** If everything that you've tried to this point doesn't work then go back through the process again. This time try a technique that you didn't use the first time around.

As noted above, this strategy is not a hard and fast set of rules. It is only intended to guide you through the process of best determining how to do any given integral. Note as well that the only place Calculus II actually arises is the third step. Steps 1, 2, and 4 involve nothing more than manipulation of the integrand either through direct manipulation of the integrand or by using a substitution. The last two steps are simply ideas to think about in going through this strategy.

Many students go through this process and concentrate almost exclusively on Step 3 (after all this is Calculus II, so it's easy to see why they might do that...) to the exclusion of the other steps. One very large consequence of that exclusion is that often a simple manipulation or substitution is overlooked that could make the integral very easy to do.

Before moving on to the next section we will work a couple of examples illustrating a couple of not so obvious simplifications/manipulations and a not so obvious substitution.

Example 1   Strategies of Integration
Evaluate the integral

$$\int \frac{\tan x}{\sec^4 x}\, dx$$

SOLUTION   This integral almost falls into the form given in 3c. It is a quotient of tangent and secant and we know that sometimes we can use the same methods for products of tangents and secants on quotients.

The process from Section 8.2 tells us that if we have even powers of secant to save two of them and convert the rest to tangents. That won't work here. We

---

Notes:

can save two secants, but they would be in the denominator and they won't do us any good here. Remember that the point of saving them is so they could be there for the substitution $u = \tan x$. That requires them to be in the numerator. So, that won't work. We need to find another solution method.

There are in fact two solution methods to this integral depending on how you want to go about it.

Solution 1   In this solution method we could just convert everything to sines and cosines and see if that gives us an integral we can deal with.

$$\int \frac{\tan x}{\sec^4 x}\, dx = \int \frac{\sin x}{\cos x} \cos^4 x\, dx$$

$$= \int \sin x \cos^3 x\, dx \qquad \text{substitute } u = \cos x$$

$$= -\int u^3\, du$$

$$= -\frac{1}{4} \cos^4 x + C$$

Note that just converting to sines and cosines won't always work and if it does it won't always work this nicely. Often there will be a lot more work that would need to be done to complete the integral.

Solution 2   This solution method goes back to dealing with secants and tangents. Let's notice that if we had a secant in the numerator we could just use $u = \sec x$ as a substitution and it would be a fairly quick and simple substitution to use. We don't have a secant in the numerator. However we could very easily get a secant in the numerator by multiplying the numerator and denominator by secant (i.e. we multiply the integrand by "1").

$$\int \frac{\tan x}{\sec^4 x}\, dx = \int \frac{\tan x \sec x}{\sec^5 x}\, dx \qquad \text{substitute } u = \sec x$$

$$= \int \frac{1}{u^5}\, du$$

$$= -\frac{1}{4} \frac{1}{\sec^4 x} + C$$

$$= -\frac{1}{4} \cos^4 x + C$$

In the previous example we saw two "simplifications" that allowed us to evaluate the integral. The first was using identities to rewrite the integral into terms we could deal with and the second involved multiplying the numerator and denominator by something to again put the integral into terms we could deal with.

Notes:

Using identities to rewrite an integral is an important "simplification" and we should not forget about it. Integrals can often be greatly simplified or at least put into a form that can be dealt with by using an identity.

The second "simplification" is not used as often, but does show up on occasion so again, it's best to remember it. In fact, let's take another look at an example in which multiplying the integrand by "1" will allow us to evaluate an integral.

Example 2   **Strategy for Integration**
Evaluate the integral

$$\int \frac{1}{1 + \sin x}\, dx$$

SOLUTION   This is an integral which if we just concentrate on the third step we won't get anywhere. This integral doesn't appear to be any of the kinds of integrals that we worked on in this chapter. We can evaluate the integral however, if we do the following,

$$\int \frac{1}{1 + \sin x}\, dx = \int \frac{1}{1 + \sin x} \frac{1 - \sin x}{1 - \sin x}\, dx$$
$$= \int \frac{1 - \sin x}{1 - \sin^2 x}\, dx$$

This does not appear to have done anything for us. However, if we now remember the first "simplification" we looked at above we will notice that we can use an identity to rewrite the denominator. Once we do that we can further manipulate the integrand into something we can evaluate.

$$\int \frac{1}{1 + \sin x}\, dx = \int \frac{1 - \sin x}{\cos^2 x}\, dx$$
$$= \int \frac{1}{\cos^2 x} - \frac{\sin x}{\cos x} \frac{1}{\cos x}\, dx$$
$$= \int \sec^2 x - \tan x \sec x\, dx$$
$$= \tan x - \sec x + C$$

So, we've just seen once again that multiplying by a helpful form of "1" can put the integral into a form we can integrate. Notice as well that this example also showed that "simplifications" do not necessarily put an integral into a simpler form. They only put the integrand into a form that is easier to integrate.

Let's now take a quick look at an example of a substitution that is not so obvious.

Notes:

Example 3   **Strategy for Integration**

Evaluate the integral

$$\int \cos \sqrt{x} \, dx$$

SOLUTION   We introduced this integral by saying that the substitution was not so obvious. However, this is really an integral that falls into the form given by 3e in Key Idea 30. Many people miss that form and so don't think about it. So, let's try the following substitution.

$$u = \sqrt{x} \qquad x = u^2 \qquad dx = 2u \, du$$

With this substitution the integral becomes,

$$\int \cos \sqrt{x} \, dx = 2 \int u \cos u \, du$$

This is now an integration by parts. Remember that often we will need to use more than one technique to completely do the integral. This is a fairly simple integration by parts problem so we'll leave the remainder of the details for you to check.

$$\int \cos \sqrt{x} \, dx = 2(\cos \sqrt{x} + \sqrt{x} \sin \sqrt{x}) + C.$$

It will be possible to integrate every integral assigned in this class, but it is important to note that there are integrals that just can't be evaluated. We should also note that after we look at series in Chapter 9 we will be able to write down a series representation of many of these types of integrals.

Notes:

Exercises 8.5

Problems

In Exercises 1–51, compute the indefinite integral.

1. $\displaystyle\int \sin^{-1} x \, dx$

2. $\displaystyle\int \cos^3 2x \, \sin^2 2x \, dx$

3. $\displaystyle\int \frac{4x^2 - 12x - 10}{(x-2)(x^2 - 4x + 3)} \, dx$

4. $\displaystyle\int \tan x \, \sec^5 x \, dx$

5. $\displaystyle\int \frac{1}{(x^2 + 25)^{3/2}} \, dx$

6. $\displaystyle\int \frac{\sqrt{4 - x^2}}{x} \, dx$

7. $\displaystyle\int \frac{x^3 + 1}{x(x-1)^3} \, dx$

8. $\displaystyle\int \frac{x}{\sqrt{4 + 4x - x^2}} \, dx$

9. $\displaystyle\int x^3 e^{x^2} \, dx$

10. $\displaystyle\int \frac{\sqrt[3]{x + 8}}{x} \, dx$

11. $\displaystyle\int e^{2x} \, \sin^2 3x \, dx$

12. $\displaystyle\int \cos^3 x \, \sin^3 x \, dx$

13. $\displaystyle\int \frac{x}{\sqrt{4 - x^2}} \, dx$

14. $\displaystyle\int \frac{x^5 - x^3 + 1}{x^3 + 2x^2} \, dx$

15. $\displaystyle\int \frac{1}{x^{3/2} + x^{1/2}} \, dx$

16. $\displaystyle\int e^x \, \sec e^x \, dx$

17. $\displaystyle\int x^2 \, \sin 3x \, dx$

18. $\displaystyle\int \sin^3 x \sqrt{\cos x} \, dx$

19. $\displaystyle\int e^x \sqrt{e^x + 1} \, dx$

20. $\displaystyle\int \frac{x^2}{\sqrt{4x^2 + 9}} \, dx$

21. $\displaystyle\int \sec^2 x \, \tan^2 x \, dx$

22. $\displaystyle\int x \csc x \, \cot x \, dx$

23. $\displaystyle\int x^2 (8 - x^3)^{1/3} \, dx$

24. $\displaystyle\int \sin \sqrt{x} \, dx$

25. $\displaystyle\int x\sqrt{3 - 2x} \, dx$

26. $\displaystyle\int \frac{e^{3x}}{1 + e^x} \, dx$

27. $\displaystyle\int \frac{x^2 - 4x + 3}{\sqrt{x}} \, dx$

28. $\displaystyle\int \frac{x^3}{\sqrt{16 - x^2}} \, dx$

29. $\displaystyle\int \frac{1 - 2x}{x^2 + 12x + 35} \, dx$

30. $\displaystyle\int \tan^{-1} 5x \, dx$

31. $\displaystyle\int \frac{e^{\tan x}}{\cos^2 x} \, dx$

32. $\displaystyle\int \frac{1}{\sqrt{7 + 5x^2}} \, dx$

33. $\displaystyle\int \cot^6 x \, dx$

34. $\displaystyle\int x^3 \sqrt{x^2 - 25} \, dx$

35. $\displaystyle\int (x^2 - \operatorname{sech}^2 4x) \, dx$

36. $\displaystyle\int x^2 e^{-4x} \, dx$

37. $\displaystyle\int \frac{3}{\sqrt{11 - 10x - x^2}} \, dx$

38. $\displaystyle\int \frac{x^3 - 20x^2 - 63x - 198}{x^4 - 1} \, dx$

39. $\displaystyle\int \tan 7x \, \cos 7x \, dx$

40. $\displaystyle\int \tan^3 x \, \sec x \, dx$

41. $\displaystyle\int (x^3 + 1) \, \cos x \, dx$

42. $\displaystyle\int \frac{\sqrt{9 - 4x^2}}{x^2} \, dx$

43. $\displaystyle\int (x - \cot 3x)^2 \, dx$

44. $\displaystyle\int \frac{1}{x(\sqrt{x} - \sqrt[4]{x})} \, dx$

45. $\displaystyle\int \frac{\sin x}{\sqrt{1 + \cos x}} \, dx$

46. $\displaystyle\int \frac{x^2}{(25 + x^2)^2} \, dx$

47. $\displaystyle\int \frac{2x^3 + 4x^2 + 10x + 13}{x^4 + 9x^2 + 20} \, dx$

48. $\displaystyle\int \frac{(x^2 - 2)^2}{x} \, dx$

49. $\displaystyle\int x^{3/2} \ln x \, dx$

50. $\displaystyle\int \frac{x^2}{\sqrt[3]{2x + 3}} \, dx$

51. $\displaystyle\int \frac{x e^x}{(x + 1)^2} \, dx$

8.6 Improper Integration

We begin this section by considering the following definite integrals:

- $\displaystyle\int_0^{100} \frac{1}{1+x^2}\, dx \approx 1.5608,$

- $\displaystyle\int_0^{1000} \frac{1}{1+x^2}\, dx \approx 1.5698,$

- $\displaystyle\int_0^{10,000} \frac{1}{1+x^2}\, dx \approx 1.5707.$

Notice how the integrand is $1/(1+x^2)$ in each integral (which is sketched in Figure 8.6). As the upper bound gets larger, one would expect the "area under the curve" would also grow. While the definite integrals do increase in value as the upper bound grows, they are not increasing by much. In fact, consider:

$$\int_0^b \frac{1}{1+x^2}\, dx = \tan^{-1}x \Big|_0^b = \tan^{-1}b - \tan^{-1}0 = \tan^{-1}b.$$

As $b \to \infty$, $\tan^{-1}b \to \pi/2$. Therefore it seems that as the upper bound b grows, the value of the definite integral $\displaystyle\int_0^b \frac{1}{1+x^2}\, dx$ approaches $\pi/2 \approx 1.5708$. This should strike the reader as being a bit amazing: even though the curve extends "to infinity," it has a finite amount of area underneath it.

When we defined the definite integral $\displaystyle\int_a^b f(x)\, dx$, we made two stipulations:

1. The interval over which we integrated, $[a, b]$, was a finite interval, and

2. The function $f(x)$ was continuous on $[a, b]$ (ensuring that the range of f was finite).

In this section we consider integrals where one or both of the above conditions do not hold. Such integrals are called **improper integrals.**
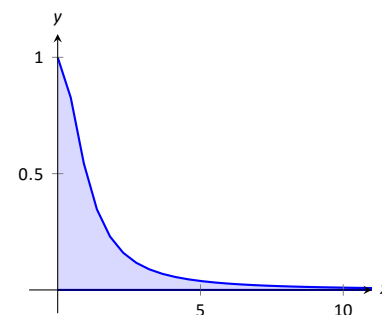
Figure 8.6: Graphing $f(x) = \dfrac{1}{1+x^2}$.

Notes:

Improper Integrals with Infinite Bounds

Definition 28   **Improper Integrals with Infinite Bounds**

1. Let f be a continuous function on $[a, \infty)$. For $t \geq a$ let

$$\int_a^\infty f(x) \, dx = \lim_{t \to \infty} \int_a^t f(x) \, dx.$$

2. Let f be a continuous function on $(-\infty, b]$. For $t \leq b$ let

$$\int_{-\infty}^b f(x) \, dx = \lim_{t \to -\infty} \int_t^b f(x) \, dx.$$

3. Let f be a continuous function on $(-\infty, \infty)$. For any real number c (which one doesn't matter), let

$$\int_{-\infty}^\infty f(x) \, dx = \lim_{a \to -\infty} \int_a^c f(x) \, dx \; + \; \lim_{b \to \infty} \int_c^b f(x) \, dx.$$

An improper integral is said to **converge** if its corresponding limit exists; otherwise, it **diverges**. The improper integral in part 3 converges if and only if both of its limits exist.

Watch the video:
Improper Integral — Infinity in Upper and Lower Limits at
`https://youtu.be/f6cGotvktxs`

Example 1   **Evaluating improper integrals**
Evaluate the following improper integrals.

1. $\displaystyle\int_1^\infty \frac{1}{x^2} \, dx$

2. $\displaystyle\int_1^\infty \frac{1}{x} \, dx$

3. $\displaystyle\int_{-\infty}^0 e^x \, dx$

4. $\displaystyle\int_{-\infty}^\infty \frac{1}{1+x^2} \, dx$

Notes:

SOLUTION

1.
$$\int_1^\infty \frac{1}{x^2}\, dx = \lim_{t\to\infty} \int_1^t \frac{1}{x^2}\, dx$$
$$= \lim_{t\to\infty} \left. \frac{-1}{x} \right|_1^t$$
$$= \lim_{t\to\infty} \frac{-1}{t} + 1$$
$$= 1.$$

A graph of the area defined by this integral is given in Figure 8.7.

2.
$$\int_1^\infty \frac{1}{x}\, dx = \lim_{t\to\infty} \int_1^t \frac{1}{x}\, dx$$
$$= \lim_{t\to\infty} \left. \ln|x| \right|_1^t$$
$$= \lim_{t\to\infty} \ln(t)$$
$$= \infty.$$

The limit does not exist, hence the improper integral $\int_1^\infty \frac{1}{x}\, dx$ diverges.

Compare the graphs in Figures 8.7 and 8.8; notice how the values of $f(x) = 1/x$ are noticeably larger than those of $f(x) = 1/x^2$. This difference is enough to cause the improper integral to diverge.

3.
$$\int_{-\infty}^0 e^x\, dx = \lim_{t\to-\infty} \int_t^0 e^x\, dx$$
$$= \lim_{t\to-\infty} \left. e^x \right|_t^0$$
$$= \lim_{t\to-\infty} e^0 - e^t$$
$$= 1.$$

A graph of the area defined by this integral is given in Figure 8.9.
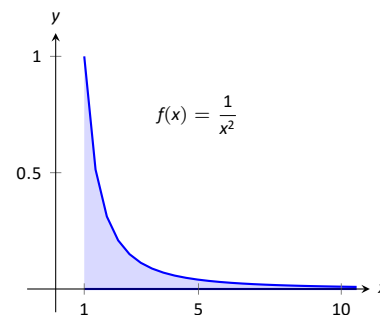
Figure 8.7: A graph of $f(x) = \frac{1}{x^2}$ in Example 1 part 1.
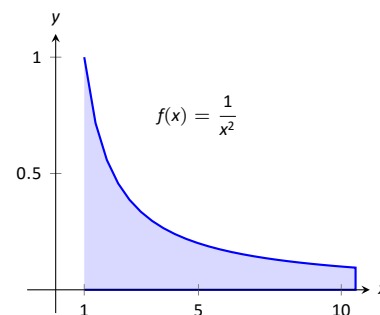
Figure 8.8: A graph of $f(x) = \frac{1}{x}$ in Example 1 part 2.
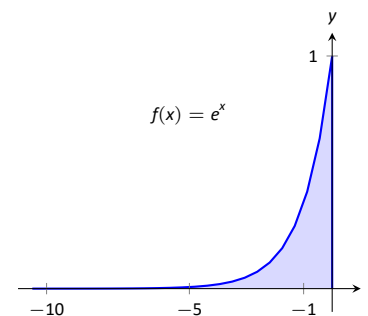
Figure 8.9: A graph of $f(x) = e^x$ in Example 1 part 3.

Notes:

433
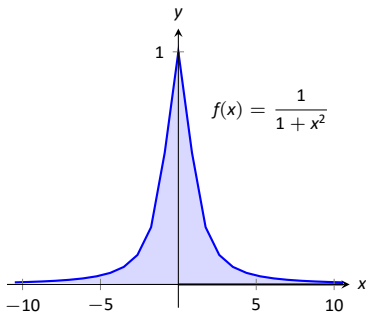
Figure 8.10: A graph of $f(x) = \frac{1}{1+x^2}$ in Example 1 part 4.

4. We will need to break this into two improper integrals and choose a value of c as in part 3 of Definition 28. Any value of c is fine; we choose $c = 0$.

$$\int_{-\infty}^{\infty} \frac{1}{1+x^2}\, dx = \lim_{t\to-\infty}\int_{t}^{0}\frac{1}{1+x^2}\, dx + \lim_{t\to\infty}\int_{0}^{t}\frac{1}{1+x^2}\, dx$$

$$= \lim_{t\to-\infty} \tan^{-1}x\Big|_{t}^{0} + \lim_{t\to\infty}\tan^{-1}x\Big|_{0}^{t}$$

$$= \lim_{t\to-\infty}\left(\tan^{-1}0 - \tan^{-1}t\right) + \lim_{t\to\infty}\left(\tan^{-1}t - \tan^{-1}0\right)$$

$$= \left(0 - \frac{-\pi}{2}\right) + \left(\frac{\pi}{2} - 0\right).$$

$$= \pi.$$

A graph of the area defined by this integral is given in Figure 8.10.

Section 7.5 introduced L'Hôpital's Rule, a method of evaluating limits that return indeterminate forms. It is not uncommon for the limits resulting from improper integrals to need this rule as demonstrated next.

Example 2   **Improper integration and L'Hôpital's Rule**
Evaluate the improper integral $\displaystyle\int_{1}^{\infty} \frac{\ln x}{x^2}\, dx$.

SOLUTION   This integral will require the use of Integration by Parts. Let $u = \ln x$ and $dv = 1/x^2\, dx$. Then

$$\int_{1}^{\infty} \frac{\ln x}{x^2}\, dx = \lim_{t\to\infty}\int_{1}^{t}\frac{\ln x}{x^2}\, dx$$

$$= \lim_{t\to\infty}\left(-\frac{\ln x}{x}\Big|_{1}^{t} + \int_{1}^{t}\frac{1}{x^2}\, dx\right)$$

$$= \lim_{t\to\infty}\left(-\frac{\ln x}{x} - \frac{1}{x}\right)\Big|_{1}^{t}$$

$$= \lim_{t\to\infty}\left(-\frac{\ln t}{t} - \frac{1}{t} - (-\ln 1 - 1)\right).$$

Figure 8.11: A graph of $f(x) = \frac{\ln x}{x^2}$ in Example 2.

The $1/t$ goes to 0, and $\ln 1 = 0$, leaving $\displaystyle\lim_{t\to\infty}\frac{\ln t}{t}$ with L'Hôpital's Rule. We have:

$$\lim_{t\to\infty}\frac{\ln t}{t} \overset{\text{by LHR}}{=} \lim_{t\to\infty}\frac{1/t}{1} = 0.$$

Notes:

Thus the improper integral evaluates as:

$$\int_1^\infty \frac{\ln x}{x^2}\, dx = 1.$$

Improper Integrals with Infinite Range

We have just considered definite integrals where the interval of integration was infinite. We now consider another type of improper integration, where the range of the integrand is infinite.

Definition 29 Improper Integration with Infinite Range
Let $f(x)$ be a continuous function on $[a, b]$ except at c, $a \le c \le b$, where $x = c$ is a vertical asymptote of f. Define

$$\int_a^b f(x)\, dx = \lim_{t\to c^-} \int_a^t f(x)\, dx + \lim_{t\to c^+} \int_t^b f(x)\, dx.$$

Note that c can be one of the endpoints (a or b). In that case, there is only one limit to consider as part of the definition.

Example 3 Improper integration of functions with infinite range
Evaluate the following improper integrals:

$$1. \int_0^1 \frac{1}{\sqrt{x}}\, dx \qquad 2. \int_{-1}^1 \frac{1}{x^2}\, dx.$$

SOLUTION

1. A graph of $f(x) = 1/\sqrt{x}$ is given in Figure 8.12. Notice that f has a vertical asymptote at $x = 0$. In some sense, we are trying to compute the area of a region that has no "top." Could this have a finite value?

$$\int_0^1 \frac{1}{\sqrt{x}}\, dx = \lim_{t\to 0^+} \int_t^1 \frac{1}{\sqrt{x}}\, dx$$
$$= \lim_{t\to 0^+} 2\sqrt{x}\Big|_t^1$$
$$= \lim_{t\to 0^+} 2\left(\sqrt{1} - \sqrt{t}\right)$$
$$= 2.$$
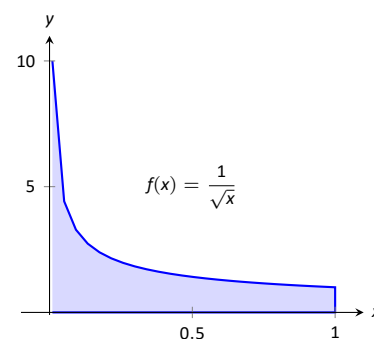
Figure 8.12: A graph of $f(x) = \frac{1}{\sqrt{x}}$ in Example 3.

Notes:

It turns out that the region does have a finite area even though it has no upper bound (strange things can occur in mathematics when considering the infinite).

2. The function $f(x) = 1/x^2$ has a vertical asymptote at $x = 0$, as shown in Figure 8.13, so this integral is an improper integral. Let's eschew using limits for a moment and proceed without recognizing the improper nature of the integral. This leads to:

$$\int_{-1}^{1} \frac{1}{x^2}\, dx = -\frac{1}{x}\Big|_{-1}^{1}$$
$$= -1 - (1)$$
$$= -2.$$

Clearly the area in question is above the x-axis, yet the area is supposedly negative. In this example we noted the discontinuity of the integrand on $[-1, 1]$ (its improper nature) but continued anyway to apply the Fundamental Theorem of Calculus. Violating the hypothesis of the FTC led us to an incorrect area of -2. If we now evaluate the integral using Definition 29 we will see that the area is unbounded.

$$\int_{-1}^{1} \frac{1}{x^2}\, dx = \lim_{t \to 0^-} \int_{-1}^{t} \frac{1}{x^2}\, dx + \lim_{t \to 0^+} \int_{t}^{1} \frac{1}{x^2}\, dx$$
$$= \lim_{t \to 0^-} -\frac{1}{x}\Big|_{-1}^{t} + \lim_{t \to 0^+} -\frac{1}{x}\Big|_{t}^{1}$$
$$= \lim_{t \to 0^-} -\frac{1}{t} + 1 + \lim_{t \to 0^+} -1 + \frac{1}{t}.$$

Neither limit converges hence the original improper integral diverges. The nonsensical answer we obtained by ignoring the improper nature of the integral is just that: nonsensical.

Figure 8.13: A graph of $f(x) = \frac{1}{x^2}$ in Example 3.

Understanding Convergence and Divergence

Oftentimes we are interested in knowing simply whether or not an improper integral converges, and not necessarily the value of a convergent integral. We provide here several tools that help determine the convergence or divergence of improper integrals without integrating.

Our first tool is knowing the behavior of functions of the form $\dfrac{1}{x^p}$.

Notes:

Example 4   **Improper integration of** $1/x^p$

Determine the values of p for which $\displaystyle\int_1^\infty \frac{1}{x^p}\, dx$ converges.

SOLUTION   We begin by integrating and then evaluating the limit.

$$\int_1^\infty \frac{1}{x^p}\, dx = \lim_{t\to\infty} \int_1^t \frac{1}{x^p}\, dx$$

$$= \lim_{t\to\infty} \int_1^t x^{-p}\, dx \qquad (\text{assume } p \neq 1)$$

$$= \lim_{t\to\infty} \frac{1}{-p+1} x^{-p+1}\Big|_1^t$$

$$= \lim_{t\to\infty} \frac{1}{1-p}\left(t^{1-p} - 1^{1-p}\right).$$

When does this limit converge – i.e., when is this limit *not* ∞? This limit converges precisely when the power of b is less than 0: when $1 - p < 0 \Rightarrow 1 < p$.

Our analysis shows that if $p > 1$, then $\displaystyle\int_1^\infty \frac{1}{x^p}\, dx$ converges. When $p < 1$ the improper integral diverges; we showed in Example 1 that when $p = 1$ the integral also diverges.

Figure 8.14 graphs $y = 1/x$ with a dashed line, along with graphs of $y = 1/x^p$, $p < 1$, and $y = 1/x^q$, $q > 1$. Somehow the dashed line forms a dividing line between convergence and divergence.

The result of Example 4 provides an important tool in determining the convergence of other integrals. A similar result is proved in the exercises about improper integrals of the form $\displaystyle\int_0^1 \frac{1}{x^p}\, dx$. These results are summarized in the following Key Idea.
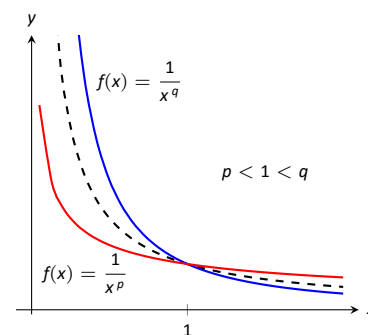
Figure 8.14: Plotting functions of the form $1/x^p$ in Example 4.

Key Idea 31   **Convergence of Improper Integrals** $\displaystyle\int_1^\infty \frac{1}{x^p}\, dx$ **and** $\displaystyle\int_0^1 \frac{1}{x^p}\, dx.$

1. The improper integral $\displaystyle\int_1^\infty \frac{1}{x^p}\, dx$ converges when $p > 1$ and diverges when $p \le 1$.

2. The improper integral $\displaystyle\int_0^1 \frac{1}{x^p}\, dx$ converges when $p < 1$ and diverges when $p \ge 1$.

Notes:

A basic technique in determining convergence of improper integrals is to compare an integrand whose convergence is unknown to an integrand whose convergence is known. We often use integrands of the form $1/x^p$ in comparisons as their convergence on certain intervals is known. This is described in the following theorem.

Note: We used the upper and lower bound of "1" in Key Idea 31 for convenience. It can be replaced by any a where $a > 0$.

Theorem 55   Direct Comparison Test for Improper Integrals

Let f and g be continuous on $[a, \infty)$ where $0 \le f(x) \le g(x)$ for all x in $[a, \infty)$.

1. If $\displaystyle\int_a^\infty g(x)\, dx$ converges, then $\displaystyle\int_a^\infty f(x)\, dx$ converges.

2. If $\displaystyle\int_a^\infty f(x)\, dx$ diverges, then $\displaystyle\int_a^\infty g(x)\, dx$ diverges.

Example 5   Determining convergence of improper integrals

Determine the convergence of the following improper integrals.

$$1. \int_1^\infty e^{-x^2}\, dx \qquad\qquad 2. \int_3^\infty \frac{1}{\sqrt{x^2 - x}}\, dx$$

SOLUTION

1. The function $f(x) = e^{-x^2}$ does not have an antiderivative expressible in terms of elementary functions, so we cannot integrate directly. It is comparable to $g(x) = 1/x^2$, and as demonstrated in Figure 8.15, $e^{-x^2} < 1/x^2$ on $[1, \infty)$. We know from Key Idea 31 that $\displaystyle\int_1^\infty \frac{1}{x^2}\, dx$ converges, hence $\displaystyle\int_1^\infty e^{-x^2}\, dx$ also converges.

2. Note that for large values of x, $\dfrac{1}{\sqrt{x^2 - x}} \approx \dfrac{1}{\sqrt{x^2}} = \dfrac{1}{x}$. We know from Key Idea 31 and the subsequent note that $\displaystyle\int_3^\infty \frac{1}{x}\, dx$ diverges, so we seek to compare the original integrand to $1/x$.

   It is easy to see that when $x > 0$, we have $x = \sqrt{x^2} > \sqrt{x^2 - x}$. Taking reciprocals reverses the inequality, giving

   $$\frac{1}{x} < \frac{1}{\sqrt{x^2 - x}}.$$
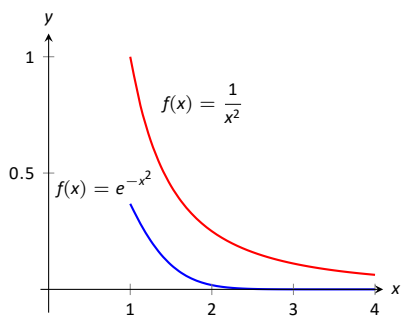
Figure 8.15: Graphs of $f(x) = e^{-x^2}$ and $f(x) = 1/x^2$ in Example 5.

Notes:

Using Theorem 55, we conclude that since $\int_3^\infty \frac{1}{x}\,dx$ diverges, $\int_3^\infty \frac{1}{\sqrt{x^2-x}}\,dx$ diverges as well. Figure 8.16 illustrates this.

Being able to compare "unknown" integrals to "known" integrals is very useful in determining convergence. However, some of our examples were a little "too nice." For instance, it was convenient that $\frac{1}{x} < \frac{1}{\sqrt{x^2-x}}$, but what if the "$-x$" were replaced with a "$+2x+5$"? That is, what can we say about the convergence of $\int_3^\infty \frac{1}{\sqrt{x^2+2x+5}}\,dx$? We have $\frac{1}{x} > \frac{1}{\sqrt{x^2+2x+5}}$, so we cannot use Theorem 55.

In cases like this (and many more) it is useful to employ the following theorem.

Figure 8.16: Graphs of $f(x) = 1/\sqrt{x^2-x}$ and $f(x) = 1/x$ in Example 5.

Theorem 56 Limit Comparison Test for Improper Integrals

Let f and g be continuous functions on $[a, \infty)$ where $f(x) > 0$ and $g(x) > 0$ for all x. If

$$\lim_{x\to\infty} \frac{f(x)}{g(x)} = L, \qquad 0 < L < \infty,$$

then

$$\int_a^\infty f(x)\,dx \quad \text{and} \quad \int_a^\infty g(x)\,dx$$

either both converge or both diverge.

Example 6 Determining convergence of improper integrals

Determine the convergence of $\int_3^\infty \frac{1}{\sqrt{x^2+2x+5}}\,dx$.

SOLUTION As x gets large, the quadratic function will begin to behave much like $y = x$. So we compare $\frac{1}{\sqrt{x^2+2x+5}}$ to $\frac{1}{x}$ with the Limit Comparison Test:

$$\lim_{x\to\infty} \frac{1/\sqrt{x^2+2x+5}}{1/x} = \lim_{x\to\infty} \frac{x}{\sqrt{x^2+2x+5}}.$$

The immediate evaluation of this limit returns ∞/∞, an indeterminate form. Using L'Hôpital's Rule seems appropriate, but in this situation, it does not lead

---

Notes:
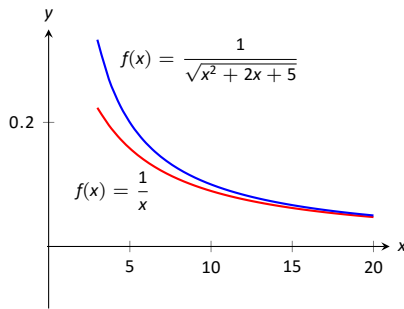
Figure 8.17: Graphing $f(x) = \frac{1}{\sqrt{x^2+2x+5}}$
and $f(x) = \frac{1}{x}$ in Example 6.

to useful results. (We encourage the reader to employ L'Hôpital's Rule at least once to verify this.)

The trouble is the square root function. We determine the limit by using a technique we learned in Calculus I:

$$\lim_{x\to\infty} \frac{x}{\sqrt{x^2+2x+5}} = \lim_{x\to\infty} \frac{\frac{x}{x}}{\sqrt{\frac{x^2+2x+5}{x^2}}} = \lim_{x\to\infty} \frac{1}{\sqrt{1+\frac{2}{x}+\frac{5}{x^2}}} = 1$$

Since we know that $\int_3^\infty \frac{1}{x}\, dx$ diverges, by the Limit Comparison Test we know that $\int_3^\infty \frac{1}{\sqrt{x^2+2x+5}}\, dx$ also diverges. Figure 8.17 graphs $f(x) = 1/\sqrt{x^2+2x+5}$ and $f(x) = 1/x$, illustrating that as x gets large, the functions become indistinguishable.

Both the Direct and Limit Comparison Tests were given in terms of integrals over an infinite interval. There are versions that apply to improper integrals with an infinite range, but as they are a bit wordy and a little more difficult to employ, they are omitted from this text.

This chapter has explored many integration techniques. We learned Integration by Parts, which reverses the Product Rule of differentiation. We also learned specialized techniques for handling trigonometric and rational functions. All techniques effectively have this goal in common: rewrite the integrand in a new way so that the integration step is easier to see and implement.

As stated before, integration is, in general, hard. It is easy to write a function whose antiderivative is impossible to write in terms of elementary functions, and even when a function does have an antiderivative expressible by elementary functions, it may be really hard to discover what it is. The powerful computer algebra system *Mathematica*® has approximately 1,000 pages of code dedicated to integration.

Do not let this difficulty discourage you. There is great value in learning integration techniques, as they allow one to manipulate an integral in ways that can illuminate a concept for greater understanding. There is also great value in understanding the need for good numerical techniques: the Trapezoidal and Simpson's Rules are just the beginning of powerful techniques for approximating the value of integration.

The next chapter stresses the uses of integration. We generally do not find antiderivatives for antiderivative's sake, but rather because they provide the solution to some type of problem. The following chapter introduces us to several different problems whose solution is provided by integration.

Notes:

Exercises 8.6

Terms and Concepts

1. The definite integral was defined with what two stipulations?

2. If $\lim\limits_{b \to \infty} \int_0^b f(x)\,dx$ exists, then the integral $\int_0^\infty f(x)\,dx$ is said to _____.

3. If $\int_1^\infty f(x)\,dx = 10$, and $0 \le g(x) \le f(x)$ for all x, then we know that $\int_1^\infty g(x)\,dx$ _____.

4. For what values of p will $\int_1^\infty \dfrac{1}{x^p}\,dx$ converge?

5. For what values of p will $\int_{10}^\infty \dfrac{1}{x^p}\,dx$ converge?

6. For what values of p will $\int_0^1 \dfrac{1}{x^p}\,dx$ converge?

Problems

In Exercises 7–37, evaluate the given improper integral.

7. $\int_0^\infty e^{5-2x}\,dx$

8. $\int_1^\infty \dfrac{1}{x^3}\,dx$

9. $\int_1^\infty x^{-4}\,dx$

10. $\int_{-\infty}^\infty \dfrac{1}{x^2+9}\,dx$

11. $\int_{-\infty}^0 2^x\,dx$

12. $\int_{-\infty}^0 \left(\dfrac{1}{2}\right)^x\,dx$

13. $\int_{-\infty}^\infty \dfrac{x}{x^2+1}\,dx$

14. $\int_{-\infty}^\infty \dfrac{x}{x^2+4}\,dx$

15. $\int_2^\infty \dfrac{1}{(x-1)^2}\,dx$

16. $\int_1^2 \dfrac{1}{(x-1)^2}\,dx$

17. $\int_2^\infty \dfrac{1}{x-1}\,dx$

18. $\int_1^2 \dfrac{1}{x-1}\,dx$

19. $\int_0^3 \dfrac{1}{x}\,dx$

20. $\int_{-1}^1 \dfrac{1}{x}\,dx$

21. $\int_2^5 \dfrac{dx}{\sqrt{x-2}}$

22. $\int_1^9 \dfrac{dx}{\sqrt[3]{9-x}}$

23. $\int_1^3 \dfrac{1}{x-2}\,dx$

24. $\int_0^\pi \sec^2 x\,dx$

25. $\int_0^{\frac{\pi}{2}} \sec x\,dx$

26. $\int_{-2}^1 \dfrac{1}{\sqrt{|x|}}\,dx$

27. $\int_0^\infty xe^{-x}\,dx$

28. $\int_0^\infty xe^{-x^2}\,dx$

29. $\int_{-\infty}^\infty xe^{-x^2}\,dx$

30. $\int_{-\infty}^\infty \dfrac{1}{e^x+e^{-x}}\,dx$

31. $\int_0^1 x\ln x\,dx$

32. $\int_1^\infty \dfrac{\ln x}{x}\,dx$

33. $\int_0^1 \ln x\,dx$

34. $\int_1^\infty \dfrac{\ln x}{x^2}\,dx$

35. $\int_1^\infty \dfrac{\ln x}{\sqrt{x}}\,dx$

36. $\int_0^\infty e^{-x}\sin x\,dx$

37. $\int_0^\infty e^{-x}\cos x\,dx$

In Exercises 38–47, use the Direct Comparison Test or the Limit Comparison Test to determine whether the given definite integral converges or diverges. Clearly state what test is being used and what function the integrand is being compared to.

38. $\int_{10}^\infty \dfrac{3}{\sqrt{3x^2+2x-5}}\,dx$

39. $\int_2^\infty \dfrac{4}{\sqrt{7x^3-x}}\,dx$

40. $\int_0^\infty \dfrac{\sqrt{x+3}}{\sqrt{x^3-x^2+x+1}}\,dx$

41. $\int_1^\infty e^{-x}\ln x\,dx$

42. $\int_5^\infty e^{-x^2+3x+1}\,dx$

43. $\int_0^\infty \dfrac{\sqrt{x}}{e^x}\,dx$

44. $\int_2^\infty \dfrac{1}{x^2+\sin x}\,dx$

45. $\int_0^\infty \dfrac{x}{x^2+\cos x}\,dx$

46. $\int_0^\infty \dfrac{1}{x+e^x}\,dx$

47. $\int_0^\infty \dfrac{1}{e^x-x}\,dx$

8.7   Numerical Integration

The Fundamental Theorem of Calculus gives a concrete technique for finding the exact value of a definite integral. That technique is based on computing antiderivatives. Despite the power of this theorem, there are still situations where we must *approximate* the value of the definite integral instead of finding its exact value. The first situation we explore is where we *cannot* compute an antiderivative of the integrand. The second case is when we actually do not know the integrand, but only its value when evaluated at certain points.

An **elementary function** is any function that is a combination of polynomials, n^{th} roots, rational, exponential, logarithmic and trigonometric functions and their inverses. We can compute the derivative of any elementary function, but there are many elementary functions of which we cannot compute an antiderivative. For example, the following functions do not have antiderivatives that we can express with elementary functions:

$$e^{-x^2}, \quad \sin(x^3) \quad \text{and} \quad \frac{\sin x}{x}.$$

The simplest way to refer to the antiderivatives of e^{-x^2} is to simply write $\int e^{-x^2}\, dx.$

This section outlines three common methods of approximating the value of definite integrals. We describe each as a systematic method of approximating area under a curve. By approximating this area accurately, we find an accurate approximation of the corresponding definite integral.

We will apply the methods we learn in this section to the following definite integrals:

$$\int_0^1 e^{-x^2}\, dx, \quad \int_{-\frac{\pi}{4}}^{\frac{\pi}{2}} \sin(x^3)\, dx, \quad \text{and} \quad \int_{0.5}^{4\pi} \frac{\sin(x)}{x}\, dx,$$

as pictured in Figure 8.18.

The Left and Right Hand Rule Methods

In Section 5.3 we addressed the problem of evaluating definite integrals by approximating the area under the curve using rectangles. We revisit those ideas here before introducing other methods of approximating definite integrals.

We start with a review of notation. Let f be a continuous function on the interval $[a,b]$. We wish to approximate $\int_a^b f(x)\, dx$. We partition $[a,b]$ into n
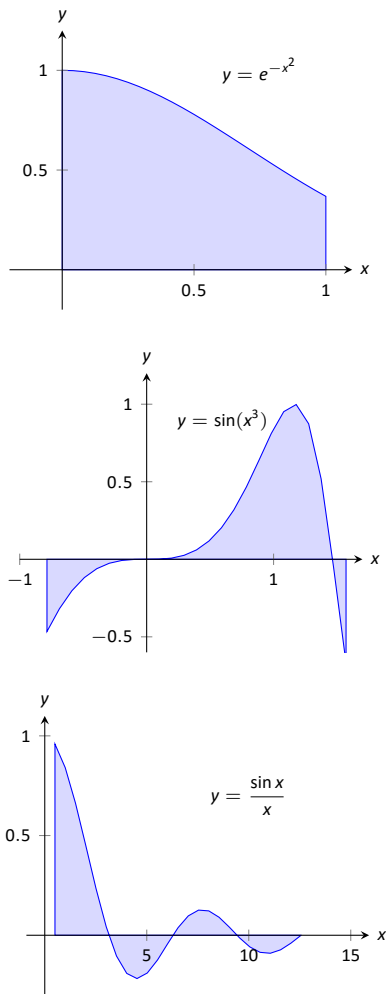
Figure 8.18: Graphically representing three definite integrals that cannot be evaluated using antiderivatives.

Notes:

equally spaced subintervals, each of length $\Delta x = \dfrac{b-a}{n}$. The endpoints of these subintervals are labeled as

$$x_0 = a, \; x_1 = a + \Delta x, \; x_2 = a + 2\Delta x, \; \ldots, \; x_i = a + i\Delta x, \; \ldots, \; x_n = b.$$

Section 5.3 showed that to use the Left Hand Rule we use the summation $\displaystyle\sum_{i=1}^{n} f(x_{i-1})\Delta x$ and to use the Right Hand Rule we use $\displaystyle\sum_{i=1}^{n} f(x_i)\Delta x$. We review the use of these rules in the context of examples.

Example 1   Approximating definite integrals with rectangles

Approximate $\displaystyle\int_0^1 e^{-x^2}\,dx$ using the Left and Right Hand Rules with 5 equally spaced subintervals.

SOLUTION   We begin by partitioning the interval $[0,1]$ into 5 equally spaced intervals. We have $\Delta x = \frac{1-0}{5} = 1/5 = 0.2$, so

$$x_0 = 0, \; x_1 = 0.2, \; x_2 = 0.4, \; x_3 = 0.6, \; x_4 = 0.8, \text{ and } x_5 = 1.$$

Using the Left Hand Rule, we have:

$$\sum_{i=1}^{n} f(x_{i-1})\Delta x = \big(f(x_0) + f(x_1) + f(x_2) + f(x_3) + f(x_4)\big)\Delta x$$

$$= \big(f(0) + f(0.2) + f(0.4) + f(0.6) + f(0.8)\big)\Delta x$$

$$\approx \big(1 + 0.961 + 0.852 + 0.698 + 0.527\big)(0.2)$$

$$\approx 0.808.$$

Using the Right Hand Rule, we have:

$$\sum_{i=1}^{n} f(x_i)\Delta x = \big(f(x_1) + f(x_2) + f(x_3) + f(x_4) + f(x_5)\big)\Delta x$$

$$= \big(f(0.2) + f(0.4) + f(0.6) + f(0.8) + f(1)\big)\Delta x$$

$$\approx \big(0.961 + 0.852 + 0.698 + 0.527 + 0.368\big)(0.2)$$

$$\approx 0.681.$$

Figure 8.19 shows the rectangles used in each method to approximate the definite integral. These graphs show that in this particular case, the Left Hand Rule is an over approximation and the Right Hand Rule is an under approximation. To get a better approximation, we could use more rectangles, as we did in
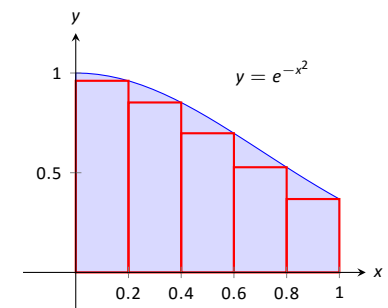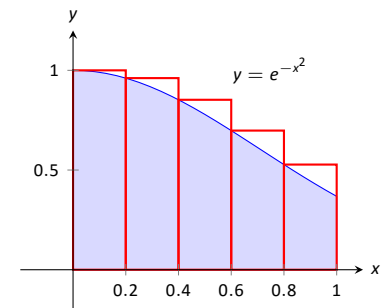
Figure 8.19: Approximating $\displaystyle\int_0^1 e^{-x^2}\,dx$ in Example 1 using (top) the left hand rule and (bottom) the right hand rule.

Notes:

x_i	Exact	Approx.	$\sin(x_i^3)$
x_0	$-\pi/4$	-0.785	-0.466
x_1	$-7\pi/40$	-0.550	-0.165
x_2	$-\pi/10$	-0.314	-0.031
x_3	$-\pi/40$	-0.0785	0
x_4	$\pi/20$	0.157	0.004
x_5	$\pi/8$	0.393	0.061
x_6	$\pi/5$	0.628	0.246
x_7	$11\pi/40$	0.864	0.601
x_8	$7\pi/20$	1.10	0.971
x_9	$17\pi/40$	1.34	0.690
x_{10}	$\pi/2$	1.57	-0.670

Figure 8.20: Table of values used to approximate $\int_{-\frac{\pi}{4}}^{\frac{\pi}{2}} \sin(x^3)\, dx$ in Example 2.
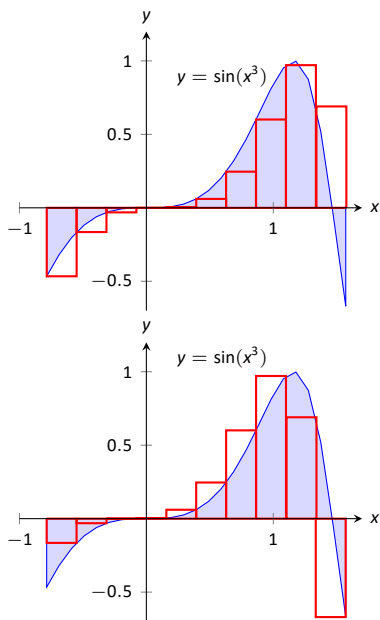
Figure 8.21: Approximating $\int_{-\frac{\pi}{4}}^{\frac{\pi}{2}} \sin(x^3)\, dx$ in Example 2 using (top) the left hand rule and (bottom) the right hand rule.

Section 5.3. We could also average the Left and Right Hand Rule results together, giving

$$\frac{0.808 + 0.681}{2} = 0.7445.$$

The actual answer, accurate to 4 places after the decimal, is 0.7468, showing our average is a good approximation.

Example 2   **Approximating definite integrals with rectangles**

Approximate $\int_{-\frac{\pi}{4}}^{\frac{\pi}{2}} \sin(x^3)\, dx$ using the Left and Right Hand Rules with 10 equally spaced subintervals.

    SOLUTION   We begin by finding Δx:

$$\frac{b - a}{n} = \frac{\pi/2 - (-\pi/4)}{10} = \frac{3\pi}{40} \approx 0.236.$$

It is useful to write out the endpoints of the subintervals in a table; in Figure 8.20, we give the exact values of the endpoints, their decimal approximations, and decimal approximations of $\sin(x^3)$ evaluated at these points.

    Once this table is created, it is straightforward to approximate the definite integral using the Left and Right Hand Rules. (Note: the table itself is easy to create, especially with a standard spreadsheet program on a computer. The last two columns are all that are needed.) The Left Hand Rule sums the first 10 values of $\sin(x_i^3)$ and multiplies the sum by Δx; the Right Hand Rule sums the last 10 values of $\sin(x_i^3)$ and multiplies by Δx. Therefore we have:

$$\text{Left Hand Rule: } \int_{-\frac{\pi}{4}}^{\frac{\pi}{2}} \sin(x^3)\, dx \approx (1.91)(0.236) = 0.451.$$

$$\text{Right Hand Rule: } \int_{-\frac{\pi}{4}}^{\frac{\pi}{2}} \sin(x^3)\, dx \approx (1.71)(0.236) = 0.404.$$

    The average of the Left and Right Hand Rules is 0.4275. The actual answer, accurate to 3 places after the decimal, is 0.460. Our approximations were once again fairly good. The rectangles used in each approximation are shown in Figure 8.21. It is clear from the graphs that using more rectangles (and hence, narrower rectangles) should result in a more accurate approximation.

Notes:

The Trapezoidal Rule

In Example 1 we approximated the value of $\int_0^1 e^{-x^2}\, dx$ with 5 rectangles of equal width. Figure 8.19 showed the rectangles used in the Left and Right Hand Rules. These graphs clearly show that rectangles do not match the shape of the graph all that well, and that accurate approximations will only come by using lots of rectangles.

Instead of using rectangles to approximate the area, we can instead use *trapezoids.* In Figure 8.22, we show the region under $f(x) = e^{-x^2}$ on $[0, 1]$ approximated with 5 trapezoids of equal width; the top "corners" of each trapezoid lies on the graph of $f(x)$. It is clear from this figure that these trapezoids more accurately approximate the area under f and hence should give a better approximation of $\int_0^1 e^{-x^2}\, dx$. (In fact, these trapezoids seem to give a *great* approximation of the area.)

Figure 8.22: Approximating $\int_0^1 e^{-x^2}\, dx$ using 5 trapezoids of equal widths.

Watch the video:
The Trapezoid Rule for Approximating Integrals at
`https://youtu.be/8z6JRFvjkpc`

The formula for the area of a trapezoid is given in Figure 8.23. We approximate $\int_0^1 e^{-x^2}\, dx$ with these trapezoids in the following example.

Figure 8.23: The area of a trapezoid.

Example 3   Approximating definite integrals using trapezoids

Use 5 trapezoids of equal width to approximate $\int_0^1 e^{-x^2}\, dx$.

SOLUTION   To compute the areas of the 5 trapezoids in Figure 8.22, it will again be useful to create a table of values as shown in Figure 8.24.

The leftmost trapezoid has legs of length 1 and 0.961 and a height of 0.2. Thus, by our formula, the area of the leftmost trapezoid is:

$$\frac{1 + 0.961}{2}(0.2) = 0.1961.$$

Moving right, the next trapezoid has legs of length 0.961 and 0.852 and a height of 0.2. Thus its area is:

$$\frac{0.961 + 0.852}{2}(0.2) = 0.1813.$$

x_i	$e^{-x_i^2}$
0	1
0.2	0.961
0.4	0.852
0.6	0.698
0.8	0.527
1	0.368

Figure 8.24: A table of values of e^{-x^2}.

Notes:

The sum of the areas of all 5 trapezoids is:

$$\frac{1+0.961}{2}(0.2) + \frac{0.961+0.852}{2}(0.2) + \frac{0.852+0.698}{2}(0.2)$$

$$+ \frac{0.698+0.527}{2}(0.2) + \frac{0.527+0.368}{2}(0.2) = 0.7445.$$

We approximate $\int_0^1 e^{-x^2}\, dx \approx 0.7445$.

There are many things to observe in this example. Note how each term in the final summation was multiplied by both 1/2 and by $\Delta x = 0.2$. We can factor these coefficients out, leaving a more concise summation as:

$$\frac{1}{2}(0.2)\Big[(1+0.961)+(0.961+0.852)+(0.852+0.698)+(0.698+0.527)+(0.527+0.368)\Big].$$

Now notice that all numbers except for the first and the last are added twice. Therefore we can write the summation even more concisely as

$$\frac{0.2}{2}\Big[1 + 2(0.961 + 0.852 + 0.698 + 0.527) + 0.368\Big].$$

This is the heart of the **Trapezoidal Rule**, wherein a definite integral $\int_a^b f(x)\, dx$ is approximated by using trapezoids of equal widths to approximate the corresponding area under f. Using n equally spaced subintervals with endpoints x_0, $x_1, \ldots, x_n$, we again have $\Delta x = \dfrac{b-a}{n}$. Thus:

$$\int_a^b f(x)\, dx \approx \sum_{i=1}^n \frac{f(x_{i-1}) + f(x_i)}{2} \Delta x$$

$$= \frac{\Delta x}{2} \sum_{i=1}^n \big(f(x_{i-1}) + f(x_i)\big)$$

$$= \frac{\Delta x}{2} \left[f(x_0) + 2 \sum_{i=1}^{n-1} f(x_i) + f(x_n) \right].$$

Example 4   Using the Trapezoidal Rule

Revisit Example 2 and approximate $\int_{-\frac{\pi}{4}}^{\frac{\pi}{2}} \sin(x^3)\, dx$ using the Trapezoidal Rule and 10 equally spaced subintervals.

Notes:

Solution   We refer back to Figure 8.20 for the table of values of $\sin(x^3)$. Recall that $\Delta x = 3\pi/40 \approx 0.236$. Thus we have:

$$\int_{-\frac{\pi}{4}}^{\frac{\pi}{2}} \sin(x^3)\,dx$$
$$\approx \frac{0.236}{2}\Big[-0.466 + 2\big(-0.165 + (-0.031) + \cdots + 0.69\big) + (-0.67)\Big]$$
$$= 0.4275.$$

Notice how "quickly" the Trapezoidal Rule can be implemented once the table of values is created. This is true for all the methods explored in this section; the real work is creating a table of x_i and $f(x_i)$ values. Once this is completed, approximating the definite integral is not difficult. Again, using technology is wise. Spreadsheets can make quick work of these computations and make using lots of subintervals easy.

Also notice the approximations the Trapezoidal Rule gives. It is the average of the approximations given by the Left and Right Hand Rules! This effectively renders the Left and Right Hand Rules obsolete. They are useful when first learning about definite integrals, but if a real approximation is needed, one is generally better off using the Trapezoidal Rule instead of either the Left or Right Hand Rule.

We will also show that the Trapezoidal Rule makes using the Midpoint Rule obsolete as well. With much more work, it will turn out that the Midpoint Rule has only a marginal gain in accuracy. But we will include it in our results for the sake of completeness.

How can we improve on the Trapezoidal Rule, apart from using more and more trapezoids? The answer is clear once we look back and consider what we have *really* done so far. The Left Hand Rule is not *really* about using rectangles to approximate area. Instead, it approximates a function f with constant functions on small subintervals and then computes the definite integral of these constant functions. The Trapezoidal Rule is really approximating a function f with a linear function on a small subinterval, then computes the definite integral of this linear function. In both of these cases the definite integrals are easy to compute in geometric terms.

So we have a progression: we start by approximating f with a constant function and then with a linear function. What is next? A quadratic function. By approximating the curve of a function with lots of parabolas, we generally get an even better approximation of the definite integral. We call this process **Simpson's Rule**, named after Thomas Simpson (1710-1761), even though others had used this rule as much as 100 years prior.

Notes:

Simpson's Rule

Given one point, we can create a constant function that goes through that point. Given two points, we can create a linear function that goes through those points. Given three points, we can create a quadratic function that goes through those three points (given that no two have the same x–value).

Consider three points (x_1, y_1), (x_2, y_2) and (x_3, y_3) whose x–values are equally spaced and $x_1 < x_2 < x_3$. Let f be the quadratic function that goes through these three points. An exercise will ask you to show that

$$\int_{x_1}^{x_3} f(x)\, dx = \frac{x_3 - x_1}{6} \left(y_1 + 4y_2 + y_3 \right). \tag{8.7}$$

Consider Figure 8.25. A function f goes through the 3 points shown and the parabola g that also goes through those points is graphed with a dashed line. Using our equation from above, we know exactly that

$$\int_1^3 g(x)\, dx = \frac{3-1}{6} \left(3 + 4(1) + 2 \right) = 3.$$

Since g is a good approximation for f on $[1, 3]$, we can state that

$$\int_1^3 f(x)\, dx \approx 3.$$

Notice how the interval $[1, 3]$ was split into two subintervals as we needed 3 points. Because of this, whenever we use Simpson's Rule, we need to break the interval into an even number of subintervals.

In general, to approximate $\int_a^b f(x)\, dx$ using Simpson's Rule, subdivide $[a, b]$ into n subintervals, where n is even and each subinterval has width $\Delta x = (b - a)/n$. We approximate f with $n/2$ parabolic curves, using Equation (8.7) to compute the area under these parabolas. Adding up these areas gives the formula:

$$\int_a^b f(x)\, dx \approx \frac{\Delta x}{3} \left[f(x_0) + 4f(x_1) + 2f(x_2) + 4f(x_3) + \cdots + 2f(x_{n-2}) + 4f(x_{n-1}) + f(x_n) \right].$$

Note how the coefficients of the terms in the summation have the pattern 1, 4, 2, 4, 2, 4, ..., 2, 4, 1.

Let's demonstrate Simpson's Rule with a concrete example.

Example 5   Using Simpson's Rule

Approximate $\int_0^1 e^{-x^2}\, dx$ using Simpson's Rule and 4 equally spaced subintervals.

Figure 8.25: A graph of a function f and a parabola that approximates it well on $[1, 3]$.

Notes:

SOLUTION   We begin by making a table of values as we have in the past, as shown in Figure 8.26(a). Simpson's Rule states that

$$\int_0^1 e^{-x^2}\, dx \approx \frac{0.25}{3}\left[1 + 4(0.939) + 2(0.779) + 4(0.570) + 0.368\right] = 0.7468\overline{3}.$$

Recall in Example 1 we stated that the correct answer, accurate to 4 places after the decimal, was 0.7468. Our approximation with Simpson's Rule, with 4 subintervals, is better than our approximation with the Trapezoidal Rule using 5.

Figure 8.26(b) shows $f(x) = e^{-x^2}$ along with its approximating parabolas, demonstrating how good our approximation is. The approximating curves are nearly indistinguishable from the actual function.

Example 6   Using Simpson's Rule

Approximate $\int_{-\frac{\pi}{4}}^{\frac{\pi}{2}} \sin(x^3)\, dx$ using Simpson's Rule and 10 equally spaced intervals.

SOLUTION   Figure 8.27(a) shows the table of values that we used in the past for this problem, shown here again for convenience. Again, $\Delta x = (\pi/2 + \pi/4)/10 \approx 0.236$.

Simpson's Rule states that

$$\int_{-\frac{\pi}{4}}^{\frac{\pi}{2}} \sin(x^3)\, dx \approx \frac{0.236}{3}\Big[(-0.466) + 4(-0.165) + 2(-0.031) + \cdots$$

$$\cdots + 2(0.971) + 4(0.69) + (-0.67)\Big]$$

$$= 0.4701$$

Recall that the actual value, accurate to 3 decimal places, is 0.460. Our approximation is within one $1/100^{\text{th}}$ of the correct value. The graph in Figure 8.27(b) shows how closely the parabolas match the shape of the graph.

Summary and Error Analysis

We summarize the key concepts of this section thus far in the following Key Idea.

Notes:

x_i	$e^{-x_i^2}$
0	1
0.25	0.939
0.5	0.779
0.75	0.570
1	0.368

(a)

(b)
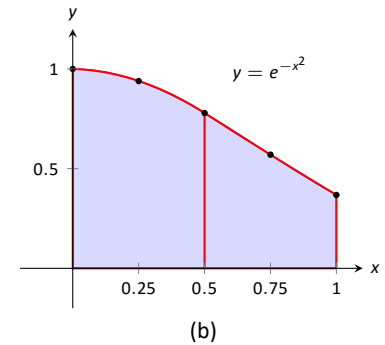
Figure 8.26: A table of values to approximate $\int_0^1 e^{-x^2}\, dx$ in Example 5, along with a graph of the function.

x_i	$\sin(x_i^3)$
−0.785	−0.466
−0.550	−0.165
−0.314	−0.031
−0.0785	0
0.157	0.004
0.393	0.061
0.628	0.246
0.864	0.601
1.10	0.971
1.34	0.690
1.57	−0.670

(a)

(b)
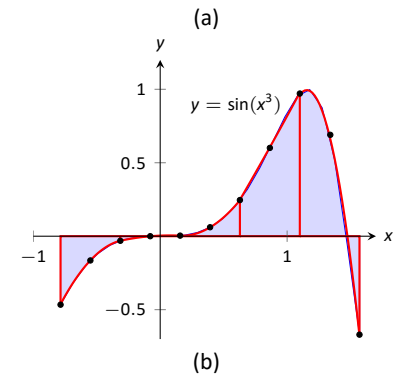
Figure 8.27: A table of values to approximate $\int_{-\frac{\pi}{4}}^{\frac{\pi}{2}} \sin(x^3)\, dx$ in Example 6, along with a graph of the function.

Key Idea 32   Numerical Integration

Let f be a continuous function on $[a,b]$, let n be a positive integer, and let $\Delta x = \dfrac{b-a}{n}$. Set $x_0 = a$, $x_1 = a + \Delta x, ..., x_i = a + i\Delta x, x_n = b$. Consider $\displaystyle\int_a^b f(x)\, dx$.

Left Hand Rule: $\displaystyle\int_a^b f(x)\, dx \approx \Delta x \Big[f(x_0) + f(x_1) + \cdots + f(x_{n-1}) \Big]$.

Right Hand Rule: $\displaystyle\int_a^b f(x)\, dx \approx \Delta x \Big[f(x_1) + f(x_2) + \cdots + f(x_n) \Big]$.

Midpoint Rule: $\displaystyle\int_a^b f(x)\, dx \approx \Delta x \Big[f\big(\tfrac{x_0 + x_1}{2}\big) + f\big(\tfrac{x_1 + x_2}{2}\big) + \cdots + f\big(\tfrac{x_{n-1} + x_n}{2}\big) \Big]$.

Trapezoidal Rule: $\displaystyle\int_a^b f(x)\, dx \approx \dfrac{\Delta x}{2} \Big[f(x_0) + 2f(x_1) + 2f(x_2) + \cdots + 2f(x_{n-1}) + f(x_n) \Big]$.

Simpson's Rule: $\displaystyle\int_a^b f(x)\, dx \approx \dfrac{\Delta x}{3} \Big[f(x_0) + 4f(x_1) + 2f(x_2) + \cdots + 4f(x_{n-1}) + f(x_n) \Big]$ (n even).

In our examples, we approximated the value of a definite integral using a given method then compared it to the "right" answer. This should have raised several questions in the reader's mind, such as:

1. How was the "right" answer computed?

2. If the right answer can be found, what is the point of approximating?

3. If there is value to approximating, how are we supposed to know if the approximation is any good?

These are good questions, and their answers are educational. In the examples, *the* right answer was never computed. Rather, an approximation accurate to a certain number of places after the decimal was given. In Example 1, we do not know the *exact* answer, but we know it starts with 0.7468. These more accurate approximations were computed using numerical integration but with more precision (i.e., more subintervals and the help of a computer).

Since the exact answer cannot be found, approximation still has its place. How are we to tell if the approximation is any good?

"Trial and error" provides one way. Using technology, make an approximation with, say, 10, 100, and 200 subintervals. This likely will not take much time at all, and a trend should emerge. If a trend does not emerge, try using yet more subintervals. Keep in mind that trial and error is never foolproof; you might stumble upon a problem in which a trend will not emerge.

A second method is to use Error Analysis. While the details are beyond the scope of this text, there are some formulas that give *bounds* for how good your

---

Notes:

approximation will be. For instance, the formula might state that the approximation is within 0.1 of the correct answer. If the approximation is 1.58, then one knows that the correct answer is between 1.48 and 1.68. By using lots of subintervals, one can get an approximation as accurate as one likes. Theorem 57 states what these bounds are.

Theorem 57   Error Bounds in Numerical Integration

Suppose that M_n is an upper bound on $\left|f^{(n)}(x)\right|$ on $[a, b]$. Then a bound for the error of the numerical method of integration is given by:

Method	Error Bound
Left/Right Hand Rule	$\dfrac{M_1(b-a)^2}{2n}$
Midpoint Rule	$\dfrac{M_2(b-a)^3}{24n^2}$
Trapezoidal Rule	$\dfrac{M_2(b-a)^3}{12n^2}$
Simpson's Rule	$\dfrac{M_4(b-a)^5}{180n^4}$

There are some key things to note about this theorem.

1. The larger the interval, the larger the error. This should make sense intuitively.

2. The error shrinks as more subintervals are used (i.e., as n gets larger).

3. When n doubles, the Left and Right Hand Rules double in accuracy, the Midpoint and Trapezoidal Rules quadruple in accuracy, and Simpson's Rule is 16 times more accurate.

4. The error in Simpson's Rule has a term relating to the 4$^\text{th}$ derivative of f. Consider a cubic polynomial: its 4$^\text{th}$ derivative is 0. Therefore, the error in approximating the definite integral of a cubic polynomial with Simpson's Rule is 0 – Simpson's Rule computes the exact answer!

We revisit Examples 3 and 5 and compute the error bounds using Theorem 57 in the following example.

Example 7   Computing error bounds

Find the error bounds when approximating $\displaystyle\int_0^1 e^{-x^2}\,dx$ using the Trapezoidal Rule and 5 subintervals, and using Simpson's Rule with 4 subintervals.

Notes:
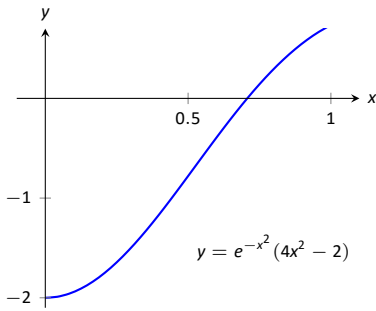
Figure 8.28: Graphing $f''(x)$ in Example 7 to help establish error bounds.
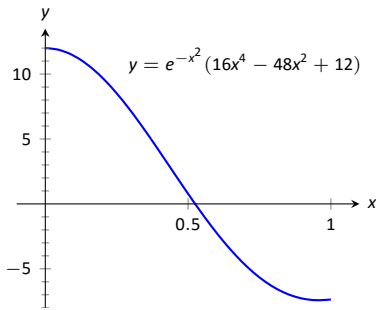
Figure 8.29: Graphing $f^{(4)}(x)$ in Example 7 to help establish error bounds.

SOLUTION   **Trapezoidal Rule with $n = 5$:**
We start by computing the 2nd derivative of $f(x) = e^{-x^2}$:

$$f''(x) = e^{-x^2}(4x^2 - 2).$$

Figure 8.28 shows a graph of $f''(x)$ on $[0, 1]$. It is clear that the largest value of f'', in absolute value, is 2. Thus we let $M = 2$ and apply the error formula from Theorem 57.

$$E_T = \frac{(1-0)^3}{12 \cdot 5^2} \cdot 2 = 0.00\overline{6}.$$

Our error estimation formula states that our approximation of 0.7445 found in Example 3 is within 0.0067 of the correct answer, hence we know that

$$0.7445 - 0.0067 = .7378 \le \int_0^1 e^{-x^2}\, dx \le 0.7512 = 0.7445 + 0.0067.$$

We had earlier computed the exact answer, correct to 4 decimal places, to be 0.7468, affirming the validity of Theorem 57.

Simpson's Rule with $n = 4$:
We start by computing the 4th derivative of $f(x) = e^{-x^2}$:

$$f^{(4)}(x) = e^{-x^2}(16x^4 - 48x^2 + 12).$$

Figure 8.29 shows a graph of $f^{(4)}(x)$ on $[0, 1]$. It is clear that the largest value of $f^{(4)}$, in absolute value, is 12. Thus we let $M = 12$ and apply the error formula from Theorem 57.

$$E_S = \frac{(1-0)^5}{180 \cdot 4^4} \cdot 12 = 0.00026.$$

Our error estimation formula states that our approximation of $0.7468\overline{3}$ found in Example 5 is within 0.00026 of the correct answer, hence we know that

$$0.74683 - 0.00026 = .74657 \le \int_0^1 e^{-x^2}\, dx \le 0.74709 = 0.74683 + 0.00026.$$

Once again we affirm the validity of Theorem 57.

At the beginning of this section we mentioned two main situations where numerical integration was desirable. We have considered the case where an antiderivative of the integrand cannot be computed. We now investigate the situation where the integrand is not known. This is, in fact, the most widely used application of Numerical Integration methods. "Most of the time" we observe behavior but do not know "the" function that describes it. We instead collect data about the behavior and make approximations based off of this data. We demonstrate this in an example.

Notes:

Example 8   **Approximating distance traveled**

One of the authors drove his daughter home from school while she recorded their speed every 30 seconds. The data is given in Figure 8.30. Approximate the distance they traveled.

    SOLUTION   Recall that by integrating a speed function we get distance traveled. We have information about $v(t)$; we will use Simpson's Rule to approximate $\int_a^b v(t)\, dt$.

    The most difficult aspect of this problem is converting the given data into the form we need it to be in. The speed is measured in miles per hour, whereas the time is measured in 30 second increments.

    We need to compute $\Delta x = (b-a)/n$. Clearly, $n = 24$. What are a and b? Since we start at time $t = 0$, we have that $a = 0$. The final recorded time came after 24 periods of 30 seconds, which is 12 minutes or 1/5 of an hour. Thus we have

$$\Delta x = \frac{b-a}{n} = \frac{1/5 - 0}{24} = \frac{1}{120}; \quad \frac{\Delta x}{3} = \frac{1}{360}.$$

    Thus the distance traveled is approximately:

$$\int_0^{0.2} v(t)\, dt \approx \frac{1}{360}\Big[f(x_1) + 4f(x_2) + 2f(x_3) + \cdots + 4f(x_n) + f(x_{n+1})\Big]$$

$$= \frac{1}{360}\Big[0 + 4 \cdot 25 + 2 \cdot 22 + \cdots + 2 \cdot 40 + 4 \cdot 23 + 0\Big]$$

$$\approx 6.2167 \text{ miles.}$$

We approximate the author drove 6.2 miles. (Because we are sure the reader wants to know, the author's odometer recorded the distance as about 6.05 miles.)

Time	Speed (mph)
0	0
1	25
2	22
3	19
4	39
5	0
6	43
7	59
8	54
9	51
10	43
11	35
12	40
13	43
14	30
15	0
16	0
17	28
18	40
19	42
20	40
21	39
22	40
23	23
24	0

Figure 8.30: Speed data collected at 30 second intervals for Example 8.

Notes:

Exercises 8.7

Terms and Concepts

1. T/F: Simpson's Rule is a method of approximating antiderivatives.
2. What are the two basic situations where approximating the value of a definite integral is necessary?
3. Why are the Left and Right Hand Rules rarely used?
4. Why is the Midpoint Rule rarely used?

Problems

In Exercises 5–12, a definite integral is given.

(a) Approximate the definite integral with the Trapezoidal Rule and $n = 4$.

(b) Approximate the definite integral with Simpson's Rule and $n = 4$.

(c) Find the exact value of the integral.

5. $\int_{-1}^{1} x^2 \, dx$

6. $\int_{0}^{10} 5x \, dx$

7. $\int_{0}^{\pi} \sin x \, dx$

8. $\int_{0}^{4} \sqrt{x} \, dx$

9. $\int_{0}^{3} (x^3 + 2x^2 - 5x + 7) \, dx$

10. $\int_{0}^{1} x^4 \, dx$

11. $\int_{0}^{2\pi} \cos x \, dx$

12. $\int_{-3}^{3} \sqrt{9 - x^2} \, dx$

In Exercises 13–20, approximate the definite integral with the Trapezoidal Rule and Simpson's Rule, with $n = 6$.

13. $\int_{0}^{1} \cos \left(x^2 \right) \, dx$

14. $\int_{-1}^{1} e^{x^2} \, dx$

15. $\int_{0}^{5} \sqrt{x^2 + 1} \, dx$

16. $\int_{0}^{\pi} x \sin x \, dx$

17. $\int_{0}^{\pi/2} \sqrt{\cos x} \, dx$

18. $\int_{1}^{4} \ln x \, dx$

19. $\int_{-1}^{1} \frac{1}{\sin x + 2} \, dx$

20. $\int_{0}^{6} \frac{1}{\sin x + 2} \, dx$

In Exercises 21–24, find n such that the error in approximating the given definite integral is less than 0.0001 when using:

(a) the Trapezoidal Rule

(b) Simpson's Rule

21. $\int_{0}^{\pi} \sin x \, dx$

22. $\int_{1}^{4} \frac{1}{\sqrt{x}} \, dx$

23. $\int_{0}^{\pi} \cos \left(x^2 \right) \, dx$

24. $\int_{0}^{5} x^4 \, dx$

In Exercises 25–26, a region is given. Find the area of the region using Simpson's Rule:

(a) where the measurements are in centimeters, taken in 1 cm increments, and

(b) where the measurements are in hundreds of yards, taken in 100 yd increments.

25.

26.
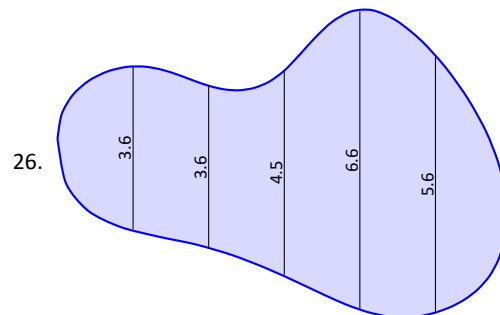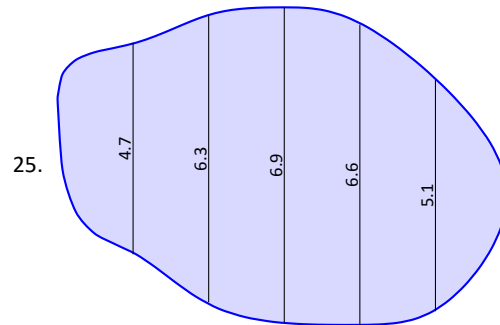
27. Let f be the quadratic function that goes through the points (x_1, y_1), $(x_1 + \Delta x, y_2)$ and $(x_1 + 2\Delta x, y_3)$. Show that
$$\int_{x_1}^{x_1 + 2\Delta x} f(x) \, dx = \frac{\Delta x}{3} (y_1 + 4y_2 + y_3).$$

9: Sequences and Series

This chapter introduces **sequences** and **series**, important mathematical constructions that are useful when solving a large variety of mathematical problems. The content of this chapter is considerably different from the content of the chapters before it. While the material we learn here definitely falls under the scope of "calculus," we will make very little use of derivatives or integrals. Limits are extremely important, though, especially limits that involve infinity.

One of the problems addressed by this chapter is this: suppose we know information about a function and its derivatives at a point, such as $f(1) = 3$, $f'(1) = 1, f''(1) = -2, f'''(1) = 7$, and so on. What can I say about $f(x)$ itself? Is there any reasonable approximation of the value of $f(2)$? The topic of Taylor Series addresses this problem, and allows us to make excellent approximations of functions when limited knowledge of the function is available.

9.1 Sequences

We commonly refer to a set of events that occur one after the other as a *sequence* of events. In mathematics, we use the word *sequence* to refer to an ordered set of numbers, i.e., a set of numbers that "occur one after the other."

For instance, the numbers 2, 4, 6, 8, ..., form a sequence. The order is important; the first number is 2, the second is 4, etc. It seems natural to seek a formula that describes a given sequence, and often this can be done. For instance, the sequence above could be described by the function $a(n) = 2n$, for the values of $n = 1, 2, \ldots$ (it could also be described by $n^4 - 10n^3 + 35n^2 - 48n + 24$, to give one of infinitely many other options). To find the 10th term in the sequence, we would compute $a(10)$. This leads us to the following, formal definition of a sequence.

> **Definition 30   Sequence**
>
> A **sequence** is a function $a(n)$ whose domain is $\mathbb{N}$. The **range** of a sequence is the set of all distinct values of $a(n)$.
>
> The **terms** of a sequence are the values $a(1), a(2), \ldots$, which are usually denoted with subscripts as $a_1, a_2, \ldots$.
>
> A sequence $a(n)$ is often denoted as $\{a_n\}$.

Factorial: The expression $3!$ refers to the number $3 \cdot 2 \cdot 1 = 6$.
In general, $n! = n \cdot (n-1) \cdot (n-2) \cdots 2 \cdot 1$, where n is a natural number.
We define $0! = 1$. While this does not immediately make sense, it makes many mathematical formulas work properly.

(a)

(b)

(c)
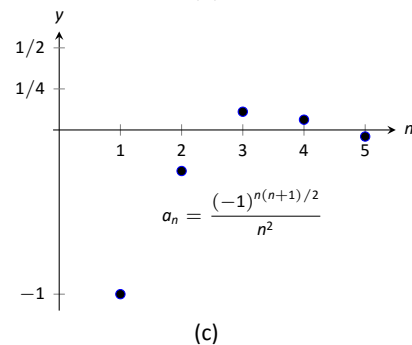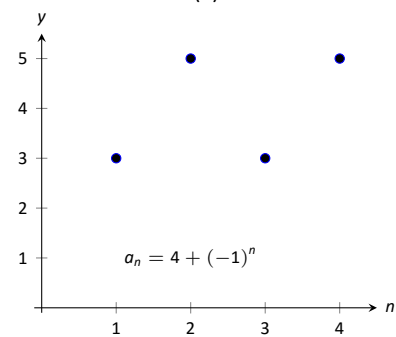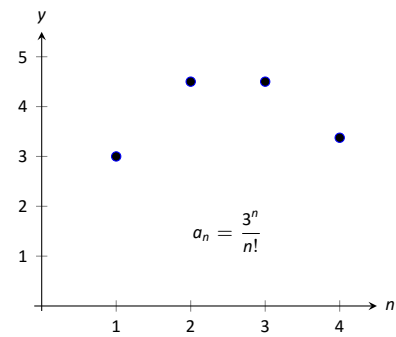
Figure 9.1: Plotting sequences in Example 1.

Example 1   **Listing terms of a sequence**
List the first four terms of the following sequences.

1. $\{a_n\} = \left\{ \dfrac{3^n}{n!} \right\}$ 2. $\{a_n\} = \{4 + (-1)^n\}$ 3. $\{a_n\} = \left\{ \dfrac{(-1)^{n(n+1)/2}}{n^2} \right\}$

SOLUTION

1. $a_1 = \dfrac{3^1}{1!} = 3; \qquad a_2 = \dfrac{3^2}{2!} = \dfrac{9}{2}; \qquad a_3 = \dfrac{3^3}{3!} = \dfrac{9}{2}; \qquad a_4 = \dfrac{3^4}{4!} = \dfrac{27}{8}$

   We can plot the terms of a sequence with a scatter plot. The "x"-axis is used for the values of n, and the values of the terms are plotted on the y-axis. To visualize this sequence, see Figure 9.1(a).

2. $a_1 = 4 + (-1)^1 = 3; \qquad a_2 = 4 + (-1)^2 = 5;$

   $a_3 = 4 + (-1)^3 = 3; \qquad a_4 = 4 + (-1)^4 = 5.$

   Note that the range of this sequence is finite, consisting of only the values 3 and 5. This sequence is plotted in Figure 9.1(b).

3. $a_1 = \dfrac{(-1)^{1(2)/2}}{1^2} = -1; \qquad a_2 = \dfrac{(-1)^{2(3)/2}}{2^2} = -\dfrac{1}{4}$

   $a_3 = \dfrac{(-1)^{3(4)/2}}{3^2} = \dfrac{1}{9} \qquad a_4 = \dfrac{(-1)^{4(5)/2}}{4^2} = \dfrac{1}{16};$

   $a_5 = \dfrac{(-1)^{5(6)/2}}{5^2} = -\dfrac{1}{25}.$

   We gave one extra term to begin to show the pattern of signs is "$-, -, +, +, -, -, \ldots$", due to the fact that the exponent of -1 is a special quadratic. This sequence is plotted in Figure 9.1(c).

Example 2   **Determining a formula for a sequence**
Find the nth term of the following sequences, i.e., find a function that describes each of the given sequences.

Notes:

1. 2, 5, 8, 11, 14, . . .

2. 2, −5, 10, −17, 26, −37, . . .

3. 1, 1, 2, 6, 24, 120, 720, . . .

4. $\dfrac{5}{2}, \dfrac{5}{2}, \dfrac{15}{8}, \dfrac{5}{4}, \dfrac{25}{32}, \dots$

   Solution   We should first note that there is never exactly one function that describes a finite set of numbers as a sequence. There are many sequences that start with 2, then 5, as our first example does. We are looking for a simple formula that describes the terms given, knowing there is possibly more than one answer.

1. Note how each term is 3 more than the previous one. This implies a linear function would be appropriate: $a(n) = a_n = 3n + b$ for some appropriate value of b. As we want $a_1 = 2$, we set $b = -1$. Thus $a_n = 3n - 1$.

2. First notice how the sign changes from term to term. This is most commonly accomplished by multiplying the terms by either $(-1)^n$ or $(-1)^{n+1}$. Using $(-1)^n$ multiplies the odd terms by (-1); using $(-1)^{n+1}$ multiplies the even terms by (-1). As this sequence has negative even terms, we will multiply by $(-1)^{n+1}$.

   After this, we might feel a bit stuck as to how to proceed. At this point, we are just looking for a pattern of some sort: what do the numbers 2, 5, 10, 17, etc., have in common? There are many correct answers, but the one that we'll use here is that each is one more than a perfect square. That is, $2 = 1^2 + 1$, $5 = 2^2 + 1$, $10 = 3^2 + 1$, etc. Thus our formula is $a_n = (-1)^{n+1}(n^2 + 1)$.

3. One who is familiar with the factorial function will readily recognize these numbers. They are 0!, 1!, 2!, 3!, etc. Since our sequences start with $n = 1$, we cannot write $a_n = n!$, for this misses the 0! term. Instead, we shift by 1, and write $a_n = (n - 1)!$.

4. This one may appear difficult, especially as the first two terms are the same, but a little "sleuthing" will help. Notice how the terms in the numerator are always multiples of 5, and the terms in the denominator are always powers of 2. Does something as simple as $a_n = \dfrac{5n}{2^n}$ work?

   When $n = 1$, we see that we indeed get 5/2 as desired. When $n = 2$, we get $10/4 = 5/2$. Further checking shows that this formula indeed matches the other terms of the sequence.

Notes:

A common mathematical endeavor is to create a new mathematical object (for instance, a sequence) and then apply previously known mathematics to the new object. We do so here. The fundamental concept of calculus is the limit, so we will investigate what it means to find the limit of a sequence.

Definition 31   Limit of a Sequence, Convergent, Divergent

Let $\{a_n\}$ be a sequence and let L be a real number. Given any $\varepsilon > 0$, if an m can be found such that $|a_n - L| < \varepsilon$ for all $n > m$, then we say the **limit of** $\{a_n\}$**, as** n **approaches infinity, is** L, denoted

$$\lim_{n \to \infty} a_n = L.$$

If $\lim_{n \to \infty} a_n$ exists, we say the sequence **converges**; otherwise, the sequence **diverges**.

This definition states, informally, that if the limit of a sequence is L, then if you go far enough out along the sequence, all subsequent terms will be *really close* to L. Of course, the terms "far enough" and "really close" are subjective terms, but hopefully the intent is clear.

This definition is reminiscent of the ε–δ proofs of Chapter 1. In that chapter we developed other tools to evaluate limits apart from the formal definition; we do so here as well.

Theorem 58   Limit of a Sequence

Let $\{a_n\}$ be a sequence and let $f(x)$ be a function whose domain contains the positive real numbers where $f(n) = a_n$ for all n in $\mathbb{N}$.

If $\lim_{x \to \infty} f(x) = L$, then $\lim_{n \to \infty} a_n = L$.

Theorem 58 allows us, in certain cases, to apply the tools developed in Chapter 1 to limits of sequences. Note two things *not* stated by the theorem:

1. If $\lim_{x \to \infty} f(x)$ does not exist, we cannot conclude that $\lim_{n \to \infty} a_n$ does not exist. It may, or may not, exist. For instance, we can define a sequence $\{a_n\} = \{\cos(2\pi n)\}$. Let $f(x) = \cos(2\pi x)$. Since the cosine function oscillates over the real numbers, the limit $\lim_{x \to \infty} f(x)$ does not exist.

   However, for every positive integer n, $\cos(2\pi n) = 1$, so $\lim_{n \to \infty} a_n = 1$.

Notes:

2. If we cannot find a function $f(x)$ whose domain contains the positive real numbers where $f(n) = a_n$ for all n in $\mathbb{N}$, we cannot conclude $\lim\limits_{n\to\infty} a_n$ does not exist. It may, or may not, exist.

Example 3   **Determining convergence/divergence of a sequence**

Determine the convergence or divergence of the following sequences.

1. $\{a_n\} = \left\{ \dfrac{3n^2 - 2n + 1}{n^2 - 1000} \right\}$ 2. $\{a_n\} = \{\cos n\}$ 3. $\{a_n\} = \left\{ \dfrac{(-1)^n}{n} \right\}$

SOLUTION

1. Using Key Idea 2, we can state that $\lim\limits_{x\to\infty} \dfrac{3x^2 - 2x + 1}{x^2 - 1000} = 3$. (We could have also directly applied L'Hôpital's Rule.) Thus the sequence $\{a_n\}$ converges, and its limit is 3. A scatter plot of every 5 values of a_n is given in Figure 9.2 (a). The values of a_n vary widely near $n = 30$, ranging from about -73 to 125, but as n grows, the values approach 3.

2. The limit $\lim\limits_{x\to\infty} \cos x$ does not exist, as $\cos x$ oscillates (and takes on every value in $[-1, 1]$ infinitely many times). Thus we cannot apply Theorem 58.

   The fact that the cosine function oscillates strongly hints that $\cos n$, when n is restricted to $\mathbb{N}$, will also oscillate. Figure 9.2 (b), where the sequence is plotted, shows that this is true. Because only discrete values of cosine are plotted, it does not bear strong resemblance to the familiar cosine wave.

   Based on the graph, we suspect that $\lim\limits_{n\to\infty} a_n$ does not exist, but we have not decisively proven it yet.

3. We cannot actually apply Theorem 58 here, as the function $f(x) = (-1)^x/x$ is not well defined. (What does $(-1)^{\sqrt{2}}$ mean? In actuality, there is an answer, but it involves *complex analysis*, beyond the scope of this text.) So for now we say that we cannot determine the limit. (But we will be able to very soon.) By looking at the plot in Figure 9.2 (c), we would like to conclude that the sequence converges to 0. That is true, but at this point we are unable to decisively say so.

It seems that $\{(-1)^n/n\}$ converges to 0 but we lack the formal tool to prove it. The following theorem gives us that tool.

Theorem 59   **Absolute Value Theorem**

Let $\{a_n\}$ be a sequence. If $\lim\limits_{n\to\infty} |a_n| = 0$, then $\lim\limits_{n\to\infty} a_n = 0$

$a_n = \dfrac{3n^2 - 2n + 1}{n^2 - 1000}$

(a)

$a_n = \cos n$

(b)

$a_n = \dfrac{(-1)^n}{n}$
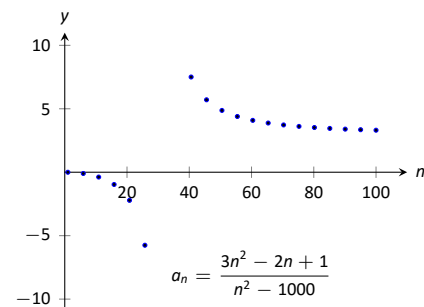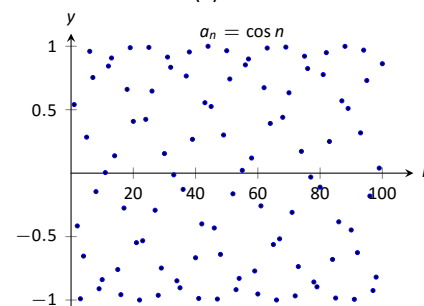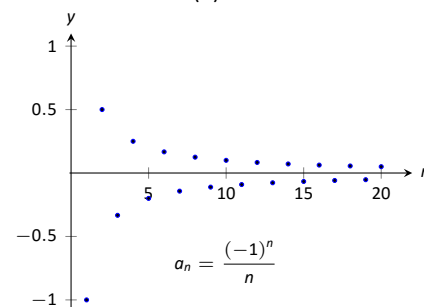
(c)

Figure 9.2: Scatter plots of the sequences in Example 3.

Notes:

Proof

We know $-|a_n| \le a_n \le |a_n|$ and $\lim\limits_{n \to \infty} (-|a_n|) = -\lim\limits_{n \to \infty} |a_n| = 0$. Thus by the Squeeze Theorem $\lim\limits_{n \to \infty} a_n = 0$. $\qquad\qquad\qquad\qquad\qquad\qquad \square$

Example 4   **Determining the convergence/divergence of a sequence**

Determine the convergence or divergence of the following sequences.

$$1.\ \{a_n\} = \left\{ \frac{(-1)^n}{n} \right\} \qquad 2.\ \{a_n\} = \left\{ \frac{(-1)^n(n+1)}{n} \right\}$$

SOLUTION

1. This appeared in Example 3. We want to apply Theorem 59, so consider the limit of $\{|a_n|\}$:

$$\lim_{n\to\infty} |a_n| = \lim_{n\to\infty} \left| \frac{(-1)^n}{n} \right|$$
$$= \lim_{n\to\infty} \frac{1}{n}$$
$$= 0.$$

Since this limit is 0, we can apply Theorem 59 and state that $\lim\limits_{n\to\infty} a_n = 0$.

2. Because of the alternating nature of this sequence (i.e., every other term is multiplied by -1), we cannot simply look at the limit $\lim\limits_{x\to\infty} \dfrac{(-1)^x(x+1)}{x}$. We can try to apply the techniques of Theorem 59:

$$\lim_{n\to\infty} |a_n| = \lim_{n\to\infty} \left| \frac{(-1)^n(n+1)}{n} \right|$$
$$= \lim_{n\to\infty} \frac{n+1}{n}$$
$$= 1.$$

We have concluded that when we ignore the alternating sign, the sequence approaches 1. This means we cannot apply Theorem 59; it states the the limit must be 0 in order to conclude anything.

Since we know that the signs of the terms alternate *and* we know that the limit of $|a_n|$ is 1, we know that as n approaches infinity, the terms will alternate between values close to 1 and -1, meaning the sequence diverges. A plot of this sequence is given in Figure 9.3.
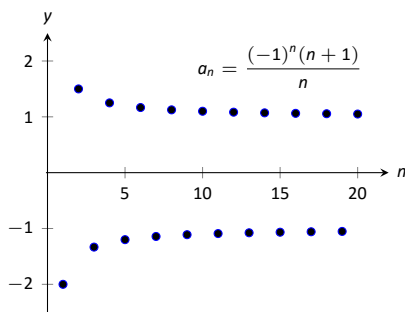
Figure 9.3: A plot of a sequence in Example 4, part 2.

Notes:

We continue our study of the limits of sequences by considering some of the properties of these limits.

> **Theorem 60   Properties of the Limits of Sequences**
>
> Let $\{a_n\}$ and $\{b_n\}$ be sequences such that $\lim\limits_{n\to\infty} a_n = L$, $\lim\limits_{n\to\infty} b_n = K$, and let c be a real number.
>
> 1. $\lim\limits_{n\to\infty} (a_n \pm b_n) = L \pm K$
>
> 2. $\lim\limits_{n\to\infty} (a_n \cdot b_n) = L \cdot K$
>
> 3. $\lim\limits_{n\to\infty} (a_n/b_n) = L/K,\ K \neq 0$
>
> 4. $\lim\limits_{n\to\infty} c \cdot a_n = c \cdot L$

Example 5   Applying properties of limits of sequences

Let the following limits be given:

- $\lim\limits_{n\to\infty} a_n = 0$;

- $\lim\limits_{n\to\infty} b_n = e$; and

- $\lim\limits_{n\to\infty} c_n = 5$.

Evaluate the following limits.

1. $\lim\limits_{n\to\infty} (a_n + b_n)$ 2. $\lim\limits_{n\to\infty} (b_n \cdot c_n)$ 3. $\lim\limits_{n\to\infty} (1000 \cdot a_n)$

SOLUTION   We will use Theorem 60 to answer each of these.

1. Since $\lim\limits_{n\to\infty} a_n = 0$ and $\lim\limits_{n\to\infty} b_n = e$, we conclude that $\lim\limits_{n\to\infty} (a_n + b_n) = 0 + e = e$. So even though we are adding something to each term of the sequence b_n, we are adding something so small that the final limit is the same as before.

2. Since $\lim\limits_{n\to\infty} b_n = e$ and $\lim\limits_{n\to\infty} c_n = 5$, we conclude that $\lim\limits_{n\to\infty} (b_n \cdot c_n) = e \cdot 5 = 5e$.

3. Since $\lim\limits_{n\to\infty} a_n = 0$, we have $\lim\limits_{n\to\infty} 1000a_n = 1000 \cdot 0 = 0$. It does not matter that we multiply each term by 1000; the sequence still approaches 0. (It just takes longer to get close to 0.)

Notes:

Definition 32   Geometric Sequence

For a constant r, the sequence $\{r^n\}$ is known as a **geometric sequence**.

Theorem 61   Convergence of Geometric Sequences

The sequence $\{r^n\}$ is convergent if $-1 < r \le 1$ and divergent for all other values of r. Furthermore,

$$\lim_{n\to\infty} r^n = \begin{cases} 0 & -1 < r < 1 \\ 1 & r = 1 \end{cases}$$

Proof

We can see from Key Idea 18 and by letting $a = r$ that

$$\lim_{n\to\infty} r^n = \begin{cases} \infty & r > 1 \\ 0 & 0 < r < 1. \end{cases}$$

We also know that $\lim_{x\to\infty} 1^n = 1$ and $\lim_{x\to\infty} 0^n = 0$. If $-1 < r < 0$, we know $0 < |r| < 1$ so $\lim_{x\to\infty} |r^n| = \lim_{x\to\infty} |r|^n = 0$ and thus by Theorem 59, $\lim_{x\to\infty} r^n = 0$. If $r \le -1$, $\lim_{x\to\infty} r^n$ does not exist. Therefore, the sequence $\{r^n\}$ is convergent if $-1 < r \le 1$ and divergent for all other values of r. $\square$

There is more to learn about sequences than just their limits. We will also study their range and the relationships terms have with the terms that follow. We start with some definitions describing properties of the range.

Definition 33   Bounded and Unbounded Sequences

A sequence $\{a_n\}$ is said to be **bounded** if there exists real numbers m and M such that $m < a_n < M$ for all n in $\mathbb{N}$.

A sequence $\{a_n\}$ is said to be **unbounded** if it is not bounded.

A sequence $\{a_n\}$ is said to be **bounded above** if there exists an M such that $a_n < M$ for all n in $\mathbb{N}$; it is **bounded below** if there exists an m such that $m < a_n$ for all n in $\mathbb{N}$.

Notes:

It follows from this definition that an unbounded sequence may be bounded above or bounded below; a sequence that is both bounded above and below is simply a bounded sequence.

Example 6   **Determining boundedness of sequences**
Determine the boundedness of the following sequences.

1. $\{a_n\} = \left\{\dfrac{1}{n}\right\}$ 2. $\{a_n\} = \{2^n\}$

SOLUTION

1. The terms of this sequence are always positive but are decreasing, so we have $0 < a_n < 2$ for all n. Thus this sequence is bounded. Figure 9.4(a) illustrates this.

2. The terms of this sequence obviously grow without bound. However, it is also true that these terms are all positive, meaning $0 < a_n$. Thus we can say the sequence is unbounded, but also bounded below. Figure 9.4(b) illustrates this.

The previous example produces some interesting concepts. First, we can recognize that the sequence $\{1/n\}$ converges to 0. This says, informally, that "most" of the terms of the sequence are "really close" to 0. This implies that the sequence is bounded, using the following logic. First, "most" terms are near 0, so we could find some sort of bound on these terms (using Definition 31, the bound is ε). That leaves a "few" terms that are not near 0 (i.e., a *finite* number of terms). A finite list of numbers is always bounded.

This logic suggests that if a sequence converges, it must be bounded. This is indeed true, as stated by the following theorem.

Theorem 62   **Convergent Sequences are Bounded**
Let $\{a_n\}$ be a convergent sequence. Then $\{a_n\}$ is bounded.

In Example 7.5.3 part 1, we found that $\displaystyle\lim_{x\to\infty}(1 + 1/x)^x = e$. If we consider the sequence $\{b_n\} = \{(1 + 1/n)^n\}$, we see that $\displaystyle\lim_{n\to\infty} b_n = e$. Even though it may be difficult to intuitively grasp the behavior of this sequence, we know immediately that it is bounded.

(a)

(b)
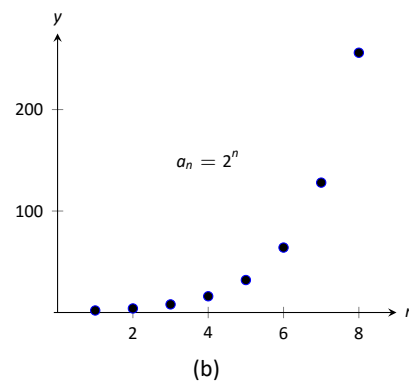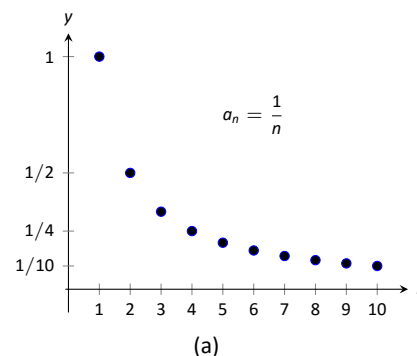
Figure 9.4: A plot of $\{a_n\} = \{1/n\}$ and $\{a_n\} = \{2^n\}$ from Example 6.

Note: Keep in mind what Theorem 62 does *not* say. It does not say that bounded sequences must converge, nor does it say that if a sequence does not converge, it is not bounded.

Notes:

Another interesting concept to come out of Example 6 again involves the sequence $\{1/n\}$. We stated, without proof, that the terms of the sequence were decreasing. That is, that $a_{n+1} < a_n$ for all n. (This is easy to show. Clearly $n < n + 1$. Taking reciprocals flips the inequality: $1/n > 1/(n + 1)$. This is the same as $a_n > a_{n+1}$.) Sequences that either steadily increase or decrease are important, so we give this property a name.

Definition 34   Monotonic Sequences

1. A sequence $\{a_n\}$ is **monotonically increasing** if $a_n \leq a_{n+1}$ for all n, i.e.,
$$a_1 \leq a_2 \leq a_3 \leq \cdots a_n \leq a_{n+1} \cdots$$

2. A sequence $\{a_n\}$ is **monotonically decreasing** if $a_n \geq a_{n+1}$ for all n, i.e.,
$$a_1 \geq a_2 \geq a_3 \geq \cdots a_n \geq a_{n+1} \cdots$$

3. A sequence is **monotonic** if it is monotonically increasing or monotonically decreasing.

Note: It is sometimes useful to call a monotonically increasing sequence *strictly increasing* if $a_n < a_{n+1}$ for all n; i.e, we remove the possibility that subsequent terms are equal.
A similar statement holds for *strictly decreasing*.

Example 7   Determining monotonicity
Determine the monotonicity of the following sequences.

1. $\{a_n\} = \left\{ \dfrac{n+1}{n} \right\}$

2. $\{a_n\} = \left\{ \dfrac{n^2 + 1}{n + 1} \right\}$

3. $\{a_n\} = \left\{ \dfrac{n^2 - 9}{n^2 - 10n + 26} \right\}$

4. $\{a_n\} = \left\{ \dfrac{n^2}{n!} \right\}$

SOLUTION   In each of the following, we will examine $a_{n+1} - a_n$. If $a_{n+1} - a_n \geq 0$, we conclude that $a_n \leq a_{n+1}$ and hence the sequence is increasing. If $a_{n+1} - a_n \leq 0$, we conclude that $a_n \geq a_{n+1}$ and the sequence is decreasing. Of course, a sequence need not be monotonic and perhaps neither of the above will apply.

We also give a scatter plot of each sequence. These are useful as they suggest a pattern of monotonicity, but analytic work should be done to confirm a graphical trend.

Notes:

1.

$$a_{n+1} - a_n = \frac{n+2}{n+1} - \frac{n+1}{n}$$
$$= \frac{(n+2)(n) - (n+1)^2}{(n+1)n}$$
$$= \frac{-1}{n(n+1)}$$
$$< 0 \quad \text{for all } n.$$

Since $a_{n+1} - a_n < 0$ for all n, we conclude that the sequence is decreasing.

2.

$$a_{n+1} - a_n = \frac{(n+1)^2 + 1}{n+2} - \frac{n^2 + 1}{n+1}$$
$$= \frac{((n+1)^2 + 1)(n+1) - (n^2 + 1)(n+2)}{(n+1)(n+2)}$$
$$= \frac{n^2 + 4n + 1}{(n+1)(n+2)}$$
$$> 0 \quad \text{for all } n.$$

Since $a_{n+1} - a_n > 0$ for all n, we conclude the sequence is increasing.

3. We can clearly see in Figure 9.7, where the sequence is plotted, that it is not monotonic. However, it does seem that after the first 4 terms it is decreasing. To understand why, perform the same analysis as done before:

$$a_{n+1} - a_n = \frac{(n+1)^2 - 9}{(n+1)^2 - 10(n+1) + 26} - \frac{n^2 - 9}{n^2 - 10n + 26}$$
$$= \frac{n^2 + 2n - 8}{n^2 - 8n + 17} - \frac{n^2 - 9}{n^2 - 10n + 26}$$
$$= \frac{(n^2 + 2n - 8)(n^2 - 10n + 26) - (n^2 - 9)(n^2 - 8n + 17)}{(n^2 - 8n + 17)(n^2 - 10n + 26)}$$
$$= \frac{-10n^2 + 60n - 55}{(n^2 - 8n + 17)(n^2 - 10n + 26)}.$$

We want to know when this is greater than, or less than, 0. The denominator is always positive, therefore we are only concerned with the numerator. Using the quadratic formula, we can determine that $-10n^2 + 60n - 55 = 0$ when $n \approx 1.13, 4.87$. So for $n < 1.13$, the sequence is decreasing. Since we are only dealing with the natural numbers, this means that $a_1 > a_2$.
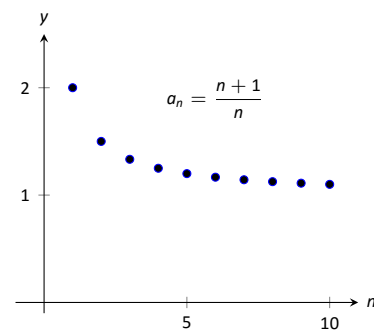
Figure 9.5: A plot of $\{a_n\} = \{\frac{n+1}{n}\}$ in Example 7(a).
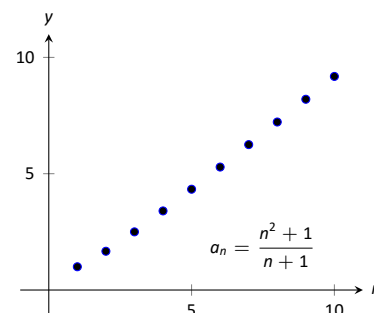
Figure 9.6: A plot of $\{a_n\} = \{\frac{n^2+1}{n+1}\}$ in Example 7(b).
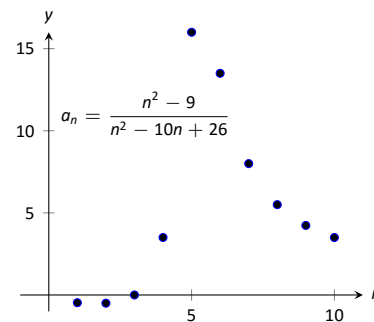
Figure 9.7: A plot of $\{a_n\} = \{\frac{n^2-9}{n^2-10n+26}\}$ in Example 7(c).

Notes:
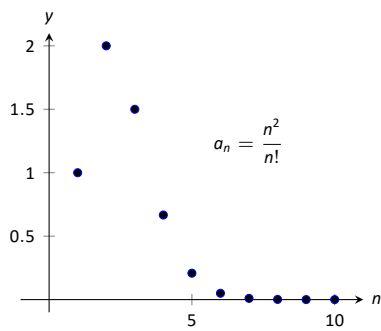
Figure 9.8: A plot of $\{a_n\} = \{n^2/n!\}$ in Example 7(d).

Between 1.13 and 4.87, i.e., for $n = 2$, 3 and 4, we have that $a_{n+1} > a_n$ and the sequence is increasing. (That is, when $n = 2$, 3 and 4, the numerator $-10n^2 + 60n + 55$ from the fraction above is > 0.)

When $n > 4.87$, i.e, for $n \geq 5$, we have that $-10n^2 + 60n + 55 < 0$, hence $a_{n+1} - a_n < 0$, so the sequence is decreasing.

In short, the sequence is simply not monotonic. However, it is useful to note that for $n \geq 5$, the sequence is monotonically decreasing.

4. Again, the plot in Figure 9.8 shows that the sequence is not monotonic, but it suggests that it is monotonically decreasing after the first term. Instead of looking at $a_{n+1} - a_n$, this time we'll look at a_n/a_{n+1}:

$$\frac{a_n}{a_{n+1}} = \frac{n^2}{n!}\frac{(n+1)!}{(n+1)^2}$$

$$= \frac{n^2}{n+1}$$

$$= n - 1 + \frac{1}{n+1}$$

When $n = 1$, the above expression is < 1; for $n \geq 2$, the above expression is > 1. Thus this sequence is not monotonic, but it is monotonically decreasing after the first term.

Knowing that a sequence is monotonic can be useful. In particular, if we know that a sequence is bounded and monotonic, we can conclude it converges. Consider, for example, a sequence that is monotonically decreasing and is bounded below. We know the sequence is always getting smaller, but that there is a bound to how small it can become. This is enough to prove that the sequence will converge, as stated in the following theorem.

> **Theorem 63   Bounded Monotonic Sequences are Convergent**
> Let $\{a_n\}$ be a bounded, monotonic sequence. Then $\{a_n\}$ converges; i.e., $\lim\limits_{n\to\infty} a_n$ exists.

Consider once again the sequence $\{a_n\} = \{1/n\}$. It is easy to show it is monotonically decreasing and that it is always positive (i.e., bounded below by 0). Therefore we can conclude by Theorem 63 that the sequence converges. We already knew this by other means, but in the following section this theorem will become very useful.

Notes:

Convergence of a sequence does not depend on the first N terms of a sequence. For example, we could adapt the sequence of the previous paragraph to be

$$1,\ 10,\ 100,\ 1000,\ \frac{1}{5},\ \frac{1}{6},\ \frac{1}{7},\ \frac{1}{8},\ \frac{1}{9},\ \frac{1}{10},\cdots$$

Because we only changed three of the first 4 terms, we have not affected whether the sequence converges or diverges.

Sequences are a great source of mathematical inquiry. The On-Line Encyclopedia of Integer Sequences (`http://oeis.org`) contains thousands of sequences and their formulae. (As of this writing, there are 257,537 sequences in the database.) Perusing this database quickly demonstrates that a single sequence can represent several different "real life" phenomena.

Interesting as this is, our interest actually lies elsewhere. We are more interested in the *sum* of a sequence. That is, given a sequence $\{a_n\}$, we are very interested in $a_1 + a_2 + a_3 + \cdots$. Of course, one might immediately counter with "Doesn't this just add up to 'infinity'?" Many times, yes, but there are many important cases where the answer is no. This is the topic of *series*, which we begin to investigate in the next section.

Notes:

Exercises 9.1

Terms and Concepts

1. Use your own words to define a *sequence.*

2. The domain of a sequence is the _____ numbers.

3. Use your own words to describe the *range* of a sequence.

4. Describe what it means for a sequence to be *bounded.*

Problems

In Exercises 5–8, give the first five terms of the given sequence.

5. $\{a_n\} = \left\{ \dfrac{4^n}{(n+1)!} \right\}$

6. $\{b_n\} = \left\{ \left(-\dfrac{3}{2} \right)^n \right\}$

7. $\{c_n\} = \left\{ -\dfrac{n^{n+1}}{n+2} \right\}$

8. $\{d_n\} = \left\{ \dfrac{1}{\sqrt{5}} \left(\left(\dfrac{1+\sqrt{5}}{2} \right)^n - \left(\dfrac{1-\sqrt{5}}{2} \right)^n \right) \right\}$

In Exercises 9–12, determine the nth term of the given sequence.

9. 4, 7, 10, 13, 16, . . .

10. $3, \ -\dfrac{3}{2}, \ \dfrac{3}{4}, \ -\dfrac{3}{8}, \ \ldots$

11. 10, 20, 40, 80, 160, . . .

12. $1, 1, \ \dfrac{1}{2}, \ \dfrac{1}{6}, \ \dfrac{1}{24}, \ \dfrac{1}{120}, \ \ldots$

In Exercises 13–16, use the following information to determine the limit of the given sequences.

- $\{a_n\} = \left\{ \dfrac{2^n - 20}{2^n} \right\};$ $\lim\limits_{n \to \infty} a_n = 1$

- $\{b_n\} = \left\{ \left(1 + \dfrac{2}{n} \right)^n \right\};$ $\lim\limits_{n \to \infty} b_n = e^2$

- $\{c_n\} = \{\sin(3/n)\};$ $\lim\limits_{n \to \infty} c_n = 0$

13. $\{a_n\} = \left\{ \dfrac{2^n - 20}{7 \cdot 2^n} \right\}$

14. $\{a_n\} = \{3b_n - a_n\}$

15. $\{a_n\} = \left\{ \sin(3/n) \left(1 + \dfrac{2}{n} \right)^n \right\}$

16. $\{a_n\} = \left\{ \left(1 + \dfrac{2}{n} \right)^{2n} \right\}$

In Exercises 17–39, determine whether the sequence converges or diverges. If convergent, give the limit of the sequence.

17. $\{a_n\} = \left\{ (-1)^n \dfrac{n}{n+1} \right\}$

18. $\{a_n\} = \left\{ \dfrac{4n^2 - n + 5}{3n^2 + 1} \right\}$

19. $\{a_n\} = \left\{ \dfrac{4^n}{5^n} \right\}$

20. $\{a_n\} = \left\{ \dfrac{(n-3)!}{(n+1)!} \right\}$

21. $\{a_n\} = \left\{ \dfrac{n-1}{n} - \dfrac{n}{n-1} \right\}, n \geq 2$

22. $\{a_n\} = \left\{ \dfrac{6^{n+3}}{8^n} \right\}$

23. $\{a_n\} = \{\ln(n)\}$

24. $\{a_n\} = \left\{ \dfrac{3n}{\sqrt{n^2 + 1}} \right\}$

25. $\{a_n\} = \left\{ \left(1 + \dfrac{1}{n} \right)^n \right\}$

26. $\{a_n\} = \left\{ \dfrac{(2n+1)!}{(2n-1)!} \right\}$

27. $\{a_n\} = \left\{ 5 - \dfrac{1}{n} \right\}$

28. $\{a_n\} = \left\{ \dfrac{(-1)^{n+1}}{n} \right\}$

29. $\{a_n\} = \left\{ \dfrac{1.1^n}{n} \right\}$

30. $\{a_n\} = \left\{ \dfrac{2n}{n+1} \right\}$

31. $\{a_n\} = \left\{ (-1)^n \dfrac{n^2}{2^n - 1} \right\}$

32. $\{a_n\} = \left\{ 2 + \dfrac{9^n}{8^n} \right\}$

33. $\{a_n\} = \left\{ \dfrac{(n-1)!}{(n+1)!} \right\}$

34. $\{a_n\} = \{\ln(3n+2) - \ln n\}$

35. $\{a_n\} = \{\ln(2n^2 + 3n + 1) - \ln(n^2 + 1)\}$

36. $\{a_n\} = \left\{ n \sin\left(\dfrac{1}{n} \right) \right\}$

37. $\{a_n\} = \left\{ \dfrac{\cos^2 n}{2^n} \right\}$

38. $\{a_n\} = \left\{ \dfrac{e^n + e^{-n}}{e^{2n} - 1} \right\}$

39. $\{a_n\} = \left\{ \dfrac{\ln n}{\ln 2n} \right\}$

In Exercises 40–43, determine whether the sequence is bounded, bounded above, bounded below, or none of the above.

40. $\{a_n\} = \{\sin n\}$

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

42. $\{a_n\} = \left\{ \dfrac{3n^2 - 1}{n} \right\}$

43. $\{a_n\} = \{2^n - n!\}$

In Exercises 44–49, determine whether the sequence is monotonically increasing or decreasing. If it is not, determine if there is an m such that it is monotonic for all $n \geq m$.

44. $\{a_n\} = \left\{ \dfrac{n}{n+2} \right\}$

45. $\{a_n\} = \left\{ \dfrac{n^2 - 6n + 9}{n} \right\}$

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

47. $\{a_n\} = \left\{ \dfrac{n^2}{2^n} \right\}$

48. $\{a_n\} = \left\{ \cos\left(\dfrac{n\pi}{2} \right) \right\}$

49. $\{a_n\} = \{ne^{-n}\}$

50. Prove Theorem 59; that is, use the definition of the limit of a sequence to show that if $\lim\limits_{n \to \infty} |a_n| = 0$, then $\lim\limits_{n \to \infty} a_n = 0$.

51. Let $\{a_n\}$ and $\{b_n\}$ be sequences such that $\lim\limits_{n \to \infty} a_n = L$ and $\lim\limits_{n \to \infty} b_n = K$.

   (a) Show that if $a_n < b_n$ for all n, then $L \leq K$.

   (b) Give an example where $L = K$.

52. Prove the Squeeze Theorem for sequences: Let $\{a_n\}$ and $\{b_n\}$ be such that $\lim\limits_{n \to \infty} a_n = L$ and $\lim\limits_{n \to \infty} b_n = L$, and let $\{c_n\}$ be such that $a_n \leq c_n \leq b_n$ for all n. Then $\lim\limits_{n \to \infty} c_n = L$

9.2   Infinite Series

Given the sequence $\{a_n\} = \{1/2^n\} = 1/2,\ 1/4,\ 1/8,\ \ldots$, consider the following sums:

$$
\begin{array}{rcccl}
a_1 & = & 1/2 & = & 1/2 \\
a_1 + a_2 & = & 1/2 + 1/4 & = & 3/4 \\
a_1 + a_2 + a_3 & = & 1/2 + 1/4 + 1/8 & = & 7/8 \\
a_1 + a_2 + a_3 + a_4 & = & 1/2 + 1/4 + 1/8 + 1/16 & = & 15/16
\end{array}
$$

Later, we will be able to show that

$$
a_1 + a_2 + a_3 + \cdots + a_n = \frac{2^n - 1}{2^n} = 1 - \frac{1}{2^n}.
$$

Let S_n be the sum of the first n terms of the sequence $\{1/2^n\}$. From the above, we see that $S_1 = 1/2$, $S_2 = 3/4$, and that $S_n = 1 - 1/2^n$.

Now consider the following limit: $\lim\limits_{n\to\infty} S_n = \lim\limits_{n\to\infty} \left(1 - 1/2^n\right) = 1$. This limit can be interpreted as saying something amazing: *the sum of* all *the terms of the sequence* $\{1/2^n\}$ *is 1.*

This example illustrates some interesting concepts that we explore in this section. We begin this exploration with some definitions.

Definition 35   Infinite Series, n^{th} Partial Sums, Convergence, Divergence

Let $\{a_n\}$ be a sequence.

1. The sum $\displaystyle\sum_{n=1}^{\infty} a_n$ is an **infinite series** (or, simply **series**).

2. Let $\displaystyle S_n = \sum_{i=1}^{n} a_i$; the sequence $\{S_n\}$ is the sequence of nth **partial sums** of $\{a_n\}$.

3. If the sequence $\{S_n\}$ converges to L, we say the series $\displaystyle\sum_{n=1}^{\infty} a_n$ **converges** to L, and we write $\displaystyle\sum_{n=1}^{\infty} a_n = L$.

4. If the sequence $\{S_n\}$ diverges, the series $\displaystyle\sum_{n=1}^{\infty} a_n$ **diverges**.

Notes:

Using our new terminology, we can state that the series $\sum_{n=1}^{\infty} 1/2^n$ converges,

and $\sum_{n=1}^{\infty} 1/2^n = 1$.

We will explore a variety of series in this section. We start with two series that diverge, showing how we might discern divergence.

Example 1　　　**Showing series diverge**

1. Let $\{a_n\} = \{n^2\}$. Show $\sum_{n=1}^{\infty} a_n$ diverges.

2. Let $\{b_n\} = \{(-1)^{n+1}\}$. Show $\sum_{n=1}^{\infty} b_n$ diverges.

SOLUTION

1. Consider S_n, the nth partial sum.

$$
\begin{aligned}
S_n &= a_1 + a_2 + a_3 + \cdots + a_n \\
&= 1^2 + 2^2 + 3^2 \cdots + n^2 \\
&= \frac{n(n+1)(2n+1)}{6}. \qquad \text{by Theorem 35}
\end{aligned}
$$

Since $\lim_{n\to\infty} S_n = \infty$, we conclude that the series $\sum_{n=1}^{\infty} n^2$ diverges. It is instructive to write $\sum_{n=1}^{\infty} n^2 = \infty$ for this tells us *how* the series diverges: it grows without bound.

A scatter plot of the sequences $\{a_n\}$ and $\{S_n\}$ is given in Figure 9.9. The terms of $\{a_n\}$ are growing, so the terms of the partial sums $\{S_n\}$ are growing even faster, illustrating that the series diverges.
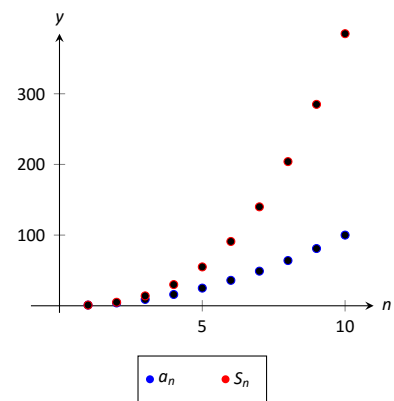
Figure 9.9: Scatter plots relating to the series of Example 1 part 1.

Notes:
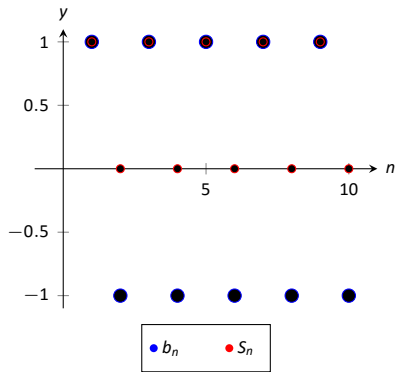
Figure 9.10: Scatter plots relating to the series of Example 1 part 2.

2. The sequence $\{b_n\}$ starts with $1, -1, 1, -1, \ldots$. Consider some of the partial sums S_n of $\{b_n\}$:

$$S_1 = 1$$
$$S_2 = 0$$
$$S_3 = 1$$
$$S_4 = 0$$

This pattern repeats; we find that $S_n = \begin{cases} 1 & n \text{ is odd} \\ 0 & n \text{ is even} \end{cases}$. As $\{S_n\}$ oscillates, repeating $1, 0, 1, 0, \ldots$, we conclude that $\lim_{n\to\infty} S_n$ does not exist, hence $\sum_{n=1}^{\infty} (-1)^{n+1}$ diverges.

A scatter plot of the sequence $\{b_n\}$ and the partial sums $\{S_n\}$ is given in Figure 9.10. When n is odd, $b_n = S_n$ so the marks for b_n are drawn oversized to show they coincide.

While it is important to recognize when a series diverges, we are generally more interested in the series that converge. In this section we will demonstrate a few general techniques for determining convergence; later sections will delve deeper into this topic.

Geometric Series

One important type of series is a *geometric series*.

> **Definition 36   Geometric Series**
> A **geometric series** is a series of the form
> $$\sum_{n=0}^{\infty} ar^n = a + ar + ar^2 + ar^3 + \cdots + ar^n + \cdots$$
> Note that the index starts at $n = 0$, if the index starts at $n = 1$ we have $\sum_{n=1}^{\infty} ar^{n-1}$.

We started this section with a geometric series, although we dropped the first term of 1. One reason geometric series are important is that they have nice convergence properties.

Notes:

Theorem 64   Convergence of Geometric Series

Consider the geometric series $\sum\limits_{n=0}^{\infty} ar^n$.

1. If $r \neq 1$, the nth partial sum is: $S_n = \dfrac{a(1 - r^n)}{1 - r}$.

2. The series converges if, and only if, $|r| < 1$. When $|r| < 1$,

$$\sum_{n=0}^{\infty} ar^n = \frac{a}{1 - r}.$$

Proof

If $r = 1$, then $S_n = a + a + a + \cdots + a = na$. Since $\lim\limits_{n \to \infty} S_n = \pm\infty$, the geometric series diverges.

If $r \neq 1$, we have

$$S_n = a + ar + ar^2 + \cdots + ar^{n-1}.$$

Multiply each term by r and we have

$$rS_n = ar + ar^2 + ar^3 \cdots + ar^n.$$

Subtract these two equations and solve for S_n.

$$S_n - rS_n = a - ar^n$$
$$S_n = \frac{a(1 - r^n)}{1 - r}$$

From Theorem 61, we know that if $-1 < r < 1$, then $\lim\limits_{n \to \infty} r^n = 0$ so

$$\lim_{n \to \infty} S_n = \lim_{n \to \infty} = \frac{a(1 - r^n)}{1 - r} = \frac{a}{1 - r} - \frac{a}{1 - r} \lim_{n \to \infty} r^n = \frac{a}{1 - r}.$$

So when $|r| < 1$ the geometric series converges and its sum is $\dfrac{a}{1 - r}$.

If either $r \leq -1$ or $r > 1$, the sequence $\{r^n\}$ is divergent by Theorem 61. Thus $\lim\limits_{n \to \infty} S_n$ does not exist, so the geometric series diverges if $r \leq -1$ or $r > 1$. $\square$

Notes:

According to Theorem 64, the series

$$\sum_{n=0}^{\infty} \frac{1}{2^n} = \sum_{n=0}^{\infty} \left(\frac{1}{2}\right)^2 = 1 + \frac{1}{2} + \frac{1}{4} + \cdots$$

converges as $r = 1/2$, and $\sum_{n=0}^{\infty} \frac{1}{2^n} = \frac{1}{1 - 1/2} = 2$. This concurs with our intro-

ductory example; while there we got a sum of 1, we skipped the first term of 1.

Example 2   Exploring geometric series
Check the convergence of the following series. If the series converges, find its sum.

$$1.\ \sum_{n=2}^{\infty} \left(\frac{3}{4}\right)^n \qquad 2.\ \sum_{n=0}^{\infty} \left(\frac{-1}{2}\right)^n \qquad 3.\ \sum_{n=0}^{\infty} 3^n$$

SOLUTION

1. Since $r = 3/4 < 1$, this series converges. By Theorem 64, we have that

$$\sum_{n=0}^{\infty} \left(\frac{3}{4}\right)^n = \frac{1}{1 - 3/4} = 4.$$

However, note the subscript of the summation in the given series: we are to start with $n = 2$. Therefore we subtract off the first two terms, giving:

$$\sum_{n=2}^{\infty} \left(\frac{3}{4}\right)^n = 4 - 1 - \frac{3}{4} = \frac{9}{4}.$$

This is illustrated in Figure 9.11(a).

2. Since $|r| = 1/2 < 1$, this series converges, and by Theorem 64,

$$\sum_{n=0}^{\infty} \left(\frac{-1}{2}\right)^n = \frac{1}{1 - (-1/2)} = \frac{2}{3}.$$

The partial sums of this series are plotted in Figure 9.11(b). Note how the partial sums are not purely increasing as some of the terms of the sequence $\{(-1/2)^n\}$ are negative.

3. Since $r > 1$, the series diverges. (This makes "common sense"; we expect the sum

$$1 + 3 + 9 + 27 + 81 + 243 + \cdots$$

to diverge.) This is illustrated in Figure 9.11(c).

Notes:

(a)

(b)

(c)
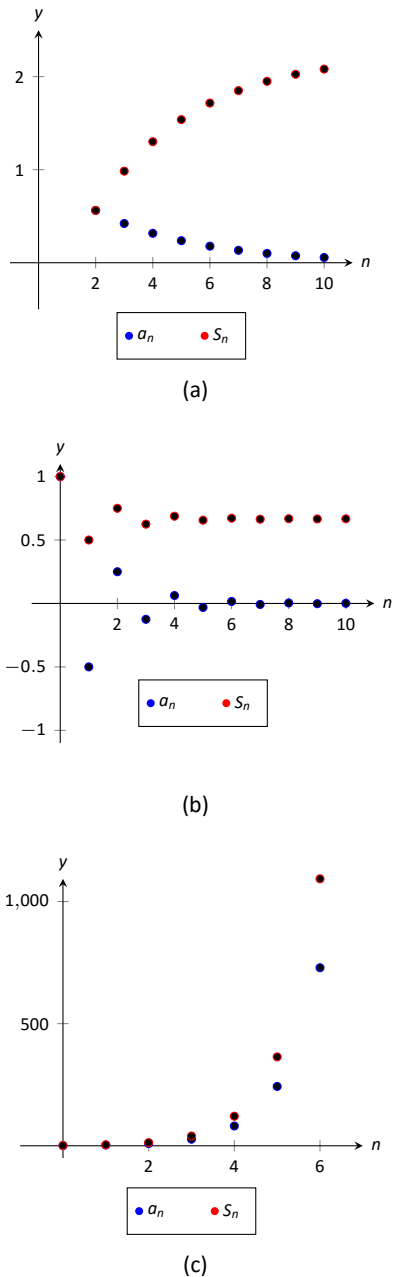
Figure 9.11: Scatter plots relating to the series in Example 2.

Later sections will provide tests by which we can determine whether or not a given series converges. This, in general, is much easier than determining *what* a given series converges to. There are many cases, though, where the sum can be determined.

Example 3   **Telescoping series**

Evaluate the sum $\displaystyle\sum_{n=1}^{\infty} \left(\frac{1}{n} - \frac{1}{n+1} \right)$.

SOLUTION   It will help to write down some of the first few partial sums of this series.

$$S_1 = \frac{1}{1} - \frac{1}{2} \qquad\qquad = 1 - \frac{1}{2}$$

$$S_2 = \left(\frac{1}{1} - \frac{1}{2}\right) + \left(\frac{1}{2} - \frac{1}{3}\right) \qquad\qquad = 1 - \frac{1}{3}$$

$$S_3 = \left(\frac{1}{1} - \frac{1}{2}\right) + \left(\frac{1}{2} - \frac{1}{3}\right) + \left(\frac{1}{3} - \frac{1}{4}\right) \qquad = 1 - \frac{1}{4}$$

$$S_4 = \left(\frac{1}{1} - \frac{1}{2}\right) + \left(\frac{1}{2} - \frac{1}{3}\right) + \left(\frac{1}{3} - \frac{1}{4}\right) + \left(\frac{1}{4} - \frac{1}{5}\right) = 1 - \frac{1}{5}$$

Note how most of the terms in each partial sum subtract out. In general, we see that $S_n = 1 - \dfrac{1}{n+1}$. The sequence $\{S_n\}$ converges, as $\displaystyle\lim_{n\to\infty} S_n = \lim_{n\to\infty}\left(1 - \frac{1}{n+1}\right) = 1$, and so we conclude that $\displaystyle\sum_{n=1}^{\infty}\left(\frac{1}{n} - \frac{1}{n+1}\right) = 1$. Partial sums of the series are plotted in Figure 9.12.

The series in Example 3 is an example of a **telescoping series**. Informally, a telescoping series is one in which the partial sums reduce to just a finite number of terms. The partial sum S_n did not contain n terms, but rather just two: 1 and $1/(n+1)$.

When possible, seek a way to write an explicit formula for the nth partial sum S_n. This makes evaluating the limit $\displaystyle\lim_{n\to\infty} S_n$ much more approachable. We do so in the next example.

Figure 9.12: Scatter plots relating to the series of Example 3.

Example 4   **Evaluating series**

Evaluate each of the following infinite series.

1. $\displaystyle\sum_{n=1}^{\infty} \frac{2}{n^2 + 2n}$ 2. $\displaystyle\sum_{n=1}^{\infty} \ln\left(\frac{n+1}{n}\right)$

Notes:

SOLUTION

1. We can decompose the fraction $2/(n^2 + 2n)$ as

$$\frac{2}{n^2 + 2n} = \frac{1}{n} - \frac{1}{n+2}.$$

(See Section 8.4, Partial Fraction Decomposition, to recall how this is done, if necessary.)

Expressing the terms of $\{S_n\}$ is now more instructive:

$$S_1 = 1 - \frac{1}{3} \qquad\qquad\qquad\qquad\qquad\qquad = 1 - \frac{1}{3}$$

$$S_2 = \left(1 - \frac{1}{3}\right) + \left(\frac{1}{2} - \frac{1}{4}\right) \qquad\qquad\qquad = 1 + \frac{1}{2} - \frac{1}{3} - \frac{1}{4}$$

$$S_3 = \left(1 - \frac{1}{3}\right) + \left(\frac{1}{2} - \frac{1}{4}\right) + \left(\frac{1}{3} - \frac{1}{5}\right) \qquad = 1 + \frac{1}{2} - \frac{1}{4} - \frac{1}{5}$$

$$S_4 = \left(1 - \frac{1}{3}\right) + \left(\frac{1}{2} - \frac{1}{4}\right) + \left(\frac{1}{3} - \frac{1}{5}\right) + \left(\frac{1}{4} - \frac{1}{6}\right) \qquad = 1 + \frac{1}{2} - \frac{1}{5} - \frac{1}{6}$$

$$S_5 = \left(1 - \frac{1}{3}\right) + \left(\frac{1}{2} - \frac{1}{4}\right) + \left(\frac{1}{3} - \frac{1}{5}\right) + \left(\frac{1}{4} - \frac{1}{6}\right) + \left(\frac{1}{5} - \frac{1}{7}\right) \quad = 1 + \frac{1}{2} - \frac{1}{6} - \frac{1}{7}$$

We again have a telescoping series. In each partial sum, most of the terms pair up to add to zero and we obtain the formula $S_n = 1 + \dfrac{1}{2} - \dfrac{1}{n+1} - \dfrac{1}{n+2}$. Taking limits allows us to determine the convergence of the series:

$$\lim_{n\to\infty} S_n = \lim_{n\to\infty} \left(1 + \frac{1}{2} - \frac{1}{n+1} - \frac{1}{n+2}\right) = \frac{3}{2}, \quad \text{so} \quad \sum_{n=1}^{\infty} \frac{1}{n^2 + 2n} = \frac{3}{2}.$$

This is illustrated in Figure 9.13.

2. We begin by writing the first few partial sums of the series:

$$S_1 = \ln(2)$$

$$S_2 = \ln(2) + \ln\left(\frac{3}{2}\right)$$

$$S_3 = \ln(2) + \ln\left(\frac{3}{2}\right) + \ln\left(\frac{4}{3}\right)$$

$$S_4 = \ln(2) + \ln\left(\frac{3}{2}\right) + \ln\left(\frac{4}{3}\right) + \ln\left(\frac{5}{4}\right)$$

Figure 9.13: Scatter plots relating to the series of Example 4 part 1.

Notes:

At first, this does not seem helpful, but recall the logarithmic identity: $\ln x + \ln y = \ln(xy)$. Applying this to S_4 gives:

$$S_4 = \ln(2) + \ln\left(\frac{3}{2}\right) + \ln\left(\frac{4}{3}\right) + \ln\left(\frac{5}{4}\right) = \ln\left(\frac{2}{1} \cdot \frac{3}{2} \cdot \frac{4}{3} \cdot \frac{5}{4}\right) = \ln(5).$$

We must generalize this for S_n.

$$S_n = \ln(2) + \ln\left(\frac{3}{2}\right) + \cdots + \ln\left(\frac{n+1}{n}\right) = \ln\left(\frac{2}{1} \cdot \frac{3}{2} \cdots \frac{n}{n-1} \cdot \frac{n+1}{n}\right) = \ln(n+1)$$

We can conclude that $\{S_n\} = \{\ln(n+1)\}$. This sequence does not converge, as $\lim\limits_{n \to \infty} S_n = \infty$. Therefore $\sum\limits_{n=1}^{\infty} \ln\left(\frac{n+1}{n}\right) = \infty$; the series diverges. Note in Figure 9.14 how the sequence of partial sums grows slowly; after 100 terms, it is not yet over 5. Graphically we may be fooled into thinking the series converges, but our analysis above shows that it does not.
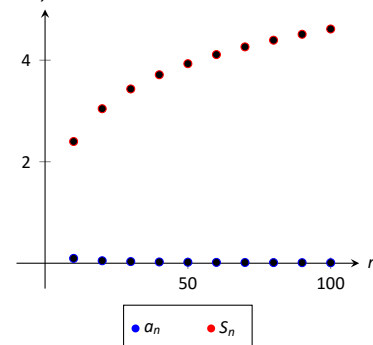
Figure 9.14: Scatter plots relating to the series of Example 4 part 2.

We are learning about a new mathematical object, the series. As done before, we apply "old" mathematics to this new topic.

Theorem 65   Properties of Infinite Series

Suppose that $\sum\limits_{n=1}^{\infty} a_n$ and $\sum\limits_{n=1}^{\infty} b_n$ are convergent series, and that

$\sum\limits_{n=1}^{\infty} a_n = L$, $\sum\limits_{n=1}^{\infty} b_n = K$, and c is a constant.

1. Constant Multiple Rule: $\sum\limits_{n=1}^{\infty} c \cdot a_n = c \cdot \sum\limits_{n=1}^{\infty} a_n = c \cdot L.$

2. Sum/Difference Rule: $\sum\limits_{n=1}^{\infty} (a_n \pm b_n) = \sum\limits_{n=1}^{\infty} a_n \pm \sum\limits_{n=1}^{\infty} b_n = L \pm K.$

Before using this theorem, we will consider the harmonic series $\sum\limits_{n=1}^{\infty} \frac{1}{n}$.

Notes:

Example 5 **Divergence of the Harmonic Series**

Show that the harmonic series $\sum\limits_{n=1}^{\infty} \dfrac{1}{n}$ diverges.

SOLUTION We will use a proof by contradiction here. Suppose the harmonic series converges to S. That is

$$S = 1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \frac{1}{5} + \frac{1}{6} + \frac{1}{7} + \frac{1}{8} + \cdots$$

We then have

$$S \geq 1 + \frac{1}{2} + \frac{1}{4} + \frac{1}{4} + \frac{1}{6} + \frac{1}{6} + \frac{1}{8} + \frac{1}{8} + \cdots$$

$$= 1 + \frac{1}{2} + \frac{1}{2} \quad + \frac{1}{3} \quad + \frac{1}{4} \quad + \cdots$$

$$= \frac{1}{2} + S$$

This gives us $S \geq \frac{1}{2} + S$ which can never be true, thus our assumption that the harmonic series converges must be false. Therefore, the harmonic series diverges.

It may take a while before one is comfortable with this statement, whose truth lies at the heart of the study of infinite series: *it is possible that the sum of an infinite list of nonzero numbers is finite.* We have seen this repeatedly in this section, yet it still may "take some getting used to."

As one contemplates the behavior of series, a few facts become clear.

1. In order to add an infinite list of nonzero numbers and get a finite result, "most" of those numbers must be "very near" 0.

2. If a series diverges, it means that the sum of an infinite list of numbers is not finite (it may approach $\pm\infty$ or it may oscillate), and:

    (a) The series will still diverge if the first term is removed.

    (b) The series will still diverge if the first 10 terms are removed.

    (c) The series will still diverge if the first 1,000,000 terms are removed.

    (d) The series will still diverge if any finite number of terms from anywhere in the series are removed.

These concepts are very important and lie at the heart of the next two theorems.

Notes:

> **Theorem 66**   **Convergence of Sequence**
>
> If the series $\displaystyle\sum_{n=1}^{\infty} a_n$ converges, then $\displaystyle\lim_{n\to\infty} a_n = 0$.

Proof

Let $S_n = a_1 + a_2 + \cdots + a_n$. We have

$$S_n = a_1 + a_2 + \cdots + a_{n-1} + a_n$$
$$S_n = S_{n-1} + a_n$$
$$a_n = S_n - S_{n-1}$$

Since $\displaystyle\sum_{n\to\infty} a_n$ converges, the sequence $\{S_n\}$ converges. Let $\displaystyle\lim_{n\to\infty} S_n = S$. As $n \to \infty$, $n - 1$ also goes to ∞, so $\displaystyle\lim_{n\to\infty} S_{n-1} = S$. We now have

$$\lim_{n\to\infty} a_n = \lim_{n\to\infty} (S_n - S_{n-1})$$
$$= \lim_{n\to\infty} S_n - \lim_{n\to\infty} S_{n-1}$$
$$= S - S = 0 \qquad \qquad \square$$

> **Theorem 67**   **Test for Divergence**
>
> If $\displaystyle\lim_{n\to\infty} a_n$ does not exist or $\displaystyle\lim_{n\to\infty} a_n \neq 0$, then the series $\displaystyle\sum_{n=1}^{\infty} a_n$ diverges.

   The Test for Divergence follows from Theorem 66. If the series does not diverge, it must converge and therefore $\displaystyle\lim_{n\to\infty} a_n = 0$.

   Note that the two statements in Theorems 66 and 67 are really the same. In order to converge, the limit of the terms of the sequence must approach 0; if they do not, the series will not converge.

   Looking back, we can apply this theorem to the series in Example 1. In that example, we had $\{a_n\} = \{n^2\}$ and $\{b_n\} = \{(-1)^{n+1}\}$.

$$\lim_{n\to\infty} a_n = \lim_{n\to\infty} n^2 = \infty$$

and

$$\lim_{n\to\infty} b_n = \lim_{n\to\infty} (-1)^{n+1} \text{ which does not exist.}$$

Notes:

Thus by the Test for Divergence, both series will diverge.

Important! This theorem *does not state* that if $\lim\limits_{n\to\infty} a_n = 0$ then $\sum\limits_{n=1}^{\infty} a_n$ converges. The standard example of this is the Harmonic Series, as given in Example 5. The Harmonic Sequence, $\{1/n\}$, converges to 0; the Harmonic Series, $\sum\limits_{n=1}^{\infty} 1/n$, diverges.

Theorem 68   Infinite Nature of Series

The convergence or divergence remains unchanged by the insertion or deletion of any finite number of terms. That is:

1. A divergent series will remain divergent with the insertion or deletion of any finite number of terms.

2. A convergent series will remain convergent with the insertion or deletion of any finite number of terms. (Of course, the *sum* will likely change.)

In other words, when we are only interested in the convergence or divergence of a series, it is safe to ignore the first few billion terms.

Example 6   Removing Terms from the Harmonic Series

Consider once more the Harmonic Series $\sum\limits_{n=1}^{\infty} \dfrac{1}{n}$ which diverges; that is, the partial sums $S_N = \sum\limits_{n=1}^{N} \dfrac{1}{n}$ grow (very, very slowly) without bound. One might think that by removing the "large" terms of the sequence that perhaps the series will converge. This is simply not the case. For instance, the sum of the first 10 million terms of the Harmonic Series is about 16.7. Removing the first 10 million terms from the Harmonic Series changes the partial sums, effectively subtracting 16.7 from the sum. However, a sequence that is growing without bound will still grow without bound when 16.7 is subtracted from it.

The equation below illustrates this. Even though we have subtracted off the first 10 million terms, this only subtracts a constant off of an expression that is

---

Notes:

still growing to infinity. Therefore, the modified series is still growing to infinity.

$$\sum_{n=10,000,001}^{\infty} \frac{1}{n} = \lim_{N\to\infty} \sum_{n=10,000,001}^{N} \frac{1}{n} = \lim_{N\to\infty} \sum_{n=1}^{N} \frac{1}{n} - \sum_{n=1}^{10,000,001} \frac{1}{n}$$

$$= \lim_{N\to\infty} \sum_{n=1}^{N} \frac{1}{n} - 16.7 = \infty.$$

This section introduced us to series and defined a few special types of series whose convergence properties are well known. We know when a geometric series converges or diverges. Most series that we encounter are not one of these types, but we are still interested in knowing whether or not they converge. The next three sections introduce tests that help us determine whether or not a given series converges.

Notes:

Exercises 9.2

Terms and Concepts

1. Use your own words to describe how sequences and series are related.
2. Use your own words to define a *partial sum*.
3. Given a series $\sum\limits_{n=1}^{\infty} a_n$, describe the two sequences related to the series that are important.
4. Use your own words to explain what a geometric series is.
5. T/F: If $\{a_n\}$ is convergent, then $\sum\limits_{n=1}^{\infty} a_n$ is also convergent.

Problems

In Exercises 6–13, a series $\sum\limits_{n=1}^{\infty} a_n$ is given.

   (a) Give the first 5 partial sums of the series.

   (b) Give a graph of the first 5 terms of a_n and S_n on the same axes.

6. $\sum\limits_{n=1}^{\infty} \dfrac{(-1)^n}{n}$

7. $\sum\limits_{n=1}^{\infty} \dfrac{1}{n^2}$

8. $\sum\limits_{n=1}^{\infty} \cos(\pi n)$

9. $\sum\limits_{n=1}^{\infty} n$

10. $\sum\limits_{n=1}^{\infty} \dfrac{1}{n!}$

11. $\sum\limits_{n=1}^{\infty} \dfrac{1}{3^n}$

12. $\sum\limits_{n=1}^{\infty} \left(-\dfrac{9}{10}\right)^n$

13. $\sum\limits_{n=1}^{\infty} \left(\dfrac{1}{10}\right)^n$

In Exercises 14–29, state whether the given series converges or diverges and provide justification for your conclusion.

14. $\sum\limits_{n=0}^{\infty} \dfrac{1}{5^n}$

15. $\sum\limits_{n=1}^{\infty} \dfrac{3n^2}{n(n+2)}$

16. $\sum\limits_{n=0}^{\infty} \dfrac{6^n}{5^n}$

17. $\sum\limits_{n=1}^{\infty} \dfrac{2^n}{n^2}$

18. $\sum\limits_{n=1}^{\infty} \sqrt{n}$

19. $\sum\limits_{n=1}^{\infty} \dfrac{n!}{10^n}$

20. $\sum\limits_{n=1}^{\infty} \dfrac{5^n - n^5}{5^n + n^5}$

21. $\sum\limits_{n=1}^{\infty} \left(\dfrac{1}{n!} + \dfrac{1}{n}\right)$

22. $\sum\limits_{n=1}^{\infty} \dfrac{1}{2n}$

23. $\sum\limits_{n=1}^{\infty} \dfrac{2^n + 1}{2^{n+1}}$

24. $\sum\limits_{n=1}^{\infty} \dfrac{1}{2n - 1}$

25. $\sum\limits_{n=1}^{\infty} \sqrt[n]{3}$

26. $\sum\limits_{n=1}^{\infty} \left(1 + \dfrac{1}{n}\right)^n$

27. $\sum\limits_{n=1}^{\infty} \dfrac{\pi^n}{3^{n+1}}$

28. $\sum\limits_{n=1}^{\infty} \dfrac{3^n + 2^n}{6^n}$

29. $\sum\limits_{n=1}^{\infty} \dfrac{3}{n(n+1)} + \dfrac{5}{4^n}$

In Exercises 30–45, a series is given.

   (a) Find a formula for S_n, the nth partial sum of the series.

   (b) Determine whether the series converges or diverges. If it converges, state what it converges to.

30. $\sum\limits_{n=0}^{\infty} \dfrac{1}{4^n}$

31. $1^3 + 2^3 + 3^3 + 4^3 + \cdots$

32. $\sum\limits_{n=1}^{\infty} (-1)^n n$

33. $\sum\limits_{n=0}^{\infty} \dfrac{5}{2^n}$

34. $\sum\limits_{n=1}^{\infty} e^{-n}$

35. $1 - \dfrac{1}{3} + \dfrac{1}{9} - \dfrac{1}{27} + \dfrac{1}{81} + \cdots$

36. $\sum\limits_{n=1}^{\infty} \dfrac{1}{n(n+1)}$

37. $\sum\limits_{n=1}^{\infty} \dfrac{3}{n(n+2)}$

38. $\displaystyle\sum_{n=1}^{\infty} \frac{1}{(2n-1)(2n+1)}$

39. $\displaystyle\sum_{n=1}^{\infty} \ln\left(\frac{n}{n+1}\right)$

40. $\displaystyle\sum_{n=1}^{\infty} \frac{2n+1}{n^2(n+1)^2}$

41. $\dfrac{1}{1\cdot 4} + \dfrac{1}{2\cdot 5} + \dfrac{1}{3\cdot 6} + \dfrac{1}{4\cdot 7} + \cdots$

42. $2 + \left(\dfrac{1}{2} + \dfrac{1}{3}\right) + \left(\dfrac{1}{4} + \dfrac{1}{9}\right) + \left(\dfrac{1}{8} + \dfrac{1}{27}\right) + \cdots$

43. $\displaystyle\sum_{n=2}^{\infty} \frac{1}{n^2-1}$

44. $\displaystyle\sum_{n=0}^{\infty} \left(\sin 1\right)^n$

45. $\displaystyle\sum_{n=1}^{\infty} \left(\frac{2}{n(n+2)} + \frac{5}{4^n}\right)$

In Exercises 46–49, find the values of x for which the series converges.

46. $\displaystyle\sum_{n=1}^{\infty} \frac{x^n}{3^n}$

47. $\displaystyle\sum_{n=1}^{\infty} \frac{(x+3)^n}{2^n}$

48. $\displaystyle\sum_{n=1}^{\infty} \frac{4^n}{x^n}$

49. $\displaystyle\sum_{n=1}^{\infty} (x+2)^n$

50. Show the series $\displaystyle\sum_{n=1}^{\infty} \frac{n}{(2n-1)(2n+1)}$ diverges.

9.3 The Integral Test

Knowing whether or not a series converges is very important, especially when we discuss Power Series in Section 9.8. Theorem 64 gives criteria for when Geometric series converge and Theorem 67 gives a quick test to determine if a series diverges. There are many important series whose convergence cannot be determined by these theorems, though, so we introduce a set of tests that allow us to handle a broad range of series. We start with the Integral Test.

Integral Test

We stated in Section 9.1 that a sequence $\{a_n\}$ is a function $a(n)$ whose domain is $\mathbb{N}$, the set of natural numbers. If we can extend $a(n)$ to have the domain of $\mathbb{R}$, the real numbers, and it is both positive and decreasing on $[1, \infty)$, then the convergence of $\sum\limits_{n=1}^{\infty} a_n$ is the same as $\int\limits_{1}^{\infty} a(x)\, dx$.

Theorem 69   Integral Test

Let a sequence $\{a_n\}$ be defined by $a_n = a(n)$, where $a(n)$ is continuous, positive and decreasing on $[1, \infty)$. Then $\sum\limits_{n=1}^{\infty} a_n$ converges, if, and only if, $\int\limits_{1}^{\infty} a(x)\, dx$ converges. In other words:

1. If $\int\limits_{1}^{\infty} a(x)\, dx$ is convergent, then $\sum\limits_{n=1}^{\infty} a_n$ is convergent.

2. If $\int\limits_{1}^{\infty} a(x)\, dx$ is divergent, then $\sum\limits_{n=1}^{\infty} a_n$ is divergent.

Note: Theorem 69 does not state that the integral and the summation have the same value.

Note that it is not necessary to start the series or the integral at $n = 1$. We may use any interval $[n, \infty]$ on which $a(n)$ is continuous, positive and decreasing. Also the sequence $\{a_n\}$ does not have to be strictly decreasing. It must be *ultimately decreasing* which means it is decreasing for all n larger than some number N.

We can demonstrate the truth of the Integral Test with two simple graphs. In Figure 9.15(a), the height of each rectangle is $a(n) = a_n$ for $n = 1, 2, \ldots,$

Notes:

and clearly the rectangles enclose more area than the area under $y = a(x)$. Therefore we can conclude that

$$\int_1^\infty a(x)\, dx < \sum_{n=1}^\infty a_n. \qquad (9.1)$$

In Figure 9.15(b), we draw rectangles under $y = a(x)$ with the Right-Hand rule, starting with $n = 2$. This time, the area of the rectangles is less than the area under $y = a(x)$, so $\sum_{n=2}^\infty a_n < \int_1^\infty a(x)\, dx$. Note how this summation starts with $n = 2$; adding a_1 to both sides lets us rewrite the summation starting with $n = 1$:

$$\sum_{n=1}^\infty a_n < a_1 + \int_1^\infty a(x)\, dx. \qquad (9.2)$$

Combining Equations (9.1) and (9.2), we have

$$\sum_{n=1}^\infty a_n < a_1 + \int_1^\infty a(x)\, dx < a_1 + \sum_{n=1}^\infty a_n. \qquad (9.3)$$

From Equation (9.3) we can make the following two statements:

1. If $\displaystyle\sum_{n=1}^\infty a_n$ diverges, so does $\displaystyle\int_1^\infty a(x)\, dx$ (because $\displaystyle\sum_{n=1}^\infty a_n < a_1 + \int_1^\infty a(x)\, dx$)

2. If $\displaystyle\sum_{n=1}^\infty a_n$ converges, so does $\displaystyle\int_1^\infty a(x)\, dx$ (because $\displaystyle\int_1^\infty a(x)\, dx < \sum_{n=1}^\infty a_n.$)

Therefore the series and integral either both converge or both diverge. Theorem 68 allows us to extend this theorem to series where $a(n)$ is positive and decreasing on $[b, \infty)$ for some $b > 1$.

Watch the video:
Integral Test for Series: Why It Works at
https://youtu.be/ObiRjUFHJHo

(a)

(b)

Figure 9.15: Illustrating the truth of the Integral Test.

Example 1   Using the Integral Test

Determine the convergence of $\displaystyle\sum_{n=1}^\infty \frac{\ln n}{n^2}$. (The terms of the sequence $\{a_n\} = \{\ln n/n^2\}$ and the n$^{\text{th}}$ partial sums are given in Figure 9.16.)

Notes:
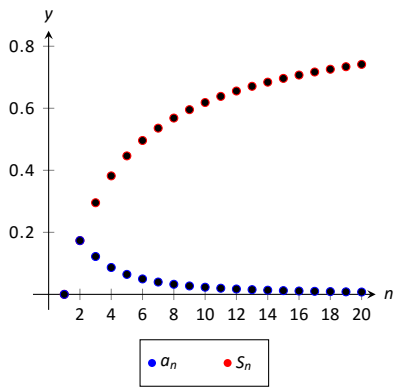
Figure 9.16: Plotting the sequence and series in Example 1.

SOLUTION   Figure 9.16 implies that $a(n) = (\ln n)/n^2$ is positive and decreasing on $[2, \infty)$. We can determine this analytically, too. We know $a(n)$ is positive as both $\ln n$ and n^2 are positive on $[2, \infty)$. To determine that $a(n)$ is decreasing, consider $a'(n) = (1 - 2\ln n)/n^3$, which is negative for $n \geq 2$. Since $a'(n)$ is negative, $a(n)$ is decreasing.

Applying the Integral Test, we test the convergence of $\displaystyle\int_1^\infty \frac{\ln x}{x^2}\,dx$. Integrating this improper integral requires the use of Integration by Parts, with $u = \ln x$ and $dv = 1/x^2\,dx$.

$$
\int_1^\infty \frac{\ln x}{x^2}\,dx = \lim_{t\to\infty} \int_1^t \frac{\ln x}{x^2}\,dx
$$

$$
= \lim_{t\to\infty}\left(-\frac{1}{x}\ln x\Big|_1^t + \int_1^t \frac{1}{x^2}\,dx \right)
$$

$$
= \lim_{t\to\infty}\left(-\frac{1}{x}\ln x - \frac{1}{x}\Big|_1^t \right)
$$

$$
= \lim_{t\to\infty}\left(1 - \frac{1}{t} - \frac{\ln t}{t} \right). \quad \text{Apply L'Hôpital's Rule:}
$$

$$
= 1 - 0 - \lim_{t\to\infty}\frac{1}{t}
$$

$$
= 1
$$

Since $\displaystyle\int_1^\infty \frac{\ln x}{x^2}\,dx$ converges, so does $\displaystyle\sum_{n=1}^\infty \frac{\ln n}{n^2}$.

Notes:

p–Series

Another important type of series is the *p-series*.

Definition 37   p–Series, General p–Series

1. A *p*–**series** is a series of the form

$$\sum_{n=1}^{\infty} \frac{1}{n^p}.$$

2. A **general** *p*–**series** is a series of the form

$$\sum_{n=1}^{\infty} \frac{1}{(an+b)^p},$$

where a and b are real numbers, and $an+b \neq 0$ for all n.

Like geometric series, one of the nice things about p–series is that they have easy to determine convergence properties.

Theorem 70   Convergence of General p–Series

Assume a and b are real numbers and $an+b \neq 0$ for all n.

A general *p*–series $\sum_{n=1}^{\infty} \frac{1}{(an+b)^p}$ will converge if and only if, $p > 1$.

Proof

Consider the integral $\int_1^{\infty} \frac{1}{(ax+b)^p}\, dx$; assuming $p \neq 1$,

$$\int_1^{\infty} \frac{1}{(ax+b)^p}\, dx = \lim_{t\to\infty} \int_1^t \frac{1}{(ax+b)^p}\, dx$$

$$= \lim_{t\to\infty} \frac{1}{a(1-p)} (ax+b)^{1-p}\Big|_1^t$$

$$= \lim_{t\to\infty} \frac{1}{a(1-p)} \left((at+b)^{1-p} - (a+b)^{1-p} \right).$$

Notes:

This limit converges if and only if, $p > 1$. It is easy to show that the integral also diverges in the case of $p = 1$. (This result is similar to the work preceding Key Idea 31.)

Therefore $\sum\limits_{n=1}^{\infty} \dfrac{1}{(an + b)^p}$ converges if, and only if, $p > 1$.   □

Example 2   **Determining convergence of series**
Determine the convergence of the following series.

1. $\sum\limits_{n=1}^{\infty} \dfrac{1}{n}$

2. $\sum\limits_{n=1}^{\infty} \dfrac{1}{n^2}$

3. $\sum\limits_{n=1}^{\infty} \dfrac{1}{\sqrt{n}}$

4. $\sum\limits_{n=1}^{\infty} \dfrac{(-1)^n}{n}$

5. $\sum\limits_{n=11}^{\infty} \dfrac{1}{\left(\frac{1}{2}n - 5\right)^3}$

6. $\sum\limits_{n=1}^{\infty} \dfrac{1}{2^n}$

SOLUTION

1. This is a p–series with $p = 1$. By Theorem 70, this series diverges.

    This series is a famous series, called the *Harmonic Series*, so named because of its relationship to *harmonics* in the study of music and sound.

2. This is a p–series with $p = 2$. By Theorem 70, it converges. Note that the theorem does not give a formula by which we can determine *what* the series converges to; we just know it converges. A famous, unexpected result is that this series converges to $\pi^2/6$.

3. This is a p–series with $p = 1/2$; the theorem states that it diverges.

4. This is not a p–series; the definition does not allow for alternating signs. Therefore we cannot apply Theorem 70. We will consider this series again in Section 9.5. (Another famous result states that this series, the *Alternating Harmonic Series*, converges to ln 2.)

5. This is a general p–series with $p = 3$, therefore it converges.

6. This is not a p–series, but a geometric series with $r = 1/2$. It converges.

   In the next section we consider two more convergence tests, both comparison tests. That is, we determine the convergence of one series by comparing it to another series with known convergence.

Notes:

Exercises 9.3

Terms and Concepts

1. In order to apply the Integral Test to a sequence $\{a_n\}$, the function $a(n) = a_n$ must be _____, _____ and _____.

2. T/F: The Integral Test can be used to determine the sum of a convergent series.

Problems

In Exercises 3–10, use the Integral Test to determine the convergence of the given series.

3. $\displaystyle\sum_{n=1}^{\infty} \frac{1}{2^n}$

4. $\displaystyle\sum_{n=1}^{\infty} \frac{1}{n^4}$

5. $\displaystyle\sum_{n=1}^{\infty} \frac{n}{n^2+1}$

6. $\displaystyle\sum_{n=2}^{\infty} \frac{1}{n\ln n}$

7. $\displaystyle\sum_{n=1}^{\infty} \frac{1}{n^2+1}$

8. $\displaystyle\sum_{n=2}^{\infty} \frac{1}{n(\ln n)^2}$

9. $\displaystyle\sum_{n=1}^{\infty} \frac{n}{2^n}$

10. $\displaystyle\sum_{n=1}^{\infty} \frac{\ln n}{n^3}$

In Exercises 11–14, find the value(s) of p for which the series is convergent.

11. $\displaystyle\sum_{i=2}^{\infty} \frac{1}{n(\ln n)^p}$

12. $\displaystyle\sum_{i=1}^{\infty} n(1+n^2)^p$

13. $\displaystyle\sum_{i=1}^{\infty} \frac{\ln n}{n^p}$

14. $\displaystyle\sum_{i=3}^{\infty} \frac{1}{n\ln n[\ln(\ln n)]^p}$

9.4   Comparison Tests

In this section we will be comparing a given series with series that we know either converge or diverge.

Theorem 71   Direct Comparison Test
Let $\{a_n\}$ and $\{b_n\}$ be positive sequences where $a_n \leq b_n$ for all $n \geq N$, for some $N \geq 1$.

1. If $\displaystyle\sum_{n=1}^{\infty} b_n$ converges, then $\displaystyle\sum_{n=1}^{\infty} a_n$ converges.

2. If $\displaystyle\sum_{n=1}^{\infty} a_n$ diverges, then $\displaystyle\sum_{n=1}^{\infty} b_n$ diverges.

Note: A sequence $\{a_n\}$ is a **positive sequence** if $a_n > 0$ for all n.

Because of Theorem 68, any theorem that relies on a positive sequence still holds true when $a_n > 0$ for all but a finite number of values of n.

Proof
First consider the partial sums of each series.

$$S_n = \sum_{i=1}^{n} a_i \quad \text{and} \quad T_n = \sum_{i=1}^{n} b_i$$

Since both series have positive terms we know that

$$S_n \leq S_n + a_{n+1} = \sum_{i=1}^{n} a_i + a_{n+1} = \sum_{i=1}^{n+1} a_i = S_{n+1}$$

and

$$T_n \leq T_n + b_{n+1} = \sum_{i=1}^{n} b_i + b_{n+1} = \sum_{i=1}^{n+1} b_i = T_{n+1}$$

Therefore, both of the sequences of partial sums, $\{S_n\}$ and $\{T_n\}$, are increasing. We also know that because $a_n \leq b_n$ for all $n \geq N$ that we must have $S_n \leq T_n$ for all $n \geq N$.

For the first part, assume that $\displaystyle\sum_{n=1}^{\infty} b_n$ converges. Since $b_n \geq 0$ we know that

$$T_n = \sum_{i=1}^{n} b_i \leq \sum_{i=1}^{\infty} b_i$$

Notes:

From above we know that $S_n \leq T_n$ for all $n \geq N$ so we also have

$$S_n \leq \sum_{i=1}^{\infty} b_i$$

Because $\sum_{i=1}^{\infty} b_i$ converges it must have a finite value and $\{S_n\}$ is bounded above. We also showed that $\{S_n\}$ is increasing so by Theorem 63 we know $\{S_n\}$ converges and so $\sum_{n=1}^{\infty} a_n$ converges.

For the second part, assume that $\sum_{n=1}^{\infty} a_n$ diverges. Because $a_n \geq 0$ we must have $\lim_{n \to \infty} S_n = \infty$. We also know that for all n, $S_n \leq T_n$ and so we also know that $\lim_{n \to \infty} T_n = \infty$. Therefore, T_n is a divergent sequence and so $\sum_{i=1}^{\infty} b_n$ diverges. $\square$

Watch the video:
Direct Comparison Test / Limit Comparison Test for Series — Basic Info at
`https://youtu.be/LAHKu3B-1zE`

Example 1   **Applying the Direct Comparison Test**

Determine the convergence of $\sum_{n=1}^{\infty} \dfrac{1}{3^n + n^2}$.

SOLUTION   This series is neither a geometric or p-series, but seems related. We predict it will converge, so we look for a series with larger terms that converges. (Note too that the Integral Test seems difficult to apply here.)

Since $3^n < 3^n + n^2$, $\dfrac{1}{3^n} > \dfrac{1}{3^n + n^2}$ for all $n \geq 1$. The series $\sum_{n=1}^{\infty} \dfrac{1}{3^n}$ is a convergent geometric series; by Theorem 71, $\sum_{n=1}^{\infty} \dfrac{1}{3^n + n^2}$ converges.

Notes:

Example 2   **Applying the Direct Comparison Test**

Determine the convergence of $\displaystyle\sum_{n=1}^{\infty} \frac{n^3}{n^4 - 1}$.

SOLUTION   We know the Harmonic Series $\displaystyle\sum_{n=1}^{\infty} \frac{1}{n}$ diverges, and it seems that the given series is closely related to it, hence we predict it will diverge.

We have $\dfrac{n^3}{n^4 - 1} > \dfrac{n^3}{n^4} = \dfrac{1}{n}$ for all $n \geq 1$.

The Harmonic Series, $\displaystyle\sum_{n=1}^{\infty} \frac{1}{n}$, diverges, so we conclude that $\displaystyle\sum_{n=1}^{\infty} \frac{n^3}{n^4 - 1}$ diverges as well.

The concept of direct comparison is powerful and often relatively easy to apply. Practice helps one develop the necessary intuition to quickly pick a proper series with which to compare. However, it is easy to construct a series for which it is difficult to apply the Direct Comparison Test.

Consider $\displaystyle\sum_{n=1}^{\infty} \frac{n^3}{n^4 + 1}$. It is very similar to the divergent series given in Example 2. We suspect that it also diverges, as $\dfrac{1}{n} \approx \dfrac{n^3}{n^4 + 1}$ for large n. However, the inequality that we naturally want to use "goes the wrong way": since $\dfrac{n^3}{n^4 + 1} < \dfrac{n^3}{n^4} = \dfrac{1}{n}$ for all $n \geq 1$. The given series has terms *less than* the terms of a divergent series, and we cannot conclude anything from this.

Fortunately, we can apply another test to the given series to determine its convergence.

Limit Comparison Test

Notes:

Theorem 72   Limit Comparison Test

Let $\{a_n\}$ and $\{b_n\}$ be positive sequences.

1. If $\displaystyle\lim_{n\to\infty} \frac{a_n}{b_n} = L$, where L is a positive real number, then $\displaystyle\sum_{n=1}^{\infty} a_n$ and

    $\displaystyle\sum_{n=1}^{\infty} b_n$ either both converge or both diverge.

2. If $\displaystyle\lim_{n\to\infty} \frac{a_n}{b_n} = 0$, then if $\displaystyle\sum_{n=1}^{\infty} b_n$ converges, then so does $\displaystyle\sum_{n=1}^{\infty} a_n$.

3. If $\displaystyle\lim_{n\to\infty} \frac{a_n}{b_n} = \infty$, then if $\displaystyle\sum_{n=1}^{\infty} b_n$ diverges, then so does $\displaystyle\sum_{n=1}^{\infty} a_n$.

Proof

1. We have $0 < L < \infty$ so we can find two positive numbers, m and M such
    that $m < L < M$. Because $L = \displaystyle\lim_{n\to\infty} \frac{a_n}{b_n}$ we know that for large enough n
    the quotient $\frac{a_n}{b_n}$ must be close to L. So there must be a positive integer N
    such that if $n > N$ we also have $m < \dfrac{a_n}{b_n} < M$. Multiply by b_n and we have

    $mb_n < a_n < Mb_n$ for $n > N$. If $\displaystyle\sum_{n=1}^{\infty} b_n$ diverges, then so does $\displaystyle\sum_{n=1}^{\infty} mb_n$.

    Also since $mb_n < a_n$ for sufficiently large n, by the Comparison Test $\displaystyle\sum_{n=1}^{\infty} a_n$

    also diverges.

    Similarly, if $\displaystyle\sum_{n=1}^{\infty} b_n$ converges, then so does $\displaystyle\sum_{n=1}^{\infty} Mb_n$. Since $a_n < Mb_n$ for

    sufficiently large n, by the Comparison Test $\displaystyle\sum_{n=1}^{\infty} a_n$ also converges.

2. Since $\displaystyle\lim_{n\to\infty} a_n = 0$, there is a number $N > 0$ such that

    $$\left| \frac{a_n}{b_n} - 0 \right| < 1 \text{ for all } n > N$$

    $$a_n < b_n \text{ since } a_n \text{ and } b_n \text{ are positive}$$

Notes:

Now since $\displaystyle\sum_{n=1}^{\infty} b_n$ converges, $\displaystyle\sum_{n=1}^{\infty} a_n$ converges by the Comparison Test.

3. Since $\displaystyle\lim_{n\to\infty} a_n = \infty$, there is a number $N > 0$ such that

$$\frac{a_n}{b_n} > 1 \text{ for all } n > N$$

$$a_n > b_n \text{ for all } n > N$$

Now since $\displaystyle\sum_{n=1}^{\infty} b_n$ diverges, $\displaystyle\sum_{n=1}^{\infty} a_n$ diverges by the Comparison Test. $\square$

Theorem 72 is most useful when the convergence of the series from $\{b_n\}$ is known and we are trying to determine the convergence of the series from $\{a_n\}$.

We use the Limit Comparison Test in the next example to examine the series $\displaystyle\sum_{n=1}^{\infty} \frac{n^3}{n^4 + 1}$ which motivated this new test.

Example 3   **Applying the Limit Comparison Test**

Determine the convergence of $\displaystyle\sum_{n=1}^{\infty} \frac{n^3}{n^4 + 1}$ using the Limit Comparison Test.

SOLUTION   We compare the terms of $\displaystyle\sum_{n=1}^{\infty} \frac{n^3}{n^4 + 1}$ to the terms of the

Harmonic Sequence $\displaystyle\sum_{n=1}^{\infty} \frac{1}{n}$:

$$\lim_{n\to\infty} \frac{n^3/(n^4 + 1)}{1/n} = \lim_{n\to\infty} \frac{n^4}{n^4 + 1} = \lim_{n\to\infty} \frac{1}{1 + 1/n^4}$$

$$= 1.$$

Since the Harmonic Series diverges, we conclude that $\displaystyle\sum_{n=1}^{\infty} \frac{n^3}{n^4 + 1}$ diverges as well.

Example 4   **Applying the Limit Comparison Test**

Determine the convergence of $\displaystyle\sum_{n=1}^{\infty} \frac{1}{3^n - n^2}$

Notes:

SOLUTION   This series is similar to the one in Example 1, but now we are considering "$3^n - n^2$" instead of "$3^n + n^2$." This difference makes applying the Direct Comparison Test difficult.

Instead, we use the Limit Comparison Test and compare with the series $\displaystyle\sum_{n=1}^{\infty} \frac{1}{3^n}$:

$$\lim_{n\to\infty} \frac{1/(3^n - n^2)}{1/3^n} = \lim_{n\to\infty} \frac{3^n}{3^n - n^2} \overset{\text{L'H}}{=} \lim_{n\to\infty} \frac{\ln 3 \cdot 3^n}{\ln 3 \cdot 3^n - 2n}$$

$$\overset{\text{L'H}}{=} \lim_{n\to\infty} \frac{(\ln 3)^2 3^n}{(\ln 3)^2 3^n - 2}$$

$$\overset{\text{L'H}}{=} \lim_{n\to\infty} \frac{(\ln 3)^3 3^n}{(\ln 3)^3 3^n} = 1.$$

We know $\displaystyle\sum_{n=1}^{\infty} \frac{1}{3^n}$ is a convergent geometric series, hence $\displaystyle\sum_{n=1}^{\infty} \frac{1}{3^n - n^2}$ converges as well.

As mentioned before, practice helps one develop the intuition to quickly choose a series with which to compare. A general rule of thumb is to pick a series based on the dominant term in the expression of $\{a_n\}$. It is also helpful to note that factorials dominate exponentials, which dominate algebraic functions (e.g., polynomials), which dominate logarithms. In the previous example, the dominant term of $\dfrac{1}{3^n - n^2}$ was 3^n, so we compared the series to $\displaystyle\sum_{n=1}^{\infty} \frac{1}{3^n}$. It is hard to apply the Limit Comparison Test to series containing factorials, though, as we have not learned how to apply L'Hôpital's Rule to $n!$.

Example 5   **Applying the Limit Comparison Test**

Determine the convergence of $\displaystyle\sum_{n=1}^{\infty} \frac{\sqrt{n} + 3}{n^2 - n + 1}$.

SOLUTION   We naïvely attempt to apply the rule of thumb given above and note that the dominant term in the expression of the series is $1/n^2$. Knowing that $\displaystyle\sum_{n=1}^{\infty} \frac{1}{n^2}$ converges, we attempt to apply the Limit Comparison Test:

$$\lim_{n\to\infty} \frac{(\sqrt{n} + 3)/(n^2 - n + 1)}{1/n^2} = \lim_{n\to\infty} \frac{n^2(\sqrt{n} + 3)}{n^2 - n + 1}$$

$$= \infty \quad \text{(Apply L'Hôpital's Rule).}$$

Notes:

Theorem 72 part (3) only applies when $\sum\limits_{n=1}^{\infty} b_n$ diverges; in our case, it converges. Ultimately, our test has not revealed anything about the convergence of our series.

The problem is that we chose a poor series with which to compare. Since the numerator and denominator of the terms of the series are both algebraic functions, we should have compared our series to the dominant term of the numerator divided by the dominant term of the denominator.

The dominant term of the numerator is $n^{1/2}$ and the dominant term of the denominator is n^2. Thus we should compare the terms of the given series to $n^{1/2}/n^2 = 1/n^{3/2}$:

$$\lim_{n\to\infty} \frac{(\sqrt{n}+3)/(n^2-n+1)}{1/n^{3/2}} = \lim_{n\to\infty} \frac{n^{3/2}(\sqrt{n}+3)}{n^2-n+1}$$

$$= 1 \quad \text{(Apply L'Hôpital's Rule)}.$$

Since the p-series $\sum\limits_{n=1}^{\infty} \dfrac{1}{n^{3/2}}$ converges, we conclude that $\sum\limits_{n=1}^{\infty} \dfrac{\sqrt{n}+3}{n^2-n+1}$ converges as well.

The tests we have encountered so far has required that we analyze series from *positive* sequences (the absolute value of the ratio and the root tests converts the sequence into a positive sequence). The next section relaxes this restriction by considering *alternating series*, where the underlying sequence has terms that alternate between being positive and negative.

Notes:

Exercises 9.4

Terms and Concepts

1. Suppose $\sum_{n=0}^{\infty} a_n$ is convergent, and there are sequences $\{b_n\}$ and $\{c_n\}$ such that $b_n \leq a_n \leq c_n$ for all n. What can be said about the series $\sum_{n=0}^{\infty} b_n$ and $\sum_{n=0}^{\infty} c_n$?

Problems

In Exercises 2–6, use the Direct Comparison Test to determine the convergence of the given series; state what series is used for comparison.

2. $\sum_{n=1}^{\infty} \dfrac{1}{n^2 + 3n - 5}$

3. $\sum_{n=1}^{\infty} \dfrac{1}{4^n + n^2 - n}$

4. $\sum_{n=1}^{\infty} \dfrac{\ln n}{n}$

5. $\sum_{n=1}^{\infty} \dfrac{1}{n! + n}$

6. $\sum_{n=2}^{\infty} \dfrac{1}{\sqrt{n^2 - 1}}$

In Exercises 7–11, use the Limit Comparison Test to determine the convergence of the given series; state what series is used for comparison.

7. $\sum_{n=1}^{\infty} \dfrac{1}{n^2 - 3n + 5}$

8. $\sum_{n=1}^{\infty} \dfrac{1}{4^n - n^2}$

9. $\sum_{n=4}^{\infty} \dfrac{\ln n}{n - 3}$

10. $\sum_{n=1}^{\infty} \dfrac{1}{\sqrt{n^2 + n}}$

11. $\sum_{n=1}^{\infty} \dfrac{1}{n + \sqrt{n}}$

In Exercises 12–24, use the Direct Comparison Test or the Limit Comparison Test to determine the convergence of the given series. State which series is used for comparison.

12. $\sum_{n=5}^{\infty} \dfrac{1}{\sqrt{n} - 2}$

13. $\sum_{n=1}^{\infty} \dfrac{n^2 + n + 1}{n^3 - 5}$

14. $\sum_{n=1}^{\infty} \dfrac{n - 10}{n^2 + 10n + 10}$

15. $\sum_{n=1}^{\infty} \sin\left(1/n\right)$

16. $\sum_{n=1}^{\infty} \dfrac{2^n}{5^n + 10}$

17. $\sum_{n=1}^{\infty} \dfrac{n + 5}{n^3 - 5}$

18. $\sum_{n=1}^{\infty} \dfrac{n}{n^4 + 1}$

19. $\sum_{n=1}^{\infty} \dfrac{n - 1}{n4^n}$

20. $\sum_{n=2}^{\infty} \dfrac{n}{n^2 - 1}$

21. $\sum_{n=1}^{\infty} \dfrac{1}{\sqrt{n} + 100}$

22. $\sum_{n=2}^{\infty} \dfrac{1}{n^2 \ln n}$

23. $\sum_{n=1}^{\infty} \dfrac{\sqrt{n} + 3}{n^2 + 17}$

24. $\sum_{n=1}^{\infty} \dfrac{1 + \sin n}{10^n}$

25. Given that $\sum_{n=1}^{\infty} a_n$ converges, state which of the following series converges, may converge, or does not converge.

   (a) $\sum_{n=1}^{\infty} \dfrac{a_n}{n}$

   (b) $\sum_{n=1}^{\infty} a_n a_{n+1}$

   (c) $\sum_{n=1}^{\infty} (a_n)^2$

   (d) $\sum_{n=1}^{\infty} n a_n$

   (e) $\sum_{n=1}^{\infty} \dfrac{1}{a_n}$

In Exercises 26–33, determine the convergence of the given series. State the test used; more than one test may be appropriate.

26. $\sum_{n=1}^{\infty} \dfrac{n^2}{2^n}$

27. $\sum_{n=1}^{\infty} \dfrac{1}{(2n + 5)^3}$

28. $\sum_{n=1}^{\infty} \dfrac{n!}{10^n}$

29. $\displaystyle\sum_{n=1}^{\infty} \frac{\ln n}{n!}$

30. $\displaystyle\sum_{n=1}^{\infty} \frac{1}{3^n + n}$

31. $\displaystyle\sum_{n=1}^{\infty} \frac{n-2}{10n + 5}$

32. $\displaystyle\sum_{n=1}^{\infty} \frac{3^n}{n^3}$

33. $\displaystyle\sum_{n=1}^{\infty} \frac{\cos(1/n)}{\sqrt{n}}$

9.5 Alternating Series and Absolute Convergence

The series convergence tests we have used require that the underlying sequence $\{a_n\}$ be a positive sequence. (We can relax this with Theorem 68 and state that there must be an $N > 0$ such that $a_n > 0$ for all $n > N$; that is, $\{a_n\}$ is positive for all but a finite number of values of n.)

In this section we explore series whose summation includes negative terms. We start with a very specific form of series, where the terms of the summation alternate between being positive and negative.

Definition 38   Alternating Series

Let $\{b_n\}$ be a positive sequence. An **alternating series** is a series of either the form

$$\sum_{n=1}^{\infty}(-1)^n b_n \qquad \text{or} \qquad \sum_{n=1}^{\infty}(-1)^{n+1} b_n.$$

We want to think that an alternating sequence $\{a_n\}$ is related to a positive sequence $\{b_n\}$ by $a_n = (-1)^n b_n$.

Recall the terms of Harmonic Series come from the Harmonic Sequence $\{b_n\} = \{1/n\}$. An important alternating series is the **Alternating Harmonic Series**:

$$\sum_{n=1}^{\infty}(-1)^{n+1}\frac{1}{n} = 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \frac{1}{5} - \frac{1}{6} + \cdots$$

Geometric Series can also be alternating series when $r < 0$. For instance, if $r = -1/2$, the geometric series is

$$\sum_{n=0}^{\infty}\left(\frac{-1}{2}\right)^n = 1 - \frac{1}{2} + \frac{1}{4} - \frac{1}{8} + \frac{1}{16} - \frac{1}{32} + \cdots$$

Theorem 64 states that geometric series converge when $|r| < 1$ and gives the sum: $\displaystyle\sum_{n=0}^{\infty} r^n = \frac{1}{1 - r}$. When $r = -1/2$ as above, we find

$$\sum_{n=0}^{\infty}\left(\frac{-1}{2}\right)^n = \frac{1}{1 - (-1/2)} = \frac{1}{3/2} = \frac{2}{3}.$$

A powerful convergence theorem exists for other alternating series that meet a few conditions.

Notes:

> **Theorem 73   Alternating Series Test**
> Let $\{b_n\}$ be a positive, decreasing sequence where $\lim\limits_{n\to\infty} b_n = 0$. Then
>
> $$\sum_{n=1}^{\infty}(-1)^n b_n \qquad \text{and} \qquad \sum_{n=1}^{\infty}(-1)^{n+1} b_n$$
>
> converge.

The basic idea behind Theorem 73 is illustrated in Figure 9.17. A positive, decreasing sequence $\{b_n\}$ is shown along with the partial sums

$$S_n = \sum_{i=1}^{n}(-1)^{i+1} b_i = b_1 - b_2 + b_3 - b_4 + \cdots + (-1)^{n+1} b_n.$$

Because $\{b_n\}$ is decreasing, the amount by which S_n bounces up and down decreases. Moreover, the odd terms of S_n form a decreasing, bounded sequence, while the even terms of S_n form an increasing, bounded sequence. Since bounded, monotonic sequences converge (see Theorem 63) and the terms of $\{b_n\}$ approach 0, we will show below the odd and even terms of S_n converge to the same common limit L, the sum of the series.

Proof
Because $\{b_n\}$ is a decreasing sequence, we have $b_n - b_{n+1} \geq 0$. We will consider the even and odd partial sums separately. First consider the even partial sums.

$$S_2 = b_1 - b_2 \geq 0 \qquad\qquad\qquad\qquad\qquad \text{since } b_2 \leq b_1$$
$$S_4 = b_1 - b_2 + b_3 - b_4 = S_2 + b_3 - b_4 \geq S_2 \qquad \text{since } b_3 - b_4 \geq 0$$
$$S_6 = S_4 + b_5 - b_6 \geq S_4 \qquad\qquad\qquad\qquad \text{since } b_5 - b_6 \geq 0$$
$$\vdots$$
$$S_{2n} = S_{2n-2} + b_{2n-1} - b_{2n} \geq S_{2n-2} \qquad\qquad \text{since } b_{2n-1} - b_{2n} \geq 0$$

We now have

$$0 \leq S_2 \leq S_4 \leq S_6 \leq \cdots \leq S_{2n} \leq \cdots$$

so $\{S_{2n}\}$ is an increasing sequence. But we can also write

$$S_{2n} = b_1 - b_2 + b_3 - b_4 + b_5 - \cdots - b_{2n-2} + b_{2n-1} - b_{2n}$$
$$= b_1 - (b_2 - b_3) - (b_4 - b_5) - \cdots - (b_{2n-2} - b_{2n-1}) - b_{2n}$$

Figure 9.17: Illustrating convergence with the Alternating Series Test.

Notes:

Each term in parentheses is positive and b_{2n} is positive so we have $S_{2n} \leq b_1$ for all n. We now have the sequence of even partial sums, $\{S_{2n}\}$, is increasing and bounded above so by Theorem 63 $\{S_{2n}\}$ converges. Since we know it converges, we will assume it's limit is L or

$$\lim_{n\to\infty} S_{2n} = L$$

Next we determine the limit of the sequence of odd partial sums.

$$\lim_{n\to\infty} S_{2n+1} = \lim_{n\to\infty} (S_{2n} + b_{2n+1})$$
$$= \lim_{n\to\infty} S_{2n} + \lim_{n\to\infty} b_{2n+1}$$
$$= L + 0$$
$$= L$$

Both the even and odd partial sums converge to s so we have $\lim_{n\to\infty} S_n = L$, which means the series is convergent. $\qquad\square$

Watch the video:
Alternating Series — Another Example 4 at
`https://youtu.be/aOiZvfFAMW8`

Example 1   **Applying the Alternating Series Test**
Determine if the Alternating Series Test applies to each of the following series.

$$1. \; \sum_{n=1}^{\infty}(-1)^{n+1}\frac{1}{n} \qquad 2. \; \sum_{n=2}^{\infty}(-1)^{n}\frac{\ln n}{n} \qquad 3. \; \sum_{n=1}^{\infty}(-1)^{n+1}\frac{|\sin n|}{n^2}$$

SOLUTION

1. This is the Alternating Harmonic Series as seen previously. The underlying sequence is $\{b_n\} = \{1/n\}$, which is positive, decreasing, and approaches 0 as $n \to \infty$. Therefore we can apply the Alternating Series Test and conclude this series converges.

   While the test does not state what the series converges to, we will see later that $\sum_{n=1}^{\infty}(-1)^{n+1}\frac{1}{n} = \ln 2$.

Notes:

501

2. The underlying sequence is $\{b_n\} = \{\ln n / n\}$. This is positive for $n \geq 2$ and
$$\lim_{n \to \infty} \frac{\ln n}{n} = \lim_{n \to \infty} \frac{1}{n} = 0 \text{ (use L'Hôpital's Rule). However, the sequence}$$
is not decreasing for all n. It is straightforward to compute $b_1 \approx 0.347$, $b_2 \approx 0.366$, and $b_3 \approx 0.347$: the sequence is increasing for at least the first 2 terms.

We do not immediately conclude that we cannot apply the Alternating Series Test. Rather, consider the long–term behavior of $\{b_n\}$. Treating $b_n = b(n)$ as a continuous function of n defined on $[2, \infty)$, we can take its derivative:

$$b'(n) = \frac{1 - \ln n}{n^2}.$$

The derivative is negative for all $n \geq 3$ (actually, for all $n > e$), meaning $b(n) = b_n$ is decreasing on $[3, \infty)$. We can apply the Alternating Series Test to the series when we start with $n = 3$ and conclude that $\sum_{n=3}^{\infty} (-1)^n \frac{\ln n}{n}$ converges; adding the terms with $n = 2$ does not change the convergence (i.e., we apply Theorem 68).

The important lesson here is that as before, if a series fails to meet the criteria of the Alternating Series Test on only a finite number of terms, we can still apply the test.

3. The underlying sequence is $\{b_n\} = \{|\sin n| / n^2\}$. This sequence is positive and approaches 0 as $n \to \infty$. However, it is not a decreasing sequence; the value of $|\sin n|$ oscillates between 0 and 1 as $n \to \infty$. We cannot remove a finite number of terms to make $\{b_n\}$ decreasing, therefore we cannot apply the Alternating Series Test.

Keep in mind that this does not mean we conclude the series diverges; in fact, it does converge. We are just unable to conclude this based on Theorem 73.

One of the famous results of mathematics is that the Harmonic Series, $\sum_{n=1}^{\infty} \frac{1}{n}$ diverges, yet the Alternating Harmonic Series, $\sum_{n=1}^{\infty} (-1)^{n+1} \frac{1}{n}$, converges. The notion that alternating the signs of the terms in a series can make a series converge leads us to the following definitions.

Notes:

Definition 39   Absolute and Conditional Convergence

1. A series $\displaystyle\sum_{n=1}^{\infty} a_n$ **converges absolutely** if $\displaystyle\sum_{n=1}^{\infty} |a_n|$ converges.

2. A series $\displaystyle\sum_{n=1}^{\infty} a_n$ **converges conditionally** if $\displaystyle\sum_{n=1}^{\infty} a_n$ converges but $\displaystyle\sum_{n=1}^{\infty} |a_n|$ diverges.

Thus we say the Alternating Harmonic Series converges conditionally.

Note: In Definition 39, $\displaystyle\sum_{n=1}^{\infty} a_n$ is not necessarily an alternating series; it just may have some negative terms.

Example 2   Determining absolute and conditional convergence.
Determine if the following series converge absolutely, conditionally, or diverge.

1. $\displaystyle\sum_{n=1}^{\infty} (-1)^n \frac{n+3}{n^2+2n+5}$

2. $\displaystyle\sum_{n=1}^{\infty} (-1)^n \frac{n^2+2n+5}{2^n}$

3. $\displaystyle\sum_{n=3}^{\infty} (-1)^n \frac{3n-3}{5n-10}$

SOLUTION

1. We can show the series

$$\sum_{n=1}^{\infty} \left| (-1)^n \frac{n+3}{n^2+2n+5} \right| = \sum_{n=1}^{\infty} \frac{n+3}{n^2+2n+5}$$

diverges using the Limit Comparison Test, comparing with $1/n$.

The series $\displaystyle\sum_{n=1}^{\infty} (-1)^n \frac{n+3}{n^2+2n+5}$ converges using the Alternating Series Test; we conclude it converges conditionally.

2. We can show the series

$$\sum_{n=1}^{\infty} \left| (-1)^n \frac{n^2+2n+5}{2^n} \right| = \sum_{n=1}^{\infty} \frac{n^2+2n+5}{2^n}$$

converges using the Ratio Test.

Therefore we conclude $\displaystyle\sum_{n=1}^{\infty} (-1)^n \frac{n^2+2n+5}{2^n}$ converges absolutely.

Notes:

3. The series

$$\sum_{n=3}^{\infty} \left| (-1)^n \frac{3n-3}{5n-10} \right| = \sum_{n=3}^{\infty} \frac{3n-3}{5n-10}$$

diverges using the Test for Divergence, so it does not converge absolutely.

The series $\sum_{n=3}^{\infty} (-1)^n \frac{3n-3}{5n-10}$ fails the conditions of the Alternating Series Test as $(3n-3)/(5n-10)$ does not approach 0 as $n \to \infty$. We can state further that this series diverges; as $n \to \infty$, the series effectively adds and subtracts $3/5$ over and over. This causes the sequence of partial sums to oscillate and not converge.

Therefore the series $\sum_{n=1}^{\infty} (-1)^n \frac{3n-3}{5n-10}$ diverges.

Knowing that a series converges absolutely allows us to make two important statements, given in the following theorem. The first is that absolute convergence is "stronger" than regular convergence. That is, just because $\sum_{n=1}^{\infty} a_n$ converges, we cannot conclude that $\sum_{n=1}^{\infty} |a_n|$ will converge, but knowing a series converges absolutely tells us that $\sum_{n=1}^{\infty} a_n$ will converge.

Theorem 74   Absolute Convergence Theorem

Let $\sum_{n=1}^{\infty} a_n$ be a series that converges absolutely.

1. $\sum_{n=1}^{\infty} a_n$ converges.

2. Let $\{b_n\}$ be any rearrangement of the sequence $\{a_n\}$. Then

$$\sum_{n=1}^{\infty} b_n = \sum_{n=1}^{\infty} a_n.$$

One reason this is important is that our convergence tests all require that the underlying sequence of terms be positive. By taking the absolute value of the terms of a series where not all terms are positive, we are often able to apply an

Notes:

appropriate test and determine absolute convergence. This, in turn, determines that the series we are given also converges.

The second statement relates to **rearrangements** of series. When dealing with a finite set of numbers, the sum of the numbers does not depend on the order which they are added. (So $1+2+3 = 3+1+2$.) One may be surprised to find out that when dealing with an infinite set of numbers, the same statement does not always hold true: some infinite lists of numbers may be rearranged in different orders to achieve different sums. The theorem states that the terms of an absolutely convergent series can be rearranged in any way without affecting the sum.

In Example 2, we determined the series in part 2 converges absolutely. Theorem 74 tells us the series converges (which we could also determine using the Alternating Series Test).

The theorem states that rearranging the terms of an absolutely convergent series does not affect its sum. This implies that perhaps the sum of a conditionally convergent series can change based on the arrangement of terms. Indeed, it can. The Riemann Rearrangement Theorem (named after Bernhard Riemann) states that any conditionally convergent series can have its terms rearranged so that the sum is any desired value or infinity.

Before we consider an example, we state the following theorem that illustrates how the alternating structure of an alternating series is a powerful tool when approximating the sum of a convergent series.

Theorem 75   The Alternating Series Approximation Theorem

Let $\{b_n\}$ be a sequence that satisfies the hypotheses of the Alternating Series Test, and let S_n and L be the nth partial sums and sum, respectively, of either $\displaystyle\sum_{n=1}^{\infty}(-1)^n b_n$ or $\displaystyle\sum_{n=1}^{\infty}(-1)^{n+1} b_n$. Then

1. $|S_n - L| < b_{n+1}$, and

2. L is between S_n and S_{n+1}.

Part 1 of Theorem 75 states that the nth partial sum of a convergent alternating series will be within b_{n+1} of its total sum. Consider the alternating series we looked at before the statement of the theorem, $\displaystyle\sum_{n=1}^{\infty}\frac{(-1)^{n+1}}{n^2}$. Since $b_{14} = 1/14^2 \approx 0.0051$, we know that S_{13} is within 0.0051 of the total sum.

Moreover, Part 2 of the theorem states that since $S_{13} \approx 0.8252$ and $S_{14} \approx 0.8201$, we know the sum L lies between 0.8201 and 0.8252. One use of this is

---

Notes:

the knowledge that S_{14} is accurate to two places after the decimal.

Some alternating series converge slowly. In Example 1 we determined the series $\sum_{n=2}^{\infty}(-1)^{n+1}\dfrac{\ln n}{n}$ converged. With $n = 1001$, we find $\ln n/n \approx 0.0069$, meaning that $S_{1000} \approx 0.1633$ is accurate to one, maybe two, places after the decimal. Since $S_{1001} \approx 0.1564$, we know the sum L is $0.1564 \leq L \leq 0.1633$.

Example 3   **Approximating the sum of convergent alternating series**
Approximate the sum of the following series, accurate to within 0.001.

$$1. \ \sum_{n=1}^{\infty}(-1)^{n+1}\frac{1}{n^3} \qquad 2. \ \sum_{n=1}^{\infty}(-1)^{n+1}\frac{\ln n}{n}.$$

SOLUTION

1. Using Theorem 75, we want to find n where $1/n^3 < 0.001$:

$$\frac{1}{n^3} \leq 0.001 = \frac{1}{1000}$$
$$n^3 \geq 1000$$
$$n \geq \sqrt[3]{1000}$$
$$n \geq 10.$$

Let L be the sum of this series. By Part 1 of the theorem, $|S_9 - L| < b_{10} = 1/1000$. We can compute $S_9 = 0.902116$, which our theorem states is within 0.001 of the total sum.

We can use Part 2 of the theorem to obtain an even more accurate result. As we know the 10th term of the series is $-1/1000$, we can easily compute $S_{10} = 0.901116$. Part 2 of the theorem states that L is between S_9 and S_{10}, so $0.901116 < L < 0.902116$.

2. We want to find n where $\ln(n)/n < 0.001$. We start by solving $\ln(n)/n = 0.001$ for n. This cannot be solved algebraically, so we will use Newton's Method to approximate a solution.

Let $f(x) = \ln(x)/x - 0.001$; we want to know where $f(x) = 0$. We make a guess that x must be "large," so our initial guess will be $x_1 = 1000$. Recall how Newton's Method works: given an approximate solution x_n, our next approximation x_{n+1} is given by

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}.$$

Notes:

We find $f'(x) = \big(1 - \ln(x)\big)/x^2$. This gives

$$x_2 = 1000 - \frac{\ln(1000)/1000 - 0.001}{\big(1 - \ln(1000)\big)/1000^2}$$

$$= 2000.$$

Using a computer, we find that Newton's Method seems to converge to a solution $x = 9118.01$ after 8 iterations. Taking the next integer higher, we have $n = 9119$, where $\ln(9119)/9119 = 0.000999903 < 0.001$.

Again using a computer, we find $S_{9118} = -0.160369$. Part 1 of the theorem states that this is within 0.001 of the actual sum L. Already knowing the 9,119th term, we can compute $S_{9119} = -0.159369$, meaning $-0.159369 < L < -0.160369$.

Notice how the first series converged quite quickly, where we needed only 10 terms to reach the desired accuracy, whereas the second series took over 9,000 terms.

We now consider the Alternating Harmonic Series once more. We have stated that

$$\sum_{n=1}^{\infty}(-1)^{n+1}\frac{1}{n} = 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \frac{1}{5} - \frac{1}{6} + \frac{1}{7}\cdots = \ln 2,$$

(see Example 1).

Consider the rearrangement where every positive term is followed by two negative terms:

$$1 - \frac{1}{2} - \frac{1}{4} + \frac{1}{3} - \frac{1}{6} - \frac{1}{8} + \frac{1}{5} - \frac{1}{10} - \frac{1}{12}\cdots$$

(Convince yourself that these are exactly the same numbers as appear in the Alternating Harmonic Series, just in a different order.) Now group some terms and simplify:

$$\left(1 - \frac{1}{2}\right) - \frac{1}{4} + \left(\frac{1}{3} - \frac{1}{6}\right) - \frac{1}{8} + \left(\frac{1}{5} - \frac{1}{10}\right) - \frac{1}{12} + \cdots =$$

$$\frac{1}{2} - \frac{1}{4} + \frac{1}{6} - \frac{1}{8} + \frac{1}{10} - \frac{1}{12} + \cdots =$$

$$\frac{1}{2}\left(1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \frac{1}{5} - \frac{1}{6} + \cdots\right) = \frac{1}{2}\ln 2.$$

By rearranging the terms of the series, we have arrived at a different sum. (One could *try* to argue that the Alternating Harmonic Series does not actually

Notes:

converge to ln 2, because rearranging the terms of the series *shouldn't* change the sum. However, the Alternating Series Test proves this series converges to *L*, for some number *L*, and if the rearrangement does not change the sum, then $L = L/2$, implying $L = 0$. But the Alternating Series Approximation Theorem quickly shows that $L > 0$. The only conclusion is that the rearrangement *did* change the sum.) This is an incredible result.

We mentioned earlier that the Integral Test did not work well with series containing factorial terms. The next section introduces the Ratio Test, which does handle such series well. We also introduce the Root Test, which is good for series where each term is raised to a power.

Notes:

Exercises 9.5

Terms and Concepts

1. Why is $\displaystyle\sum_{n=1}^{\infty} \sin n$ not an alternating series?

2. A series $\displaystyle\sum_{n=1}^{\infty} (-1)^n a_n$ converges when $\{a_n\}$ is _____, _____ and $\displaystyle\lim_{n\to\infty} a_n = $ _____.

3. Give an example of a series where $\displaystyle\sum_{n=0}^{\infty} a_n$ converges but $\displaystyle\sum_{n=0}^{\infty} |a_n|$ does not.

4. The sum of a _____ convergent series can be changed by rearranging the order of its terms.

Problems

In Exercises 5–20, an alternating series $\displaystyle\sum_{n=i}^{\infty} a_n$ is given.

(a) Determine if the series converges or diverges.

(b) Determine if $\displaystyle\sum_{n=0}^{\infty} |a_n|$ converges or diverges.

(c) If $\displaystyle\sum_{n=0}^{\infty} a_n$ converges, determine if the convergence is conditional or absolute.

5. $\displaystyle\sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n^2}$

6. $\displaystyle\sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{\sqrt{n!}}$

7. $\displaystyle\sum_{n=0}^{\infty} (-1)^n \frac{n+5}{3n-5}$

8. $\displaystyle\sum_{n=1}^{\infty} (-1)^n \frac{2^n}{n^2}$

9. $\displaystyle\sum_{n=0}^{\infty} (-1)^{n+1} \frac{3n+5}{n^2-3n+1}$

10. $\displaystyle\sum_{n=1}^{\infty} \frac{(-1)^n}{\ln n + 1}$

11. $\displaystyle\sum_{n=2}^{\infty} (-1)^n \frac{n}{\ln n}$

12. $\displaystyle\sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{1+3+5+\cdots+(2n-1)}$

13. $\displaystyle\sum_{n=1}^{\infty} \cos\left(\pi n\right)$

14. $\displaystyle\sum_{n=1}^{\infty} \frac{\sin\left((n+1/2)\pi\right)}{n \ln n}$

15. $\displaystyle\sum_{n=0}^{\infty} \left(-\frac{2}{3}\right)^n$

16. $\displaystyle\sum_{n=0}^{\infty} (-e)^{-n}$

17. $\displaystyle\sum_{n=0}^{\infty} \frac{(-1)^n n^2}{n!}$

18. $\displaystyle\sum_{n=0}^{\infty} (-1)^n 2^{-n^2}$

19. $\displaystyle\sum_{n=1}^{\infty} \frac{(-1)^n}{\sqrt{n}}$

20. $\displaystyle\sum_{n=1}^{\infty} \frac{(-1000)^n}{n!}$

Let S_n be the nth partial sum of a series. In Exercises 21–24, a convergent alternating series is given and a value of n. Compute S_n and S_{n+1} and use these values to find bounds on the sum of the series.

21. $\displaystyle\sum_{n=1}^{\infty} \frac{(-1)^n}{\ln(n+1)}, \quad n=5$

22. $\displaystyle\sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n^4}, \quad n=4$

23. $\displaystyle\sum_{n=0}^{\infty} \frac{(-1)^n}{n!}, \quad n=6$

24. $\displaystyle\sum_{n=0}^{\infty} \left(-\frac{1}{2}\right)^n, \quad n=9$

In Exercises 25–28, a convergent alternating series is given along with its sum and a value of ε. Use Theorem 75 to find n such that the nth partial sum of the series is within ε of the sum of the series.

25. $\displaystyle\sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n^4} = \frac{7\pi^4}{720}, \quad \varepsilon = 0.001$

26. $\displaystyle\sum_{n=0}^{\infty} \frac{(-1)^n}{n!} = \frac{1}{e}, \quad \varepsilon = 0.0001$

27. $\displaystyle\sum_{n=0}^{\infty} \frac{(-1)^n}{2n+1} = \frac{\pi}{4}, \quad \varepsilon = 0.001$

28. $\displaystyle\sum_{n=0}^{\infty} \frac{(-1)^n}{(2n)!} = \cos 1, \quad \varepsilon = 10^{-8}$

9.6   Ratio and Root Tests

Theorem 67 states that if a series $\sum\limits_{n=1}^{\infty} a_n$ converges, then $\lim\limits_{n\to\infty} a_n = 0$. That is, the terms of $\{a_n\}$ must get very small. Not only must the terms approach 0, they must approach 0 "fast enough": while $\lim\limits_{n\to\infty} 1/n = 0$, the Harmonic Series $\sum\limits_{n=1}^{\infty} \dfrac{1}{n}$ diverges as the terms of $\{1/n\}$ do not approach 0 "fast enough."

The comparison tests of Section 9.4 determine convergence by comparing terms of a series to terms of another series whose convergence is known. This section introduces the Ratio and Root Tests, which determine convergence by analyzing the terms of a series to see if they approach 0 "fast enough."

Ratio Test

Theorem 76   Ratio Test

Let $\{a_n\}$ be a sequence where $\lim\limits_{n\to\infty} \left| \dfrac{a_{n+1}}{a_n} \right| = L.$

1. If $L < 1$, then $\sum\limits_{n=1}^{\infty} a_n$ converges.

2. If $L > 1$ or $L = \infty$, then $\sum\limits_{n=1}^{\infty} a_n$ diverges.

3. If $L = 1$, the Ratio Test is inconclusive.

The principle of the Ratio Test is this: if $\lim\limits_{n\to\infty} \left| \dfrac{a_{n+1}}{a_n} \right| = L < 1$, then for large n, each term of $\{a_n\}$ is significantly smaller than its previous term which is enough to ensure convergence. A full proof can be found at `http://tutorial.math.lamar.edu/Classes/CalcII/RatioTest.aspx`.

Watch the video:
Using the Ratio Test to Determine if a Series Converges #1 at
`https://youtu.be/iy8mhbZTY7g`

Notes:

Example 1   **Applying the Ratio Test**

Use the Ratio Test to determine the convergence of the following series:

$$1. \ \sum_{n=1}^{\infty} \frac{2^n}{n!} \qquad 2. \ \sum_{n=1}^{\infty} \frac{3^n}{n^3} \qquad 3. \ \sum_{n=1}^{\infty} \frac{1}{n^2+1}.$$

SOLUTION

1. $\displaystyle\sum_{n=1}^{\infty} \frac{2^n}{n!}$:

$$\lim_{n\to\infty} \frac{2^{n+1}/(n+1)!}{2^n/n!} = \lim_{n\to\infty} \frac{2^{n+1}n!}{2^n(n+1)!}$$
$$= \lim_{n\to\infty} \frac{2}{n+1}$$
$$= 0.$$

Since the limit is $0 < 1$, by the Ratio Test $\displaystyle\sum_{n=1}^{\infty} \frac{2^n}{n!}$ converges.

2. $\displaystyle\sum_{n=1}^{\infty} \frac{3^n}{n^3}$:

$$\lim_{n\to\infty} \frac{3^{n+1}/(n+1)^3}{3^n/n^3} = \lim_{n\to\infty} \frac{3^{n+1}n^3}{3^n(n+1)^3}$$
$$= \lim_{n\to\infty} \frac{3n^3}{(n+1)^3}$$
$$= 3.$$

Since the limit is $3 > 1$, by the Ratio Test $\displaystyle\sum_{n=1}^{\infty} \frac{3^n}{n^3}$ diverges.

3. $\displaystyle\sum_{n=1}^{\infty} \frac{1}{n^2+1}$:

$$\lim_{n\to\infty} \frac{1/\big((n+1)^2+1\big)}{1/(n^2+1)} = \lim_{n\to\infty} \frac{n^2+1}{(n+1)^2+1}$$
$$= 1.$$

Since the limit is 1, the Ratio Test is inconclusive. We can easily show this series converges using the Direct or Limit Comparison Tests, with each comparing to the series $\displaystyle\sum_{n=1}^{\infty} \frac{1}{n^2}$.

The Ratio Test is not effective when the terms of a series *only* contain algebraic functions (e.g., polynomials). It is most effective when the terms contain some factorials or exponentials. The previous example also reinforces our developing intuition: factorials dominate exponentials, which dominate algebraic functions, which dominate logarithmic functions. In Part 1 of the example,

---

Notes:

the factorial in the denominator dominated the exponential in the numerator, causing the series to converge. In Part 2, the exponential in the numerator dominated the algebraic function in the denominator, causing the series to diverge.

While we have used factorials in previous sections, we have not explored them closely and one is likely to not yet have a strong intuitive sense for how they behave. The following example gives more practice with factorials.

Example 2   **Applying the Ratio Test**

Determine the convergence of $\displaystyle\sum_{n=1}^{\infty} \frac{n!n!}{(2n)!}$.

SOLUTION   Before we begin, be sure to note the difference between $(2n)!$ and $2n!$. When $n = 4$, the former is $8! = 8 \cdot 7 \cdot \ldots \cdot 2 \cdot 1 = 40,320$, whereas the latter is $2(4 \cdot 3 \cdot 2 \cdot 1) = 48$.

Applying the Ratio Test:

$$\lim_{n\to\infty} \frac{(n+1)!(n+1)!/\big(2(n+1)\big)!}{n!n!/(2n)!} = \lim_{n\to\infty} \frac{(n+1)!(n+1)!(2n)!}{n!n!(2n+2)!}$$

Noting that $(2n+2)! = (2n+2) \cdot (2n+1) \cdot (2n)!$, we have

$$= \lim_{n\to\infty} \frac{(n+1)(n+1)}{(2n+2)(2n+1)}$$
$$= 1/4.$$

Since the limit is $1/4 < 1$, by the Ratio Test we conclude $\displaystyle\sum_{n=1}^{\infty} \frac{n!n!}{(2n)!}$ converges.

Root Test

The final test we introduce is the Root Test, which works particularly well on series where each term is raised to a power, and does not work well with terms containing factorials.

Notes:

Theorem 77 Root Test

Let $\{a_n\}$ be a sequence where $\lim\limits_{n \to \infty} |a_n|^{1/n} = L$.

1. If $L < 1$, then $\sum\limits_{n=1}^{\infty} a_n$ converges.

2. If $L > 1$ or $L = \infty$, then $\sum\limits_{n=1}^{\infty} a_n$ diverges.

3. If $L = 1$, the Root Test is inconclusive.

Example 3 Applying the Root Test

Determine the convergence of the following series using the Root Test:

1. $\displaystyle\sum_{n=1}^{\infty} \left(\frac{3n + 1}{5n - 2} \right)^n$
2. $\displaystyle\sum_{n=1}^{\infty} \frac{n^4}{(\ln n)^n}$
3. $\displaystyle\sum_{n=1}^{\infty} \frac{2^n}{n^2}$.

SOLUTION

1. $\lim\limits_{n \to \infty} \left(\left(\frac{3n + 1}{5n - 2} \right)^n \right)^{1/n} = \lim\limits_{n \to \infty} \frac{3n + 1}{5n - 2} = \frac{3}{5}$.

   Since the limit is less than 1, we conclude the series converges. Note: it is difficult to apply the Ratio Test to this series.

2. $\lim\limits_{n \to \infty} \left(\frac{n^4}{(\ln n)^n} \right)^{1/n} = \lim\limits_{n \to \infty} \frac{\left(n^{1/n} \right)^4}{\ln n}$.

   As n grows, the numerator approaches 1 (apply L'Hôpital's Rule) and the denominator grows to infinity. Thus

   $$\lim\limits_{n \to \infty} \frac{\left(n^{1/n} \right)^4}{\ln n} = 0.$$

   Since the limit is less than 1, we conclude the series converges.

3. $\lim\limits_{n \to \infty} \left(\frac{2^n}{n^2} \right)^{1/n} = \lim\limits_{n \to \infty} \frac{2}{\left(n^{1/n} \right)^2} = 2$.

   Since this is greater than 1, we conclude the series diverges.

Notes:

We end here our study of tests to determine convergence. The next section of this text provides strategies for testing series, while the back of the book contains a table summarizing the tests that one may find useful.

While series are worthy of study in and of themselves, our ultimate goal within calculus is the study of Power Series, which we will consider in Section 9.8. We will use power series to create functions where the output is the result of an infinite summation.

Notes:

Exercises 9.6

Terms and Concepts

1. The Ratio Test is not effective when the terms of a sequence only contain _____ functions.
2. The Ratio Test is most effective when the terms of a sequence contains _____ and/or _____ functions.
3. What three convergence tests do not work well with terms containing factorials?
4. The Root Test works particularly well on series where each term is _____ to a _____.

Problems

In Exercises 5–16, determine the convergence of the given series using the Ratio Test. If the Ratio Test is inconclusive, state so and determine convergence with another test.

5. $\displaystyle\sum_{n=0}^{\infty} \frac{2n}{n!}$

6. $\displaystyle\sum_{n=0}^{\infty} \frac{5^n - 3n}{4^n}$

7. $\displaystyle\sum_{n=0}^{\infty} \frac{n!10^n}{(2n)!}$

8. $\displaystyle\sum_{n=1}^{\infty} \frac{5^n + n^4}{7^n + n^2}$

9. $\displaystyle\sum_{n=1}^{\infty} \frac{1}{n}$

10. $\displaystyle\sum_{n=1}^{\infty} \frac{1}{3n^3 + 7}$

11. $\displaystyle\sum_{n=1}^{\infty} \frac{10 \cdot 5^n}{7^n - 3}$

12. $\displaystyle\sum_{n=1}^{\infty} n \cdot \left(\frac{3}{5}\right)^n$

13. $\displaystyle\sum_{n=1}^{\infty} \frac{2 \cdot 4 \cdot 6 \cdot 8 \cdots 2n}{3 \cdot 6 \cdot 9 \cdot 12 \cdots 3n}$

14. $\displaystyle\sum_{n=1}^{\infty} \frac{n!}{5 \cdot 10 \cdot 15 \cdots (5n)}$

15. $\displaystyle\sum_{n=1}^{\infty} e^{-n} n!$

16. $\displaystyle\sum_{n=1}^{\infty} \frac{e^{1/n}}{n^3}$

In Exercises 17–26, determine the convergence of the given series using the Root Test. If the Root Test is inconclusive, state so and determine convergence with another test.

17. $\displaystyle\sum_{n=1}^{\infty} \left(\frac{2n + 5}{3n + 11}\right)^n$

18. $\displaystyle\sum_{n=1}^{\infty} \left(\frac{.9n^2 - n - 3}{n^2 + n + 3}\right)^n$

19. $\displaystyle\sum_{n=1}^{\infty} \frac{2^n n^2}{3^n}$

20. $\displaystyle\sum_{n=1}^{\infty} \frac{1}{n^n}$

21. $\displaystyle\sum_{n=1}^{\infty} \frac{3^n}{n^2 2^{n+1}}$

22. $\displaystyle\sum_{n=1}^{\infty} \frac{4^{n+7}}{7^n}$

23. $\displaystyle\sum_{n=1}^{\infty} \left(\frac{n^2 - n}{n^2 + n}\right)^n$

24. $\displaystyle\sum_{n=1}^{\infty} \left(\frac{1}{n} - \frac{1}{n^2}\right)^n$

25. $\displaystyle\sum_{n=1}^{\infty} \frac{1}{\left(\ln n\right)^n}$

26. $\displaystyle\sum_{n=1}^{\infty} \frac{n^2}{\left(\ln n\right)^n}$

9.7   Strategy for testing series

We have now covered all of the tests for determining the convergence or divergence of a series, which we summarize here. Because more than one test may apply to a given series, you should always go completely through the guidelines and identify all possible tests that you can use. Once you've done this, you can identify the test that will be the easiest for you to use.

1. With a quick glance does it look like the series terms don't converge to zero in the limit, i.e. does $\lim\limits_{n\to\infty} a_n \neq 0$? If so, use the Test for Divergence. Note that you should only use the Test for Divergence if a quick glance suggests that the series terms may not converge to zero in the limit.

2. Is the series a p-series $\left(\sum n^{-p}\right)$ or a geometric series $\left(\sum ar^n\right)$? If so, use the fact that p-series will converge only if $p > 1$ and a geometric series will only converge if $|r| < 1$. Remember as well that often some algebraic manipulation is required to get a geometric series into the correct form.

3. Is the series similar to a p-series or a geometric series? If so, try the Comparison Test.

4. Is the series a rational expression involving only polynomials or polynomials under radicals? If so, try the Comparison test or the Limit Comparison Test. Remember however, that in order to use the Comparison Test and the Limit Comparison Test the series terms all need to be positive.

5. Is the series of the form $\sum (-1)^n a_n$? If so, then the Alternating Series Test may work.

6. Does the series contain factorials or constants raised to powers involving n? If so, then the Ratio Test may work. Note that if the series term contains a factorial then the only test that we have that will work is the Ratio Test.

7. Can the series terms be written in the form $a_n = (b_n)^n$? If so, then the Root Test may work.

8. If $a_n = f(n)$ for some positive, decreasing function and $\int_a^\infty f(x)\,dx$ is easy to evaluate then the Integral Test may work.

Again, remember that these are only a set of guidelines and not a set of hard and fast rules to use when trying to determine the best test to use on a series. If more that one test can be used, try to use the test that will be the easiest for you to use. These guidelines are also summarized in a table in the back of the book.

We now consider several examples.

Notes:

Example 1   **Testing Series**
Determine whether the given series converges absolutely, converges condition-
ally, or diverges.

1. $\displaystyle\sum_{n=2}^{\infty} \frac{(-1)^n n}{n^2 + 3}$

2. $\displaystyle\sum_{n=1}^{\infty} \frac{n^2 - 3n}{4n^2 - 2n + 1}$

3. $\displaystyle\sum_{n=2}^{\infty} \frac{e^n}{(n+3)!}$

SOLUTION

1. We see that this series is alternating so we use the alternating series test.
   The underlying sequence is $\{a_n\} = \{\frac{n}{n^2+3}\}$ which is positive and decreas-
   ing since $a'(n) = \frac{3-n^2}{(n^2+3)^2} < 0$ for $n \geq 2$. We also see $\displaystyle\lim_{n\to\infty} \frac{n}{n^2 + 3} = 0$
   so by the Alternating Series Test $\displaystyle\sum_{n=2}^{\infty} \frac{(-1)^n n}{n^2 + 3}$ converges. We now deter-
   mine if it converges absolutely. Consider the sequence $\displaystyle\sum_{n=2}^{\infty} \left| \frac{(-1)^n n}{n^2 + 3} \right| =$
   $\displaystyle\sum_{n=2}^{\infty} \frac{n}{n^2 + 3}$. We compare this series to $\displaystyle\sum_{n=2}^{\infty} \frac{n}{n^2} = \sum_{n=2}^{\infty} \frac{1}{n}$ which is a diver-
   gent p-series. We also have $\dfrac{n}{n^2 + 3} > \dfrac{n}{n^2} = \dfrac{1}{n}$ so by the Comparison test,
   $\displaystyle\sum_{n=2}^{\infty} \frac{n}{n^2 + 3}$ diverges. Therefore, $\displaystyle\sum_{n=2}^{\infty} \frac{(-1)^n n}{n^2 + 3}$ converges conditionally.

2. $\displaystyle\lim_{n\to\infty} \frac{n^2 - 3n}{4n^2 - 2n + 1} = \frac{1}{4}$ so by the Test for Divergence $\displaystyle\sum_{n=1}^{\infty} \frac{n^2 - 3n}{4n^2 - 2n + 1}$
   diverges.

3. We see the factorial and use the Ratio Test. All terms of the series are

---

Notes:

positive so we consider

$$\lim_{n\to\infty} \frac{a_{n+1}}{a_n} = \lim_{n\to\infty} \frac{\frac{e^{n+1}}{(n+4)!}}{\frac{e^n}{(n+3)!}}$$

$$= \lim_{n\to\infty} \frac{e^{n+1}(n+3)!}{e^n(n+4)!}$$

$$= \lim_{n\to\infty} \frac{e\cdot e^n(n+3)!}{e^n(n+4)(n+3)!}$$

$$= \lim_{n\to\infty} \frac{e}{n+4} = 0 < 1$$

So by the Ratio Test, $\displaystyle\sum_{n=2}^{\infty} \frac{e^n}{(n+3)!}$ converges. Because all of the series terms are positive it converges absolutely.

Notes:

Exercises 9.7

Problems

In Exercises 1–37, determine whether the given series converges absolutely, converges conditionally, or diverges.

1. $\displaystyle\sum_{n=1}^{\infty} \frac{1}{\sqrt[3]{n(n+2)(n+4)}}$

2. $\displaystyle\sum_{n=1}^{\infty} \left(-\frac{2}{3}\right)^{n-1}$

3. $\displaystyle\sum_{n=1}^{\infty} \frac{3^{2n+1}}{n5^{n-1}}$

4. $\displaystyle\sum_{n=1}^{\infty} n^{-2} e^{\frac{1}{n}}$

5. $\displaystyle\sum_{n=1}^{\infty} \frac{n!}{\ln(n+2)}$

6. $\displaystyle\sum_{n=1}^{\infty} (n^2 + 4)(-2)^{1-n}$

7. $\displaystyle\sum_{n=1}^{\infty} \frac{2}{n+4^n}$

8. $\displaystyle\sum_{n=1}^{\infty} \frac{e^n}{n^e}$

9. $\displaystyle\sum_{n=1}^{\infty} \frac{(-1)^n}{\sqrt[4]{n}}$

10. $\displaystyle\sum_{n=1}^{\infty} \frac{\sin\left(\frac{4\pi n}{3}\right)}{n^{4\pi/3}}$

11. $\displaystyle\sum_{n=1}^{\infty} \frac{3^n n!}{(n+2)!}$

12. $\displaystyle\sum_{n=1}^{\infty} (-1)^n \frac{\sqrt{n}}{n+1}$

13. $\displaystyle\sum_{n=1}^{\infty} \frac{1-\cos n}{n^3}$

14. $\displaystyle\sum_{n=1}^{\infty} \frac{4+3n-5n^3}{2+n^3}$

15. $\displaystyle\sum_{n=1}^{\infty} (-1)^n \frac{n^2+1}{n^4+1}$

16. $\displaystyle\sum_{n=1}^{\infty} \frac{(3n)^n}{n^{3n}}$

17. $\displaystyle\sum_{n=1}^{\infty} \frac{e^{2n}}{(2n-1)!}$

18. $\displaystyle\sum_{n=1}^{\infty} \frac{1}{x\sqrt{\ln x}}$

19. $\displaystyle\sum_{n=2}^{\infty} (-1)^{n-1} \frac{\sqrt{\ln n}}{n}$

20. $\displaystyle\sum_{n=1}^{\infty} \frac{n^2}{(-2)^n}$

21. $\displaystyle\sum_{n=1}^{\infty} \frac{(-1)^n}{(2n+5)^3}$

22. $\displaystyle\sum_{n=1}^{\infty} \frac{n!}{(-10)^n}$

23. $\displaystyle\sum_{n=1}^{\infty} \frac{\ln n}{n!}$

24. $\displaystyle\sum_{n=1}^{\infty} \frac{1}{(-3)^n + n}$

25. $\displaystyle\sum_{n=1}^{\infty} \frac{(-1)^n(n-2)}{10n+5}$

26. $\displaystyle\sum_{n=1}^{\infty} \frac{(-3)^n}{n^3}$

27. $\displaystyle\sum_{n=1}^{\infty} \frac{\cos(1/n)}{\sqrt{n}}$

28. $\displaystyle\sum_{n=1}^{\infty} \frac{(-1)^n(n^2+4n-2)}{n^3+4n^2-3n+7}$

29. $\displaystyle\sum_{n=1}^{\infty} \frac{n^4(-4)^n}{n!}$

30. $\displaystyle\sum_{n=1}^{\infty} \frac{n^2}{(-3)^n + n}$

31. $\displaystyle\sum_{n=1}^{\infty} \frac{(-1)^n n}{\sqrt{n^2+4n+1}}$

32. $\displaystyle\sum_{n=1}^{\infty} \frac{(-3)^n}{n^n}$

33. $\displaystyle\sum_{n=1}^{\infty} \frac{n!n!n!}{(3n)!}$

34. $\displaystyle\sum_{n=1}^{\infty} \frac{(-1)^n}{\ln n}$

35. $\displaystyle\sum_{n=1}^{\infty} \left(\frac{n+2}{n+1}\right)^n$

36. $\displaystyle\sum_{n=1}^{\infty} \frac{n^3}{(\ln n)^n}$

37. $\displaystyle\sum_{n=1}^{\infty} \left(\frac{1}{n} - \frac{1}{n+2}\right)$

9.8   Power Series

So far, our study of series has examined the question of "Is the sum of these infinite terms finite?," i.e., "Does the series converge?" We now approach series from a different perspective: as a function. Given a value of x, we evaluate $f(x)$ by finding the sum of a particular series that depends on x (assuming the series converges). We start this new approach to series with a definition.

Definition 40   Power Series

Let $\{a_n\}$ be a sequence, let x be a variable, and let c be a real number.

1. The **power series in** x is the series

$$\sum_{n=0}^{\infty} a_n x^n = a_0 + a_1 x + a_2 x^2 + a_3 x^3 + \cdots$$

2. The **power series in** x **centered at** c is the series

$$\sum_{n=0}^{\infty} a_n (x - c)^n = a_0 + a_1(x - c) + a_2(x - c)^2 + a_3(x - c)^3 + \cdots$$

Example 1   Examples of power series
Write out the first five terms of the following power series:

1. $\displaystyle\sum_{n=0}^{\infty} x^n$ 2. $\displaystyle\sum_{n=1}^{\infty} (-1)^{n+1} \frac{(x+1)^n}{n}$ 3. $\displaystyle\sum_{n=0}^{\infty} (-1)^{n+1} \frac{(x - \pi)^{2n}}{(2n)!}$.

SOLUTION

1. One of the conventions we adopt is that $x^0 = 1$ regardless of the value of x. Therefore

$$\sum_{n=0}^{\infty} x^n = 1 + x + x^2 + x^3 + x^4 + \cdots$$

This is a geometric series in x.

2. This series is centered at $c = -1$. Note how this series starts with $n = 1$. We could rewrite this series starting at $n = 0$ with the understanding that

---

Notes:

$a_0 = 0$, and hence the first term is 0.

$$\sum_{n=1}^{\infty}(-1)^{n+1}\frac{(x+1)^n}{n} = (x+1) - \frac{(x+1)^2}{2} + \frac{(x+1)^3}{3} - \frac{(x+1)^4}{4} + \frac{(x+1)^5}{5} \cdots$$

3. This series is centered at $c = \pi$. Recall that $0! = 1$.

$$\sum_{n=0}^{\infty}(-1)^{n+1}\frac{(x-\pi)^{2n}}{(2n)!} = -1 + \frac{(x-\pi)^2}{2} - \frac{(x-\pi)^4}{24} + \frac{(x-\pi)^6}{6!} - \frac{(x-\pi)^8}{8!} \cdots$$

We introduced power series as a type of function, where a value of x is given and the sum of a series is returned. Of course, not every series converges. For instance, in part 1 of Example 1, we recognized the series $\sum_{n=0}^{\infty} x^n$ as a geometric series in x. Theorem 64 states that this series converges only when $|x| < 1$.

This raises the question: "For what values of x will a given power series converge?," which leads us to a theorem and definition.

Theorem 78   Convergence of Power Series

Let a power series $\sum_{n=0}^{\infty} a_n(x-c)^n$ be given. Then one of the following is true:

1. The series converges only at $x = c$.

2. There is an $R > 0$ such that the series converges for all x in $(c - R, c + R)$ and diverges for all $x < c - R$ and $x > c + R$.

3. The series converges for all x.

The value of R is important when understanding a power series, hence it is given a name in the following definition. Also, note that part 2 of Theorem 78 makes a statement about the interval $(c - R, c + R)$, but the not the endpoints of that interval. A series may or may not converge at these endpoints.

Notes:

Definition 41   Radius and Interval of Convergence

1. The number R given in Theorem 78 is the **radius of convergence** of a given series. When a series converges for only $x = c$, we say the radius of convergence is 0, i.e., $R = 0$. When a series converges for all x, we say the series has an infinite radius of convergence, i.e., $R = \infty$.

2. The **interval of convergence** is the set of all values of x for which the series converges.

To find the values of x for which a given series converges, we will use the convergence tests we studied previously (especially the Ratio Test). However, the tests all required that the terms of a series be positive. The following theorem gives us a work–around to this problem.

Theorem 79   The Radius of Convergence of a Series and Absolute Convergence

The series $\displaystyle\sum_{n=0}^{\infty} a_n(x-c)^n$ and $\displaystyle\sum_{n=0}^{\infty} |a_n(x-c)^n|$ have the same radius of convergence R.

Theorem 79 allows us to find the radius of convergence R of a series by applying the Ratio Test (or any applicable test) to the absolute value of the terms of the series. We practice this in the following example.

Watch the video:
Power Series — Finding the Interval of Convergence
at
`https://youtu.be/01LzAU__J-0`

Notes:

Example 2   **Determining the radius and interval of convergence.**

Find the radius and interval of convergence for each of the following series:

1. $\displaystyle\sum_{n=0}^{\infty} \frac{x^n}{n!}$ 2. $\displaystyle\sum_{n=1}^{\infty} (-1)^{n+1}\frac{x^n}{n}$ 3. $\displaystyle\sum_{n=0}^{\infty} 2^n(x-3)^n$ 4. $\displaystyle\sum_{n=0}^{\infty} n!x^n$

SOLUTION

1. We apply the Ratio Test to the series $\displaystyle\sum_{n=0}^{\infty} \left|\frac{x^n}{n!}\right|$:

$$\lim_{n\to\infty} \frac{\left|x^{n+1}/(n+1)!\right|}{\left|x^n/n!\right|} = \lim_{n\to\infty} \left|\frac{x^{n+1}}{x^n} \cdot \frac{n!}{(n+1)!}\right|$$

$$= \lim_{n\to\infty} \left|\frac{x}{n+1}\right|$$

$$= 0 \text{ for all } x.$$

The Ratio Test shows us that regardless of the choice of x, the series converges. Therefore the radius of convergence is $R = \infty$, and the interval of convergence is $(-\infty, \infty)$.

2. We apply the Ratio Test to the series $\displaystyle\sum_{n=1}^{\infty} \left|(-1)^{n+1}\frac{x^n}{n}\right| = \sum_{n=1}^{\infty} \left|\frac{x^n}{n}\right|$:

$$\lim_{n\to\infty} \frac{\left|x^{n+1}/(n+1)\right|}{\left|x^n/n\right|} = \lim_{n\to\infty} \left|\frac{x^{n+1}}{x^n} \cdot \frac{n}{n+1}\right|$$

$$= \lim_{n\to\infty} |x| \frac{n}{n+1}$$

$$= |x|.$$

The Ratio Test states a series converges if the limit of $|a_{n+1}/a_n| = L < 1$. We found the limit above to be $|x|$; therefore, the power series converges when $|x| < 1$, or when x is in $(-1, 1)$. Thus the radius of convergence is $R = 1$.

To determine the interval of convergence, we need to check the endpoints of $(-1, 1)$. When $x = -1$, we have the opposite of the Harmonic Series:

$$\sum_{n=1}^{\infty} (-1)^{n+1}\frac{(-1)^n}{n} = \sum_{n=1}^{\infty} \frac{(-1)^{2n+1}}{n}$$

$$= \sum_{n=1}^{\infty} \frac{-1}{n}$$

$$= -\infty.$$

Notes:

The series diverges when $x = -1$.

When $x = 1$, we have the series $\sum_{n=1}^{\infty} (-1)^{n+1} \dfrac{(1)^n}{n}$, which is the Alternating Harmonic Series, which converges. Therefore the interval of convergence is $(-1, 1]$.

3. We apply the Ratio Test to the series $\sum_{n=0}^{\infty} |2^n(x-3)^n|$:

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

$$= \lim_{n\to\infty} |2(x-3)|.$$

According to the Ratio Test, the series converges when $|2(x-3)| < 1 \implies |x-3| < 1/2$. The series is centered at 3, and x must be within $1/2$ of 3 in order for the series to converge. Therefore the radius of convergence is $R = 1/2$, and we know that the series converges absolutely for all x in $(3 - 1/2, 3 + 1/2) = (2.5, 3.5)$.

We check for convergence at the endpoints to find the interval of convergence. When $x = 2.5$, we have:

$$\sum_{n=0}^{\infty} 2^n(2.5 - 3)^n = \sum_{n=0}^{\infty} 2^n(-1/2)^n$$

$$= \sum_{n=0}^{\infty} (-1)^n,$$

which diverges. A similar process shows that the series also diverges at $x = 3.5$. Therefore the interval of convergence is $(2.5, 3.5)$.

4. We apply the Ratio Test to $\sum_{n=0}^{\infty} |n!x^n|$:

$$\lim_{n\to\infty} \frac{\left|(n+1)!x^{n+1}\right|}{\left|n!x^n\right|} = \lim_{n\to\infty} |(n+1)x|$$

$$= \infty \text{ for all } x, \text{ except } x = 0.$$

The Ratio Test shows that the series diverges for all x except $x = 0$. Therefore the radius of convergence is $R = 0$.

Notes:

Power Series as Functions

We can use a power series to define a function:

$$f(x) = \sum_{n=0}^{\infty} a_n x^n$$

where the domain of f is a subset of the interval of convergence of the power series. One can apply calculus techniques to such functions; in particular, we can find derivatives and antiderivatives.

Theorem 80   **Derivatives and Indefinite Integrals of Power Series Functions**

Let $f(x) = \sum_{n=0}^{\infty} a_n (x - c)^n$ be a function defined by a power series, with radius of convergence R.

1. $f(x)$ is continuous and differentiable on $(c - R, c + R)$.

2. $f'(x) = \sum_{n=1}^{\infty} a_n \cdot n \cdot (x - c)^{n-1}$, with radius of convergence R.

3. $\int f(x)\, dx = C + \sum_{n=0}^{\infty} a_n \frac{(x - c)^{n+1}}{n + 1}$, with radius of convergence R.

A few notes about Theorem 80:

1. The theorem states that differentiation and integration do not change the radius of convergence. It does not state anything about the *interval* of convergence. They are not always the same.

2. Notice how the summation for $f'(x)$ starts with $n = 1$. This is because the constant term a_0 of $f(x)$ goes to 0.

3. Differentiation and integration are simply calculated term–by–term using previous rules of integration and differentiation.

Example 3   **Derivatives and indefinite integrals of power series**

Let $f(x) = \sum_{n=0}^{\infty} x^n$. Find the following along with their respective intervals of convergence.

$$1. \quad f'(x) \qquad \text{and} \qquad 2. \quad F(x) = \int f(x)\, dx$$

Notes:

Solution We find the derivative and indefinite integral of $f(x)$, following Theorem 80.

1.
$$f(x) = 1 + x + x^2 + x^3 + x^4 + \cdots = \sum_{n=0}^{\infty} x^n$$

$$f'(x) = 0 + 1 + 2x + 3x^2 + 4x^3 + \cdots = \sum_{n=1}^{\infty} nx^{n-1}$$

In Example 1, we recognized that $\sum_{n=0}^{\infty} x^n$ is a geometric series in x. We know that such a geometric series converges when $|x| < 1$; that is, the interval of convergence is $(-1, 1)$.

To determine the interval of convergence of $f'(x)$, we consider the endpoints of $(-1, 1)$. When $x = -1$ we have

$$f'(-1) = \sum_{n=1}^{\infty} n(-1)^{n-1}$$

which diverges by the Test for Divergence and when $x = 1$ we have

$$f'(1) = \sum_{n=1}^{\infty} n$$

which also diverges by the Test for Divergence. Therefore, the interval of convergence of $f'(x)$ is $(-1, 1)$.

2.
$$f(x) = 1 + x + x^2 + x^3 + \cdots = \sum_{n=0}^{\infty} x^n$$

$$F(x) = \int f(x)\, dx = C + x + \frac{x^2}{2} + \frac{x^3}{3} + \frac{x^4}{4} + \cdots$$

$$= C + \sum_{n=0}^{\infty} \frac{x^{n+1}}{n+1} = C + \sum_{n=1}^{\infty} \frac{x^n}{n}$$

To find the interval of convergence of $F(x)$, we again consider the endpoints of $(-1, 1)$. When $x = -1$ we have

$$F(-1) = C + \sum_{n=1}^{\infty} \frac{(-1)^n}{n}$$

The value of C is irrelevant; notice that the rest of the series is an Alternating Series that whose terms converge to 0. By the Alternating Series

Notes:

Test, this series converges. (In fact, we can recognize that the terms of the series after C are the opposite of the Alternating Harmonic Series. We can thus say that $F(-1) = C - \ln 2$.)

$$F(1) = C + \sum_{n=1}^{\infty} \frac{1}{n}$$

Notice that this summation is $C +$ the Harmonic Series, which diverges. Since F converges for $x = -1$ and diverges for $x = 1$, the interval of convergence of $F(x)$ is $[-1, 1)$.

The previous example showed how to take the derivative and indefinite integral of a power series without motivation for why we care about such operations. We may care for the sheer mathematical enjoyment "that we can", which is motivation enough for many. However, we would be remiss to not recognize that we can learn a great deal from taking derivatives and indefinite integrals.

Recall that $f(x) = \sum_{n=0}^{\infty} x^n$ in Example 3 is a geometric series. According to Theorem 64, this series converges to $1/(1 - x)$ when $|x| < 1$. Thus we can say

$$f(x) = \sum_{n=0}^{\infty} x^n = \frac{1}{1 - x}, \quad \text{on} \quad (-1, 1).$$

Integrating the power series, (as done in Example 3,) we find

$$F(x) = C_1 + \sum_{n=0}^{\infty} \frac{x^{n+1}}{n + 1}, \tag{9.4}$$

while integrating the function $f(x) = 1/(1 - x)$ gives

$$F(x) = -\ln|1 - x| + C_2. \tag{9.5}$$

Equating Equations (9.4) and (9.5), we have

$$F(x) = C_1 + \sum_{n=0}^{\infty} \frac{x^{n+1}}{n + 1} = -\ln|1 - x| + C_2.$$

Letting $x = 0$, we have $F(0) = C_1 = C_2$. This implies that we can drop the constants and conclude

$$\sum_{n=0}^{\infty} \frac{x^{n+1}}{n + 1} = -\ln|1 - x|.$$

Notes:

We established in Example 3 that the series $\displaystyle\sum_{n=0}^{\infty} \frac{x^{n+1}}{n+1}$ converges at $x = -1$; substituting $x = -1$ on both sides of the above equality gives

$$-1 + \frac{1}{2} - \frac{1}{3} + \frac{1}{4} - \frac{1}{5} + \cdots = -\ln 2.$$

On the left we have the opposite of the Alternating Harmonic Series; on the right, we have $-\ln 2$. We conclude that

$$1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \cdots = \ln 2.$$

In Example 9.5.1 of Section 9.5 we said the Alternating Harmonic Series converges to $\ln 2$, but did not show why this was the case. The work above shows how we conclude that the Alternating Harmonic Series Converges to $\ln 2$.

We use this type of analysis in the next example.

Example 4   **Analyzing power series functions**

Let $f(x) = \displaystyle\sum_{n=0}^{\infty} \frac{x^n}{n!}$. Find $f'(x)$ and $\displaystyle\int f(x)\, dx$, and use these to analyze the behavior of $f(x)$.

SOLUTION   We start by making two notes: first, in Example 2, we found the interval of convergence of this power series is $(-\infty, \infty)$. Second, we will find it useful later to have a few terms of the series written out:

$$\sum_{n=0}^{\infty} \frac{x^n}{n!} = 1 + x + \frac{x^2}{2} + \frac{x^3}{6} + \frac{x^4}{24} + \cdots \tag{9.6}$$

We now find the derivative:

$$f'(x) = \sum_{n=1}^{\infty} n \frac{x^{n-1}}{n!}$$

$$= \sum_{n=1}^{\infty} \frac{x^{n-1}}{(n-1)!} = 1 + x + \frac{x^2}{2!} + \cdots.$$

Since the series starts at $n = 1$ and each term refers to $(n-1)$, we can re-index the series starting with $n = 0$:

$$= \sum_{n=0}^{\infty} \frac{x^n}{n!}$$

$$= f(x).$$

Notes:

We found the derivative of $f(x)$ is $f(x)$. The only functions for which this is true are of the form $y = ce^x$ for some constant c. As $f(0) = 1$ (see Equation (9.6)), c must be 1. Therefore we conclude that

$$f(x) = \sum_{n=0}^{\infty} \frac{x^n}{n!} = e^x$$

for all x.

We can also find $\int f(x)\,dx$:

$$\int f(x)\,dx = C + \sum_{n=0}^{\infty} \frac{x^{n+1}}{n!(n+1)}$$

$$= C + \sum_{n=0}^{\infty} \frac{x^{n+1}}{(n+1)!}$$

We write out a few terms of this last series:

$$C + \sum_{n=0}^{\infty} \frac{x^{n+1}}{(n+1)!} = C + x + \frac{x^2}{2} + \frac{x^3}{6} + \frac{x^4}{24} + \cdots$$

The integral of $f(x)$ differs from $f(x)$ only by a constant, again indicating that $f(x) = e^x$.

Example 4 and the work following Example 3 established relationships between a power series function and "regular" functions that we have dealt with in the past. In general, given a power series function, it is difficult (if not impossible) to express the function in terms of elementary functions. We chose examples where things worked out nicely.

Representations of Functions with Power Series

It can be difficult to recognize an elementary function by its power series expansion. It is far easier to start with a known function, expressed in terms of elementary functions, and represent it as a power series function. One may wonder why we would bother doing so, as the latter function probably seems more complicated.

Let's start off with a series we already know how to do, although when we first ran across this series we didn't think of it as a power series nor did we acknowledge that it represented a function. Recall that the geometric series is

$$\sum_{n=0}^{\infty} ar^n = \frac{a}{1-r} \qquad \text{provided } |r| < 1.$$

Notes:

We also know that if $|r| \geq 1$ the series diverges. Now, if we take $a = 1$ and $r = x$ this becomes,

$$\sum_{n=0}^{\infty} x^n = \frac{1}{1-x} \qquad \text{provided } |x| < 1 \tag{9.7}$$

Turning this around we can see that we can represent the function

$$f(x) = \frac{1}{1-x} \tag{9.8}$$

with the power series

$$\sum_{n=0}^{\infty} x^n \qquad \text{provided } |x| < 1. \tag{9.9}$$

This provision is important. We can clearly plug any number other than $x = 1$ into the function, however, we will only get a convergent power series if $|x| < 1$. This means the equality in Equation (9.7) will only hold if $|x| < 1$. For any other value of x the equality won't hold. Note as well that we can also use this to acknowledge that the radius of convergence of this power series is $R = 1$ and the interval of convergence is $|x| < 1$.

This idea of convergence is important here. We will be representing many functions as power series and it will be important to recognize that the representations will often only be valid for a range of x's and that there may be values of x that we can plug into the function that we can't plug into the power series representation.

In this section we are going to concentrate on representing functions with power series where the function can be related back to a geometric series. In this way we will hopefully become familiar with some of the kinds of manipulations that we will sometimes need when working with power series. We will see in Section 9.10 that this strategy is useful for integrating functions that don't have elementary derivatives.

Example 5   **Finding a Power Series**

Find a power series representation for $g(x) = \dfrac{1}{1+x^3}$ and determine its interval of convergence.

SOLUTION   We want to relate this function back to Equation (9.8). This is actually easier than it might look. Recall that the x in Equation (9.8) is simply a variable and can represent anything. So, a quick rewrite of $g(x)$ gives,

$$g(x) = \frac{1}{1-(-x^3)}$$

Notes:

and so the $-x^3$ holds the same place as the x in Equation (9.8). Therefore, all we need to do is replace the x in Equation (9.9) and we've got a power series representation for $g(x)$.

$$g(x) = \sum_{n=0}^{\infty} \left(-x^3\right)^n \qquad \text{provided } \left|-x^3\right| < 1$$

Notice that we replaced both the x in the power series and in the interval of convergence. All we need to do now is a little simplification.

$$g(x) = \sum_{n=0}^{\infty} (-1)^n x^{3n} \qquad \text{provided } |x| < 1$$

So, in this case the interval of convergence is the same as the original power series. This usually won't happen. More often than not the new interval of convergence will be different from the original interval of convergence.

Example 6   **Finding a Power Series**

Find a power series representation for $h(x) = \dfrac{2x^2}{1 + x^3}$ and determine its interval of convergence.

    SOLUTION   This function is similar to the previous function, however the numerator is different. Since Equation (9.8) doesn't have an x in the numerator it appears that we can't relate this function back to that. However, now that we've worked the first example this one is actually very simple since we can use the result of the answer from that example. To see how to do this let's first rewrite the function a little.

$$h(x) = 2x^2 \frac{1}{1 + x^3}.$$

Now, from the first example we've already got a power series for the second term so let's use that to write the function as,

$$h(x) = 2x^2 \sum_{n=0}^{\infty} (-1)^n x^{3n} \qquad \text{provided } |x| < 1$$

Notice that the presence of x's outside of the series will NOT affect its convergence and so the interval of convergence remains the same. The last step is to bring the coefficient into the series and we'll be done. When we do this make sure and combine the x's as well. We typically only want a single x in a power series.

$$h(x) = \sum_{n=0}^{\infty} 2 (-1)^n x^{3n+2} \qquad \text{provided } |x| < 1.$$

Notes:

As we saw in the previous example we can often use previous results to help us out. This is an important idea to remember as it can often greatly simplify our work.

Example 7   **Finding a Power Series**

Find a power series representation for $f(x) = \dfrac{x}{5 - x}$ and determine its interval of convergence.

    SOLUTION   So again, we have an x in the numerator. As with the last example factor x out and we have $f(x) = x\dfrac{1}{5 - x}$. If we had a power series representation for $g(x) = \dfrac{1}{5 - x}$ we could get a power series representation for $f(x)$. We need the number in the denominator to be a one so we rewrite the denominator.

$$g(x) = \frac{1}{5}\frac{1}{1 - \frac{x}{5}}$$

    Now all we need to do to get a power series representation is to replace the x in Equation (9.9) with $\dfrac{x}{5}$. Doing this gives

$$g(x) = \frac{1}{5}\sum_{n=0}^{\infty}\left(\frac{x}{5}\right)^n \qquad \text{provided } \left|\frac{x}{5}\right| < 1.$$

    Now simplify the series.

$$g(x) = \frac{1}{5}\sum_{n=0}^{\infty}\frac{x^n}{5^n}$$

$$= \sum_{n=0}^{\infty}\frac{x^n}{5^{n+1}}$$

    The interval of convergence for this series is

$$\left|\frac{x}{5}\right| < 1 \qquad \Rightarrow \qquad \frac{1}{5}|x| < 1 \qquad \Rightarrow \qquad |x| < 5$$

    We now have a power series representation for $g(x)$ but we need to find a power series representation for the original function. All we need to do for this

---

Notes:

is to multiply the power series representative for $g(x)$ by x and we'll have it.

$$f(x) = x\frac{1}{5-x}$$

$$= x\sum_{n=0}^{\infty} \frac{x^n}{5^{n+1}}$$

$$= \sum_{n=0}^{\infty} \frac{x^{n+1}}{5^{n+1}}$$

The interval of convergence doesn't change and so it will be $|x| < 5$.

We now consider several examples where differentiation and integration of power series from Theorem 80 are used to write the power series for a function.

Example 8   **Differentiating a Power Series**

Find a power series representation for $g(x) = \dfrac{1}{(1-x)^2}$ and determine its radius of convergence.

SOLUTION   We know that

$$\frac{1}{(1-x)^2} = \frac{d}{dx}\left(\frac{1}{1-x}\right).$$

Since we have a power series representation for $\dfrac{1}{1-x}$, we can differentiate that power series to get a power series representation for $g(x)$.

$$g(x) = \frac{1}{1-x}$$

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

$$= \frac{d}{dx}\left(\sum_{n=0}^{\infty} x^n\right)$$

$$= \sum_{n=1}^{\infty} nx^{n-1}$$

Since the original power series had a radius of convergence of $R = 1$ the derivative, and hence $g(x)$, will also have a radius of convergence of $R = 1$.

Notes:

Example 9   **Integrating a Power Series**

Find a power series representation for $h(x) = \ln(5-x)$ and determine its radius of convergence.

SOLUTION   In this case we need the fact that

$$\int \frac{1}{5-x}\,dx = -\ln(5-x).$$

Recall that we found a power series representation for $\dfrac{1}{5-x}$ in Example 7. We now have

$$
\begin{aligned}
\ln(5-x) &= -\int \frac{1}{5-x}\,dx \\
&= -\int \sum_{n=0}^{\infty} \frac{x^n}{5^{n+1}}\,dx \qquad \text{where } |x| < 5 \\
&= C - \sum_{n=0}^{\infty} \frac{x^{n+1}}{(n+1)5^{n+1}} \qquad \text{where } |x| < 5
\end{aligned}
$$

We can find the constant of integration, C, by substituting in a value of x. A good choice is $x = 0$ as the series is usually easy to evaluate there.

$$
\begin{aligned}
\ln(5-0) &= C - \sum_{n=0}^{\infty} \frac{0^{n+1}}{(n+1)5^{n+1}} \\
\ln(5-0) &= C
\end{aligned}
$$

So, the final answer is,

$$\ln(5-x) = \ln(5) - \sum_{n=0}^{\infty} \frac{x^{n+1}}{(n+1)5^{n+1}},$$

and the radius of convergence is 5. Notice that $x = -5$ allows for convergence so the interval of convergence is $[-5, 5)$.

Notes:

Exercises 9.8

Terms and Concepts

1. We adopt the convention that $x^0 =$____, regardless of the value of x.

2. What is the difference between the radius of convergence and the interval of convergence?

3. If the radius of convergence of $\sum_{n=0}^{\infty} a_n x^n$ is 5, what is the radius of convergence of $\sum_{n=1}^{\infty} n \cdot a_n x^{n-1}$?

4. If the radius of convergence of $\sum_{n=0}^{\infty} a_n x^n$ is 5, what is the radius of convergence of $\sum_{n=0}^{\infty} (-1)^n a_n x^n$?

Problems

In Exercises 5–8, write out the sum of the first 5 terms of the given power series.

5. $\sum_{n=0}^{\infty} 2^n x^n$

6. $\sum_{n=1}^{\infty} \frac{1}{n^2} x^n$

7. $\sum_{n=0}^{\infty} \frac{1}{n!} x^n$

8. $\sum_{n=0}^{\infty} \frac{(-1)^n}{(2n)!} x^{2n}$

In Exercises 9–28, a power series is given.

 (a) Find the radius of convergence.

 (b) Find the interval of convergence.

9. $\sum_{n=0}^{\infty} \frac{(-1)^{n+1}}{n!} x^n$

10. $\sum_{n=0}^{\infty} n x^n$

11. $\sum_{n=1}^{\infty} \frac{(-1)^n (x-3)^n}{n}$

12. $\sum_{n=0}^{\infty} \frac{(x+4)^n}{n!}$

13. $\sum_{n=0}^{\infty} \frac{x^n}{2^n}$

14. $\sum_{n=0}^{\infty} \frac{(-1)^n (x-5)^n}{10^n}$

15. $\sum_{n=0}^{\infty} 5^n (x-1)^n$

16. $\sum_{n=0}^{\infty} (-2)^n x^n$

17. $\sum_{n=0}^{\infty} \sqrt{n} x^n$

18. $\sum_{n=0}^{\infty} \frac{n}{3^n} x^n$

19. $\sum_{n=0}^{\infty} \frac{3^n}{n!} (x-5)^n$

20. $\sum_{n=0}^{\infty} (-1)^n n! (x-10)^n$

21. $\sum_{n=1}^{\infty} \frac{x^n}{n^2}$

22. $\sum_{n=1}^{\infty} \frac{(x+2)^n}{n^3}$

23. $\sum_{n=0}^{\infty} n! \left(\frac{x}{10}\right)^n$

24. $\sum_{n=0}^{\infty} n^2 \left(\frac{x+4}{4}\right)^n$

25. $\sum_{n=1}^{\infty} \frac{(3x-2)^n}{n3^n}$

26. $\sum_{n=1}^{\infty} \frac{x^n}{5^n n^5}$

27. $\sum_{n=2}^{\infty} \frac{x^n}{(\ln n)^n}$

28. $\sum_{n=1}^{\infty} (-1)^n \frac{x^{2n+1}}{(2n+1)!}$

In Exercises 29–33, write the following functions as a power series and give the radius of convergence.

29. $f(x) = \frac{x}{1-8x}$

30. $f(x) = \frac{x^7}{8+x^3}$

31. $f(x) = \frac{6}{1+7x^4}$

32. $f(x) = \frac{x^3}{3-x^2}$

33. $f(x) = \frac{3x^2}{5-2\sqrt[3]{x}}$

34.
 (a) Use differentiation to find a power series representation for $f(x) = \frac{1}{(1+x)^2}$. What is the radius of convergence?

 (b) Use part (a) to find a power series for $f(x) = \frac{1}{(1+x)^3}$.

 (c) Use part (b) to find a power series for $f(x) = \frac{x^2}{(1+x)^3}$.

In Exercises 35–41, find a power series representation for the function and determine the radius of convergence.

35. $f(x) = \ln(3 - x)$

36. $f(x) = \dfrac{x}{(1 + 9x)^2}$

37. $f(x) = \ln\left(\dfrac{1 + x}{1 - x}\right)$

38. $f(x) = \tan^{-1} x$

39. $f(x) = x^2 \tan^{-1}(x^3)$

40. $f(x) = \dfrac{1 + x}{(1 - x)^2}$

41. $f(x) = \left(\dfrac{x}{2 - x}\right)^3$

9.9   Taylor Polynomials

Consider a function $y = f(x)$ and a point $\big(c, f(c)\big)$. The derivative, $f'(c)$, gives the instantaneous rate of change of f at $x = c$. Of all lines that pass through the point $\big(c, f(c)\big)$, the line that best approximates f at this point is the tangent line; that is, the line whose slope (rate of change) is $f'(c)$.

In Figure 9.18, we see a function $y = f(x)$ graphed. The table below the graph shows that $f(0) = 2$ and $f'(0) = 1$; therefore, the tangent line to f at $x = 0$ is $p_1(x) = 1(x-0) + 2 = x + 2$. The tangent line is also given in the figure. Note that "near" $x = 0$, $p_1(x) \approx f(x)$; that is, the tangent line approximates f well.

One shortcoming of this approximation is that the tangent line only matches the slope of f; it does not, for instance, match the concavity of f. We can find a polynomial, $p_2(x)$, that does match the concavity without much difficulty, though. The table in Figure 9.18 gives the following information:

$$f(0) = 2 \qquad f'(0) = 1 \qquad f''(0) = 2.$$

Therefore, we want our polynomial $p_2(x)$ to have these same properties. That is, we need

$$p_2(0) = 2 \qquad p_2'(0) = 1 \qquad p_2''(0) = 2.$$

This is simply an initial–value problem. We can solve this using the techniques first described in Section 5.1. To keep $p_2(x)$ as simple as possible, we'll assume that not only $p_2''(0) = 2$, but that $p_2''(x) = 2$. That is, the second derivative of p_2 is constant.

If $p_2''(x) = 2$, then $p_2'(x) = 2x + C$ for some constant C. Since we have determined that $p_2'(0) = 1$, we find that $C = 1$ and so $p_2'(x) = 2x + 1$. Finally, we can compute $p_2(x) = x^2 + x + C$. Using our initial values, we know $p_2(0) = 2$ so $C = 2$. We conclude that $p_2(x) = x^2 + x + 2$. This function is plotted with f in Figure 9.19.

We can repeat this approximation process by creating polynomials of higher degree that match more of the derivatives of f at $x = 0$. In general, a polynomial of degree n can be created to match the first n derivatives of f. Figure 9.19 also shows $p_4(x) = -x^4/2 - x^3/6 + x^2 + x + 2$, whose first four derivatives at 0 match those of f. (Using the table in Figure 9.18, start with $p_4^{(4)}(x) = -12$ and solve the related initial–value problem.)

As we use more and more derivatives, our polynomial approximation to f gets better and better. In this example, the interval on which the approximation is "good" gets bigger and bigger. Figure 9.20 shows $p_{13}(x)$; we can visually affirm that this polynomial approximates f very well on $[-2, 3]$. The polynomial $p_{13}(x)$ is fairly complicated:
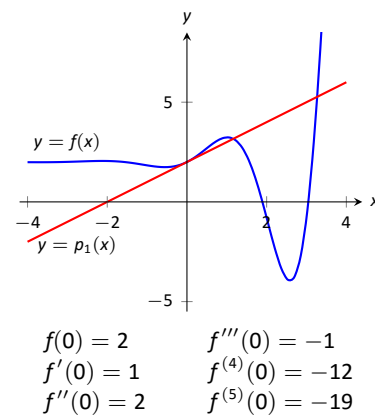
Notes:

Figure 9.18: Plotting $y = f(x)$ and a table of derivatives of f evaluated at 0.

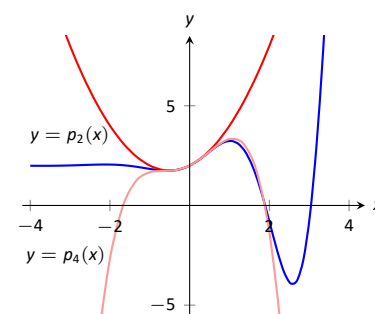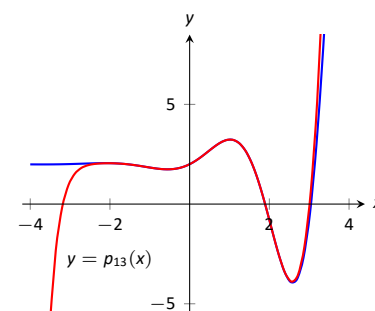|	
$f(0) = 2$	$f'''(0) = -1$
$f'(0) = 1$	$f^{(4)}(0) = -12$
$f''(0) = 2$	$f^{(5)}(0) = -19$

Figure 9.19: Plotting f, p_2, and p_4.

Figure 9.20: Plotting f and p_{13}.

$$\frac{16901x^{13}}{6227020800} + \frac{13x^{12}}{1209600} - \frac{1321x^{11}}{39916800} - \frac{779x^{10}}{1814400} - \frac{359x^9}{362880} + \frac{x^8}{240} + \frac{139x^7}{5040} + \frac{11x^6}{360} - \frac{19x^5}{120} - \frac{x^4}{2} - \frac{x^3}{6} + x^2 + x + 2.$$

The polynomials we have created are examples of *Taylor polynomials*, named after the British mathematician Brook Taylor who made important discoveries about such functions. While we created the above Taylor polynomials by solving initial–value problems, it can be shown that Taylor polynomials follow a general pattern that make their formation much more direct. This is described in the following definition.

Definition 42   Taylor Polynomial, Maclaurin Polynomial
Let f be a function whose first n derivatives exist at $x = c$.

1. The **Taylor polynomial of degree n of f at $x = c$** is

$$p_n(x) = f(c) + f'(c)(x - c) + \frac{f''(c)}{2!}(x - c)^2 + \frac{f'''(c)}{3!}(x - c)^3 + \cdots + \frac{f^{(n)}(c)}{n!}(x - c)^n$$

$$= \sum_{k=0}^{n} \frac{f^{(k)}(c)}{k!}(x - c)^k.$$

2. A special case of the Taylor polynomial is the Maclaurin polynomial, where $c = 0$. That is, the **Maclaurin polynomial of degree n of f** is

$$p_n(x) = f(0) + f'(0)x + \frac{f''(0)}{2!}x^2 + \frac{f'''(0)}{3!}x^3 + \cdots + \frac{f^{(n)}(0)}{n!}x^n$$

$$= \sum_{k=0}^{n} \frac{f^{(k)}(0)}{k!}x^k.$$

Note: The summations in this definition use the convention that $x^0 = 1$ even when $x = 0$ and that $f^{(0)} = f$. They also use the definition that $0! = 1$.

Generally, we order the terms of a polynomial to have decreasing degrees, and that is how we began this section. This definition, and the rest of this chapter, reverses this order to reflect the greater importance of the lower degree terms in the polynomials that we will be finding.

Watch the video:
Taylor Polynomial to Approximate a Function, Ex 3 at
`https://youtu.be/UINFWGOErSA`

Notes:

We will practice creating Taylor and Maclaurin polynomials in the following examples.

Example 1   Finding and using Maclaurin polynomials

1. Find the nth Maclaurin polynomial for $f(x) = e^x$.

2. Use $p_5(x)$ to approximate the value of e.

SOLUTION

1. We start with creating a table of the derivatives of e^x evaluated at $x = 0$. In this particular case, this is relatively simple, as shown in Figure 9.21. By the definition of the Maclaurin series, we have

$$p_n(x) = \sum_{k=0}^{n} \frac{f^{(k)}(0)}{k!} x^k = \sum_{k=0}^{n} \frac{1}{k!} x^k.$$

2. Using our answer from part 1, we have

$$p_5(x) = 1 + x + \frac{1}{2}x^2 + \frac{1}{6}x^3 + \frac{1}{24}x^4 + \frac{1}{120}x^5.$$

To approximate the value of e, note that $e = e^1 = f(1) \approx p_5(1)$. It is very straightforward to evaluate $p_5(1)$:

$$p_5(1) = 1 + 1 + \frac{1}{2} + \frac{1}{6} + \frac{1}{24} + \frac{1}{120} = \frac{163}{60} \approx 2.71667.$$

This is an error of about 0.0016, or 0.06%.

A plot of $f(x) = e^x$ and $p_5(x)$ is given in Figure 9.22.

$$
\begin{aligned}
f(x) &= e^x & \Rightarrow & \quad f(0) = 1 \\
f'(x) &= e^x & \Rightarrow & \quad f'(0) = 1 \\
f''(x) &= e^x & \Rightarrow & \quad f''(0) = 1 \\
&\;\vdots & & \quad\quad \vdots \\
f^{(n)}(x) &= e^x & \Rightarrow & \quad f^{(n)}(0) = 1
\end{aligned}
$$

Figure 9.21: The derivatives of $f(x) = e^x$ evaluated at $x = 0$.
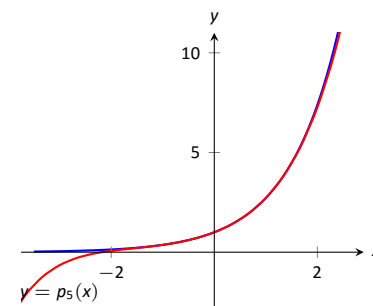
Figure 9.22: A plot of $f(x) = e^x$ and its 5th degree Maclaurin polynomial $p_5(x)$.

Example 2   Finding and using Taylor polynomials

1. Find the nth Taylor polynomial of $y = \ln x$ at $x = 1$.

2. Use $p_6(x)$ to approximate the value of $\ln 1.5$.

3. Use $p_6(x)$ to approximate the value of $\ln 2$.

Notes:

$$f(x) = \ln x \quad \Rightarrow \quad f(1) = 0$$
$$f'(x) = 1/x \quad \Rightarrow \quad f'(1) = 1$$
$$f''(x) = -1/x^2 \quad \Rightarrow \quad f''(1) = -1$$
$$f'''(x) = 2/x^3 \quad \Rightarrow \quad f'''(1) = 2$$
$$f^{(4)}(x) = -6/x^4 \quad \Rightarrow \quad f^{(4)}(1) = -6$$
$$\vdots \qquad\qquad \vdots$$
$$f^{(n)}(x) = \quad \Rightarrow \quad f^{(n)}(1) =$$
$$\frac{(-1)^{n+1}(n-1)!}{x^n} \qquad (-1)^{n+1}(n-1)!$$
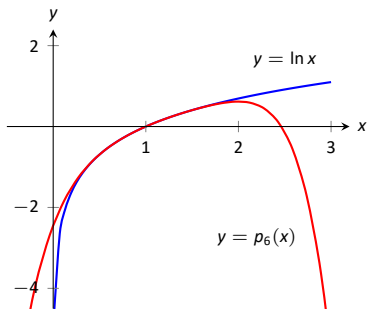
Figure 9.23: Derivatives of $\ln x$ evaluated at $x = 1$.

Figure 9.24: A plot of $y = \ln x$ and its 6th degree Taylor polynomial at $x = 1$.
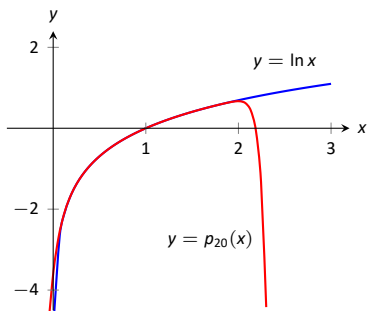
Figure 9.25: A plot of $y = \ln x$ and its 20th degree Taylor polynomial at $x = 1$.

SOLUTION

1. We begin by creating a table of derivatives of $\ln x$ evaluated at $x = 1$. While this is not as straightforward as it was in the previous example, a pattern does emerge, as shown in Figure 9.23.

   Using Definition 42, we have

   $$p_n(x) = \sum_{k=0}^{n} \frac{f^{(k)}(c)}{k!}(x-c)^k = \sum_{k=1}^{n} \frac{1}{k}(x-1)^k.$$

2. We can compute $p_6(x)$ using our work above:

   $$p_6(x) = (x-1) - \frac{1}{2}(x-1)^2 + \frac{1}{3}(x-1)^3 - \frac{1}{4}(x-1)^4 + \frac{1}{5}(x-1)^5 - \frac{1}{6}(x-1)^6.$$

   Since $p_6(x)$ approximates $\ln x$ well near $x = 1$, we approximate $\ln 1.5 \approx p_6(1.5)$:

   $$p_6(1.5) = (1.5 - 1) - \frac{1}{2}(1.5-1)^2 + \frac{1}{3}(1.5-1)^3$$
   $$- \frac{1}{4}(1.5-1)^4 + \frac{1}{5}(1.5-1)^5 - \frac{1}{6}(1.5-1)^6$$
   $$= \frac{259}{640}$$
   $$\approx 0.404688.$$

   This is a good approximation as a calculator shows that $\ln 1.5 \approx 0.4055$. Figure 9.24 plots $y = \ln x$ with $y = p_6(x)$. We can see that $\ln 1.5 \approx p_6(1.5)$.

3. We approximate $\ln 2$ with $p_6(2)$:

   $$p_6(2) = (2-1) - \frac{1}{2}(2-1)^2 + \frac{1}{3}(2-1)^3$$
   $$- \frac{1}{4}(2-1)^4 + \frac{1}{5}(2-1)^5 - \frac{1}{6}(2-1)^6$$
   $$= 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \frac{1}{5} - \frac{1}{6}$$
   $$= \frac{37}{60}$$
   $$\approx 0.616667.$$

   This approximation is not terribly impressive: a hand held calculator shows

Notes:

that $\ln 2 \approx 0.693147$. The graph in Figure 9.24 shows that $p_6(x)$ provides less accurate approximations of $\ln x$ as x gets close to 0 or 2.

Surprisingly enough, even the 20th degree Taylor polynomial fails to approximate $\ln x$ for $x > 2$, as shown in Figure 9.25. We'll soon discuss why this is.

Taylor polynomials are used to approximate functions $f(x)$ in mainly two situations:

1. When $f(x)$ is known, but perhaps "hard" to compute directly. For instance, we can define $y = \cos x$ as either the ratio of sides of a right triangle ("adjacent over hypotenuse") or with the unit circle. However, neither of these provides a convenient way of computing $\cos 2$. A Taylor polynomial of sufficiently high degree can provide a reasonable method of computing such values using only operations usually hard–wired into a computer ($+$, $-$, $\times$ and $\div$).

2. When $f(x)$ is not known, but information about its derivatives is known. This occurs more often than one might think, especially in the study of differential equations.

In both situations, a critical piece of information to have is "How good is my approximation?" If we use a Taylor polynomial to compute $\cos 2$, how do we know how accurate the approximation is?

We had the same problem when studying Numerical Integration. Theorem 57 provided bounds on the error when using, say, Simpson's Rule to approximate a definite integral. These bounds allowed us to determine that, for example, using 10 subintervals provided an approximation within $\pm.01$ of the exact value. The following theorem gives similar bounds for Taylor (and hence Maclaurin) polynomials.

Note: Even though Taylor polynomials *could* be used in calculators and computers to calculate values of trigonometric functions, in practice they generally aren't. Other more efficient and accurate methods have been developed, such as the CORDIC algorithm.

Notes:

Theorem 81   Taylor's Theorem

1. Let f be a function whose $(n+1)^{\text{th}}$ derivative exists on an interval I and let c be in I. Then, for each x in I, there exists z_x between x and c such that

$$f(x) = \sum_{k=0}^{n} \frac{f^{(k)}(x-c)}{k!} + R_n(x),$$

where $R_n(x) = \dfrac{f^{(n+1)}(z_x)}{(n+1)!}(x-c)^{n+1}$.

2. $|R_n(x)| \leq \dfrac{\max \left|f^{(n+1)}(z)\right|}{(n+1)!} |x-c|^{n+1}$.

The first part of Taylor's Theorem states that $f(x) = p_n(x) + R_n(x)$, where $p_n(x)$ is the nth order Taylor polynomial and $R_n(x)$ is the remainder, or error, in the Taylor approximation. The second part gives bounds on how big that error can be. If the $(n+1)^{\text{th}}$ derivative is large, the error may be large; if x is far from c, the error may also be large. However, the $(n+1)!$ term in the denominator tends to ensure that the error gets smaller as n increases.

The following example computes error estimates for the approximations of ln 1.5 and ln 2 made in Example 2.

Example 3   Finding error bounds of a Taylor polynomial
Use Theorem 81 to find error bounds when approximating ln 1.5 and ln 2 with $p_6(x)$, the Taylor polynomial of degree 6 of $f(x) = \ln x$ at $x = 1$, as calculated in Example 2.

SOLUTION

1. We start with the approximation of ln 1.5 with $p_6(1.5)$. The theorem references an open interval I that contains both x and c. The smaller the interval we use the better; it will give us a more accurate (and smaller) approximation of the error. We let $I = (0.9, 1.6)$, as this interval contains both $c = 1$ and $x = 1.5$.

   The theorem references $\max \left|f^{(n+1)}(z)\right|$. In our situation, this is asking "How big can the 7th derivative of $y = \ln x$ be on the interval $(0.9, 1.6)$?" The seventh derivative is $y = -6!/x^7$. The largest value it attains on I is

Notes:

about 1506. Thus we can bound the error as:

$$|R_6(1.5)| \leq \frac{\max \left|f^{(7)}(z)\right|}{7!} |1.5 - 1|^7$$

$$\leq \frac{1506}{5040} \cdot \frac{1}{2^7}$$

$$\approx 0.0023.$$

We computed $p_6(1.5) = 0.404688$; using a calculator, we find $\ln 1.5 \approx 0.405465$, so the actual error is about 0.000778 (or 0.2%), which is less than our bound of 0.0023. This affirms Taylor's Theorem; the theorem states that our approximation would be within about 2 thousandths of the actual value, whereas the approximation was actually closer.

2. We again find an interval I that contains both $c = 1$ and $x = 2$; we choose $I = (0.9, 2.1)$. The maximum value of the seventh derivative of f on this interval is again about 1506 (as the largest values come near $x = 0.9$). Thus

$$|R_6(2)| \leq \frac{\max \left|f^{(7)}(z)\right|}{7!} |2 - 1|^7$$

$$\leq \frac{1506}{5040} \cdot 1^7$$

$$\approx 0.30.$$

This bound is not as nearly as good as before. Using the degree 6 Taylor polynomial at $x = 1$ will bring us within 0.3 of the correct answer. As $p_6(2) \approx 0.61667$, our error estimate guarantees that the actual value of $\ln 2$ is somewhere between 0.31667 and 0.91667. These bounds are not particularly useful.

In reality, our approximation was only off by about 0.07 (or 11%). However, we are approximating ostensibly because we do not know the real answer. In order to be assured that we have a good approximation, we would have to resort to using a polynomial of higher degree.

We practice again. This time, we use Taylor's theorem to find n that guarantees our approximation is within a certain amount.

Example 4   Finding sufficiently accurate Taylor polynomials
Find n such that the nth Taylor polynomial of $f(x) = \cos x$ at $x = 0$ approximates $\cos 2$ to within 0.001 of the actual answer. What is $p_n(2)$?

Notes:

Solution Following Taylor's theorem, we need bounds on the size of the derivatives of $f(x) = \cos x$. In the case of this trigonometric function, this is easy. All derivatives of cosine are $\pm \sin x$ or $\pm \cos x$. In all cases, these functions are never greater than 1 in absolute value. We want the error to be less than 0.001. To find the appropriate n, consider the following inequalities:

$$\frac{\max \left| f^{(n+1)}(z) \right|}{(n+1)!} |2 - 0|^{n+1} \leq 0.001$$

$$\frac{1}{(n+1)!} \cdot 2^{n+1} \leq 0.001$$

We find an n that satisfies this last inequality with trial–and–error. When $n = 8$, we have $\dfrac{2^{8+1}}{(8+1)!} \approx 0.0014$; when $n = 9$, we have $\dfrac{2^{9+1}}{(9+1)!} \approx 0.000282 < 0.001$. Thus we want to approximate $\cos 2$ with $p_9(2)$.

We now set out to compute $p_9(x)$. We again need a table of the derivatives of $f(x) = \cos x$ evaluated at $x = 0$. A table of these values is given in Figure 9.26. Notice how the derivatives, evaluated at $x = 0$, follow a certain pattern. All the odd powers of x in the Taylor polynomial will disappear as their coefficient is 0. While our error bounds state that we need $p_9(x)$, our work shows that this will be the same as $p_8(x)$.

Since we are forming our polynomial at $x = 0$, we are creating a Maclaurin polynomial, and:

$$p_8(x) = \sum_{k=0}^{8} \frac{f^{(k)}(0)}{k!} x^k = 1 - \frac{1}{2!}x^2 + \frac{1}{4!}x^4 - \frac{1}{6!}x^6 + \frac{1}{8!}x^8$$

We finally approximate $\cos 2$:

$$\cos 2 \approx p_8(2) = -\frac{131}{315} \approx -0.41587.$$

Our error bound guarantee that this approximation is within 0.001 of the correct answer. Technology shows us that our approximation is actually within about 0.0003 (or 0.07%) of the correct answer.

Figure 9.27 shows a graph of $y = p_8(x)$ and $y = \cos x$. Note how well the two functions agree on about $(-\pi, \pi)$.

$$
\begin{aligned}
f(x) &= \cos x &\Rightarrow&\quad f(0) = 1 \\
f'(x) &= -\sin x &\Rightarrow&\quad f'(0) = 0 \\
f''(x) &= -\cos x &\Rightarrow&\quad f''(0) = -1 \\
f'''(x) &= \sin x &\Rightarrow&\quad f'''(0) = 0 \\
f^{(4)}(x) &= \cos x &\Rightarrow&\quad f^{(4)}(0) = 1 \\
f^{(5)}(x) &= -\sin x &\Rightarrow&\quad f^{(5)}(0) = 0 \\
f^{(6)}(x) &= -\cos x &\Rightarrow&\quad f^{(6)}(0) = -1 \\
f^{(7)}(x) &= \sin x &\Rightarrow&\quad f^{(7)}(0) = 0 \\
f^{(8)}(x) &= \cos x &\Rightarrow&\quad f^{(8)}(0) = 1 \\
f^{(9)}(x) &= -\sin x &\Rightarrow&\quad f^{(9)}(0) = 0
\end{aligned}
$$

Figure 9.26: A table of the derivatives of $f(x) = \cos x$ evaluated at $x = 0$.
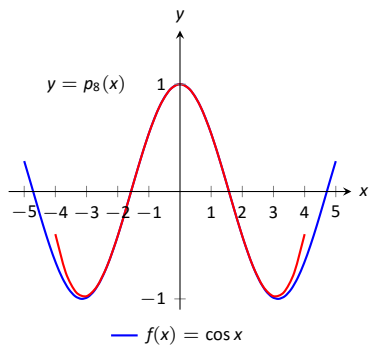
Figure 9.27: A graph of $f(x) = \cos x$ and its degree 8 Maclaurin polynomial.

Example 5 Finding and using Taylor polynomials

1. Find the degree 4 Taylor polynomial, $p_4(x)$, for $f(x) = \sqrt{x}$ at $x = 4$.

2. Use $p_4(x)$ to approximate $\sqrt{3}$.

3. Find bounds on the error when approximating $\sqrt{3}$ with $p_4(3)$.

Notes:

SOLUTION

1. We begin by evaluating the derivatives of f at $x = 4$. This is done in Figure 9.28. These values allow us to form the Taylor polynomial $p_4(x)$:

$$p_4(x) = 2 + \frac{1}{4}(x-4) + \frac{-1/32}{2!}(x-4)^2 + \frac{3/256}{3!}(x-4)^3 + \frac{-15/2048}{4!}(x-4)^4.$$

2. As $p_4(x) \approx \sqrt{x}$ near $x = 4$, we approximate $\sqrt{3}$ with $p_4(3) = 1.73212$.

3. To find a bound on the error, we need an open interval that contains $x = 3$ and $x = 4$. We set $I = (2.9, 4.1)$. The largest value the fifth derivative of $f(x) = \sqrt{x}$ takes on this interval is near $x = 2.9$, at about 0.0273. Thus

$$|R_4(3)| \leq \frac{0.0273}{5!}|3-4|^5 \approx 0.00023.$$

This shows our approximation is accurate to at least the first 2 places after the decimal. It turns out that our approximation has an error of 0.00007, or 0.004%. A graph of $f(x) = \sqrt{x}$ and $p_4(x)$ is given in Figure 9.29. Note how the two functions are nearly indistinguishable on $(2, 7)$.

$$f(x) = \sqrt{x} \qquad \Rightarrow \qquad f(4) = 2$$

$$f'(x) = \frac{1}{2\sqrt{x}} \qquad \Rightarrow \qquad f'(4) = \frac{1}{4}$$

$$f''(x) = \frac{-1}{4x^{3/2}} \qquad \Rightarrow \qquad f''(4) = \frac{-1}{32}$$

$$f'''(x) = \frac{3}{8x^{5/2}} \qquad \Rightarrow \qquad f'''(4) = \frac{3}{256}$$

$$f^{(4)}(x) = \frac{-15}{16x^{7/2}} \qquad \Rightarrow \qquad f^{(4)}(4) = \frac{-15}{2048}$$

Figure 9.28: A table of the derivatives of $f(x) = \sqrt{x}$ evaluated at $x = 4$.
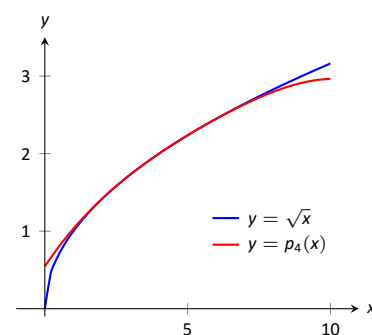
Figure 9.29: A graph of $f(x) = \sqrt{x}$ and its degree 4 Taylor polynomial at $x = 4$.

Most of this chapter has been devoted to the study of infinite series. This section has stepped aside from this study, focusing instead on finite summation of terms. In the next section, we will combine power series and Taylor polynomials into **Taylor Series**, where we represent a function with an infinite series.

Notes:

Exercises 9.9

Terms and Concepts

1. What is the difference between a Taylor polynomial and a Maclaurin polynomial?

2. T/F: In general, $p_n(x)$ approximates $f(x)$ better and better as n gets larger.

3. For some function $f(x)$, the Maclaurin polynomial of degree 4 is $p_4(x) = 6 + 3x - 4x^2 + 5x^3 - 7x^4$. What is $p_2(x)$?

4. For some function $f(x)$, the Maclaurin polynomial of degree 4 is $p_4(x) = 6 + 3x - 4x^2 + 5x^3 - 7x^4$. What is $f'''(0)$?

Problems

In Exercises 5–12, find the Maclaurin polynomial of degree n for the given function.

5. $f(x) = e^{-x}, \quad n = 3$

6. $f(x) = \sin x, \quad n = 8$

7. $f(x) = x \cdot e^x, \quad n = 5$

8. $f(x) = \tan x, \quad n = 6$

9. $f(x) = e^{2x}, \quad n = 4$

10. $f(x) = \dfrac{1}{1-x}, \quad n = 4$

11. $f(x) = \dfrac{1}{1+x}, \quad n = 4$

12. $f(x) = \dfrac{1}{1+x}, \quad n = 7$

In Exercises 13–20, find the Taylor polynomial of degree n, at $x = c$, for the given function.

13. $f(x) = \sqrt{x}, \quad n = 4, \quad c = 1$

14. $f(x) = \ln(x+1), \quad n = 4, \quad c = 1$

15. $f(x) = \cos x, \quad n = 6, \quad c = \pi/4$

16. $f(x) = \sin x, \quad n = 5, \quad c = \pi/6$

17. $f(x) = \dfrac{1}{x}, \quad n = 5, \quad c = 2$

18. $f(x) = \dfrac{1}{x^2}, \quad n = 8, \quad c = 1$

19. $f(x) = \dfrac{1}{x^2+1}, \quad n = 3, \quad c = -1$

20. $f(x) = x^2 \cos x, \quad n = 2, \quad c = \pi$

In Exercises 21–24, approximate the function value with the indicated Taylor polynomial and give approximate bounds on the error.

21. Approximate $\sin 0.1$ with the Maclaurin polynomial of degree 3.

22. Approximate $\cos 1$ with the Maclaurin polynomial of degree 4.

23. Approximate $\sqrt{10}$ with the Taylor polynomial of degree 2 centered at $x = 9$.

24. Approximate $\ln 1.5$ with the Taylor polynomial of degree 3 centered at $x = 1$.

Exercises 25–28 ask for an n to be found such that $p_n(x)$ approximates $f(x)$ within a certain bound of accuracy.

25. Find n such that the Maclaurin polynomial of degree n of $f(x) = e^x$ approximates e within 0.0001 of the actual value.

26. Find n such that the Taylor polynomial of degree n of $f(x) = \sqrt{x}$, centered at $x = 4$, approximates $\sqrt{3}$ within 0.0001 of the actual value.

27. Find n such that the Maclaurin polynomial of degree n of $f(x) = \cos x$ approximates $\cos \pi/3$ within 0.0001 of the actual value.

28. Find n such that the Maclaurin polynomial of degree n of $f(x) = \sin x$ approximates $\cos \pi$ within 0.0001 of the actual value.

In Exercises 29–33, find the nth term of the indicated Taylor polynomial.

29. Find a formula for the nth term of the Maclaurin polynomial for $f(x) = e^x$.

30. Find a formula for the nth term of the Maclaurin polynomial for $f(x) = \cos x$.

31. Find a formula for the nth term of the Maclaurin polynomial for $f(x) = \dfrac{1}{1-x}$.

32. Find a formula for the nth term of the Maclaurin polynomial for $f(x) = \dfrac{1}{1+x}$.

33. Find a formula for the nth term of the Taylor polynomial for $f(x) = \ln x$.

In Exercises 34–36, approximate the solution to the given differential equation with a degree 4 Maclaurin polynomial.

34. $y' = y, \qquad y(0) = 1$

35. $y' = 5y, \qquad y(0) = 3$

36. $y' = \dfrac{2}{y}, \qquad y(0) = 1$

9.10   Taylor Series

In Section 9.8, we showed how certain functions can be represented by a power series function. In Section 9.9, we showed how we can approximate functions with polynomials, given that enough derivative information is available. In this section we combine these concepts: if a function $f(x)$ is infinitely differentiable, we show how to represent it with a power series function.

Definition 43   Taylor and Maclaurin Series

Let $f(x)$ have derivatives of all orders at $x = c$.

1. The **Taylor Series of $f(x)$, centered at c** is

$$\sum_{n=0}^{\infty} \frac{f^{(n)}(c)}{n!}(x - c)^n.$$

2. Setting $c = 0$ gives the **Maclaurin Series of $f(x)$**:

$$\sum_{n=0}^{\infty} \frac{f^{(n)}(0)}{n!}x^n.$$

Watch the video:
Taylor and Maclaurin Series — Example 2 at
`https://youtu.be/0s80tXFBLkY`

The difference between a Taylor polynomial and a Taylor series is the former is a polynomial, containing only a finite number of terms, whereas the latter is a series, a summation of an infinite set of terms. When creating the Taylor polynomial of degree n for a function $f(x)$ at $x = c$, we needed to evaluate f, and the first n derivatives of f, at $x = c$. When creating the Taylor series of f, we need to find a pattern that describes the nth derivative of f at $x = c$. We demonstrate this in the next two examples.

Example 1   The Maclaurin series of $f(x) = \cos x$
Find the Maclaurin series of $f(x) = \cos x$.

Notes:

$$f(x) = \cos x \quad \Rightarrow \quad f(0) = 1$$
$$f'(x) = -\sin x \quad \Rightarrow \quad f'(0) = 0$$
$$f''(x) = -\cos x \quad \Rightarrow \quad f''(0) = -1$$
$$f'''(x) = \sin x \quad \Rightarrow \quad f'''(0) = 0$$
$$f^{(4)}(x) = \cos x \quad \Rightarrow \quad f^{(4)}(0) = 1$$
$$f^{(5)}(x) = -\sin x \quad \Rightarrow \quad f^{(5)}(0) = 0$$
$$f^{(6)}(x) = -\cos x \quad \Rightarrow \quad f^{(6)}(0) = -1$$
$$f^{(7)}(x) = \sin x \quad \Rightarrow \quad f^{(7)}(0) = 0$$
$$f^{(8)}(x) = \cos x \quad \Rightarrow \quad f^{(8)}(0) = 1$$
$$f^{(9)}(x) = -\sin x \quad \Rightarrow \quad f^{(9)}(0) = 0$$

Figure 9.30: A table of the derivatives of $f(x) = \cos x$ evaluated at $x = 0$.

SOLUTION   In Example 9.9.4 we found the 8th degree Maclaurin polynomial of $\cos x$. In doing so, we created the table shown in Figure 9.30. Notice how $f^{(n)}(0) = 0$ when n is odd, $f^{(n)}(0) = 1$ when n is divisible by 4, and $f^{(n)}(0) = -1$ when n is even but not divisible by 4. Thus the Maclaurin series of $\cos x$ is

$$1 - \frac{x^2}{2} + \frac{x^4}{4!} - \frac{x^6}{6!} + \frac{x^8}{8!} - \cdots$$

We can go further and write this as a summation. Since we only need the terms where the power of x is even, we write the power series in terms of x^{2n}:

$$\sum_{n=0}^{\infty} (-1)^n \frac{x^{2n}}{(2n)!}.$$

Example 2   **The Taylor series of $f(x) = \ln x$ at $x = 1$**
Find the Taylor series of $f(x) = \ln x$ centered at $x = 1$.

SOLUTION   Figure 9.31 shows the nth derivative of $\ln x$ evaluated at $x = 1$ for $n = 0, \dots, 5$, along with an expression for the nth term:

$$f^{(n)}(1) = (-1)^{n+1}(n-1)! \quad \text{for } n \geq 1.$$

Remember that this is what distinguishes Taylor series from Taylor polynomials; we are very interested in finding a pattern for the nth term, not just finding a finite set of coefficients for a polynomial. Since $f(1) = \ln 1 = 0$, we skip the first term and start the summation with $n = 1$, giving the Taylor series for $\ln x$, centered at $x = 1$, as

$$f(x) = \ln x \quad \Rightarrow \quad f(1) = 0$$
$$f'(x) = 1/x \quad \Rightarrow \quad f'(1) = 1$$
$$f''(x) = -1/x^2 \quad \Rightarrow \quad f''(1) = -1$$
$$f'''(x) = 2/x^3 \quad \Rightarrow \quad f'''(1) = 2$$
$$f^{(4)}(x) = -6/x^4 \quad \Rightarrow \quad f^{(4)}(1) = -6$$
$$f^{(5)}(x) = 24/x^5 \quad \Rightarrow \quad f^{(5)}(1) = 24$$
$$\vdots \qquad\qquad\qquad \vdots$$
$$f^{(n)}(x) = \quad \Rightarrow \quad f^{(n)}(1) =$$
$$\frac{(-1)^{n+1}(n-1)!}{x^n} \qquad (-1)^{n+1}(n-1)!$$

Figure 9.31: Derivatives of $\ln x$ evaluated at $x = 1$.

$$\sum_{n=1}^{\infty} (-1)^{n+1}(n-1)! \frac{1}{n!}(x-1)^n = \sum_{n=1}^{\infty} (-1)^{n+1} \frac{(x-1)^n}{n}.$$

It is important to note that Definition 43 defines a Taylor series given a function $f(x)$; however, we *cannot* yet state that $f(x)$ *is equal* to its Taylor series. We will find that "most of the time" they are equal, but we need to consider the conditions that allow us to conclude this.

Theorem 81 states that the error between a function $f(x)$ and its nth–degree Taylor polynomial $p_n(x)$ is $R_n(x)$, where

$$|R_n(x)| \leq \frac{\max\left|f^{(n+1)}(z)\right|}{(n+1)!} |x - c|^{n+1}.$$

If $R_n(x)$ goes to 0 for each x in an interval I as n approaches infinity, we conclude that the function is equal to its Taylor series expansion.

Notes:

> **Theorem 82   Function and Taylor Series Equality**
>
> Let $f(x)$ have derivatives of all orders at $x = c$, let $R_n(x)$ be as stated in Theorem 81, and let I be an interval on which the Taylor series of $f(x)$ converges. If $\lim_{n\to\infty} R_n(x) = 0$ for all x in I, then
>
> $$f(x) = \sum_{n=0}^{\infty} \frac{f^{(n)}(c)}{n!}(x - c)^n \text{ on } I.$$

We demonstrate the use of this theorem in an example.

Example 3   Establishing equality of a function and its Taylor series
Show that $f(x) = \cos x$ is equal to its Maclaurin series, as found in Example 1, for all x.

SOLUTION   Given a value x, the magnitude of the error term $R_n(x)$ is bounded by

$$|R_n(x)| \leq \frac{\max\left|f^{(n+1)}(z)\right|}{(n+1)!} |x|^{n+1}.$$

Since all derivatives of $\cos x$ are $\pm \sin x$ or $\pm \cos x$, whose magnitudes are bounded by 1, we can state

$$|R_n(x)| \leq \frac{1}{(n+1)!} |x|^{n+1}$$

which implies

$$-\frac{|x|^{n+1}}{(n+1)!} \leq R_n(x) \leq \frac{|x|^{n+1}}{(n+1)!}. \tag{9.10}$$

For any x, $\lim_{n\to\infty} \frac{x^{n+1}}{(n+1)!} = 0$. Applying the Squeeze Theorem to Equation (9.10), we conclude that $\lim_{n\to\infty} R_n(x) = 0$ for all x, and hence

$$\cos x = \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n}}{(2n)!} \quad \text{for all } x.$$

It is natural to assume that a function is equal to its Taylor series on the series' interval of convergence, but this is not the case. In order to properly establish equality, one must use Theorem 82. This is a bit disappointing, as we developed beautiful techniques for determining the interval of convergence of a power series, and proving that $R_n(x) \to 0$ can be cumbersome as it deals with high order derivatives of the function.

Notes:

There is good news. A function $f(x)$ that is equal to its Taylor series, centered at any point the domain of $f(x)$, is said to be an **analytic function**, and most, if not all, functions that we encounter within this course are analytic functions. Generally speaking, any function that one creates with elementary functions (polynomials, exponentials, trigonometric functions, etc.) that is not piecewise defined is probably analytic. For most functions, we assume the function is equal to its Taylor series on the series' interval of convergence and only use Theorem 82 when we suspect something may not work as expected.

We develop the Taylor series for one more important function, then give a table of the Taylor series for a number of common functions.

Example 4   The Binomial Series
Find the Maclaurin series of $f(x) = (1 + x)^k$, $k \neq 0$.

SOLUTION   When k is a positive integer, the Maclaurin series is finite. For instance, when $k = 4$, we have

$$f(x) = (1 + x)^4 = 1 + 4x + 6x^2 + 4x^3 + x^4.$$

The coefficients of x when k is a positive integer are known as the *binomial coefficients*, giving the series we are developing its name.

When $k = 1/2$, we have $f(x) = \sqrt{1 + x}$. Knowing a series representation of this function would give a useful way of approximating $\sqrt{1.3}$, for instance.

To develop the Maclaurin series for $f(x) = (1 + x)^k$ for any value of $k \neq 0$, we consider the derivatives of f evaluated at $x = 0$:

$$
\begin{aligned}
f(x) &= (1 + x)^k & f(0) &= 1 \\
f'(x) &= k(1 + x)^{k-1} & f'(0) &= k \\
f''(x) &= k(k - 1)(1 + x)^{k-2} & f''(0) &= k(k - 1) \\
f'''(x) &= k(k - 1)(k - 2)(1 + x)^{k-3} & f'''(0) &= k(k - 1)(k - 2) \\
&\ \ \vdots & &\ \ \vdots \\
f^{(n)}(x) &= k(k-1)\cdots\big(k - (n-1)\big)(1 + x)^{k-n} & f^{(n)}(0) &= k(k-1)\cdots\big(k - (n-1)\big)
\end{aligned}
$$

Thus the Maclaurin series for $f(x) = (1 + x)^k$ is

$$1 + k + \frac{k(k - 1)}{2!} + \frac{k(k - 1)(k - 2)}{3!} + \cdots + \frac{k(k-1)\cdots\big(k - (n-1)\big)}{n!} + \cdots$$

It is important to determine the interval of convergence of this series. With

$$a_n = \frac{k(k-1)\cdots\big(k - (n-1)\big)}{n!}x^n,$$

Notes:

we apply the Ratio Test:

$$\lim_{n\to\infty} \frac{|a_{n+1}|}{|a_n|} = \lim_{n\to\infty} \frac{\left|\frac{k(k-1)\cdots(k-n)}{(n+1)!}x^{n+1}\right|}{\left|\frac{k(k-1)\cdots\left(k-(n-1)\right)}{n!}x^n\right|}$$

$$= \lim_{n\to\infty} \left|\frac{k-n}{n}x\right|$$

$$= |x|.$$

The series converges absolutely when the limit of the Ratio Test is less than 1; therefore, we have absolute convergence when $|x| < 1$.

While outside the scope of this text, the interval of convergence depends on the value of k. When $k > 0$, the interval of convergence is $[-1, 1]$. When $-1 < k < 0$, the interval of convergence is $[-1, 1)$. If $k \leq -1$, the interval of convergence is $(-1, 1)$.

We learned that Taylor polynomials offer a way of approximating a "difficult to compute" function with a polynomial. Taylor series offer a way of exactly representing a function with a series. One probably can see the use of a good approximation; is there any use of representing a function exactly as a series?

While we should not overlook the mathematical beauty of Taylor series (which is reason enough to study them), there are practical uses as well. They provide a valuable tool for solving a variety of problems, including problems relating to integration and differential equations.

In Key Idea 33 (on the following page) we give a table of the Maclaurin series of a number of common functions. We then give a theorem about the "algebra of power series," that is, how we can combine power series to create power series of new functions. This allows us to find the Taylor series of functions like $f(x) = e^x \cos x$ by knowing the Taylor series of e^x and $\cos x$.

Before we investigate combining functions, consider the Taylor series for the arctangent function (see Key Idea 33). Knowing that $\tan^{-1}(1) = \pi/4$, we can use this series to approximate the value of π:

$$\frac{\pi}{4} = \tan^{-1}(1) = 1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} + \frac{1}{9} - \cdots$$

$$\pi = 4\left(1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} + \frac{1}{9} - \cdots\right)$$

Unfortunately, this particular expansion of π converges very slowly. The first 100 terms approximate π as 3.13159, which is not particularly good.

Notes:

Key Idea 33   Important Maclaurin Series Expansions

Function and Series	First Few Terms	Interval of Convergence
$e^x = \displaystyle\sum_{n=0}^{\infty} \frac{x^n}{n!}$	$1 + x + \dfrac{x^2}{2!} + \dfrac{x^3}{3!} + \cdots$	$(-\infty, \infty)$
$\sin x = \displaystyle\sum_{n=0}^{\infty} (-1)^n \frac{x^{2n+1}}{(2n+1)!}$	$x - \dfrac{x^3}{3!} + \dfrac{x^5}{5!} - \dfrac{x^7}{7!} + \cdots$	$(-\infty, \infty)$
$\cos x = \displaystyle\sum_{n=0}^{\infty} (-1)^n \frac{x^{2n}}{(2n)!}$	$1 - \dfrac{x^2}{2!} + \dfrac{x^4}{4!} - \dfrac{x^6}{6!} + \cdots$	$(-\infty, \infty)$
$\ln(x + 1) = \displaystyle\sum_{n=1}^{\infty} (-1)^{n+1} \frac{x^n}{n}$	$x - \dfrac{x^2}{2} + \dfrac{x^3}{3} - \cdots$	$(-1, 1]$
$\dfrac{1}{1 - x} = \displaystyle\sum_{n=0}^{\infty} x^n$	$1 + x + x^2 + x^3 + \cdots$	$(-1, 1)$
$(1 + x)^k = \displaystyle\sum_{n=0}^{\infty} \frac{k(k - 1) \cdots (k - (n - 1))}{n!} x^n$	$1 + kx + \dfrac{k(k - 1)}{2!} x^2 + \cdots$	$\begin{cases} (-1, 1) & k \le -1 \\ [-1, 1) & -1 < k < 0 \\ [-1, 1] & 0 < k \end{cases}$
$\tan^{-1} x = \displaystyle\sum_{n=0}^{\infty} (-1)^n \frac{x^{2n+1}}{2n + 1}$	$x - \dfrac{x^3}{3} + \dfrac{x^5}{5} - \dfrac{x^7}{7} + \cdots$	$[-1, 1]$

Notes:

Theorem 83   **Algebra of Power Series**

Let $f(x) = \sum\limits_{n=0}^{\infty} a_n x^n$ and $g(x) = \sum\limits_{n=0}^{\infty} b_n x^n$ converge absolutely for $|x| <$
R, and let $h(x)$ be continuous.

1. $f(x) \pm g(x) = \sum\limits_{n=0}^{\infty} (a_n \pm b_n) x^n$ for $|x| < R$.

2. $f(x)g(x) = \left(\sum\limits_{n=0}^{\infty} a_n x^n \right) \left(\sum\limits_{n=0}^{\infty} b_n x^n \right) =$

      $\sum\limits_{n=0}^{\infty} (a_0 b_n + a_1 b_{n-1} + \cdots + a_n b_0) x^n$ for $|x| < R$.

3. $f\big(h(x)\big) = \sum\limits_{n=0}^{\infty} a_n \big(h(x)\big)^n$ for $|h(x)| < R$.

Example 5   **Combining Taylor series**
Write out the first 3 terms of the Maclaurin Series for $f(x) = e^x \cos x$ using Key
Idea 33 and Theorem 83.

    SOLUTION   Key Idea 33 informs us that

$$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots \quad \text{and} \quad \cos x = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} + \cdots .$$

Applying Theorem 83, we find that

$$e^x \cos x = \left(1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots \right) \left(1 - \frac{x^2}{2!} + \frac{x^4}{4!} + \cdots \right).$$

Distribute the right hand expression across the left:

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

Notes:

553

Distribute again and collect like terms.

$$= 1 + x - \frac{x^3}{3} - \frac{x^4}{6} - \frac{x^5}{30} + \frac{x^7}{630} + \cdots$$

While this process is a bit tedious, it is much faster than evaluating all the necessary derivatives of $e^x \cos x$ and computing the Taylor series directly.

Because the series for e^x and $\cos x$ both converge on $(-\infty, \infty)$, so does the series expansion for $e^x \cos x$.

Example 6   **Creating new Taylor series**
Use Theorem 83 to create the Taylor series for $y = \sin(x^2)$ centered at $x = 0$ and a series for $y = \ln(\sqrt{x})$ centered at $c = 1$.

SOLUTION   Given that

$$\sin x = \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n+1}}{(2n+1)!} = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \cdots,$$

we simply substitute x^2 for x in the series, giving

$$\sin(x^2) = \sum_{n=0}^{\infty} (-1)^n \frac{(x^2)^{2n+1}}{(2n+1)!} = \sum_{n=0}^{\infty} (-1)^n \frac{x^{4n+2}}{(2n+1)!} = x^2 - \frac{x^6}{3!} + \frac{x^{10}}{5!} - \frac{x^{14}}{7!} \cdots.$$

Since the Taylor series for $\sin x$ has an infinite radius of convergence, so does the Taylor series for $\sin(x^2)$.

The Taylor expansion for $\ln x$ given in Key Idea 33 is centered at $x = 1$, so we will center the series for $\ln(\sqrt{x})$ at $x = 1$ as well. With

$$\ln x = \sum_{n=1}^{\infty} (-1)^{n+1} \frac{(x-1)^n}{n} = (x-1) - \frac{(x-1)^2}{2} + \frac{(x-1)^3}{3} - \cdots,$$

we substitute $\sqrt{x}$ for x to obtain

$$\ln(\sqrt{x}) = \sum_{n=1}^{\infty} (-1)^{n+1} \frac{(\sqrt{x}-1)^n}{n} = (\sqrt{x}-1) - \frac{(\sqrt{x}-1)^2}{2} + \frac{(\sqrt{x}-1)^3}{3} - \cdots.$$

While this is not strictly a power series because of the $\sqrt{x}$, it is a series that allows us to study the function $\ln(\sqrt{x})$. Since the interval of convergence of $\ln x$ is $(0, 2]$, and the range of $\sqrt{x}$ on $(0, 4]$ is $(0, 2]$, the interval of convergence of this series expansion of $\ln(\sqrt{x})$ is $(0, 4]$.

Note: In Example 6, one could create a series for $\ln(\sqrt{x})$ by simply recognizing that $\ln(\sqrt{x}) = \ln(x^{1/2}) = 1/2 \ln x$, and hence multiplying the Taylor series for $\ln x$ by $1/2$. This example was chosen to demonstrate other aspects of series, such as the fact that the interval of convergence changes.

Notes:

Example 7   **Using Taylor series to evaluate definite integrals**

Use the Taylor series of e^{-x^2} to evaluate $\displaystyle\int_0^1 e^{-x^2}\, dx$.

     SOLUTION   We learned, when studying Numerical Integration, that e^{-x^2} does not have an antiderivative expressible in terms of elementary functions. This means any definite integral of this function must have its value approximated, and not computed exactly.

     We can quickly write out the Taylor series for e^{-x^2} using the Taylor series of e^x:

$$e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!} = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots$$

and so

$$
\begin{aligned}
e^{-x^2} &= \sum_{n=0}^{\infty} \frac{(-x^2)^n}{n!} \\
&= \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n}}{n!} \\
&= 1 - x^2 + \frac{x^4}{2!} - \frac{x^6}{3!} + \cdots .
\end{aligned}
$$

We use Theorem 80 to integrate:

$$\int e^{-x^2}\, dx = C + x - \frac{x^3}{3} + \frac{x^5}{5\cdot 2!} - \frac{x^7}{7\cdot 3!} + \cdots + (-1)^n \frac{x^{2n+1}}{(2n+1)n!} + \cdots$$

This *is* the antiderivative of e^{-x^2}; while we can write it out as a series, we cannot write it out in terms of elementary functions. We can evaluate the definite integral $\displaystyle\int_0^1 e^{-x^2}\, dx$ using this antiderivative; substituting 1 and 0 for x and subtracting gives

$$\int_0^1 e^{-x^2}\, dx = 1 - \frac{1}{3} + \frac{1}{5\cdot 2!} - \frac{1}{7\cdot 3!} + \frac{1}{9\cdot 4!} - \cdots .$$

Summing the 5 terms shown above give the approximation of 0.74749. Since this is an alternating series, we can use the Alternating Series Approximation Theorem, (Theorem 75), to determine how accurate this approximation is. The next term of the series is $1/(11\cdot 5!) \approx 0.00075758$. Thus we know our approximation is within 0.00075758 of the actual value of the integral. This is arguably much less work than using Simpson's Rule to approximate the value of the integral.

Notes:

Another advantage to using Taylor series instead of Simpson's Rule is for making subsequent approximations. We found in Example 8.7.7 that the error in using Simpson's Rule for $\int_0^1 e^{-x^2}\, dx$ with four intervals was 0.00026. If we wanted to decrease that error, we would need to use more intervals, essentially starting the problem over. Using a Taylor series, if we wanted a more accurate approximation, we can just subtract the next term $1/(11 \cdot 5!)$ to get an approximation of 0.7467, with an error of at most $1/(13 \cdot 6!) \approx 0.0001$.

Finding a pattern in the coefficients that match the series expansion of a known function, such as those shown in Key Idea 33, can be difficult. What if the coefficients are given in their reduced form; how could we still recover the function?

Suppose that all we know is that

$$a_0 = 1, \quad a_1 = 2, \quad a_2 = 2, \quad a_3 = \frac{4}{3}, \quad a_4 = \frac{2}{3}.$$

Definition 43 states that each term of the Taylor expansion of a function includes an $n!$. This allows us to say that

$$a_2 = 2 = \frac{b_2}{2!}, \quad a_3 = \frac{4}{3} = \frac{b_3}{3!}, \quad \text{and} \quad a_4 = \frac{2}{3} = \frac{b_4}{4!}$$

for some values b_2, b_3 and b_4. Solving for these values, we see that $b_2 = 4$, $b_3 = 8$ and $b_4 = 16$. That is, we are recovering the pattern $b_n = 2^n$, allowing us to write

$$f(x) = \sum_{n=0}^{\infty} a_n x^n = \sum_{n=0}^{\infty} \frac{b_n}{n!} x^n$$

$$= 1 + 2x + \frac{4}{2!}x^2 + \frac{8}{3!}x^3 + \frac{16}{4!}x^4 + \cdots$$

From here it is easier to recognize that the series is describing an exponential function.

This chapter introduced sequences, which are ordered lists of numbers, followed by series, wherein we add up the terms of a sequence. We quickly saw that such sums do not always add up to "infinity," but rather converge. We studied tests for convergence, then ended the chapter with a formal way of defining functions based on series. Such "series–defined functions" are a valuable tool in solving a number of different problems throughout science and engineering.

Coming in the next chapters are new ways of defining curves in the plane apart from using functions of the form $y = f(x)$. Curves created by these new methods can be beautiful, useful, and important.

Notes:

Exercises 9.10

Terms and Concepts

1. What is the difference between a Taylor polynomial and a Taylor series?
2. What theorem must we use to show that a function is equal to its Taylor series?

Problems

Key Idea 33 gives the nth term of the Taylor series of common functions. In Exercises 3–6, verify the formula given in the Key Idea by finding the first few terms of the Taylor series of the given function and identifying a pattern.

3. $f(x) = e^x; \quad c = 0$

4. $f(x) = \sin x; \quad c = 0$

5. $f(x) = 1/(1-x); \quad c = 0$

6. $f(x) = \tan^{-1} x; \quad c = 0$

In Exercises 7–12, find a formula for the nth term of the Taylor series of $f(x)$, centered at c, by finding the coefficients of the first few powers of x and looking for a pattern. (The formulas for several of these are found in Key Idea 33; show work verifying these formula.)

7. $f(x) = \cos x; \quad c = \pi/2$

8. $f(x) = 1/x; \quad c = 1$

9. $f(x) = e^{-x}; \quad c = 0$

10. $f(x) = \ln(1+x); \quad c = 0$

11. $f(x) = x/(x+1); \quad c = 1$

12. $f(x) = \sin x; \quad c = \pi/4$

In Exercises 13–16, show that the Taylor series for $f(x)$, as given in Key Idea 33, is equal to $f(x)$ by applying Theorem 82; that is, show $\lim\limits_{n \to \infty} R_n(x) = 0$.

13. $f(x) = e^x$

14. $f(x) = \sin x$

15. $f(x) = \ln(x+1)$ (show equality only on $(0,1)$).

16. $f(x) = 1/(1-x)$ (show equality only on $(-1, 0)$)

In Exercises 17–20, use the Taylor series given in Key Idea 33 to verify the given identity.

17. $\cos(-x) = \cos x$

18. $\sin(-x) = -\sin x$

19. $\frac{d}{dx}\left(\sin x \right) = \cos x$

20. $\frac{d}{dx}\left(\cos x \right) = -\sin x$

In Exercises 21–24, write out the first 5 terms of the Binomial series with the given k-value.

21. $k = 1/2$

22. $k = -1/2$

23. $k = 1/3$

24. $k = 4$

In Exercises 25–30, use the Taylor series given in Key Idea 33 to create the Taylor series of the given functions.

25. $f(x) = \cos\left(x^2\right)$

26. $f(x) = e^{-x}$

27. $f(x) = \sin\left(2x + 3\right)$

28. $f(x) = \tan^{-1}\left(x/2\right)$

29. $f(x) = e^x \sin x$ (only find the first non-zero 4 terms)

30. $f(x) = (1+x)^{1/2} \cos x$ (only find the first non-zero 4 terms)

In Exercises 31–32, approximate the value of the given definite integral by using the first 4 nonzero terms of the integrand's Taylor series.

31. $\displaystyle\int_0^{\sqrt{\pi}} \sin\left(x^2\right) dx$

32. $\displaystyle\int_0^{\pi^2/4} \cos\left(\sqrt{x}\right) dx$

10.0   Chapter Prerequisites — Conic Sections

The material in this section provides a basic review of and practice problems for pre-calculus skills essential to your success in Calculus. You should take time to review this section and work the suggested problems (checking your answers against those in the back of the book). Since this content is a pre-requisite for Calculus, reviewing and mastering these skills are considered your responsibility. This means that minimal, and in some cases no, class time will be devoted to this section. When you identify areas that you need help with we strongly urge you to seek assistance outside of class from your instructor or other student tutoring service.

The ancient Greeks recognized that interesting shapes can be formed by intersecting a plane with a *double napped* cone (i.e., two identical cones placed tip–to–tip as shown in the following figures). As these shapes are formed as sections of conics, they have earned the official name "conic sections."

The three "most interesting" conic sections are given in the top row of Figure 10.1. They are the parabola, the ellipse (which includes circles) and the hyperbola. In each of these cases, the plane does not intersect the tips of the cones (usually taken to be the origin).

Parabola   Ellipse   Circle   Hyperbola

Point   Line   Crossed Lines
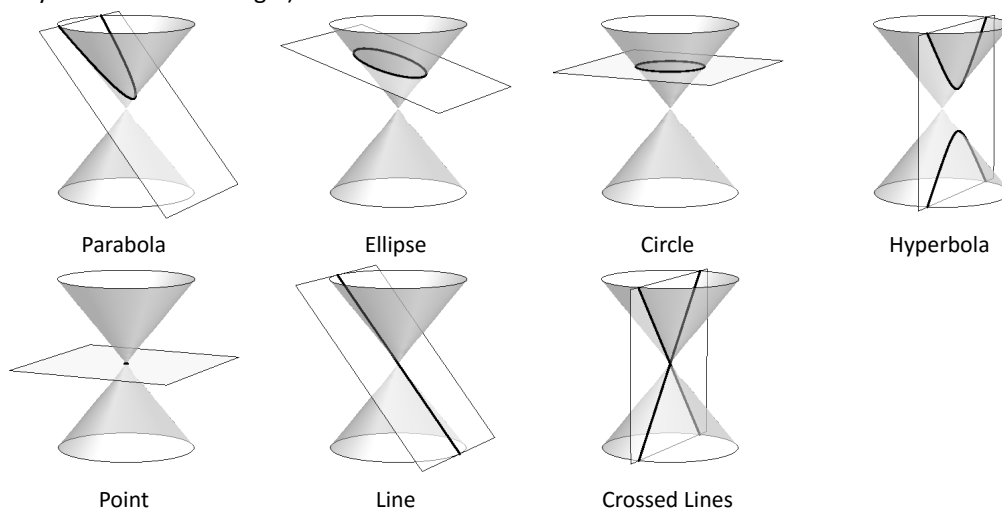
Figure 10.1: Conic Sections

When the plane does contain the origin, three **degenerate** cones can be formed as shown the bottom row of Figure 10.1: a point, a line, and crossed lines. We focus here on the nondegenerate cases.

While the above geometric constructs define the conics in an intuitive, visual way, these constructs are not very helpful when trying to analyze the shapes

---

Notes:

algebraically or consider them as the graph of a function. It can be shown that all conics can be defined by the general second–degree equation

$$Ax^2 + Bxy + Cy^2 + Dx + Ey + F = 0.$$

While this algebraic definition has its uses, most find another geometric perspective of the conics more beneficial.

Each nondegenerate conic can be defined as the **locus**, or set, of points that satisfy a certain distance property. These distance properties can be used to generate an algebraic formula, allowing us to study each conic as the graph of a function.

Parabolas

> **Definition 44   Parabola**
> A **parabola** is the locus of all points equidistant from a point (called a **focus**) and a line (called the **directrix**) that does not contain the focus.
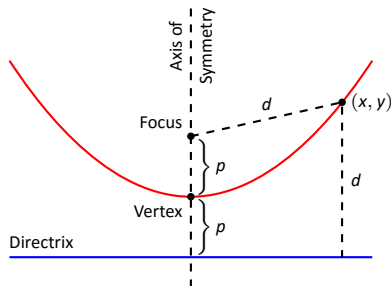
Figure 10.2: Illustrating the definition of the parabola and establishing an algebraic formula.

Figure 10.2 illustrates this definition. The point halfway between the focus and the directrix is the **vertex**. The line through the focus, perpendicular to the directrix, is the **axis of symmetry**, as the portion of the parabola on one side of this line is the mirror–image of the portion on the opposite side.

The geometric definition of the parabola and distance formula can be used to derive the quadratic function whose graph is a parabola with vertex at the origin.

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

Applying transformations of functions we get the following standard form of the parabola.

Notes:

Key Idea 34   General Equation of a Parabola

1. **Vertical Axis of Symmetry:** The equation of the parabola with vertex at (h, k), directrix $y = k - p$, and focus at $(h, k + p)$ in standard form is
$$y = \frac{1}{4p}(x - h)^2 + k.$$

2. **Horizontal Axis of Symmetry:** The equation of the parabola with vertex at (h, k), directrix $x = h - p$, and focus at $(h + p, k)$ in standard form is
$$x = \frac{1}{4p}(y - k)^2 + h.$$

Note: p is not necessarily a positive number.

Example 1   Finding the equation of a parabola
Give the equation of the parabola with focus at $(1, 2)$ and directrix at $y = 3$.

SOLUTION   The vertex is located halfway between the focus and directrix, so $(h, k) = (1, 2.5)$. This gives $p = -0.5$. Using Key Idea 34 we have the equation of the parabola as

$$y = \frac{1}{4(-0.5)}(x - 1)^2 + 2.5 = -\frac{1}{2}(x - 1)^2 + 2.5.$$

The parabola is sketched in Figure 10.3.

Ellipses

Definition 45   Ellipse
An **ellipse** is the locus of all points whose sum of distances from two fixed points, each a **focus** of the ellipse, is constant.

An easy way to visualize this construction of an ellipse is to pin both ends of a string to a board. The pins become the foci. Holding a pencil tight against the
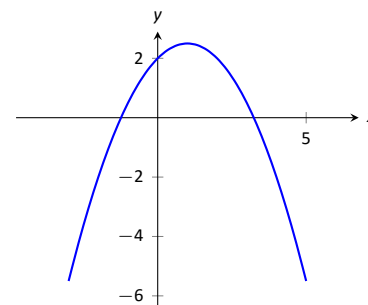
Figure 10.3: The parabola described in Example 1.

Notes:

string places the pencil on the ellipse; the sum of distances from the pencil to the pins is constant: the length of the string. See Figure 10.4.

As shown in Figure 10.5, the values of a and b have meaning. In general, the two foci of an ellipse lie on the **major axis** of the ellipse, and the midpoint of the segment joining the two foci is the **center**. The major axis intersects the ellipse at two points, each of which is a **vertex**. The line segment through the center and perpendicular to the major axis is the **minor axis**. The "constant sum of distances" that defines the ellipse is the length of the major axis, i.e., $2a$.

Allowing for the shifting of the ellipse gives the following standard equations.

Figure 10.4: Illustrating the construction of an ellipse with pins, pencil and string.
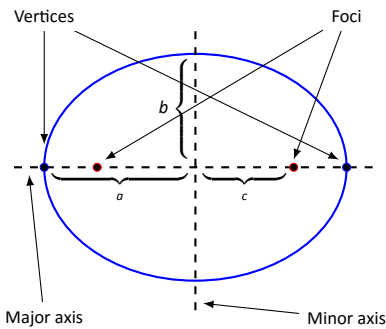
Figure 10.5: Labeling the significant features of an ellipse.

Key Idea 35   Standard Equation of the Ellipse

The equation of an ellipse centered at (h, k) with major axis of length $2a$ and minor axis of length $2b$ in standard form is:

1. **Horizontal major axis:** $\dfrac{(x-h)^2}{a^2} + \dfrac{(y-k)^2}{b^2} = 1.$

2. **Vertical major axis:** $\dfrac{(x-h)^2}{b^2} + \dfrac{(y-k)^2}{a^2} = 1.$

The foci lie along the major axis, c units from the center, where $c^2 = a^2 - b^2$.

Example 2   Finding the equation of an ellipse

Find the general equation of the ellipse graphed in Figure 10.6.

SOLUTION   The center is located at $(-3, 1)$. The distance from the center to a vertex is 5 units, hence $a = 5$. The minor axis seems to have length 4, so $b = 2$. Thus the equation of the ellipse is

$$\frac{(x+3)^2}{4} + \frac{(y-1)^2}{25} = 1.$$

Figure 10.6: The ellipse used in Example 2.

Example 3   Graphing an ellipse

Graph the ellipse defined by $4x^2 + 9y^2 - 8x - 36y = -4$.

SOLUTION   It is simple to graph an ellipse once it is in standard form. In order to put the given equation in standard form, we must complete the square with both the x and y terms. We first rewrite the equation by regrouping:

$$4x^2 + 9y^2 - 8x - 36y = -4 \quad \Rightarrow \quad (4x^2 - 8x) + (9y^2 - 36y) = -4.$$

Notes:

Now we complete the squares.

$$(4x^2 - 8x) + (9y^2 - 36y) = -4$$
$$4(x^2 - 2x) + 9(y^2 - 4y) = -4$$
$$4(x^2 - 2x + 1 - 1) + 9(y^2 - 4y + 4 - 4) = -4$$
$$4\big((x - 1)^2 - 1\big) + 9\big((y - 2)^2 - 4\big) = -4$$
$$4(x - 1)^2 - 4 + 9(y - 2)^2 - 36 = -4$$
$$4(x - 1)^2 + 9(y - 2)^2 = 36$$
$$\frac{(x - 1)^2}{9} + \frac{(y - 2)^2}{4} = 1.$$

We see the center of the ellipse is at $(1, 2)$. We have $a = 3$ and $b = 2$; the major axis is horizontal, so the vertices are located at $(-2, 2)$ and $(4, 2)$. We find $c = \sqrt{9 - 4} = \sqrt{5} \approx 2.24$. The foci are located along the major axis, approximately 2.24 units from the center, at $(1 \pm 2.24, 2)$. This is all graphed in Figure 10.7.

Hyperbolas

The definition of a hyperbola is very similar to the definition of an ellipse; we essentially just change the word "sum" to "difference."

> **Definition 46   Hyperbola**
> A **hyperbola** is the locus of all points where the absolute value of the difference of distances from two fixed points, each a focus of the hyperbola, is constant.

We do not have a convenient way of visualizing the construction of a hyperbola as we did for the ellipse. The geometric definition does allow us to find an algebraic expression that describes it. It will be useful to define some terms first.

The two foci lie on the **transverse axis** of the hyperbola; the midpoint of the line segment joining the foci is the **center** of the hyperbola. The transverse axis intersects the hyperbola at two points, each a **vertex** of the hyperbola. The line through the center and perpendicular to the transverse axis is the **conjugate axis.** This is illustrated in Figure 10.8. It is easy to show that the constant difference of distances used in the definition of the hyperbola is the distance between the vertices, i.e., $2a$.

Notes:

Figure 10.7: Graphing the ellipse in Example 3.
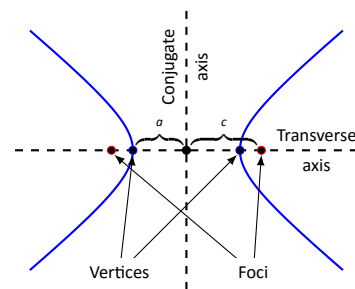
Figure 10.8: Labeling the significant features of a hyperbola.

> **Key Idea 36   Standard Equation of a Hyperbola**
>
> The equation of a hyperbola centered at (h, k) in standard form is:
>
> 1. **Horizontal Transverse Axis:** $\dfrac{(x-h)^2}{a^2} - \dfrac{(y-k)^2}{b^2} = 1.$
>
> 2. **Vertical Transverse Axis:** $\dfrac{(y-k)^2}{a^2} - \dfrac{(x-h)^2}{b^2} = 1.$
>
> The vertices are located a units from the center and the foci are located c units from the center, where $c^2 = a^2 + b^2$.
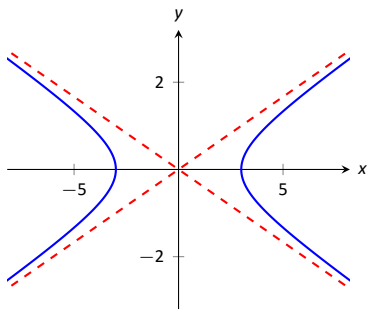
Figure 10.9: Graphing the hyperbola $\frac{x^2}{9} - \frac{y^2}{1} = 1$ along with its asymptotes, $y = \pm x/3$.
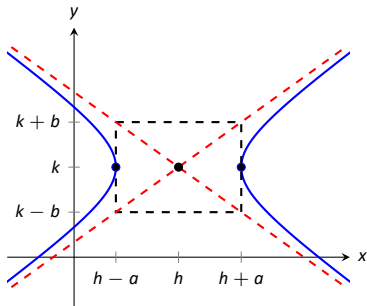
Figure 10.10: Using the asymptotes of a hyperbola as a graphing aid.
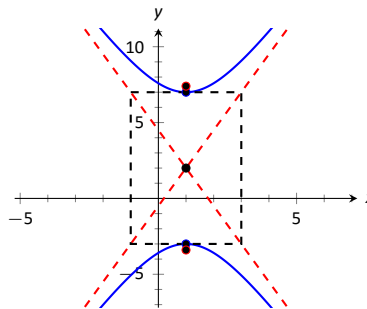
Figure 10.11: Graphing the hyperbola in Example 4.

Graphing Hyperbolas

Consider the hyperbola $\frac{x^2}{9} - \frac{y^2}{1} = 1$. Solving for y, we find $y = \pm\sqrt{x^2/9 - 1}$. As x grows large, the "-1" part of the equation for y becomes less significant and $y \approx \pm\sqrt{x^2/9} = \pm x/3$. That is, as x gets large, the graph of the hyperbola looks very much like the lines $y = \pm x/3$. These lines are asymptotes of the hyperbola, as shown in Figure 10.9.

This is a valuable tool in sketching. Given the equation of a hyperbola in general form, draw a rectangle centered at (h, k) with sides of length $2a$ parallel to the transverse axis and sides of length $2b$ parallel to the conjugate axis. (See Figure 10.10 for an example with a horizontal transverse axis.) The diagonals of the rectangle lie on the asymptotes.

These lines pass through (h, k). When the transverse axis is horizontal, the slopes are $\pm b/a$; when the transverse axis is vertical, their slopes are $\pm a/b$. This gives equations:

| Horizontal | Vertical |
| Transverse Axis | Transverse Axis |

$$y = \pm\frac{b}{a}(x - h) + k \qquad y = \pm\frac{a}{b}(x - h) + k.$$

Example 4   Graphing a hyperbola

Sketch the hyperbola given by $\dfrac{(y-2)^2}{25} - \dfrac{(x-1)^2}{4} = 1.$

SOLUTION   The hyperbola is centered at $(1, 2)$; $a = 5$ and $b = 2$. In Figure 10.11 we draw the prescribed rectangle centered at $(1, 2)$ along with the asymptotes defined by its diagonals. The hyperbola has a vertical transverse

Notes:

axis, so the vertices are located at $(1, 7)$ and $(1, -3)$. This is enough to make a good sketch.

We also find the location of the foci: as $c^2 = a^2 + b^2$, we have $c = \sqrt{29} \approx 5.4$. Thus the foci are located at $(1, 2 \pm 5.4)$ as shown in the figure.

Example 5   **Graphing a hyperbola**

Sketch the hyperbola given by $9x^2 - y^2 + 2y = 10$.

SOLUTION   We must complete the square to put the equation in general form. (We recognize this as a hyperbola since it is a general quadratic equation and the x^2 and y^2 terms have opposite signs.)

$$9x^2 - y^2 + 2y = 10$$
$$9x^2 - (y^2 - 2y) = 10$$
$$9x^2 - (y^2 - 2y + 1 - 1) = 10$$
$$9x^2 - \left((y-1)^2 - 1\right) = 10$$
$$9x^2 - (y-1)^2 = 9$$
$$x^2 - \frac{(y-1)^2}{9} = 1$$

We see the hyperbola is centered at $(0, 1)$, with a horizontal transverse axis, where $a = 1$ and $b = 3$. The appropriate rectangle is sketched in Figure 10.12 along with the asymptotes of the hyperbola. The vertices are located at $(\pm 1, 1)$. We have $c = \sqrt{10} \approx 3.2$, so the foci are located at $(\pm 3.2, 1)$ as shown in Figure 10.12.

This chapter explores curves in the plane, in particular curves that cannot be described by functions of the form $y = f(x)$. In this section, we learned of ellipses and hyperbolas that are defined implicitly, not explicitly. In the following sections, we will learn completely new ways of describing curves in the plane, using *parametric equations* and *polar coordinates*, then study these curves using calculus techniques.
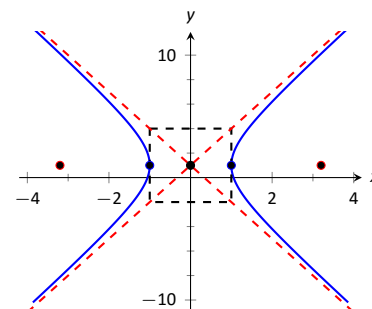
Figure 10.12: Graphing the hyperbola in Example 5.

Notes:

Exercises 10.0

Problems

In Exercises 1–8, find the equation of the parabola defined by the given information. Sketch the parabola.

1. Focus: $(3, 2)$; directrix: $y = 1$

2. Focus: $(-1, -4)$; directrix: $y = 2$

3. Focus: $(1, 5)$; directrix: $x = 3$

4. Focus: $(1/4, 0)$; directrix: $x = -1/4$

5. Focus: $(1, 1)$; vertex: $(1, 2)$

6. Focus: $(-3, 0)$; vertex: $(0, 0)$

7. Vertex: $(0, 0)$; directrix: $y = -1/16$

8. Vertex: $(2, 3)$; directrix: $x = 4$

In Exercises 9–10, sketch the ellipse defined by the given equation. Label the center, foci and vertices.

9. $\dfrac{(x-1)^2}{3} + \dfrac{(y-2)^2}{5} = 1$

10. $\dfrac{1}{25}x^2 + \dfrac{1}{9}(y+3)^2 = 1$

In Exercises 11–12, find the equation of the ellipse shown in the graph.

11.

12.
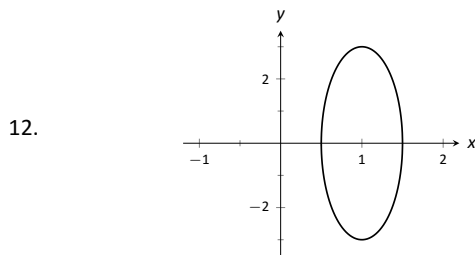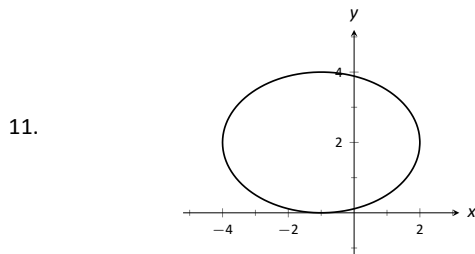
In Exercises 13–16, write the equation of the given ellipse in standard form.

13. $x^2 - 2x + 2y^2 - 8y = -7$

14. $5x^2 + 3y^2 = 15$

15. $3x^2 + 2y^2 - 12y + 6 = 0$

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

In Exercises 17–20, find the equation of the hyperbola shown in the graph.

17.

18.

19.

20.
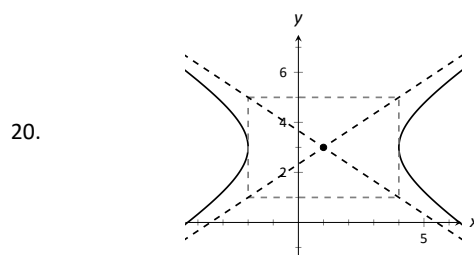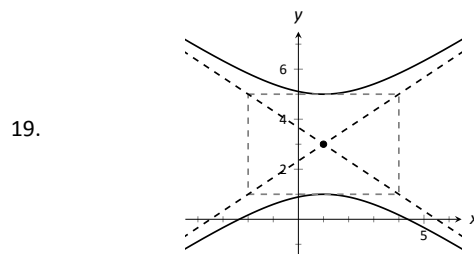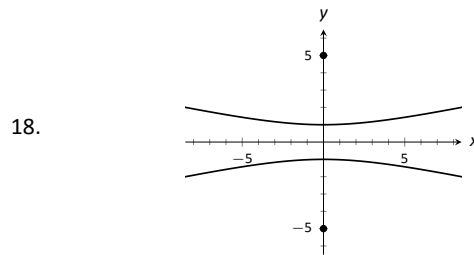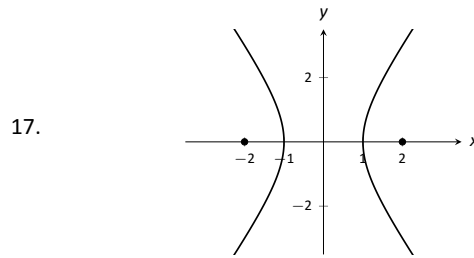
In Exercises 21–22, sketch the hyperbola defined by the given equation. Label the center.

21. $\dfrac{(x-1)^2}{16} - \dfrac{(y+2)^2}{9} = 1$

22. $(y-4)^2 - \dfrac{(x+1)^2}{25} = 1$

In Exercises 23–26, write the equation of the hyperbola in standard form.

23. $3x^2 - 4y^2 = 12$

24. $3x^2 - y^2 + 2y = 10$

25. $x^2 - 10y^2 + 40y = 30$

26. $(4y - x)(4y + x) = 4$

10: Curves in the Plane

We have explored functions of the form $y = f(x)$ closely throughout this text. We have explored their limits, their derivatives and their antiderivatives; we have learned to identify key features of their graphs, such as relative maxima and minima, inflection points and asymptotes; we have found equations of their tangent lines, the areas between portions of their graphs and the x-axis, and the volumes of solids generated by revolving portions of their graphs about a horizontal or vertical axis.

Despite all this, the graphs created by functions of the form $y = f(x)$ are limited. Since each x-value can correspond to only 1 y-value, common shapes like circles cannot be fully described by a function in this form. Fittingly, the "vertical line test" excludes vertical lines from being functions of x, even though these lines are important in mathematics.

In this chapter we'll explore new ways of drawing curves in the plane. We'll still work within the framework of functions, as an input will still only correspond to one output. However, our new techniques of drawing curves will render the vertical line test pointless, and allow us to create important — and beautiful — new curves. Once these curves are defined, we'll apply the concepts of calculus to them, continuing to find equations of tangent lines and the areas of enclosed regions.

One aspect that we'll be interested in is "how long is this curve?" Before we explore that idea for these new ways to draw curves, we'll start by exploring how long a curve is when we've gotten it from a regular $y = f(x)$ function.

10.1 Arc Length and Surface Area

In previous sections we have used integration to answer the following questions:

1. Given a region, what is its area?

2. Given a solid, what is its volume?

In this section, we address a related question: Given a curve, what is its length? This is often referred to as **arc length**.

Consider the graph of $y = \sin x$ on $[0, \pi]$ given in Figure 10.13 (a). How long is this curve? That is, if we were to use a piece of string to exactly match the shape of this curve, how long would the string be?

As we have done in the past, we start by approximating; later, we will refine our answer using limits to get an exact solution.

(a)

(b)
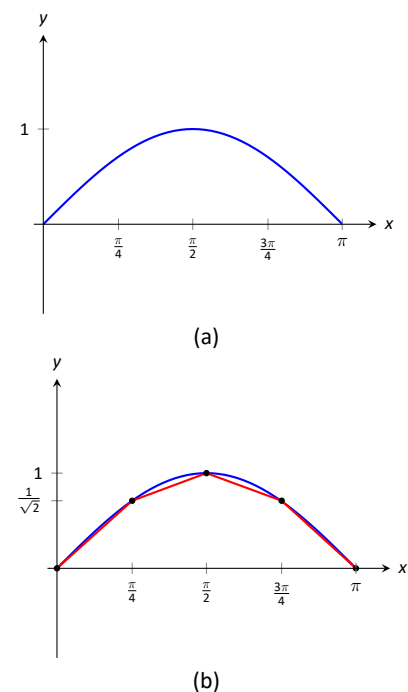
Figure 10.13: Graphing $y = \sin x$ on $[0, \pi]$ and approximating the curve with line segments.

The length of straight–line segments is easy to compute using the Distance Formula. We can approximate the length of the given curve by approximating the curve with straight lines and measuring their lengths.

In Figure 10.13 (b), the curve $y = \sin x$ has been approximated with 4 line segments (the interval $[0, \pi]$ has been divided into 4 equally–lengthed subintervals). It is clear that these four line segments approximate $y = \sin x$ very well on the first and last subinterval, though not so well in the middle. Regardless, the sum of the lengths of the line segments is 3.79, so we approximate the arc length of $y = \sin x$ on $[0, \pi]$ to be 3.79.

In general, we can approximate the arc length of $y = f(x)$ on $[a, b]$ in the following manner. Let $a = x_0 < x_1 < \ldots < x_{n-1} < x_n = b$ be a partition of $[a, b]$ into n subintervals. Let Δx_i represent the length of the ith subinterval $[x_{i-1}, x_i]$.

Figure 10.14 zooms in on the ith subinterval where $y = f(x)$ is approximated by a straight line segment. The dashed lines show that we can view this line segment as the hypotenuse of a right triangle whose sides have length Δx_i and Δy_i. Using the Pythagorean Theorem, the length of this line segment is $\sqrt{(\Delta x_i)^2 + (\Delta y_i)^2}$. Summing over all subintervals gives an arc length approximation

$$L \approx \sum_{i=1}^{n} \sqrt{(\Delta x_i)^2 + (\Delta y_i)^2}.$$

Figure 10.14: Zooming in on the ith subinterval $[x_{i-1}, x_i]$ of a partition of $[a, b]$.

As it is written, this is *not* a Riemann Sum. While we could conclude that taking a limit as the subinterval length goes to zero gives the exact arc length, we would not be able to compute the answer with a definite integral. We need first to do a little algebra.

In the above expression factor out a Δx_i^2 term:

$$\sum_{i=1}^{n} \sqrt{(\Delta x_i)^2 + (\Delta y_i)^2} = \sum_{i=1}^{n} \sqrt{(\Delta x_i)^2 \left(1 + \frac{(\Delta y_i)^2}{(\Delta x_i)^2}\right)}.$$

Now pull the $(\Delta x_i)^2$ term out of the square root:

$$L \approx \sum_{i=1}^{n} \sqrt{1 + \frac{(\Delta y_i)^2}{(\Delta x_i)^2}}\, \Delta x_i.$$

This is nearly a Riemann Sum. Consider the $(\Delta y_i)^2/(\Delta x_i)^2$ term. The expression $\Delta y_i/\Delta x_i$ measures the "change in y/change in x," that is, the "rise over run" of f on the ith subinterval. The Mean Value Theorem of Differentiation (Theorem 24) states that there is a c_i in the ith subinterval where $f'(c_i) = \Delta y_i/\Delta x_i$. Thus we can rewrite our above expression as:

$$L \approx \sum_{i=1}^{n} \sqrt{1 + [f'(c_i)]^2}\, \Delta x_i.$$

Notes:

This *is* a Riemann Sum. As long as f' is continuous on $[a, b]$, we can invoke Theorem 36 and conclude

$$L = \int_a^b \sqrt{1 + [f'(x)]^2} \, dx.$$

Key Idea 37   Arc Length

Let f be differentiable on an open interval containing $[a, b]$, where f' is also continuous on $[a, b]$. Then the arc length of f from $x = a$ to $x = b$ is

$$L = \int_a^b \sqrt{1 + [f'(x)]^2} \, dx.$$

Watch the video:
Arc Length at
`https://youtu.be/PwmCZAWeRNE`

As the integrand contains a square root, it is often difficult to use the formula in Key Idea 37 to find the length exactly. When exact answers are difficult to come by, we resort to using numerical methods of approximating definite integrals. The following examples will demonstrate this.

Example 1   Finding arc length
Find the arc length of $f(x) = x^{3/2}$ from $x = 0$ to $x = 4$.

SOLUTION   A graph of f is given in Figure 10.15. We begin by finding
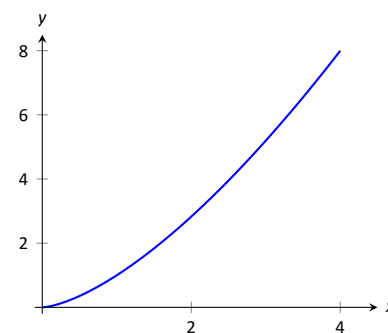
Figure 10.15: A graph of $f(x) = x^{3/2}$ from Example 1.

Notes:

$f'(x) = \frac{3}{2}x^{1/2}$. Using the formula, we find the arc length L as

$$L = \int_0^4 \sqrt{1 + \left(\frac{3}{2}x^{1/2}\right)^2}\, dx$$

$$= \int_0^4 \sqrt{1 + \frac{9}{4}x}\, dx$$

$$= \int_0^4 \left(1 + \frac{9}{4}x\right)^{1/2} dx$$

$$= \frac{2}{3}\frac{4}{9}\left(1 + \frac{9}{4}x\right)^{3/2}\Big|_0^4$$

$$= \frac{8}{27}\left(10^{3/2} - 1\right) \text{ units.}$$

Example 2   **Finding arc length**

Find the arc length of $f(x) = \frac{1}{8}x^2 - \ln x$ from $x = 1$ to $x = 2$.

SOLUTION   A graph of f is given in Figure 10.16; the portion of the curve measured in this problem is in bold. This function was chosen specifically because the resulting integral can be evaluated exactly. We begin by finding $f'(x) = x/4 - 1/x$. The arc length is

$$L = \int_1^2 \sqrt{1 + \left(\frac{x}{4} - \frac{1}{x}\right)^2}\, dx$$

$$= \int_1^2 \sqrt{1 + \frac{x^2}{16} - \frac{1}{2} + \frac{1}{x^2}}\, dx$$

$$= \int_1^2 \sqrt{\frac{x^2}{16} + \frac{1}{2} + \frac{1}{x^2}}\, dx$$

$$= \int_1^2 \sqrt{\left(\frac{x}{4} + \frac{1}{x}\right)^2}\, dx$$

$$= \int_1^2 \left(\frac{x}{4} + \frac{1}{x}\right)\, dx$$

$$= \left(\frac{x^2}{8} + \ln x\right)\Big|_1^2$$

$$= \frac{3}{8} + \ln 2 \text{ units.}$$
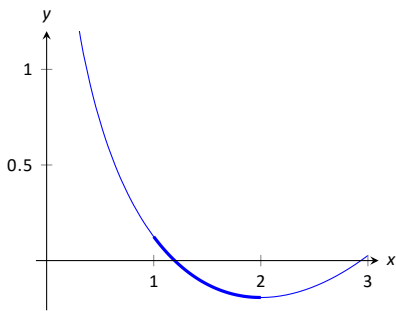
Figure 10.16: A graph of $f(x) = \frac{1}{8}x^2 - \ln x$ from Example 2.

Notes:

The previous examples found the arc length exactly through careful choice of the functions. In general, exact answers are much more difficult to come by and numerical approximations are necessary.

Example 3   **Approximating arc length numerically**
Find the length of the sine curve from $x = 0$ to $x = \pi$.

SOLUTION   This is somewhat of a mathematical curiosity; in Example 5.4.3 we found the area under one "hump" of the sine curve is 2 square units; now we are measuring its arc length.
     The setup is straightforward: $f(x) = \sin x$ and $f'(x) = \cos x$. Thus

$$L = \int_0^\pi \sqrt{1 + \cos^2 x}\, dx.$$

This integral *cannot* be evaluated in terms of elementary functions so we will approximate it with Simpson's Method with $n = 4$. Figure 10.17 gives $\sqrt{1 + \cos^2 x}$ evaluated at 5 evenly spaced points in $[0, \pi]$. Simpson's Rule then states that

$$\int_0^\pi \sqrt{1 + \cos^2 x}\, dx \approx \frac{\pi - 0}{4 \cdot 3}\left(\sqrt{2} + 4\sqrt{3/2} + 2(1) + 4\sqrt{3/2} + \sqrt{2} \right)$$

$$\approx 3.82918.$$

Using a computer with $n = 100$ the approximation is $L \approx 3.8202$; our approximation with $n = 4$ is quite good. Our approximation of 3.79 from the beginning of this section isn't as close.

x	$\sqrt{1 + \cos^2 x}$
0	$\sqrt{2}$
$\pi/4$	$\sqrt{3/2}$
$\pi/2$	1
$3\pi/4$	$\sqrt{3/2}$
π	$\sqrt{2}$

Figure 10.17: A table of values of $y = \sqrt{1 + \cos^2 x}$ to evaluate a definite integral in Example 3.

Surface Area of Solids of Revolution

We have already seen how a curve $y = f(x)$ on $[a, b]$ can be revolved around an axis to form a solid. Instead of computing its volume, we now consider its surface area.
     We begin as we have in the previous sections: we partition the interval $[a, b]$ with n subintervals, where the ith subinterval is $[x_i, x_{i+1}]$. On each subinterval, we can approximate the curve $y = f(x)$ with a straight line that connects $f(x_i)$ and $f(x_{i+1})$ as shown in Figure 10.18(a). Revolving this line segment about the x-axis creates part of a cone (called a *frustum* of a cone) as shown in Figure 10.18(b). The surface area of a frustum of a cone is

$$A = 2\pi r_{\text{avg}} L,$$

where r_{avg} is the average of R_1 and R_2. The length is given by L; we use the material just covered by arc length to state that

$$L \approx \sqrt{1 + [f'(c_i)]^2}\, \Delta x_i$$

(a)

(b)
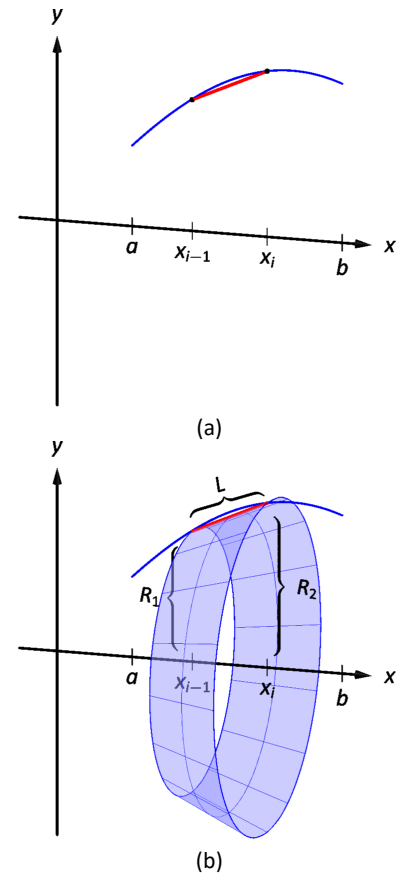
Figure 10.18: Establishing the formula for surface area.

Notes:

for some c_i in the i th subinterval. The radii are just the function evaluated at the endpoints of the interval: $f(x_{i-1})$ and $f(x_i)$. Thus the surface area of this sample frustum of the cone is approximately

$$2\pi\frac{f(x_{i-1})+f(x_i)}{2}\sqrt{1+[f'(c_i)]^2}\Delta x_i.$$

Since f is a continuous function, the Intermediate Value Theorem states there is some d_i in $[x_{i-1}, x_i]$ such that $f(d_i) = \dfrac{f(x_{i-1})+f(x_i)}{2}$; we can use this to rewrite the above equation as

$$2\pi f(d_i)\sqrt{1+[f'(c_i)]^2}\Delta x_i.$$

Summing over all the subintervals we get the total surface area to be approximately

$$\text{Surface Area} \approx \sum_{i=1}^{n} 2\pi f(d_i)\sqrt{1+[f'(c_i)]^2}\Delta x_i,$$

which is a Riemann Sum. Taking the limit as the subinterval lengths go to zero gives us the exact surface area, given in the upcoming Key Idea.

If instead we revolve $y = f(x)$ about the y-axis, the radii of the resulting frustum are x_{i-1} and x_i; their average value is simply the midpoint of the interval. In the limit, this midpoint is just x. This gives the second part of Key Idea 38.

Key Idea 38   Surface Area of a Solid of Revolution

Let f be differentiable on an open interval containing $[a, b]$ where f' is also continuous on $[a, b]$.

1. The surface area of the solid formed by revolving the graph of $y = f(x)$, where $f(x) \geq 0$, about the x-axis is

$$\text{Surface Area} = 2\pi \int_a^b f(x)\sqrt{1+[f'(x)]^2}\,dx.$$

2. The surface area of the solid formed by revolving the graph of $y = f(x)$ about the y-axis, where $a, b \geq 0$, is

$$\text{Surface Area} = 2\pi \int_a^b x\sqrt{1+[f'(x)]^2}\,dx.$$

Notes:

Example 4   **Finding surface area of a solid of revolution**
Find the surface area of the solid formed by revolving $y = \sin x$ on $[0, \pi]$ around the x-axis, as shown in Figure 10.19.

SOLUTION   The setup turns out to be easier than the resulting integral. Using Key Idea 38, we have the surface area SA is:

$$SA = 2\pi \int_0^\pi \sin x \sqrt{1 + \cos^2 x}\, dx$$

$$= -2\pi \int_1^{-1} \sqrt{1 + u^2}\, du \qquad \text{substitute } u = \cos x$$

$$= 2\pi \int_{-\pi/4}^{\pi/4} \sec^3 \theta\, d\theta \qquad \text{substitute } u = \tan \theta$$

$$= \pi \left(\sec \theta \tan \theta + \ln |\sec \theta + \tan \theta|\right)\big|_{-\pi/4}^{\pi/4} \qquad \text{by Example 8.2.6}$$

$$= \pi \left(\sqrt{2} + \ln\left(\sqrt{2} + 1\right) - \left(-\sqrt{2} + \ln\left(\sqrt{2} - 1\right)\right)\right)$$

$$= \pi \left(2\sqrt{2} + \ln\left(\frac{\sqrt{2} + 1}{\sqrt{2} - 1}\right)\right)$$

$$= 2\pi \left(\sqrt{2} + \ln\left(\sqrt{2} + 1\right)\right) \text{ units}^2 \qquad \text{rationalize the denominator.}$$

It is interesting to see that the surface area of a solid, whose shape is defined by a trigonometric function, involves both a square root and a natural logarithm.

Figure 10.19: Revolving $y = \sin x$ on $[0, \pi]$ about the x-axis.

Example 5   **Finding surface area of a solid of revolution**
Find the surface area of the solid formed by revolving the curve $y = x^2$ on $[0, 1]$ about:

     1.   the x-axis   2.   the y-axis.

SOLUTION

1. The solid formed by revolving $y = x^2$ around the x-axis is graphed in Figure 10.20(a). Like the integral in Example 4, this integral is easier to setup than to actually integrate. While it is possible to use a trigonometric substitution to evaluate this integral, it is significantly more difficult than a solution employing the hyperbolic sine:

$$SA = 2\pi \int_0^1 x^2 \sqrt{1 + (2x)^2}\, dx.$$

$$= \frac{\pi}{32} \left(2(8x^3 + x)\sqrt{1 + 4x^2} - \sinh^{-1}(2x)\right)\Big|_0^1$$

$$= \frac{\pi}{32} \left(18\sqrt{5} - \sinh^{-1} 2\right) \text{ units}^2.$$

(a)

(b)

Figure 10.20: The solids used in Example 5.

Notes:

2. Since we are revolving around the *y*-axis, the "radius" of the solid is not *f(x)* but rather *x*. Thus the integral to compute the surface area is:

$$SA = 2\pi \int_0^1 x\sqrt{1 + (2x)^2}\, dx$$

$$= \frac{\pi}{4} \int_1^5 \sqrt{u}\, du \qquad \text{substitute } u = 1 + 4x^2$$

$$= \frac{\pi}{4}\frac{2}{3}u^{3/2}\Big|_1^5$$

$$= \frac{\pi}{6}\left(5\sqrt{5} - 1\right) \text{ units}^2.$$

The solid formed by revolving $y = x^2$ about the *y*-axis is graphed in Figure 10.20 (b).

Our final example is a famous mathematical "paradox."

Example 6 **The surface area and volume of Gabriel's Horn**
Consider the solid formed by revolving $y = 1/x$ about the *x*-axis on $[1, \infty)$. Find the volume and surface area of this solid. (This shape, as graphed in Figure 10.21, is known as "Gabriel's Horn" since it looks like a very long horn that only a supernatural person, such as an angel, could play.)

SOLUTION To compute the volume it is natural to use the Disk Method. We have:

$$V = \pi \int_1^\infty \frac{1}{x^2}\, dx$$

$$= \lim_{b \to \infty} \pi \int_1^b \frac{1}{x^2}\, dx$$

$$= \lim_{b \to \infty} \pi \left(\frac{-1}{x}\right)\Big|_1^b$$

$$= \lim_{b \to \infty} \pi \left(1 - \frac{1}{b}\right)$$

$$= \pi \text{ units}^3.$$

Gabriel's Horn has a finite volume of π cubic units. Since we have already seen that regions with infinite length can have a finite area, this is not too difficult to accept.
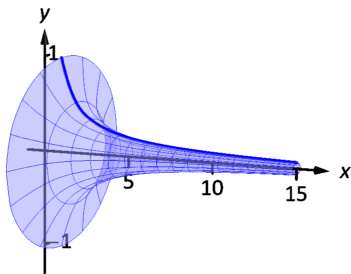
Figure 10.21: A graph of Gabriel's Horn.

Notes:

We now consider its surface area. The integral is straightforward to setup:

$$SA = 2\pi \int_1^\infty \frac{1}{x}\sqrt{1 + 1/x^4}\,dx.$$

Integrating this expression is not trivial. We can, however, compare it to other improper integrals. Since $1 < \sqrt{1 + 1/x^4}$ on $[1, \infty)$, we can state that

$$2\pi \int_1^\infty \frac{1}{x}\,dx < 2\pi \int_1^\infty \frac{1}{x}\sqrt{1 + 1/x^4}\,dx.$$

By Key Idea 31, the improper integral on the left diverges. Since the integral on the right is larger, we conclude it also diverges, meaning Gabriel's Horn has infinite surface area.

Hence the "paradox": we can fill Gabriel's Horn with a finite amount of paint, but since it has infinite surface area, we can never paint it.

Somehow this paradox is striking when we think about it in terms of volume and area. However, we have seen a similar paradox before, as referenced above. We know that the area under the curve $y = 1/x^2$ on $[1, \infty)$ is finite, yet the shape has an infinite perimeter. Strange things can occur when we deal with the infinite.

Notes:

Exercises 10.1

Terms and Concepts

1. T/F: The integral formula for computing Arc Length was found by first approximating arc length with straight line segments.

2. T/F: The integral formula for computing Arc Length includes a square–root, meaning the integration is probably easy.

Problems

In Exercises 3–12, find the arc length of the function on the given interval.

3. $f(x) = x$ on $[0, 1]$.

4. $f(x) = \sqrt{8}x$ on $[-1, 1]$.

5. $f(x) = \dfrac{1}{3}x^{3/2} - x^{1/2}$ on $[0, 1]$.

6. $f(x) = \dfrac{1}{12}x^3 + \dfrac{1}{x}$ on $[1, 4]$.

7. $f(x) = 2x^{3/2} - \dfrac{1}{6}\sqrt{x}$ on $[0, 9]$.

8. $f(x) = \cosh x$ on $[-\ln 2, \ln 2]$.

9. $f(x) = \dfrac{1}{2}\left(e^x + e^{-x}\right)$ on $[0, \ln 5]$.

10. $f(x) = \dfrac{1}{12}x^5 + \dfrac{1}{5x^3}$ on $[.1, 1]$.

11. $f(x) = \ln\left(\sin x\right)$ on $[\pi/6, \pi/2]$.

12. $f(x) = \ln\left(\cos x\right)$ on $[0, \pi/4]$.

In Exercises 13–20, set up the integral to compute the arc length of the function on the given interval. Do not evaluate the integral.

13. $f(x) = x^2$ on $[0, 1]$.

14. $f(x) = x^{10}$ on $[0, 1]$.

15. $f(x) = \sqrt{x}$ on $[0, 1]$.

16. $f(x) = \ln x$ on $[1, e]$.

17. $f(x) = \sqrt{1 - x^2}$ on $[-1, 1]$. (Note: this describes the top half of a circle with radius 1.)

18. $f(x) = \sqrt{1 - x^2/9}$ on $[-3, 3]$. (Note: this describes the top half of an ellipse with a major axis of length 6 and a minor axis of length 2.)

19. $f(x) = \dfrac{1}{x}$ on $[1, 2]$.

20. $f(x) = \sec x$ on $[-\pi/4, \pi/4]$.

In Exercises 21–28, use Simpson's Rule, with $n = 4$, to approximate the arc length of the function on the given interval. Note: these are the same problems as in Exercises 13–20.

21. $f(x) = x^2$ on $[0, 1]$.

22. $f(x) = x^{10}$ on $[0, 1]$.

23. $f(x) = \sqrt{x}$ on $[0, 1]$. (Note: $f'(x)$ is not defined at $x = 0$.)

24. $f(x) = \ln x$ on $[1, e]$.

25. $f(x) = \sqrt{1 - x^2}$ on $[-1, 1]$. (Note: $f'(x)$ is not defined at the endpoints.)

26. $f(x) = \sqrt{1 - x^2/9}$ on $[-3, 3]$. (Note: $f'(x)$ is not defined at the endpoints.)

27. $f(x) = \dfrac{1}{x}$ on $[1, 2]$.

28. $f(x) = \sec x$ on $[-\pi/4, \pi/4]$.

In Exercises 29–33, find the surface area of the described solid of revolution.

29. The solid formed by revolving $y = 2x$ on $[0, 1]$ about the x-axis.

30. The solid formed by revolving $y = x^2$ on $[0, 1]$ about the y-axis.

31. The solid formed by revolving $y = x^3$ on $[0, 1]$ about the x-axis.

32. The solid formed by revolving $y = \sqrt{x}$ on $[0, 1]$ about the x-axis.

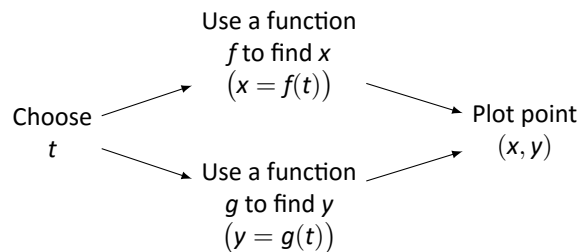33. The sphere formed by revolving $y = \sqrt{1 - x^2}$ on $[-1, 1]$ about the x-axis.

10.2 Parametric Equations

We are familiar with sketching shapes, such as parabolas, by following this basic procedure:

$$\text{Choose } x \quad \xrightarrow{\quad} \quad \begin{array}{c} \text{Use a function} \\ f \text{ to find } y \\ \left(y = f(x)\right) \end{array} \quad \xrightarrow{\quad} \quad \begin{array}{c} \text{Plot point} \\ (x, y) \end{array}$$
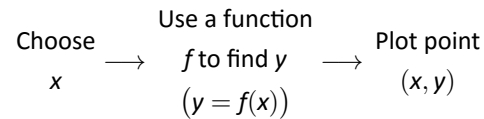
In the rectangular coordinate system, the **rectangular equation** $y = f(x)$ works well for some shapes like a parabola with a vertical axis of symmetry, but in precalculus and the review of conic sections in Section 10.0, we encountered several shapes that could not be sketched in this manner. (To plot an ellipse using the above procedure, we need to plot the "top" and "bottom" separately.)

In this section we introduce a new sketching procedure:

$$\text{Choose } t \quad \begin{array}{c} \nearrow \quad \begin{array}{c} \text{Use a function} \\ f \text{ to find } x \\ \left(x = f(t)\right) \end{array} \quad \searrow \\ \\ \searrow \quad \begin{array}{c} \text{Use a function} \\ g \text{ to find } y \\ \left(y = g(t)\right) \end{array} \quad \nearrow \end{array} \quad \begin{array}{c} \text{Plot point} \\ (x, y) \end{array}$$

Here, x and y are found separately but then plotted together. This leads us to a definition.

Definition 47　　Parametric Equations and Curves

Let f and g be continuous functions on an interval I. The **graph** of the **parametric equations** $x = f(t)$ and $y = g(t)$ is the set of all points $(x, y) = \left(f(t), g(t)\right)$ in the Cartesian plane, as the **parameter** t varies over I. A **curve** is a graph along with the parametric equations that define it.

Notes:

t	x	y
−2	4	−1
−1	1	0
0	0	1
1	1	2
2	4	3

(a)
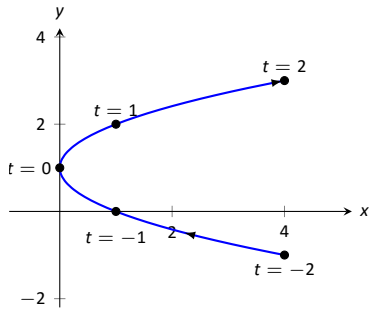
(b)

Figure 10.22: A table of values of the parametric equations in Example 1 along with a sketch of their graph.

t	x	y
0	1	2
$\pi/4$	1/2	$1 + \sqrt{2}/2$
$\pi/2$	0	1
$3\pi/4$	1/2	$1 - \sqrt{2}/2$
π	1	0

(a)
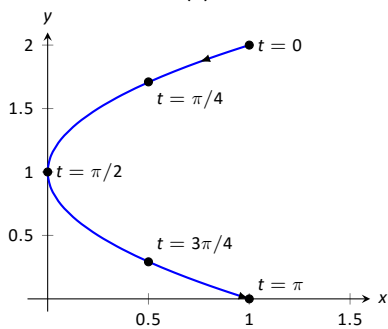
(b)

Figure 10.23: A table of values of the parametric equations in Example 2 along with a sketch of their graph.

Watch the video:
Parametric Equations — Some basic questions at
`https://youtu.be/9kKZHQtYP7g`

This is a formal definition of the word *curve*. When a curve lies in a plane (such as the Cartesian plane), it is often referred to as a **plane curve**. Examples will help us understand the concepts introduced in the definition.

Example 1   Plotting parametric functions
Plot the graph of the parametric equations $x = t^2$, $y = t + 1$ for t in $[-2, 2]$.

Solution   We plot the graphs of parametric equations in much the same manner as we plotted graphs of functions like $y = f(x)$: we make a table of values, plot points, then connect these points with a "reasonable" looking curve. Figure 10.22(a) shows such a table of values; note how we have 3 columns.

The points (x, y) from the table are plotted in Figure 10.22(b). The points have been connected with a smooth curve. Each point has been labeled with its corresponding t-value. These values, along with the two arrows along the curve, are used to indicate the **orientation** of the graph. This information describes the **path** of a particle traveling along the curve.

We often use the letter t as the parameter as we often regard t as representing *time*. Certainly there are many contexts in which the parameter is not time, but it can be helpful to think in terms of time as one makes sense of parametric plots and their orientation (for instance, "At time $t = 0$ the position is $(1, 2)$ and at time $t = 3$ the position is $(5, 1)$.").

Example 2   Plotting parametric functions
Sketch the graph of the parametric equations $x = \cos^2 t$, $y = \cos t + 1$ for t in $[0, \pi]$.

Solution   We again start by making a table of values in Figure 10.23(a), then plot the points (x, y) on the Cartesian plane in Figure 10.23(b).

The curves in Examples 1 and 2 are portions of the same parabola $(y - 1)^2 + x = 1$. While the *parabola* is the same, the *curves* are different. In Example 1, if we let t vary over all real numbers, we'd obtain the entire parabola. In this example, letting t vary over all real numbers would still produce the same graph; this portion of the parabola would be traced, and re–traced, infinitely often. The orientation shown in Figure 10.23 shows the orientation on $[0, \pi]$, but this orientation is reversed on $[\pi, 2\pi]$.

Notes:

Converting between rectangular and parametric equations

It is sometimes useful to transform rectangular form equations (i.e., $y = f(x)$) into parametric form equations, and vice–versa. Converting from rectangular to parametric can be very simple: given $y = f(x)$, the parametric equations $x = t$, $y = f(t)$ produce the same graph. As an example, given $y = x^2 - x - 6$, the parametric equations $x = t, y = t^2 - t - 6$ produce the same parabola. However, other parameterizations can be used. The following example demonstrates one possible alternative.

Example 3   **Converting from rectangular to parametric**
Find parametric equations for $f(x) = x^2 - x - 6$.

SOLUTION   Solution 1: For any choice for x we can determine the corresponding y by substitution. If we choose $x = t - 1$ then $y = (t-1)^2 - (t-1) - 6 = t^2 - 3t - 4$. Thus $f(x)$ can be represented by the parametric equations

$$x = t - 1 \qquad y = t^2 - 3t - 4.$$

On the graph of this parameterization (Figure 10.24(a)) the points have been labeled with the corresponding t-values and arrows indicate the path of a particle traveling on this curve. The particle would move from the upper left, down to the vertex at $(.5, -5.75)$ and then up to the right.

Solution 2: If we choose $x = 3 - t$ then $y = (3-t)^2 - (3-t) - 6 = t^2 - 5t$. Thus $f(x)$ can also be represented by the parametric equations

$$x = 3 - t \qquad y = t^2 - 5t.$$

On the graph of this parameterization (Figure 10.24(b)) the points have been labeled with the corresponding t−values and arrows indicate the path of a particle traveling on this curve. The particle would move down from the upper right, to the vertex at $(.5, -5.75)$ and then up to the left.

Solution 3: We can also parameterize any $y = f(x)$ by setting $t = \frac{dy}{dx}$. That is, $t = a$ corresponds to the point on the graph whose tangent line has a slope a. Computing $\frac{dy}{dx}, f'(x) = 2x - 1$ we set $t = 2x - 1$. Solving for x we find $x = \frac{t+1}{2}$ and by substitution $y = \frac{1}{4}t^2 - \frac{25}{4}$. Thus $f(x)$ can be represented by the parametric equations

$$x = \frac{t+1}{2} \qquad y = \frac{1}{4}t^2 - \frac{25}{4}.$$

The graph of this parameterization is shown in Figure 10.24(c). To find the point where the tangent line has a slope of 0, we set $t = 0$. This gives us the point $(-.5, -5.75)$ which is the vertex of $f(x)$.

(a)

(b)

(c)
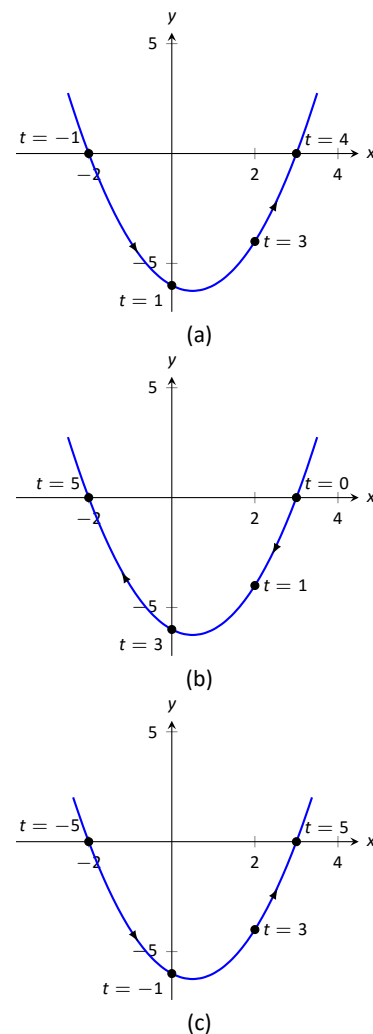
Figure 10.24: The equation $f(x) = x^2 - x - 6$ with different parameterizations.

Notes:

Example 4   **Converting from rectangular to parametric**
Find parametric equations for the circle $x^2 + y^2 = 4$.

SOLUTION   We will present three different approaches:

Solution 1: Consider the equivalent equation $\left(\dfrac{x}{2}\right)^2 + \left(\dfrac{y}{2}\right)^2 = 1$ and the Pythagorean Identity, $\sin^2 t + \cos^2 t = 1$. We set $\cos t = \frac{x}{2}$ and $\sin t = \frac{y}{2}$, which gives $x = 2\cos t$ and $y = 2\sin t$. To trace the circle once, we must have $0 \le t \le 2\pi$. Note that when $t = 0$ a particle tracing the curve would be at the point $(2, 0)$ and would move in a counterclockwise direction.

Solution 2: Another parameterization of the same circle would be $x = 2\sin t$ and $y = 2\cos t$ for $0 \le t \le 2\pi$. When $t = 0$ a particle would be at the point $(0, 2)$ and would move in a clockwise direction.

Solution 3: We could let $x = -2\sin t$ and $y = 2\cos t$ for $0 \le t \le 2\pi$. Also note that we could use $x = 2\cos 2t$ and $y = 2\sin 2t$ for $0 \le t \le \pi$.

As we have shown in the previous examples, there are many different ways to parameterize any given curve. We sometimes choose the parameter to accurately model physical behavior.

Example 5   **Converting from rectangular to parametric**
Find a parameterization that traces the ellipse $\dfrac{(x-2)^2}{9} + \dfrac{(y+3)^2}{4} = 1$ starting at the point $(-1, -3)$ in a clockwise direction.

SOLUTION   Applying the Pythagorean Identity, $\cos^2 t + \sin^2 t = 1$, we set $\cos^2 t = \dfrac{(x-2)^2}{9}$ and $\sin^2 t = \dfrac{(y+3)^2}{4}$. Solving these equations for x and y we set $x = -3\cos t + 2$ and $y = 2\sin t - 3$ for $0 \le t \le 2\pi$.

Example 6   **Converting from rectangular to parametric**
Find a parameterization for the hyperbola $\dfrac{(x-2)^2}{9} - \dfrac{(y-3)^2}{4} = 1$.

SOLUTION   We use an alternative form of the Pythagorean Identity, $\sec^2 t - \tan^2 t = 1$. We let $\sec^2 t = \dfrac{(x-2)^2}{9}$ and $\tan^2 t = \dfrac{(y-3)^2}{4}$. Solving these equations for x and y we have $x = 3\sec t + 2$ and $y = 2\tan t + 3$ for $0 \le t \le 2\pi$.

Notes:

Example 7   **Converting from rectangular to parametric**

An object is fired from a height of 0ft and lands 6 seconds later, 192ft away. Assuming ideal projectile motion, the height, in feet, of the object can be described by $h(x) = -x^2/64 + 3x$, where x is the distance in feet from the initial location. (Thus $h(0) = h(192) = 0$ft.) Find parametric equations $x = f(t)$, $y = g(t)$ for the path of the projectile where x is the horizontal distance the object has traveled at time t (in seconds) and y is the height at time t.

SOLUTION   Physics tells us that the horizontal motion of the projectile is linear; that is, the horizontal speed of the projectile is constant. Since the object travels 192ft in 6s, we deduce that the object is moving horizontally at a rate of 32ft/s, giving the equation $x = 32t$. As $y = -x^2/64 + 3x$, we find $y = -16t^2 + 96t$. We can quickly verify that $y'' = -32$ft/s^2, the acceleration due to gravity, and that the projectile reaches its maximum at $t = 3$, halfway along its path.

These parametric equations make certain determinations about the object's location easy: 2 seconds into the flight the object is at the point $(x(2), y(2)) = (64, 128)$. That is, it has traveled horizontally 64ft and is at a height of 128ft, as shown in Figure 10.25.

It is sometimes necessary to convert given parametric equations into rectangular form. This can be decidedly more difficult, as some "simple" looking parametric equations can have very "complicated" rectangular equations. This conversion is often referred to as "eliminating the parameter," as we are looking for a relationship between x and y that does not involve the parameter t.
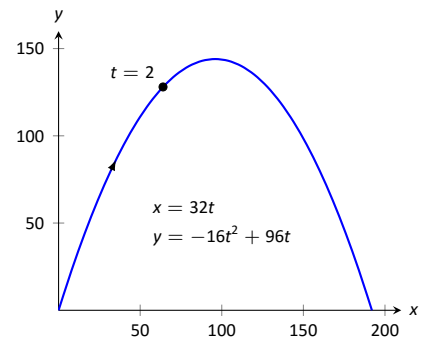
Figure 10.25: Graphing projectile motion in Example 7.

Example 8   **Eliminating the parameter**

Find a rectangular equation for the curve described by

$$x = \frac{1}{t^2 + 1} \quad \text{and} \quad y = \frac{t^2}{t^2 + 1}.$$

SOLUTION   There is not a set way to eliminate a parameter. One method is to solve for t in one equation and then substitute that value in the second. We use that technique here, then show a second, simpler method.

Starting with $x = 1/(t^2 + 1)$, solve for t: $t = \pm\sqrt{1/x - 1}$. Substitute this

---

Notes:

value for t in the equation for y:

$$y = \frac{t^2}{t^2+1}$$
$$= \frac{1/x - 1}{1/x - 1 + 1}$$
$$= \frac{1/x - 1}{1/x}$$
$$= \left(\frac{1}{x} - 1\right) \cdot x$$
$$= 1 - x.$$

Thus $y = 1 - x$. One may have recognized this earlier by manipulating the equation for y:

$$y = \frac{t^2}{t^2+1} = 1 - \frac{1}{t^2+1} = 1 - x.$$

This is a shortcut that is very specific to this problem; sometimes shortcuts exist and are worth looking for.

We should be careful to limit the domain of the function $y = 1 - x$. The parametric equations limit x to values in $(0, 1]$, thus to produce the same graph we should limit the domain of $y = 1 - x$ to the same.

The graphs of these functions are given in Figure 10.26. The portion of the graph defined by the parametric equations is given in a thick line; the graph defined by $y = 1 - x$ with unrestricted domain is given in a thin line.

Figure 10.26: Graphing parametric and rectangular equations for a graph in Example 8.

Example 9   **Eliminating the parameter**
Eliminate the parameter in $x = 4\cos t + 3$, $y = 2\sin t + 1$

SOLUTION   We should not try to solve for t in this situation as the resulting algebra/trig would be messy. Rather, we solve for $\cos t$ and $\sin t$ in each equation, respectively. This gives

$$\cos t = \frac{x - 3}{4} \quad \text{and} \quad \sin t = \frac{y - 1}{2}.$$

The Pythagorean Theorem gives $\cos^2 t + \sin^2 t = 1$, so:

$$\cos^2 t + \sin^2 t = 1$$
$$\left(\frac{x-3}{4}\right)^2 + \left(\frac{y-1}{2}\right)^2 = 1$$
$$\frac{(x-3)^2}{16} + \frac{(y-1)^2}{4} = 1$$

Notes:

This final equation should look familiar — it is the equation of an ellipse. Figure 10.27 plots the parametric equations, demonstrating that the graph is indeed of an ellipse with a horizontal major axis and center at $(3, 1)$.

Graphs of Parametric Equations

These examples begin to illustrate the powerful nature of parametric equations. Their graphs are far more diverse than the graphs of functions produced by "$y = f(x)$" functions.

One nice feature of parametric equations is that their graphs are easy to shift. While this is not too difficult in the "$y = f(x)$" context, the resulting function can look rather messy. (Plus, to shift to the right by two, we replace x with $x - 2$, which is counterintuitive.) The following example demonstrates this.

Figure 10.27: Graphing the parametric equations $x = 4\cos t + 3$, $y = 2\sin t + 1$ in Example 9.

Example 10   **Shifting the graph of parametric functions**
Sketch the graph of the parametric equations $x = t^2 + t$, $y = t^2 - t$. Find new parametric equations that shift this graph to the right 3 units and down 2.

SOLUTION   The graph of the parametric equations is given in Figure 10.28 (a). It is a parabola with an axis of symmetry along the line $y = x$; the vertex is at $(0, 0)$. It should be noted that finding the vertex is not a trivial matter and not something you will be asked to do in this text.

In order to shift the graph to the right 3 units, we need to increase the x-value by 3 for every point. The straightforward way to accomplish this is simply to add 3 to the function defining x: $x = t^2 + t + 3$. To shift the graph down by 2 units, we wish to decrease each y-value by 2, so we subtract 2 from the function defining y: $y = t^2 - t - 2$. Thus our parametric equations for the shifted graph are $x = t^2 + t + 3$, $y = t^2 - t - 2$. This is graphed in Figure 10.28 (b). Notice how the vertex is now at $(3, -2)$.

Because the x- and y-values of a graph are determined independently, the graphs of parametric functions often possess features not seen on "$y = f(x)$" type graphs. The next example demonstrates how such graphs can arrive at the same point more than once.

Example 11   **Graphs that cross themselves**
Plot the parametric functions $x = t^3 - 5t^2 + 3t + 11$ and $y = t^2 - 2t + 3$ and determine the t-values where the graph crosses itself.

(a)

(b)

Figure 10.28: Illustrating how to shift graphs in Example 10.

Notes:

SOLUTION   Using the methods developed in this section, we again plot points and graph the parametric equations as shown in Figure 10.29. It appears that the graph crosses itself at the point $(2, 6)$, but we'll need to analytically determine this.

We are looking for two different values, say, s and t, where $x(s) = x(t)$ and $y(s) = y(t)$. That is, the x-values are the same precisely when the y-values are the same. This gives us a system of 2 equations with 2 unknowns:
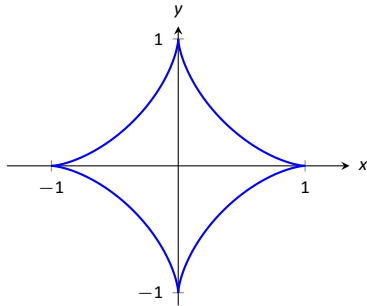
$$s^3 - 5s^2 + 3s + 11 = t^3 - 5t^2 + 3t + 11$$
$$s^2 - 2s + 3 = t^2 - 2t + 3$$

Solving this system is not trivial but involves only algebra. Using the quadratic formula, one can solve for t in the second equation and find that $t = 1 \pm \sqrt{s^2 - 2s + 1}$. This can be substituted into the first equation, revealing that the graph crosses itself at $t = -1$ and $t = 3$. We confirm our result by computing $x(-1) = x(3) = 2$ and $y(-1) = y(3) = 6$.

We now present a small gallery of "interesting" and "famous" curves along with parametric equations that produce them.

Figure 10.29: A graph of the parametric equations from Example 11.

The figure shows the curve with equations $x = t^3 - 5t^2 + 3t + 11$ and $y = t^2 - 2t + 3$.

Notes:

Astroid
$x = \cos^3 t$
$y = \sin^3 t$
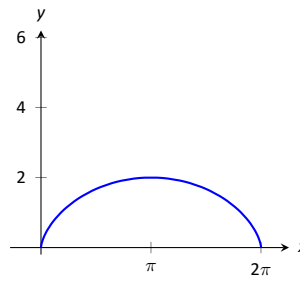
Cycloid
$x = r(t - \sin t)$
$y = r(1 - \cos t)$
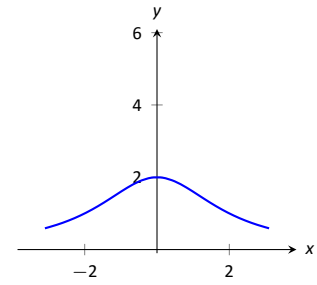
Witch of Agnesi
$x = 2at$
$y = 2a/(1 + t^2)$
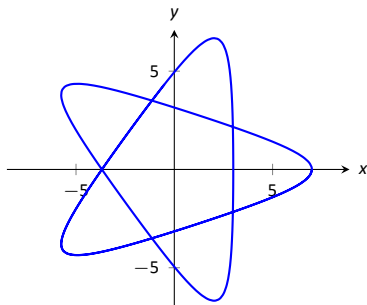
Hypotrochoid
$x = 2\cos(t) + 5\cos(2t/3)$
$y = 2\sin(t) - 5\sin(2t/3)$
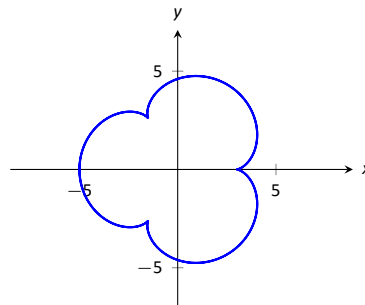
Epicycloid
$x = 4\cos(t) - \cos(4t)$
$y = 4\sin(t) - \sin(4t)$
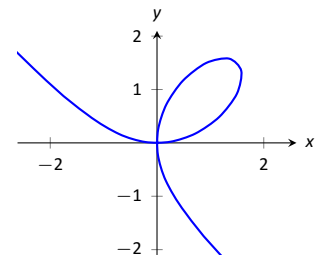
Folium of Descartes
$x = 3at/(1 + t^3)$
$y = 3at^2/(1 + t^3)$

One might note a feature shared by three of these graphs: "sharp corners," or **cusps**. We have seen graphs with cusps before and determined that such functions are not differentiable at these points. This leads us to a definition.

> **Definition 48   Smooth**
> A curve C defined by $x = f(t), y = g(t)$ is **smooth** on an interval I if f' and g' are continuous on I and not simultaneously 0 (except possibly at the endpoints of I). A curve is **piecewise smooth** on I if I can be partitioned into subintervals where C is smooth on each subinterval.

Consider the astroid, given by $x = \cos^3 t$, $y = \sin^3 t$. Taking derivatives, we have:
$$x' = -3\cos^2 t \sin t \quad \text{and} \quad y' = 3\sin^2 t \cos t.$$

Notes:

It is clear that each is 0 when $t = 0$, $\pi/2$, $\pi, \ldots$. Thus the astroid is not smooth at these points, corresponding to the cusps seen in the figure. However, by restricting the domain of the astroid to all reals except $t = \frac{k\pi}{2}$ for $k \in \mathbb{Z}$ we have a piecewise smooth curve.

We demonstrate this once more.

Example 12   **Determine where a curve is not smooth**
Let a curve C be defined by the parametric equations $x = t^3 - 12t + 17$ and $y = t^2 - 4t + 8$. Determine the points, if any, where it is not smooth.

     SOLUTION   We begin by taking derivatives.

$$x' = 3t^2 - 12, \quad y' = 2t - 4.$$

We set each equal to 0:

$$x' = 0 \Rightarrow 3t^2 - 12 = 0 \Rightarrow t = \pm 2$$
$$y' = 0 \Rightarrow 2t - 4 = 0 \Rightarrow t = 2$$

We consider only the value of $t = 2$ since both x' and y' must be 0. Thus C is not smooth at $t = 2$, corresponding to the point $(1, 4)$. The curve is graphed in Figure 10.30, illustrating the cusp at $(1, 4)$.
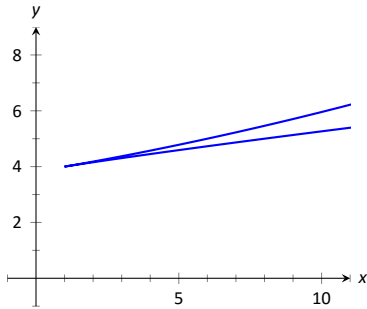
Figure 10.30: Graphing the curve in Example 12; note it is not smooth at $(1, 4)$.

If a curve is not smooth at $t = t_0$, it means that $x'(t_0) = y'(t_0) = 0$ as defined. This, in turn, means that rate of change of x (and y) is 0; that is, at that instant, neither x nor y is changing. If the parametric equations describe the path of some object, this means the object is at rest at t_0. An object at rest can make a "sharp" change in direction, whereas moving objects tend to change direction in a "smooth" fashion.

Example 13   **The Cycloid**
A well-known parametric curve is the cycloid. Fix r, and let $x = r(t - \sin t)$, $y = r(1 - \cos t)$. This represents the path traced out by a point on a wheel of radius r as starts rolling to the right. We can think of t as the angle through which the point has rotated.
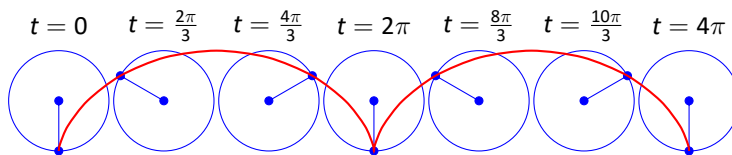
Figure 10.31: A cycloid traced through two revolutions.

Notes:

Figure 10.31 shows a cycloid sketched out with the wheel shown at various places. The dot on the rim is the point on the wheel that we're using to trace out the curve.

From this sketch we can see that one arch of the cycloid is traced out in the range $0 \leq t \leq 2\pi$. This makes sense when you consider that the point will be back on the ground after it has rotated through an angle of 2π.

One should be careful to note that a "sharp corner" does not have to occur when a curve is not smooth. For instance, one can verify that $x = t^3$, $y = t^6$ produce the familiar $y = x^2$ parabola. However, in this parametrization, the curve is not smooth. A particle traveling along the parabola according to the given parametric equations comes to rest at $t = 0$, though no sharp point is created.

Our previous experience with cusps taught us that a function was not differentiable at a cusp. This can lead us to wonder about derivatives in the context of parametric equations and the application of other calculus concepts. Given a curve defined parametrically, how do we find the slopes of tangent lines? Can we determine concavity? We explore these concepts and more in the next section.

Notes:

Exercises 10.2

Terms and Concepts

1. T/F: When sketching the graph of parametric equations, the x and y values are found separately, then plotted together.

2. The direction in which a graph is "moving" is called the _____ of the graph.

3. An equation written as $y = f(x)$ is written in _____ form.

4. Create parametric equations $x = f(t)$, $y = g(t)$ and sketch their graph. Explain any interesting features of your graph based on the functions f and g.

Problems

In Exercises 5–8, sketch the graph of the given parametric equations **by hand**, making a table of points to plot. Be sure to indicate the orientation of the graph.

5. $x = t^2 + t$, $\quad y = 1 - t^2$, $\quad -3 \le t \le 3$

6. $x = 1$, $\quad y = 5 \sin t$, $\quad -\pi/2 \le t \le \pi/2$

7. $x = t^2$, $\quad y = 2$, $\quad -2 \le t \le 2$

8. $x = t^3 - t + 3$, $\quad y = t^2 + 1$, $\quad -2 \le t \le 2$

In Exercises 9–17, sketch the graph of the given parametric equations; using a graphing utility is advisable. Be sure to indicate the orientation of the graph.

9. $x = t^3 - 2t^2$, $\quad y = t^2$, $\quad -2 \le t \le 3$

10. $x = 1/t$, $\quad y = \sin t$, $\quad 0 < t \le 10$

11. $x = 3 \cos t$, $\quad y = 5 \sin t$, $\quad 0 \le t \le 2\pi$

12. $x = 3 \cos t + 2$, $\quad y = 5 \sin t + 3$, $\quad 0 \le t \le 2\pi$

13. $x = \cos t$, $\quad y = \cos(2t)$, $\quad 0 \le t \le \pi$

14. $x = \cos t$, $\quad y = \sin(2t)$, $\quad 0 \le t \le 2\pi$

15. $x = 2 \sec t$, $\quad y = 3 \tan t$, $\quad -\pi/2 < t < \pi/2$

16. $x = \cos t + \frac{1}{4} \cos(8t)$, $\quad y = \sin t + \frac{1}{4} \sin(8t)$, $\quad 0 \le t \le 2\pi$

17. $x = \cos t + \frac{1}{4} \sin(8t)$, $\quad y = \sin t + \frac{1}{4} \cos(8t)$, $\quad 0 \le t \le 2\pi$

In Exercises 18–19, four sets of parametric equations are given. Describe how their graphs are similar and different. Be sure to discuss orientation and ranges.

18.
    (a) $x = t \quad y = t^2$, $\quad -\infty < t < \infty$
    (b) $x = \sin t \quad y = \sin^2 t$, $\quad -\infty < t < \infty$
    (c) $x = e^t \quad y = e^{2t}$, $\quad -\infty < t < \infty$
    (d) $x = -t \quad y = t^2$, $\quad -\infty < t < \infty$

19.
    (a) $x = \cos t \quad y = \sin t$, $\quad 0 \le t \le 2\pi$
    (b) $x = \cos(t^2) \quad y = \sin(t^2)$, $\quad 0 \le t \le 2\pi$
    (c) $x = \cos(1/t) \quad y = \sin(1/t)$, $\quad 0 < t < 1$
    (d) $x = \cos(\cos t) \quad y = \sin(\cos t)$, $\quad 0 \le t \le 2\pi$

In Exercises 20–22, find a parameterization for the curve.

20. $y = 9 - 4x$

21. $4x - y^2 = 5$

22. $(x + 9)^2 + (y - 4)^2 = 49$

In Exercises 23–26, find a parametric equation and a parameter interval.

23. The line segment with endpoints $(-1, -3)$ and $(4, 1)$

24. The line segment with endpoints $(-1, 3)$ and $(3, -2)$

25. The left half of the parabola $y = x^2 + 2x$

26. The lower half of the parabola $x = 1 - y^2$

In Exercises 27–30, find parametric equations for the given rectangular equation using the parameter $t = \dfrac{dy}{dx}$. Verify that at $t = 1$, the point on the graph has a tangent line with slope of 1.

27. $y = 3x^2 - 11x + 2$

28. $y = e^x$

29. $y = \sin x$ on $[0, \pi]$

30. $y = \sqrt{x}$ on $[0, \infty)$

31. Find parametric equations and a parameter interval for the motion of a particle that starts at $(1, 0)$ and traces the circle $x^2 + y^2 = 1$
    (a) once clockwise
    (b) once counter-clockwise
    (c) twice clockwise
    (d) twice counter-clockwise

32. Find parametric equations and a parameter interval for the motion of a particle that starts at $(a, 0)$ and traces the ellipse $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$
    (a) once clockwise
    (b) once counter-clockwise
    (c) twice clockwise
    (d) twice counter-clockwise

In Exercises 33–41, find parametric equations that describe the given situation.

33. A projectile is fired from a height of 0ft, landing 16ft away in 4s.

34. A projectile is fired from a height of 0ft, landing 200ft away in 4s.

35. A projectile is fired from a height of 0ft, landing 200ft away in 20s.

36. A circle of radius 2, centered at the origin, that is traced clockwise once on $[0, 2\pi]$.

37. A circle of radius 3, centered at $(1, 1)$, that is traced once counter–clockwise on $[0, 1]$.

38. An ellipse centered at $(1, 3)$ with vertical major axis of length 6 and minor axis of length 2.

39. An ellipse with foci at $(\pm 1, 0)$ and vertices at $(\pm 5, 0)$.

40. A hyperbola with foci at $(5, -3)$ and $(-1, -3)$, and with vertices at $(1, -3)$ and $(3, -3)$.

41. A hyperbola with vertices at $(0, \pm 6)$ and asymptotes $y = \pm 3x$.

In Exercises 42–51, eliminate the parameter in the given parametric equations.

42. $x = 2t + 5, \quad y = -3t + 1$

43. $x = \sec t, \quad y = \tan t$

44. $x = 4 \sin t + 1, \quad y = 3 \cos t - 2$

45. $x = t^2, \quad y = t^3$

46. $x = \dfrac{1}{t + 1}, \quad y = \dfrac{3t + 5}{t + 1}$

47. $x = e^t, \quad y = e^{3t} - 3$

48. $x = \ln t, \quad y = t^2 - 1$

49. $x = \cot t, \quad y = \csc t$

50. $x = \cosh t, \quad y = \sinh t$

51. $x = \cos(2t), \quad y = \sin t$

In Exercises 52–55, eliminate the parameter in the given parametric equations. Describe the curve defined by the parametric equations based on its rectangular form.

52. $x = at + x_0, \quad y = bt + y_0$

53. $x = r \cos t, \quad y = r \sin t$

54. $x = a \cos t + h, \quad y = b \sin t + k$

55. $x = a \sec t + h, \quad y = b \tan t + k$

In Exercises 56–59, find the values of t where the graph of the parametric equations crosses itself.

56. $x = t^3 - t + 3, \quad y = t^2 - 3$

57. $x = t^3 - 4t^2 + t + 7, \quad y = t^2 - t$

58. $x = \cos t, \quad y = \sin(2t)$ on $[0, 2\pi]$

59. $x = \cos t \cos(3t), \quad y = \sin t \cos(3t)$ on $[0, \pi]$

In Exercises 60–63, find the value(s) of t where the curve defined by the parametric equations is not smooth.

60. $x = t^3 + t^2 - t, \quad y = t^2 + 2t + 3$

61. $x = t^2 - 4t, \quad y = t^3 - 2t^2 - 4t$

62. $x = \cos t, \quad y = 2 \cos t$

63. $x = 2 \cos t - \cos(2t), \quad y = 2 \sin t - \sin(2t)$

10.3   Calculus and Parametric Equations

The previous section defined curves based on parametric equations. In this section we'll employ the techniques of calculus to study these curves.

We are still interested in lines tangent to points on a curve. They describe how the *y*-values are changing with respect to the *x*-values, they are useful in making approximations, and they indicate instantaneous direction of travel.

The slope of the tangent line is still $\frac{dy}{dx}$, and the Chain Rule allows us to calculate this in the context of parametric equations. If $x = f(t)$ and $y = g(t)$, the Chain Rule states that

$$\frac{dy}{dt} = \frac{dy}{dx} \cdot \frac{dx}{dt}.$$

Solving for $\frac{dy}{dx}$, we get

$$\frac{dy}{dx} = \frac{dy/dt}{dx/dt} = \frac{g'(t)}{f'(t)},$$

provided that $f'(t) \neq 0$. This is important so we label it a Key Idea.

Key Idea 39   Finding $\frac{dy}{dx}$ with Parametric Equations.

Let $x = f(t)$ and $y = g(t)$, where f and g are differentiable on some open interval I and $f'(t) \neq 0$ on I. Then

$$\frac{dy}{dx} = \frac{dy/dt}{dx/dt} = \frac{g'(t)}{f'(t)}.$$

We use this to define the tangent line.

Definition 49   Tangent Lines

Let a curve C be parameterized by $x = f(t)$ and $y = g(t)$, where f and g are differentiable functions on some interval I containing $t = t_0$. The **tangent line** to C at $t = t_0$ is the line through $(f(t_0), g(t_0))$ with slope $m = \frac{g'(t_0)}{f'(t_0)}$, provided $f'(t_0) \neq 0$.

It is possible for parametric curves to have horizontal and vertical tangents. As expected a horizontal tangent occurs whenever $\frac{dy}{dx} = 0$ or when $\frac{dy}{dt} = 0$ (provided $\frac{dx}{dt} \neq 0$). Similarly, a vertical tangent occurs whenever $\frac{dy}{dx}$ is undefined or when $\frac{dx}{dt} = 0$ (provided $\frac{dy}{dt} \neq 0$).

Notes:

> **Definition 50 Normal Lines**
>
> The **normal line** to a curve C at a point P is the line through P and perpendicular to the tangent line at P. For $t = t_0$ the normal line is the line through $(f(t_0), g(t_0))$ with slope $m = -\frac{f'(t_0)}{g'(t_0)}$, provided $g'(t_0) \neq 0$.

As with the tangent line we note that it is possible for a normal line to be vertical or horizontal. A horizontal normal line occurs whenever $\frac{dy}{dx}$ is undefined or when $\frac{dx}{dt} = 0$ (provided $\frac{dy}{dt} \neq 0$). Similarly, a vertical normal line occurs whenever $\frac{dy}{dx} = 0$ or when $\frac{dy}{dt} = 0$ (provided $\frac{dx}{dt} \neq 0$). In other words, if the curve C has a vertical tangent at $(f(t_0), g(t_0))$ the normal line will be horizontal and if the tangent is horizontal the normal line will be a vertical line.

Watch the video:
Derivatives of Parametric Functions at
`https://youtu.be/k5QnaGVk1JI`

Example 1   Tangent and Normal Lines to Curves
Let $x = 5t^2 - 6t + 4$ and $y = t^2 + 6t - 1$, and let C be the curve defined by these equations.

1. Find the equations of the tangent and normal lines to C at $t = 3$.

2. Find where C has vertical and horizontal tangent lines.

SOLUTION

1. We start by computing $f'(t) = 10t - 6$ and $g'(t) = 2t + 6$. Thus

$$\frac{dy}{dx} = \frac{2t + 6}{10t - 6}.$$

   Make note of something that might seem unusual: $\frac{dy}{dx}$ is a function of t, not x. Just as points on the curve are found in terms of t, so are the slopes of the tangent lines.

   The point on C at $t = 3$ is $(31, 26)$. The slope of the tangent line is $m = 1/2$ and the slope of the normal line is $m = -2$. Thus,

Notes:
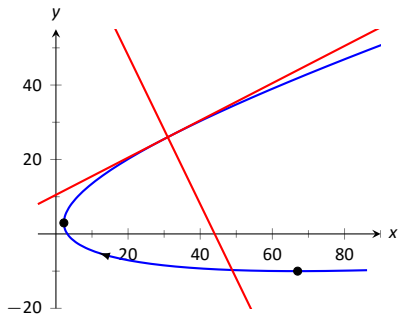
Figure 10.32: Graphing tangent and normal lines in Example 1.

- the equation of the tangent line is $y = \dfrac{1}{2}(x - 31) + 26$, and

- the equation of the normal line is $y = -2(x - 31) + 26$.

This is illustrated in Figure 10.32.

2. To find where C has a horizontal tangent line, we set $\frac{dy}{dx} = 0$ and solve for t. In this case, this amounts to setting $g'(t) = 0$ and solving for t (and making sure that $f'(t) \neq 0$).

$$g'(t) = 0 \quad \Rightarrow \quad 2t + 6 = 0 \quad \Rightarrow \quad t = -3.$$

The point on C corresponding to $t = -3$ is $(67, -10)$; the tangent line at that point is horizontal (hence with equation $y = -10$).

To find where C has a vertical tangent line, we find where it has a horizontal normal line, and set $-\frac{f'(t)}{g'(t)} = 0$. This amounts to setting $f'(t) = 0$ and solving for t (and making sure that $g'(t) \neq 0$).

$$f'(t) = 0 \quad \Rightarrow \quad 10t - 6 = 0 \quad \Rightarrow \quad t = 0.6.$$

The point on C corresponding to $t = 0.6$ is $(2.2, 2.96)$. The tangent line at that point is $x = 2.2$.

The points where the tangent lines are vertical and horizontal are indicated on the graph in Figure 10.32.

Example 2   **Tangent and Normal Lines to a Circle**

1. Find where the unit circle, defined by $x = \cos t$ and $y = \sin t$ on $[0, 2\pi]$, has vertical and horizontal tangent lines.

2. Find the equation of the normal line at $t = t_0$.

SOLUTION

1. We compute the derivative following Key Idea 39:

$$\frac{dy}{dx} = \frac{g'(t)}{f'(t)} = -\frac{\cos t}{\sin t}.$$

The derivative is 0 when $\cos t = 0$; that is, when $t = \pi/2, \ 3\pi/2$. These are the points $(0, 1)$ and $(0, -1)$ on the circle.

The normal line is horizontal (and hence, the tangent line is vertical) when $\sin t = 0$; that is, when $t = 0, \ \pi, \ 2\pi$, corresponding to the points $(-1, 0)$ and $(0, 1)$ on the circle. These results should make intuitive sense.

Notes:

2. The slope of the normal line at $t = t_0$ is $m = \dfrac{\sin t_0}{\cos t_0} = \tan t_0$. This normal line goes through the point $(\cos t_0, \sin t_0)$, giving the line

$$y = \frac{\sin t_0}{\cos t_0}(x - \cos t_0) + \sin t_0$$
$$= (\tan t_0)x,$$

as long as $\cos t_0 \neq 0$. It is an important fact to recognize that the normal lines to a circle pass through its center, as illustrated in Figure 10.33. Stated in another way, any line that passes through the center of a circle intersects the circle at right angles.
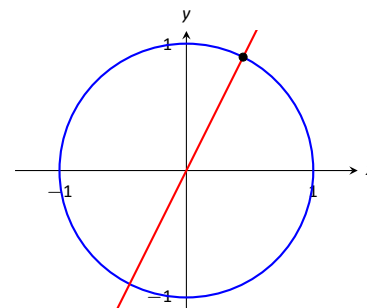
Figure 10.33: Illustrating how a circle's normal lines pass through its center.

Example 3 **Tangent lines when $\frac{dy}{dx}$ is not defined**

Find the equation of the tangent line to the astroid $x = \cos^3 t$, $y = \sin^3 t$ at $t = 0$, shown in Figure 10.34.

SOLUTION We start by finding $x'(t)$ and $y'(t)$:

$$x'(t) = -3 \sin t \cos^2 t, \qquad y'(t) = 3 \cos t \sin^2 t.$$

Note that both of these are 0 at $t = 0$; the curve is not smooth at $t = 0$ forming a cusp on the graph. Evaluating $\frac{dy}{dx}$ at this point returns the indeterminate form of "0/0".

We can, however, examine the slopes of tangent lines near $t = 0$, and take the limit as $t \to 0$.

$$\lim_{t \to 0} \frac{y'(t)}{x'(t)} = \lim_{t \to 0} \frac{3 \cos t \sin^2 t}{-3 \sin t \cos^2 t} \quad \text{(We can reduce as } t \neq 0.)$$
$$= \lim_{t \to 0} -\frac{\sin t}{\cos t}$$
$$= 0.$$

Figure 10.34: A graph of an astroid.

We have accomplished something significant. When the derivative $\frac{dy}{dx}$ returns an indeterminate form at $t = t_0$, we can define its value by setting it to be $\lim\limits_{t \to t_0} \dfrac{dy}{dx}$, if that limit exists. This allows us to find slopes of tangent lines at cusps, which can be very beneficial.

We found the slope of the tangent line at $t = 0$ to be 0; therefore the tangent line is $y = 0$, the x-axis.

Notes:

Concavity

We continue to analyze curves in the plane by considering their concavity; that is, we are interested in $\frac{d^2y}{dx^2}$, "the second derivative of y with respect to x." To find this, we need to find the derivative of $\frac{dy}{dx}$ with respect to x; that is,

$$\frac{d^2y}{dx^2} = \frac{d}{dx}\left[\frac{dy}{dx}\right],$$

but recall that $\frac{dy}{dx}$ is a function of t, not x, making this computation not straight-forward.

To make the upcoming notation a bit simpler, let $h(t) = \frac{dy}{dx}$. We want $\frac{d}{dx}[h(t)]$; that is, we want $\frac{dh}{dx}$. We again appeal to the Chain Rule. Note:

$$\frac{dh}{dt} = \frac{dh}{dx} \cdot \frac{dx}{dt} \quad \Rightarrow \quad \frac{dh}{dx} = \frac{dh/dt}{dx/dt}.$$

In words, to find $\frac{d^2y}{dx^2}$, we first take the derivative of $\frac{dy}{dx}$ *with respect to t*, then divide by $x'(t)$. We restate this as a Key Idea.

Key Idea 40   Finding $\frac{d^2y}{dx^2}$ with Parametric Equations

Let $x = f(t)$ and $y = g(t)$ be twice differentiable functions on an open interval I, where $f'(t) \neq 0$ on I. Then

$$\frac{d^2y}{dx^2} = \frac{\dfrac{d}{dt}\left[\dfrac{dy}{dx}\right]}{\dfrac{dx}{dt}} = \frac{\dfrac{d}{dt}\left[\dfrac{dy}{dx}\right]}{f'(t)}.$$

Examples will help us understand this Key Idea.

Example 4   Concavity of Plane Curves

Let $x = 5t^2 - 6t + 4$ and $y = t^2 + 6t - 1$ as in Example 1. Determine the t-intervals on which the graph is concave up/down.

SOLUTION   Concavity is determined by the second derivative of y with respect to x, $\frac{d^2y}{dx^2}$, so we compute that here following Key Idea 40.

Notes:

In Example 1, we found $\dfrac{dy}{dx} = \dfrac{2t+6}{10t-6}$ and $f'(t) = 10t - 6$. So:

$$\frac{d^2y}{dx^2} = \frac{\frac{d}{dt}\left[\frac{2t+6}{10t-6}\right]}{10t-6}$$

$$= \frac{-\frac{72}{(10t-6)^2}}{10t-6}$$

$$= -\frac{72}{(10t-6)^3}$$

$$= -\frac{9}{(5t-3)^3}$$

The graph of the parametric functions is concave up when $\frac{d^2y}{dx^2} > 0$ and concave down when $\frac{d^2y}{dx^2} < 0$. We determine the intervals when the second derivative is greater/less than 0 by first finding when it is 0 or undefined.

As the numerator of $-\dfrac{9}{(5t-3)^3}$ is never 0, $\frac{d^2y}{dx^2} \neq 0$ for all t. It is undefined when $5t - 3 = 0$; that is, when $t = 3/5$. Following the work established in Section 3.4, we look at values of t greater or less than $3/5$ on a number line:

Reviewing Example 1, we see that when $t = 3/5 = 0.6$, the graph of the parametric equations has a vertical tangent line. This point is also a point of inflection for the graph, illustrated in Figure 10.35.
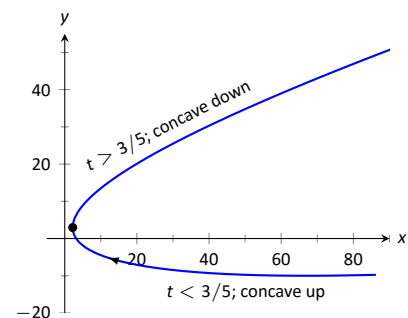
Figure 10.35: Graphing the parametric equations in Example 4 to demonstrate concavity.

Example 5   Concavity of Plane Curves
Find the points of inflection of the graph of the parametric equations $x = \sqrt{t}$, $y = \sin t$, for $0 \leq t \leq 16$.

SOLUTION   We need to compute $\frac{dy}{dx}$ and $\frac{d^2y}{dx^2}$.

$$\frac{dy}{dx} = \frac{y'(t)}{x'(t)} = \frac{\cos t}{1/(2\sqrt{t})} = 2\sqrt{t}\cos t.$$

$$\frac{d^2y}{dx^2} = \frac{\frac{d}{dt}\left[\frac{dy}{dx}\right]}{x'(t)} = \frac{\cos t/\sqrt{t} - 2\sqrt{t}\sin t}{1/(2\sqrt{t})} = 2\cos t - 4t\sin t.$$

Notes:
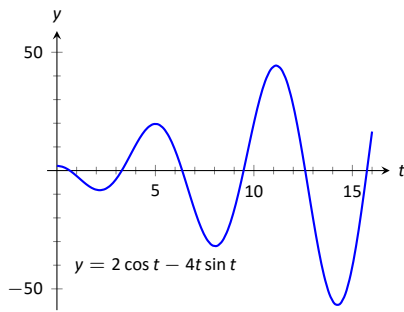
(a)
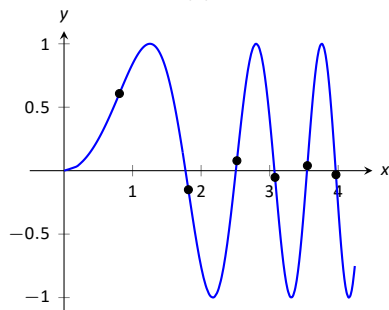
(b)

Figure 10.36: In (a), a graph of $\frac{d^2y}{dx^2}$, showing where it is approximately 0. In (b), graph of the parametric equations in Example 5 along with the points of inflection.

The possible points of inflection are found by setting $\frac{d^2y}{dx^2} = 0$. This is not trivial, as equations that mix polynomials and trigonometric functions generally do not have "nice" solutions.

In Figure 10.36(a) we see a plot of the second derivative. It shows that it has zeros at approximately $t = 0.5,\ 3.5,\ 6.5,\ 9.5,\ 12.5$ and 16. These approximations are not very good, made only by looking at the graph. Newton's Method provides more accurate approximations. Accurate to 2 decimal places, we have:

$$t = 0.65,\ 3.29,\ 6.36,\ 9.48,\ 12.61 \text{ and } 15.74.$$

The corresponding points have been plotted on the graph of the parametric equations in Figure 10.36(b). Note how most occur near the x-axis, but not exactly on the axis.

Area with Parametric Equations

We will now find a formula for determining the area under a parametric curve given by the parametric equations

$$x = f(t) \qquad y = g(t).$$

We will also need to further add in the assumption that the curve is traced out exactly once as t increases from α to β. First, recall how to find the area under $y = F(x)$ on $a \leq x \leq b$:

$$A = \int_a^b F(x)\,dx.$$

Now think of the parametric equation $x = f(t)$ as a substitution in the integral, assuming that $a = f(\alpha)$ and $b = f(\beta)$ for the purposes of this formula. (There is actually no reason to assume that this will always be the case and so we'll give a corresponding formula later if it's the opposite case ($b = f(\alpha)$ and $a = f(\beta)$).)

In order to substitute, we'll need $dx = f'(t)\,dt$. Plugging this into the area formula above and making sure to change the limits to their corresponding t values gives us

$$A = \int_\alpha^\beta F(f(t))f'(t)\,dt.$$

Since we don't know what $F(x)$ is, we'll use the fact that

$$y = F(x) = F(f(t)) = g(t)$$

and arrive at the formula that we want.

Notes:

Key Idea 41   Area Under a Parametric Curve

The area under the parametric curve given by $x = f(t)$, $y = g(t)$, for $f(\alpha) = a < x < b = f(\beta)$ is

$$A = \int_\alpha^\beta g(t)f'(t)\, dt.$$

On the other hand, if we should happen to have $b = f(\alpha)$ and $a = f(\beta)$, then the formula would be

$$A = \int_\beta^\alpha g(t)f'(t)\, dt.$$

Let's work an example.

Example 6   Finding the area under a parametric curve

Determine the area under the cycloid given by the parametric equations

$$x = 6(\theta - \sin\theta) \qquad y = 6(1 - \cos\theta) \qquad 0 \le \theta \le 2\pi.$$

Solution   First, notice that we've switched the parameter to θ for this problem. This is to make sure that we don't get too locked into always having t as the parameter.

Now, we could graph this to verify that the curve is traced out exactly once for the given range if we wanted to.

There really isn't too much to this example other than plugging the parametric equations into the formula. We'll first need the derivative of the parametric equation for x however.

$$\frac{dx}{d\theta} = 6(1 - \cos\theta).$$

The area is then

$$A = \int_0^{2\pi} 36(1 - \cos\theta)^2\, d\theta$$

$$= 36 \int_0^{2\pi} 1 - 2\cos\theta + \cos^2\theta\, d\theta$$

$$= 36 \int_0^{2\pi} \frac{3}{2} - 2\cos\theta + \frac{1}{2}\cos(2\theta)\, d\theta$$

$$= 36 \left[\frac{3}{2}\theta - 2\sin\theta + \frac{1}{4}\sin(2\theta) \right]_0^{2\pi}$$

$$= 108\pi.$$

Notes:

Arc Length

We continue our study of the features of the graphs of parametric equations by computing their arc length.

Recall in Section 10.1 we found the arc length of the graph of a function, from $x = a$ to $x = b$, to be

$$L = \int_a^b \sqrt{1 + \left(\frac{dy}{dx}\right)^2}\, dx.$$

We can use this equation and convert it to the parametric equation context. Letting $x = f(t)$ and $y = g(t)$, we know that $\frac{dy}{dx} = g'(t)/f'(t)$. It will also be useful to calculate the differential of x:

$$dx = f'(t)dt \qquad \Rightarrow \qquad dt = \frac{1}{f'(t)} \cdot dx.$$

Starting with the arc length formula above, consider:

$$L = \int_a^b \sqrt{1 + \left(\frac{dy}{dx}\right)^2}\, dx$$

$$= \int_a^b \sqrt{1 + \frac{[g'(t)]^2}{[f'(t)]^2}}\, dx$$

$$= \int_a^b \sqrt{[f'(t)]^2 + [g'(t)]^2} \cdot \underbrace{\frac{1}{f'(t)}}_{=dt}\, dx \qquad \text{Factor out the } [f'(t)]^2$$

$$= \int_{t_1}^{t_2} \sqrt{[f'(t)]^2 + [g'(t)]^2}\, dt.$$

Note the new bounds (no longer "x" bounds, but "t" bounds). They are found by finding t_1 and t_2 such that $a = f(t_1)$ and $b = f(t_2)$. This formula is important, so we restate it as a theorem.

> **Theorem 84   Arc Length of Parametric Curves**
>
> Let $x = f(t)$ and $y = g(t)$ be parametric equations with f' and g' continuous on some open interval I containing t_1 and t_2 on which the graph traces itself only once. The arc length of the graph, from $t = t_1$ to $t = t_2$, is
>
> $$L = \int_{t_1}^{t_2} \sqrt{[f'(t)]^2 + [g'(t)]^2}\, dt.$$

Notes:

As before, these integrals are often not easy to compute. We start with a simple example, then give another where we approximate the solution.

Example 7   **Arc Length of a Circle**
Find the arc length of the circle parametrized by $x = 3\cos t$, $y = 3\sin t$ on $[0, 3\pi/2]$.

SOLUTION   By direct application of Theorem 84, we have

$$L = \int_0^{3\pi/2} \sqrt{(-3\sin t)^2 + (3\cos t)^2}\, dt.$$

Apply the Pythagorean Theorem.

$$= \int_0^{3\pi/2} 3\, dt$$

$$= 3t\Big|_0^{3\pi/2} = 9\pi/2.$$

This should make sense; we know from geometry that the circumference of a circle with radius 3 is 6π; since we are finding the arc length of $3/4$ of a circle, the arc length is $3/4 \cdot 6\pi = 9\pi/2$.

Example 8   **Arc Length of a Parametric Curve**
The graph of the parametric equations $x = t(t^2 - 1)$, $y = t^2 - 1$ crosses itself as shown in Figure 10.37, forming a "teardrop." Find the arc length of the teardrop.

SOLUTION   We can see by the parameterizations of x and y that when $t = \pm 1$, $x = 0$ and $y = 0$. This means we'll integrate from $t = -1$ to $t = 1$. Applying Theorem 84, we have

$$L = \int_{-1}^{1} \sqrt{(3t^2 - 1)^2 + (2t)^2}\, dt$$

$$= \int_{-1}^{1} \sqrt{9t^4 - 2t^2 + 1}\, dt.$$

Unfortunately, the integrand does not have an antiderivative expressible by elementary functions. We turn to numerical integration to approximate its value. Using 4 subintervals, Simpson's Rule approximates the value of the integral as 2.65051. Using a computer, more subintervals are easy to employ, and $n = 20$ gives a value of 2.71559. Increasing n shows that this value is stable and a good approximation of the actual value.
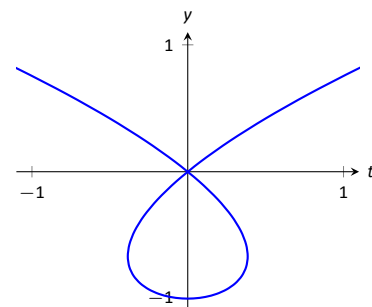
Figure 10.37: A graph of the parametric equations in Example 8, where the arc length of the teardrop is calculated.

Notes:

Surface Area of a Solid of Revolution

Related to the formula for finding arc length is the formula for finding surface area. We can adapt the formula found in Key Idea 38 from Section 10.1 in a similar way as done to produce the formula for arc length done before.

Key Idea 42   Surface Area of a Solid of Revolution

Consider the graph of the parametric equations $x = f(t)$ and $y = g(t)$, where f' and g' are continuous on an open interval I containing t_1 and t_2 on which the graph does not cross itself.

1. The surface area of the solid formed by revolving the graph about the x-axis is (where $g(t) \geq 0$ on $[t_1, t_2]$):

$$\text{Surface Area} = 2\pi \int_{t_1}^{t_2} g(t)\sqrt{[f'(t)]^2 + [g'(t)]^2}\, dt.$$

2. The surface area of the solid formed by revolving the graph about the y-axis is (where $f(t) \geq 0$ on $[t_1, t_2]$):

$$\text{Surface Area} = 2\pi \int_{t_1}^{t_2} f(t)\sqrt{[f'(t)]^2 + [g'(t)]^2}\, dt.$$
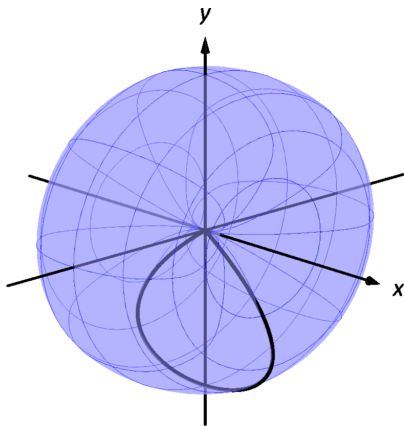
Figure 10.38: Rotating a teardrop shape about the x-axis in Example 9.

Example 9   Surface Area of a Solid of Revolution

Consider the teardrop shape formed by the parametric equations $x = t(t^2 - 1)$, $y = t^2 - 1$ as seen in Example 8. Find the surface area if this shape is rotated about the x-axis, as shown in Figure 10.38.

Solution   The teardrop shape is formed between $t = -1$ and $t = 1$. Using Key Idea 42, we see we need for $g(t) \geq 0$ on $[-1, 1]$, and this is not the case. To fix this, we simplify replace $g(t)$ with $-g(t)$, which flips the whole graph about the x-axis (and does not change the surface area of the resulting solid). The surface area is:

$$\text{Area } S = 2\pi \int_{-1}^{1} (1 - t^2)\sqrt{(3t^2 - 1)^2 + (2t)^2}\, dt$$

$$= 2\pi \int_{-1}^{1} (1 - t^2)\sqrt{9t^4 - 2t^2 + 1}\, dt.$$

Once again we arrive at an integral that we cannot compute in terms of elementary functions. Using Simpson's Rule with $n = 20$, we find the area to be

---

Notes:

$S = 9.44$. Using larger values of n shows this is accurate to 2 places after the decimal.

After defining a new way of creating curves in the plane, in this section we have applied calculus techniques to the parametric equation defining these curves to study their properties. In the next section, we define another way of forming curves in the plane. To do so, we create a new coordinate system, called *polar coordinates*, that identifies points in the plane in a manner different than from measuring distances from the *y*- and *x*- axes.

Notes:

Exercises 10.3

Terms and Concepts

1. T/F: Given parametric equations $x = f(t)$ and $y = g(t)$, $\frac{dy}{dx} = f'(t)/g'(t)$, as long as $g'(t) \neq 0$.

2. Given parametric equations $x = f(t)$ and $y = g(t)$, the derivative $\frac{dy}{dx}$ as given in Key Idea 39 is a function of _____?

3. T/F: Given parametric equations $x = f(t)$ and $y = g(t)$, to find $\frac{d^2y}{dx^2}$, one simply computes $\frac{d}{dt}\left(\frac{dy}{dx}\right)$.

4. T/F: If $\frac{dy}{dx} = 0$ at $t = t_0$, then the normal line to the curve at $t = t_0$ is a vertical line.

Problems

In Exercises 5–12, parametric equations for a curve are given.

(a) Find $\dfrac{dy}{dx}$.

(b) Find the equations of the tangent and normal line(s) at the point(s) given.

(c) Sketch the graph of the parametric functions along with the found tangent and normal lines.

5. $x = t, y = t^2;\quad t = 1$

6. $x = \sqrt{t}, y = 5t + 2;\quad t = 4$

7. $x = t^2 - t, y = t^2 + t;\quad t = 1$

8. $x = t^2 - 1, y = t^3 - t;\quad t = 0$ and $t = 1$

9. $x = \sec t, y = \tan t$ on $(-\pi/2, \pi/2);\quad t = \pi/4$

10. $x = \cos t, y = \sin(2t)$ on $[0, 2\pi];\quad t = \pi/4$

11. $x = \cos t \sin(2t), y = \sin t \sin(2t)$ on $[0, 2\pi];\quad t = 3\pi/4$

12. $x = e^{t/10} \cos t, y = e^{t/10} \sin t;\quad t = \pi/2$

In Exercises 13–20, find t-values where the curve defined by the given parametric equations has a horizontal tangent line. Note: these are the same equations as in Exercises 5 – 12.

13. $x = t, y = t^2$

14. $x = \sqrt{t}, y = 5t + 2$

15. $x = t^2 - t, y = t^2 + t$

16. $x = t^2 - 1, y = t^3 - t$

17. $x = \sec t, y = \tan t$ on $(-\pi/2, \pi/2)$

18. $x = \cos t, y = \sin(2t)$ on $[0, 2\pi]$

19. $x = \cos t \sin(2t), y = \sin t \sin(2t)$ on $[0, 2\pi]$

20. $x = e^{t/10} \cos t, y = e^{t/10} \sin t$

In Exercises 21–24, find $t = t_0$ where the graph of the given parametric equations is not smooth, then find $\displaystyle\lim_{t \to t_0} \frac{dy}{dx}$.

21. $x = \dfrac{1}{t^2 + 1},\quad y = t^3$

22. $x = -t^3 + 7t^2 - 16t + 13,\quad y = t^3 - 5t^2 + 8t - 2$

23. $x = t^3 - 3t^2 + 3t - 1,\quad y = t^2 - 2t + 1$

24. $x = \cos^2 t,\quad y = 1 - \sin^2 t$

In Exercises 25–32, parametric equations for a curve are given. Find $\frac{d^2y}{dx^2}$, then determine the intervals on which the graph of the curve is concave up/down. Note: these are the same equations as in Exercises 5 – 12.

25. $x = t,\quad y = t^2$

26. $x = \sqrt{t},\quad y = 5t + 2$

27. $x = t^2 - t,\quad y = t^2 + t$

28. $x = t^2 - 1,\quad y = t^3 - t$

29. $x = \sec t,\quad y = \tan t$ on $(-\pi/2, \pi/2)$

30. $x = \cos t,\quad y = \sin(2t)$ on $[0, 2\pi]$

31. $x = \cos t \sin(2t),\quad y = \sin t \sin(2t)$ on $[-\pi/2, \pi/2]$

32. $x = e^{t/10} \cos t,\quad y = e^{t/10} \sin t$

In Exercises 33–40, find the arc length of the graph of the parametric equations on the given interval(s).

33. $x = -3\sin(2t),\quad y = 3\cos(2t)$ on $[0, \pi]$

34. $x = e^{t/10} \cos t,\quad y = e^{t/10} \sin t$ on $[0, 2\pi]$ and $[2\pi, 4\pi]$

35. $x = 5t + 2,\quad y = 1 - 3t$ on $[-1, 1]$

36. $x = 2t^{3/2},\quad y = 3t$ on $[0, 1]$

37. $x = \cos t,\quad y = \sin t$ on $[0, 2\pi]$

38. $x = 1 + 3t^2,\quad y = 4 + 2t^3$ on $[0, 1]$

39. $x = \dfrac{t}{1+t},\quad y = \ln(1 + t)$ on $[0, 2]$

40. $x = e^t - t,\quad y = 4e^{-t/2}$ on $[-8, 3]$

In Exercises 41–44, numerically approximate the given arc length.

41. Approximate the arc length of one petal of the rose curve $x = \cos t \cos(2t),\quad y = \sin t \cos(2t)$ using Simpson's Rule and $n = 4$.

42. Approximate the arc length of the "bow tie curve" $x = \cos t,\quad y = \sin(2t)$ using Simpson's Rule and $n = 6$.

43. Approximate the arc length of the parabola $x = t^2 - t$, $y = t^2 + t$ on $[-1, 1]$ using Simpson's Rule and $n = 4$.

44. A common approximate of the circumference of an ellipse given by $x = a \cos t,\quad y = b \sin t$ is $C \approx 2\pi\sqrt{\dfrac{a^2 + b^2}{2}}$. Use this formula to approximate the circumference of $x = 5 \cos t,\quad y = 3 \sin t$ and compare this to the approximation given by Simpson's Rule and $n = 6$.

In Exercises 45–51, a solid of revolution is described. Find or approximate its surface area as specified.

45. Find the surface area of the sphere formed by rotating the circle $x = 2 \cos t,\quad y = 2 \sin t$ about:

   (a) the x-axis and

   (b) the y-axis.

46. Find the surface area of the torus (or "donut") formed by rotating the circle $x = \cos t + 2,\quad y = \sin t$ about the y-axis.

47. Find the surface area of the solid formed by rotating the curve $x = a\cos^3\theta$, $y = a\sin^3\theta$ on $[0, \pi/2]$ about the x−axis

48. Find the surface area of the solid formed by rotating the curve $x = t^3$, $y = t^2$ on $[0, 1]$ about the x−axis

49. Find the surface area of the solid formed by rotating the curve $x = 3t^2$, $y = 2t^3$ on $[0, 5]$ about the y−axis

50. Approximate the surface area of the solid formed by rotating the "upper right half" of the bow tie curve $x = \cos t$, $y = \sin(2t)$ on $[0, \pi/2]$ about the x-axis, using Simpson's Rule and $n = 4$.

51. Approximate the surface area of the solid formed by rotating the one petal of the rose curve $x = \cos t \cos(2t)$, $y = \sin t \cos(2t)$ on $[0, \pi/4]$ about the x-axis, using Simpson's Rule and $n = 4$.

10.4   Introduction to Polar Coordinates

We are generally introduced to the idea of graphing curves by relating x-values to y-values through a function f. That is, we set $y = f(x)$, and plot lots of point pairs (x, y) to get a good notion of how the curve looks. This method is useful but has limitations, not least of which is that curves that "fail the vertical line test" cannot be graphed without using multiple functions.

The previous two sections introduced and studied a new way of plotting points in the x, y-plane. Using parametric equations, x and y values are computed independently and then plotted together. This method allows us to graph an extraordinary range of curves. This section introduces yet another way to plot points in the plane: using **polar coordinates**.

Polar Coordinates

Figure 10.39: Illustrating polar coordinates.

Start with a point O in the plane called the **pole** (we will always identify this point with the origin). From the pole, draw a ray, called the **initial ray** (we will always draw this ray horizontally, identifying it with the positive x-axis). A point P in the plane is determined by the distance r that P is from O, and the angle θ formed between the initial ray and the segment $\overline{OP}$ (measured counter-clockwise). We record the distance and angle as an ordered pair (r, θ).

Watch the video:
Polar Coordinates — The Basics at
`https://youtu.be/r0fv9V9GHdo`

Practice will make this process more clear.

Example 1   **Plotting Polar Coordinates**
Plot the following polar coordinates:

$$A(1, \pi/4) \quad B(1.5, \pi) \quad C(2, -\pi/3) \quad D(-1, \pi/4)$$

SOLUTION   To aid in the drawing, a polar grid is provided at the bottom of this page.   To place the point A, go out 1 unit along the initial ray (putting you on the inner circle shown on the grid), then rotate counter-clockwise $\pi/4$ radians (or $45°$). Alternately, one can consider the rotation first: think about the ray from O that forms an angle of $\pi/4$ with the initial ray, then move out 1 unit along this ray (again placing you on the inner circle of the grid).
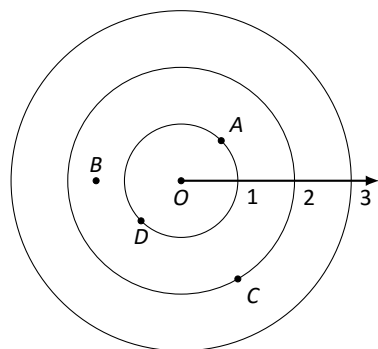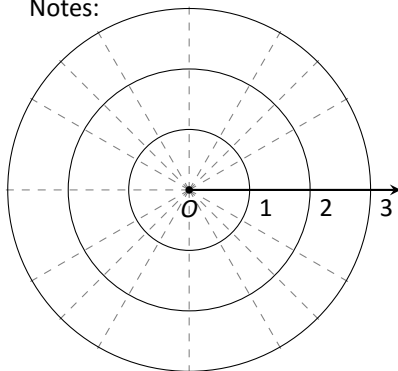
Figure 10.40: Plotting polar points in Example 1.

Notes:

To plot *B*, go out 1.5 units along the initial ray and rotate π radians (180°).

To plot *C*, go out 2 units along the initial ray then rotate *clockwise* $\pi/3$ radians, as the angle given is negative.

To plot *D*, move along the initial ray "-1" units – in other words, "back up" 1 unit, then rotate counter-clockwise by $\pi/4$. The results are given in Figure 10.40.

Consider the following two points: $A(1, \pi)$ and $B(-1, 0)$. To locate *A*, go out 1 unit on the initial ray then rotate π radians; to locate *B*, go out -1 units on the initial ray and don't rotate. One should see that *A* and *B* are located at the same point in the plane. We can also consider $C(1, 3\pi)$, or $D(1, -\pi)$; all four of these points share the same location.

This ability to identify a point in the plane with multiple polar coordinates is both a "blessing" and a "curse." We will see that it is beneficial as we can plot beautiful functions that intersect themselves (much like we saw with parametric functions). The unfortunate part of this is that it can be difficult to determine when this happens. We'll explore this more later in this section.

Polar to Rectangular Conversion

It is useful to recognize both the rectangular (or, Cartesian) coordinates of a point in the plane and its polar coordinates. Figure 10.41 shows a point *P* in the plane with rectangular coordinates (x, y) and polar coordinates (r, θ). Using trigonometry, we can make the identities given in the following Key Idea.
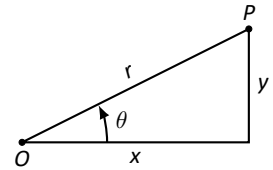
Figure 10.41: Converting between rectangular and polar coordinates.

> **Key Idea 43   Converting Between Rectangular and Polar Coordinates**
>
> Given the polar point $P(r, \theta)$, the rectangular coordinates are determined by
> $$x = r\cos\theta \qquad y = r\sin\theta.$$
>
> Given the rectangular coordinates (x, y), the polar coordinates are determined by
> $$r^2 = x^2 + y^2 \qquad \tan\theta = \frac{y}{x}.$$

Example 2   Converting Between Polar and Rectangular Coordinates

1. Convert the polar coordinates $A(2, 2\pi/3)$ and $B(-1, 5\pi/4)$ to rectangular coordinates.

2. Convert the rectangular coordinates $(1, 2)$ and $(-1, 1)$ to polar coordinates.

Notes:

605

SOLUTION

1. (a) We start with $A(2, 2\pi/3)$. Using Key Idea 43, we have

$$x = 2\cos(2\pi/3) = -1 \qquad y = 2\sin(2\pi/3) = \sqrt{3}.$$

So the rectangular coordinates are $(-1, \sqrt{3}) \approx (-1, 1.732)$.

(b) The polar point $B(-1, 5\pi/4)$ is converted to rectangular with:

$$x = -1\cos(5\pi/4) = \sqrt{2}/2 \qquad y = -1\sin(5\pi/4) = \sqrt{2}/2.$$

So the rectangular coordinates are $(\sqrt{2}/2, \sqrt{2}/2) \approx (0.707, 0.707)$.

These points are plotted in Figure 10.42 (a). The rectangular coordinate system is drawn lightly under the polar coordinate system so that the relationship between the two can be seen.

2. (a) To convert the rectangular point $(1, 2)$ to polar coordinates, we use the Key Idea to form the following two equations:

$$1^2 + 2^2 = r^2 \qquad \tan\theta = \frac{2}{1}.$$

The first equation tells us that $r = \sqrt{5}$. Using the inverse tangent function, we find

$$\tan\theta = 2 \quad \Rightarrow \quad \theta = \tan^{-1}2 \approx 1.11 \text{ radians} \approx 63.43°.$$

Thus polar coordinates of $(1, 2)$ are $(\sqrt{5}, 1.11)$.

(b) To convert $(-1, 1)$ to polar coordinates, we form the equations

$$(-1)^2 + 1^2 = r^2 \qquad \tan\theta = \frac{1}{-1}.$$

Thus $r = \sqrt{2}$. We need to be careful in computing θ: using the inverse tangent function, we have

$$\tan\theta = -1 \quad \Rightarrow \quad \theta = \tan^{-1}(-1) = -\pi/4.$$

This is not the angle we desire. The range of $\tan^{-1}x$ is $(-\pi/2, \pi/2)$; that is, it returns angles that lie in the 1st and 4th quadrants. To find locations in the 2nd and 3rd quadrants, add π to the result of $\tan^{-1}x$. So $\pi + (-\pi/4)$ puts the angle at $3\pi/4$. Thus the polar point is $(\sqrt{2}, 3\pi/4)$.

An alternate method is to use the angle θ given by arctangent, but change the sign of r. Thus we could also refer to $(-1, 1)$ as $(-\sqrt{2}, -\pi/4)$.

These points are plotted in Figure 10.42 (b). The polar system is drawn lightly under the rectangular grid with rays to demonstrate the angles used.

(a)

(b)

Figure 10.42: Plotting rectangular and polar points in Example 2.

Notes:

Polar Functions and Polar Graphs

Defining a new coordinate system allows us to create a new kind of function, a **polar function.** Rectangular coordinates lent themselves well to creating functions that related x and y, such as $y = x^2$. Polar coordinates allow us to create functions that relate r and θ. Normally these functions look like $r = f(\theta)$, although we can create functions of the form $\theta = f(r)$. The following examples introduce us to this concept.

Example 3   Introduction to Graphing Polar Functions
Describe the graphs of the following polar functions.

1. $r = 1.5$

2. $\theta = \pi/4$

SOLUTION

1. The equation $r = 1.5$ describes all points that are 1.5 units from the pole; as the angle is not specified, any θ is allowable. All points 1.5 units from the pole describes a circle of radius 1.5.

   We can consider the rectangular equivalent of this equation; using $r^2 = x^2 + y^2$, we see that $1.5^2 = x^2 + y^2$, which we recognize as the equation of a circle centered at $(0, 0)$ with radius 1.5. This is sketched in Figure 10.43.

2. The equation $\theta = \pi/4$ describes all points such that the line through them and the pole make an angle of $\pi/4$ with the initial ray. As the radius r is not specified, it can be any value (even negative). Thus $\theta = \pi/4$ describes the line through the pole that makes an angle of $\pi/4 = 45°$ with the initial ray.

   We can again consider the rectangular equivalent of this equation. Combine $\tan \theta = y/x$ and $\theta = \pi/4$:

   $$\tan \pi/4 = y/x \quad \Rightarrow \quad x \tan \pi/4 = y \quad \Rightarrow \quad y = x.$$

   This graph is also plotted in Figure 10.43.

Figure 10.43: Plotting standard polar plots.

The basic rectangular equations of the form $x = h$ and $y = k$ create vertical and horizontal lines, respectively; the basic polar equations $r = h$ and $\theta = \alpha$ create circles and lines through the pole, respectively. With this as a foundation, we can create more complicated polar functions of the form $r = f(\theta)$. The input is an angle; the output is a length, how far in the direction of the angle to go out.

We sketch these functions much like we sketch rectangular and parametric functions: we plot lots of points and "connect the dots" with curves. We demonstrate this in the following example.

Notes:

θ	$r = 1 + \cos\theta$
0	2
$\pi/6$	$1 + \sqrt{3}/2$
$\pi/4$	$1 + 1/\sqrt{2}$
$\pi/3$	$3/2$
$\pi/2$	1
$2\pi/3$	$1/2$
$3\pi/4$	$1 - 1/\sqrt{2}$
$5\pi/6$	$1 - \sqrt{3}/2$
π	0
$7\pi/6$	$1 - \sqrt{3}/2$
$5\pi/4$	$1 - 1/\sqrt{2}$
$4\pi/3$	$1/2$
$3\pi/2$	1
$5\pi/3$	$3/2$
$7\pi/4$	$1 + 1/\sqrt{2}$
$11\pi/6$	$1 + \sqrt{3}/2$
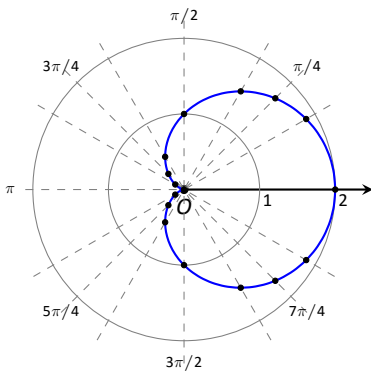
Figure 10.44: Graph of the polar function in Example 4 by plotting points.

Figure 10.45: Polar plots from Example 5.

Example 4   Sketching Polar Functions
Sketch the polar function $r = 1 + \cos\theta$ on $[0, 2\pi]$ by plotting points.

SOLUTION   A common question when sketching curves by plotting points is "Which points should I plot?" With rectangular equations, we often chose "easy" values — integers, then added more if needed. When plotting polar equations, start with the "common" angles — multiples of $\pi/6$ and $\pi/4$. Figure 10.44 gives a table of just a few values of θ in $[0, \pi]$.

Consider the point $(2, 0)$ determined by the first line of the table. The angle is 0 radians – we do not rotate from the initial ray – then we go out 2 units from the pole. When $\theta = \pi/6$, $r = 1 + \sqrt{3}/2$; so rotate by $\pi/6$ radians and go out $1 + \sqrt{3}/2$ units.

Example 5   Sketching Polar Functions
Sketch the polar function $r = \cos(2\theta)$ on $[0, 2\pi]$ by plotting points.

SOLUTION   We start by making a table of $\cos(2\theta)$ evaluated at common angles θ, as shown in Figure 10.46. These points are then plotted in Figure 10.45. This particular graph "moves" around quite a bit and one can easily forget which points should be connected to each other. To help us with this, we numbered each point in the table and on the graph.

Pt.	θ	$\cos(2\theta)$	Pt.	θ	$\cos(2\theta)$
1	0	1	10	$7\pi/6$	0.5
2	$\pi/6$	0.5	11	$5\pi/4$	0
3	$\pi/4$	0	12	$4\pi/3$	−0.5
4	$\pi/3$	−0.5	13	$3\pi/2$	−1
5	$\pi/2$	−1	14	$5\pi/3$	−0.5
6	$2\pi/3$	−0.5	15	$7\pi/4$	0
7	$3\pi/4$	0	16	$11\pi/6$	0.5
8	$5\pi/6$	0.5	17	2π	1
9	π	1			

Figure 10.46: Tables of points for plotting a polar curve.

This plot is an example of a *rose curve*.

It is sometimes desirable to refer to a graph via a polar equation, and other times by a rectangular equation. Therefore it is necessary to be able to convert between polar and rectangular functions, which we practice in the following example. We will make frequent use of the identities found in Key Idea 43.

Notes:

Example 6   **Converting between rectangular and polar equations.**

Convert from rectangular to polar.   Convert from polar to rectangular.

1. $y = x^2$

3. $r = \dfrac{2}{\sin\theta - \cos\theta}$

2. $xy = 1$

4. $r = 2\cos\theta$

SOLUTION

1. Replace y with $r\sin\theta$ and replace x with $r\cos\theta$, giving:

$$y = x^2$$
$$r\sin\theta = r^2\cos^2\theta$$
$$\frac{\sin\theta}{\cos^2\theta} = r$$

We have found that $r = \sin\theta/\cos^2\theta = \tan\theta\sec\theta$. The domain of this polar function is $(-\pi/2, \pi/2)$; plot a few points to see how the familiar parabola is traced out by the polar equation.

2. We again replace x and y using the standard identities and work to solve for r:

$$xy = 1$$
$$r\cos\theta \cdot r\sin\theta = 1$$
$$r^2 = \frac{1}{\cos\theta\sin\theta}$$
$$r = \frac{1}{\sqrt{\cos\theta\sin\theta}}$$

This function is valid only when the product of $\cos\theta\sin\theta$ is positive. This occurs in the first and third quadrants, meaning the domain of this polar function is $(0, \pi/2) \cup (\pi, 3\pi/2)$.

We can rewrite the original rectangular equation $xy = 1$ as $y = 1/x$. This is graphed in Figure 10.47; note how it only exists in the first and third quadrants.

3. There is no set way to convert from polar to rectangular; in general, we look to form the products $r\cos\theta$ and $r\sin\theta$, and then replace these with x and y, respectively. We start in this problem by multiplying both sides
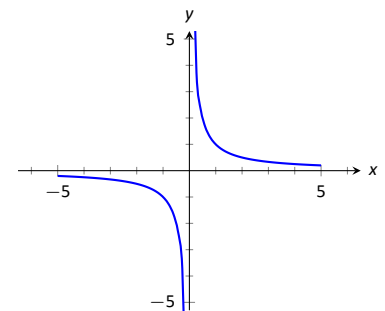
Figure 10.47: Graphing $xy = 1$ from Example 6.

Notes:

by $\sin\theta - \cos\theta$:

$$r = \frac{2}{\sin\theta - \cos\theta}$$
$$r(\sin\theta - \cos\theta) = 2$$
$$r\sin\theta - r\cos\theta = 2. \qquad \text{Now replace with } y \text{ and } x:$$
$$y - x = 2$$
$$y = x + 2.$$

The original polar equation, $r = 2/(\sin\theta - \cos\theta)$ does not easily reveal that its graph is simply a line. However, our conversion shows that it is. The upcoming gallery of polar curves gives the general equations of lines in polar form.

4. By multiplying both sides by r, we obtain both an r^2 term and an $r\cos\theta$ term, which we replace with $x^2 + y^2$ and x, respectively.

$$r = 2\cos\theta$$
$$r^2 = 2r\cos\theta$$
$$x^2 + y^2 = 2x.$$

We recognize this as a circle; by completing the square we can find its radius and center.

$$x^2 - 2x + y^2 = 0$$
$$(x - 1)^2 + y^2 = 1.$$

The circle is centered at $(1, 0)$ and has radius 1. The upcoming gallery of polar curves gives the equations of *some* circles in polar form; circles with arbitrary centers have a complicated polar equation that we do not consider here.

Some curves have very simple polar equations but rather complicated rectangular ones. For instance, the equation $r = 1 + \cos\theta$ describes a *cardioid* (a shape important to the sensitivity of microphones, among other things; one is graphed in the gallery in the Limaçon section). It's rectangular form is not nearly as simple; it is the implicit equation $x^4 + y^4 + 2x^2y^2 - 2xy^2 - 2x^3 - y^2 = 0$. The conversion is not "hard," but takes several steps, and is left as an exercise.

Notes:

Gallery of Polar Curves

There are a number of basic and "classic" polar curves, famous for their beauty and/or applicability to the sciences. This section ends with a small gallery of some of these graphs. We encourage the reader to understand how these graphs are formed, and to investigate with technology other types of polar functions.

Lines

Through the origin:

$\theta = \alpha$
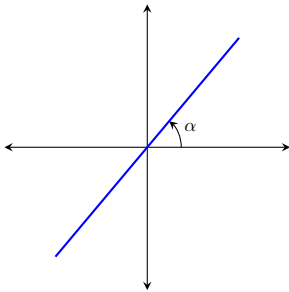
Horizontal line:

$r = a \csc \theta$
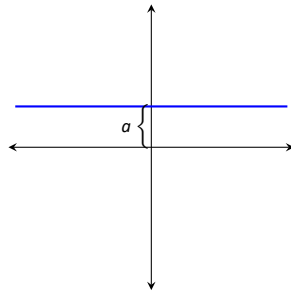
Vertical line:

$r = a \sec \theta$
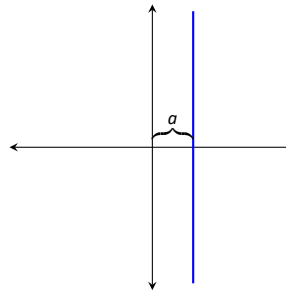
Not through origin:

$r = \dfrac{b}{\sin \theta - m \cos \theta}$
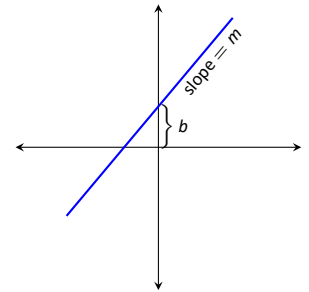
Notes:

Circles

Centered on origin:

$r = a$
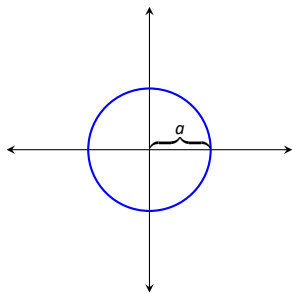
$(x - \frac{a}{2})^2 + y^2 = \frac{a^2}{4}$

$r = a \cos\theta$

$x^2 + (y - \frac{a}{2})^2 = \frac{a^2}{4}$

$r = a \sin\theta$
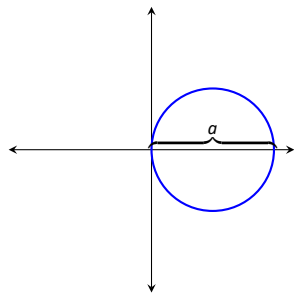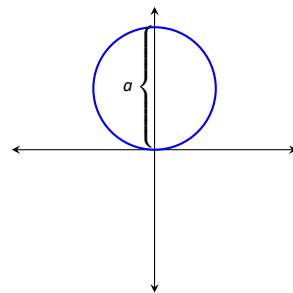
Sprial

Archimedean spiral

$r = \theta$
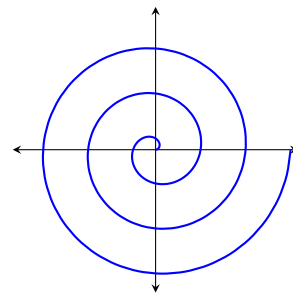
Limaçons

Symmetric about x-axis: $r = a \pm b \cos\theta$;   Symmetric about y-axis: $r = a \pm b \sin\theta$; $a, b > 0$

With inner loop:

$\dfrac{a}{b} < 1$

Cardioid:

$\dfrac{a}{b} = 1$

Dimpled:

$1 < \dfrac{a}{b} < 2$

Convex:

$\dfrac{a}{b} > 2$
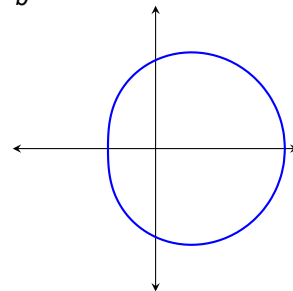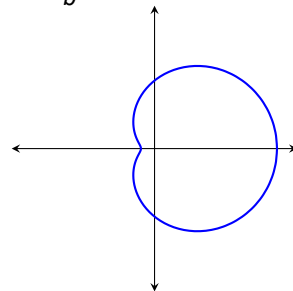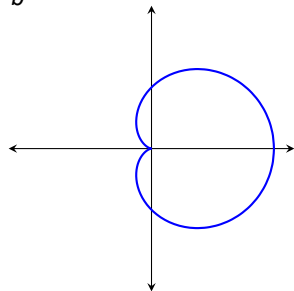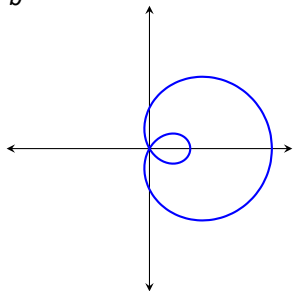
Rose Curves

Symmetric about x-axis: $r = a \cos(n\theta)$;   Symmetric about y-axis: $r = a \sin(n\theta)$

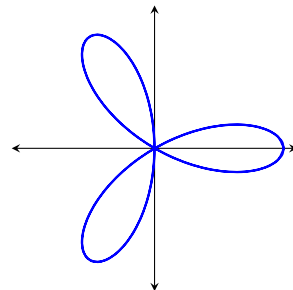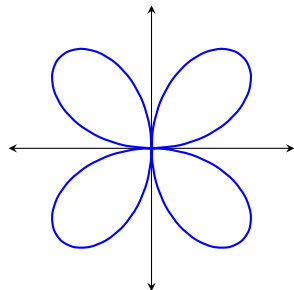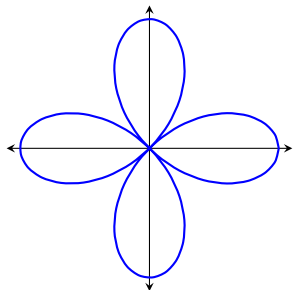Curve contains 2n petals when n is even and n petals when n is odd.

$r = a \cos(2\theta)$
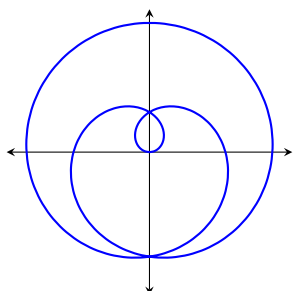
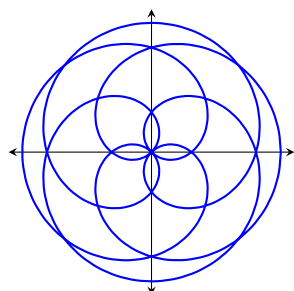$r = a \sin(2\theta)$

$r = a \cos(3\theta)$

$r = a \sin(3\theta)$
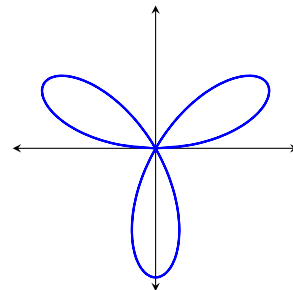
Special Curves

Rose curves

$r = a \sin(\theta/5)$
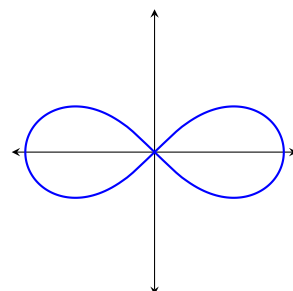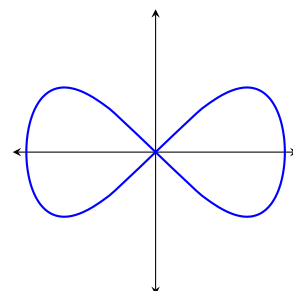
$r = a \sin(2\theta/5)$

Lemniscate:

$r^2 = a^2 \cos(2\theta)$

Eight Curve:

$r^2 = a^2 \sec^4\theta \cos(2\theta)$

Earlier we discussed how each point in the plane does not have a unique representation in polar form. This can be a "good" thing, as it allows for the beautiful and interesting curves seen in the preceding gallery. However, it can also be a "bad" thing, as it can be difficult to determine where two curves intersect.

Example 7   **Finding points of intersection with polar curves**
Determine where the graphs of the polar equations $r = 1 + 3\cos\theta$ and $r = \cos\theta$ intersect.

SOLUTION   As technology is generally readily available, it is usually a good idea to start with a graph. We have graphed the two functions in Figure 10.48(a); to better discern the intersection points, part (b) of the figure zooms in around the origin. We start by setting the two functions equal to each other and solving for θ:

$$1 + 3\cos\theta = \cos\theta$$
$$2\cos\theta = -1$$
$$\cos\theta = -\frac{1}{2}$$
$$\theta = \frac{2\pi}{3}, \frac{4\pi}{3}.$$

(There are, of course, infinite solutions to the equation $\cos\theta = -1/2$; as the limaçon is traced out once on $[0, 2\pi]$, we restrict our solutions to this interval.)

We need to analyze this solution. When $\theta = 2\pi/3$ we obtain the point of intersection that lies in the 4th quadrant. When $\theta = 4\pi/3$, we get the point of intersection that lies in the 2nd quadrant. There is more to say about this second intersection point, however. The circle defined by $r = \cos\theta$ is traced out once on $[0, \pi]$, meaning that this point of intersection occurs while tracing out the circle a second time. It seems strange to pass by the point once and then recognize it as a point of intersection only when arriving there a "second time." The first time the circle arrives at this point is when $\theta = \pi/3$. It is key to understand that these two points are the same: $(\cos\pi/3, \pi/3)$ and $(\cos 4\pi/3, 4\pi/3)$.

To summarize what we have done so far, we have found two points of intersection: when $\theta = 2\pi/3$ and when $\theta = 4\pi/3$. When referencing the circle $r = \cos\theta$, the latter point is better referenced as when $\theta = \pi/3$.

There is yet another point of intersection: the pole (or, the origin). We did not recognize this intersection point using our work above as each graph arrives at the pole at a different θ value.

A graph intersects the pole when $r = 0$. Considering the circle $r = \cos\theta$, $r = 0$ when $\theta = \pi/2$ (and odd multiples thereof, as the circle is repeatedly traced).
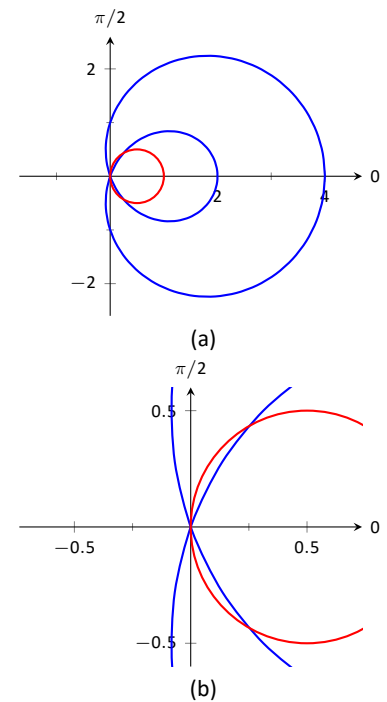
(a)

(b)

Figure 10.48: Graphs to help determine the points of intersection of the polar functions given in Example 7.

Notes:

The limaçon intersects the pole when $1 + 3\cos\theta = 0$; this occurs when $\cos\theta = -1/3$, or for $\theta = \cos^{-1}(-1/3)$. This is a nonstandard angle, approximately $\theta = 1.9106$ radians $\approx 109.47°$. The limaçon intersects the pole twice in $[0, 2\pi]$; the other angle at which the limaçon is at the pole is the reflection of the first angle across the x-axis. That is, $\theta = 4.3726 \approx 250.53°$.

If all one is concerned with is the (x, y) coordinates at which the graphs intersect, much of the above work is extraneous. We know they intersect at $(0, 0)$; we might not care at what θ value. Likewise, using $\theta = 2\pi/3$ and $\theta = 4\pi/3$ can give us the needed rectangular coordinates. However, in the next section we apply calculus concepts to polar functions. When computing the area of a region bounded by polar curves, understanding the nuances of the points of intersection becomes important.

Notes:

Exercises 10.4

Terms and Concepts

1. In your own words, describe how to plot the polar point $P(r, \theta)$.

2. T/F: When plotting a point with polar coordinate $P(r, \theta)$, r must be positive.

3. T/F: Every point in the Cartesian plane can be represented by a polar coordinate.

4. T/F: Every point in the Cartesian plane can be represented uniquely by a polar coordinate.
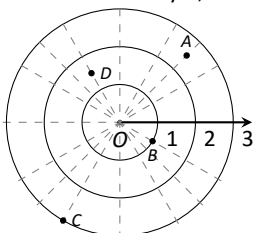
Problems

5. Plot the points with the given polar coordinates.

   (a) $A(2, 0)$ (c) $C(-2, \pi/2)$

   (b) $B(1, \pi)$ (d) $D(1, \pi/4)$

6. Plot the points with the given polar coordinates.

   (a) $A(2, 3\pi)$ (c) $C(1, 2)$

   (b) $B(1, -\pi)$ (d) $D(1/2, 5\pi/6)$

7. For each of the given points give two sets of polar coordinates that identify it, where $0 \leq \theta \leq 2\pi$.

8. For each of the given points give two sets of polar coordinates that identify it, where $-\pi \leq \theta \leq \pi$.
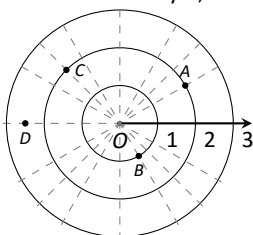
9. Convert the polar coordinates A and B to rectangular, and the rectangular coordinates C and D to polar.

   (a) $A(2, \pi/4)$ (c) $C(2, -1)$

   (b) $B(2, -\pi/4)$ (d) $D(-2, 1)$

10. Convert the polar coordinates A and B to rectangular, and the rectangular coordinates C and D to polar.

   (a) $A(3, \pi)$ (c) $C(0, 4)$

   (b) $B(1, 2\pi/3)$ (d) $D(1, -\sqrt{3})$

In Exercises 11–32, graph the polar function on the given interval.

11. $r = 2$, $0 \leq \theta \leq \pi/2$

12. $\theta = \pi/6$, $-1 \leq r \leq 2$

13. $r = 1 - \cos\theta$, $[0, 2\pi]$

14. $r = 2 + \sin\theta$, $[0, 2\pi]$

15. $r = 2 - \sin\theta$, $[0, 2\pi]$

16. $r = 1 - 2\sin\theta$, $[0, 2\pi]$

17. $r = 1 + 2\sin\theta$, $[0, 2\pi]$

18. $r = \cos(2\theta)$, $[0, 2\pi]$

19. $r = \sin(3\theta)$, $[0, \pi]$

20. $r = \cos(\theta/3)$, $[0, 3\pi]$

21. $r = \cos(2\theta/3)$, $[0, 6\pi]$

22. $r = \theta/2$, $[0, 4\pi]$

23. $r = 3\sin(\theta)$, $[0, \pi]$

24. $r = -4\sin(\theta)$, $[0, \pi]$

25. $r = -2\cos(\theta)$, $[0, \pi]$

26. $r = \dfrac{3}{2}\cos(\theta)$, $[0, \pi]$

27. $r = \cos\theta \sin\theta$, $[0, 2\pi]$

28. $r = \theta^2 - (\pi/2)^2$, $[-\pi, \pi]$

29. $r = \dfrac{3}{5\sin\theta - \cos\theta}$, $[0, 2\pi]$

30. $r = \dfrac{-2}{3\cos\theta - 2\sin\theta}$, $[0, 2\pi]$

31. $r = 3\sec\theta$, $(-\pi/2, \pi/2)$

32. $r = 3\csc\theta$, $(0, \pi)$

In Exercises 33–43, convert the polar equation to a rectangular equation.

33. $r = 2\cos\theta$

34. $r = -4\sin\theta$

35. $r = 3\sin(\theta)$

36. $r = -\dfrac{3}{2}\cos(\theta)$

37. $r = \cos\theta + \sin\theta$

38. $r = \dfrac{7}{5\sin\theta - 2\cos\theta}$

39. $r = \dfrac{3}{\cos\theta}$

40. $r = \dfrac{4}{\sin\theta}$

41. $r = \tan\theta$

42. $r = 2$

43. $\theta = \dfrac{\pi}{6}$

In Exercises 44–51, convert the rectangular equation to a polar equation.

44. $y = x$

45. $y = 4x + 7$

46. $x = 5$

47. $y = 5$

48. $x = y^2$

49. $x^2 y = 1$

50. $x^2 + y^2 = 7$

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

In Exercises 52–59, find the points of intersection of the polar graphs.

52. $r = \sin(2\theta)$ and $r = \cos\theta$ on $[0, \pi]$

53. $r = \cos(2\theta)$ and $r = \cos\theta$ on $[0, \pi]$

54. $r = 2\cos\theta$ and $r = 2\sin\theta$ on $[0, \pi]$

55. $r = \sin\theta$ and $r = \sqrt{3} + 3\sin\theta$ on $[0, 2\pi]$

56. $r = \sin(3\theta)$ and $r = \cos(3\theta)$ on $[0, \pi]$

57. $r = 3\cos\theta$ and $r = 1 + \cos\theta$ on $[-\pi, \pi]$

58. $r = 1$ and $r = 2\sin(2\theta)$ on $[0, 2\pi]$

59. $r = 1 - \cos\theta$ and $r = 1 + \sin\theta$ on $[0, 2\pi]$

60. Pick a integer value for n, where $n \neq 2, 3$, and use technology to plot $r = \sin\left(\dfrac{m}{n}\theta\right)$ for three different integer values of m. Sketch these and determine a minimal interval on which the entire graph is shown.

61. Create your own polar function, $r = f(\theta)$ and sketch it. Describe why the graph looks as it does.

10.5 Calculus and Polar Functions

The previous section defined polar coordinates, leading to polar functions. We investigated plotting these functions and solving a fundamental question about their graphs, namely, where do two polar graphs intersect?

We now turn our attention to answering other questions, whose solutions require the use of calculus. A basis for much of what is done in this section is the ability to turn a polar function $r = f(\theta)$ into a set of parametric equations. Using the identities $x = r\cos\theta$ and $y = r\sin\theta$, we can create the parametric equations $x = f(\theta)\cos\theta$, $y = f(\theta)\sin\theta$ and apply the concepts of Section 10.3.

Polar Functions and $\dfrac{dy}{dx}$

We are interested in the lines tangent to a given graph, regardless of whether that graph is produced by rectangular, parametric, or polar equations. In each of these contexts, the slope of the tangent line is $\frac{dy}{dx}$. Given $r = f(\theta)$, we are generally *not* concerned with $r' = f'(\theta)$; that describes how fast r changes with respect to θ. Instead, we will use $x = f(\theta)\cos\theta$, $y = f(\theta)\sin\theta$ to compute $\frac{dy}{dx}$.

Using Key Idea 39 we have

$$\frac{dy}{dx} = \frac{dy}{d\theta} \bigg/ \frac{dx}{d\theta}.$$

Each of the two derivatives on the right hand side of the equality requires the use of the Product Rule. We state the important result as a Key Idea.

Key Idea 44   Finding $\frac{dy}{dx}$ with Polar Functions

Let $r = f(\theta)$ be a polar function. With $x = f(\theta)\cos\theta$ and $y = f(\theta)\sin\theta$,

$$\frac{dy}{dx} = \frac{\frac{dy}{d\theta}}{\frac{dx}{d\theta}} = \frac{f'(\theta)\sin\theta + f(\theta)\cos\theta}{f'(\theta)\cos\theta - f(\theta)\sin\theta}.$$

Watch the video:
The Slope of Tangent Lines to Polar Curves at
`https://youtu.be/QTa9OZ4iGPo`

Notes:

Example 1   **Finding $\frac{dy}{dx}$ with polar functions.**
Consider the limaçon $r = 1 + 2\sin\theta$ on $[0, 2\pi]$.

1. Find the rectangular equations of the tangent and normal lines to the graph at $\theta = \pi/4$.

2. Find where the graph has vertical and horizontal tangent lines.

SOLUTION

1. We start by computing $\frac{dy}{dx}$. With $f'(\theta) = 2\cos\theta$, we have

$$\frac{dy}{dx} = \frac{2\cos\theta\sin\theta + \cos\theta(1 + 2\sin\theta)}{2\cos^2\theta - \sin\theta(1 + 2\sin\theta)}$$

$$= \frac{\cos\theta(4\sin\theta + 1)}{2(\cos^2\theta - \sin^2\theta) - \sin\theta}.$$

When $\theta = \pi/4$, $\frac{dy}{dx} = -2\sqrt{2} - 1$ (this requires a bit of simplification). In rectangular coordinates, the point on the graph at $\theta = \pi/4$ is $(1 + \sqrt{2}/2, 1 + \sqrt{2}/2)$. Thus the rectangular equation of the line tangent to the limaçon at $\theta = \pi/4$ is

$$y = (-2\sqrt{2} - 1)\left(x - (1 + \sqrt{2}/2)\right) + 1 + \sqrt{2}/2 \approx -3.83x + 8.24.$$

The limaçon and the tangent line are graphed in Figure 10.49.

The normal line has the opposite–reciprocal slope as the tangent line, so its equation is

$$y \approx \frac{1}{3.83}x + 1.26.$$

2. To find the horizontal lines of tangency, we find where $\frac{dy}{dx} = 0$ (when the denominator does not equal 0); thus we find where the numerator of our equation for $\frac{dy}{dx}$ is 0.

$$\cos\theta(4\sin\theta + 1) = 0 \quad \Rightarrow \quad \cos\theta = 0 \quad \text{or} \quad 4\sin\theta + 1 = 0.$$

On $[0, 2\pi]$, $\cos\theta = 0$ when $\theta = \pi/2$, $3\pi/2$.

Setting $4\sin\theta + 1 = 0$ gives $\theta = \sin^{-1}(-1/4) \approx -0.2527 = -14.48°$. We want the results in $[0, 2\pi]$; we also recognize there are two solutions, one in the 3rd quadrant and one in the 4th. Using reference angles, we have our two solutions as $\theta = 3.39$ and 6.03 radians. The four points we obtained where the limaçon has a horizontal tangent line are given in Figure 10.49 with black–filled dots.

Notes:

To find the vertical lines of tangency, we determine where $\frac{dy}{dx}$ is undefined by setting the denominator of $\frac{dy}{dx} = 0$ (when the numerator does not equal 0).

$$2(\cos^2 \theta - \sin^2 \theta) - \sin \theta = 0.$$

Convert the $\cos^2 \theta$ term to $1 - \sin^2 \theta$:

$$2(1 - \sin^2 \theta - \sin^2 \theta) - \sin \theta = 0$$
$$4 \sin^2 \theta + \sin \theta - 2 = 0.$$

Recognize this as a quadratic in the variable $\sin \theta$. Using the quadratic formula, we have

$$\sin \theta = \frac{-1 \pm \sqrt{33}}{8}.$$

We solve $\sin \theta = \frac{-1+\sqrt{33}}{8}$ and $\sin \theta = \frac{-1-\sqrt{33}}{8}$:

$$\sin \theta = \frac{-1 + \sqrt{33}}{8} \qquad \sin \theta = \frac{-1 - \sqrt{33}}{8}$$
$$\theta = \sin^{-1}\left(\frac{-1 + \sqrt{33}}{8}\right) \qquad \theta = \sin^{-1}\left(\frac{-1 - \sqrt{33}}{8}\right)$$
$$\theta \approx 0.6349 \qquad \theta \approx -1.0030$$

In each of the solutions above, we only get one of the possible two solutions as $\sin^{-1} x$ only returns solutions in $[-\pi/2, \pi/2]$, the 4th and 1st quadrants. Again using reference angles, we have:

$$\sin \theta = \frac{-1 + \sqrt{33}}{8} \quad \Rightarrow \quad \theta \approx 0.6349,\ 2.5067 \text{ radians}$$

and

$$\sin \theta = \frac{-1 - \sqrt{33}}{8} \quad \Rightarrow \quad \theta \approx 4.1446,\ 5.2802 \text{ radians}.$$

These points are also shown in Figure 10.49 with white–filled dots.

When the graph of the polar function $r = f(\theta)$ intersects the pole, it means that $f(\alpha) = 0$ for some angle α. Making this substitution in the formula for $\frac{dy}{dx}$ given in Key Idea 44 we see

$$\frac{dy}{dx} = \frac{f'(\alpha) \sin \alpha + f(\alpha) \cos \alpha}{f'(\alpha) \cos \alpha + f(\alpha) \sin \alpha} = \frac{\sin \alpha}{\cos \alpha} = \tan \alpha.$$
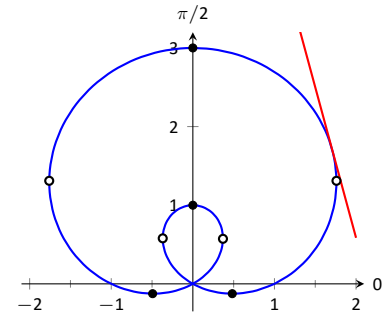

Figure 10.49: The limaçon in Example 1 with its tangent line at $\theta = \pi/4$ and points of vertical and horizontal tangency.

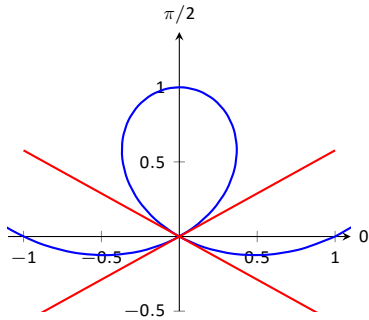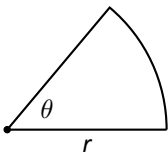
Notes:

This equation makes an interesting point. It tells us the slope of the tangent line at the pole is $\tan\alpha$; some of our previous work (see, for instance, Example 10.4.3) shows us that the line through the pole with slope $\tan\alpha$ has polar equation $\theta = \alpha$. Thus when a polar graph touches the pole at $\theta = \alpha$, the equation of the tangent line at the pole is $\theta = \alpha$.

Example 2   Finding tangent lines at the pole
Let $r = 1 + 2\sin\theta$, a limaçon. Find the equations of the lines tangent to the graph at the pole.

SOLUTION   We need to know when $r = 0$.

$$1 + 2\sin\theta = 0$$
$$\sin\theta = -1/2$$
$$\theta = \frac{7\pi}{6}, \ \frac{11\pi}{6}.$$

Thus the equations of the tangent lines, in polar, are $\theta = 7\pi/6$ and $\theta = 11\pi/6$. In rectangular form, the tangent lines are $y = \tan(7\pi/6)x = \frac{x}{\sqrt{3}}$ and $y = \tan(11\pi/6)x = -\frac{x}{\sqrt{3}}$. The full limaçon can be seen in Figure 10.49; we zoom in on the tangent lines in Figure 10.50.

Figure 10.50: Graphing the tangent lines at the pole in Example 2.

Area

When using rectangular coordinates, the equations $x = h$ and $y = k$ defined vertical and horizontal lines, respectively, and combinations of these lines create rectangles (hence the name "rectangular coordinates"). It is then somewhat natural to use rectangles to approximate area as we did when learning about the definite integral.

When using polar coordinates, the equations $\theta = \alpha$ and $r = c$ form lines through the origin and circles centered at the origin, respectively, and combinations of these curves form sectors of circles. It is then somewhat natural to calculate the area of regions defined by polar functions by first approximating with sectors of circles.

Consider Figure 10.51 (a) where a region defined by $r = f(\theta)$ on $[\alpha, \beta]$ is given. (Note how the "sides" of the region are the lines $\theta = \alpha$ and $\theta = \beta$, whereas in rectangular coordinates the "sides" of regions were often the vertical lines $x = a$ and $x = b$.)

Partition the interval $[\alpha, \beta]$ into n equally spaced subintervals as $\alpha = \theta_0 < \theta_1 < \ldots < \theta_n = \beta$. The radian of each subinterval is $\Delta\theta = (\beta - \alpha)/n$, representing a small change in angle. The area of the region defined by the ith

Note: Recall that the area of a sector of a circle with radius r subtended by an angle θ is $A = \frac{1}{2}\theta r^2$.

Notes:

subinterval $[\theta_{i-1}, \theta_i]$ can be approximated with a sector of a circle with radius $f(c_i)$, for some c_i in $[\theta_{i-1}, \theta_i]$. The area of this sector is $\frac{1}{2}[f(c_i)]^2\Delta\theta$. This is shown in part (b) of the figure, where $[\alpha, \beta]$ has been divided into 4 subintervals. We approximate the area of the whole region by summing the areas of all sectors:

$$\text{Area} \approx \sum_{i=1}^{n} \frac{1}{2}[f(c_i)]^2\Delta\theta.$$

This is a Riemann sum. By taking the limit of the sum as $n \to \infty$, we find the exact area of the region in the form of a definite integral.

> **Theorem 85   Area of a Polar Region**
> Let f be continuous and non-negative on $[\alpha, \beta]$, where $0 \le \beta - \alpha \le 2\pi$. The area A of the region bounded by the curve $r = f(\theta)$ and the lines $\theta = \alpha$ and $\theta = \beta$ is
>
> $$A = \frac{1}{2}\int_{\alpha}^{\beta} [f(\theta)]^2\, d\theta = \frac{1}{2}\int_{\alpha}^{\beta} r^2\, d\theta$$

The theorem states that $0 \le \beta - \alpha \le 2\pi$. This ensures that region does not overlap itself, which would give a result that does not correspond directly to the area.

Example 3   Area of a polar region
Find the area of the circle defined by $r = \cos\theta$. (Recall this circle has radius $1/2$.)

SOLUTION   This is a direct application of Theorem 85. The circle is traced out on $[0, \pi]$, leading to the integral

$$\begin{aligned}
\text{Area} &= \frac{1}{2}\int_{0}^{\pi} \cos^2\theta\, d\theta \\
&= \frac{1}{2}\int_{0}^{\pi} \frac{1 + \cos(2\theta)}{2}\, d\theta \\
&= \frac{1}{4}\left(\theta + \frac{1}{2}\sin(2\theta)\right)\Big|_{0}^{\pi} \\
&= \frac{\pi}{4}.
\end{aligned}$$

Of course, we already knew the area of a circle with radius $1/2$. We did this example to demonstrate that the area formula is correct.

Notes:
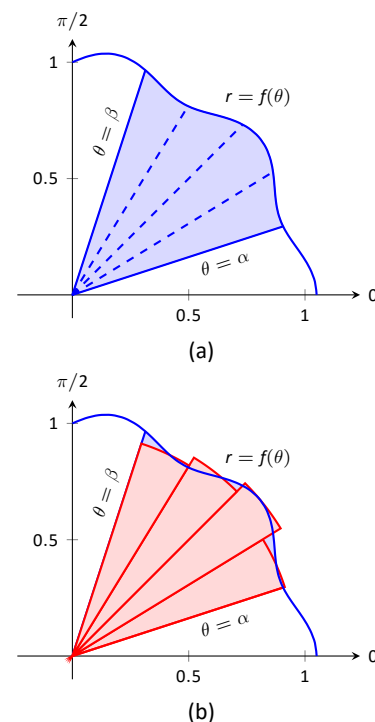
Figure 10.51: Computing the area of a polar region.

Note: Example 3 requires the use of the integral $\int \cos^2\theta\, d\theta$. This is handled well by using the half angle formula as found in the back of this text. Due to the nature of the area formula, integrating $\cos^2\theta$ and $\sin^2\theta$ is required often. We offer here these indefinite integrals as a time–saving measure.

$$\int \cos^2\theta\, d\theta = \frac{1}{2}\theta + \frac{1}{4}\sin(2\theta) + C$$

$$\int \sin^2\theta\, d\theta = \frac{1}{2}\theta - \frac{1}{4}\sin(2\theta) + C$$

Example 4 **Area of a polar region**
Find the area of the cardioid $r = 1 + \cos\theta$ bound between $\theta = \pi/6$ and $\theta = \pi/3$, as shown in Figure 10.52.

SOLUTION This is again a direct application of Theorem 85.

$$\text{Area} = \frac{1}{2} \int_{\pi/6}^{\pi/3} (1 + \cos\theta)^2 \, d\theta$$

$$= \frac{1}{2} \int_{\pi/6}^{\pi/3} (1 + 2\cos\theta + \cos^2\theta) \, d\theta$$

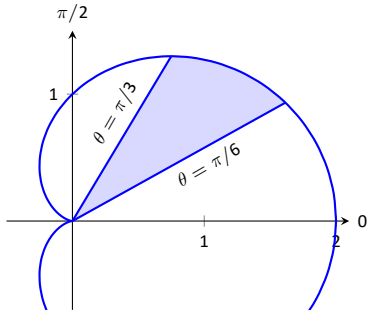$$= \frac{1}{2} \left[\theta + 2\sin\theta + \frac{1}{2}\theta + \frac{1}{4}\sin(2\theta) \right]_{\pi/6}^{\pi/3}$$

$$= \frac{1}{8} \left(\pi + 4\sqrt{3} - 4 \right).$$

Figure 10.52: Finding the area of the shaded region of a cardioid in Example 4.

Area Between Curves

Our study of area in the context of rectangular functions led naturally to finding area bounded between curves. We consider the same in the context of polar functions.

Consider the shaded region shown in Figure 10.53. We can find the area of this region by computing the area bounded by $r_2 = f_2(\theta)$ and subtracting the area bounded by $r_1 = f_1(\theta)$ on $[\alpha, \beta]$. Thus

$$\text{Area} = \frac{1}{2} \int_\alpha^\beta r_2^2 \, d\theta - \frac{1}{2} \int_\alpha^\beta r_1^2 \, d\theta = \frac{1}{2} \int_\alpha^\beta \left(r_2^2 - r_1^2 \right) d\theta.$$
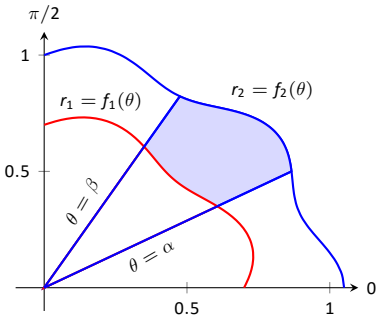
Figure 10.53: Illustrating area bound between two polar curves.

Key Idea 45 **Area Between Polar Curves**
The area A of the region bounded by $r_1 = f_1(\theta)$ and $r_2 = f_2(\theta)$, $\theta = \alpha$ and $\theta = \beta$, where $f_1(\theta) \le f_2(\theta)$ on $[\alpha, \beta]$, is

$$A = \frac{1}{2} \int_\alpha^\beta [f_2(\theta)]^2 - [f_1(\theta)]^2 \, d\theta = \frac{1}{2} \int_\alpha^\beta \left(r_2^2 - r_1^2 \right) d\theta.$$

Example 5 **Area between polar curves**
Find the area bounded between the curves $r = 1 + \cos\theta$ and $r = 3\cos\theta$, as shown in Figure 10.54.
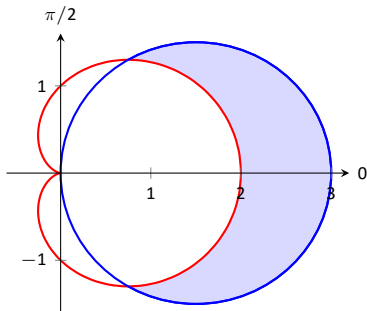

Figure 10.54: Finding the area between polar curves in Example 5.

Notes:

SOLUTION We need to find the points of intersection between these two functions. Setting them equal to each other, we find:

$$1 + \cos\theta = 3\cos\theta$$
$$\cos\theta = 1/2$$
$$\theta = \pm\pi/3$$

Thus we integrate $\frac{1}{2}\left((3\cos\theta)^2 - (1+\cos\theta)^2\right)$ on $[-\pi/3, \pi/3]$.

$$\begin{aligned}
\text{Area} &= \frac{1}{2}\int_{-\pi/3}^{\pi/3}\left((3\cos\theta)^2 - (1+\cos\theta)^2\right)\,d\theta \\
&= \frac{1}{2}\int_{-\pi/3}^{\pi/3}\left(8\cos^2\theta - 2\cos\theta - 1\right)\,d\theta \\
&= \frac{1}{2}\left(2\sin(2\theta) - 2\sin\theta + 3\theta\right)\Big|_{-\pi/3}^{\pi/3} \\
&= \pi.
\end{aligned}$$

Amazingly enough, the area between these curves has a "nice" value.

Example 6 **Area defined by polar curves**
Find the area bounded between the polar curves $r = 1$ and $r = 2\cos(2\theta)$, as shown in Figure 10.55 (a).

SOLUTION We need to find the point of intersection between the two curves. Setting the two functions equal to each other, we have

$$2\cos(2\theta) = 1 \quad\Rightarrow\quad \cos(2\theta) = \frac{1}{2} \quad\Rightarrow\quad 2\theta = \pi/3 \quad\Rightarrow\quad \theta = \pi/6.$$

In part (b) of the figure, we zoom in on the region and note that it is not really bounded *between* two polar curves, but rather *by* two polar curves, along with $\theta = 0$. The dashed line breaks the region into its component parts. Below the dashed line, the region is defined by $r = 1$, $\theta = 0$ and $\theta = \pi/6$. (Note: the dashed line lies on the line $\theta = \pi/6$.) Above the dashed line the region is bounded by $r = 2\cos(2\theta)$ and $\theta = \pi/6$. Since we have two separate regions, we find the area using two separate integrals.

Call the area below the dashed line A_1 and the area above the dashed line A_2. They are determined by the following integrals:

$$A_1 = \frac{1}{2}\int_0^{\pi/6}(1)^2\,d\theta \qquad A_2 = \frac{1}{2}\int_{\pi/6}^{\pi/4}\left(2\cos(2\theta)\right)^2\,d\theta.$$
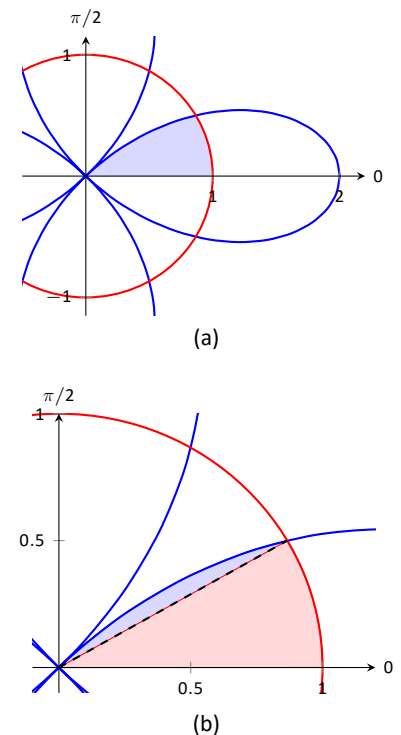
(a)

(b)

Figure 10.55: Graphing the region bounded by the functions in Example 6.

Notes:

(The upper bound of the integral computing A_2 is $\pi/4$ as $r = 2\cos(2\theta)$ is at the pole when $\theta = \pi/4$.)

We omit the integration details and let the reader verify that $A_1 = \pi/12$ and $A_2 = \pi/12 - \sqrt{3}/8$; the total area is $A = \pi/6 - \sqrt{3}/8$.

Arc Length

As we have already considered the arc length of curves defined by rectangular and parametric equations, we now consider it in the context of polar equations. Recall that the arc length L of the graph defined by the parametric equations $x = f(t), y = g(t)$ on $[a, b]$ is

$$L = \int_a^b \sqrt{[f'(t)]^2 + [g'(t)]^2}\, dt = \int_a^b \sqrt{[x'(t)]^2 + [y'(t)]^2}\, dt. \qquad (10.1)$$

Now consider the polar function $r = f(\theta)$. We again use the identities $x = f(\theta)\cos\theta$ and $y = f(\theta)\sin\theta$ to create parametric equations based on the polar function. We compute $x'(\theta)$ and $y'(\theta)$ as done before when computing $\frac{dy}{dx}$, then apply Equation (10.1).

The expression $[x'(\theta)]^2 + [y'(\theta)]^2$ can be simplified a great deal; we leave this as an exercise and state that

$$[x'(\theta)]^2 + [y'(\theta)]^2 = [f'(\theta)]^2 + [f(\theta)]^2.$$

This leads us to the arc length formula.

Key Idea 46   Arc Length of Polar Curves

Let $r = f(\theta)$ be a polar function with f' continuous on an open interval I containing $[\alpha, \beta]$, on which the graph traces itself only once. The arc length L of the graph on $[\alpha, \beta]$ is

$$L = \int_\alpha^\beta \sqrt{[f'(\theta)]^2 + [f(\theta)]^2}\, d\theta = \int_\alpha^\beta \sqrt{(r')^2 + r^2}\, d\theta.$$

Example 7   Arc Length of Polar Curves
Find the arc length of the cardioid $r = 1 + \cos\theta$.

Notes:

SOLUTION With $r = 1 + \cos\theta$, we have $r' = -\sin\theta$. The cardioid is traced out once on $[0, 2\pi]$, giving us our bounds of integration. Applying Key Idea 46 we have

$$L = \int_0^{2\pi} \sqrt{(-\sin\theta)^2 + (1 + \cos\theta)^2}\, d\theta$$

$$= \int_0^{2\pi} \sqrt{\sin^2\theta + (1 + 2\cos\theta + \cos\theta)}\, d\theta$$

$$= \int_0^{2\pi} \sqrt{2 + 2\cos\theta}\, d\theta$$

$$= \int_0^{2\pi} \sqrt{2 + 2\cos\theta}\, \frac{\sqrt{2 - 2\cos\theta}}{\sqrt{2 - 2\cos\theta}}\, d\theta$$

$$= \int_0^{2\pi} \frac{\sqrt{4 - 4\cos^2\theta}}{\sqrt{2 - 2\cos\theta}}\, d\theta$$

$$= 2\int_0^{2\pi} \frac{\sqrt{1 - \cos^2\theta}}{\sqrt{2 - 2\cos\theta}}\, d\theta$$

$$= 2\int_0^{2\pi} \frac{|\sin\theta|}{\sqrt{2 - 2\cos\theta}}\, d\theta$$

Since the $\sin\theta > 0$ on $[0, \pi]$ and $\sin\theta < 0$ on $[\pi, 2\pi]$ we separate the integral into two parts

$$2\int_0^{\pi} \frac{\sin\theta}{\sqrt{2 - 2\cos\theta}}\, d\theta - 2\int_{\pi}^{2\pi} \frac{\sin\theta}{\sqrt{2 - 2\cos\theta}}\, d\theta$$

Using the symmetry of the cardioid and u-substitution ($u = 2 - 2\cos\theta$) we simplify the integration to

$$L = 4\int_0^{\pi} \frac{\sin\theta}{\sqrt{2 - 2\cos\theta}}\, d\theta$$

$$= 2\int_0^{4} \frac{1}{\sqrt{u}}\, du$$

$$= 4u^{1/2}\Big|_0^{4} = 8.$$

Example 8 **Arc length of a limaçon**

Find the arc length of the limaçon $r = 1 + 2\sin t$.

Notes:

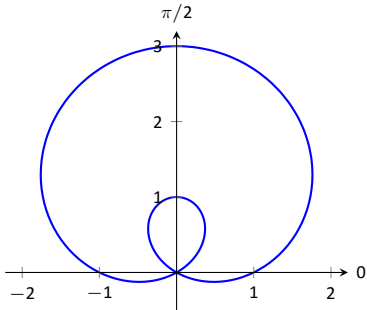**SOLUTION**   With $r = 1 + 2\sin t$, we have $r' = 2\cos t$. The limaçon is traced out once on $[0, 2\pi]$, giving us our bounds of integration. Applying Key Idea 46, we have

$$L = \int_0^{2\pi} \sqrt{(2\cos\theta)^2 + (1 + 2\sin\theta)^2}\, d\theta$$

$$= \int_0^{2\pi} \sqrt{4\cos^2\theta + 4\sin^2\theta + 4\sin\theta + 1}\, d\theta$$

$$= \int_0^{2\pi} \sqrt{4\sin\theta + 5}\, d\theta$$

$$\approx 13.3649.$$

The final integral cannot be solved in terms of elementary functions, so we resorted to a numerical approximation. (Simpson's Rule, with $n = 4$, approximates the value with 13.0608. Using $n = 22$ gives the value above, which is accurate to 4 places after the decimal.)

Figure 10.56: The limaçon in Example 8 whose arc length is measured.

Surface Area

The formula for arc length leads us to a formula for surface area. The following Key Idea is based on Key Idea 42.

Key Idea 47   Surface Area of a Solid of Revolution

Consider the graph of the polar equation $r = f(\theta)$, where f' is continuous on an open interval containing $[\alpha, \beta]$ on which the graph does not cross itself.

1. The surface area of the solid formed by revolving the graph about the initial ray ($\theta = 0$) is:

$$\text{Surface Area} = 2\pi \int_\alpha^\beta f(\theta)\sin\theta\sqrt{[f'(\theta)]^2 + [f(\theta)]^2}\, d\theta.$$

2. The surface area of the solid formed by revolving the graph about the line $\theta = \pi/2$ is:

$$\text{Surface Area} = 2\pi \int_\alpha^\beta f(\theta)\cos\theta\sqrt{[f'(\theta)]^2 + [f(\theta)]^2}\, d\theta.$$

Notes:

Example 9 **Surface area determined by a polar curve**
Find the surface area formed by revolving one petal of the rose curve $r = \cos(2\theta)$ about its central axis (see Figure 10.57).

SOLUTION We choose, as implied by the figure, to revolve the portion of the curve that lies on $[0, \pi/4]$ about the initial ray. Using Key Idea 47 and the fact that $f'(\theta) = -2\sin(2\theta)$, we have

$$\text{Surface Area} = 2\pi \int_0^{\pi/4} \cos(2\theta)\sin(\theta)\sqrt{\big(-2\sin(2\theta)\big)^2 + \big(\cos(2\theta)\big)^2}\, d\theta$$

$$\approx 1.36707.$$

The integral is another that cannot be evaluated in terms of elementary functions. Simpson's Rule, with $n = 4$, approximates the value at 1.36751.

This chapter has been about curves in the plane. While there is great mathematics to be discovered in the two dimensions of a plane, we live in a three dimensional world and hence we should also look to do mathematics in 3D — that is, in *space*. The next chapter begins our exploration into space by introducing the topic of *vectors*, which are incredibly useful and powerful mathematical objects.
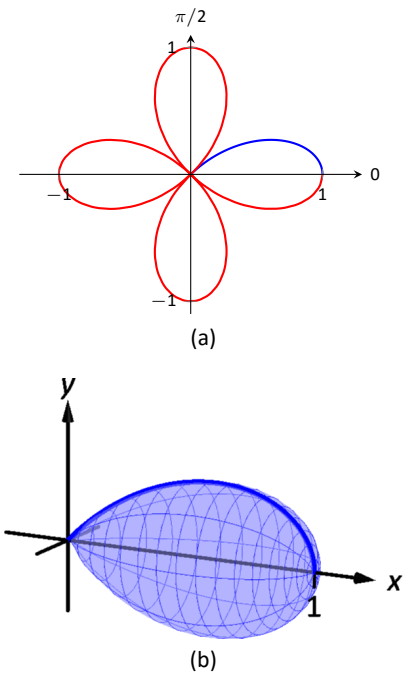
(a)

(b)

Figure 10.57: Finding the surface area of a rose–curve petal that is revolved around its central axis.

Notes:

Exercises 10.5

Terms and Concepts

1. Given polar equation $r = f(\theta)$, how can one create parametric equations of the same curve?

2. With rectangular coordinates, it is natural to approximate area with _____; with polar coordinates, it is natural to approximate area with _____.

Problems

In Exercises 3–10, find:

(a) $\dfrac{dy}{dx}$

(b) the equation of the tangent and normal lines to the curve at the indicated θ–value.

3. $r = 1; \quad \theta = \pi/4$

4. $r = \cos\theta; \quad \theta = \pi/4$

5. $r = 1 + \sin\theta; \quad \theta = \pi/6$

6. $r = 1 - 3\cos\theta; \quad \theta = 3\pi/4$

7. $r = \theta; \quad \theta = \pi/2$

8. $r = \cos(3\theta); \quad \theta = \pi/6$

9. $r = \sin(4\theta); \quad \theta = \pi/3$

10. $r = \dfrac{1}{\sin\theta - \cos\theta}; \quad \theta = \pi$

In Exercises 11–14, find the values of θ in the given interval where the graph of the polar function has horizontal and vertical tangent lines.

11. $r = 3; \quad [0, 2\pi]$

12. $r = 2\sin\theta; \quad [0, \pi]$

13. $r = \cos(2\theta); \quad [0, 2\pi]$

14. $r = 1 + \cos\theta; \quad [0, 2\pi]$

In Exercises 15–18, find the equation of the lines tangent to the graph at the pole.

15. $r = \sin\theta; \quad [0, \pi]$

16. $r = \cos 3\theta; \quad [0, \pi]$

17. $r = \cos 2\theta; \quad [0, 2\pi]$

18. $r = \sin 2\theta; \quad [0, 2\pi]$

In Exercises 19–29, find the area of the described region.

19. Enclosed by the circle: $r = 4\sin\theta, \quad \frac{\pi}{3} \leq \theta \leq \frac{2\pi}{3}$

20. Enclosed by the circle $r = 5$

21. Enclosed by one petal of $r = \sin(3\theta)$

22. Enclosed by the cardioid $r = 1 - \sin\theta$

23. Enclosed by the inner loop of the limaçon $r = 1 + 2\cos\theta$

24. Enclosed by the outer loop of the limaçon $r = 1 + 2\cos\theta$ (including area enclosed by the inner loop)

25. Enclosed between the inner and outer loop of the limaçon $r = 1 + 2\cos\theta$

26. Enclosed by $r = 2\cos\theta$ and $r = 2\sin\theta$, as shown:
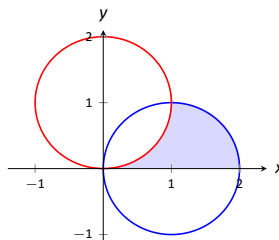
27. Enclosed by $r = \cos(3\theta)$ and $r = \sin(3\theta)$, as shown:
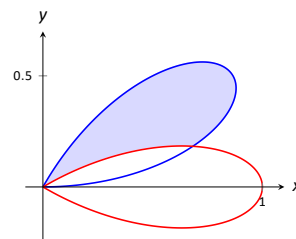
28. Enclosed by $r = \cos\theta$ and $r = \sin(2\theta)$, as shown:
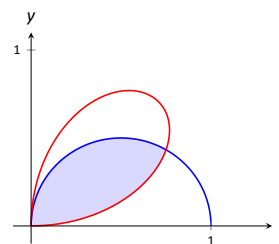
29. Enclosed by $r = \cos\theta$ and $r = 1 - \cos\theta$, as shown:
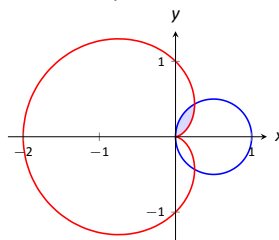
In Exercises 30–34, answer the questions involving arc length.

30. Let $x(\theta) = f(\theta)\cos\theta$ and $y(\theta) = f(\theta)\sin\theta$. Show, as suggested by the text, that

$$x'(\theta)^2 + y'(\theta)^2 = f'(\theta)^2 + f(\theta)^2.$$

31. Use the arc length formula to compute the arc length of the circle $r = 2$.
32. Use the arc length formula to compute the arc length of the circle $r = 4 \sin \theta$.
33. Approximate the arc length of one petal of the rose curve $r = \sin(3\theta)$ with Simpson's Rule and $n = 4$.
34. Approximate the arc length of the cardioid $r = 1 + \cos \theta$ with Simpson's Rule and $n = 6$.

In Exercises 35–39, answer the questions involving surface area.

35. Use Key Idea 47 to find the surface area of the sphere formed by revolving the circle $r = 2$ about the initial ray.

36. Use Key Idea 47 to find the surface area of the sphere formed by revolving the circle $r = 2 \cos \theta$ about the initial ray.

37. Find the surface area of the solid formed by revolving the cardioid $r = 1 + \cos \theta$ about the initial ray.

38. Find the surface area of the solid formed by revolving the circle $r = 2 \cos \theta$ about the line $\theta = \pi/2$.

39. Find the surface area of the solid formed by revolving the line $r = 3 \sec \theta$, $-\pi/4 \le \theta \le \pi/4$, about the line $\theta = \pi/2$.

Notes:

Calculus III

11: VECTORS

This chapter introduces a new mathematical object, the **vector**. Defined in Section 11.2, we will see that vectors provide a powerful language for describing quantities that have magnitude and direction. A simple example of such a quantity is force: when applying a force, one is generally interested in how much force is applied (i.e., the magnitude of the force) and the direction in which the force is applied. Vectors will play an important role in many of the subsequent chapters in this text.

This chapter begins with moving our mathematics out of the plane and into "space." That is, we begin to think mathematically not only in two dimensions, but in three. With this foundation, we can explore vectors both in the plane and in space.

11.1 Introduction to Cartesian Coordinates in Space

Up to this point in this text we have considered mathematics in a 2–dimensional world. We have plotted graphs on the *x-y* plane using rectangular and polar coordinates and found the area of regions in the plane. We have considered properties of *solid* objects, such as volume and surface area, but only by first defining a curve in the plane and then rotating it out of the plane.

While there is wonderful mathematics to explore in "2D," we live in a "3D" world and eventually we will want to apply mathematics involving this third dimension. In this section we introduce Cartesian coordinates in space and explore basic surfaces. This will lay a foundation for much of what we do in the remainder of the text.

Each point P in space can be represented with an ordered triple, $P = (a, b, c)$, where a, b and c represent the relative position of P along the *x*-, *y*- and *z*-axes, respectively. Each axis is perpendicular to the other two.

Visualizing points in space on paper can be problematic, as we are trying to represent a 3-dimensional concept on a 2–dimensional medium. We cannot draw three lines representing the three axes in which each line is perpendicular to the other two. Despite this issue, standard conventions exist for plotting shapes in space that we will discuss that are more than adequate.

One convention is that the axes must conform to the **right hand rule**. This rule states that when the fingers of the right hand extend in the direction of the positive *x*-axis and curve toward the positive *y*-axis, then the extended thumb will point in the direction of the positive *z*-axis. (It may take some thought to verify this, but this system is inherently different from the one created by using

Figure 11.1: Illustrating the right hand rule. Figure courtesy of user:Schorschi2 / Wikimedia Commons / Public Domain.
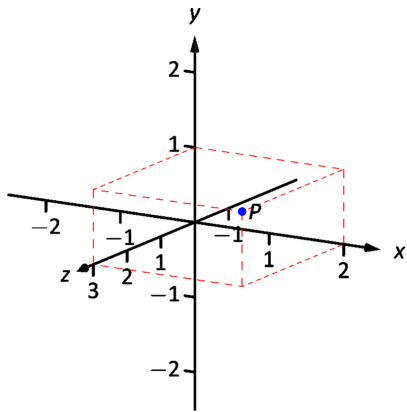
the "left hand rule.") Another way to view the rule is that when the index finger of the right hand extends in the direction of the positive *x*-axis, and the middle finger (bent "inward" so it is perpendicular to the palm) points along the positive *y*-axis, then the extended thumb will point in the direction of the positive *z*-axis.

As long as the coordinate axes are positioned so that they follow this rule, it does not matter how the axes are drawn on paper. There are two popular methods that we briefly discuss.

In Figure 11.2 we see the point $P = (2, 1, 3)$ plotted on a set of axes. The basic convention here is that the *x*-*y* plane is drawn in its standard way, with the *z*-axis down to the left. The perspective is that the paper represents the *x*-*y* plane and the positive *z* axis is coming up, off the page. This method is preferred by many engineers. Because it can be hard to tell where a single point lies in relation to all the axes, dashed lines have been added to let one see how far along each axis the point lies.

One can also consider the *x*-*y* plane as being a horizontal plane in, say, a room, where the positive *z*-axis is pointing up. When one steps back and looks at this room, one might draw the axes as shown in Figure 11.3. The same point *P* is drawn, again with dashed lines. This point of view is preferred by most mathematicians, and is the convention adopted by this text.

Figure 11.2: Plotting the point $P = (2, 1, 3)$ in space.

Measuring Distances

It is of critical importance to know how to measure distances between points in space. The formula for doing so is based on measuring distance in the plane, and is known (in both contexts) as the Euclidean measure of distance.

> **Definition 51   Distance In Space**
> Let $P = (x_1, y_1, z_1)$ and $Q = (x_2, y_2, z_2)$ be points in space. The distance *D* between *P* and *Q* is
> $$D = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2 + (z_2 - z_1)^2}.$$

We refer to the line segment that connects points *P* and *Q* in space as $\overline{PQ}$, and refer to the length of this segment as $\left\lVert \overline{PQ} \right\rVert$. The above distance formula allows us to compute the length of this segment.

Figure 11.3: Plotting the point $P = (2, 1, 3)$ in space with a perspective used in this text.

Example 1   Length of a line segment
Let $P = (1, 4, -1)$ and let $Q = (2, 1, 1)$. Draw the line segment $\overline{PQ}$ and find its length.

Notes:

SOLUTION   The points P and Q are plotted in Figure 11.4; no special consideration needs to be made to draw the line segment connecting these two points; simply connect them with a straight line. One *cannot* actually measure this line on the page and deduce anything meaningful; its true length must be measured analytically. Applying Definition 51, we have

$$\left\|\overline{PQ}\right\| = \sqrt{(2-1)^2 + (1-4)^2 + (1-(-1))^2} = \sqrt{14}.$$

Spheres

Just as a circle is the set of all points in the *plane* equidistant from a given point (its center), a sphere is the set of all points in *space* that are equidistant from a given point. Definition 51 allows us to write an equation of the sphere.

We start with a point $C = (a, b, c)$ which is to be the center of a sphere with radius r. If a point $P = (x, y, z)$ lies on the sphere, then P is r units from C; that is,

$$\left\|\overline{PC}\right\| = \sqrt{(x-a)^2 + (y-b)^2 + (z-c)^2} = r.$$

Squaring both sides, we get the standard equation of a sphere in space with center at $C = (a, b, c)$ with radius r, as given in the following Key Idea.

Figure 11.4: Plotting points P and Q in Example 1.

Key Idea 48   Standard Equation of a Sphere in Space

The standard equation of the sphere with radius r, centered at $C = (a, b, c)$, is

$$(x-a)^2 + (y-b)^2 + (z-c)^2 = r^2.$$

Watch the video:
Example of Equation of a Sphere at
`https://youtu.be/fE_PWxyohXQ`

Example 2   Equation of a sphere
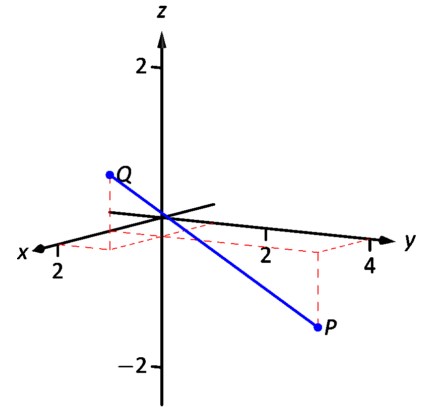Find the center and radius of the sphere defined by $x^2 + 2x + y^2 - 4y + z^2 - 6z = 2$.

Notes:

SOLUTION   To determine the center and radius, we must put the equation in standard form. This requires us to complete the square (three times).

$$x^2 + 2x + y^2 - 4y + z^2 - 6z = 2$$
$$(x^2 + 2x + 1) + (y^2 - 4y + 4) + (z^2 - 6z + 9) - 14 = 2$$
$$(x+1)^2 + (y-2)^2 + (z-3)^2 = 16$$

The sphere is centered at $(-1, 2, 3)$ and has a radius of 4.

The equation of a sphere is an example of an implicit function defining a surface in space. In the case of a sphere, the variables x, y and z are all used. We now consider situations where surfaces are defined where one or two of these variables are absent.

Introduction to Planes in Space

The coordinate axes naturally define three planes (shown in Figure 11.5), the **coordinate planes**: the x-y plane, the y-z plane and the x-z plane. The x-y plane is characterized as the set of all points in space where the z-value is 0.   This, in fact, gives us an equation that describes this plane: $z = 0$. Likewise, the x-z plane is all points where the y-value is 0, characterized by $y = 0$.

the x-y plane   the y-z plane   the x-z plane
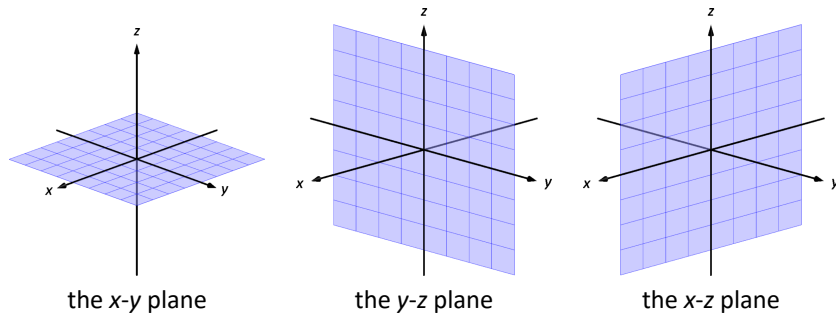
Figure 11.5: The coordinate planes.
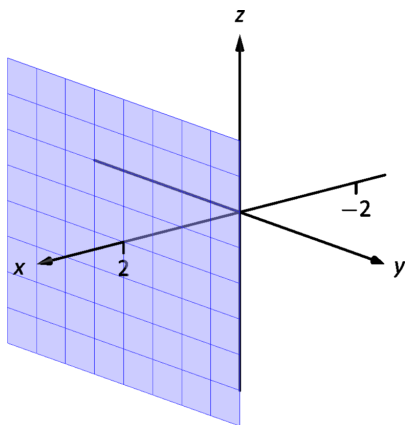
Example 3   **A plane in three dimensions**
The equation $x = 2$ describes all points in space where the x-value is 2. This is a plane, parallel to the y-z coordinate plane, shown in Figure 11.6.

Example 4   **Regions defined by planes**
Sketch the region defined by the inequalities $-1 \leq y \leq 2$.

Figure 11.6: The plane $x = 2$.

Notes:

SOLUTION   The region is all points between the planes $y = -1$ and $y = 2$. These planes are sketched in Figure 11.7, which are parallel to the x-z plane. Thus the region extends infinitely in the x and z directions, and is bounded by planes in the y direction.

Cylinders

The equation $x = 1$ obviously lacks the y and z variables, meaning it defines points where the y and z coordinates can take on any value. Now consider the equation $x^2 + y^2 = 1$ *in space*. In *the plane*, this equation describes a circle of radius 1, centered at the origin. In space, the z coordinate is not specified, meaning it can take on any value. In Figure 11.8 (a), we show part of the graph of the equation $x^2 + y^2 = 1$ by sketching 3 circles: the bottom one has a constant z-value of -1.5, the middle one has a z-value of 0 and the top circle has a z-value of 1. By plotting *all* possible z-values, we get the surface shown in Figure 11.8 (b). This surface looks like a "tube," or a "cylinder", which leads to our next definition.

> **Definition 52   Cylinder**
>
> Let C be a curve in a plane and let L be a line not parallel to C. A **cylinder** is the set of all lines parallel to L that pass through C. The curve C is the **directrix** of the cylinder, and the lines are the **rulings**.

In this text, we consider curves C that lie in planes parallel to one of the coordinate planes, and lines L that are perpendicular to these planes, forming **right cylinders**. Thus the directrix can be defined using equations involving 2 variables, and the rulings will be parallel to the axis of the 3rd variable.

In the example preceding the definition, the curve $x^2 + y^2 = 1$ in the x-y plane is the directrix and the rulings are lines parallel to the z-axis. (Any circle shown in Figure 11.8 can be considered a directrix; we simply choose the one where $z = 0$.) Sample rulings can also be viewed in part (b) of the figure. More examples will help us understand this definition.

Example 5   Graphing cylinders
Graph the cylinder following cylinders.

$$1. \quad z = y^2 \qquad 2. \quad x = \sin z$$

Notes:

(a)

(b)
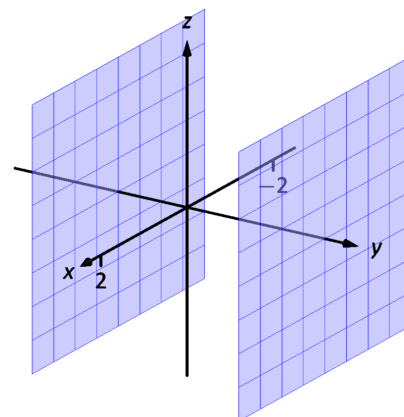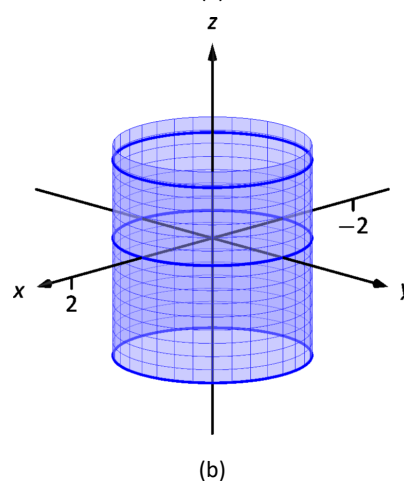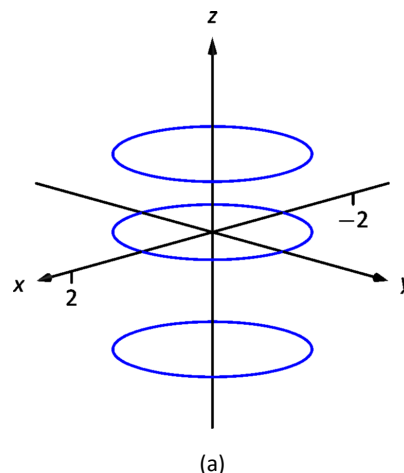
Figure 11.8: Sketching $x^2 + y^2 = 1$.

Solution

1. We can view the equation $z = y^2$ as a parabola in the y-z plane, as illustrated in Figure 11.9 (a). As x does not appear in the equation, the rulings are lines through this parabola parallel to the x-axis, shown in (b). These rulings give a general idea as to what the surface looks like, drawn in (c).

(a)   (b)   (c)
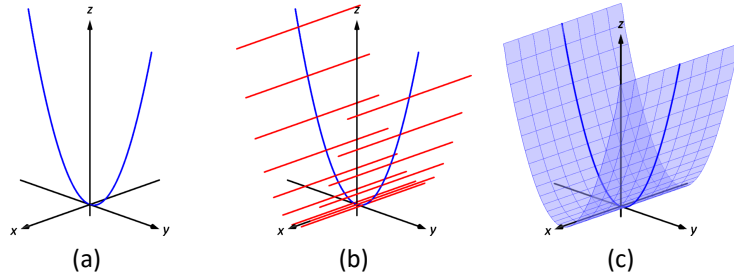
Figure 11.9: Sketching the cylinder defined by $z = y^2$.

2. We can view the equation $x = \sin z$ as a sine curve that exists in the x-z plane, as shown in Figure 11.10 (a). The rules are parallel to the y axis as the variable y does not appear in the equation $x = \sin z$; some of these are shown in part (b). The surface is shown in part (c) of the figure.

(a)   (b)   (c)
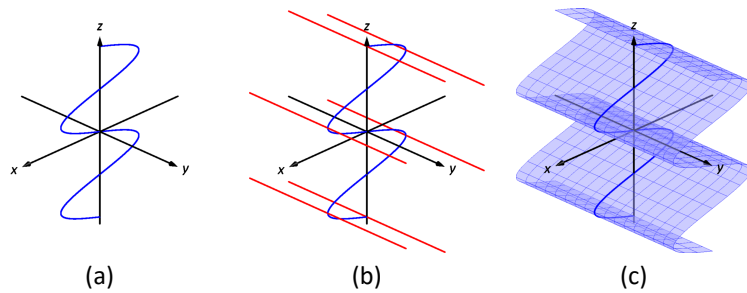
Figure 11.10: Sketching the cylinder defined by $x = \sin z$.

Surfaces of Revolution

One of the applications of integration we learned previously was to find the volume of solids of revolution – solids formed by revolving a curve about a horizontal or vertical axis. We now consider how to find the equation of the surface of such a solid.

Consider the surface formed by revolving $y = \sqrt{x}$ about the x-axis. Cross–sections of this surface parallel to the y-z plane are circles, as shown in Figure 11.11(a). Each circle has equation of the form $y^2 + z^2 = r^2$ for some radius

Notes:

r. The radius is a function of *x*; in fact, it is $r(x) = \sqrt{x}$. Thus the equation of the surface shown in Figure 11.11(b) is $y^2 + z^2 = (\sqrt{x})^2$.

We generalize the above principles to give the equations of surfaces formed by revolving curves about the coordinate axes.

Key Idea 49   Surfaces of Revolution, Part 1

Let *r* be a radius function.

1. The equation of the surface formed by revolving $y = r(x)$ or $z = r(x)$ about the *x*-axis is $y^2 + z^2 = r(x)^2$.

2. The equation of the surface formed by revolving $x = r(y)$ or $z = r(y)$ about the *y*-axis is $x^2 + z^2 = r(y)^2$.

3. The equation of the surface formed by revolving $x = r(z)$ or $y = r(z)$ about the *z*-axis is $x^2 + y^2 = r(z)^2$.

Example 6   Finding equation of a surface of revolution
Let $y = \sin z$ on $[0, \pi]$. Find the equation of the surface of revolution formed by revolving $y = \sin z$ about the *z*-axis.

SOLUTION   Using Key Idea 49, we find the surface has equation $x^2 + y^2 = \sin^2 z$. The curve is sketched in Figure 11.12(a) and the surface is drawn in Figure 11.12(b).

Note how the surface (and hence the resulting equation) is the same if we began with the curve $x = \sin z$, which is also drawn in Figure 11.12(a).

This particular method of creating surfaces of revolution is limited. For instance, in Example 6.3.5 of Section 6.3 we found the volume of the solid formed by revolving $y = \sin x$ about the *y*-axis. Our current method of forming surfaces can only rotate $y = \sin x$ about the *x*-axis. Trying to rewrite $y = \sin x$ as a function of *y* is not trivial, as simply writing $x = \sin^{-1} y$ only gives part of the region we desire.

What we desire is a way of writing the surface of revolution formed by rotating $y = f(x)$ about the *y*-axis. We start by first recognizing this surface is the same as revolving $z = f(x)$ about the *z*-axis. This will give us a more natural way of viewing the surface.

A value of *x* is a measurement of distance from the *z*-axis. At the distance *r*, we plot a *z*-height of $f(r)$. When rotating $f(x)$ about the *z*-axis, we want all points a distance of *r* from the *z*-axis in the *x*-*y* plane to have a *z*-height of $f(r)$. All such

Notes:

(a)

(b)
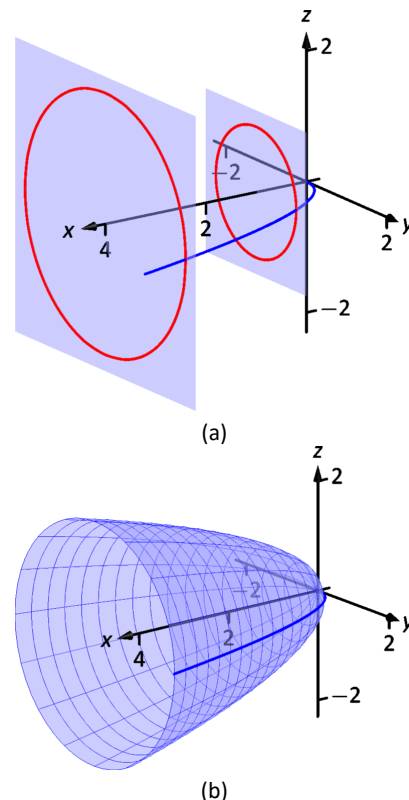
Figure 11.11: Introducing surfaces of revolution.

(a)

(b)
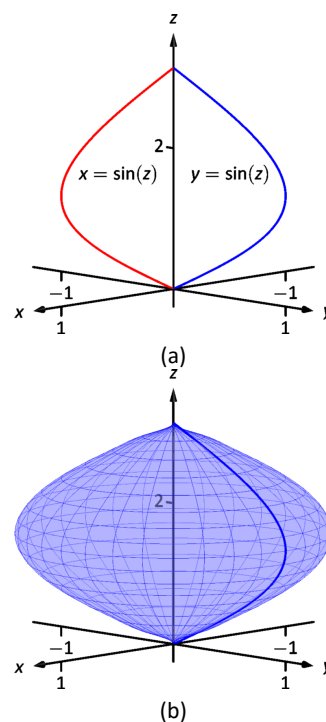
Figure 11.12: Revolving $y = \sin z$ about the *z*-axis in Example 6.

639

points satisfy the equation $r^2 = x^2 + y^2$; hence $r = \sqrt{x^2 + y^2}$. Replacing r with $\sqrt{x^2 + y^2}$ in $f(r)$ gives $z = f(\sqrt{x^2 + y^2})$. This is the equation of the surface.

Key Idea 50 **Surfaces of Revolution, Part 2**

Let $z = f(x)$, $x \geq 0$, be a curve in the x-z plane. The surface formed by revolving this curve about the z-axis has equation $z = f(\sqrt{x^2 + y^2})$.

Example 7 **Finding equation of surface of revolution**

Find the equation of the surface found by revolving $z = \sin x$ about the z-axis.

SOLUTION Using Key Idea 50, the surface has equation $z = \sin\left(\sqrt{x^2 + y^2}\right)$. The curve and surface are graphed in Figure 11.13.

Quadric Surfaces

Spheres, planes and cylinders are important surfaces to understand. We now consider one last type of surface, a **quadric surface**. The definition may look intimidating, but we will show how to analyze these surfaces in an illuminating way.

Definition 53 **Quadric Surface**

A **quadric surface** is the graph of the general second–degree equation in three variables:

$$Ax^2 + By^2 + Cz^2 + Dxy + Exz + Fyz + Gx + Hy + Iz + J = 0.$$

When the coefficients D, E or F are not zero, the basic shapes of the quadric surfaces are rotated in space. We will focus on quadric surfaces where these coefficients are 0; we will not consider rotations. There are six basic quadric surfaces: the elliptic paraboloid, elliptic cone, ellipsoid, hyperboloid of one sheet, hyperboloid of two sheets, and the hyperbolic paraboloid.

We study each shape by considering **traces**, that is, intersections of each surface with a plane parallel to a coordinate plane. For instance, consider the elliptic paraboloid $z = x^2/4 + y^2$, shown in Figure 11.14. If we intersect this shape with the plane $z = d$ (i.e., replace z with d), we have the equation:

$$d = \frac{x^2}{4} + y^2.$$

(a)

(b)

Figure 11.13: Revolving $z = \sin x$ about the z-axis in Example 7.
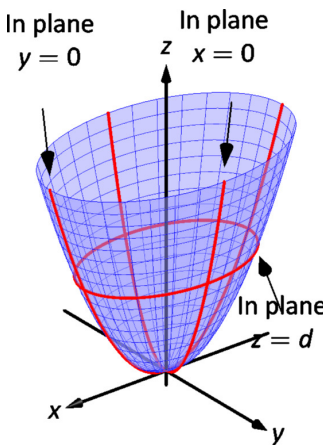
Figure 11.14: The elliptic paraboloid $z = x^2/4 + y^2$.

Notes:

Divide both sides by d:

$$1 = \frac{x^2}{4d} + \frac{y^2}{d}.$$

This describes an ellipse – so cross sections parallel to the x-y coordinate plane are ellipses. This ellipse is drawn in the figure.

Now consider cross sections parallel to the x-z plane. For instance, letting $y = 0$ gives the equation $z = x^2/4$, clearly a parabola. Intersecting with the plane $x = 0$ gives a cross section defined by $z = y^2$, another parabola. These parabolas are also sketched in the figure.

Thus we see where the elliptic paraboloid gets its name: some cross sections are ellipses, and others are parabolas.

Such an analysis can be made with each of the quadric surfaces. We give a sample equation of each, provide a sketch with representative traces, and describe these traces.

Notes:

Elliptic Paraboloid, $\quad z = \dfrac{x^2}{a^2} + \dfrac{y^2}{b^2}$

In plane
$y = 0$

In plane
$z \quad x = 0$
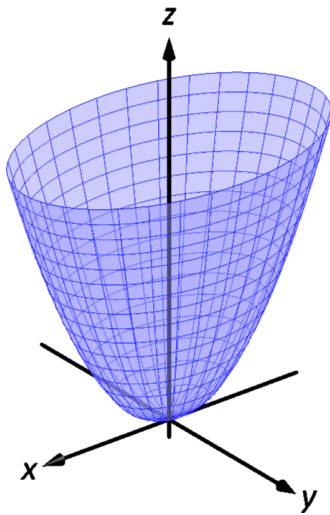
Plane	Trace
$x = d$	Parabola
$y = d$	Parabola
$z = d$	Ellipse

In plane
$z = d$
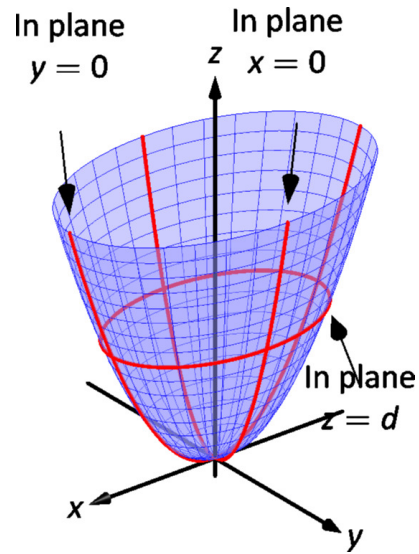
One variable in the equation of the elliptic paraboloid will be raised to the first power; above, this is the z variable. The paraboloid will "open" in the direction of this variable's axis. Thus $x = y^2/a^2 + z^2/b^2$ is an elliptic paraboloid that opens along the x-axis.

Multiplying the right hand side by (-1) defines an elliptic paraboloid that "opens" in the opposite direction.

Elliptic Cone, $\quad z^2 = \dfrac{x^2}{a^2} + \dfrac{y^2}{b^2}$
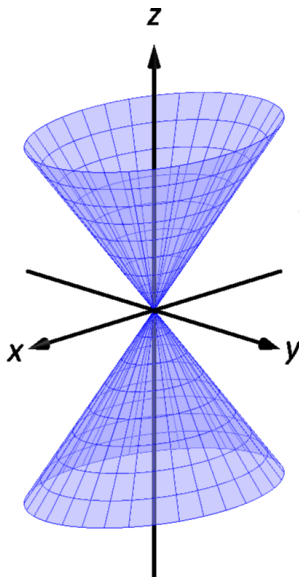
Plane	Trace
$x = 0$	Crossed Lines
$y = 0$	Crossed Lines
$x = d$	Hyperbola
$y = d$	Hyperbola
$z = d$	Ellipse

In plane
$y = 0$

In plane
$z = d$

In plane
$y = d$
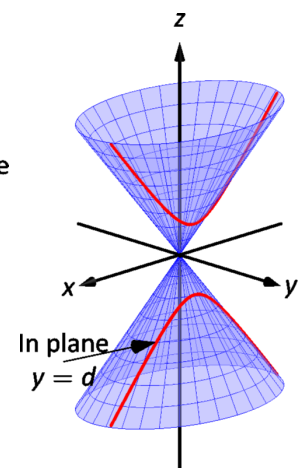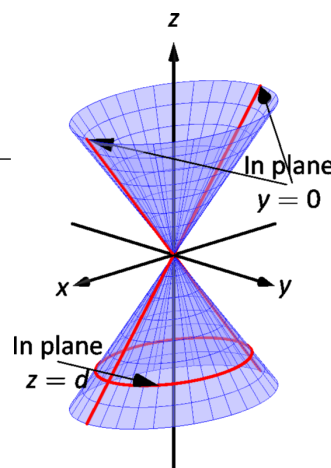
One can rewrite the equation as $z^2 - x^2/a^2 - y^2/b^2 = 0$. The one variable with a positive coefficient corresponds to the axis that the cones "open" along.

Ellipsoid, $\dfrac{x^2}{a^2} + \dfrac{y^2}{b^2} + \dfrac{z^2}{c^2} = 1$
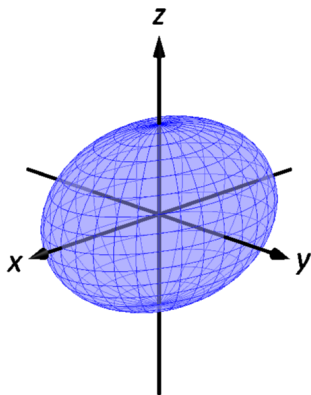
Plane	Trace
$x = d$	Ellipse
$y = d$	Ellipse
$z = d$	Ellipse
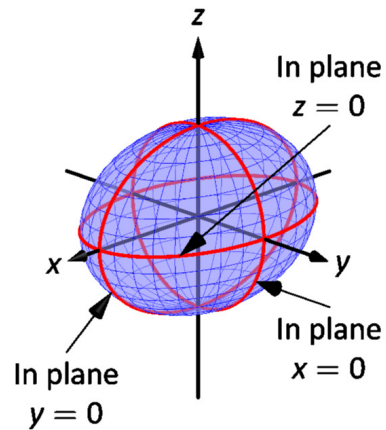
If $a = b = c \neq 0$, the ellipsoid is a sphere with radius a; compare to Key Idea 48.

Hyperboloid of One Sheet, $\dfrac{x^2}{a^2} + \dfrac{y^2}{b^2} - \dfrac{z^2}{c^2} = 1$
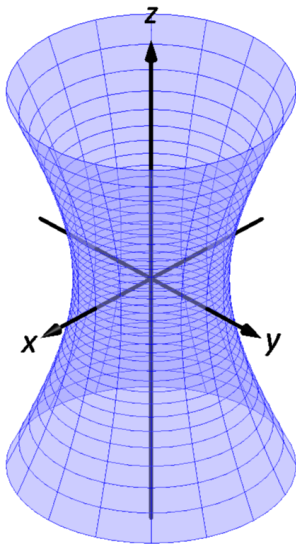
Plane	Trace
$x = d$	Hyperbola
$y = d$	Hyperbola
$z = d$	Ellipse
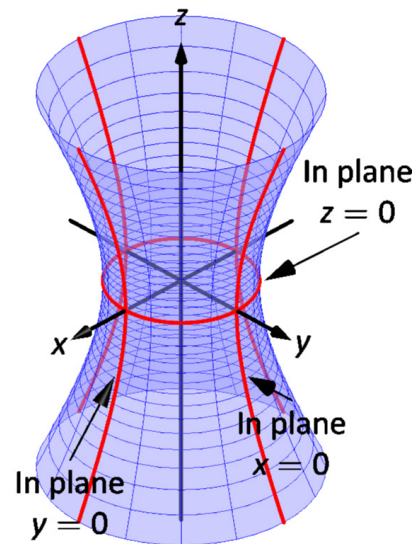
The one variable with a negative coefficient corresponds to the axis along which the hyperboloid "opens".

Hyperboloid of Two Sheets, $\dfrac{z^2}{c^2} - \dfrac{x^2}{a^2} - \dfrac{y^2}{b^2} = 1$
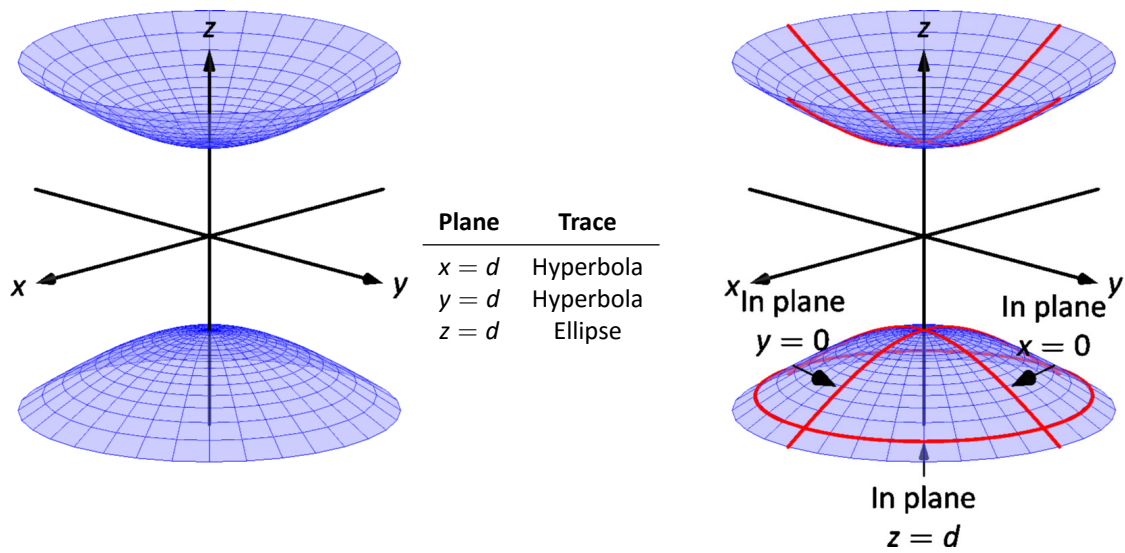
Plane	Trace
$x = d$	Hyperbola
$y = d$	Hyperbola
$z = d$	Ellipse

The one variable with a positive coefficient corresponds to the axis along which the hyperboloid "opens". In the case illustrated, when $|d| < |c|$, there is no trace in the plane $z = d$.

Hyperbolic Paraboloid, $z = \dfrac{x^2}{a^2} - \dfrac{y^2}{b^2}$
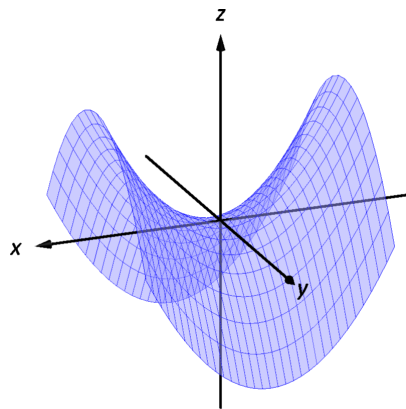
Plane	Trace
$x = d$	Parabola
$y = d$	Parabola
$z = d$	Hyperbola
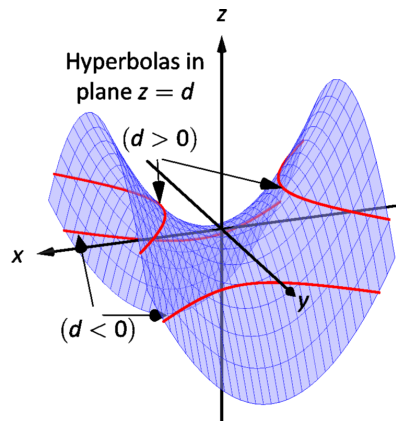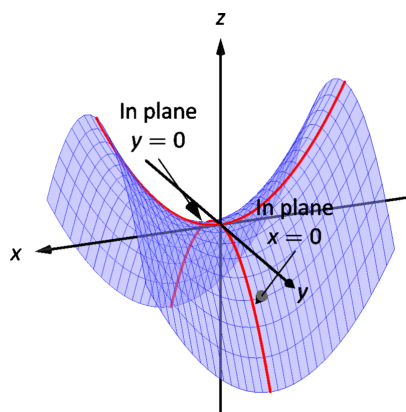
The parabolic traces will open along the axis of the one variable that is raised to the first power.

Watch the video:
Introduction to Quadric Surfaces at
`https://youtu.be/x6c2DdOrkQI`

Example 8   **Sketching quadric surfaces**

Sketch the quadric surface defined by the given equation.

1. $y = \dfrac{x^2}{4} + \dfrac{z^2}{16}$ 2. $x^2 + \dfrac{y^2}{9} + \dfrac{z^2}{4} = 1$ 3. $z = y^2 - x^2$.

SOLUTION

1. $y = \dfrac{x^2}{4} + \dfrac{z^2}{16}$:

   We first identify the quadric by pattern–matching with the equations given previously. Only two surfaces have equations where one variable is raised to the first power, the elliptic paraboloid and the hyperbolic paraboloid. In the latter case, the other variables have different signs, so we conclude that this describes a hyperbolic paraboloid. As the variable with the first power is y, we note the paraboloid opens along the y-axis.

   To make a decent sketch by hand, we need only draw a few traces. In this case, the traces $x = 0$ and $z = 0$ form parabolas that outline the shape.

   $x = 0$: The trace is the parabola $y = z^2/16$

   $z = 0$: The trace is the parabola $y = x^2/4$.

   Graphing each trace in the respective plane creates a sketch as shown in Figure 11.15(a). This is enough to give an idea of what the paraboloid looks like. The surface is filled in in (b).

2. $x^2 + \dfrac{y^2}{9} + \dfrac{z^2}{4} = 1$:

   This is an ellipsoid. We can get a good idea of its shape by drawing the traces in the coordinate planes.

   $x = 0$: The trace is the ellipse $\dfrac{y^2}{9} + \dfrac{z^2}{4} = 1$. The major axis is along the y–axis with length 6 (as $b = 3$, the length of the axis is 6); the minor axis is along the z-axis with length 4.

   $y = 0$: The trace is the ellipse $x^2 + \dfrac{z^2}{4} = 1$. The major axis is along the z-axis, and the minor axis has length 2 along the x-axis.

Notes:

(a)

(b)
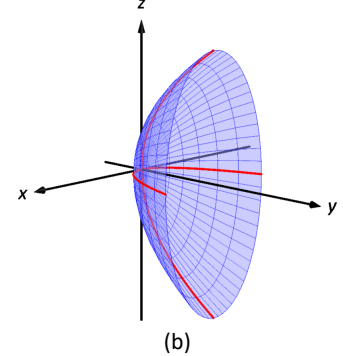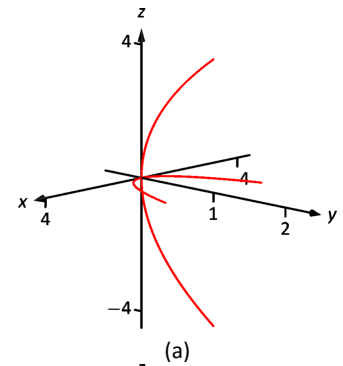
Figure 11.15: Sketching an elliptic paraboloid.

(a)

(b)
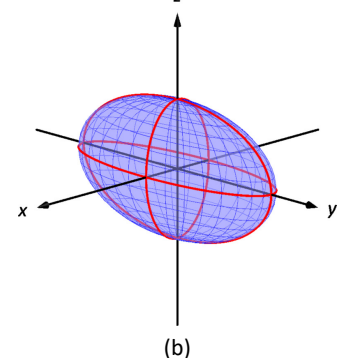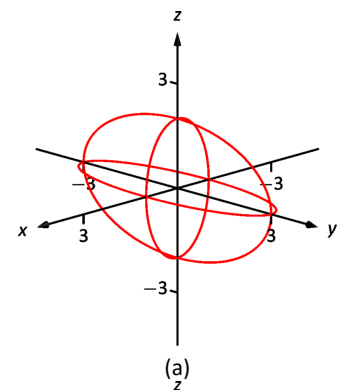
Figure 11.16: Sketching an ellipsoid.

645

(a)

(b)
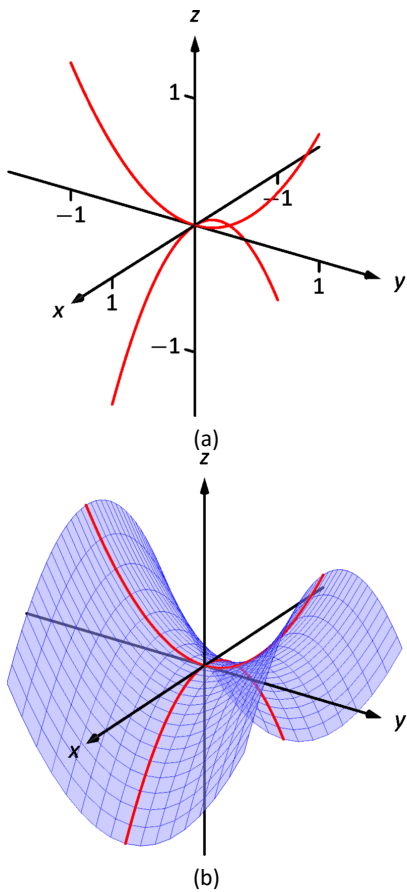
Figure 11.17: Sketching a hyperbolic paraboloid.
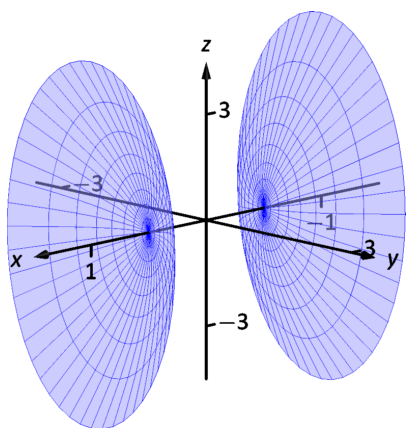
Figure 11.18: A possible equation of this quadric surface is found in Example 9.

$z = 0$: The trace is the ellipse $x^2 + \dfrac{y^2}{9} = 1$, with major axis along the y-axis.

Graphing each trace in the respective plane creates a sketch as shown in Figure 11.16(a). Filling in the surface gives Figure 11.16(b).

3. $z = y^2 - x^2$:
This defines a hyperbolic paraboloid, very similar to the one shown in the gallery of quadric sections. Consider the traces in the $y-z$ and $x-z$ planes:

$x = 0$: The trace is $z = y^2$, a parabola opening up in the $y - z$ plane.

$y = 0$: The trace is $z = -x^2$, a parabola opening down in the $x - z$ plane.

Sketching these two parabolas gives a sketch like that in Figure 11.17 (a), and filling in the surface gives a sketch like (b).

Example 9   **Identifying quadric surfaces**
Consider the quadric surface shown in Figure 11.18. Which of the following equations best fits this surface?

(a) $x^2 - y^2 - \dfrac{z^2}{9} = 0$ (c) $z^2 - x^2 - y^2 = 1$

(b) $x^2 - y^2 - z^2 = 1$ (d) $4x^2 - y^2 - \dfrac{z^2}{9} = 1$

SOLUTION   The image clearly displays a hyperboloid of two sheets. The gallery informs us that the equation will have a form similar to $\frac{z^2}{c^2} - \frac{x^2}{a^2} - \frac{y^2}{b^2} = 1$.
We can immediately eliminate option (a), as the constant in that equation is not 1.

The hyperboloid "opens" along the x-axis, meaning x must be the only variable with a positive coefficient, eliminating (c).

The hyperboloid is wider in the z-direction than in the y-direction, so we need an equation where $c > b$. This eliminates (b), leaving us with (d). We should verify that the equation given in (d), $4x^2 - y^2 - \frac{z^2}{9} = 1$, fits.

We already established that this equation describes a hyperboloid of two sheets that opens in the x-direction and is wider in the z-direction than in the y. Now note the coefficient of the x-term. Rewriting $4x^2$ in standard form, we have: $4x^2 = \dfrac{x^2}{(1/2)^2}$. Thus when $y = 0$ and $z = 0$, x must be $1/2$; i.e., each hyperboloid "starts" at $x = 1/2$. This matches our figure.

We conclude that $4x^2 - y^2 - \dfrac{z^2}{9} = 1$ best fits the graph.

Notes:

This section has introduced points in space and shown how equations can describe surfaces. The next sections explore *vectors*, an important mathematical object that we'll use to explore curves in space.

Notes:

Exercises 11.1

Terms and Concepts

1. Axes drawn in space must conform to the _____ _____ rule.

2. In the plane, the equation $x = 2$ defines a _____; in space, $x = 2$ defines a _____.

3. In the plane, the equation $y = x^2$ defines a _____; in space, $y = x^2$ defines a _____.

4. Which quadric surface looks like a Pringles® chip?

5. Consider the hyperbola $x^2 - y^2 = 1$ in the plane. If this hyperbola is rotated about the x-axis, what quadric surface is formed?

6. Consider the hyperbola $x^2 - y^2 = 1$ in the plane. If this hyperbola is rotated about the y-axis, what quadric surface is formed?

Problems

7. The points $A = (1, 4, 2)$, $B = (2, 6, 3)$ and $C = (4, 3, 1)$ form a triangle in space. Find the distances between each pair of points and determine if the triangle is a right triangle.

8. The points $A = (1, 1, 3)$, $B = (3, 2, 7)$, $C = (2, 0, 8)$ and $D = (0, -1, 4)$ form a quadrilateral $ABCD$ in space. Is this a parallelogram?

9. Find the center and radius of the sphere defined by $x^2 - 8x + y^2 + 2y + z^2 + 8 = 0$.

10. Find the center and radius of the sphere defined by $x^2 + y^2 + z^2 + 4x - 2y - 4z + 4 = 0$.

In Exercises 11–14, describe the region in space defined by the inequalities.

11. $x^2 + y^2 + z^2 < 1$

12. $0 \le x \le 3$

13. $x \ge 0$, $y \ge 0$, $z \ge 0$

14. $y \ge 3$

In Exercises 15–18, sketch the cylinder in space.

15. $z = x^3$

16. $y = \cos z$

17. $\dfrac{x^2}{4} + \dfrac{y^2}{9} = 1$

18. $y = \dfrac{1}{x}$

In Exercises 19–22, give the equation of the surface of revolution described.

19. Revolve $z = \dfrac{1}{1 + y^2}$ about the y-axis.

20. Revolve $y = x^2$ about the x-axis.

21. Revolve $z = x^2$ about the z-axis.

22. Revolve $z = 1/x$ about the z-axis.

In Exercises 23–26, a quadric surface is sketched. Determine which of the given equations best fits the graph.

23.

(a) $x = y^2 + \dfrac{z^2}{9}$ (b) $x = y^2 + \dfrac{z^2}{3}$

24.

(a) $x^2 - y^2 - z^2 = 0$ (b) $x^2 - y^2 + z^2 = 0$

25.

(a) $x^2 + \dfrac{y^2}{3} + \dfrac{z^2}{2} = 1$ (b) $x^2 + \dfrac{y^2}{9} + \dfrac{z^2}{4} = 1$

26.

(a) $y^2 - x^2 - z^2 = 1$ (b) $y^2 + x^2 - z^2 = 1$
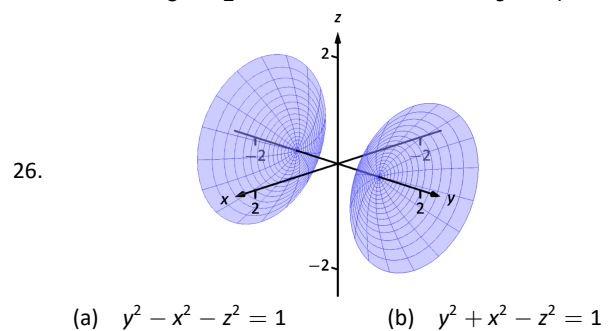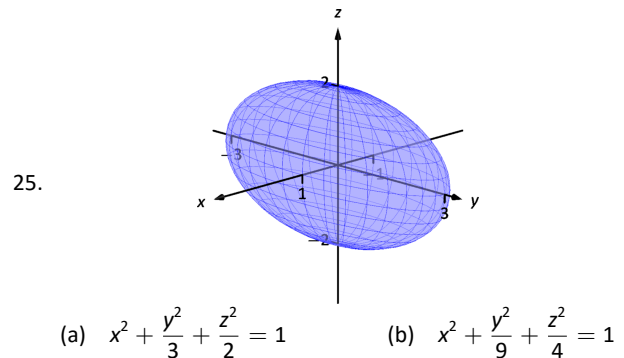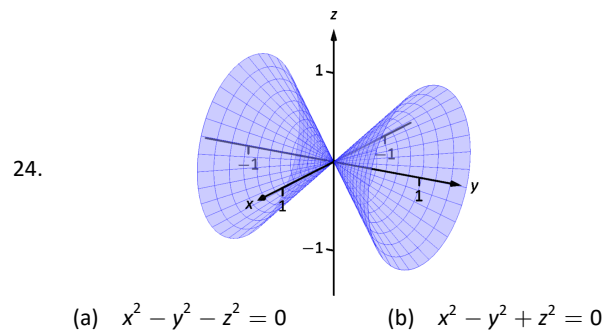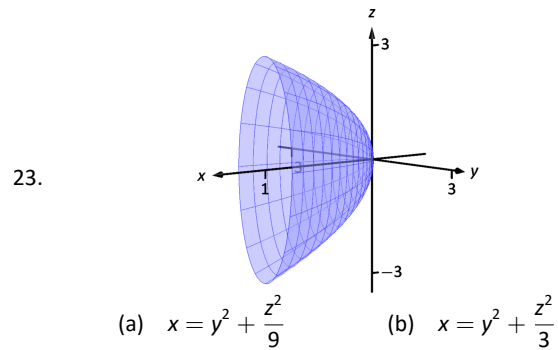
In Exercises 27–32, sketch the quadric surface.

27. $z - y^2 + x^2 = 0$

28. $z^2 = x^2 + \dfrac{y^2}{4}$

29. $x = -y^2 - z^2$

30. $16x^2 - 16y^2 - 16z^2 = 1$

31. $\dfrac{x^2}{9} - y^2 + \dfrac{z^2}{25} = 1$

32. $4x^2 + 2y^2 + z^2 = 4$

11.2   An Introduction to Vectors

Many quantities we think about daily can be described by a single number: temperature, speed, cost, weight and height. There are also many other concepts we encounter daily that cannot be described with just one number. For instance, a weather forecaster often describes wind with its speed and its direction ("...with winds from the southeast gusting up to 30 mph ..."). When applying a force, we are concerned with both the magnitude and direction of that force. In both of these examples, *direction* is important. Because of this, we study *vectors*, mathematical objects that convey both magnitude and direction information.

One "bare–bones" definition of a vector is based on what we wrote above: "a vector is a mathematical object with magnitude and direction parameters." This definition leaves much to be desired, as it gives no indication as to how such an object is to be used. Several other definitions exist; we choose here a definition rooted in a geometric visualization of vectors. It is very simplistic but readily permits further investigation.

Figure 11.19: Drawing the same vector with different initial points.

> **Definition 54   Vector**
> A **vector** is a directed line segment.
>
> Given points P and Q (either in the plane or in space), we denote with $\overrightarrow{PQ}$ the vector *from P to Q*. The point P is said to be the **initial point** of the vector, and the point Q is the **terminal point**.
>
> The **magnitude**, **length** or **norm** of $\overrightarrow{PQ}$ is the length of the line segment $\overline{PQ}$: $\left\|\overrightarrow{PQ}\right\| = \left\|\overline{PQ}\right\|$.
>
> Two vectors are **equal** if they have the same magnitude and direction.

Figure 11.19 shows multiple instances of the same vector. Each directed line segment has the same direction and length (magnitude), hence each is the same vector.

We use $\mathbb{R}^2$ (pronounced "r two") to represent all the vectors in the plane, and use $\mathbb{R}^3$ (pronounced "r three") to represent all the vectors in space.

Consider the vectors $\overrightarrow{PQ}$ and $\overrightarrow{RS}$ as shown in Figure 11.20. The vectors look to be equal; that is, they seem to have the same length and direction. Indeed, they are. Both vectors move 2 units to the right and 1 unit up from the initial point
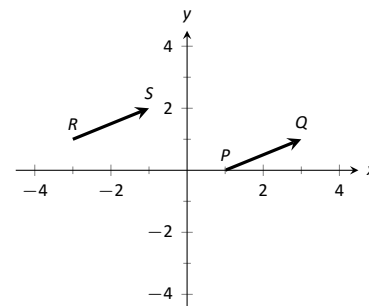
Figure 11.20: Illustrating how equal vectors have the same displacement.

Notes:

to reach the terminal point. One can analyze this movement to measure the magnitude of the vector, and the movement itself gives direction information (one could also measure the slope of the line passing through P and Q or R and S). Since they have the same length and direction, these two vectors are equal.

This demonstrates that inherently all we care about is *displacement*; that is, how far in the x, y and possibly z directions the terminal point is from the initial point. Both the vectors $\overrightarrow{PQ}$ and $\overrightarrow{RS}$ in Figure 11.20 have an x-displacement of 2 and a y-displacement of 1. This suggests a standard way of describing vectors in the plane. A vector whose x-displacement is a and whose y-displacement is b will have terminal point (a, b) when the initial point is the origin, $(0, 0)$. This leads us to a definition of a standard and concise way of referring to vectors.

Definition 55   Component Form of a Vector

1. The **component form** of a vector $\vec{v}$ in $\mathbb{R}^2$, whose terminal point is (a, b) when its initial point is $(0, 0)$, is $\langle a, b \rangle$.

2. The **component form** of a vector $\vec{v}$ in $\mathbb{R}^3$, whose terminal point is (a, b, c) when its initial point is $(0, 0, 0)$, is $\langle a, b, c \rangle$.

The numbers a, b (and c, respectively) are the **components** of $\vec{v}$.

Note: Instead of $\vec{v}$, some texts use boldface: **v**. The advantage of **v** is that it tends to be easier to read. The advantage of $\vec{v}$ is that it's easier to write.

It follows from the definition that the component form of the vector $\overrightarrow{PQ}$, where $P = (x_1, y_1)$ and $Q = (x_2, y_2)$ is

$$\overrightarrow{PQ} = \langle x_2 - x_1, y_2 - y_1 \rangle ;$$

in space, where $P = (x_1, y_1, z_1)$ and $Q = (x_2, y_2, z_2)$, the component form of $\overrightarrow{PQ}$ is

$$\overrightarrow{PQ} = \langle x_2 - x_1, y_2 - y_1, z_2 - z_1 \rangle .$$

Watch the video:
An Introduction to Vectors, Part 1 at
`https://youtu.be/60btq9PN8IM`

We practice using this notation in the following example.

Notes:

Example 1   **Using component form notation for vectors**

1. Sketch the vector $\vec{v} = \langle 2, -1 \rangle$ starting at $P = (3, 2)$ and find its magnitude.

2. Find the component form of the vector $\vec{w}$ whose initial point is $R = (-3, -2)$ and whose terminal point is $S = (-1, 2)$.

3. Sketch the vector $\vec{u} = \langle 2, -1, 3 \rangle$ starting at the point $Q = (1, 1, 1)$ and find its magnitude.

SOLUTION

1. Using P as the initial point, we move 2 units in the positive x-direction and -1 units in the positive y-direction to arrive at the terminal point $P' = (5, 1)$, as drawn in Figure 11.21(a).

   The magnitude of $\vec{v}$ is determined directly from the component form:

   $$\|\vec{v}\| = \sqrt{2^2 + (-1)^2} = \sqrt{5}.$$

2. Using the paragraph following Definition 55, we have

   $$\overrightarrow{RS} = \langle -1 - (-3), 2 - (-2) \rangle = \langle 2, 4 \rangle.$$

   One can readily see from Figure 11.21(a) that the x- and y-displacement of $\overrightarrow{RS}$ is 2 and 4, respectively, as the component form suggests.

3. Using Q as the initial point, we move 2 units in the positive x-direction, -1 unit in the positive y-direction, and 3 units in the positive z-direction to arrive at the terminal point $Q' = (3, 0, 4)$, illustrated in Figure 11.21(b).

   The magnitude of $\vec{u}$ is:

   $$\|\vec{u}\| = \sqrt{2^2 + (-1)^2 + 3^2} = \sqrt{14}.$$

Now that we have defined vectors, and have created a nice notation by which to describe them, we start considering how vectors interact with each other. That is, we define an *algebra* on vectors.

(a)

(b)
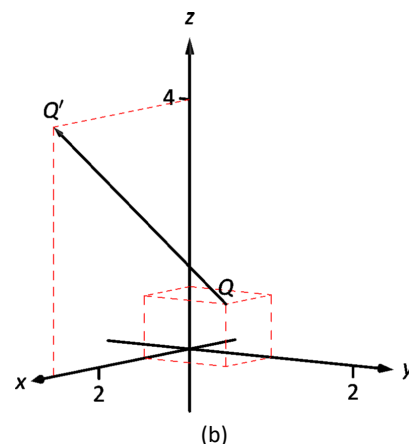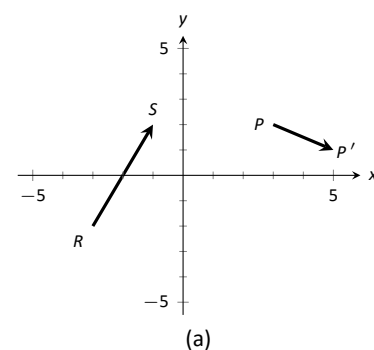
Figure 11.21: Graphing vectors in Example 1.

Notes:

Definition 56   Vector Algebra

1. Let $\vec{u} = \langle u_1, u_2 \rangle$ and $\vec{v} = \langle v_1, v_2 \rangle$ be vectors in $\mathbb{R}^2$, and let c be a scalar.

    (a) The addition, or sum, of the vectors $\vec{u}$ and $\vec{v}$ is the vector

    $$\vec{u} + \vec{v} = \langle u_1 + v_1, u_2 + v_2 \rangle .$$

    (b) The scalar product of c and $\vec{v}$ is the vector

    $$c\vec{v} = c \langle v_1, v_2 \rangle = \langle cv_1, cv_2 \rangle .$$

2. Let $\vec{u} = \langle u_1, u_2, u_3 \rangle$ and $\vec{v} = \langle v_1, v_2, v_3 \rangle$ be vectors in $\mathbb{R}^3$, and let c be a scalar.

    (a) The addition, or sum, of the vectors $\vec{u}$ and $\vec{v}$ is the vector

    $$\vec{u} + \vec{v} = \langle u_1 + v_1, u_2 + v_2, u_3 + v_3 \rangle .$$

    (b) The scalar product of c and $\vec{v}$ is the vector

    $$c\vec{v} = c \langle v_1, v_2, v_3 \rangle = \langle cv_1, cv_2, cv_3 \rangle .$$

In short, we say addition and scalar multiplication are computed "component–wise."

Example 2   Adding vectors
Sketch the vectors $\vec{u} = \langle 1, 3 \rangle$, $\vec{v} = \langle 2, 1 \rangle$ and $\vec{u} + \vec{v}$ all with initial point at the origin.

SOLUTION   We first compute $\vec{u} + \vec{v}$.

$$\vec{u} + \vec{v} = \langle 1, 3 \rangle + \langle 2, 1 \rangle$$
$$= \langle 3, 4 \rangle .$$

These are all sketched in Figure 11.22.

As vectors convey magnitude and direction information, the sum of vectors also convey length and magnitude information. Adding $\vec{u} + \vec{v}$ suggests the following idea:

"Starting at an initial point, go out $\vec{u}$, then go out $\vec{v}$."
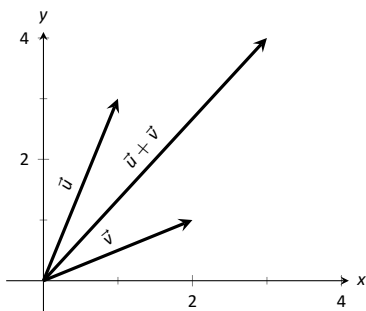
Figure 11.22: Graphing the sum of vectors in Example 2.

Notes:

This idea is sketched in Figure 11.23, where the initial point of $\vec{v}$ is the terminal point of $\vec{u}$. This is known as the "Head to Tail Rule" of adding vectors. Vector addition is very important. For instance, if the vectors $\vec{u}$ and $\vec{v}$ represent forces acting on a body, the sum $\vec{u} + \vec{v}$ gives the resulting force. Because of various physical applications of vector addition, the sum $\vec{u} + \vec{v}$ is often referred to as the **resultant vector**, or just the "resultant."

Analytically, it is easy to see that $\vec{u} + \vec{v} = \vec{v} + \vec{u}$. Figure 11.23 also gives a graphical representation of this, using gray vectors. Note that the vectors $\vec{u}$ and $\vec{v}$, when arranged as in the figure, form a parallelogram. Because of this, the Head to Tail Rule is also known as the Parallelogram Law: the vector $\vec{u} + \vec{v}$ is defined by forming the parallelogram defined by the vectors $\vec{u}$ and $\vec{v}$; the initial point of $\vec{u} + \vec{v}$ is the common initial point of parallelogram, and the terminal point of the sum is the common terminal point of the parallelogram.

While not illustrated here, the Head to Tail Rule and Parallelogram Law hold for vectors in $\mathbb{R}^3$ as well.

It follows from the properties of the real numbers and Definition 56 that

$$\vec{u} - \vec{v} = \vec{u} + (-1)\vec{v}.$$

The Parallelogram Law gives us a good way to visualize this subtraction. We demonstrate this in the following example.

Figure 11.23: Illustrating how to add vectors using the Head to Tail Rule and Parallelogram Law.

Example 3 Vector Subtraction
Let $\vec{u} = \langle 3, 1 \rangle$ and $\vec{v} = \langle 1, 2 \rangle$. Compute and sketch $\vec{u} - \vec{v}$.

SOLUTION The computation of $\vec{u} - \vec{v}$ is straightforward, and we show all steps below. Usually the formal step of multiplying by (-1) is omitted and we "just subtract."

$$\begin{aligned}\vec{u} - \vec{v} &= \vec{u} + (-1)\vec{v} \\ &= \langle 3, 1 \rangle + \langle -1, -2 \rangle \\ &= \langle 2, -1 \rangle.\end{aligned}$$

Figure 11.24 illustrates, using the Head to Tail Rule, how the subtraction can be viewed as the sum $\vec{u}+(-\vec{v})$. The figure also illustrates how $\vec{u}-\vec{v}$ can be obtained by looking only at the terminal points of $\vec{u}$ and $\vec{v}$ (when their initial points are the same).

Example 4 Scaling vectors

1. Sketch the vectors $\vec{v} = \langle 2, 1 \rangle$ and $2\vec{v}$ with initial point at the origin.

2. Compute the magnitudes of $\vec{v}$ and $2\vec{v}$.

Figure 11.24: Illustrating how to subtract vectors graphically.

Notes:
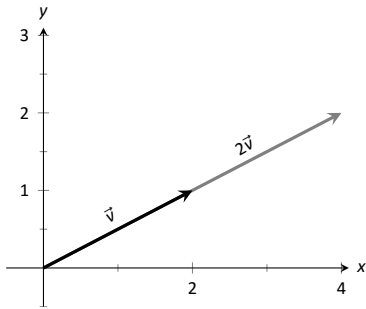
Figure 11.25: Graphing vectors $\vec{v}$ and $2\vec{v}$ in Example 4.

SOLUTION

1. We compute $2\vec{v}$:

$$2\vec{v} = 2\langle 2, 1 \rangle$$
$$= \langle 4, 2 \rangle.$$

Both $\vec{v}$ and $2\vec{v}$ are sketched in Figure 11.25. Make note that $2\vec{v}$ does not start at the terminal point of $\vec{v}$; rather, its initial point is also the origin.

2. The figure suggests that $2\vec{v}$ is twice as long as $\vec{v}$. We compute their magnitudes to confirm this.

$$\|\vec{v}\| = \sqrt{2^2 + 1^2}$$
$$= \sqrt{5}.$$
$$\|2\vec{v}\| = \sqrt{4^2 + 2^2}$$
$$= \sqrt{20}$$
$$= \sqrt{4 \cdot 5} = 2\sqrt{5}.$$

As we suspected, $2\vec{v}$ is twice as long as $\vec{v}$.

The **zero vector** is the vector whose initial point is also its terminal point. It is denoted by $\vec{0}$. Its component form, in $\mathbb{R}^2$, is $\langle 0, 0 \rangle$; in $\mathbb{R}^3$, it is $\langle 0, 0, 0 \rangle$. Usually the context makes is clear whether $\vec{0}$ is referring to a vector in the plane or in space.

Our examples have illustrated key principles in vector algebra: how to add and subtract vectors and how to multiply vectors by a scalar. The following theorem states formally the properties of these operations.

Notes:

Theorem 86   Properties of Vector Operations

The following are true for all scalars c and d, and for all vectors $\vec{u}$, $\vec{v}$ and $\vec{w}$, where $\vec{u}$, $\vec{v}$ and $\vec{w}$ are all in $\mathbb{R}^2$ or where $\vec{u}$, $\vec{v}$ and $\vec{w}$ are all in $\mathbb{R}^3$:

1. $\vec{u} + \vec{v} = \vec{v} + \vec{u}$ Commutative Property

2. $(\vec{u} + \vec{v}) + \vec{w} = \vec{u} + (\vec{v} + \vec{w})$ Associative Property

3. $\vec{v} + \vec{0} = \vec{v}$ Additive Identity

4. $(cd)\vec{v} = c(d\vec{v})$

5. $c(\vec{u} + \vec{v}) = c\vec{u} + c\vec{v}$ Distributive Property

6. $(c + d)\vec{v} = c\vec{v} + d\vec{v}$ Distributive Property

7. $0\vec{v} = \vec{0}$

8. $\|c\vec{v}\| = |c| \cdot \|\vec{v}\|$

9. $\|\vec{u}\| = 0$ if, and only if, $\vec{u} = \vec{0}$.

As stated before, each vector $\vec{v}$ conveys magnitude and direction information. We have a method of extracting the magnitude, which we write as $\|\vec{v}\|$. *Unit vectors* are a way of extracting just the direction information from a vector.

Definition 57   Unit Vector

A **unit vector** is a vector $\vec{v}$ with a magnitude of 1; that is,

$$\|\vec{v}\| = 1.$$

Consider this scenario: you are given a vector $\vec{v}$ and are told to create a vector of length 10 in the direction of $\vec{v}$. How does one do that? If we knew that $\vec{u}$ was the unit vector in the direction of $\vec{v}$, the answer would be easy: $10\vec{u}$. So how do we find $\vec{u}$?

Property 8 of Theorem 86 holds the key. If we divide $\vec{v}$ by its magnitude, it becomes a vector of length 1. Consider:

$$\left\| \frac{1}{\|\vec{v}\|} \vec{v} \right\| = \frac{1}{\|\vec{v}\|} \|\vec{v}\| \qquad \text{(we can pull out } \frac{1}{\|\vec{v}\|} \text{ as it is a scalar)}$$
$$= 1.$$

Notes:

So the vector of length 10 in the direction of $\vec{v}$ is $10\dfrac{1}{\|\vec{v}\|}\vec{v}$. An example will make this more clear.

Example 5　　　**Using Unit Vectors**

Let $\vec{v} = \langle 3, 1 \rangle$ and let $\vec{w} = \langle 1, 2, 2 \rangle$.

1. Find the unit vector in the direction of $\vec{v}$.

2. Find the unit vector in the direction of $\vec{w}$.

3. Find the vector in the direction of $\vec{v}$ with magnitude 5.

SOLUTION

1. We find $\|\vec{v}\| = \sqrt{10}$. So the unit vector $\vec{u}$ in the direction of $\vec{v}$ is

$$\vec{u} = \frac{1}{\sqrt{10}}\vec{v} = \left\langle \frac{3}{\sqrt{10}}, \frac{1}{\sqrt{10}} \right\rangle.$$

2. We find $\|\vec{w}\| = 3$, so the unit vector $\vec{z}$ in the direction of $\vec{w}$ is

$$\vec{u} = \frac{1}{3}\vec{w} = \left\langle \frac{1}{3}, \frac{2}{3}, \frac{2}{3} \right\rangle.$$

3. To create a vector with magnitude 5 in the direction of $\vec{v}$, we multiply the unit vector $\vec{u}$ by 5. Thus $5\vec{u} = \langle 15/\sqrt{10}, 5/\sqrt{10} \rangle$ is the vector we seek. This is sketched in Figure 11.26.

The basic formation of the unit vector $\vec{u}$ in the direction of a vector $\vec{v}$ leads to a interesting equation. It is:

$$\vec{v} = \|\vec{v}\|\,\frac{1}{\|\vec{v}\|}\vec{v}.$$

We rewrite the equation with parentheses to make a point:

$$\vec{v} = \underbrace{\|\vec{v}\|}_{\text{magnitude}} \cdot \underbrace{\left(\frac{1}{\|\vec{v}\|}\vec{v}\right)}_{\text{direction}}.$$

This equation illustrates the fact that a vector has both magnitude and direction, where we view a unit vector as supplying *only* direction information. Identifying unit vectors with direction allows us to define **parallel vectors**.
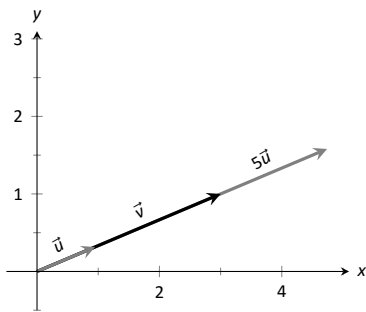
Figure 11.26: Graphing vectors in Example 5. All vectors shown have their initial point at the origin.

Notes:

Definition 58   Parallel Vectors

1. Unit vectors $\vec{u}_1$ and $\vec{u}_2$ are **parallel** if $\vec{u}_1 = \pm\vec{u}_2$.

2. Nonzero vectors $\vec{v}_1$ and $\vec{v}_2$ are **parallel** if their respective unit vectors are parallel.

It is equivalent to say that vectors $\vec{v}_1$ and $\vec{v}_2$ are parallel if there is a scalar $c \neq 0$ such that $\vec{v}_1 = c\vec{v}_2$ (see marginal note).

If one graphed all unit vectors in $\mathbb{R}^2$ with the initial point at the origin, then the terminal points would all lie on the unit circle. Based on what we know from trigonometry, we can then say that the component form of all unit vectors in $\mathbb{R}^2$ is $\langle \cos\theta, \sin\theta \rangle$ for some angle θ.

A similar construction in $\mathbb{R}^3$ shows that the terminal points all lie on the unit sphere. These vectors also have a particular component form, but its derivation is not as straightforward as the one for unit vectors in $\mathbb{R}^2$. Important concepts about unit vectors are given in Key Idea 51 below.

Note: $\vec{0}$ is directionless; because $\left\|\vec{0}\right\| = 0$, there is no unit vector in the "direction" of $\vec{0}$.

Some texts define two vectors as being parallel if one is a scalar multiple of the other. By this definition, $\vec{0}$ is parallel to all vectors as $\vec{0} = 0\vec{v}$ for all $\vec{v}$.

We prefer the given definition of parallel as it is grounded in the fact that unit vectors provide direction information. One may adopt the convention that $\vec{0}$ is parallel to all vectors if they desire. (See also the marginal note on page 680.)

Key Idea 51   Unit Vectors

1. The unit vector in the direction of $\vec{v}$ is

$$\vec{u} = \frac{1}{\|\vec{v}\|}\vec{v}.$$

2. A vector $\vec{u}$ in $\mathbb{R}^2$ is a unit vector if, and only if, its component form is $\langle \cos\theta, \sin\theta \rangle$ for some angle θ.

3. A vector $\vec{u}$ in $\mathbb{R}^3$ is a unit vector if, and only if, its component form is $\langle \sin\theta\cos\varphi, \sin\theta\sin\varphi, \cos\theta \rangle$ for some angles θ and φ.

These formulas can come in handy in a variety of situations, especially the formula for unit vectors in the plane.

Example 6   Finding Component Forces
Consider a weight of 50lb hanging from two chains, as shown in Figure 11.27. One chain makes an angle of $30°$ with the vertical, and the other an angle of $45°$. Find the force applied to each chain.
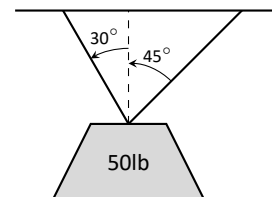
Figure 11.27: A diagram of a weight hanging from 2 chains in Example 6.

Notes:
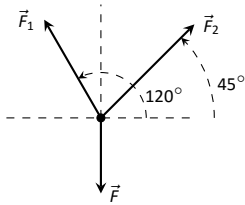
SOLUTION   Knowing that gravity is pulling the 50lb weight straight down, we can create a vector $\vec{F}$ to represent this force.

$$\vec{F} = 50 \langle 0, -1 \rangle = \langle 0, -50 \rangle.$$

We can view each chain as "pulling" the weight up, preventing it from falling. We can represent the force from each chain with a vector. Let $\vec{F}_1$ represent the force from the chain making an angle of 30° with the vertical, and let $\vec{F}_2$ represent the force form the other chain. Convert all angles to be measured from the horizontal (as shown in Figure 11.28), and apply Key Idea 51. As we do not yet know the magnitudes of these vectors, (that is the problem at hand), we use m_1 and m_2 to represent them.

$$\vec{F}_1 = m_1 \langle \cos 120°, \sin 120° \rangle$$
$$\vec{F}_2 = m_2 \langle \cos 45°, \sin 45° \rangle$$

As the weight is not moving, we know the sum of the forces is $\vec{0}$. This gives:

$$\vec{F} + \vec{F}_1 + \vec{F}_2 = \vec{0}$$
$$\langle 0, -50 \rangle + m_1 \langle \cos 120°, \sin 120° \rangle + m_2 \langle \cos 45°, \sin 45° \rangle = \vec{0}$$

The sum of the entries in the first component is 0, and the sum of the entries in the second component is also 0. This leads us to the following two equations:

$$m_1 \cos 120° + m_2 \cos 45° = 0$$
$$m_1 \sin 120° + m_2 \sin 45° = 50$$

This is a simple 2-equation, 2-unkown system of linear equations. We leave it to the reader to verify that the solution is

$$m_1 = 50(\sqrt{3} - 1)\text{lb}; \qquad m_2 = \frac{50\sqrt{2}}{1 + \sqrt{3}}\text{lb}.$$

It might seem odd that the sum of the forces applied to the chains is more than 50lb. We leave it to a physics class to discuss the full details, but offer this short explanation. Our equations were established so that the *vertical* components of each force sums to 50lb, thus supporting the weight. Since the chains are at an angle, they also pull against each other, creating an "additional" horizontal force while holding the weight in place.

Unit vectors were very important in the previous calculation; they allowed us to define a vector in the proper direction but with an unknown magnitude. Our computations were then computed component–wise. Because such calculations are often necessary, the *standard unit vectors* can be useful.

Figure 11.28: A diagram of the force vectors from Example 6.

Notes:

Definition 59   Standard Unit Vectors

1. In $\mathbb{R}^2$, the standard unit vectors are

$$\vec{\imath} = \langle 1, 0 \rangle \quad \text{and} \quad \vec{\jmath} = \langle 0, 1 \rangle\,.$$

2. In $\mathbb{R}^3$, the standard unit vectors are

$$\vec{\imath} = \langle 1, 0, 0 \rangle \quad \text{and} \quad \vec{\jmath} = \langle 0, 1, 0 \rangle \quad \text{and} \quad \vec{k} = \langle 0, 0, 1 \rangle\,.$$

Example 7   Using standard unit vectors

1. Rewrite $\vec{v} = \langle 2, -3 \rangle$ using the standard unit vectors.

2. Rewrite $\vec{w} = 4\vec{\imath} - 5\vec{\jmath} + 2\vec{k}$ in component form.

SOLUTION

1.
$$\begin{aligned}
\vec{v} &= \langle 2, -3 \rangle \\
&= \langle 2, 0 \rangle + \langle 0, -3 \rangle \\
&= 2\langle 1, 0 \rangle - 3\langle 0, 1 \rangle \\
&= 2\vec{\imath} - 3\vec{\jmath}
\end{aligned}$$

2.
$$\begin{aligned}
\vec{w} &= 4\vec{\imath} - 5\vec{\jmath} + 2\vec{k} \\
&= \langle 4, 0, 0 \rangle + \langle 0, -5, 0 \rangle + \langle 0, 0, 2 \rangle \\
&= \langle 4, -5, 2 \rangle
\end{aligned}$$

These two examples demonstrate that converting between component form and the standard unit vectors is rather straightforward. Many mathematicians prefer component form, and it is the preferred notation in this text. Many engineers prefer using the standard unit vectors, and many engineering texts use that notation.

Example 8   Finding Component Force
A weight of 25lb is suspended from a chain of length 2ft while a wind pushes the weight to the right with constant force of 5lb as shown in Figure 11.29. What angle will the chain make with the vertical as a result of the wind's pushing? How much higher will the weight be?
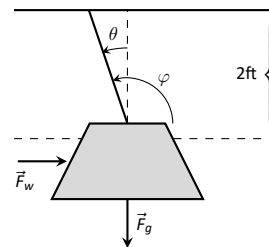
Figure 11.29: A figure of a weight being pushed by the wind in Example 8.

Notes:

SOLUTION   The force of the wind is represented by the vector $\vec{F}_w = 5\vec{\imath}$. The force of gravity on the weight is represented by $\vec{F}_g = -25\vec{\jmath}$. The direction and magnitude of the vector representing the force on the chain are both unknown. We represent this force with

$$\vec{F}_c = m\,\langle\cos\varphi, \sin\varphi\rangle = m\cos\varphi\,\vec{\imath} + m\sin\varphi\,\vec{\jmath}$$

for some magnitude m and some angle with the horizontal φ. (Note: θ is the angle the chain makes with the *vertical*; φ is the angle with the *horizontal*.)

As the weight is at equilibrium, the sum of the forces is $\vec{0}$:

$$\vec{F}_c + \vec{F}_w + \vec{F}_g = \vec{0}$$
$$m\cos\varphi\,\vec{\imath} + m\sin\varphi\,\vec{\jmath} + 5\vec{\imath} - 25\vec{\jmath} = \vec{0}$$

Thus the sum of the $\vec{\imath}$ and $\vec{\jmath}$ components are 0, leading us to the following system of equations:

$$5 + m\cos\varphi = 0 \qquad\qquad\qquad (11.1)$$
$$-25 + m\sin\varphi = 0 \qquad\qquad\qquad (11.2)$$

This is enough to determine $\vec{F}_c$ already, as we know $m\cos\varphi = -5$ and $m\sin\varphi = 25$. Thus $F_c = \langle -5, 25\rangle$. We can use this to find the magnitude m:

$$m = \sqrt{(-5)^2 + 25^2} = 5\sqrt{26}\text{lb}.$$

We can then use either equality from Equation (11.2) to solve for φ. We choose the first equality as using arccosine will return an angle in the 2nd quadrant:

$$5 + 5\sqrt{26}\cos\varphi = 0 \quad \Rightarrow \quad \varphi = \cos^{-1}\left(\frac{-5}{5\sqrt{26}}\right) \approx 1.7682 \approx 101.31°.$$

Subtracting $90°$ from this angle gives us an angle of $11.31°$ with the vertical.

We can now use trigonometry to find out how high the weight is lifted. Figure 11.29 shows that a right triangle is formed with the 2ft chain as the hypotenuse. We have found that the interior angle is $11.31°$. The length of the adjacent side (in the diagram, the dashed vertical line) is $2\cos 11.31° \approx 1.96$ft. Thus the weight is lifted by about 0.04ft, almost 1/2in.

The algebra we have applied to vectors is already demonstrating itself to be very useful. There are two more fundamental operations we can perform with vectors, the *dot product* and the *cross product*. The next two sections explore each in turn.

Notes:

Exercises 11.2

Terms and Concepts

1. Name two different things that cannot be described with just one number, but rather need 2 or more numbers to fully describe them.

2. What is the difference between $(1, 2)$ and $\langle 1, 2 \rangle$?

3. What is a unit vector?

4. What does it mean for two vectors to be parallel?

5. What effect does multiplying a vector by -2 have?

Problems

In Exercises 6–9, points P and Q are given. Write the vector $\overrightarrow{PQ}$ in component form and using the standard unit vectors.

6. $P = (2, -1)$, $Q = (3, 5)$

7. $P = (3, 2)$, $Q = (7, -2)$

8. $P = (0, 3, -1)$, $Q = (6, 2, 5)$

9. $P = (2, 1, 2)$, $Q = (4, 3, 2)$

10. Let $\vec{u} = \langle 1, -2 \rangle$ and $\vec{v} = \langle 1, 1 \rangle$.

    (a) Find $\vec{u} + \vec{v}, \vec{u} - \vec{v}, 2\vec{u} - 3\vec{v}$.

    (b) Sketch the above vectors on the same axes, along with $\vec{u}$ and $\vec{v}$.

    (c) Find $\vec{x}$ where $\vec{u} + \vec{x} = 2\vec{v} - \vec{x}$.

11. Let $\vec{u} = \langle 1, 1, -1 \rangle$ and $\vec{v} = \langle 2, 1, 2 \rangle$.

    (a) Find $\vec{u} + \vec{v}, \vec{u} - \vec{v}, \pi\vec{u} - \sqrt{2}\vec{v}$.

    (b) Sketch the above vectors on the same axes, along with $\vec{u}$ and $\vec{v}$.

    (c) Find $\vec{x}$ where $\vec{u} + \vec{x} = \vec{v} + 2\vec{x}$.

In Exercises 12–15, sketch $\vec{u}, \vec{v}, \vec{u} + \vec{v}$ and $\vec{u} - \vec{v}$ on the same axes.

12.

13.

14.

15.
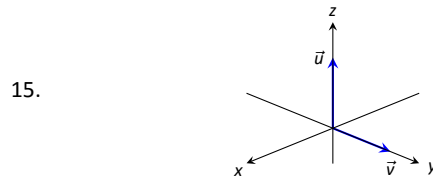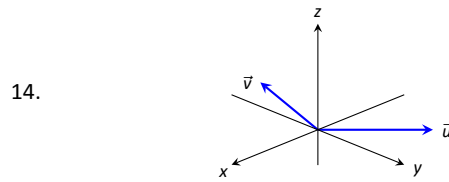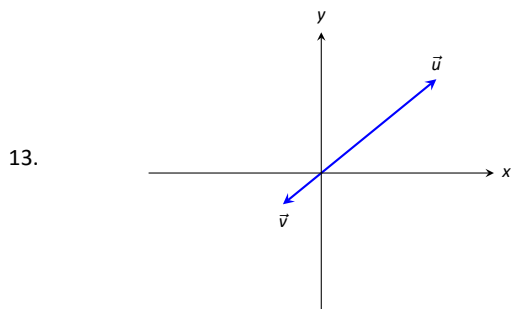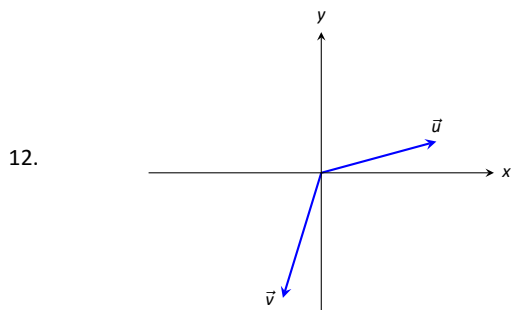
In Exercises 16–19, find $\|\vec{u}\|$, $\|\vec{v}\|$, $\|\vec{u} + \vec{v}\|$ and $\|\vec{u} - \vec{v}\|$.

16. $\vec{u} = \langle 2, 1 \rangle$, $\vec{v} = \langle 3, -2 \rangle$

17. $\vec{u} = \langle -3, 2, 2 \rangle$, $\vec{v} = \langle 1, -1, 1 \rangle$

18. $\vec{u} = \langle 1, 2 \rangle$, $\vec{v} = \langle -3, -6 \rangle$

19. $\vec{u} = \langle 2, -3, 6 \rangle$, $\vec{v} = \langle 10, -15, 30 \rangle$

20. Under what conditions is $\|\vec{u}\| + \|\vec{v}\| = \|\vec{u} + \vec{v}\|$?

In Exercises 21–24, find the unit vector $\vec{u}$ in the direction of $\vec{v}$.

21. $\vec{v} = \langle 3, 7 \rangle$

22. $\vec{v} = \langle 6, 8 \rangle$

23. $\vec{v} = \langle 1, -2, 2 \rangle$

24. $\vec{v} = \langle 2, -2, 2 \rangle$

25. Find the unit vector in the first quadrant of $\mathbb{R}^2$ that makes a $50°$ angle with the x-axis.

26. Find the unit vector in the second quadrant of $\mathbb{R}^2$ that makes a $30°$ angle with the y-axis.

27. Verify, from Key Idea 51, that $\vec{u} = \langle \sin\theta\cos\varphi, \sin\theta\sin\varphi, \cos\theta \rangle$ is a unit vector for all angles θ and φ.

A weight of 100lb is suspended from two chains, making angles with the vertical of θ and φ as shown in the figure below.
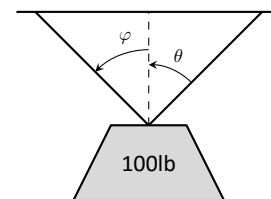
In Exercises 28–31, angles θ and φ are given. Find the force applied to each chain.

28. $\theta = 30°$, $\varphi = 30°$

29. $\theta = 60°$, $\varphi = 60°$

30. $\theta = 20°$, $\varphi = 15°$

31. $\theta = 0°$, $\varphi = 0°$

A weight of $p$lb is suspended from a chain of length ℓ while a constant force of $\vec{F}_w$ pushes the weight to the right, making an angle of θ with the vertical, as shown in the figure below.
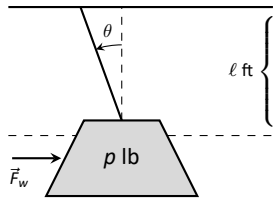
In Exercises 32–35, a force $\vec{F}_w$ and length ℓ are given. Find the

angle θ and the height the weight is lifted as it moves to the right.

32. $\vec{F}_w = 1\text{lb}, \quad \ell = 1\text{ft}, \quad p = 1\text{lb}$

33. $\vec{F}_w = 1\text{lb}, \quad \ell = 1\text{ft}, \quad p = 10\text{lb}$

34. $\vec{F}_w = 1\text{lb}, \quad \ell = 10\text{ft}, \quad p = 1\text{lb}$

35. $\vec{F}_w = 10\text{lb}, \quad \ell = 10\text{ft}, \quad p = 1\text{lb}$

11.3 The Dot Product

The previous section introduced vectors and described how to add them to-gether and how to multiply them by scalars. This section introduces *a* multi-plication on vectors called the **dot product**.

Definition 60   **Dot Product**

1. Let $\vec{u} = \langle u_1, u_2 \rangle$ and $\vec{v} = \langle v_1, v_2 \rangle$ in $\mathbb{R}^2$. The **dot product** of $\vec{u}$ and $\vec{v}$, denoted $\vec{u} \cdot \vec{v}$, is

$$\vec{u} \cdot \vec{v} = u_1 v_1 + u_2 v_2.$$

2. Let $\vec{u} = \langle u_1, u_2, u_3 \rangle$ and $\vec{v} = \langle v_1, v_2, v_3 \rangle$ in $\mathbb{R}^3$. The **dot product** of $\vec{u}$ and $\vec{v}$, denoted $\vec{u} \cdot \vec{v}$, is

$$\vec{u} \cdot \vec{v} = u_1 v_1 + u_2 v_2 + u_3 v_3.$$

Note how this product of vectors returns a *scalar*, not another vector.

We practice evaluating a dot product in the following example, then we will discuss why this product is useful.

Example 1   **Evaluating dot products**

1. Let $\vec{u} = \langle 1, 2 \rangle$, $\vec{v} = \langle 3, -1 \rangle$ in $\mathbb{R}^2$. Find $\vec{u} \cdot \vec{v}$.

2. Let $\vec{x} = \langle 2, -2, 5 \rangle$ and $\vec{y} = \langle -1, 0, 3 \rangle$ in $\mathbb{R}^3$. Find $\vec{x} \cdot \vec{y}$.

SOLUTION

1. Using Definition 60, we have

$$\vec{u} \cdot \vec{v} = 1(3) + 2(-1) = 1.$$

2. Using the definition, we have

$$\vec{x} \cdot \vec{y} = 2(-1) - 2(0) + 5(3) = 13.$$

The dot product, as shown by the preceding example, is very simple to eval-uate. It is only the sum of products. While the definition gives no hint as to why we would care about this operation, there is an amazing connection between the dot product and angles formed by the vectors. Before stating this connec-tion, we give a theorem stating some of the properties of the dot product.

Notes:

Theorem 87   Properties of the Dot Product

Let $\vec{u}$, $\vec{v}$ and $\vec{w}$ be vectors in $\mathbb{R}^2$ or $\mathbb{R}^3$ and let c be a scalar.

1. $\vec{u} \cdot \vec{v} = \vec{v} \cdot \vec{u}$ Commutative Property

2. $\vec{u} \cdot (\vec{v} + \vec{w}) = \vec{u} \cdot \vec{v} + \vec{u} \cdot \vec{w}$ Distributive Property

3. $c(\vec{u} \cdot \vec{v}) = (c\vec{u}) \cdot \vec{v} = \vec{u} \cdot (c\vec{v})$

4. $\vec{0} \cdot \vec{v} = 0$

5. $\vec{v} \cdot \vec{v} = \|\vec{v}\|^2$

The last statement of the theorem makes a handy connection between the magnitude of a vector and the dot product with itself. Our definition and theorem give properties of the dot product, but we are still likely wondering "What does the dot product *mean*?" It is helpful to understand that the dot product of a vector with itself is connected to its magnitude.

The next theorem extends this understanding by connecting the dot product to magnitudes and angles. Given vectors $\vec{u}$ and $\vec{v}$ in the plane, an angle θ is clearly formed when $\vec{u}$ and $\vec{v}$ are drawn with the same initial point as illustrated in Figure 11.30(a). (We always take θ to be the angle in $[0, \pi]$ as two angles are actually created.)

The same is also true of 2 vectors in space: given $\vec{u}$ and $\vec{v}$ in $\mathbb{R}^3$ with the same initial point, there is a plane that contains both $\vec{u}$ and $\vec{v}$. (When $\vec{u}$ and $\vec{v}$ are colinear, there are infinite planes that contain both vectors.) In that plane, we can again find an angle θ between them (and again, $0 \le \theta \le \pi$). This is illustrated in Figure 11.30(b).

The following theorem connects this angle θ to the dot product of $\vec{u}$ and $\vec{v}$.

Theorem 88   The Dot Product and Angles

Let $\vec{u}$ and $\vec{v}$ be vectors in $\mathbb{R}^2$ or $\mathbb{R}^3$. Then

$$\vec{u} \cdot \vec{v} = \|\vec{u}\| \ \|\vec{v}\| \cos\theta,$$

where θ, $0 \le \theta \le \pi$, is the angle between $\vec{u}$ and $\vec{v}$.

When θ is an acute angle (i.e., $0 \le \theta < \pi/2$), $\cos\theta$ is positive; when $\theta = \pi/2$, $\cos\theta = 0$; when θ is an obtuse angle ($\pi/2 < \theta \le \pi$), $\cos\theta$ is negative. Thus the sign of the dot product gives a general indication of the angle between the vectors, illustrated in Figure 11.31.
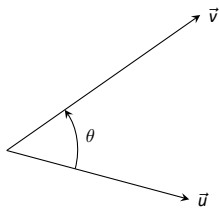
(a)
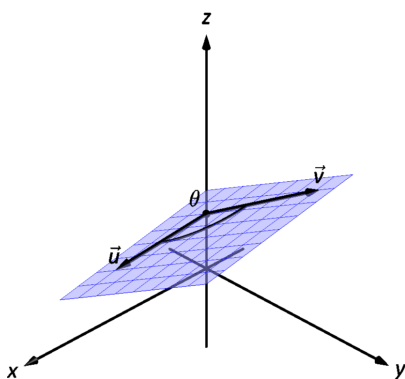
(b)

Figure 11.30:   Illustrating the angle formed by two vectors with the same initial point.

Notes:
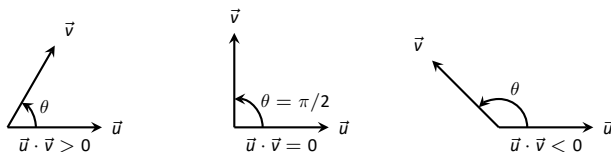
Figure 11.31: Illustrating the relationship between the angle between vectors and the sign of their dot product.

We *can* use Theorem 88 to compute the dot product, but generally this theorem is used to find the angle between known vectors (since the dot product is generally easy to compute). To this end, we rewrite the theorem's equation as

$$\cos\theta = \frac{\vec{u}\cdot\vec{v}}{\|\vec{u}\|\,\|\vec{v}\|} \quad \Leftrightarrow \quad \theta = \cos^{-1}\left(\frac{\vec{u}\cdot\vec{v}}{\|\vec{u}\|\,\|\vec{v}\|}\right).$$

Watch the video:
Vectors: The Dot Product at
`https://youtu.be/98C7iv8OcnI`

We practice using this theorem in the following example.

Example 2   **Using the dot product to find angles**
Let $\vec{u} = \langle 3, 1\rangle$, $\vec{v} = \langle -2, 6\rangle$ and $\vec{w} = \langle -4, 3\rangle$, as shown in Figure 11.32. Find the angles α, β and θ.

Figure 11.32: Vectors used in Example 2.

SOLUTION   We start by computing the magnitude of each vector.

$$\|\vec{u}\| = \sqrt{10}; \quad \|\vec{v}\| = 2\sqrt{10}; \quad \|\vec{w}\| = 5.$$

We now apply Theorem 88 to find the angles.

$$\alpha = \cos^{-1}\left(\frac{\vec{u}\cdot\vec{v}}{(\sqrt{10})(2\sqrt{10})}\right)$$
$$= \cos^{-1}(0) = \frac{\pi}{2} = 90°.$$

Notes:

$$\beta = \cos^{-1}\left(\frac{\vec{v} \cdot \vec{w}}{(2\sqrt{10})(5)}\right)$$

$$= \cos^{-1}\left(\frac{26}{10\sqrt{10}}\right)$$

$$\approx 0.6055 \approx 34.7°.$$

$$\theta = \cos^{-1}\left(\frac{\vec{u} \cdot \vec{w}}{(\sqrt{10})(5)}\right)$$

$$= \cos^{-1}\left(\frac{-9}{5\sqrt{10}}\right)$$

$$\approx 2.1763 \approx 124.7°$$

We see from our computation that $\alpha + \beta = \theta$, as indicated by Figure 11.32. While we knew this should be the case, it is nice to see that this non-intuitive formula indeed returns the results we expected.

We do a similar example next in the context of vectors in space.

Example 3   **Using the dot product to find angles**

Let $\vec{u} = \langle 1, 1, 1\rangle$, $\vec{v} = \langle -1, 3, -2\rangle$ and $\vec{w} = \langle -5, 1, 4\rangle$, as illustrated in Figure 11.33. Find the angle between each pair of vectors.

SOLUTION

1. Between $\vec{u}$ and $\vec{v}$:

$$\theta = \cos^{-1}\left(\frac{\vec{u} \cdot \vec{v}}{\|\vec{u}\|\,\|\vec{v}\|}\right)$$

$$= \cos^{-1}\left(\frac{0}{\sqrt{3}\sqrt{14}}\right)$$

$$= \frac{\pi}{2}.$$

2. Between $\vec{u}$ and $\vec{w}$:

$$\theta = \cos^{-1}\left(\frac{\vec{u} \cdot \vec{w}}{\|\vec{u}\|\,\|\vec{w}\|}\right)$$

$$= \cos^{-1}\left(\frac{0}{\sqrt{3}\sqrt{42}}\right)$$

$$= \frac{\pi}{2}.$$

Figure 11.33: Vectors used in Example 3.

Notes:

3. Between $\vec{v}$ and $\vec{w}$:

$$\theta = \cos^{-1}\left(\frac{\vec{v}\cdot\vec{w}}{\|\vec{v}\|\,\|\vec{w}\|}\right)$$
$$= \cos^{-1}\left(\frac{0}{\sqrt{14}\sqrt{42}}\right)$$
$$= \frac{\pi}{2}.$$

While our work shows that each angle is $\pi/2$, i.e., 90°, none of these angles looks to be a right angle in Figure 11.33. Such is the case when drawing three–dimensional objects on the page.

All three angles between these vectors was $\pi/2$, or 90°. We know from geometry and everyday life that 90° angles are "nice" for a variety of reasons, so it should seem significant that these angles are all $\pi/2$. Notice the common feature in each calculation (and also the calculation of α in Example 2): the dot products of each pair of angles was 0. We use this as a basis for a definition of the term **orthogonal**, which is essentially synonymous to *perpendicular*.

> **Definition 61   Orthogonal**
> Vectors $\vec{u}$ and $\vec{v}$ are **orthogonal** if their dot product is 0.

Note: The term *perpendicular* originally referred to lines. As mathematics progressed, the concept of "being at right angles to" was applied to other objects, such as vectors and planes, and the term *orthogonal* was introduced. It is especially used when discussing objects that are hard, or impossible, to visualize: two vectors in 5-dimensional space are orthogonal if their dot product is 0. It is not wrong to say they are *perpendicular*, but common convention gives preference to the word *orthogonal*.

Example 4   Finding orthogonal vectors
Let $\vec{u} = \langle 3,5\rangle$ and $\vec{v} = \langle 1,2,3\rangle$.

1. Find two vectors in $\mathbb{R}^2$ that are orthogonal to $\vec{u}$.

2. Find two non–parallel vectors in $\mathbb{R}^3$ that are orthogonal to $\vec{v}$.

SOLUTION

1. Recall that a line perpendicular to a line with slope m has slope $-1/m$, the "opposite reciprocal slope." We can think of the slope of $\vec{u}$ as 5/3, its "rise over run." A vector orthogonal to $\vec{u}$ will have slope $-3/5$. There are many such choices, though all parallel:

$$\langle -5,3\rangle \quad \text{or} \quad \langle 5,-3\rangle \quad \text{or} \quad \langle -10,6\rangle \quad \text{or} \quad \langle 15,-9\rangle, \text{etc.}$$

Notes:

2. There are infinite directions in space orthogonal to any given direction, so there are an infinite number of non–parallel vectors orthogonal to $\vec{v}$. Since there are so many, we have great leeway in finding some.

   One way is to arbitrarily pick values for the first two components, leaving the third unknown. For instance, let $\vec{v}_1 = \langle 2, 7, z \rangle$. If $\vec{v}_1$ is to be orthogonal to $\vec{v}$, then $\vec{v}_1 \cdot \vec{v} = 0$, so

   $$2 + 14 + 3z = 0 \quad \Rightarrow \quad z = \frac{-16}{3}.$$

   So $\vec{v}_1 = \langle 2, 7, -16/3 \rangle$ is orthogonal to $\vec{v}$. We can apply a similar technique by leaving the first or second component unknown.

   Another method of finding a vector orthogonal to $\vec{v}$ mirrors what we did in part 1. Let $\vec{v}_2 = \langle -2, 1, 0 \rangle$. Here we switched the first two components of $\vec{v}$, changing the sign of one of them (similar to the "opposite reciprocal" concept before). Letting the third component be 0 effectively ignores the third component of $\vec{v}$, and it is easy to see that

   $$\vec{v}_2 \cdot \vec{v} = \langle -2, 1, 0 \rangle \cdot \langle 1, 2, 3 \rangle = 0.$$

   Clearly $\vec{v}_1$ and $\vec{v}_2$ are not parallel.

An important construction is illustrated in Figure 11.34, where vectors $\vec{u}$ and $\vec{v}$ are sketched. In part (a), a dotted line is drawn from the tip of $\vec{u}$ to the line containing $\vec{v}$, where the dotted line is orthogonal to $\vec{v}$. In part (b), the dotted line is replaced with the vector $\vec{z}$ and $\vec{w}$ is formed, parallel to $\vec{v}$. It is clear by the diagram that $\vec{u} = \vec{w} + \vec{z}$. What is important about this construction is this: $\vec{u}$ is *decomposed* as the sum of two vectors, one of which is parallel to $\vec{v}$ and one that is perpendicular to $\vec{v}$. It is hard to overstate the importance of this construction (as we'll see in upcoming examples).

The vectors $\vec{w}$, $\vec{z}$ and $\vec{u}$ as shown in Figure 11.34 (b) form a right triangle, where the angle between $\vec{v}$ and $\vec{u}$ is labeled θ. We can find $\vec{w}$ in terms of $\vec{v}$ and $\vec{u}$.

Using trigonometry, we can state that

$$\|\vec{w}\| = \|\vec{u}\| \cos\theta. \tag{11.3}$$

We also know that $\vec{w}$ is parallel to to $\vec{v}$; that is, the direction of $\vec{w}$ is the direction of $\vec{v}$, described by the unit vector $\frac{1}{\|\vec{v}\|}\vec{v}$. The vector $\vec{w}$ is the vector in the direction $\frac{1}{\|\vec{v}\|}\vec{v}$ with magnitude $\|\vec{u}\| \cos\theta$:

$$\vec{w} = \left(\|\vec{u}\| \cos\theta \right) \frac{1}{\|\vec{v}\|}\vec{v}.$$

(a)

(b)

Figure 11.34: Developing the construction of the *orthogonal projection*.

Notes:

Replace $\cos\theta$ using Theorem 88:

$$= \left(\|\vec{u}\| \frac{\vec{u}\cdot\vec{v}}{\|\vec{u}\|\,\|\vec{v}\|}\right)\frac{1}{\|\vec{v}\|}\vec{v}$$

$$= \frac{\vec{u}\cdot\vec{v}}{\|\vec{v}\|^2}\vec{v}.$$

Now apply Theorem 87.

$$= \frac{\vec{u}\cdot\vec{v}}{\vec{v}\cdot\vec{v}}\vec{v}.$$

Since this construction is so important, it is given a special name.

Definition 62   Orthogonal Projection

Let $\vec{u}$ and $\vec{v}$ be given. The **orthogonal projection of $\vec{u}$ onto $\vec{v}$**, denoted $\mathrm{proj}_{\vec{v}}\,\vec{u}$, is

$$\mathrm{proj}_{\vec{v}}\,\vec{u} = \frac{\vec{u}\cdot\vec{v}}{\vec{v}\cdot\vec{v}}\vec{v}.$$

Example 5   Computing the orthogonal projection

1. Let $\vec{u} = \langle -2, 1\rangle$ and $\vec{v} = \langle 3, 1\rangle$. Find $\mathrm{proj}_{\vec{v}}\,\vec{u}$, and sketch all three vectors with initial points at the origin.

2. Let $\vec{w} = \langle 2, 1, 3\rangle$ and $\vec{x} = \langle 1, 1, 1\rangle$. Find $\mathrm{proj}_{\vec{x}}\,\vec{w}$, and sketch all three vectors with initial points at the origin.

SOLUTION

1. Applying Definition 62, we have

$$\mathrm{proj}_{\vec{v}}\,\vec{u} = \frac{\vec{u}\cdot\vec{v}}{\vec{v}\cdot\vec{v}}\vec{v}$$

$$= \frac{-5}{10}\langle 3, 1\rangle$$

$$= \left\langle -\frac{3}{2}, -\frac{1}{2}\right\rangle.$$

Vectors $\vec{u}$, $\vec{v}$ and $\mathrm{proj}_{\vec{v}}\,\vec{u}$ are sketched in Figure 11.35(a). Note how the projection is parallel to $\vec{v}$; that is, it lies on the same line through the origin as $\vec{v}$, although it points in the opposite direction. That is because the angle between $\vec{u}$ and $\vec{v}$ is obtuse (i.e., greater than $90°$).

(a)

(b)
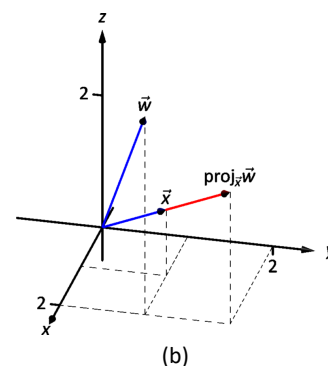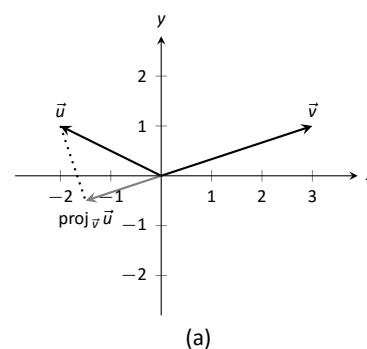
Figure 11.35: Graphing the vectors used in Example 5.

Notes:

2. Apply the definition:

$$\text{proj}_{\,\vec{x}}\,\vec{w} = \frac{\vec{w} \cdot \vec{x}}{\vec{x} \cdot \vec{x}}\vec{x}$$
$$= \frac{6}{3}\langle 1, 1, 1 \rangle$$
$$= \langle 2, 2, 2 \rangle .$$

These vectors are sketched in Figure 11.35(b).

Consider Figure 11.36 where the concept of the orthogonal projection is again illustrated. It is clear that

$$\vec{u} = \text{proj}_{\,\vec{v}}\,\vec{u} + \vec{z}. \tag{11.4}$$

As we know what $\vec{u}$ and $\text{proj}_{\,\vec{v}}\,\vec{u}$ are, we can solve for $\vec{z}$ and state that

$$\vec{z} = \vec{u} - \text{proj}_{\,\vec{v}}\,\vec{u}.$$

This leads us to rewrite Equation (11.4) in a seemingly silly way:

$$\vec{u} = \text{proj}_{\,\vec{v}}\,\vec{u} + (\vec{u} - \text{proj}_{\,\vec{v}}\,\vec{u}).$$

This is not nonsense, as pointed out in the following Key Idea. (Notation note: the expression "$\parallel \vec{y}$" means "is parallel to $\vec{y}$." We can use this notation to state "$\vec{x} \parallel \vec{y}$" which means "$\vec{x}$ is parallel to $\vec{y}$." The expression "$\perp \vec{y}$" means "is orthogonal to $\vec{y}$," and is used similarly.)

Key Idea 52   **Orthogonal Decomposition of Vectors**

Let $\vec{u}$ and $\vec{v}$ be given. Then $\vec{u}$ can be written as the sum of two vectors, one of which is parallel to $\vec{v}$, and one of which is orthogonal to $\vec{v}$:

$$\vec{u} = \underbrace{\text{proj}_{\,\vec{v}}\,\vec{u}}_{\parallel\,\vec{v}} + \underbrace{(\vec{u} - \text{proj}_{\,\vec{v}}\,\vec{u})}_{\perp\,\vec{v}}.$$

We illustrate the use of this equality in the following example.

Example 6   **Orthogonal decomposition of vectors**

1. Let $\vec{u} = \langle -2, 1 \rangle$ and $\vec{v} = \langle 3, 1 \rangle$ as in Example 5. Decompose $\vec{u}$ as the sum of a vector parallel to $\vec{v}$ and a vector orthogonal to $\vec{v}$.

2. Let $\vec{w} = \langle 2, 1, 3 \rangle$ and $\vec{x} = \langle 1, 1, 1 \rangle$ as in Example 5. Decompose $\vec{w}$ as the sum of a vector parallel to $\vec{x}$ and a vector orthogonal to $\vec{x}$.

Figure 11.36: Illustrating the orthogonal projection.

Notes:

SOLUTION

1. In Example 5, we found that $\text{proj}_{\vec{v}}\,\vec{u} = \langle -1.5, -0.5 \rangle$. Let

$$\vec{z} = \vec{u} - \text{proj}_{\vec{v}}\,\vec{u} = \langle -2, 1 \rangle - \langle -1.5, -0.5 \rangle = \langle -0.5, 1.5 \rangle\,.$$

   Is $\vec{z}$ orthogonal to $\vec{v}$? (I.e, is $\vec{z} \perp \vec{v}$?) We check for orthogonality with the dot product:

$$\vec{z} \cdot \vec{v} = \langle -0.5, 1.5 \rangle \cdot \langle 3, 1 \rangle = 0.$$

   Since the dot product is 0, we know $\vec{z} \perp \vec{v}$. Thus:

$$\vec{u} = \text{proj}_{\vec{v}}\,\vec{u} \ + \ (\vec{u} - \text{proj}_{\vec{v}}\,\vec{u})$$

$$\langle -2, 1 \rangle = \underbrace{\langle -1.5, -0.5 \rangle}_{\parallel\,\vec{v}} + \underbrace{\langle -0.5, 1.5 \rangle}_{\perp\,\vec{v}}.$$

2. We found in Example 5 that $\text{proj}_{\vec{x}}\,\vec{w} = \langle 2, 2, 2 \rangle$. Applying the Key Idea 52, we have:

$$\vec{z} = \vec{w} - \text{proj}_{\vec{x}}\,\vec{w} = \langle 2, 1, 3 \rangle - \langle 2, 2, 2 \rangle = \langle 0, -1, 1 \rangle\,.$$

   We check to see if $\vec{z} \perp \vec{x}$:

$$\vec{z} \cdot \vec{x} = \langle 0, -1, 1 \rangle \cdot \langle 1, 1, 1 \rangle = 0.$$

   Since the dot product is 0, we know the two vectors are orthogonal. We now write $\vec{w}$ as the sum of two vectors, one parallel and one orthogonal to $\vec{x}$:

$$\vec{w} = \text{proj}_{\vec{x}}\,\vec{w} \ + \ (\vec{w} - \text{proj}_{\vec{x}}\,\vec{w})$$

$$\langle 2, 1, 3 \rangle = \underbrace{\langle 2, 2, 2 \rangle}_{\parallel\,\vec{x}} + \underbrace{\langle 0, -1, 1 \rangle}_{\perp\,\vec{x}}$$

We give an example of where this decomposition is useful.

Example 7   **Orthogonally decomposing a force vector**
Consider Figure 11.37(a), showing a box weighing 50 lb on a ramp that rises 5 ft over a span of 20 ft. Find the components of force, and their magnitudes, acting on the box (as sketched in part (b) of the figure):

1. in the direction of the ramp, and

2. orthogonal to the ramp.

Notes:

(a)

(b)
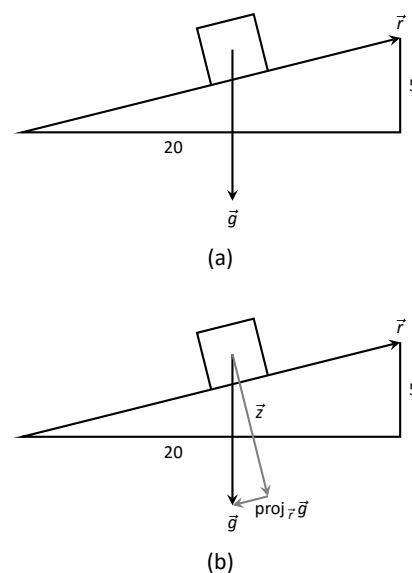
Figure 11.37: Sketching the ramp and box in Example 7. Note: *The vectors are not drawn to scale.*

671

Solution   As the ramp rises 5 ft over a horizontal distance of 20 ft, we can represent the direction of the ramp with the vector $\vec{r} = \langle 20, 5 \rangle$. Gravity pulls down with a force of 50 lb, which we represent with $\vec{g} = \langle 0, -50 \rangle$.

1. To find the force of gravity in the direction of the ramp, we compute $\text{proj}_{\vec{r}}\,\vec{g}$:

$$\text{proj}_{\vec{r}}\,\vec{g} = \frac{\vec{g} \cdot \vec{r}}{\vec{r} \cdot \vec{r}}\vec{r}$$
$$= \frac{-250}{425}\langle 20, 5 \rangle$$
$$= \left\langle -\frac{200}{17}, -\frac{50}{17} \right\rangle.$$

The magnitude of $\text{proj}_{\vec{r}}\,\vec{g}$ is $\|\text{proj}_{\vec{r}}\,\vec{g}\| = 50/\sqrt{17} \approx 12.13$ lb. Though the box weighs 50 lb, a force of about 12 lb is enough to keep the box from sliding down the ramp.

2. To find the component $\vec{z}$ of gravity orthogonal to the ramp, we use Key Idea 52.

$$\vec{z} = \vec{g} - \text{proj}_{\vec{r}}\,\vec{g}$$
$$= \left\langle \frac{200}{17}, -\frac{800}{17} \right\rangle.$$

The magnitude of this force is $\|\vec{z}\| = 200/\sqrt{17}$ lb. In physics and engineering, knowing this force is important when computing things like static frictional force. (For instance, we could easily compute if the static frictional force alone was enough to keep the box from sliding down the ramp.)

Application to Work

In physics, the application of a force F to move an object in a straight line a distance d produces *work*; the amount of work W is $W = Fd$, (where F is in the direction of travel). The orthogonal projection allows us to compute work when the force is not in the direction of travel.

Consider Figure 11.38, where a force $\vec{F}$ is being applied to an object moving in the direction of $\vec{d}$. (The distance the object travels is the magnitude of $\vec{d}$.) The work done is the amount of force in the direction of $\vec{d}$, $\|\text{proj}_{\vec{d}}\,\vec{F}\|$, times $\|\vec{d}\|$:
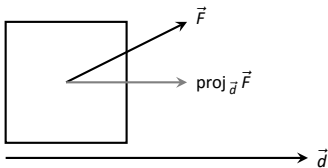
Figure 11.38: Finding work when the force and direction of travel are given as vectors.

Notes:

$$\left\| \text{proj}_{\vec{d}}\, \vec{F} \right\| \cdot \left\| \vec{d} \right\| = \left\| \frac{\vec{F} \cdot \vec{d}}{\vec{d} \cdot \vec{d}}\vec{d} \right\| \cdot \left\| \vec{d} \right\|$$

$$= \left| \frac{\vec{F} \cdot \vec{d}}{\left\| \vec{d} \right\|^2} \right| \cdot \left\| \vec{d} \right\| \cdot \left\| \vec{d} \right\|$$

$$= \frac{\left| \vec{F} \cdot \vec{d} \right|}{\left\| \vec{d} \right\|^2} \left\| \vec{d} \right\|^2$$

$$= \left| \vec{F} \cdot \vec{d} \right|.$$

The expression $\vec{F} \cdot \vec{d}$ will be positive if the angle between $\vec{F}$ and $\vec{d}$ is acute; when the angle is obtuse (hence $\vec{F} \cdot \vec{d}$ is negative), the force is causing motion in the opposite direction of $\vec{d}$, resulting in "negative work." We want to capture this sign, so we drop the absolute value and find that $W = \vec{F} \cdot \vec{d}$.

> **Definition 63   Work**
> Let $\vec{F}$ be a constant force that moves an object in a straight line from point P to point Q. Let $\vec{d} = \overrightarrow{PQ}$. The **work** W done by $\vec{F}$ along $\vec{d}$ is $W = \vec{F} \cdot \vec{d}$.

Example 8   Computing work
A man slides a box along a ramp that rises 3 ft over a distance of 15 ft by applying 50 lb of force as shown in Figure 11.39. Compute the work done.

Figure 11.39: Computing work when sliding a box up a ramp in Example 8.

SOLUTION   The figure indicates that the force applied makes a 30° angle with the horizontal, so $\vec{F} = 50 \left\langle \cos 30°, \sin 30° \right\rangle \left\langle 25\sqrt{3}, 25 \right\rangle$. The ramp is represented by $\vec{d} = \langle 15, 3 \rangle$. The work done is simply

$$\vec{F} \cdot \vec{d} = \left\langle 25\sqrt{3}, 25 \right\rangle \cdot \langle 15, 3 \rangle = 375\sqrt{3} + 75 \text{ ft–lb.}$$

Note how we did not actually compute the distance the object traveled, nor the magnitude of the force in the direction of travel; this is all inherently computed by the dot product!

Notes:

The dot product is a powerful way of evaluating computations that depend on angles without actually using angles. The next section explores another "product" on vectors, the *cross product.* Once again, angles play an important role, though in a much different way.

Notes:

Exercises 11.3

Terms and Concepts

1. The dot product of two vectors is a _____, not a vector.
2. How are the concepts of the dot product and vector magnitude related?
3. How can one quickly tell if the angle between two vectors is acute or obtuse?
4. Give a synonym for "orthogonal."

Problems

In Exercises 5–10, find the dot product of the given vectors.

5. $\vec{u} = \langle 2, -4 \rangle$, $\vec{v} = \langle 3, 7 \rangle$
6. $\vec{u} = \langle 5, 3 \rangle$, $\vec{v} = \langle 6, 1 \rangle$
7. $\vec{u} = \langle 1, -1, 2 \rangle$, $\vec{v} = \langle 2, 5, 3 \rangle$
8. $\vec{u} = \langle 3, 5, -1 \rangle$, $\vec{v} = \langle 4, -1, 7 \rangle$
9. $\vec{u} = \langle 1, 1 \rangle$, $\vec{v} = \langle 1, 2, 3 \rangle$
10. $\vec{u} = \langle 1, 2, 3 \rangle$, $\vec{v} = \langle 0, 0, 0 \rangle$
11. Create your own vectors $\vec{u}$, $\vec{v}$ and $\vec{w}$ in $\mathbb{R}^2$ and show that $\vec{u} \cdot (\vec{v} + \vec{w}) = \vec{u} \cdot \vec{v} + \vec{u} \cdot \vec{w}$.
12. Create your own vectors $\vec{u}$ and $\vec{v}$ in $\mathbb{R}^3$ and scalar c and show that $c(\vec{u} \cdot \vec{v}) = \vec{u} \cdot (c\vec{v})$.

In Exercises 13–16, find the measure of the angle between the two vectors in both radians and degrees.

13. $\vec{u} = \langle 1, 1 \rangle$, $\vec{v} = \langle 1, 2 \rangle$
14. $\vec{u} = \langle -2, 1 \rangle$, $\vec{v} = \langle 3, 5 \rangle$
15. $\vec{u} = \langle 8, 1, -4 \rangle$, $\vec{v} = \langle 2, 2, 0 \rangle$
16. $\vec{u} = \langle 1, 7, 2 \rangle$, $\vec{v} = \langle 4, -2, 5 \rangle$

In Exercises 17–20, a vector $\vec{v}$ is given. Give two vectors that are orthogonal to $\vec{v}$.

17. $\vec{v} = \langle 4, 7 \rangle$
18. $\vec{v} = \langle -3, 5 \rangle$
19. $\vec{v} = \langle 1, 1, 1 \rangle$
20. $\vec{v} = \langle 1, -2, 3 \rangle$

In Exercises 21–26, vectors $\vec{u}$ and $\vec{v}$ are given. Find $\text{proj}_{\vec{v}}\, \vec{u}$, the orthogonal projection of $\vec{u}$ onto $\vec{v}$, and sketch all three vectors on the same axes.

21. $\vec{u} = \langle 1, 2 \rangle$, $\vec{v} = \langle -1, 3 \rangle$
22. $\vec{u} = \langle 5, 5 \rangle$, $\vec{v} = \langle 1, 3 \rangle$
23. $\vec{u} = \langle -3, 2 \rangle$, $\vec{v} = \langle 1, 1 \rangle$
24. $\vec{u} = \langle -3, 2 \rangle$, $\vec{v} = \langle 2, 3 \rangle$
25. $\vec{u} = \langle 1, 5, 1 \rangle$, $\vec{v} = \langle 1, 2, 3 \rangle$
26. $\vec{u} = \langle 3, -1, 2 \rangle$, $\vec{v} = \langle 2, 2, 1 \rangle$

In Exercises 27–32, vectors $\vec{u}$ and $\vec{v}$ are given. Write $\vec{u}$ as the sum of two vectors, one of which is parallel to $\vec{v}$ and one of which is perpendicular to $\vec{v}$. Note: these are the same pairs of vectors as found in Exercises 21 – 26.

27. $\vec{u} = \langle 1, 2 \rangle$, $\vec{v} = \langle -1, 3 \rangle$
28. $\vec{u} = \langle 5, 5 \rangle$, $\vec{v} = \langle 1, 3 \rangle$
29. $\vec{u} = \langle -3, 2 \rangle$, $\vec{v} = \langle 1, 1 \rangle$
30. $\vec{u} = \langle -3, 2 \rangle$, $\vec{v} = \langle 2, 3 \rangle$
31. $\vec{u} = \langle 1, 5, 1 \rangle$, $\vec{v} = \langle 1, 2, 3 \rangle$
32. $\vec{u} = \langle 3, -1, 2 \rangle$, $\vec{v} = \langle 2, 2, 1 \rangle$
33. A 10lb box sits on a ramp that rises 4ft over a distance of 20ft. How much force is required to keep the box from sliding down the ramp?
34. A 10lb box sits on a 15ft ramp that makes a $30°$ angle with the horizontal. How much force is required to keep the box from sliding down the ramp?
35. How much work is performed in moving a box horizontally 10ft with a force of 20lb applied at an angle of $45°$ to the horizontal?
36. How much work is performed in moving a box horizontally 10ft with a force of 20lb applied at an angle of $10°$ to the horizontal?
37. How much work is performed in moving a box up the length of a ramp that rises 2ft over a distance of 10ft, with a force of 50lb applied horizontally?
38. How much work is performed in moving a box up the length of a ramp that rises 2ft over a distance of 10ft, with a force of 50lb applied at an angle of $45°$ to the horizontal?
39. How much work is performed in moving a box up the length of a 10ft ramp that makes a $5°$ angle with the horizontal, with 50lb of force applied in the direction of the ramp?

11.4   The Cross Product

"Orthogonality" is immensely important. A quick scan of your current environment will undoubtedly reveal numerous surfaces and edges that are perpendicular to each other (including the edges of this page). The dot product provides a quick test for orthogonality: vectors $\vec{u}$ and $\vec{v}$ are perpendicular if, and only if, $\vec{u} \cdot \vec{v} = 0$.

Given two non–parallel, nonzero vectors $\vec{u}$ and $\vec{v}$ in space, it is very useful to find a vector $\vec{w}$ that is perpendicular to both $\vec{u}$ and $\vec{v}$. There is a operation, called the **cross product**, that creates such a vector. This section defines the cross product, then explores its properties and applications.

Definition 64   **Cross Product**

Let $\vec{u} = \langle u_1, u_2, u_3 \rangle$ and $\vec{v} = \langle v_1, v_2, v_3 \rangle$ be vectors in $\mathbb{R}^3$. The **cross product of $\vec{u}$ and $\vec{v}$**, denoted $\vec{u} \times \vec{v}$, is the vector

$$\vec{u} \times \vec{v} = \langle u_2 v_3 - u_3 v_2, -(u_1 v_3 - u_3 v_1), u_1 v_2 - u_2 v_1 \rangle .$$

This definition can be a bit cumbersome to remember. After an example we will give a convenient method for computing the cross product. For now, careful examination of the products and differences given in the definition should reveal a pattern that is not too difficult to remember. (For instance, in the first component only 2 and 3 appear as subscripts; in the second component, only 1 and 3 appear as subscripts. Further study reveals the order in which they appear.)

Watch the video:
Cross Product at
`https://youtu.be/qsgK1d-_8ik`

Let's practice using this definition by computing a cross product.

Example 1   **Computing a cross product**

Let $\vec{u} = \langle 2, -1, 4 \rangle$ and $\vec{v} = \langle 3, 2, 5 \rangle$. Find $\vec{u} \times \vec{v}$, and verify that it is orthogonal to both $\vec{u}$ and $\vec{v}$.

Notes:

SOLUTION   Using Definition 64, we have

$$\vec{u} \times \vec{v} = \langle (-1)5 - (4)2, -((2)5 - (4)3), (2)2 - (-1)3 \rangle = \langle -13, 2, 7 \rangle.$$

(We encourage the reader to compute this product on their own, then verify their result.)

We test whether or not $\vec{u} \times \vec{v}$ is orthogonal to $\vec{u}$ and $\vec{v}$ using the dot product:

$$(\vec{u} \times \vec{v}) \cdot \vec{u} = \langle -13, 2, 7 \rangle \cdot \langle 2, -1, 4 \rangle = 0,$$
$$(\vec{u} \times \vec{v}) \cdot \vec{v} = \langle -13, 2, 7 \rangle \cdot \langle 3, 2, 5 \rangle = 0.$$

Since both dot products are zero, $\vec{u} \times \vec{v}$ is indeed orthogonal to both $\vec{u}$ and $\vec{v}$.

A convenient method of computing the cross product starts with forming a particular 3×3 *matrix*, or rectangular array. The first row comprises the standard unit vectors $\vec{i}, \vec{j}$, and $\vec{k}$. The second and third rows are the vectors $\vec{u}$ and $\vec{v}$, respectively. Using $\vec{u}$ and $\vec{v}$ from Example 1, we begin with:

$$\begin{array}{ccc} \vec{i} & \vec{j} & \vec{k} \\ 2 & -1 & 4 \\ 3 & 2 & 5 \end{array}$$

Now repeat the first two columns after the original three:

$$\begin{array}{ccccc} \vec{i} & \vec{j} & \vec{k} & \vec{i} & \vec{j} \\ 2 & -1 & 4 & 2 & -1 \\ 3 & 2 & 5 & 3 & 2 \end{array}$$

This gives three full "upper left to lower right" diagonals, and three full "upper right to lower left" diagonals, as shown. Compute the products along each diagonal, then add the products on the right and subtract the products on the left:

$$\vec{u} \times \vec{v} = \left(-5\vec{i} + 12\vec{j} + 4\vec{k}\right) - \left(-3\vec{k} + 8\vec{i} + 10\vec{j}\right) = -13\vec{i} + 2\vec{j} + 7\vec{k} = \langle -13, 2, 7 \rangle.$$

This is equivalent to evaluating the determinant

$$\begin{vmatrix} \vec{i} & \vec{j} & \vec{k} \\ 2 & -1 & 4 \\ 3 & 2 & 5 \end{vmatrix} = \begin{vmatrix} -1 & 4 \\ 2 & 5 \end{vmatrix}\vec{i} - \begin{vmatrix} 2 & 4 \\ 3 & 5 \end{vmatrix}\vec{j} + \begin{vmatrix} 2 & -1 \\ 3 & 2 \end{vmatrix}\vec{k}$$

$$= (-5 - 8)\vec{i} - (10 - 12)\vec{j} + (4 - (-3))\vec{k} = -13\vec{i} + 2\vec{j} + 7\vec{k}.$$

We practice using this method.

Notes:

Example 2   Computing a cross product
Let $\vec{u} = \langle 1, 3, 6 \rangle$ and $\vec{v} = \langle -1, 2, 1 \rangle$. Compute both $\vec{u} \times \vec{v}$ and $\vec{v} \times \vec{u}$.

SOLUTION   To compute $\vec{u} \times \vec{v}$, we form the matrix as prescribed above, complete with repeated first columns:

$$
\begin{array}{ccccc}
\vec{i} & \vec{j} & \vec{k} & \vec{i} & \vec{j} \\
1 & 3 & 6 & 1 & 3 \\
-1 & 2 & 1 & -1 & 2
\end{array}
$$

We let the reader compute the products of the diagonals; we give the result:

$$
\vec{u} \times \vec{v} = \left(3\vec{i} - 6\vec{j} + 2\vec{k} \right) - \left(-3\vec{k} + 12\vec{i} + \vec{j} \right) = \langle -9, -7, 5 \rangle .
$$

To compute $\vec{v} \times \vec{u}$, we switch the second and third rows of the above matrix, then multiply along diagonals and subtract:

$$
\begin{array}{ccccc}
\vec{i} & \vec{j} & \vec{k} & \vec{i} & \vec{j} \\
-1 & 2 & 1 & -1 & 2 \\
1 & 3 & 6 & 1 & 3
\end{array}
$$

Note how with the rows being switched, the products that once appeared on the right now appear on the left, and vice–versa. Thus the result is:

$$
\vec{v} \times \vec{u} = \left(12\vec{i} + \vec{j} - 3\vec{k} \right) - \left(2\vec{k} + 3\vec{i} - 6\vec{j} \right) = \langle 9, 7, -5 \rangle ,
$$

which is the opposite of $\vec{u} \times \vec{v}$. We leave it to the reader to verify that each of these vectors is orthogonal to $\vec{u}$ and $\vec{v}$.

Properties of the Cross Product

It is not coincidence that $\vec{v} \times \vec{u} = -(\vec{u} \times \vec{v})$ in the preceding example; one can show using Definition 64 that this will always be the case. The following theorem states several useful properties of the cross product, each of which can be verified by referring to the definition.

Notes:

Theorem 89   Properties of the Cross Product

Let $\vec{u}$, $\vec{v}$ and $\vec{w}$ be vectors in $\mathbb{R}^3$ and let c be a scalar. The following identities hold:

1. $\vec{u} \times \vec{v} = -(\vec{v} \times \vec{u})$ Anticommutative Property

2. (a) $(\vec{u} + \vec{v}) \times \vec{w} = \vec{u} \times \vec{w} + \vec{v} \times \vec{w}$ Distributive Properties
   (b) $\vec{u} \times (\vec{v} + \vec{w}) = \vec{u} \times \vec{v} + \vec{u} \times \vec{w}$

3. $c(\vec{u} \times \vec{v}) = (c\vec{u}) \times \vec{v} = \vec{u} \times (c\vec{v})$

4. (a) $(\vec{u} \times \vec{v}) \cdot \vec{u} = 0$ Orthogonality Properties
   (b) $(\vec{u} \times \vec{v}) \cdot \vec{v} = 0$

5. $\vec{u} \times \vec{u} = \vec{0}$

6. $\vec{u} \times \vec{0} = \vec{0}$

7. $\vec{u} \cdot (\vec{v} \times \vec{w}) = (\vec{u} \times \vec{v}) \cdot \vec{w}$ Triple Scalar Product

We introduced the cross product as a way to find a vector orthogonal to two given vectors, but we did not give a proof that the construction given in Definition 64 satisfies this property. Theorem 89 asserts this property holds; we leave the verification to Exercise 42.

Property 5 from the theorem is also left to the reader to prove in Exercise 43, but it reveals something more interesting than "the cross product of a vector with itself is $\vec{0}$." Let $\vec{u}$ and $\vec{v}$ be parallel vectors; that is, let there be a scalar c such that $\vec{v} = c\vec{u}$. Consider their cross product:

$$\begin{aligned}
\vec{u} \times \vec{v} &= \vec{u} \times (c\vec{u}) \\
&= c(\vec{u} \times \vec{u}) \quad \text{(by Property 3 of Theorem 89)} \\
&= \vec{0}. \quad\quad\quad\; \text{(by Property 5 of Theorem 89)}
\end{aligned}$$

We have just shown that the cross product of parallel vectors is $\vec{0}$. This hints at something deeper. Theorem 88 related the angle between two vectors and their dot product; there is a similar relationship relating the cross product of two vectors and the angle between them, given by the following theorem.

Notes:

> **Theorem 90   The Cross Product and Angles**
> Let $\vec{u}$ and $\vec{v}$ be vectors in $\mathbb{R}^3$. Then
>
> $$\|\vec{u} \times \vec{v}\| = \|\vec{u}\| \, \|\vec{v}\| \sin \theta,$$
>
> where θ, $0 \leq \theta \leq \pi$, is the angle between $\vec{u}$ and $\vec{v}$.

Note: Definition 61 (through Theorem 88) defines $\vec{u}$ and $\vec{v}$ to be orthogonal if $\vec{u} \cdot \vec{v} = 0$. We could use Theorem 90 to define $\vec{u}$ and $\vec{v}$ are parallel if $\vec{u} \times \vec{v} = 0$. By such a definition, $\vec{0}$ would be both orthogonal and parallel to every vector. Apparent paradoxes such as this are not uncommon in mathematics and can be very useful. (See also the marginal note on page 657.)

Note that this theorem makes a statement about the *magnitude* of the cross product. When the angle between $\vec{u}$ and $\vec{v}$ is 0 or π (i.e., the vectors are parallel), the magnitude of the cross product is 0. The only vector with a magnitude of 0 is $\vec{0}$ (see Property 9 of Theorem 86), hence the cross product of parallel vectors is $\vec{0}$.

We demonstrate the truth of this theorem in the following example.

Example 3   The cross product and angles
Let $\vec{u} = \langle 1, 3, 6 \rangle$ and $\vec{v} = \langle -1, 2, 1 \rangle$ as in Example 2. Verify Theorem 90 by finding θ, the angle between $\vec{u}$ and $\vec{v}$, and the magnitude of $\vec{u} \times \vec{v}$.

SOLUTION   We use Theorem 88 to find the angle between $\vec{u}$ and $\vec{v}$.

$$\theta = \cos^{-1} \left(\frac{\vec{u} \cdot \vec{v}}{\|\vec{u}\| \, \|\vec{v}\|} \right)$$

$$= \cos^{-1} \left(\frac{11}{\sqrt{46}\sqrt{6}} \right)$$

$$\approx 0.8471 = 48.54°.$$

Our work in Example 2 showed that $\vec{u} \times \vec{v} = \langle -9, -7, 5 \rangle$, hence $\|\vec{u} \times \vec{v}\| = \sqrt{155}$. Is $\|\vec{u} \times \vec{v}\| = \|\vec{u}\| \, \|\vec{v}\| \sin \theta$? Using numerical approximations, we find:

$$\|\vec{u} \times \vec{v}\| = \sqrt{155} \qquad \|\vec{u}\| \, \|\vec{v}\| \sin \theta = \sqrt{46}\sqrt{6} \sin 0.8471$$

$$\approx 12.45. \qquad\qquad\qquad \approx 12.45.$$

Numerically, they seem equal. Using a right triangle, one can show that

$$\sin \left(\cos^{-1} \left(\frac{11}{\sqrt{46}\sqrt{6}} \right) \right) = \frac{\sqrt{155}}{\sqrt{46}\sqrt{6}},$$

which allows us to verify the theorem exactly.

Notes:

Right Hand Rule

The anticommutative property of the cross product demonstrates that $\vec{u} \times \vec{v}$ and $\vec{v} \times \vec{u}$ differ only by a sign — these vectors have the same magnitude but point in the opposite direction. When seeking a vector perpendicular to $\vec{u}$ and $\vec{v}$, we essentially have two directions to choose from, one in the direction of $\vec{u} \times \vec{v}$ and one in the direction of $\vec{v} \times \vec{u}$. Does it matter which we choose? How can we tell which one we will get without graphing, etc.?

Another property of the cross product, as defined, is that it follows the **right hand rule.** Given $\vec{u}$ and $\vec{v}$ in $\mathbb{R}^3$ with the same initial point, point the index finger of your right hand in the direction of $\vec{u}$ and let your middle finger point in the direction of $\vec{v}$ (much as we did when establishing the right hand rule for the 3-dimensional coordinate system). Your thumb will naturally extend in the direction of $\vec{u} \times \vec{v}$. One can "practice" this using Figure 11.40. If you switch, and point the index finder in the direction of $\vec{v}$ and the middle finger in the direction of $\vec{u}$, your thumb will now point in the opposite direction, allowing you to "visualize" the anticommutative property of the cross product.

Applications of the Cross Product

There are a number of ways in which the cross product is useful in mathematics, physics and other areas of science beyond "just" finding a vector perpendicular to two others. We highlight a few here.

Area of a Parallelogram

It is a standard geometry fact that the area of a parallelogram is $A = bh$, where b is the length of the base and h is the height of the parallelogram, as illustrated in Figure 11.41(a). As shown when defining the Parallelogram Law of vector addition, two vectors $\vec{u}$ and $\vec{v}$ define a parallelogram when drawn from the same initial point, as illustrated in Figure 11.41(b). Trigonometry tells us that $h = \|\vec{u}\| \sin\theta$, hence the area of the parallelogram is

$$A = \|\vec{u}\| \, \|\vec{v}\| \sin\theta = \|\vec{u} \times \vec{v}\|, \tag{11.5}$$

where the second equality comes from Theorem 90. We illustrate using Equation (11.5) in the following example.

Example 4   Finding the area of a parallelogram

1. Find the area of the parallelogram defined by the vectors $\vec{u} = \langle 2, 1 \rangle$ and $\vec{v} = \langle 1, 3 \rangle$.

Notes:

Figure 11.40: Illustrating the Right Hand Rule of the cross product.
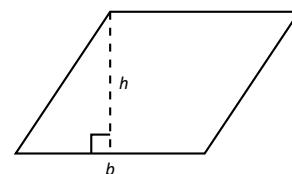
(a)
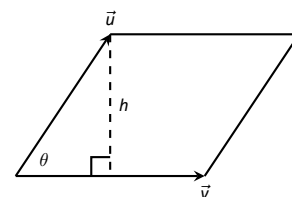
(b)

Figure 11.41: Using the cross product to find the area of a parallelogram.

(a)

(b)
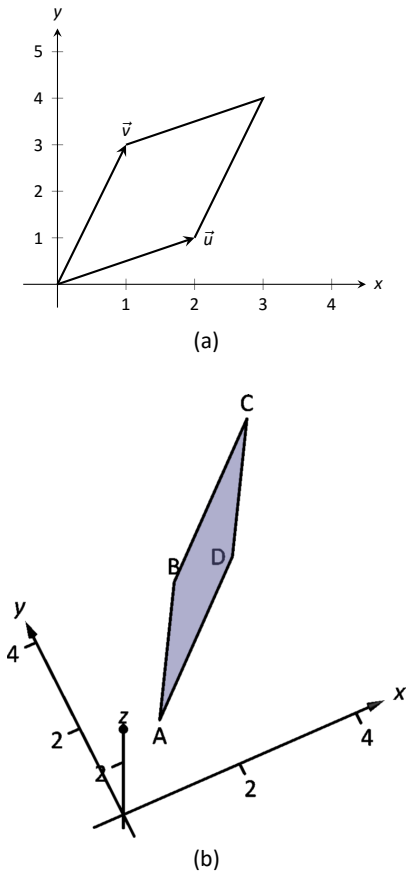
Figure 11.42: Sketching the parallelograms in Example 4.
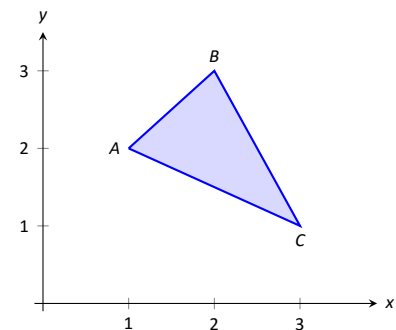
Figure 11.43: Finding the area of a triangle in Example 5.

2. Verify that the points $A = (1, 1, 1)$, $B = (2, 3, 2)$, $C = (4, 5, 3)$ and $D = (3, 3, 2)$ are the vertices of a parallelogram. Find the area of the parallelogram.

SOLUTION

1. Figure 11.42(a) sketches the parallelogram defined by the vectors $\vec{u}$ and $\vec{v}$. We have a slight problem in that our vectors exist in $\mathbb{R}^2$, not $\mathbb{R}^3$, and the cross product is only defined on vectors in $\mathbb{R}^3$. We skirt this issue by viewing $\vec{u}$ and $\vec{v}$ as vectors in the $x-y$ plane of $\mathbb{R}^3$, and rewrite them as $\vec{u} = \langle 2, 1, 0 \rangle$ and $\vec{v} = \langle 1, 3, 0 \rangle$. We can now compute the cross product. It is easy to show that $\vec{u} \times \vec{v} = \langle 0, 0, 5 \rangle$; therefore the area of the parallelogram is $A = \| \vec{u} \times \vec{v} \| = 5$.

2. To show that the quadrilateral $ABCD$ is a parallelogram (shown in Figure 11.42(b)), we need to show that the opposite sides are parallel. We can quickly show that $\overrightarrow{AB} = \overrightarrow{DC} = \langle 1, 2, 1 \rangle$ and $\overrightarrow{BC} = \overrightarrow{AD} = \langle 2, 2, 1 \rangle$. We find the area by computing the magnitude of the cross product of $\overrightarrow{AB}$ and $\overrightarrow{BC}$:

$$\overrightarrow{AB} \times \overrightarrow{BC} = \langle 0, 1, -2 \rangle \quad \Rightarrow \quad \left\| \overrightarrow{AB} \times \overrightarrow{BC} \right\| = \sqrt{5}.$$

This application is more commonly used to find the area of a triangle (because triangles are used more often than parallelograms). We illustrate this in the following example.

Example 5   **Area of a triangle**
Find the area of the triangle with vertices $A = (1, 2)$, $B = (2, 3)$ and $C = (3, 1)$, as pictured in Figure 11.43.

SOLUTION   We found the area of this triangle in Example 6.1.5 to be 1.5 using integration. There we discussed the fact that finding the area of a triangle can be inconvenient using the "$\frac{1}{2}bh$" formula as one has to compute the height, which generally involves finding angles, etc. Using a cross product is much more direct.

We can choose any two sides of the triangle to use to form vectors; we choose $\overrightarrow{AB} = \langle 1, 1 \rangle$ and $\overrightarrow{AC} = \langle 2, -1 \rangle$. As in the previous example, we will rewrite these vectors with a third component of 0 so that we can apply the cross product. The area of the triangle is

$$\frac{1}{2} \left\| \overrightarrow{AB} \times \overrightarrow{AC} \right\| = \frac{1}{2} \| \langle 1, 1, 0 \rangle \times \langle 2, -1, 0 \rangle \| = \frac{1}{2} \| \langle 0, 0, -3 \rangle \| = \frac{3}{2}.$$

We arrive at the same answer as before with less work.

_____

Notes:

Volume of a Parallelepiped

The three dimensional analogue to the parallelogram is the **parallelepiped**. Each face is parallel to the opposite face, as illustrated in Figure 11.44. By crossing $\vec{v}$ and $\vec{w}$, one gets a vector whose magnitude is the area of the base. Dotting this vector with $\vec{u}$ computes the volume of parallelepiped! (Up to a sign; take the absolute value.)

Thus the volume of a parallelepiped defined by vectors $\vec{u}$, $\vec{v}$ and $\vec{w}$ is

$$V = |\vec{u} \cdot (\vec{v} \times \vec{w})| . \tag{11.6}$$

Note how this is the Triple Scalar Product, first seen in Theorem 89. Applying the identities given in the theorem shows that we can apply the Triple Scalar Product in any "order" we choose to find the volume. That is,

$$V = |\vec{u} \cdot (\vec{v} \times \vec{w})| = |\vec{u} \cdot (\vec{w} \times \vec{v})| = |(\vec{u} \times \vec{v}) \cdot \vec{w}| , \quad \text{etc.}$$

As with the cross product, we can also write $\vec{u}\cdot(\vec{v}\times\vec{w})$ in terms of a determinant:

$$\vec{u} \cdot (\vec{v} \times \vec{w}) = \begin{vmatrix} u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \\ w_1 & w_2 & w_3 \end{vmatrix} .$$

Because the volume is the absolute value of the determinant, the order of the rows doesn't matter.

Example 6   Finding the volume of parallelepiped
Find the volume of the parallelepiped defined by the vectors $\vec{u} = \langle 1, 1, 0 \rangle$, $\vec{v} = \langle -1, 1, 0 \rangle$ and $\vec{w} = \langle 0, 1, 1 \rangle$.

Solution   We apply Equation (11.6). We first find $\vec{v} \times \vec{w} = \langle 1, 1, -1 \rangle$. Then

$$|\vec{u} \cdot (\vec{v} \times \vec{w})| = |\langle 1, 1, 0 \rangle \cdot \langle 1, 1, -1 \rangle| = 2.$$

So the volume of the parallelepiped is 2 cubic units. In terms of determinants, we have

$$\begin{vmatrix} 1 & 1 & 0 \\ -1 & 1 & 0 \\ 0 & 1 & 1 \end{vmatrix} = \begin{vmatrix} 1 & 0 \\ 1 & 1 \end{vmatrix} 1 - \begin{vmatrix} -1 & 0 \\ 0 & 1 \end{vmatrix} 1 + \begin{vmatrix} -1 & 1 \\ 0 & 1 \end{vmatrix} 0 (1-0)1 - (-1-0)1 = 1+1 = 2,$$

and the absolute value of this determinant is again 2.

While this application of the Triple Scalar Product is interesting, it is not used all that often: parallelepipeds are not a common shape in physics and engineering. The last application of the cross product is very applicable in engineering.
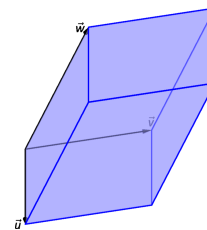
Figure 11.44: A parallelepiped is the three dimensional analogue to the parallelogram.
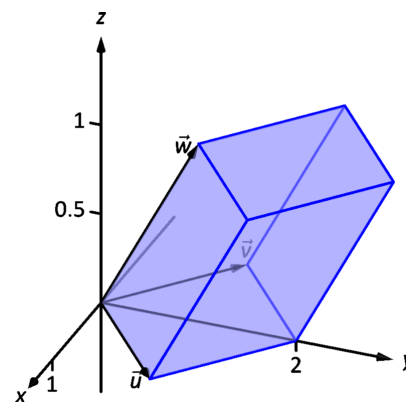
Figure 11.45: A parallelepiped in Example 6.

Notes:

Torque

Torque is a measure of the turning force applied to an object. A classic scenario involving torque is the application of a wrench to a bolt. When a force is applied to the wrench, the bolt turns. When we represent the force and wrench with vectors $\vec{F}$ and $\vec{\ell}$, we see that the bolt moves (because of the threads) in a direction orthogonal to $\vec{F}$ and $\vec{\ell}$. Torque is usually represented by the Greek letter τ, or tau, and has units of N·m, a Newton–meter, or ft·lb, a foot–pound.

   While a full understanding of torque is beyond the purposes of this book, when a force $\vec{F}$ is applied to a lever arm $\vec{\ell}$, the resulting torque is

$$\vec{\tau} = \vec{\ell} \times \vec{F}. \tag{11.7}$$

Example 7　　　**Computing torque**

A lever of length 2ft makes an angle with the horizontal of $45°$. Find the resulting torque when a force of 10lb is applied to the end of the level where:

1. the force is perpendicular to the lever, and

2. the force makes an angle of $60°$ with the lever, as shown in Figure 11.46.

SOLUTION

1. We start by determining vectors for the force and lever arm. Since the lever arm makes a $45°$ angle with the horizontal and is 2ft long, we can state that $\vec{\ell} = 2 \langle \cos 45°, \sin 45° \rangle = \langle \sqrt{2}, \sqrt{2} \rangle$.

   Since the force vector is perpendicular to the lever arm (as seen in the left hand side of Figure 11.46), we can conclude it is making an angle of $-45°$ with the horizontal. As it has a magnitude of 10lb, we can state $\vec{F} = 10 \langle \cos(-45°), \sin(-45°) \rangle = \langle 5\sqrt{2}, -5\sqrt{2} \rangle$.

   Using Equation (11.7) to find the torque requires a cross product. We again let the third component of each vector be 0 and compute the cross product:

$$\vec{\tau} = \vec{\ell} \times \vec{F}$$
$$= \left\langle \sqrt{2}, \sqrt{2}, 0 \right\rangle \times \left\langle 5\sqrt{2}, -5\sqrt{2}, 0 \right\rangle$$
$$= \langle 0, 0, -20 \rangle$$

   This clearly has a magnitude of 20 ft-lb.

   We can view the force and lever arm vectors as lying "on the page"; our computation of $\vec{\tau}$ shows that the torque goes "into the page." This follows the Right Hand Rule of the cross product, and it also matches well with the example of the wrench turning the bolt. Turning a bolt clockwise moves it in.
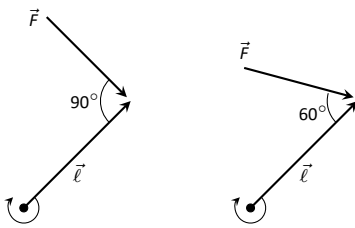
Figure 11.46: Showing a force being applied to a lever in Example 7.

Notes:

2. Our lever arm can still be represented by $\vec{\ell} = \langle \sqrt{2}, \sqrt{2} \rangle$. As our force vector makes a $60°$ angle with $\vec{\ell}$, we can see (referencing the right hand side of the figure) that $\vec{F}$ makes a $-15°$ angle with the horizontal. Thus

$$\vec{F} = 10 \langle \cos -15°, \sin -15° \rangle = \left\langle \frac{5(1+\sqrt{3})}{\sqrt{2}}, \frac{5(1-\sqrt{3})}{\sqrt{2}} \right\rangle.$$

We again make the third component 0 and take the cross product to find the torque:

$$\vec{\tau} = \vec{\ell} \times \vec{F}$$

$$= \left\langle \sqrt{2}, \sqrt{2}, 0 \right\rangle \times \left\langle \frac{5(1+\sqrt{3})}{\sqrt{2}}, \frac{5(1-\sqrt{3})}{\sqrt{2}}, 0 \right\rangle$$

$$= \left\langle 0, 0, -10\sqrt{3} \right\rangle.$$

As one might expect, when the force and lever arm vectors *are* orthogonal, the magnitude of force is greater than when the vectors *are not* orthogonal.

While the cross product has a variety of applications (as noted in this chapter), its fundamental use is finding a vector perpendicular to two others. Knowing a vector is orthogonal to two others is of incredible importance, as it allows us to find the equations of lines and planes in a variety of contexts. The importance of the cross product, in some sense, relies on the importance of lines and planes, which see widespread use throughout engineering, physics and mathematics. We study lines and planes in the next two sections.

Notes:

Exercises 11.4

Terms and Concepts

1. The cross product of two vectors is a _____, not a scalar.
2. One can visualize the direction of $\vec{u} \times \vec{v}$ using the
   _____ _____ _____.
3. Give a synonym for "orthogonal."
4. T/F: A fundamental principle of the cross product is that $\vec{u} \times \vec{v}$ is orthogonal to $\vec{u}$ and $\vec{v}$.
5. _____ is a measure of the turning force applied to an object.

Problems

6. State whether each expression is meaningful. If not, explain why. If so, state whether it is a vector or a scalar.
   (a) $\vec{a} \cdot (\vec{b} \times \vec{c})$
   (b) $\vec{a} \times (\vec{b} \times \vec{c})$
   (c) $(\vec{a} \cdot \vec{b}) \times (\vec{c} \cdot \vec{d})$
   (d) $\vec{a} \times (\vec{b} \cdot \vec{c})$
   (e) $(\vec{a} \times \vec{b})(\vec{c} \times \vec{d})$
   (f) $(\vec{a} \times \vec{b}) \cdot (\vec{c} \times \vec{d})$

In Exercises 7–15, vectors $\vec{u}$ and $\vec{v}$ are given. Compute $\vec{u} \times \vec{v}$ and show this is orthogonal to both $\vec{u}$ and $\vec{v}$.

7. $\vec{u} = \langle 3, 2, -2 \rangle$, $\vec{v} = \langle 0, 1, 5 \rangle$
8. $\vec{u} = \langle 5, -4, 3 \rangle$, $\vec{v} = \langle 2, -5, 1 \rangle$
9. $\vec{u} = \langle 4, -5, -5 \rangle$, $\vec{v} = \langle 3, 3, 4 \rangle$
10. $\vec{u} = \langle -4, 7, -10 \rangle$, $\vec{v} = \langle 4, 4, 1 \rangle$
11. $\vec{u} = \langle 1, 0, 1 \rangle$, $\vec{v} = \langle 5, 0, 7 \rangle$
12. $\vec{u} = \langle 1, 5, -4 \rangle$, $\vec{v} = \langle -2, -10, 8 \rangle$
13. $\vec{u} = \vec{\imath}$, $\vec{v} = \vec{\jmath}$
14. $\vec{u} = \vec{\imath}$, $\vec{v} = \vec{k}$
15. $\vec{u} = \vec{\jmath}$, $\vec{v} = \vec{k}$

16. Pick any vectors $\vec{u}$, $\vec{v}$ and $\vec{w}$ in $\mathbb{R}^3$ and show that $\vec{u} \times (\vec{v} + \vec{w}) = \vec{u} \times \vec{v} + \vec{u} \times \vec{w}$.

17. Pick any vectors $\vec{u}$, $\vec{v}$ and $\vec{w}$ in $\mathbb{R}^3$ and show that $\vec{u} \cdot (\vec{v} \times \vec{w}) = (\vec{u} \times \vec{v}) \cdot \vec{w}$.

In Exercises 18–21, the magnitudes of vectors $\vec{u}$ and $\vec{v}$ in $\mathbb{R}^3$ are given, along with the angle θ between them. Use this information to find the magnitude of $\vec{u} \times \vec{v}$.

18. $\|\vec{u}\| = 2$, $\|\vec{v}\| = 5$, $\theta = 30°$
19. $\|\vec{u}\| = 3$, $\|\vec{v}\| = 7$, $\theta = \pi/2$
20. $\|\vec{u}\| = 3$, $\|\vec{v}\| = 4$, $\theta = \pi$
21. $\|\vec{u}\| = 2$, $\|\vec{v}\| = 5$, $\theta = 5\pi/6$

In Exercises 22–25, find the area of the parallelogram defined by the given vectors.

22. $\vec{u} = \langle 1, 1, 2 \rangle$, $\vec{v} = \langle 2, 0, 3 \rangle$
23. $\vec{u} = \langle -2, 1, 5 \rangle$, $\vec{v} = \langle -1, 3, 1 \rangle$
24. $\vec{u} = \langle 1, 2 \rangle$, $\vec{v} = \langle 2, 1 \rangle$
25. $\vec{u} = \langle 2, 0 \rangle$, $\vec{v} = \langle 0, 3 \rangle$

In Exercises 26–29, find the area of the triangle with the given vertices.

26. Vertices: $(0, 0, 0)$, $(1, 3, -1)$ and $(2, 1, 1)$.
27. Vertices: $(5, 2, -1)$, $(3, 6, 2)$ and $(1, 0, 4)$.
28. Vertices: $(1, 1)$, $(1, 3)$ and $(2, 2)$.
29. Vertices: $(3, 1)$, $(1, 2)$ and $(4, 3)$.

In Exercises 30–31, find the area of the quadrilateral with the given vertices. (Hint: break the quadrilateral into 2 triangles.)

30. Vertices: $(0, 0)$, $(1, 2)$, $(3, 0)$ and $(4, 3)$.
31. Vertices: $(0, 0, 0)$, $(2, 1, 1)$, $(-1, 2, -8)$ and $(1, -1, 5)$.

In Exercises 32–33, find the volume of the parallelepiped defined by the given vectors.

32. $\vec{u} = \langle 1, 1, 1 \rangle$, $\vec{v} = \langle 1, 2, 3 \rangle$, $\vec{w} = \langle 1, 0, 1 \rangle$
33. $\vec{u} = \langle -1, 2, 1 \rangle$, $\vec{v} = \langle 2, 2, 1 \rangle$, $\vec{w} = \langle 3, 1, 3 \rangle$

In Exercises 34–37, find a unit vector orthogonal to both $\vec{u}$ and $\vec{v}$.

34. $\vec{u} = \langle 1, 1, 1 \rangle$, $\vec{v} = \langle 2, 0, 1 \rangle$
35. $\vec{u} = \langle 1, -2, 1 \rangle$, $\vec{v} = \langle 3, 2, 1 \rangle$
36. $\vec{u} = \langle 5, 0, 2 \rangle$, $\vec{v} = \langle -3, 0, 7 \rangle$
37. $\vec{u} = \langle 1, -2, 1 \rangle$, $\vec{v} = \langle -2, 4, -2 \rangle$

38. A bicycle rider applies 150lb of force, straight down, onto a pedal that extends 7in horizontally from the crankshaft. Find the magnitude of the torque applied to the crankshaft.

39. A bicycle rider applies 150lb of force, straight down, onto a pedal that extends 7in from the crankshaft, making a $30°$ angle with the horizontal. Find the magnitude of the torque applied to the crankshaft.

40. To turn a stubborn bolt, 80lb of force is applied to a 10in wrench. What is the maximum amount of torque that can be applied to the bolt?

41. To turn a stubborn bolt, 80lb of force is applied to a 10in wrench in a confined space, where the direction of applied force makes a $10°$ angle with the wrench. How much torque is subsequently applied to the wrench?

42. Show, using the definition of the Cross Product, that $\vec{u} \cdot (\vec{u} \times \vec{v}) = 0$; that is, that $\vec{u}$ is orthogonal to the cross product of $\vec{u}$ and $\vec{v}$.

43. Show, using the definition of the Cross Product, that $\vec{u} \times \vec{u} = \vec{0}$.

11.5   Lines

To find the equation of a line in the *x-y* plane, we need two pieces of information: a point and the slope. The slope conveys *direction* information. As vertical lines have an undefined slope, the following statement is more accurate:

> To define a line, one needs a point on the line and the direction of the line.

This holds true for lines in space.

Let *P* be a point in space, let $\vec{p}$ be the vector with initial point at the origin and terminal point at *P* (i.e., $\vec{p}$ "points" to *P*), and let $\vec{d}$ be a vector. Consider the points on the line through *P* in the direction of $\vec{d}$.

Clearly one point on the line is *P*; we can say that the *vector* $\vec{p}$ lies at this point on the line. To find another point on the line, we can start at $\vec{p}$ and move in a direction parallel to $\vec{d}$. For instance, starting at $\vec{p}$ and traveling one length of $\vec{d}$ places one at another point on the line. Consider Figure 11.47 where certain points along the line are indicated.

The figure illustrates how every point on the line can be obtained by starting with $\vec{p}$ and moving a certain distance in the direction of $\vec{d}$. That is, we can define the line as a function of *t*:

$$\vec{\ell}(t) = \vec{p} + t\,\vec{d}. \tag{11.8}$$

In many ways, this is *not* a new concept. Compare Equation (11.8) to the familiar "$y = mx + b$" equation of a line:
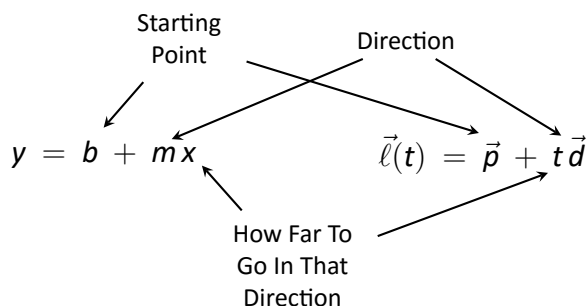
Figure 11.48: Understanding the vector equation of a line.

Figure 11.47: Defining a line in space.

The equations exhibit the same structure: they give a starting point, define a direction, and state how far in that direction to travel.

Notes:

Equation (11.8) is an example of a **vector–valued function**; the input of the function is a real number and the output is a vector. We will cover vector–valued functions extensively in the next chapter.

There are other ways to represent a line. Let $\vec{p} = \langle x_0, y_0, z_0 \rangle$ and let $\vec{d} = \langle a, b, c \rangle$. Then the equation of the line through $\vec{p}$ in the direction of $\vec{d}$ is:

$$
\begin{aligned}
\vec{\ell}(t) &= \vec{p} + t\vec{d} \\
&= \langle x_0, y_0, z_0 \rangle + t \langle a, b, c \rangle \\
&= \langle x_0 + at, y_0 + bt, z_0 + ct \rangle .
\end{aligned}
$$

The last line states the the x values of the line are given by $x = x_0 + at$, the y values are given by $y = y_0 + bt$, and the z values are given by $z = z_0 + ct$. These three equations, taken together, are the **parametric equations of the line** through $\vec{p}$ in the direction of $\vec{d}$.

Finally, each of the equations for x, y and z above contain the variable t. We can solve for t in each equation:

$$
\begin{aligned}
x = x_0 + at &\quad\Rightarrow\quad t = \frac{x - x_0}{a}, \\
y = y_0 + bt &\quad\Rightarrow\quad t = \frac{y - y_0}{b}, \\
z = z_0 + ct &\quad\Rightarrow\quad t = \frac{z - z_0}{c},
\end{aligned}
$$

assuming $a, b, c \neq 0$. Since t is equal to each expression on the right, we can set these equal to each other, forming the **symmetric equations of the line** through $\vec{p}$ in the direction of $\vec{d}$:

$$
\frac{x - x_0}{a} = \frac{y - y_0}{b} = \frac{z - z_0}{c}.
$$

Each representation has its own advantages, depending on the context. We summarize these three forms in the following definition, then give examples of their use.

Notes:

> **Definition 65**   **Equations of Lines in Space**
> Consider the line in space that passes through $\vec{p} = \langle x_0, y_0, z_0 \rangle$ in the direction of $\vec{d} = \langle a, b, c \rangle$.
>
> 1. The **vector equation** of the line is
> $$\vec{\ell}(t) = \vec{p} + t\vec{d}.$$
>
> 2. The **parametric equations** of the line are
> $$x = x_0 + at, \quad y = y_0 + bt, \quad z = z_0 + ct.$$
>
> 3. The **symmetric equations** of the line are
> $$\frac{x - x_0}{a} = \frac{y - y_0}{b} = \frac{z - z_0}{c}.$$

Watch the video:
Example of Symmetric Equations of a Line at
`https://youtu.be/q4wDcrCkkfQ`

Example 1   **Finding the equation of a line**
Give all three equations, as given in Definition 65, of the line through $P = (2, 3, 1)$ in the direction of $\vec{d} = \langle -1, 1, 2 \rangle$. Does the point $Q = (-1, 6, 6)$ lie on this line?

SOLUTION   We identify the point $P = (2, 3, 1)$ with the vector $\vec{p} = \langle 2, 3, 1 \rangle$. Following the definition, we have

- the vector equation of the line is $\vec{\ell}(t) = \langle 2, 3, 1 \rangle + t \langle -1, 1, 2 \rangle$;

- the parametric equations of the line are
$$x = 2 - t, \quad y = 3 + t, \quad z = 1 + 2t; \text{ and}$$

- the symmetric equations of the line are
$$\frac{x - 2}{-1} = \frac{y - 3}{1} = \frac{z - 1}{2}.$$

Notes:
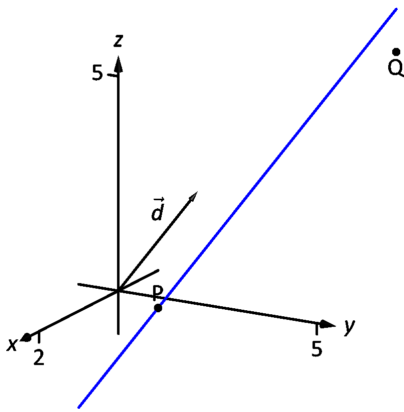
Figure 11.49: Graphing a line in Example 1.

The first two equations of the line are useful when a t value is given: one can immediately find the corresponding point on the line. These forms are good when calculating with a computer; most software programs easily handle equations in these formats.

Does the point $Q = (-1, 6, 6)$ lie on the line? The graph in Figure 11.49 makes it clear that it does not. We can answer this question without the graph using any of the three equation forms. Of the three, the symmetric equations are probably best suited for this task. Simply plug in the values of x, y and z and see if equality is maintained:

$$\frac{-1-2}{-1} \overset{?}{=} \frac{6-3}{1} \overset{?}{=} \frac{6-1}{2} \quad \Rightarrow \quad 3 = 3 \neq 2.5.$$

We see that Q does not lie on the line as it did not satisfy the symmetric equations.

Example 2　　　**Finding the equation of a line through two points**
Find the parametric equations of the line through the points $P = (2, -1, 2)$ and $Q = (1, 3, -1)$.

SOLUTION　　　Recall the statement made at the beginning of this section: to find the equation of a line, we need a point and a direction. We have *two* points; either one will suffice. The direction of the line can be found by the vector with initial point P and terminal point Q: $\overrightarrow{PQ} = \langle -1, 4, -3 \rangle$.

The parametric equations of the line ℓ through P in the direction of $\overrightarrow{PQ}$ are:

$$\ell: \quad x = 2 - t \quad y = -1 + 4t \quad z = 2 - 3t.$$

A graph of the points and line are given in Figure 11.50. Note how in the given parametrization of the line, $t = 0$ corresponds to the point P, and $t = 1$ corresponds to the point Q. This relates to the understanding of the vector equation of a line described in Figure 11.48. The parametric equations "start" at the point P, and t determines how far in the direction of $\overrightarrow{PQ}$ to travel. When $t = 0$, we travel 0 lengths of $\overrightarrow{PQ}$; when $t = 1$, we travel one length of $\overrightarrow{PQ}$, resulting in the point Q.
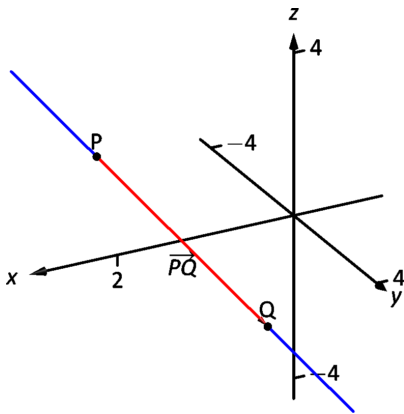
Figure 11.50: A graph of the line in Example 2.

Parallel, Intersecting and Skew Lines

In the plane, two *distinct* lines can either be parallel or they will intersect at exactly one point. In space, given equations of two lines, it can sometimes be difficult to tell whether the lines are distinct or not (i.e., the same line can be represented in different ways). Given lines $\vec{\ell_1}(t) = \vec{p_1} + t\vec{d_1}$ and $\vec{\ell_2}(t) = \vec{p_2} + t\vec{d_2}$, we have four possibilities: $\vec{\ell_1}$ and $\vec{\ell_2}$ are

---

Notes:

the same line	they share all points;
intersecting lines	share only 1 point;
parallel lines	$\vec{d}_1 \parallel \vec{d}_2$, no points in common; or
skew lines	$\vec{d}_1 \nparallel \vec{d}_2$, no points in common.

The next two examples investigate these possibilities.

Example 3   **Comparing lines**

Consider lines ℓ_1 and ℓ_2, given in parametric equation form:

$$\ell_1 : \begin{array}{rcl} x &=& 1 + 3t \\ y &=& 2 - t \\ z &=& t \end{array} \qquad \ell_2 : \begin{array}{rcl} x &=& -2 + 4s \\ y &=& 3 + s \\ z &=& 5 + 2s. \end{array}$$

Determine whether ℓ_1 and ℓ_2 are the same line, intersect, are parallel, or skew.

SOLUTION   We start by looking at the directions of each line. Line ℓ_1 has the direction given by $\vec{d}_1 = \langle 3, -1, 1 \rangle$ and line ℓ_2 has the direction given by $\vec{d}_2 = \langle 4, 1, 2 \rangle$. It should be clear that $\vec{d}_1$ and $\vec{d}_2$ are not parallel, hence ℓ_1 and ℓ_2 are not the same line, nor are they parallel. Figure 11.51 verifies this fact (where the points and directions indicated by the equations of each line are identified).

We next check to see if they intersect (if they do not, they are skew lines). To find if they intersect, we look for t and s values such that the respective x, y and z values are the same. That is, we want s and t such that:

$$\begin{array}{rcl} 1 + 3t &=& -2 + 4s \\ 2 - t &=& 3 + s \\ t &=& 5 + 2s. \end{array}$$

This is a relatively simple system of linear equations. Since the last equation is already solved for t, substitute that value of t into the equation above it:

$$2 - (5 + 2s) = 3 + s \quad \Rightarrow \quad s = -2, \ t = 1.$$

A key to remember is that we have *three* equations; we need to check if $s = -2$, $t = 1$ satisfies the first equation as well:

$$1 + 3(1) \neq -2 + 4(-2).$$

It does not. Therefore, we conclude that the lines ℓ_1 and ℓ_2 are skew.
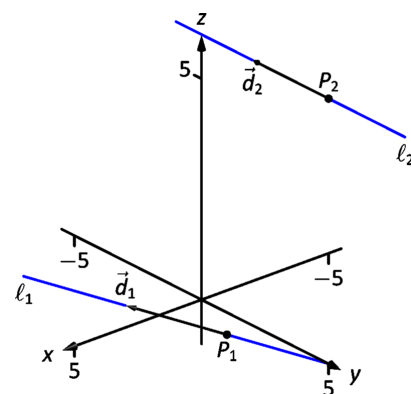
Figure 11.51: Sketching the lines from Example 3.

Example 4   **Comparing lines**

Consider lines ℓ_1 and ℓ_2, given in parametric equation form:

$$\ell_1 : \begin{array}{rcl} x &=& -0.7 + 1.6t \\ y &=& 4.2 + 2.72t \\ z &=& 2.3 - 3.36t \end{array} \qquad \ell_2 : \begin{array}{rcl} x &=& 2.8 - 2.9s \\ y &=& 10.15 - 4.93s \\ z &=& -5.05 + 6.09s. \end{array}$$

Determine whether ℓ_1 and ℓ_2 are the same line, intersect, are parallel, or skew.

Notes:

SOLUTION   It is obviously very difficult to simply look at these equations and discern anything. This is done intentionally. In the "real world," most equations that are used do not have nice, integer coefficients.

We again start by deciding whether or not each line has the same direction. The direction of ℓ_1 is given by $\vec{d}_1 = \langle 1.6, 2.72, -3.36 \rangle$ and the direction of ℓ_2 is given by $\vec{d}_2 = \langle -2.9, -4.93, 6.09 \rangle$. When it is not clear through observation whether two vectors are parallel or not, the standard way of determining this is by comparing their respective unit vectors. Using a calculator, we find:

$$\vec{u}_1 = \frac{\vec{d}_1}{\left\| \vec{d}_1 \right\|} = \langle 0.3471, 0.5901, -0.7289 \rangle$$

$$\vec{u}_2 = \frac{\vec{d}_2}{\left\| \vec{d}_2 \right\|} = \langle -0.3471, -0.5901, 0.7289 \rangle .$$

The two vectors seem to be parallel (at least, their components are equal to 4 decimal places). In most situations, it would suffice to conclude that the lines are at least parallel, if not the same. One way to be sure is to rewrite $\vec{d}_1$ and $\vec{d}_2$ in terms of fractions, not decimals. We have

$$\vec{d}_1 = \left\langle \frac{16}{10}, \frac{272}{100}, -\frac{336}{100} \right\rangle \qquad \vec{d}_2 = \left\langle -\frac{29}{10}, -\frac{493}{100}, \frac{609}{100} \right\rangle .$$

One can then find the magnitudes of each vector in terms of fractions, then compute the unit vectors likewise. After a lot of manual arithmetic (or after briefly using a computer algebra system), one finds that

$$\vec{u}_1 = \left\langle \sqrt{\frac{10}{83}}, \frac{17}{\sqrt{830}}, -\frac{21}{\sqrt{830}} \right\rangle \qquad \vec{u}_2 = \left\langle -\sqrt{\frac{10}{83}}, -\frac{17}{\sqrt{830}}, \frac{21}{\sqrt{830}} \right\rangle .$$

We can now say without equivocation that these lines are parallel.

Are they the same line? The parametric equations for a line describe one point that lies on the line, so we know that the point $P_1 = (-0.7, 4.2, 2.3)$ lies on ℓ_1. To determine if this point also lies on ℓ_2, plug in the x, y and z values of P_1 into the symmetric equations for ℓ_2:

$$\frac{(-0.7) - 2.8}{-2.9} \stackrel{?}{=} \frac{(4.2) - 10.15}{-4.93} \stackrel{?}{=} \frac{(2.3) - (-5.05)}{6.09}$$

$$\Rightarrow \quad 1.2069 = 1.2069 = 1.2069 .$$

The point P_1 lies on both lines, so we conclude they are the same line, just parametrized differently. Figure 11.52 graphs this line along with the points and vectors described by the parametric equations. Note how $\vec{d}_1$ and $\vec{d}_2$ are parallel, though point in opposite directions (as indicated by their unit vectors above).
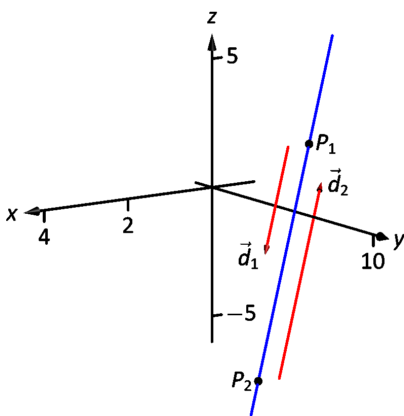
Figure 11.52: Graphing the lines in Example 4.

Notes:

Distances

Given a point Q and a line $\vec{\ell}(t) = \vec{p} + t\vec{d}$ in space, it is often useful to know the distance from the point to the line. (Here we use the standard definition of "distance," i.e., the length of the shortest line segment from the point to the line.) Identifying $\vec{p}$ with the point P, Figure 11.53 will help establish a general method of computing this distance h.

From trigonometry, we know $h = \left\| \overrightarrow{PQ} \right\| \sin \theta$. We have a similar identity involving the cross product: $\left\| \overrightarrow{PQ} \times \vec{d} \right\| = \left\| \overrightarrow{PQ} \right\| \left\| \vec{d} \right\| \sin \theta$. Divide both sides of this latter equation by $\left\| \vec{d} \right\|$ to obtain h:

$$h = \frac{\left\| \overrightarrow{PQ} \times \vec{d} \right\|}{\left\| \vec{d} \right\|}. \tag{11.9}$$

Figure 11.53: Establishing the distance from a point to a line.

It is also useful to determine the distance between lines, which we define as the length of the shortest line segment that connects the two lines (an argument from geometry shows that this line segments is perpendicular to both lines). Let lines $\vec{\ell}_1(t) = \vec{p}_1 + t\vec{d}_1$ and $\vec{\ell}_2(t) = \vec{p}_2 + t\vec{d}_2$ be given, as shown in Figure 11.54. To find the direction orthogonal to both $\vec{d}_1$ and $\vec{d}_2$, we take the cross product: $\vec{c} = \vec{d}_1 \times \vec{d}_2$. The magnitude of the orthogonal projection of $\overrightarrow{P_1P_2}$ onto $\vec{c}$ is the distance h we seek:

$$\begin{aligned} h &= \left\| \operatorname{proj}_{\vec{c}} \overrightarrow{P_1P_2} \right\| \\ &= \left\| \frac{\overrightarrow{P_1P_2} \cdot \vec{c}}{\vec{c} \cdot \vec{c}} \vec{c} \right\| \\ &= \frac{\left| \overrightarrow{P_1P_2} \cdot \vec{c} \right|}{\|\vec{c}\|^2} \|\vec{c}\| \\ &= \frac{\left| \overrightarrow{P_1P_2} \cdot \vec{c} \right|}{\|\vec{c}\|}. \end{aligned}$$

Figure 11.54: Establishing the distance between lines.

Exercise 30 shows that this distance is 0 when the lines intersect. Note the use of the Triple Scalar Product: $\overrightarrow{P_1P_2} \cdot \vec{c} = \overrightarrow{P_1P_2} \cdot (\vec{d}_1 \times \vec{d}_2)$.

The following Key Idea restates these two distance formulas.

Notes:

Key Idea 53   Distances to Lines

1. Let P be a point on a line ℓ that is parallel to $\vec{d}$. The distance h from a point Q to the line ℓ is:

$$h = \frac{\left\| \overrightarrow{PQ} \times \vec{d} \right\|}{\left\| \vec{d} \right\|}.$$

2. Let P_1 be a point on line ℓ_1 that is parallel to $\vec{d}_1$, and let P_2 be a point on line ℓ_2 parallel to $\vec{d}_2$, and let $\vec{c} = \vec{d}_1 \times \vec{d}_2$, where lines ℓ_1 and ℓ_2 are not parallel. The distance h between the two lines is:

$$h = \frac{\left| \overrightarrow{P_1 P_2} \cdot \vec{c} \right|}{\left\| \vec{c} \right\|}.$$

Example 5   Finding the distance from a point to a line
Find the distance from the point $Q = (1, 1, 3)$ to the line $\vec{\ell}(t) = \langle 1, -1, 1 \rangle + t \langle 2, 3, 1 \rangle$.

SOLUTION   The equation of the line gives us the point $P = (1, -1, 1)$ that lies on the line, hence $\overrightarrow{PQ} = \langle 0, 2, 2 \rangle$. The equation also gives $\vec{d} = \langle 2, 3, 1 \rangle$. Following Key Idea 53, we have the distance as

$$h = \frac{\left\| \overrightarrow{PQ} \times \vec{d} \right\|}{\left\| \vec{d} \right\|}$$

$$= \frac{\left\| \langle -4, 4, -4 \rangle \right\|}{\sqrt{14}}$$

$$= \frac{4\sqrt{3}}{\sqrt{14}}.$$

The point Q is approximately 1.852 units from the line $\vec{\ell}(t)$.

Example 6   Finding the distance between lines
Find the distance between the lines

$$\ell_1 : \begin{array}{rcl} x &=& 1 + 3t \\ y &=& 2 - t \\ z &=& t \end{array} \qquad \ell_2 : \begin{array}{rcl} x &=& -2 + 4s \\ y &=& 3 + s \\ z &=& 5 + 2s. \end{array}$$

Notes:

SOLUTION   These are the sames lines as given in Example 3, where we showed them to be skew. The equations allow us to identify the following points and vectors:

$$P_1 = (1, 2, 0) \quad P_2 = (-2, 3, 5) \quad \Rightarrow \quad \overrightarrow{P_1 P_2} = \langle -3, 1, 5 \rangle .$$

$$\vec{d_1} = \langle 3, -1, 1 \rangle \quad \vec{d_2} = \langle 4, 1, 2 \rangle \quad \Rightarrow \quad \vec{c} = \vec{d_1} \times \vec{d_2} = \langle -3, -2, 7 \rangle .$$

From Key Idea 53 we have the distance h between the two lines is

$$h = \frac{\left| \overrightarrow{P_1 P_2} \cdot \vec{c} \right|}{\| \vec{c} \|}$$

$$= \frac{42}{\sqrt{62}} .$$

The lines are approximately 5.334 units apart.

One of the key points to understand from this section is this: to describe a line, we need a point and a direction. Whenever a problem is posed concerning a line, one needs to take whatever information is offered and glean point and direction information. Many questions can be asked (and *are* asked in the exercises) whose answer immediately follows from this understanding.

Lines are one of two fundamental objects of study in space. The other fundamental object is the *plane*, which we study in detail in the next section. Many complex three dimensional objects are studied by approximating their surfaces with lines and planes.

Notes:

Exercises 11.5

Terms and Concepts

1. To find an equation of a line, what two pieces of information are needed?
2. Two distinct lines in the plane can intersect or be _____.
3. Two distinct lines in space can intersect, be _____ or be _____.
4. Use your own words to describe what it means for two lines in space to be skew.

Problems

In Exercises 5–14, write the vector, parametric and symmetric equations of the lines described.

5. Passes through $P = (2, -4, 1)$, parallel to $\vec{d} = \langle 9, 2, 5 \rangle$.

6. Passes through $P = (6, 1, 7)$, parallel to $\vec{d} = \langle -3, 2, 5 \rangle$.

7. Passes through $P = (2, 1, 5)$ and $Q = (7, -2, 4)$.

8. Passes through $P = (1, -2, 3)$ and $Q = (5, 5, 5)$.

9. Passes through $P = (0, 1, 2)$ and orthogonal to both $\vec{d_1} = \langle 2, -1, 7 \rangle$ and $\vec{d_2} = \langle 7, 1, 3 \rangle$.

10. Passes through $P = (5, 1, 9)$ and orthogonal to both $\vec{d_1} = \langle 1, 0, 1 \rangle$ and $\vec{d_2} = \langle 2, 0, 3 \rangle$.

11. Passes through the point of intersection of $\vec{\ell_1}(t)$ and $\vec{\ell_2}(t)$ and orthogonal to both lines, where $\vec{\ell_1}(t) = \langle 2, 1, 1 \rangle + t \langle 5, 1, -2 \rangle$ and $\vec{\ell_2}(t) = \langle -2, -1, 2 \rangle + t \langle 3, 1, -1 \rangle$.

12. Passes through the point of intersection of $\ell_1(t)$ and $\ell_2(t)$ and orthogonal to both lines, where
$$\ell_1 = \begin{cases} x = t \\ y = -2 + 2t \\ z = 1 + t \end{cases} \quad \text{and} \quad \ell_2 = \begin{cases} x = 2 + t \\ y = 2 - t \\ z = 3 + 2t \end{cases}.$$

13. Passes through $P = (1, 1)$, parallel to $\vec{d} = \langle 2, 3 \rangle$.

14. Passes through $P = (-2, 5)$, parallel to $\vec{d} = \langle 0, 1 \rangle$.

In Exercises 15–22, determine if the described lines are the same line, parallel lines, intersecting or skew lines. If intersecting, give the point of intersection.

15. $\vec{\ell_1}(t) = \langle 1, 2, 1 \rangle + t \langle 2, -1, 1 \rangle$,
$\vec{\ell_2}(t) = \langle 3, 3, 3 \rangle + t \langle -4, 2, -2 \rangle$.

16. $\vec{\ell_1}(t) = \langle 2, 1, 1 \rangle + t \langle 5, 1, 3 \rangle$,
$\vec{\ell_2}(t) = \langle 14, 5, 9 \rangle + t \langle 1, 1, 1 \rangle$.

17. $\vec{\ell_1}(t) = \langle 3, 4, 1 \rangle + t \langle 2, -3, 4 \rangle$,
$\vec{\ell_2}(t) = \langle -3, 3, -3 \rangle + t \langle 3, -2, 4 \rangle$.

18. $\vec{\ell_1}(t) = \langle 1, 1, 1 \rangle + t \langle 3, 1, 3 \rangle$,
$\vec{\ell_2}(t) = \langle 7, 3, 7 \rangle + t \langle 6, 2, 6 \rangle$.

19. $\ell_1 = \begin{cases} x = 1 + 2t \\ y = 3 - 2t \\ z = t \end{cases} \quad \text{and} \quad \ell_2 = \begin{cases} x = 3 - t \\ y = 3 + 5t \\ z = 2 + 7t \end{cases}$

20. $\ell_1 = \begin{cases} x = 1.1 + 0.6t \\ y = 3.77 + 0.9t \\ z = -2.3 + 1.5t \end{cases} \quad \text{and}$
$\ell_2 = \begin{cases} x = 3.11 + 3.4t \\ y = 2 + 5.1t \\ z = 2.5 + 8.5t \end{cases}$

21. $\ell_1 = \begin{cases} x = 0.2 + 0.6t \\ y = 1.33 - 0.45t \\ z = -4.2 + 1.05t \end{cases} \quad \text{and}$
$\ell_2 = \begin{cases} x = 0.86 + 9.2t \\ y = 0.835 - 6.9t \\ z = -3.045 + 16.1t \end{cases}$

22. $\ell_1 = \begin{cases} x = 0.1 + 1.1t \\ y = 2.9 - 1.5t \\ z = 3.2 + 1.6t \end{cases} \quad \text{and}$
$\ell_2 = \begin{cases} x = 4 - 2.1t \\ y = 1.8 + 7.2t \\ z = 3.1 + 1.1t \end{cases}$

In Exercises 23–26, find the distance from the point to the line.

23. $P = (1, 1, 1), \quad \vec{\ell}(t) = \langle 2, 1, 3 \rangle + t \langle 2, 1, -2 \rangle$

24. $P = (2, 5, 6), \quad \vec{\ell}(t) = \langle -1, 1, 1 \rangle + t \langle 1, 0, 1 \rangle$

25. $P = (0, 3), \quad \vec{\ell}(t) = \langle 2, 0 \rangle + t \langle 1, 1 \rangle$

26. $P = (1, 1), \quad \vec{\ell}(t) = \langle 4, 5 \rangle + t \langle -4, 3 \rangle$

In Exercises 27–28, find the distance between the two lines.

27. $\vec{\ell_1}(t) = \langle 1, 2, 1 \rangle + t \langle 2, -1, 1 \rangle$,
$\vec{\ell_2}(t) = \langle 3, 3, 3 \rangle + t \langle 4, 2, -2 \rangle$.

28. $\vec{\ell_1}(t) = \langle 0, 0, 1 \rangle + t \langle 1, 0, 0 \rangle$,
$\vec{\ell_2}(t) = \langle 0, 0, 3 \rangle + t \langle 0, 1, 0 \rangle$.

Exercises 29–31 explore special cases of the distance formulas found in Key Idea 53.

29. Let Q be a point on the line $\vec{\ell}(t)$. Show why the distance formula correctly gives the distance from the point to the line as 0.

30. Let lines $\vec{\ell_1}(t)$ and $\vec{\ell_2}(t)$ be intersecting lines. Show why the distance formula correctly gives the distance between these lines as 0.

31. Let lines $\vec{\ell_1}(t)$ and $\vec{\ell_2}(t)$ be parallel.

    (a) Show why the distance formula for distance between lines cannot be used as stated to find the distance between the lines.

    (b) Show why letting $\vec{c} = (\vec{P_1 P_2} \times \vec{d_2}) \times \vec{d_2}$ allows one to the use the formula.

    (c) Show how one can use the formula for the distance between a point and a line to find the distance between parallel lines.

11.6 Planes

Any flat surface, such as a wall, table top or stiff piece of cardboard can be thought of as representing part of a plane. Consider a piece of cardboard with a point P marked on it. One can take a nail and stick it into the cardboard at P such that the nail is perpendicular to the cardboard; see Figure 11.55.

This nail provides a "handle" for the cardboard. Moving the cardboard around moves P to different locations in space. Tilting the nail (but keeping P fixed) tilts the cardboard. Both moving and tilting the cardboard defines a different plane in space. In fact, we can define a plane by: 1) the location of P in space, and 2) the direction of the nail.

The previous section showed that one can define a line given a point on the line and the direction of the line (usually given by a vector). One can make a similar statement about planes: we can define a plane in space given a point on the plane and the direction the plane "faces" (using the description above, the direction of the nail). Once again, the direction information will be supplied by a vector, called a **normal vector**, that is orthogonal to the plane.

What exactly does "orthogonal to the plane" mean? Choose any two points P and Q in the plane, and consider the vector $\vec{PQ}$. We say a vector $\vec{n}$ is orthogonal to the plane if $\vec{n}$ is perpendicular to $\vec{PQ}$ for all choices of P and Q; that is, if $\vec{n} \cdot \vec{PQ} = 0$ for all P and Q.

This gives us way of writing an equation describing the plane. Let $P = (x_0, y_0, z_0)$ be a point in the plane and let $\vec{n} = \langle a, b, c \rangle$ be a normal vector to the plane. A point $Q = (x, y, z)$ lies in the plane defined by P and $\vec{n}$ if, and only if, $\vec{PQ}$ is orthogonal to $\vec{n}$. Knowing $\vec{PQ} = \langle x - x_0, y - y_0, z - z_0 \rangle$, consider:

$$\vec{PQ} \cdot \vec{n} = 0$$
$$\langle x - x_0, y - y_0, z - z_0 \rangle \cdot \langle a, b, c \rangle = 0$$
$$a(x - x_0) + b(y - y_0) + c(z - z_0) = 0 \tag{11.10}$$

Equation (11.10) defines an *implicit* function describing the plane. More algebra produces:

$$ax + by + cz = ax_0 + by_0 + cz_0.$$

The right hand side is just a number, so we replace it with d:

$$ax + by + cz = d. \tag{11.11}$$

As long as $c \neq 0$, we can solve for z:

$$z = \frac{1}{c}(d - ax - by). \tag{11.12}$$
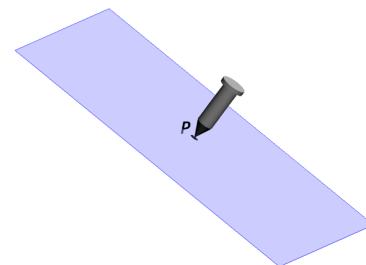
Figure 11.55: Illustrating defining a plane with a sheet of cardboard and a nail.

Notes:

Equation (11.12) is especially useful as many computer programs can graph functions in this form. Equations (11.10) and (11.11) have specific names, given next.

Definition 66   **Equations of a Plane in Standard and General Forms**
The plane passing through the point $P = (x_0, y_0, z_0)$ with normal vector $\vec{n} = \langle a, b, c \rangle$ can be described by an equation with **standard form**

$$a(x - x_0) + b(y - y_0) + c(z - z_0) = 0;$$

the equation's **general form** is

$$ax + by + cz = d;$$

the equation's **vector form** is

$$\langle x, y, z \rangle \cdot \vec{n} = \langle x_0, y_0, z_0 \rangle \cdot \vec{n} = d.$$

A key to remember throughout this section is this: to find the equation of a plane, we need a point and a normal vector. We will give several examples of finding the equation of a plane, and in each one different types of information are given. In each case, we need to use the given information to find a point on the plane and a normal vector.

Watch the video:
Equation of a Plane Through Three Points at
`https://youtu.be/0MECmEjR2WU`

Example 1   **Finding the equation of a plane.**
Write the equation of the plane that passes through the points $P = (1, 1, 0)$, $Q = (1, 2, -1)$ and $R = (0, 1, 2)$ in standard form.

SOLUTION   We need a vector $\vec{n}$ that is orthogonal to the plane. Since P, Q and R are in the plane, so are the vectors $\vec{PQ}$ and $\vec{PR}$; $\vec{PQ} \times \vec{PR}$ is orthogonal to $\vec{PQ}$ and $\vec{PR}$ and hence the plane itself.
It is straightforward to compute $\vec{n} = \vec{PQ} \times \vec{PR} = \langle 2, 1, 1 \rangle$. We can use any point we wish in the plane (any of P, Q or R will do) and we arbitrarily choose P.

Notes:

Figure 11.56: Sketching the plane in Example 1.

Following Definition 66, the equation of the plane in standard form is

$$2(x-1)+(y-1)+z=0.$$

The plane is sketched in Figure 11.56.

We have just demonstrated the fact that any three non-collinear points define a plane. (This is why a three-legged stool does not "rock;" it's three feet always lie in a plane. A four-legged stool will rock unless all four feet lie in the same plane.)

Example 2   Finding the equation of a plane.
Verify that lines ℓ_1 and ℓ_2, whose parametric equations are given below, intersect, then give the equation of the plane that contains these two lines in general form.

$$\ell_1 : \begin{aligned} x &= -5 + 2s \\ y &= 1 + s \\ z &= -4 + 2s \end{aligned} \qquad \ell_2 : \begin{aligned} x &= 2 + 3t \\ y &= 1 - 2t \\ z &= 1 + t \end{aligned}$$

SOLUTION   The lines clearly are not parallel. If they do not intersect, they are skew, meaning there is not a plane that contains them both. If they do intersect, there is such a plane.

To find their point of intersection, we set the *x*, *y* and *z* equations equal to each other and solve for *s* and *t*:

$$\begin{aligned} -5 + 2s &= 2 + 3t \\ 1 + s &= 1 - 2t \\ -4 + 2s &= 1 + t \end{aligned} \quad \Rightarrow \quad s = 2, \quad t = -1.$$

When $s = 2$ and $t = -1$, the lines intersect at the point $P = (-1, 3, 0)$.

Let $\vec{d}_1 = \langle 2, 1, 2 \rangle$ and $\vec{d}_2 = \langle 3, -2, 1 \rangle$ be the directions of lines ℓ_1 and ℓ_2, respectively. A normal vector to the plane containing these the two lines will also be orthogonal to $\vec{d}_1$ and $\vec{d}_2$. Thus we find a normal vector $\vec{n}$ by computing $\vec{n} = \vec{d}_1 \times \vec{d}_2 = \langle 5, 4 - 7 \rangle$.

We can pick any point in the plane with which to write our equation; each line gives us infinite choices of points. We choose *P*, the point of intersection. We follow Definition 66 to write the plane's equation in general form:

$$\begin{aligned} 5(x + 1) + 4(y - 3) - 7z &= 0 \\ 5x + 5 + 4y - 12 - 7z &= 0 \\ 5x + 4y - 7z &= 7. \end{aligned}$$

The plane's equation in general form is $5x + 4y - 7z = 7$; it is sketched in Figure 11.57.
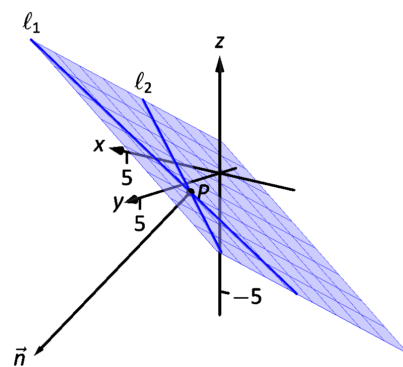
Notes:

Figure 11.57: Sketching the plane in Example 2.

Example 3 **Finding the equation of a plane**

Give the equation, in standard form, of the plane that passes through the point $P = (-1, 0, 1)$ and is orthogonal to the line with vector equation $\vec{\ell}(t) = \langle -1, 0, 1 \rangle + t \langle 1, 2, 2 \rangle$.

SOLUTION As the plane is to be orthogonal to the line, the plane must be orthogonal to the direction of the line given by $\vec{d} = \langle 1, 2, 2 \rangle$. We use this as our normal vector. Thus the plane's equation, in standard form, is

$$(x + 1) + 2y + 2(z - 1) = 0.$$

The line and plane are sketched in Figure 11.58.

Example 4 **Finding the intersection of two planes**

Give the parametric equations of the line that is the intersection of the planes p_1 and p_2, where:

$$p_1 : x - (y - 2) + (z - 1) = 0$$
$$p_2 : -2(x - 2) + (y + 1) + (z - 3) = 0$$

SOLUTION To find an equation of a line, we need a point on the line and the direction of the line.

We can find a point on the line by solving each equation of the planes for z:

$$p_1 : z = -x + y - 1$$
$$p_2 : z = 2x - y - 2$$

We can now set these two equations equal to each other (i.e., we are finding values of x and y where the planes have the same z value):

$$-x + y - 1 = 2x - y - 2$$
$$2y = 3x - 1$$
$$y = \frac{1}{2}(3x - 1)$$

We can choose any value for x; we choose $x = 1$. This determines that $y = 1$. We can now use the equations of either plane to find z: when $x = 1$ and $y = 1$, $z = -1$ on both planes. We have found a point P on the line: $P = (1, 1, -1)$.

We now need the direction of the line. Since the line lies in each plane, its direction is orthogonal to a normal vector for each plane. Considering the equations for p_1 and p_2, we can quickly determine their normal vectors. For p_1,

Figure 11.58: The line and plane in Example 3.
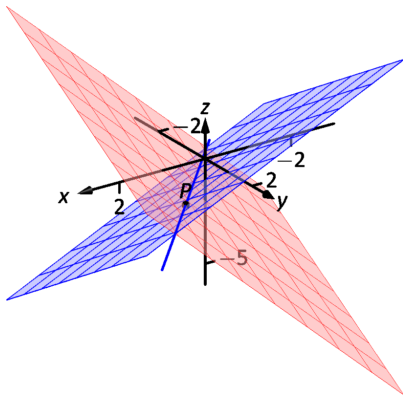
Figure 11.59: Graphing the planes and their line of intersection in Example 4.

Notes:

$\vec{n}_1 = \langle 1, -1, 1 \rangle$ and for p_2, $\vec{n}_2 = \langle -2, 1, 1 \rangle$. A direction orthogonal to both of these directions is their cross product: $\vec{d} = \vec{n}_1 \times \vec{n}_2 = \langle -2, -3, -1 \rangle$.

The parametric equations of the line through $P = (1, 1, -1)$ in the direction of $d = \langle -2, -3, -1 \rangle$ is:

$$\ell: \quad x = -2t + 1 \quad y = -3t + 1 \quad z = -t - 1.$$

The planes and line are graphed in Figure 11.59.

Example 5   **Finding the intersection of a plane and a line**

Find the point of intersection, if any, of the line $\ell(t) = \langle 3, -3, -1 \rangle + t \langle -1, 2, 1 \rangle$ and the plane with equation in general form $2x + y + z = 4$.

SOLUTION   The equation of the plane shows that the vector $\vec{n} = \langle 2, 1, 1 \rangle$ is a normal vector to the plane, and the equation of the line shows that the line moves parallel to $\vec{d} = \langle -1, 2, 1 \rangle$. Since these are not orthogonal, we know there is a point of intersection. (If there were orthogonal, it would mean that the plane and line were parallel to each other, either never intersecting or the line was in the plane itself.)

To find the point of intersection, we need to find a t value such that $\ell(t)$ satisfies the equation of the plane. Rewriting the equation of the line with parametric equations will help:

$$\ell(t) = \begin{cases} x = 3 - t \\ y = -3 + 2t \\ z = -1 + t \end{cases}.$$

Replacing x, y and z in the equation of the plane with the expressions containing t found in the equation of the line allows us to determine a t value that indicates the point of intersection:

$$2x + y + z = 4$$
$$2(3 - t) + (-3 + 2t) + (-1 + t) = 4$$
$$t = 2.$$

When $t = 2$, the point on the line satisfies the equation of the plane; that point is $\ell(2) = \langle 1, 1, 1 \rangle$. Thus the point $(1, 1, 1)$ is the point of intersection between the plane and the line, illustrated in Figure 11.60.
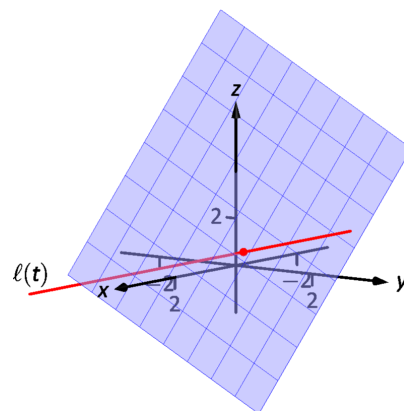
Figure 11.60: Illustrating the intersection of a line and a plane in Example 5.

Notes:

Distances

Just as it was useful to find distances between points and lines in the previous section, it is also often necessary to find the distance from a point to a plane.

Consider Figure 11.61, where a plane with normal vector $\vec{n}$ is sketched containing a point P and a point Q, not on the plane, is given. We measure the distance from Q to the plane by measuring the length of the projection of $\overrightarrow{PQ}$ onto $\vec{n}$. That is, we want:

$$\left\| \text{proj}_{\vec{n}}\, \overrightarrow{PQ} \right\| = \left\| \frac{\vec{n} \cdot \overrightarrow{PQ}}{\|\vec{n}\|^2}\vec{n} \right\| = \frac{|\vec{n} \cdot \overrightarrow{PQ}|}{\|\vec{n}\|} \tag{11.13}$$

Equation (11.13) is important as it does more than just give the distance between a point and a plane. We will see how it allows us to find several other distances as well: the distance between parallel planes and the distance from a line and a plane. Because Equation (11.13) is important, we restate it as a Key Idea.
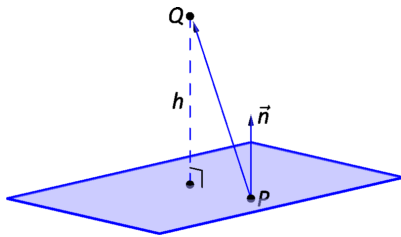
Figure 11.61: Illustrating finding the distance from a point to a plane.

Key Idea 54   Distance from a Point to a Plane

Let a plane with normal vector $\vec{n}$ be given, and let Q be a point. The distance h from Q to the plane is

$$h = \frac{|\vec{n} \cdot \overrightarrow{PQ}|}{\|\vec{n}\|},$$

where P is any point in the plane.

Example 6   Distance between a point and a plane
Find the distance between the point $Q = (2, 1, 4)$ and the plane with equation $2x - 5y + 6z = 9$.

Solution   Using the equation of the plane, we find the normal vector $\vec{n} = \langle 2, -5, 6 \rangle$. To find a point on the plane, we can let x and y be anything we choose, then let z be whatever satisfies the equation. Letting x and y be 0 seems simple; this makes $z = 1.5$. Thus we let $P = \langle 0, 0, 1.5 \rangle$, and $\overrightarrow{PQ} = \langle 2, 1, 2.5 \rangle$.

Notes:

The distance h from Q to the plane is given by Key Idea 54:

$$
\begin{aligned}
h &= \frac{|\vec{n} \cdot \overrightarrow{PQ}|}{\|\vec{n}\|} \\
&= \frac{|\langle 2, -5, 6 \rangle \cdot \langle 2, 1, 2.5 \rangle|}{\|\langle 2, -5, 6 \rangle\|} \\
&= \frac{|14|}{\sqrt{65}} \\
&= \frac{14}{\sqrt{65}}.
\end{aligned}
$$

We can use Key Idea 54 to find other distances. Given two parallel planes, we can find the distance between these planes by letting P be a point on one plane and Q a point on the other. If ℓ is a line parallel to a plane, we can use the Key Idea to find the distance between them as well: again, let P be a point in the plane and let Q be any point on the line. (One can also use Key Idea 53.) The exercises contain several problems of these types.

These past two sections have not explored lines and planes in space as an exercise of mathematical curiosity. However, there are many, many applications of these fundamental concepts. Complex shapes can be modeled (or, *approximated*) using planes. For instance, part of the exterior of an aircraft may have a complex, yet smooth, shape, and engineers will want to know how air flows across this piece as well as how heat might build up due to air friction. Many equations that help determine air flow and heat dissipation are difficult to apply to arbitrary surfaces, but simple to apply to planes. By approximating a surface with millions of small planes one can more readily model the needed behavior.

Notes:

Exercises 11.6

Terms and Concepts

1. In order to find the equation of a plane, what two pieces of information must one have?
2. What is the relationship between a plane and one of its normal vectors?

Problems

In Exercises 3–6, give any two points in the given plane.

3. $2x - 4y + 7z = 2$

4. $3(x + 2) + 5(y - 9) - 4z = 0$

5. $x = 2$

6. $4(y + 2) - (z - 6) = 0$

In Exercises 7–20, give the equation of the described plane in standard and general forms.

7. Passes through $(2, 3, 4)$ and has normal vector $\vec{n} = \langle 3, -1, 7 \rangle$.

8. Passes through $(1, 3, 5)$ and has normal vector $\vec{n} = \langle 0, 2, 4 \rangle$.

9. Passes through the points $(1, 2, 3)$, $(3, -1, 4)$ and $(1, 0, 1)$.

10. Passes through the points $(5, 3, 8)$, $(6, 4, 9)$ and $(3, 3, 3)$.

11. Contains the intersecting lines
    $\vec{\ell_1}(t) = \langle 2, 1, 2 \rangle + t \langle 1, 2, 3 \rangle$ and
    $\vec{\ell_2}(t) = \langle 2, 1, 2 \rangle + t \langle 2, 5, 4 \rangle$.

12. Contains the intersecting lines
    $\vec{\ell_1}(t) = \langle 5, 0, 3 \rangle + t \langle -1, 1, 1 \rangle$ and
    $\vec{\ell_2}(t) = \langle 1, 4, 7 \rangle + t \langle 3, 0, -3 \rangle$.

13. Contains the parallel lines
    $\vec{\ell_1}(t) = \langle 1, 1, 1 \rangle + t \langle 1, 2, 3 \rangle$ and
    $\vec{\ell_2}(t) = \langle 1, 1, 2 \rangle + t \langle 1, 2, 3 \rangle$.

14. Contains the parallel lines
    $\vec{\ell_1}(t) = \langle 1, 1, 1 \rangle + t \langle 4, 1, 3 \rangle$ and
    $\vec{\ell_2}(t) = \langle 2, 2, 2 \rangle + t \langle 4, 1, 3 \rangle$.

15. Contains the point $(2, -6, 1)$ and the line
    $$\ell(t) = \begin{cases} x = 2 + 5t \\ y = 2 + 2t \\ z = -1 + 2t \end{cases}$$

16. Contains the point $(5, 7, 3)$ and the line
    $$\ell(t) = \begin{cases} x = t \\ y = t \\ z = t \end{cases}$$

17. Contains the point $(5, 7, 3)$ and is orthogonal to the line
    $\vec{\ell}(t) = \langle 4, 5, 6 \rangle + t \langle 1, 1, 1 \rangle$.

18. Contains the point $(4, 1, 1)$ and is orthogonal to the line
    $$\ell(t) = \begin{cases} x = 4 + 4t \\ y = 1 + 1t \\ z = 1 + 1t \end{cases}$$

19. Contains the point $(-4, 7, 2)$ and is parallel to the plane
    $3(x - 2) + 8(y + 1) - 10z = 0$.

20. Contains the point $(1, 2, 3)$ and is parallel to the plane
    $x = 5$.

In Exercises 21–22, give the equation of the line that is the intersection of the given planes.

21. $p1$: $3(x - 2) + (y - 1) + 4z = 0$, and
    $p2$: $2(x - 1) - 2(y + 3) + 6(z - 1) = 0$.

22. $p1$: $5(x - 5) + 2(y + 2) + 4(z - 1) = 0$, and
    $p2$: $3x - 4(y - 1) + 2(z - 1) = 0$.

In Exercises 23–26, find the point of intersection between the line and the plane.

23. line: $\langle 5, 1, -1 \rangle + t \langle 2, 2, 1 \rangle$,
    plane: $5x - y - z = -3$

24. line: $\langle 4, 1, 0 \rangle + t \langle 1, 0, -1 \rangle$,
    plane: $3x + y - 2z = 8$

25. line: $\langle 1, 2, 3 \rangle + t \langle 3, 5, -1 \rangle$,
    plane: $3x - 2y - z = 4$

26. line: $\langle 1, 2, 3 \rangle + t \langle 3, 5, -1 \rangle$,
    plane: $3x - 2y - z = -4$

In Exercises 27–30, find the given distances.

27. The distance from the point $(1, 2, 3)$ to the plane
    $3(x - 1) + (y - 2) + 5(z - 2) = 0$.

28. The distance from the point $(2, 6, 2)$ to the plane
    $2(x - 1) - y + 4(z + 1) = 0$.

29. The distance between the parallel planes
    $x + y + z = 0$ and
    $(x - 2) + (y - 3) + (z + 4) = 0$

30. The distance between the parallel planes
    $2(x - 1) + 2(y + 1) + (z - 2) = 0$ and
    $2(x - 3) + 2(y - 1) + (z - 3) = 0$

31. Show why if the point Q lies in a plane, then the distance formula correctly gives the distance from the point to the plane as 0.

32. How is Exercise 30 in Section 11.5 easier to answer once we have an understanding of planes?

11.7   Curvilinear Coordinates

The Cartesian coordinates of a point (x, y, z) are determined by following straight paths starting from the origin: first along the x-axis, then parallel to the y-axis, then parallel to the z-axis, as in Figure 11.62(a). In *curvilinear coordinate systems*, these paths can be curved. The two types of curvilinear coordinates which we will consider are cylindrical and spherical coordinates. Instead of referencing a point in terms of sides of a rectangular parallelepiped, as with Cartesian coordinates, we will think of the point as lying on a cylinder or sphere. Cylindrical coordinates are often used when there is symmetry around the z-axis; spherical coordinates are useful when there is symmetry about the origin.

Let $P = (x, y, z)$ be a point in Cartesian coordinates in $\mathbb{R}^3$, and let $P_0 = (x, y, 0)$ be the projection of P upon the xy-plane. Treating (x, y) as a point in $\mathbb{R}^2$, let (r, θ) be its polar coordinates (see Figure 11.62(b)). Let ρ be the length of the line segment from the origin to P, and let ϕ be the angle between that line segment and the positive z-axis (see Figure 11.62(c)), which is called the *zenith angle*. Then the **cylindrical coordinates** (r, θ, z) and the **spherical coordinates** (ρ, θ, ϕ) of $P(x, y, z)$ are defined as follows:

Key Idea 55   Cylindrical coordinates (r, θ, z)

$$x = r\cos\theta \qquad r = \sqrt{x^2 + y^2}$$
$$y = r\sin\theta \qquad \theta = \tan^{-1}\left(\frac{y}{x}\right)$$
$$z = z \qquad z = z$$

where $0 \le \theta \le \pi$ if $y \ge 0$ and $\pi < \theta < 2\pi$ if $y < 0$.

Key Idea 56   Spherical coordinates (ρ, θ, ϕ)

$$x = \rho\sin\phi\,\cos\theta \qquad \rho = \sqrt{x^2 + y^2 + z^2}$$
$$y = \rho\sin\phi\,\sin\theta \qquad \theta = \tan^{-1}\left(\frac{y}{x}\right)$$
$$z = \rho\cos\phi \qquad \phi = \cos^{-1}\left(\frac{z}{\sqrt{x^2+y^2+z^2}}\right)$$

where $0 \le \theta \le \pi$ if $y \ge 0$ and $\pi < \theta < 2\pi$ if $y < 0$.
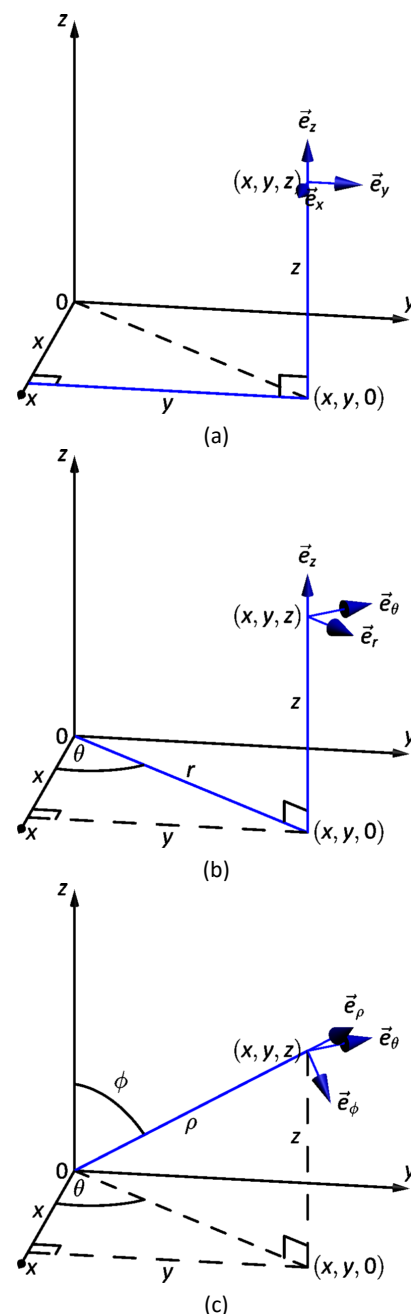
Notes:

Figure 11.62: Cartesian (top), cylindrical (middle), and spherical (bottom) coordinate systems

This "standard" definition of spherical coordinates used by mathematicians results in a left-handed system. For this reason, physicists usually switch the definitions of θ and ϕ to make (ρ, θ, ϕ) a right-handed system.

Both θ and ϕ are measured in radians. Note that $r \geq 0$, $0 \leq \theta < 2\pi$, $\rho \geq 0$ and $0 \leq \phi \leq \pi$. Also, θ is undefined when $(x, y) = (0, 0)$, and ϕ is undefined when $(x, y, z) = (0, 0, 0)$.

Watch the video:
Conversion From Rectangular Coordinates at
`https://youtu.be/w9wCdLuklw8`

Example 1   **Converting Between Coordinate Systems**
Convert the point $(-2, -2, 1)$ from Cartesian coordinates to 1. cylindrical and 2. spherical coordinates.

SOLUTION

1. $r = \sqrt{(-2)^2 + (-2)^2} = 2\sqrt{2}$ and $\theta = \tan^{-1}\left(\frac{-2}{-2}\right) = \tan^{-1}(1) = \frac{5\pi}{4}$, since $y = -2 < 0$. Therefore $(r, \theta, z) = \left(2\sqrt{2}, \frac{5\pi}{4}, 1\right)$.

2. $\rho = \sqrt{(-2)^2 + (-2)^2 + 1^2} = \sqrt{9} = 3$ and $\phi = \cos^{-1}\left(\frac{1}{3}\right) \approx 1.23$ radians. Therefore $(\rho, \theta, \phi) = \left(3, \frac{5\pi}{4}, 1.23\right)$.

For cylindrical coordinates (r, θ, z), and constants r_0, θ_0 and z_0, we see from Figure 11.63 that the surface $r = r_0$ is a cylinder of radius r_0 centered along the z-axis, the surface $\theta = \theta_0$ is a half-plane emanating from the z-axis, and the surface $z = z_0$ is a plane parallel to the xy-plane.

(a) $r = r_0$ (b) $\theta = \theta_0$ (c) $z = z_0$
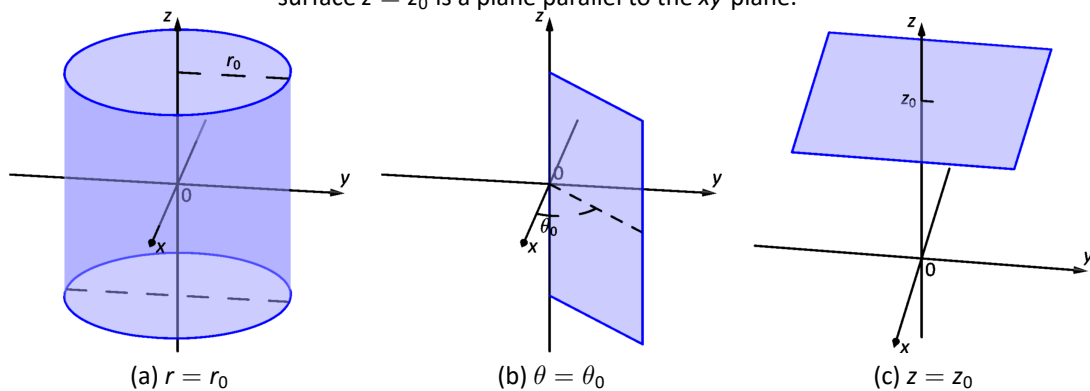
Figure 11.63: Cylindrical coordinate surfaces

Notes:

For spherical coordinates (ρ, θ, ϕ), and constants ρ_0, θ_0 and ϕ_0, we see from Figure 11.64 that the surface $\rho = \rho_0$ is a sphere of radius ρ_0 centered at the origin, the surface $\theta = \theta_0$ is a half-plane emanating from the z-axis, and the surface $\phi = \phi_0$ is a circular cone whose vertex is at the origin.

(a) $\rho = \rho_0$ (b) $\theta = \theta_0$ (c) $\phi = \phi_0$
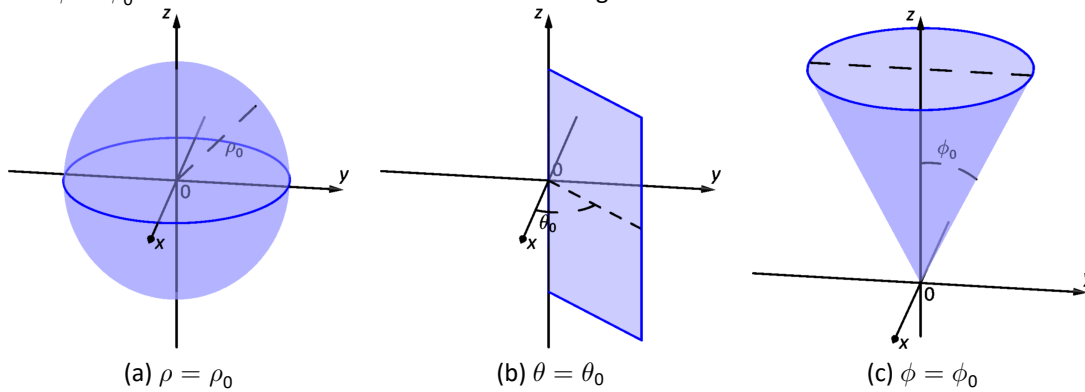
Figure 11.64: Spherical coordinate surfaces

Figures 11.63(a) and 11.64(a) show how these coordinate systems got their names.

Sometimes the equation of a surface in Cartesian coordinates can be transformed into a simpler equation in some other coordinate system, as in the following example.

Example 2   **Converting an Equation in Coordinate Systems**
Write the equation of the cylinder $x^2 + y^2 = 4$ in cylindrical coordinates.

     SOLUTION   Since $r = \sqrt{x^2 + y^2}$, then the equation in cylindrical coordinates is $r = 2$.

Using spherical coordinates to write the equation of a sphere does not necessarily make the equation simpler, if the sphere is not centered at the origin.

Example 3   **Converting an Equation to Spherical Coordinates**
Write the equation $(x - 2)^2 + (y - 1)^2 + z^2 = 9$ in spherical coordinates.

     SOLUTION   Multiplying the equation out gives

$$x^2 + y^2 + z^2 - 4x - 2y + 5 = 9 \text{ , so we get}$$
$$\rho^2 - 4\rho \sin \phi \cos \theta - 2\rho \sin \phi \sin \theta - 4 = 0 \text{ , or}$$
$$\rho^2 - 2 \sin \phi (2 \cos \theta - \sin \theta)\rho - 4 = 0$$

Notes:

707

after combining terms. Note that this actually makes it more difficult to figure out what the surface is, as opposed to the Cartesian equation where you could immediately identify the surface as a sphere of radius 3 centered at $(2, 1, 0)$.

Example 4   **Identifying a Surface**
Describe the surface given by $\theta = z$ in cylindrical coordinates.

SOLUTION   This surface is called a *helicoid*. As the (vertical) z coordinate increases, so does the angle θ, while the radius r is unrestricted. So this sweeps out a (ruled!) surface shaped like a spiral staircase, where the spiral has an infinite radius. Figure 11.65 shows a section of this surface restricted to $0 \le z \le 4\pi$ and $0 \le r \le 2$.
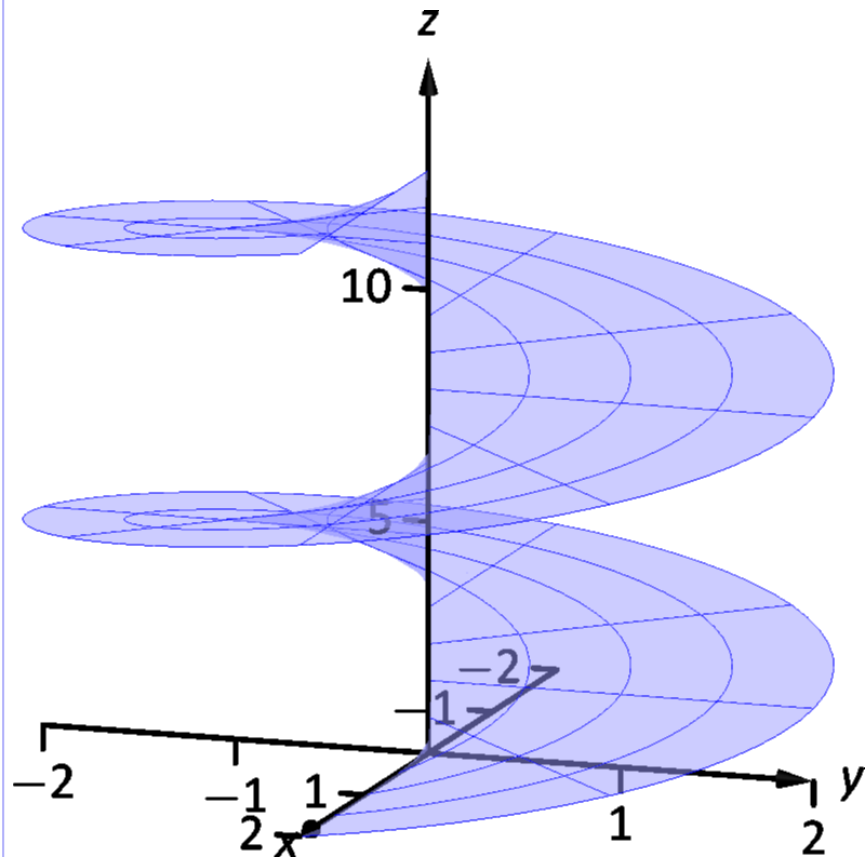
Figure 11.65: Helicoid $\theta = z$

Notes:

Exercises 11.7

Problems

In Exercises 1–4, find the (a) cylindrical and (b) spherical coordinates of the point whose Cartesian coordinates are given.

1. $(2, 2\sqrt{3}, -1)$
2. $(-5, 5, 6)$
3. $(\sqrt{21}, -\sqrt{7}, 0)$
4. $(0, \sqrt{2}, 2)$

In Exercises 5–7, write the given equation in (a) cylindrical and (b) spherical coordinates.

5. $x^2 + y^2 + z^2 = 25$
6. $x^2 + y^2 = 2y$
7. $x^2 + y^2 + 9z^2 = 36$
8. Describe the intersection of the surfaces whose equations in spherical coordinates are $\theta = \frac{\pi}{2}$ and $\phi = \frac{\pi}{4}$.
9. Show that for $a \neq 0$, the equation $\rho = 2a \sin\phi \, \cos\theta$ in spherical coordinates describes a sphere centered at $(a, 0, 0)$ with radius $|a|$.
10. Let $P = (a, \theta, \phi)$ be a point in spherical coordinates, with $a > 0$ and $0 < \phi < \pi$. Then P lies on the sphere $\rho = a$. Since $0 < \phi < \pi$, the line segment from the origin to P can be extended to intersect the cylinder given by $r = a$ (in cylindrical coordinates). Find the cylindrical coordinates of that point of intersection.

11. Let P_1 and P_2 be points whose spherical coordinates are $(\rho_1, \theta_1, \phi_1)$ and $(\rho_2, \theta_2, \phi_2)$, respectively. Let $\vec{v}_1$ be the vector from the origin to P_1, and let $\vec{v}_2$ be the vector from the origin to P_2. For the angle γ between $\vec{v}_1$ and $\vec{v}_2$, show that

$$\cos\gamma = \cos\phi_1 \, \cos\phi_2 + \sin\phi_1 \, \sin\phi_2 \, \cos(\theta_2 - \theta_1).$$

This formula is used in electrodynamics to prove the addition theorem for spherical harmonics, which provides a general expression for the electrostatic potential at a point due to a unit charge.

12. Show that the distance d between the points P_1 and P_2 with cylindrical coordinates (r_1, θ_1, z_1) and (r_2, θ_2, z_2), respectively, is

$$d = \sqrt{r_1^2 + r_2^2 - 2r_1 r_2 \cos(\theta_2 - \theta_1) + (z_2 - z_1)^2}\,.$$

13. Show that the distance d between the points P_1 and P_2 with spherical coordinates $(\rho_1, \theta_1, \phi_1)$ and $(\rho_2, \theta_2, \phi_2)$, respectively, is

$$d = \sqrt{\rho_1^2 + \rho_2^2 - 2\rho_1 \rho_2 [\sin\phi_1 \, \sin\phi_2 \, \cos(\theta_2 - \theta_1) + \cos\phi_1 \, \cos\phi_2]}\,.$$

12: VECTOR VALUED FUNCTIONS

In the previous chapter, we learned about vectors and were introduced to the power of vectors within mathematics. In this chapter, we'll build on this foundation to define functions whose input is a real number and whose output is a vector. We'll see how to graph these functions and apply calculus techniques to analyze their behavior. Most importantly, we'll see *why* we are interested in doing this: we'll see beautiful applications to the study of moving objects.

12.1 Vector–Valued Functions

We are very familiar with **real valued functions**, that is, functions whose output is a real number. This section introduces **vector–valued functions** — functions whose output is a vector.

> **Definition 67   Vector–Valued Functions**
>
> A **vector–valued function** is a function of the form
>
> $$\vec{r}(t) = \langle f(t), g(t) \rangle \quad \text{or} \quad \vec{r}(t) = \langle f(t), g(t), h(t) \rangle,$$
>
> where f, g and h are real valued functions.
>
> The **domain** of $\vec{r}$ is the set of all values of t for which $\vec{r}(t)$ is defined. The **range** of $\vec{r}$ is the set of all possible output vectors $\vec{r}(t)$.

Evaluating and Graphing Vector–Valued Functions

Evaluating a vector–valued function at a specific value of t is straightforward; simply evaluate each component function at that value of t. For instance, if $\vec{r}(t) = \langle t^2, t^2 + t - 1 \rangle$, then $\vec{r}(-2) = \langle 4, 1 \rangle$. We can sketch this vector, as is done in Figure 12.1(a). Plotting lots of vectors is cumbersome, though, so generally we do not sketch the whole vector but just the terminal point. The **graph** of a vector–valued function is the set of all terminal points of $\vec{r}(t)$, where the initial point of each vector is always the origin. In Figure 12.1(b) we sketch the graph of $\vec{r}$; we can indicate individual points on the graph with their respective vector, as shown.

Vector–valued functions are closely related to parametric equations of graphs. While in both methods we plot points $\big(x(t), y(t)\big)$ or $\big(x(t), y(t), z(t)\big)$ to produce

(a)

(b)
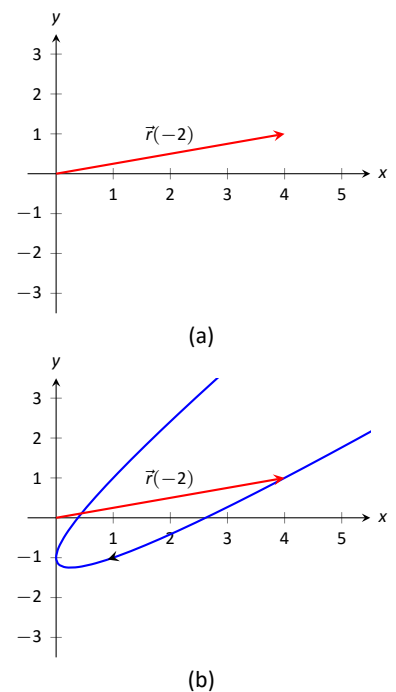
Figure 12.1: Sketching the graph of a vector–valued function.

a graph, in the context of vector–valued functions each such point represents a vector. The implications of this will be more fully realized in the next section as we apply calculus ideas to these functions.

Watch the video:
Domain of a Vector–Valued Function at
`https://youtu.be/Djtttm0C7zA`

t	$t^3 - t$	$\dfrac{1}{t^2+1}$
-2	-6	$1/5$
-1	0	$1/2$
0	0	1
1	0	$1/2$
2	6	$1/5$

(a)
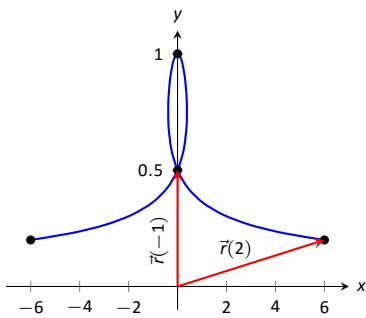
(b)

Figure 12.2: Sketching the vector–valued function of Example 1.

Example 1   **Graphing vector–valued functions**

Graph $\vec{r}(t) = \left\langle t^3 - t, \dfrac{1}{t^2+1} \right\rangle$, for $-2 \le t \le 2$. Sketch $\vec{r}(-1)$ and $\vec{r}(2)$.

SOLUTION   We start by making a table of t, x and y values as shown in Figure 12.2(a). Plotting these points gives an indication of what the graph looks like. In Figure 12.2(b), we indicate these points and sketch the full graph. We also highlight $\vec{r}(-1)$ and $\vec{r}(2)$ on the graph.

Example 2   **Graphing vector–valued functions.**

Graph $\vec{r}(t) = \langle \cos t, \sin t, t \rangle$ for $0 \le t \le 4\pi$.

SOLUTION   We can again plot points, but careful consideration of this function is very revealing. Momentarily ignoring the third component, we see the x and y components trace out a circle of radius 1 centered at the origin. Noticing that the z component is t, we see that as the graph winds around the z-axis, it is also increasing at a constant rate in the positive z direction, forming a spiral. This is graphed in Figure 12.3. In the graph $\vec{r}(7\pi/4) = \left\langle \dfrac{1}{\sqrt{2}}, -\dfrac{1}{\sqrt{2}}, \dfrac{7\pi}{4} \right\rangle$ is highlighted to help us understand the graph.

Algebra of Vector–Valued Functions
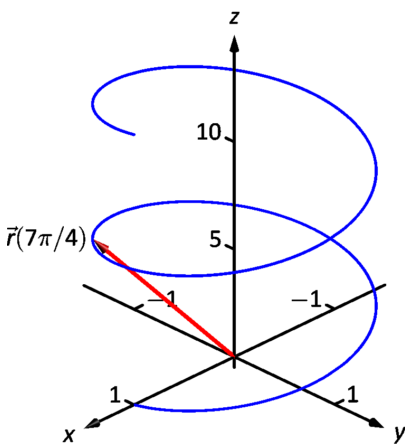
Figure 12.3: Viewing a vector–valued function, and its value at one point.

Notes:

Definition 68   Operations on Vector–Valued Functions

Let $\vec{r}_1(t) = \langle f_1(t), g_1(t) \rangle$ and $\vec{r}_2(t) = \langle f_2(t), g_2(t) \rangle$ be vector–valued functions in $\mathbb{R}^2$ and let c be a scalar. Then:

1. $\vec{r}_1(t) \pm \vec{r}_2(t) = \langle f_1(t) \pm f_2(t), g_1(t) \pm g_2(t) \rangle$.

2. $c\vec{r}_1(t) = \langle cf_1(t), cg_1(t) \rangle$.

A similar definition holds for vector–valued functions in $\mathbb{R}^3$.

This definition states that we add, subtract and scale vector-valued functions component–wise. Combining vector–valued functions in this way can be very useful (as well as create interesting graphs).

Example 3   Adding and scaling vector–valued functions.
Let $\vec{r}_1(t) = \langle 0.2t, 0.3t \rangle$, $\vec{r}_2(t) = \langle \cos t, \sin t \rangle$ and $\vec{r}(t) = \vec{r}_1(t) + \vec{r}_2(t)$. Graph $\vec{r}_1(t), \vec{r}_2(t), \vec{r}(t)$ and $5\vec{r}(t)$ on $-10 \le t \le 10$.

SOLUTION   We can graph $\vec{r}_1$ and $\vec{r}_2$ easily by plotting points (or just using technology). Let's think about each for a moment to better understand how vector–valued functions work.

We can rewrite $\vec{r}_1(t) = \langle 0.2t, 0.3t \rangle$ as $\vec{r}_1(t) = t\langle 0.2, 0.3 \rangle$. That is, the function $\vec{r}_1$ scales the vector $\langle 0.2, 0.3 \rangle$ by t. This scaling of a vector produces a line in the direction of $\langle 0.2, 0.3 \rangle$.

We are familiar with $\vec{r}_2(t) = \langle \cos t, \sin t \rangle$; it traces out a circle, centered at the origin, of radius 1. Figure 12.4(a) graphs $\vec{r}_1(t)$ and $\vec{r}_2(t)$.

Adding $\vec{r}_1(t)$ to $\vec{r}_2(t)$ produces $\vec{r}(t) = \langle \cos t + 0.2t, \sin t + 0.3t \rangle$, graphed in Figure 12.4(b). The linear movement of the line combines with the circle to create loops that move in the direction of $\langle 0.2, 0.3 \rangle$. (We encourage the reader to experiment by changing $\vec{r}_1(t)$ to $\langle 2t, 3t \rangle$, etc., and observe the effects on the loops.)

Multiplying $\vec{r}(t)$ by 5 scales the function by 5, producing $5\vec{r}(t) = \langle 5\cos t + 1, 5\sin t + 1.5 \rangle$, which is graphed in Figure 12.4(c) along with $\vec{r}(t)$. The new function is "5 times bigger" than $\vec{r}(t)$. Note how the graph of $5\vec{r}(t)$ in (c) looks identical to the graph of $\vec{r}(t)$ in (b). This is due to the fact that the x and y bounds of the plot in (c) are exactly 5 times larger than the bounds in (b).

Example 4   Adding and scaling vector–valued functions.
A **cycloid** is a graph traced by a point p on a rolling circle, as shown in Figure 12.6. Find an equation describing the cycloid, where the circle has radius 1.

Notes:

(a)

(b)

(c)
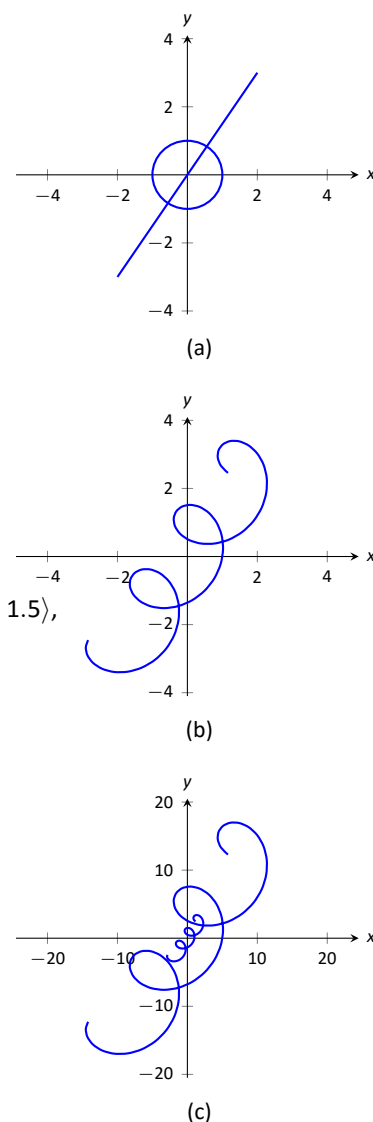
Figure 12.4: Graphing the functions in Example 3.

SOLUTION   This problem is not very difficult if we approach it in a clever way. We start by letting $\vec{p}(t)$ describe the position of the point p on the circle, where the circle is centered at the origin and only rotates clockwise (i.e., it does not roll). This is relatively simple given our previous experiences with parametric equations; $\vec{p}(t) = \langle \cos t, -\sin t \rangle$.

We now want the circle to roll. We represent this by letting $\vec{c}(t)$ represent the location of the center of the circle. It should be clear that the y component of $\vec{c}(t)$ should be 1; the center of the circle is always going to be 1 if it rolls on a horizontal surface.

The x component of $\vec{c}(t)$ is a linear function of t: $f(t) = mt$ for some scalar m. When $t = 0$, $f(t) = 0$ (the circle starts centered on the y-axis). When $t = 2\pi$, the circle has made one complete revolution, traveling a distance equal to its circumference, which is also 2π. This gives us a point on our line $f(t) = mt$, the point $(2\pi, 2\pi)$. It should be clear that $m = 1$ and $f(t) = t$. So $\vec{c}(t) = \langle t, 1 \rangle$.

We now combine $\vec{p}$ and $\vec{c}$ together to form the equation of the cycloid: $\vec{r}(t) = \vec{p}(t) + \vec{c}(t) = \langle \cos t + t, -\sin t + 1 \rangle$, which is graphed in Figure 12.5.

Displacement

A vector–valued function $\vec{r}(t)$ is often used to describe the position of a moving object at time t. At $t = t_0$, the object is at $\vec{r}(t_0)$; at $t = t_1$, the object is at $\vec{r}(t_1)$. Knowing the locations $\vec{r}(t_0)$ and $\vec{r}(t_1)$ give no indication of the path taken between them, but often we only care about the difference of the locations, $\vec{r}(t_1) - \vec{r}(t_0)$, the **displacement**.
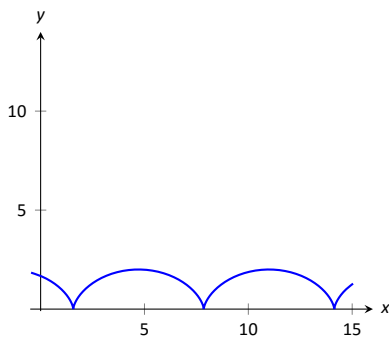
Figure 12.5: The cycloid in Example 4.

Definition 69   Displacement

Let $\vec{r}(t)$ be a vector–valued function and let $t_0 < t_1$ be values in the domain. The **displacement** $\vec{d}$ of $\vec{r}$, from $t = t_0$ to $t = t_1$, is

$$\vec{d} = \vec{r}(t_1) - \vec{r}(t_0).$$
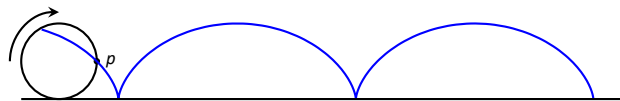
Figure 12.6: Tracing a cycloid.

Notes:

When the displacement vector is drawn with initial point at $\vec{r}(t_0)$, its terminal point is $\vec{r}(t_1)$. We think of it as the vector which points from a starting position to an ending position.

Example 5   **Finding and graphing displacement vectors**
Let $\vec{r}(t) = \left\langle \cos(\frac{\pi}{2}t), \sin(\frac{\pi}{2}t) \right\rangle$. Graph $\vec{r}(t)$ on $-1 \leq t \leq 1$, and find the displacement of $\vec{r}(t)$ on this interval.

SOLUTION   The function $\vec{r}(t)$ traces out the unit circle, though at a different rate than the "usual" $\langle \cos t, \sin t \rangle$ parametrization. At $t_0 = -1$, we have $\vec{r}(t_0) = \langle 0, -1 \rangle$; at $t_1 = 1$, we have $\vec{r}(t_1) = \langle 0, 1 \rangle$. The displacement of $\vec{r}(t)$ on $[-1, 1]$ is thus $\vec{d} = \langle 0, 1 \rangle - \langle 0, -1 \rangle = \langle 0, 2 \rangle$.
A graph of $\vec{r}(t)$ on $[-1, 1]$ is given in Figure 12.6, along with the displacement vector $\vec{d}$ on this interval.

Measuring displacement makes us contemplate related, yet very different, concepts. Considering the semi–circular path the object in Example 5 took, we can quickly verify that the object ended up a distance of 2 units from its initial location. That is, we can compute $\left\| \vec{d} \right\| = 2$. However, measuring *distance from the starting point* is different from measuring *distance traveled*. Being a semi–circle, we can measure the distance traveled by this object as π units. Knowing *distance from the starting point* allows us to compute **average rate of change.**
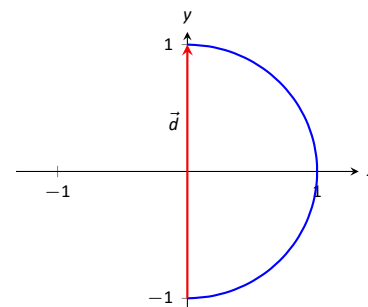
Figure 12.6: Graphing the displacement of a position function in Example 5.

> **Definition 70**   **Average Rate of Change**
> Let $\vec{r}(t)$ be a vector–valued function, where each of its component functions is continuous on its domain, and let $t_0 < t_1$. The **average rate of change** of $\vec{r}(t)$ on $[t_0, t_1]$ is
>
> $$\text{average rate of change} = \frac{\vec{r}(t_1) - \vec{r}(t_0)}{t_1 - t_0}.$$

Example 6   **Average rate of change**
Let $\vec{r}(t) = \left\langle \cos(\frac{\pi}{2}t), \sin(\frac{\pi}{2}t) \right\rangle$ as in Example 5. Find the average rate of change of $\vec{r}(t)$ on $[-1, 1]$ and on $[-1, 5]$.

SOLUTION   We computed in Example 5 that the displacement of $\vec{r}(t)$ on $[-1, 1]$ was $\vec{d} = \langle 0, 2 \rangle$. Thus the average rate of change of $\vec{r}(t)$ on $[-1, 1]$ is:

$$\frac{\vec{r}(1) - \vec{r}(-1)}{1 - (-1)} = \frac{\langle 0, 2 \rangle}{2} = \langle 0, 1 \rangle.$$

Notes:

We interpret this as follows: the object followed a semi–circular path, meaning it moved towards the right then moved back to the left, while climbing slowly, then quickly, then slowly again. *On average*, however, it progressed straight up at a constant rate of $\langle 0, 1\rangle$ per unit of time.

We can quickly see that the displacement on $[-1, 5]$ is the same as on $[-1, 1]$, so $\vec{d} = \langle 0, 2\rangle$. The average rate of change is different, though:

$$\frac{\vec{r}(5) - \vec{r}(-1)}{5 - (-1)} = \frac{\langle 0, 2\rangle}{6} = \langle 0, 1/3\rangle \, .$$

As it took "3 times as long" to arrive at the same place, this average rate of change on $[-1, 5]$ is $1/3$ the average rate of change on $[-1, 1]$.

We considered average rates of change in Sections 1.1 and 2.1 as we studied limits and derivatives. The same is true here; in the following section we apply calculus concepts to vector–valued functions as we find limits, derivatives, and integrals. Understanding the average rate of change will give us an understanding of the derivative; displacement gives us one application of integration.

Notes:

Exercises 12.1

Terms and Concepts

1. Vector–valued functions are closely related to _____ _____ of graphs.
2. When sketching vector–valued functions, technically one isn't graphing points, but rather _____.
3. It can be useful to think of _____ as a vector that points from a starting position to an ending position.

Problems

In Exercises 4–11, sketch the vector–valued function on the given interval.

4. $\vec{r}(t) = \langle t^2, t^2 - 1 \rangle$, for $-2 \le t \le 2$.
5. $\vec{r}(t) = \langle t^2, t^3 \rangle$, for $-2 \le t \le 2$.
6. $\vec{r}(t) = \langle 1/t, 1/t^2 \rangle$, for $-2 \le t \le 2$.
7. $\vec{r}(t) = \langle \frac{1}{10}t^2, \sin t \rangle$, for $-2\pi \le t \le 2\pi$.
8. $\vec{r}(t) = \langle \frac{1}{10}t^2, \sin t \rangle$, for $-2\pi \le t \le 2\pi$.
9. $\vec{r}(t) = \langle 3\sin(\pi t), 2\cos(\pi t) \rangle$, on $[0, 2]$.
10. $\vec{r}(t) = \langle 3\cos t, 2\sin(2t) \rangle$, on $[0, 2\pi]$.
11. $\vec{r}(t) = \langle 2\sec t, \tan t \rangle$, on $[-\pi, \pi]$.

In Exercises 12–15, sketch the vector–valued function on the given interval in $\mathbb{R}^3$. Technology may be useful in creating the sketch.

12. $\vec{r}(t) = \langle 2\cos t, t, 2\sin t \rangle$, on $[0, 2\pi]$.
13. $\vec{r}(t) = \langle 3\cos t, \sin t, t/\pi \rangle$ on $[0, 2\pi]$.
14. $\vec{r}(t) = \langle \cos t, \sin t, \sin t \rangle$ on $[0, 2\pi]$.
15. $\vec{r}(t) = \langle \cos t, \sin t, \sin(2t) \rangle$ on $[0, 2\pi]$.

In Exercises 16–19, find $\|\vec{r}(t)\|$.

16. $\vec{r}(t) = \langle t, t^2 \rangle$.
17. $\vec{r}(t) = \langle 5\cos t, 3\sin t \rangle$.
18. $\vec{r}(t) = \langle 2\cos t, 2\sin t, t \rangle$.
19. $\vec{r}(t) = \langle \cos t, t, t^2 \rangle$.

In Exercises 20–27, create a vector–valued function whose graph matches the given description.

20. A circle of radius 2, centered at $(1, 2)$, traced counter–clockwise once on $[0, 2\pi]$.
21. A circle of radius 3, centered at $(5, 5)$, traced clockwise once on $[0, 2\pi]$.
22. An ellipse, centered at $(0, 0)$ with vertical major axis of length 10 and minor axis of length 3, traced once counter–clockwise on $[0, 2\pi]$.
23. An ellipse, centered at $(3, -2)$ with horizontal major axis of length 6 and minor axis of length 4, traced once clockwise on $[0, 2\pi]$.
24. A line through $(2, 3)$ with a slope of 5.
25. A line through $(1, 5)$ with a slope of $-1/2$.
26. A vertically oriented helix with radius of 2 that starts at $(2, 0, 0)$ and ends at $(2, 0, 4\pi)$ after 1 revolution on $[0, 2\pi]$.
27. A vertically oriented helix with radius of 3 that starts at $(3, 0, 0)$ and ends at $(3, 0, 3)$ after 2 revolutions on $[0, 1]$.

In Exercises 28–31, find the average rate of change of $\vec{r}(t)$ on the given interval.

28. $\vec{r}(t) = \langle t, t^2 \rangle$ on $[-2, 2]$.
29. $\vec{r}(t) = \langle t, t + \sin t \rangle$ on $[0, 2\pi]$.
30. $\vec{r}(t) = \langle 3\cos t, 2\sin t, t \rangle$ on $[0, 2\pi]$.
31. $\vec{r}(t) = \langle t, t^2, t^3 \rangle$ on $[-1, 3]$.

12.2   Calculus and Vector–Valued Functions

The previous section introduced us to a new mathematical object, the vector–valued function. We now apply calculus concepts to these functions. We start with the limit, then work our way through derivatives to integrals.

Limits of Vector–Valued Functions

The initial definition of the limit of a vector–valued function is a bit intimidating, as was the definition of the limit in Definition 1. The theorem following the definition shows that in practice, taking limits of vector–valued functions is no more difficult than taking limits of real–valued functions.

Definition 71   Limits of Vector–Valued Functions
Let I be an open interval containing c, and let $\vec{r}(t)$ be a vector–valued function defined on I, except possibly at c. The **limit of $\vec{r}(t)$, as t approaches c, is $\vec{L}$,** expressed as

$$\lim_{t \to c} \vec{r}(t) = \vec{L},$$

means that given any $\varepsilon > 0$, there exists a $\delta > 0$ such that for all $t \neq c$, if $|t - c| < \delta$, we have $\left\| \vec{r}(t) - \vec{L} \right\| < \varepsilon$.

Note how the measurement of distance between real numbers is the absolute value of their difference; the measure of distance between vectors is the vector norm, or magnitude, of their difference.

Notes:

Theorem 91   Limits of Vector–Valued Functions

1. Let $\vec{r}(t) = \langle f(t), g(t) \rangle$ be a vector–valued function in $\mathbb{R}^2$ defined on an open interval I containing c. Then

$$\lim_{t \to c} \vec{r}(t) = \left\langle \lim_{t \to c} f(t), \lim_{t \to c} g(t) \right\rangle.$$

2. Let $\vec{r}(t) = \langle f(t), g(t), h(t) \rangle$ be a vector–valued function in $\mathbb{R}^3$ defined on an open interval I containing c. Then

$$\lim_{t \to c} \vec{r}(t) = \left\langle \lim_{t \to c} f(t), \lim_{t \to c} g(t), \lim_{t \to c} h(t) \right\rangle$$

Theorem 91 states that we compute limits component–wise.

Example 1   Finding limits of vector–valued functions

Let $\vec{r}(t) = \left\langle \dfrac{\sin t}{t}, t^2 - 3t + 3, \cos t \right\rangle$. Find $\lim\limits_{t \to 0} \vec{r}(t)$.

SOLUTION   We apply the theorem and compute limits component–wise.

$$\lim_{t \to 0} \vec{r}(t) = \left\langle \lim_{t \to 0} \frac{\sin t}{t}, \lim_{t \to 0} t^2 - 3t + 3, \lim_{t \to 0} \cos t \right\rangle$$
$$= \langle 1, 3, 1 \rangle.$$

Continuity

Definition 72   Continuity of Vector–Valued Functions
Let $\vec{r}(t)$ be a vector–valued function defined on an open interval I containing c.

1. $\vec{r}(t)$ is **continuous at** c if $\lim\limits_{t \to c} \vec{r}(t) = \vec{r}(c)$.

2. If $\vec{r}(t)$ is continuous at all c in I, then $\vec{r}(t)$ is **continuous on** I.

We again have a theorem that lets us evaluate continuity component–wise.

Notes:

> **Theorem 92   Continuity of Vector–Valued Functions**
> Let $\vec{r}(t)$ be a vector–valued function defined on an open interval I containing c. Then $\vec{r}(t)$ is continuous at c if, and only if, each of its component functions is continuous at c.

Example 2   Evaluating continuity of vector–valued functions
Let $\vec{r}(t) = \left\langle \dfrac{\sin t}{t}, \ t^2 - 3t + 3, \ \cos t \right\rangle$. Determine whether $\vec{r}$ is continuous at $t = 0$ and $t = 1$.

SOLUTION   While the second and third components of $\vec{r}(t)$ are defined at $t = 0$, the first component, $(\sin t)/t$, is not. Since the first component is not even defined at $t = 0$, $\vec{r}(t)$ is not defined at $t = 0$, and hence it is not continuous at $t = 0$.
   At $t = 1$ each of the component functions is continuous. Therefore $\vec{r}(t)$ is continuous at $t = 1$.

Derivatives

Consider a vector–valued function $\vec{r}$ defined on an open interval I containing t_0 and t_1. We can compute the displacement of $\vec{r}$ on $[t_0, t_1]$, as shown in Figure 12.7(a). Recall that dividing the displacement vector by $t_1 - t_0$ gives the average rate of change on $[t_0, t_1]$, as shown in (b).

(a)

(b)
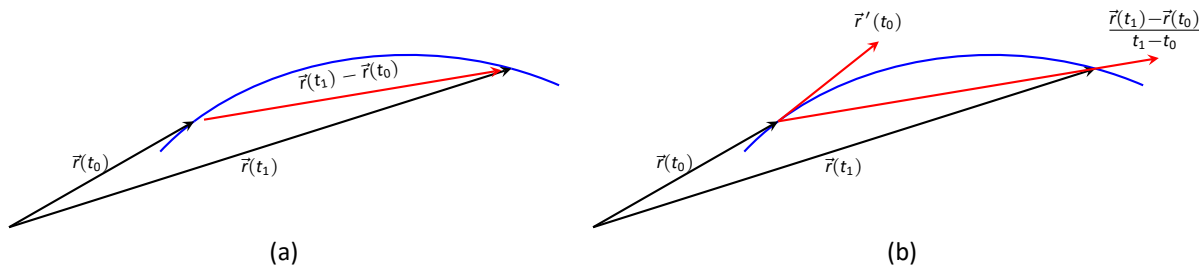
Figure 12.7: Illustrating displacement, leading to an understanding of the derivative of vector–valued functions.
   The **derivative** of a vector–valued function is a measure of the *instantaneous* rate of change, measured by taking the limit as the length of $[t_0, t_1]$ goes to 0. Instead of thinking of an interval as $[t_0, t_1]$, we think of it as $[c, c + h]$ for some

Notes:

value of h (hence the interval has length h). The *average* rate of change is

$$\frac{\vec{r}(c+h) - \vec{r}(c)}{h}$$

for any value of $h \neq 0$. We take the limit as $h \to 0$ to measure the instantaneous rate of change; this is the derivative of $\vec{r}$.

Definition 73   Derivative of a Vector–Valued Function

Let $\vec{r}(t)$ be continuous on an open interval I containing c.

1. The derivative of $\vec{r}$ at $t = c$ is

$$\vec{r}\,'(c) = \lim_{h \to 0} \frac{\vec{r}(c+h) - \vec{r}(c)}{h}.$$

2. The derivative of $\vec{r}$ is

$$\vec{r}\,'(t) = \lim_{h \to 0} \frac{\vec{r}(t+h) - \vec{r}(t)}{h}.$$

If a vector–valued function has a derivative for all c in an open interval I, we say that $\vec{r}(t)$ is **differentiable** on I.

Once again we might view this definition as intimidating, but recall that we can evaluate limits component–wise. The following theorem verifies that this means we can compute derivatives component–wise as well, making the task not too difficult.

Alternate notations for the derivative of $\vec{r}$ include:

$$\vec{r}\,'(t) = \frac{d}{dt}\big(\vec{r}(t)\big) = \frac{d\vec{r}}{dt}.$$

Theorem 93   Derivatives of Vector–Valued Functions

1. Let $\vec{r}(t) = \langle f(t), g(t) \rangle$. Then

$$\vec{r}\,'(t) = \langle f'(t), g'(t) \rangle.$$

2. Let $\vec{r}(t) = \langle f(t), g(t), h(t) \rangle$. Then

$$\vec{r}\,'(t) = \langle f'(t), g'(t), h'(t) \rangle.$$

Notes:
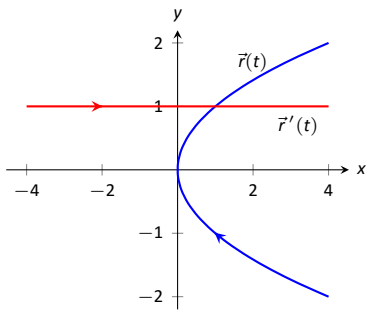
(a)
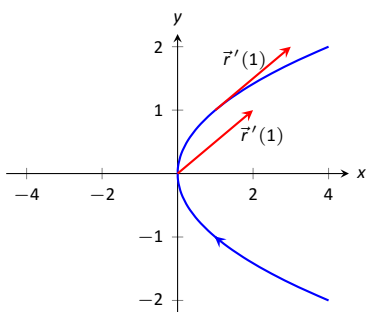
(b)

Figure 12.8: Graphing the derivative of a vector–valued function in Example 3.

Figure 12.9: Viewing a vector–valued function and its derivative at one point.

Watch the video:
Limit and Derivative of Vector Function at
`https://youtu.be/238wiC0U7NE`

Example 3   **Derivatives of vector–valued functions**
Let $\vec{r}(t) = \langle t^2, t \rangle$.

1. Sketch $\vec{r}(t)$ and $\vec{r}'(t)$ on the same axes.

2. Compute $\vec{r}'(1)$ and sketch this vector with its initial point at the origin and at $\vec{r}(1)$.

SOLUTION

1. Theorem 93 allows us to compute derivatives component–wise, so

$$\vec{r}'(t) = \langle 2t, 1 \rangle.$$

$\vec{r}(t)$ and $\vec{r}'(t)$ are graphed together in Figure 12.8(a). Note how plotting the two of these together, in this way, is not very illuminating. When dealing with real–valued functions, plotting $f(x)$ with $f'(x)$ gave us useful information as we were able to compare f and f' at the same x-values. When dealing with vector–valued functions, it is hard to tell which points on the graph of $\vec{r}'$ correspond to which points on the graph of $\vec{r}$.

2. We easily compute $\vec{r}'(1) = \langle 2, 1 \rangle$, which is drawn in Figure 12.8 with its initial point at the origin, as well as at $\vec{r}(1) = \langle 1, 1 \rangle$. These are sketched in Figure 12.8(b).

Example 4   **Derivatives of vector–valued functions**
Let $\vec{r}(t) = \langle \cos t, \sin t, t \rangle$. Compute $\vec{r}'(t)$ and $\vec{r}'(\pi/2)$. Sketch $\vec{r}'(\pi/2)$ with its initial point at the origin and at $\vec{r}(\pi/2)$.

SOLUTION   We compute $\vec{r}'$ as $\vec{r}'(t) = \langle -\sin t, \cos t, 1 \rangle$. At $t = \pi/2$, we have $\vec{r}'(\pi/2) = \langle -1, 0, 1 \rangle$. Figure 12.9 shows a graph of $\vec{r}(t)$, with $\vec{r}'(\pi/2)$ plotted with its initial point at the origin and at $\vec{r}(\pi/2)$.

Notes:

In Examples 3 and 4, sketching a particular derivative with its initial point at the origin did not seem to reveal anything significant. However, when we sketched the vector with its initial point on the corresponding point on the graph, we did see something significant: the vector appeared to be *tangent* to the graph. We have not yet defined what "tangent" means in terms of curves in space; in fact, we use the derivative to define this term.

Definition 74   Tangent Vector, Tangent Line

Let $\vec{r}(t)$ be a differentiable vector–valued function on an open interval I containing c, where $\vec{r}\,'(c) \neq \vec{0}$.

1. A vector $\vec{v}$ is **tangent to the graph of $\vec{r}(t)$ at $t = c$** if $\vec{v}$ is parallel to $\vec{r}\,'(c)$.

2. The **tangent line** to the graph of $\vec{r}(t)$ at $t = c$ is the line through $\vec{r}(c)$ with direction parallel to $\vec{r}\,'(c)$. An equation of the tangent line is
$$\vec{\ell}(t) = \vec{r}(c) + t\,\vec{r}\,'(c).$$

Example 5   Finding tangent lines to curves in space

Let $\vec{r}(t) = \langle t, t^2, t^3 \rangle$ on $[-1.5, 1.5]$. Find the vector equation of the line tangent to the graph of $\vec{r}$ at $t = -1$.

SOLUTION   To find the equation of a line, we need a point on the line and the line's direction. The point is given by $\vec{r}(-1) = \langle -1, 1, -1 \rangle$. (To be clear, $\langle -1, 1, -1 \rangle$ is a *vector*, not a point, but we use the point "pointed to" by this vector.)

The direction comes from $\vec{r}\,'(-1)$. We compute, component–wise, $\vec{r}\,'(t) = \langle 1, 2t, 3t^2 \rangle$. Thus $\vec{r}\,'(-1) = \langle 1, -2, 3 \rangle$.

The vector equation of the line is $\ell(t) = \langle -1, 1, -1 \rangle + t \langle 1, -2, 3 \rangle$. This line and $\vec{r}(t)$ are sketched in Figure 12.10.
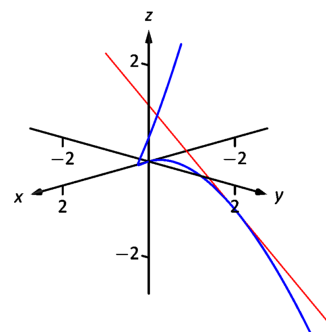
Figure 12.10: Graphing a curve in space with its tangent line.

Example 6   Finding tangent lines to curves

Find the equations of the lines tangent to $\vec{r}(t) = \langle t^3, t^2 \rangle$ at $t = -1$ and $t = 0$.

SOLUTION   We find that $\vec{r}\,'(t) = \langle 3t^2, 2t \rangle$. At $t = -1$, we have

$$\vec{r}(-1) = \langle -1, 1 \rangle \quad \text{and} \quad \vec{r}\,'(-1) = \langle 3, -2 \rangle,$$
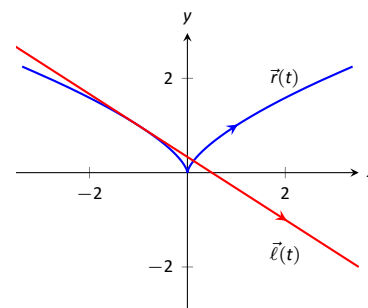
Figure 12.11: Graphing $\vec{r}(t)$ and its tangent line in Example 6.

Notes:

so the equation of the line tangent to the graph of $\vec{r}(t)$ at $t = -1$ is

$$\ell(t) = \langle -1, 1 \rangle + t \langle 3, -2 \rangle .$$

This line is graphed with $\vec{r}(t)$ in Figure 12.11.

At $t = 0$, we have $\vec{r}\,'(0) = \langle 0, 0 \rangle = \vec{0}$. This implies that the tangent line "has no direction." We cannot apply Definition 74, hence cannot find the equation of the tangent line.

We were unable to compute the equation of the tangent line to $\vec{r}(t) = \langle t^3, t^2 \rangle$ at $t = 0$ because $\vec{r}\,'(0) = \vec{0}$. The graph in Figure 12.11 shows that there is a cusp at this point. This leads us to another definition of **smooth**, previously defined by Definition 48 in Section 10.2.

Definition 75   Smooth Vector–Valued Functions
Let $\vec{r}(t)$ be a differentiable vector–valued function on an open interval I. Then $\vec{r}(t)$ is **smooth** on I if $\vec{r}\,'(t)$ is continuous and $\vec{r}\,'(t) \neq \vec{0}$ on I.

Having established derivatives of vector–valued functions, we now explore the relationships between the derivative and other vector operations. The following theorem states how the derivative interacts with vector addition and the various vector products.

Theorem 94   Properies of Derivatives of Vector–Valued Functions
Let $\vec{r}$ and $\vec{s}$ be differentiable vector–valued functions, let f be a differentiable real–valued function, and let c be a real number.

1. $\dfrac{d}{dt}\left(\vec{r}(t) \pm \vec{s}(t) \right) = \vec{r}\,'(t) \pm \vec{s}\,'(t)$

2. $\dfrac{d}{dt}\left(c\vec{r}(t) \right) = c\vec{r}\,'(t)$

3. $\dfrac{d}{dt}\left(f(t)\vec{r}(t) \right) = f'(t)\vec{r}(t) + f(t)\vec{r}\,'(t)$ **Product Rule**

4. $\dfrac{d}{dt}\left(\vec{r}(t) \cdot \vec{s}(t) \right) = \vec{r}\,'(t) \cdot \vec{s}(t) + \vec{r}(t) \cdot \vec{s}\,'(t)$ **Product Rule**

5. $\dfrac{d}{dt}\left(\vec{r}(t) \times \vec{s}(t) \right) = \vec{r}\,'(t) \times \vec{s}(t) + \vec{r}(t) \times \vec{s}\,'(t)$ **Product Rule**

6. $\dfrac{d}{dt}\left(\vec{r}(f(t)) \right) = \vec{r}\,'(f(t))f'(t)$ **Chain Rule**

Note: Because the order is important when computing a cross product, we must maintain the correct order of the functions in rule 5.

Notes:

Example 7   **Using derivative properties of vector–valued functions**

Let $\vec{r}(t) = \langle t, t^2 - 1 \rangle$ and let $\vec{u}(t)$ be the unit vector that points in the direction of $\vec{r}(t)$.

1. Graph $\vec{r}(t)$ and $\vec{u}(t)$ on the same axes, on $[-2, 2]$.

2. Find $\vec{u}\,'(t)$ and sketch $\vec{u}\,'(-2)$, $\vec{u}\,'(-1)$ and $\vec{u}\,'(0)$. Sketch each with initial point the corresponding point on the graph of $\vec{u}$.

SOLUTION

1. To form the unit vector that points in the direction of $\vec{r}$, we need to divide $\vec{r}(t)$ by its magnitude.

$$\|\vec{r}(t)\| = \sqrt{t^2 + (t^2 - 1)^2} \quad \Rightarrow \quad \vec{u}(t) = \frac{1}{\sqrt{t^2 + (t^2 - 1)^2}} \langle t, t^2 - 1 \rangle .$$

$\vec{r}(t)$ and $\vec{u}(t)$ are graphed in Figure 12.12. Note how the graph of $\vec{u}(t)$ forms part of a circle; this must be the case, as the length of $\vec{u}(t)$ is 1 for all t.

2. To compute $\vec{u}\,'(t)$, we use Theorem 94, writing

$$\vec{u}(t) = f(t)\vec{r}(t), \quad \text{where} \quad f(t) = \frac{1}{\sqrt{t^2 + (t^2 - 1)^2}} = \left(t^2 + (t^2 - 1)^2\right)^{-1/2}.$$

(We *could* write

$$\vec{u}(t) = \left\langle \frac{t}{\sqrt{t^2 + (t^2 - 1)^2}}, \frac{t^2 - 1}{\sqrt{t^2 + (t^2 - 1)^2}} \right\rangle$$

and then take the derivative. It is a matter of preference; this latter method requires two applications of the Quotient Rule where our method uses the Product and Chain Rules.)

We find $f'(t)$ using the Chain Rule:

$$f'(t) = -\frac{1}{2}\left(t^2 + (t^2 - 1)^2\right)^{-3/2}\left(2t + 2(t^2 - 1)(2t)\right)$$

$$= -\frac{2t(2t^2 - 1)}{2\left(\sqrt{t^2 + (t^2 - 1)^2}\right)^3}$$

We now find $\vec{u}\,'(t)$ using part 3 of Theorem 94:

$$\vec{u}\,'(t) = f'(t)\vec{u}(t) + f(t)\vec{u}\,'(t)$$

$$= -\frac{2t(2t^2 - 1)}{2\left(\sqrt{t^2 + (t^2 - 1)^2}\right)^3} \langle t, t^2 - 1 \rangle + \frac{1}{\sqrt{t^2 + (t^2 - 1)^2}} \langle 1, 2t \rangle .$$
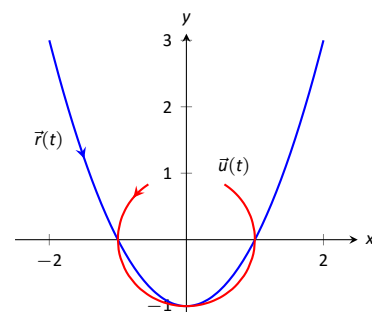
Figure 12.12: Graphing $\vec{r}(t)$ and $\vec{u}(t)$ in Example 7.

Notes:

This is admittedly very "messy;" such is usually the case when we deal with unit vectors. We can use this formula to compute $\vec{u}\,'(-2)$, $\vec{u}\,'(-1)$ and $\vec{u}\,'(0)$:

$$\vec{u}\,'(-2) = \left\langle -\frac{15}{13\sqrt{13}}, -\frac{10}{13\sqrt{13}} \right\rangle$$
$$\vec{u}\,'(-1) = \langle 0, -2 \rangle$$
$$\vec{u}\,'(0) = \langle 1, 0 \rangle$$

Each of these is sketched in Figure 12.13. Note how the length of the vector gives an indication of how quickly the circle is being traced at that point. When $t = -2$, the circle is being drawn relatively slow; when $t = -1$, the circle is being traced much more quickly.

Figure 12.13:   Graphing some of the derivatives of $\vec{u}(t)$ in Example 7.

It is a basic geometric fact that a line tangent to a circle at a point P is perpendicular to the line passing through the center of the circle and P. This is illustrated in Figure 12.13; each tangent vector is perpendicular to the line that passes through its initial point and the center of the circle. Since the center of the circle is the origin, we can state this another way: $\vec{u}\,'(t)$ is orthogonal to $\vec{u}(t)$.

Recall that the dot product serves as a test for orthogonality: if $\vec{u} \cdot \vec{v} = 0$, then $\vec{u}$ is orthogonal to $\vec{v}$. Thus in the above example, $\vec{u}(t) \cdot \vec{u}\,'(t) = 0$.

This is true of any vector–valued function that has a constant length, that is, that traces out part of a circle. It has important implications later on, so we state it as a theorem (and leave its formal proof as Exercise 42.)

Theorem 95   Vector–Valued Functions of Constant Length

Let $\vec{r}(t)$ be a differentiable vector–valued function on an open interval I of constant length. That is, $\|\vec{r}(t)\| = c$ for all t in I (equivalently, $\vec{r}(t) \cdot \vec{r}(t) = c^2$ for all t in I). Then $\vec{r}(t) \cdot \vec{r}\,'(t) = 0$ for all t in I.

Integration

Indefinite and definite integrals of vector–valued functions are also evaluated component–wise.

Notes:

Theorem 96 **Indefinite and Definite Integrals of Vector–Valued Functions**

Let $\vec{r}(t) = \langle f(t), g(t) \rangle$ be a vector–valued function in $\mathbb{R}^2$.

1. $\int \vec{r}(t) \, dt = \left\langle \int f(t) \, dt, \int g(t) \, dt \right\rangle$

2. $\int_a^b \vec{r}(t) \, dt = \left\langle \int_a^b f(t) \, dt, \int_a^b g(t) \, dt \right\rangle$

Let $\vec{r}(t) = \langle f(t), g(t), h(t) \rangle$ be a vector–valued function in $\mathbb{R}^3$.

1. $\int \vec{r}(t) \, dt = \left\langle \int f(t) \, dt, \int g(t) \, dt, \int h(t) \, dt \right\rangle$

2. $\int_a^b \vec{r}(t) \, dt = \left\langle \int_a^b f(t) \, dt, \int_a^b g(t) \, dt, \int_a^b h(t) \, dt \right\rangle$

Example 8 **Evaluating a definite integral of a vector–valued function**

Let $\vec{r}(t) = \left\langle e^{2t}, \sin t \right\rangle$. Evaluate $\int_0^1 \vec{r}(t) \, dt$.

SOLUTION We follow Theorem 96.

$$\int_0^1 \vec{r}(t) \, dt = \int_0^1 \left\langle e^{2t}, \sin t \right\rangle \, dt$$

$$= \left\langle \int_0^1 e^{2t} \, dt, \int_0^1 \sin t \, dt \right\rangle$$

$$= \left\langle \left. \frac{1}{2} e^{2t} \right|_0^1, \left. -\cos t \right|_0^1 \right\rangle$$

$$= \left\langle \frac{1}{2}(e^2 - 1), -\cos(1) + 1 \right\rangle.$$

Example 9 **Solving an initial value problem**

Let $\vec{r}''(t) = \langle 2, \cos t, 12t \rangle$. Find $\vec{r}(t)$ where:

- $\vec{r}(0) = \langle -7, -1, 2 \rangle$ and
- $\vec{r}'(0) = \langle 5, 3, 0 \rangle$.

Notes:

SOLUTION   Knowing $\vec{r}''(t) = \langle 2, \cos t, 12t\rangle$, we find $\vec{r}'(t)$ by evaluating the indefinite integral.

$$\int \vec{r}''(t)\, dt = \left\langle \int 2\, dt \,, \int \cos t\, dt \,, \int 12t\, dt \right\rangle$$
$$= \left\langle 2t + C_1, \sin t + C_2, 6t^2 + C_3 \right\rangle$$
$$= \left\langle 2t, \sin t, 6t^2 \right\rangle + \langle C_1, C_2, C_3 \rangle$$
$$= \left\langle 2t, \sin t, 6t^2 \right\rangle + \vec{C}.$$

Note how each indefinite integral creates its own constant which we collect as one constant vector $\vec{C}$. Knowing $\vec{r}'(0) = \langle 5, 3, 0\rangle$ allows us to solve for $\vec{C}$:

$$\vec{r}'(t) = \left\langle 2t, \sin t, 6t^2 \right\rangle + \vec{C}$$
$$\vec{r}'(0) = \langle 0, 0, 0 \rangle + \vec{C}$$
$$\langle 5, 3, 0 \rangle = \vec{C}.$$

So $\vec{r}'(t) = \left\langle 2t, \sin t, 6t^2 \right\rangle + \langle 5, 3, 0 \rangle = \left\langle 2t + 5, \sin t + 3, 6t^2 \right\rangle$. To find $\vec{r}(t)$, we integrate once more.

$$\int \vec{r}'(t)\, dt = \left\langle \int 2t + 5\, dt, \int \sin t + 3\, dt, \int 6t^2\, dt \right\rangle$$
$$= \left\langle t^2 + 5t, -\cos t + 3t, 2t^3 \right\rangle + \vec{C}.$$

With $\vec{r}(0) = \langle -7, -1, 2\rangle$, we solve for $\vec{C}$:

$$\vec{r}(t) = \left\langle t^2 + 5t, -\cos t + 3t, 2t^3 \right\rangle + \vec{C}$$
$$\vec{r}(0) = \langle 0, -1, 0 \rangle + \vec{C}$$
$$\langle -7, -1, 2 \rangle = \langle 0, -1, 0 \rangle + \vec{C}$$
$$\langle -7, 0, 2 \rangle = \vec{C}.$$

Therefore,

$$\vec{r}(t) = \left\langle t^2 + 5t, -\cos t + 3t, 2t^3 \right\rangle + \langle -7, 0, 2 \rangle$$
$$= \left\langle t^2 + 5t - 7, -\cos t + 3t, 2t^3 + 2 \right\rangle.$$

What does the integration of a vector–valued function *mean*? There are many applications, but none as direct as "the area under the curve" that we used in understanding the integral of a real–valued function.

Notes:

A key understanding for us comes from considering the integral of a derivative:

$$\int_a^b \vec{r}\,'(t)\ dt = \vec{r}(t)\Big|_a^b = \vec{r}(b) - \vec{r}(a).$$

Integrating a *rate of change* function gives *displacement*.

Noting that vector–valued functions are closely related to parametric equations, we can describe the arc length of the graph of a vector–valued function as an integral. Given parametric equations $x = f(t)$, $y = g(t)$, the arc length on $[a, b]$ of the graph is

$$\text{Arc Length} = \int_a^b \sqrt{f'(t)^2 + g'(t)^2}\ dt,$$

as stated in Theorem 84 in Section 10.3. If $\vec{r}(t) = \langle f(t), g(t)\rangle$, note that $\sqrt{f'(t)^2 + g'(t)^2} = \|\vec{r}\,'(t)\|$. Therefore we can express the arc length of the graph of a vector–valued function as an integral of the magnitude of its derivative.

Theorem 97   Arc Length of a Vector–Valued Function

Let $\vec{r}(t)$ be a vector–valued function where $\vec{r}\,'(t)$ is continuous on $[a, b]$. The arc length L of the graph of $\vec{r}(t)$ is

$$L = \int_a^b \|\vec{r}\,'(t)\|\ dt.$$

Note that we are actually integrating a scalar–function here, not a vector–valued function.

The next section takes what we have established thus far and applies it to objects in motion. We will let $\vec{r}(t)$ describe the path of an object in the plane or in space and will discover the information provided by $\vec{r}\,'(t)$ and $\vec{r}\,''(t)$.

Notes:

Exercises 12.2

Terms and Concepts

1. Limits, derivatives and integrals of vector–valued functions are all evaluated _____–wise.

2. The definite integral of a rate of change function gives _____.

3. Why is it generally not useful to graph both $\vec{r}(t)$ and $\vec{r}\,'(t)$ on the same axes?

Problems

In Exercises 4–7, evaluate the given limit.

4. $\lim\limits_{t \to 5} \langle 2t + 1, 3t^2 - 1, \sin t \rangle$

5. $\lim\limits_{t \to 3} \left\langle e^t, \dfrac{t^2 - 9}{t + 3} \right\rangle$

6. $\lim\limits_{t \to 0} \left\langle \dfrac{t}{\sin t}, (1 + t)^{\frac{1}{t}} \right\rangle$

7. $\lim\limits_{h \to 0} \dfrac{\vec{r}(t + h) - \vec{r}(t)}{h}$, where $\vec{r}(t) = \langle t^2, t, 1 \rangle$.

In Exercises 8–9, identify the interval(s) on which $\vec{r}(t)$ is continuous.

8. $\vec{r}(t) = \langle t^2, 1/t \rangle$

9. $\vec{r}(t) = \langle \cos t, e^t, \ln t \rangle$

In Exercises 10–14, find the derivative of the given function.

10. $\vec{r}(t) = \langle \cos t, e^t, \ln t \rangle$

11. $\vec{r}(t) = \left\langle \dfrac{1}{t}, \dfrac{2t - 1}{3t + 1}, \tan t \right\rangle$

12. $\vec{r}(t) = (t^2) \langle \sin t, 2t + 5 \rangle$

13. $\vec{r}(t) = \langle t^2 + 1, t - 1 \rangle \cdot \langle \sin t, 2t + 5 \rangle$

14. $\vec{r}(t) = \langle t^2 + 1, t - 1, 1 \rangle \times \langle \sin t, 2t + 5, 1 \rangle$

In Exercises 15–18, find $\vec{r}\,'(t)$. Sketch $\vec{r}(t)$ and $\vec{r}\,'(1)$, with the initial point of $\vec{r}\,'(1)$ at $\vec{r}(1)$.

15. $\vec{r}(t) = \langle t^2 + t, t^2 - t \rangle$

16. $\vec{r}(t) = \langle t^2 - 2t + 2, t^3 - 3t^2 + 2t \rangle$

17. $\vec{r}(t) = \langle t^2 + 1, t^3 - t \rangle$

18. $\vec{r}(t) = \langle t^2 - 4t + 5, t^3 - 6t^2 + 11t - 6 \rangle$

In Exercises 19–22, give the equation of the line tangent to the graph of $\vec{r}(t)$ at the given t value.

19. $\vec{r}(t) = \langle t^2 + t, t^2 - t \rangle$ at $t = 1$.

20. $\vec{r}(t) = \langle 3\cos t, \sin t \rangle$ at $t = \pi/4$.

21. $\vec{r}(t) = \langle 3\cos t, 3\sin t, t \rangle$ at $t = \pi$.

22. $\vec{r}(t) = \langle e^t, \tan t, t \rangle$ at $t = 0$.

In Exercises 23–26, find the value(s) of t for which $\vec{r}(t)$ is not smooth.

23. $\vec{r}(t) = \langle \cos t, \sin t - t \rangle$

24. $\vec{r}(t) = \langle t^2 - 2t + 1, t^3 + t^2 - 5t + 3 \rangle$

25. $\vec{r}(t) = \langle \cos t - \sin t, \sin t - \cos t, \cos(4t) \rangle$

26. $\vec{r}(t) = \langle t^3 - 3t + 2, -\cos(\pi t), \sin^2(\pi t) \rangle$

Exercises 27–29 ask you to verify parts of Theorem 94. In each let $f(t) = t^3$, $\vec{r}(t) = \langle t^2, t - 1, 1 \rangle$ and $\vec{s}(t) = \langle \sin t, e^t, t \rangle$. Compute the various derivatives as indicated.

27. Simplify $f(t)\vec{r}(t)$, then find its derivative; show this is the same as $f'(t)\vec{r}(t) + f(t)\vec{r}\,'(t)$.

28. Simplify $\vec{r}(t) \cdot \vec{s}(t)$, then find its derivative; show this is the same as $\vec{r}\,'(t) \cdot \vec{s}(t) + \vec{r}(t) \cdot \vec{s}\,'(t)$.

29. Simplify $\vec{r}(t) \times \vec{s}(t)$, then find its derivative; show this is the same as $\vec{r}\,'(t) \times \vec{s}(t) + \vec{r}(t) \times \vec{s}\,'(t)$.

In Exercises 30–33, evaluate the given definite or indefinite integral.

30. $\displaystyle\int \langle t^3, \cos t, te^t \rangle \; dt$

31. $\displaystyle\int \left\langle \dfrac{1}{1 + t^2}, \sec^2 t \right\rangle \; dt$

32. $\displaystyle\int_0^\pi \langle -\sin t, \cos t \rangle \; dt$

33. $\displaystyle\int_{-2}^2 \langle 2t + 1, 2t - 1 \rangle \; dt$

In Exercises 34–37, solve the given initial value problems.

34. Find $\vec{r}(t)$, given that $\vec{r}\,'(t) = \langle t, \sin t \rangle$ and $\vec{r}(0) = \langle 2, 2 \rangle$.

35. Find $\vec{r}(t)$, given that $\vec{r}\,'(t) = \langle 1/(t + 1), \tan t \rangle$ and $\vec{r}(0) = \langle 1, 2 \rangle$.

36. Find $\vec{r}(t)$, given that $\vec{r}\,''(t) = \langle t^2, t, 1 \rangle$, $\vec{r}\,'(0) = \langle 1, 2, 3 \rangle$ and $\vec{r}(0) = \langle 4, 5, 6 \rangle$.

37. Find $\vec{r}(t)$, given that $\vec{r}\,''(t) = \langle \cos t, \sin t, e^t \rangle$, $\vec{r}\,'(0) = \langle 0, 0, 0 \rangle$ and $\vec{r}(0) = \langle 0, 0, 0 \rangle$.

In Exercises 38–41, find the arc length of $\vec{r}(t)$ on the indicated interval.

38. $\vec{r}(t) = \langle 2\cos t, 2\sin t, 3t \rangle$ on $[0, 2\pi]$.

39. $\vec{r}(t) = \langle 5\cos t, 3\sin t, 4\sin t \rangle$ on $[0, 2\pi]$.

40. $\vec{r}(t) = \langle t^3, t^2, t^3 \rangle$ on $[0, 1]$.

41. $\vec{r}(t) = \langle e^{-t}\cos t, e^{-t}\sin t \rangle$ on $[0, 1]$.

42. Prove Theorem 95; that is, show if $\vec{r}(t)$ has constant length and is differentiable, then $\vec{r}(t) \cdot \vec{r}\,'(t) = 0$. (Hint: use the Product Rule to compute $\frac{d}{dt}\big(\vec{r}(t) \cdot \vec{r}(t)\big)$.)

12.3   The Calculus of Motion

A common use of vector–valued functions is to describe the motion of an object in the plane or in space. A **position function** $\vec{r}(t)$ gives the position of an object at **time** t. This section explores how derivatives and integrals are used to study the motion described by such a function.

Definition 76   Velocity, Speed and Acceleration

Let $\vec{r}(t)$ be a position function in $\mathbb{R}^2$ or $\mathbb{R}^3$.

1. **Velocity**, denoted $\vec{v}(t)$, is the instantaneous rate of position change; that is, $\vec{v}(t) = \vec{r}\,'(t)$.

2. **Speed** is the magnitude of velocity, $\|\vec{v}(t)\|$.

3. **Acceleration**, denoted $\vec{a}(t)$, is the instantaneous rate of velocity change; that is, $\vec{a}(t) = \vec{v}\,'(t) = \vec{r}\,''(t)$.

Watch the video:
Example of Position, Velocity and Acceleration in Three Space at
`https://youtu.be/gD2R4Jqw6dQ`

Example 1   Finding velocity and acceleration

An object is moving with position function $\vec{r}(t) = \left\langle t^2 - t, t^2 + t \right\rangle$, $-3 \leq t \leq 3$, where distances are measured in feet and time is measured in seconds.

1. Find $\vec{v}(t)$ and $\vec{a}(t)$.

2. Sketch $\vec{r}(t)$; plot $\vec{v}(-1)$, $\vec{a}(-1)$, $\vec{v}(1)$ and $\vec{a}(1)$, each with their initial point at their corresponding point on the graph of $\vec{r}(t)$.

3. When is the object's speed minimized?

Solution

Notes:

1. Taking derivatives, we find

$$\vec{v}(t) = \vec{r}\,'(t) = \langle 2t - 1, 2t + 1 \rangle \quad \text{and} \quad \vec{a}(t) = \vec{r}\,''(t) = \langle 2, 2 \rangle.$$

Note that acceleration is constant.

2. $\vec{v}(-1) = \langle -3, -1 \rangle$, $\vec{a}(-1) = \langle 2, 2 \rangle$; $\vec{v}(1) = \langle 1, 3 \rangle$, $\vec{a}(1) = \langle 2, 2 \rangle$. These are plotted with $\vec{r}(t)$ in Figure 12.14(a).

We can think of acceleration as "pulling" the velocity vector in a certain direction. At $t = -1$, the velocity vector points down and to the left; at $t = 1$, the velocity vector has been pulled in the $\langle 2, 2 \rangle$ direction and is now pointing up and to the right. In Figure 12.14(b) we plot more velocity/acceleration vectors, making more clear the effect acceleration has on velocity.

Since $\vec{a}(t)$ is constant in this example, as t grows large $\vec{v}(t)$ becomes almost parallel to $\vec{a}(t)$. For instance, when $t = 10$, $\vec{v}(10) = \langle 19, 21 \rangle$, which is nearly parallel to $\langle 2, 2 \rangle$.

3. The object's speed is given by

$$\|\vec{v}(t)\| = \sqrt{(2t-1)^2 + (2t+1)^2} = \sqrt{8t^2 + 2}.$$

To find the minimal speed, we could apply calculus techniques (such as set the derivative equal to 0 and solve for t, etc.) but we can find it by inspection. Inside the square root we have a quadratic which is minimized when $t = 0$. Thus the speed is minimized at $t = 0$, with a speed of $\sqrt{2}$ ft/s.

The graph in Figure 12.14(b) also implies speed is minimized here. The filled dots on the graph are located at integer values of t between -3 and 3. Dots that are far apart imply the object traveled a far distance in 1 second, indicating high speed; dots that are close together imply the object did not travel far in 1 second, indicating a low speed. The dots are closest together near $t = 0$, implying the speed is minimized near that value.

(a)

(b)

Figure 12.14: Graphing the position, velocity and acceleration of an object in Example 1.

Example 2   Analyzing Motion

Two objects follow an identical path at different rates on $[-1, 1]$. The position function for Object 1 is $\vec{r}_1(t) = \langle t, t^2 \rangle$; the position function for Object 2 is $\vec{r}_2(t) = \langle t^3, t^6 \rangle$, where distances are measured in feet and time is measured in seconds. Compare the velocity, speed and acceleration of the two objects on the path.

Notes:

SOLUTION We begin by computing the velocity and acceleration function for each object:

$$\vec{v}_1(t) = \langle 1, 2t \rangle \qquad\qquad \vec{v}_2(t) = \langle 3t^2, 6t^5 \rangle$$
$$\vec{a}_1(t) = \langle 0, 2 \rangle \qquad\qquad \vec{a}_2(t) = \langle 6t, 30t^4 \rangle$$

We immediately see that Object 1 has constant acceleration, whereas Object 2 does not.

At $t = -1$, we have $\vec{v}_1(-1) = \langle 1, -2 \rangle$ and $\vec{v}_2(-1) = \langle 3, -6 \rangle$; the velocity of Object 2 is three times that of Object 1 and so it follows that the speed of Object 2 is three times that of Object 1 ($3\sqrt{5}$ ft/s compared to $\sqrt{5}$ ft/s.)

At $t = 0$, the velocity of Object 1 is $\vec{v}(1) = \langle 1, 0 \rangle$ and the velocity of Object 2 is $\vec{0}$. This tells us that Object 2 comes to a complete stop at $t = 0$.

In Figure 12.15, we see the velocity and acceleration vectors for Object 1 plotted for $t = -1, -1/2, 0, 1/2$ and $t = 1$. Note again how the constant acceleration vector seems to "pull" the velocity vector from pointing down, right to up, right. We could plot the analogous picture for Object 2, but the velocity and acceleration vectors are rather large ($\vec{a}_2(-1) = \langle -6, 30 \rangle$).

Instead, we simply plot the locations of Object 1 and 2 on intervals of $1/5^{\text{th}}$ of a second, shown in Figure 12.16(a) and (b). Note how the x-values of Object 1 increase at a steady rate. This is because the x-component of $\vec{a}(t)$ is 0; there is no acceleration in the x-component. The dots are not evenly spaced; the object is moving faster near $t = -1$ and $t = 1$ than near $t = 0$.

In part (b) of the Figure, we see the points plotted for Object 2. Note the large change in position from $t = -1$ to $t = -0.8$; the object starts moving very quickly. However, it slows considerably at it approaches the origin, and comes to a complete stop at $t = 0$. While it looks like there are 3 points near the origin, there are in reality 5 points there.

Since the objects begin and end at the same location, they have the same displacement. Since they begin and end at the same time, with the same displacement, they have the same average rate of change (i.e, they have the same average velocity). Since they follow the same path, they have the same distance traveled. Even though these three measurements are the same, the objects obviously travel the path in very different ways.

Example 3 Analyzing the motion of a whirling ball on a string

A young boy whirls a ball, attached to a string, above his head in a counterclockwise circle. The ball follows a circular path and makes 2 revolutions per second. The string has length 2ft.

1. Find the position function $\vec{r}(t)$ that describes this situation.

Notes:
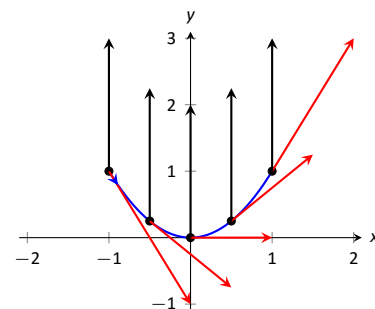
Figure 12.15: Plotting velocity and acceleration vectors for Object 1 in Example 2.
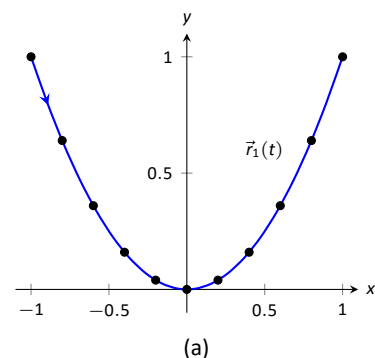
(a)
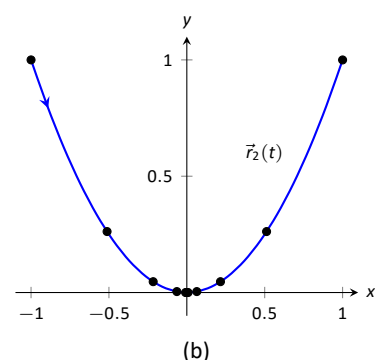
(b)

Figure 12.16: Comparing the positions of Objects 1 and 2 in Example 2.

2. Find the acceleration of the ball and derive a physical interpretation of it.

3. A tree stands 10ft in front of the boy. At what *t*-values should the boy release the string so that the ball hits the tree?

SOLUTION

1. The ball whirls in a circle. Since the string is 2ft long, the radius of the circle is 2. The position function $\vec{r}(t) = \langle 2\cos t, 2\sin t\rangle$ describes a circle with radius 2, centered at the origin, but makes a full revolution every 2π seconds, not two revolutions per second. We modify the period of the trigonometric functions to be 1/2 by multiplying t by 4π. The final position function is thus

$$\vec{r}(t) = \langle 2\cos(4\pi t), 2\sin(4\pi t)\rangle .$$

(Plot this for $0 \le t \le 1/2$ to verify that one revolution is made in 1/2 a second.)

2. To find $\vec{a}(t)$, we derive $\vec{r}(t)$ twice.

$$\vec{v}(t) = \vec{r}\,'(t) = \langle -8\pi\sin(4\pi t), 8\pi\cos(4\pi t)\rangle$$
$$\vec{a}(t) = \vec{r}\,''(t) = \langle -32\pi^2\cos(4\pi t), -32\pi^2\sin(4\pi t)\rangle$$
$$= -32\pi^2\langle\cos(4\pi t), \sin(4\pi t)\rangle .$$

Note how $\vec{a}(t)$ is parallel to $\vec{r}(t)$, but has a different magnitude and points in the opposite direction. Why is this?

Recall the classic physics equation, "Force = mass × acceleration." A force acting on a mass induces acceleration (i.e., the mass moves); acceleration acting on a mass induces a force (gravity gives our mass a *weight*). Thus force and acceleration are closely related. A moving ball "wants" to travel in a straight line. Why does the ball in our example move in a circle? It is attached to the boy's hand by a string. The string applies a force to the ball, affecting it's motion: the string *accelerates* the ball. This is not acceleration in the sense of "it travels faster;" rather, this acceleration is changing the velocity of the ball. In what direction is this force/acceleration being applied? In the direction of the string, towards the boy's hand.

The magnitude of the acceleration is related to the speed at which the ball is traveling. A ball whirling quickly is rapidly changing direction/velocity. When velocity is changing rapidly, the acceleration must be "large."

3. When the boy releases the string, the string no longer applies a force to the ball, meaning acceleration is $\vec{0}$ and the ball can now move in a straight line in the direction of $\vec{v}(t)$.

Notes:

Let $t = t_0$ be the time when the boy lets go of the string. The ball will be at $\vec{r}(t_0)$, traveling in the direction of $\vec{v}(t_0)$. We want to find t_0 so that this line contains the point $(0, 10)$ (since the tree is 10ft directly in front of the boy).

There are many ways to find this time value. We choose one that is relatively simple computationally. As shown in Figure 12.17, the vector from the release point to the tree is $\langle 0, 10 \rangle - \vec{r}(t_0)$. This line segment is tangent to the circle, which means it is also perpendicular to $\vec{r}(t_0)$ itself, so their dot product is 0.

$$\vec{r}(t_0) \cdot \big(\langle 0, 10 \rangle - \vec{r}(t_0) \big) = 0$$
$$\langle 2\cos(4\pi t_0), 2\sin(4\pi t_0) \rangle \cdot \langle -2\cos(4\pi t_0), 10 - 2\sin(4\pi t_0) \rangle = 0$$
$$-4\cos^2(4\pi t_0) + 20\sin(4\pi t_0) - 4\sin^2(4\pi t_0) = 0$$
$$20\sin(4\pi t_0) - 4 = 0$$
$$\sin(4\pi t_0) = 1/5$$
$$4\pi t_0 = \sin^{-1}(1/5)$$
$$4\pi t_0 \approx 0.2 + 2\pi n$$
$$t_0 \approx 0.016 + n/2$$

where n is an integer. This means that every 1/2 second after $t = 0.016$s the boy can release the string (since the ball makes 2 revolutions per second, he has two chances each second to release the ball).
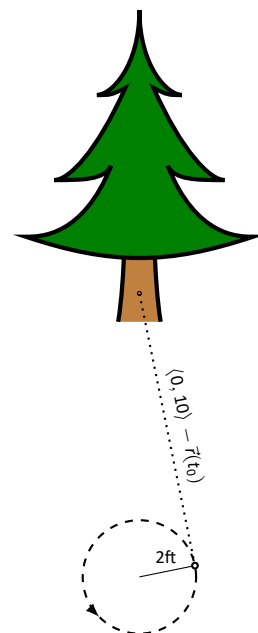
Figure 12.17: Modeling the flight of a ball in Example 3.

Example 4 Analyzing motion in space

An object moves in a spiral with position function $\vec{r}(t) = \langle \cos t, \sin t, t \rangle$, where distances are measured in meters and time is in minutes. Describe the object's speed and acceleration at time t.

SOLUTION   With $\vec{r}(t) = \langle \cos t, \sin t, t \rangle$, we have:

$$\vec{v}(t) = \langle -\sin t, \cos t, 1 \rangle \quad \text{and}$$
$$\vec{a}(t) = \langle -\cos t, -\sin t, 0 \rangle .$$

The speed of the object is $\|\vec{v}(t)\| = \sqrt{(-\sin t)^2 + \cos^2 t + 1} = \sqrt{2}$m/min; it moves at a constant speed. Note that the object does not accelerate in the z-direction, but rather moves up at a constant rate of 1m/min.

Notes:

The objects in Examples 3 and 4 traveled at a constant speed. That is, $\|\vec{v}(t)\| = c$ for some constant c. Recall Theorem 95, which states that if a vector–valued function $\vec{r}(t)$ has constant length, then $\vec{r}(t)$ is perpendicular to its derivative: $\vec{r}(t) \cdot \vec{r}\,'(t) = 0$. In these examples, the velocity function has constant length, therefore we can conclude that the velocity is perpendicular to the acceleration: $\vec{v}(t) \cdot \vec{a}(t) = 0$. A quick check verifies this.

There is an intuitive understanding of this. If acceleration is parallel to velocity, then it is only affecting the object's speed; it does not change the direction of travel. (For example, consider a dropped stone. Acceleration and velocity are parallel – straight down – and the direction of velocity never changes, though speed does increase.) If acceleration is not perpendicular to velocity, then there is some acceleration in the direction of travel, influencing the speed. If speed is constant, then acceleration must be orthogonal to velocity, as it then only affects direction, and not speed.

Key Idea 57   Objects With Constant Speed

If an object moves with constant speed, then its velocity and acceleration vectors are orthogonal. That is, $\vec{v}(t) \cdot \vec{a}(t) = 0$.

Projectile Motion

An important application of vector–valued position functions is *projectile motion*: the motion of objects under only the influence of gravity. We will measure time in seconds, and distances will either be in meters or feet. We will show that we can completely describe the path of such an object knowing its initial position and initial velocity (i.e., where it *is* and where it *is going*.)

Suppose an object has initial position $\vec{r}(0) = \langle x_0, y_0 \rangle$ and initial velocity $\vec{v}(0) = \langle v_x, v_y \rangle$. It is customary to rewrite $\vec{v}(0)$ in terms of its speed v_0 and direction $\vec{u}$, where $\vec{u}$ is a unit vector. Recall all unit vectors in $\mathbb{R}^2$ can be written as $\langle \cos\theta, \sin\theta \rangle$, where θ is an angle measure counter–clockwise from the x-axis. (We refer to θ as the **angle of elevation.**) Thus $\vec{v}(0) = v_0 \langle \cos\theta, \sin\theta \rangle$.

Since the acceleration of the object is known, namely $\vec{a}(t) = \langle 0, -g \rangle$, where g is the gravitational constant, we can find $\vec{r}(t)$ knowing our two initial conditions. We first find $\vec{v}(t)$:

Note: In this text we use $g = 32\text{ft/s}^2$ when using Imperial units, and $g = 9.8\text{m/s}^2$ when using SI units.

Notes:

$$\vec{v}(t) = \int \vec{a}(t)\, dt$$

$$\vec{v}(t) = \int \langle 0, -g \rangle\, dt$$

$$\vec{v}(t) = \langle 0, -gt \rangle + \vec{C}.$$

Knowing $\vec{v}(0) = v_0 \langle \cos\theta, \sin\theta \rangle$, we have $\vec{C} = v_0 \langle \cos\theta, \sin\theta \rangle$ and so

$$\vec{v}(t) = \langle v_0 \cos\theta, -gt + v_0 \sin\theta \rangle.$$

We integrate once more to find $\vec{r}(t)$:

$$\vec{r}(t) = \int \vec{v}(t)\, dt$$

$$\vec{r}(t) = \int \langle v_0 \cos\theta, -gt + v_0 \sin\theta \rangle\, dt$$

$$\vec{r}(t) = \left\langle \left(v_0 \cos\theta\right)t, -\frac{1}{2}gt^2 + \left(v_0 \sin\theta\right)t \right\rangle + \vec{C}.$$

Knowing $\vec{r}(0) = \langle x_0, y_0 \rangle$, we conclude $\vec{C} = \langle x_0, y_0 \rangle$ and

$$\vec{r}(t) = \left\langle \left(v_0 \cos\theta\right)t + x_0, -\frac{1}{2}gt^2 + \left(v_0 \sin\theta\right)t + y_0 \right\rangle. \qquad (12.1)$$

We demonstrate how to use this position function in the next two examples.

Example 5   **Projectile Motion**
Sydney shoots her Red Ryder® bb gun across level ground from an elevation of 4ft, where the barrel of the gun makes a 5° angle with the horizontal. Find how far the bb travels before landing, assuming the bb is fired at the advertised rate of 350ft/s and ignoring air resistance.

Solution   A direct application of Equation (12.1) gives

$$\vec{r}(t) = \langle (350\cos 5°)t, -16t^2 + (350\sin 5°)t + 4 \rangle$$
$$\approx \langle 346.67t, -16t^2 + 30.50t + 4 \rangle,$$

where we set her initial position to be $\langle 0, 4 \rangle$. We need to find *when* the bb lands, then we can find *where*. We accomplish this by setting the *y*-component equal to 0 and solving for *t*:

$$-16t^2 + 30.50t + 4 = 0$$

$$t = \frac{-30.50 \pm \sqrt{30.50^2 - 4(-16)(4)}}{-32}$$

$$t \approx 2.03s.$$

Notes:

(We discarded a negative solution that resulted from our quadratic equation.)

We have found that the bb lands 2.03s after firing; with $t = 2.03$, we find the *x*-component of our position function is $346.67(2.03) = 703.74$ft. The bb lands about 704 feet away.

Example 6   **Projectile Motion**

Alex holds his sister's bb gun at a height of 3ft and wants to shoot a target that is 6ft above the ground, 25ft away. At what angle should he hold the gun to hit his target? (We still assume the muzzle velocity is 350ft/s.)

SOLUTION   The position function for the path of Alex's bb is

$$\vec{r}(t) = \left\langle (350\cos\theta)t, -16t^2 + (350\sin\theta)t + 3 \right\rangle.$$

We need to find θ so that $\vec{r}(t) = \langle 25, 6 \rangle$ for some value of *t*. That is, we want to find θ and *t* such that

$$(350\cos\theta)t = 25 \quad \text{and} \quad -16t^2 + (350\sin\theta)t + 3 = 6.$$

This is not trivial (though not "hard"). We start by solving each equation for $\cos\theta$ and $\sin\theta$, respectively.

$$\cos\theta = \frac{25}{350t} \quad \text{and} \quad \sin\theta = \frac{3 + 16t^2}{350t}.$$

Using the Pythagorean Identity $\cos^2\theta + \sin^2\theta = 1$, we have

$$\left(\frac{25}{350t}\right)^2 + \left(\frac{3 + 16t^2}{350t}\right)^2 = 1$$

Multiply both sides by $(350t)^2$:

$$25^2 + (3 + 16t^2)^2 = 350^2 t^2$$
$$256t^4 - 122,404t^2 + 634 = 0.$$

This is a quadratic *in* t^2. That is, we can apply the quadratic formula to find t^2, then solve for *t* itself.

$$t^2 = \frac{122,404 \pm \sqrt{122,404^2 - 4(256)(634)}}{512}$$
$$t^2 = 0.0052, \ 478.135$$
$$t = \pm 0.072, \ \pm 21.866$$

Notes:

Clearly the negative t values do not fit our context, so we have $t = 0.072$ and $t = 21.866$. Using $\cos\theta = 25/(350t)$, we can solve for θ:

$$\theta = \cos^{-1}\left(\frac{25}{350 \cdot 0.072}\right) \quad \text{and} \quad \cos^{-1}\left(\frac{25}{350 \cdot 21.866}\right)$$

$$\theta = 7.03° \quad \text{and} \quad 89.8°.$$

Alex has two choices of angle. He can hold the rifle at an angle of about $7°$ with the horizontal and hit his target 0.07s after firing, or he can hold his rifle almost straight up, with an angle of $89.8°$, where he'll hit his target about 22s later. The first option is clearly the option he should choose.

Distance Traveled

Consider a driver who sets her cruise–control to 60mph, and travels at this speed for an hour. We can ask:

1. How far did the driver travel?

2. How far from her starting position is the driver?

The first is easy to answer: she traveled 60 miles. The second is impossible to answer with the given information. We do not know if she traveled in a straight line, on an oval racetrack, or along a slowly–winding highway.

This highlights an important fact: to compute distance traveled, we need only to know the speed, given by $\|\vec{v}(t)\|$.

> **Theorem 98   Distance Traveled**
> Let $\vec{v}(t)$ be a velocity function for a moving object. The distance traveled by the object on $[a, b]$ is:
>
> $$\text{distance traveled} = \int_a^b \|\vec{v}(t)\|\ dt.$$

Note that this is just a restatement of Theorem 97: arc length is the same as distance traveled, just viewed in a different context.

Example 7   Distance Traveled, Displacement, and Average Speed
A particle moves in space with position function $\vec{r}(t) = \langle t, t^2, \sin(\pi t)\rangle$ on $[-2, 2]$, where t is measured in seconds and distances are in meters. Find:

Notes:

1. The distance traveled by the particle on $[-2, 2]$.

2. The displacement of the particle on $[-2, 2]$.

3. The particle's average speed.

SOLUTION

1. We use Theorem 98 to establish the integral:

$$\text{distance traveled} = \int_{-2}^{2} \|\vec{v}(t)\| \ dt$$

$$= \int_{-2}^{2} \sqrt{1 + (2t)^2 + \pi^2 \cos^2(\pi t)} \ dt.$$

This cannot be solved in terms of elementary functions so we turn to numerical integration, finding the distance to be 12.88 m.

2. The displacement is the vector

$$\vec{r}(2) - \vec{r}(-2) = \langle 2, 4, 0 \rangle - \langle -2, 4, 0 \rangle = \langle 4, 0, 0 \rangle.$$

That is, the particle ends with an *x*-value increased by 4 and with *y*- and *z*-values the same (see Figure 12.18).

3. We found above that the particle traveled 12.88 m over 4 seconds. We can compute average speed by dividing: 12.88/4 = 3.22 m/s.

We should also consider Definition 25 of Section 5.4, which says that the average value of a function f on $[a, b]$ is $\frac{1}{b-a} \int_a^b f(x) \ dx$. In our context, the average value of the speed is

$$\text{average speed} = \frac{1}{2 - (-2)} \int_{-2}^{2} \|\vec{v}(t)\| \ dt \approx \frac{1}{4} 12.88 = 3.22 \text{m/s}.$$

Note how the physical context of a particle traveling gives meaning to a more abstract concept learned earlier.

Figure 12.18: The path of the particle in Example 7.

In Definition 25 of Chapter 5 we defined the average value of a function $f(x)$ on $[a, b]$ to be

$$\frac{1}{b - a} \int_a^b f(x) \ dx.$$

Note how in Example 7 we computed the average speed as

$$\frac{\text{distance traveled}}{\text{travel time}} = \frac{1}{2 - (-2)} \int_{-2}^{2} \|\vec{v}(t)\| \ dt;$$

Notes:

that is, we just found the average value of $\|\vec{v}(t)\|$ on $[-2, 2]$.

Likewise, given position function $\vec{r}(t)$, the average velocity on $[a, b]$ is

$$\frac{\text{displacement}}{\text{travel time}} = \frac{1}{b - a} \int_a^b \vec{r}\,'(t) \, dt = \frac{\vec{r}(b) - \vec{r}(a)}{b - a};$$

that is, it is the average value of $\vec{r}\,'(t)$, or $\vec{v}(t)$, on $[a, b]$.

Key Idea 58   Average Speed, Average Velocity

Let $\vec{r}(t)$ be a continuous position function on an open interval I containing $a < b$.

The **average speed** is:

$$\frac{\text{distance traveled}}{\text{travel time}} = \frac{\int_a^b \|\vec{v}(t)\| \; dt}{b - a} = \frac{1}{b - a} \int_a^b \|\vec{v}(t)\| \; dt.$$

The **average velocity** is:

$$\frac{\text{displacement}}{\text{travel time}} = \frac{\int_a^b \vec{r}\,'(t) \, dt}{b - a} = \frac{1}{b - a} \int_a^b \vec{r}\,'(t) \, dt.$$

The next two sections investigate more properties of the graphs of vector–valued functions and we'll apply these new ideas to what we just learned about motion.

Notes:

Exercises 12.3

Terms and Concepts

1. How is *velocity* different from *speed*?

2. What is the difference between *displacement* and *distance traveled*?

3. What is the difference between *average velocity* and *average speed*?

4. *Distance traveled* is the same as _____ _____, just viewed in a different context.

5. Describe a scenario where an object's average speed is a large number, but the magnitude of the average velocity is not a large number.

6. Explain why it is not possible to have an average velocity with a large magnitude but a small average speed.

Problems

In Exercises 7–10, a position function $\vec{r}(t)$ is given. Find $\vec{v}(t)$ and $\vec{a}(t)$.

7. $\vec{r}(t) = \langle 2t + 1, 5t - 2, 7 \rangle$

8. $\vec{r}(t) = \langle 3t^2 - 2t + 1, -t^2 + t + 14 \rangle$

9. $\vec{r}(t) = \langle \cos t, \sin t \rangle$

10. $\vec{r}(t) = \langle t/10, -\cos t, \sin t \rangle$

In Exercises 11–14, a position function $\vec{r}(t)$ is given. Sketch $\vec{r}(t)$ on the indicated interval. Find $\vec{v}(t)$ and $\vec{a}(t)$, then add $\vec{v}(t_0)$ and $\vec{a}(t_0)$ to your sketch, with their initial points at $\vec{r}(t_0)$, for the given value of t_0.

11. $\vec{r}(t) = \langle t, \sin t \rangle$ on $[0, \pi/2]$; $t_0 = \pi/4$

12. $\vec{r}(t) = \langle t^2, \sin t^2 \rangle$ on $[0, \pi/2]$; $t_0 = \sqrt{\pi/4}$

13. $\vec{r}(t) = \langle t^2 + t, -t^2 + 2t \rangle$ on $[-2, 2]$; $t_0 = 1$

14. $\vec{r}(t) = \left\langle \dfrac{2t + 3}{t^2 + 1}, t^2 \right\rangle$ on $[-1, 1]$; $t_0 = 0$

In Exercises 15–24, a position function $\vec{r}(t)$ of an object is given. Find the speed of the object in terms of t, and find where the speed is minimized/maximized on the indicated interval.

15. $\vec{r}(t) = \langle t^2, t \rangle$ on $[-1, 1]$

16. $\vec{r}(t) = \langle t^2, t^2 - t^3 \rangle$ on $[-1, 1]$

17. $\vec{r}(t) = \langle 5\cos t, 5\sin t \rangle$ on $[0, 2\pi]$

18. $\vec{r}(t) = \langle 2\cos t, 5\sin t \rangle$ on $[0, 2\pi]$

19. $\vec{r}(t) = \langle \sec t, \tan t \rangle$ on $[0, \pi/4]$

20. $\vec{r}(t) = \langle t + \cos t, 1 - \sin t \rangle$ on $[0, 2\pi]$

21. $\vec{r}(t) = \langle 12t, 5\cos t, 5\sin t \rangle$ on $[0, 4\pi]$

22. $\vec{r}(t) = \langle t^2 - t, t^2 + t, t \rangle$ on $[0, 1]$

23. $\vec{r}(t) = \left\langle t, t^2, \sqrt{1 - t^2} \right\rangle$ on $[-1, 1]$

24. **Projectile Motion:** $\vec{r}(t) =$
$$\left\langle (v_0 \cos\theta)t, -\frac{1}{2}gt^2 + (v_0 \sin\theta)t \right\rangle \text{ on } \left[0, \frac{2v_0 \sin\theta}{g}\right]$$

In Exercises 25–28, position functions $\vec{r}_1(t)$ and $\vec{r}_2(s)$ for two objects are given that follow the same path on the respective intervals.

(a) Show that the positions are the same at the indicated t_0 and s_0 values; i.e., show $\vec{r}_1(t_0) = \vec{r}_2(s_0)$.

(b) Find the velocity, speed and acceleration of the two objects at t_0 and s_0, respectively.

25. $\vec{r}_1(t) = \langle t, t^2 \rangle$ on $[0, 1]$; $t_0 = 1$
    $\vec{r}_2(s) = \langle s^2, s^4 \rangle$ on $[0, 1]$; $s_0 = 1$

26. $\vec{r}_1(t) = \langle 3\cos t, 3\sin t \rangle$ on $[0, 2\pi]$; $t_0 = \pi/2$
    $\vec{r}_2(s) = \langle 3\cos(4s), 3\sin(4s) \rangle$ on $[0, \pi/2]$; $s_0 = \pi/8$

27. $\vec{r}_1(t) = \langle 3t, 2t \rangle$ on $[0, 2]$; $t_0 = 2$
    $\vec{r}_2(s) = \langle 6s - 6, 4s - 4 \rangle$ on $[1, 2]$; $s_0 = 2$

28. $\vec{r}_1(t) = \langle t, \sqrt{t} \rangle$ on $[0, 1]$; $t_0 = 1$
    $\vec{r}_2(s) = \langle \sin t, \sqrt{\sin t} \rangle$ on $[0, \pi/2]$; $s_0 = \pi/2$

In Exercises 29–32, find the position function of an object given its acceleration and initial velocity and position.

29. $\vec{a}(t) = \langle 2, 3 \rangle$; $\vec{v}(0) = \langle 1, 2 \rangle$, $\vec{r}(0) = \langle 5, -2 \rangle$

30. $\vec{a}(t) = \langle 2, 3 \rangle$; $\vec{v}(1) = \langle 1, 2 \rangle$, $\vec{r}(1) = \langle 5, -2 \rangle$

31. $\vec{a}(t) = \langle \cos t, -\sin t \rangle$; $\vec{v}(0) = \langle 0, 1 \rangle$, $\vec{r}(0) = \langle 0, 0 \rangle$

32. $\vec{a}(t) = \langle 0, -32 \rangle$; $\vec{v}(0) = \langle 10, 50 \rangle$, $\vec{r}(0) = \langle 0, 0 \rangle$

In Exercises 33–36, find the displacement, distance traveled, average velocity and average speed of the described object on the given interval.

33. An object with position function $\vec{r}(t) = \langle 2\cos t, 2\sin t, 3t \rangle$, where distances are measured in feet and time is in seconds, on $[0, 2\pi]$.

34. An object with position function $\vec{r}(t) = \langle 5\cos t, -5\sin t \rangle$, where distances are measured in feet and time is in seconds, on $[0, \pi]$.

35. An object with velocity function $\vec{v}(t) = \langle \cos t, \sin t \rangle$, where distances are measured in feet and time is in seconds, on $[0, 2\pi]$.

36. An object with velocity function $\vec{v}(t) = \langle 1, 2, -1 \rangle$, where distances are measured in feet and time is in seconds, on $[0, 10]$.

Exercises 37–42 ask you to solve a variety of problems based on the principles of projectile motion.

37. A boy whirls a ball, attached to a 3ft string, above his head in a counter–clockwise circle. The ball makes 2 revolutions per second.
    At what *t*-values should the boy release the string so that the ball heads directly for a tree standing 10ft in front of him?

38. David faces Goliath with only a stone in a 3ft sling, which he whirls above his head at 4 revolutions per second. They stand 20ft apart.

    (a) At what t-values must David release the stone in his sling in order to hit Goliath?

    (b) What is the speed at which the stone is traveling when released?

    (c) Assume David releases the stone from a height of 6ft and Goliath's forehead is 9ft above the ground. What angle of elevation must David apply to the stone to hit Goliath's head?

39. A hunter aims at a deer which is 40 yards away. Her crossbow is at a height of 5ft, and she aims for a spot on the deer 4ft above the ground. The crossbow fires her arrows at 300ft/s.

    (a) At what angle of elevation should she hold the crossbow to hit her target?

    (b) If the deer is moving perpendicularly to her line of sight at a rate of 20mph, by approximately how much should she lead the deer in order to hit it in the desired location?

40. A baseball player hits a ball at 100mph, with an initial height of 3ft and an angle of elevation of $20°$, at Boston's Fenway Park. The ball flies towards the famed "Green Monster," a wall 37ft high located 310ft from home plate.

    (a) Show that as hit, the ball hits the wall.

    (b) Show that if the angle of elevation is $21°$, the ball clears the Green Monster.

41. A Cessna flies at 1000ft at 150mph and drops a box of supplies to the professor (and his wife) on an island. Ignoring wind resistance, how far horizontally will the supplies travel before they land?

42. A football quarterback throws a pass from a height of 6ft, intending to hit his receiver 20yds away at a height of 5ft.

    (a) If the ball is thrown at a rate of 50mph, what angle of elevation is needed to hit his intended target?

    (b) If the ball is thrown at with an angle of elevation of $8°$, what initial ball speed is needed to hit his target?

12.4   Unit Tangent and Normal Vectors

Unit Tangent Vector

Given a smooth vector–valued function $\vec{r}(t)$, we defined in Definition 74 that any vector parallel to $\vec{r}\,'(t_0)$ is *tangent* to the graph of $\vec{r}(t)$ at $t = t_0$. It is often useful to consider just the *direction* of $\vec{r}\,'(t)$ and not its magnitude. Therefore we are interested in the unit vector in the direction of $\vec{r}\,'(t)$. This leads to a definition.

Definition 77   Unit Tangent Vector

Let $\vec{r}(t)$ be a smooth function on an open interval I. The unit tangent vector $\vec{T}(t)$ is

$$\vec{T}(t) = \frac{1}{\|\vec{r}\,'(t)\|}\vec{r}\,'(t).$$

Watch the video:
Tangent Line to a Parametrized Curve at
`https://youtu.be/39LA5WyVgKY`

Example 1   Computing the unit tangent vector
Let $\vec{r}(t) = \langle 3\cos t, 3\sin t, 4t\rangle$. Find $\vec{T}(t)$ and compute $\vec{T}(0)$ and $\vec{T}(1)$.

SOLUTION   We apply Definition 77 to find $\vec{T}(t)$.

$$\vec{T}(t) = \frac{1}{\|\vec{r}\,'(t)\|}\vec{r}\,'(t)$$

$$= \frac{1}{\sqrt{\left(-3\sin t\right)^2 + \left(3\cos t\right)^2 + 4^2}}\,\langle -3\sin t, 3\cos t, 4\rangle$$

$$= \left\langle -\frac{3}{5}\sin t, \frac{3}{5}\cos t, \frac{4}{5}\right\rangle.$$

We can now easily compute $\vec{T}(0)$ and $\vec{T}(1)$:

$$\vec{T}(0) = \left\langle 0, \frac{3}{5}, \frac{4}{5}\right\rangle ; \quad \vec{T}(1) = \left\langle -\frac{3}{5}\sin 1, \frac{3}{5}\cos 1, \frac{4}{5}\right\rangle.$$

Notes:

These are plotted in Figure 12.19 with their initial points at $\vec{r}(0)$ and $\vec{r}(1)$, respectively. (They look rather "short" since they are only length 1.)

In many ways, the previous example was "too nice." It turned out that $\vec{r}\,'(t)$ was always of length 5. In the next example the length of $\vec{r}\,'(t)$ is variable, leaving us with a formula that is not as clean.

Example 2   **Computing the unit tangent vector**
Let $\vec{r}(t) = \left\langle t^2 - t, t^2 + t \right\rangle$. Find $\vec{T}(t)$ and compute $\vec{T}(0)$ and $\vec{T}(1)$.

SOLUTION   We find $\vec{r}\,'(t) = \langle 2t - 1, 2t + 1 \rangle$, and

$$\|\vec{r}\,'(t)\| = \sqrt{(2t-1)^2 + (2t+1)^2} = \sqrt{8t^2 + 2}.$$

Therefore

$$\vec{T}(t) = \frac{1}{\sqrt{8t^2 + 2}} \langle 2t - 1, 2t + 1 \rangle = \left\langle \frac{2t-1}{\sqrt{8t^2+2}}, \frac{2t+1}{\sqrt{8t^2+2}} \right\rangle.$$

When $t = 0$, we have $\vec{T}(0) = \left\langle -1/\sqrt{2}, 1/\sqrt{2} \right\rangle$; when $t = 1$, we have $\vec{T}(1) = \left\langle 1/\sqrt{10}, 3/\sqrt{10} \right\rangle$. We leave it to the reader to verify each of these is a unit vector. They are plotted in Figure 12.20

Unit Normal Vector

Just as knowing the direction tangent to a path is important, knowing a direction orthogonal to a path is important. When dealing with real-valued functions, we defined the normal line at a point to the be the line through the point that was perpendicular to the tangent line at that point. We can do a similar thing with vector–valued functions. Given $\vec{r}(t)$ in $\mathbb{R}^2$, we have 2 directions perpendicular to the tangent vector, as shown in Figure 12.21. It is good to wonder "Is one of these two directions preferable over the other?"

Given $\vec{r}(t)$ in $\mathbb{R}^3$, there are infinite vectors orthogonal to the tangent vector at a given point. Again, we might wonder "Is one of these infinite choices preferable over the others? Is one of these the 'right' choice?"

The answer in both $\mathbb{R}^2$ and $\mathbb{R}^3$ is "Yes, there is one vector that is preferable." Recall Theorem 95, which states that if $\vec{r}(t)$ has constant length, then $\vec{r}(t)$ is orthogonal to $\vec{r}\,'(t)$ for all t. We know $\vec{T}(t)$, the unit tangent vector, has constant length. Therefore $\vec{T}(t)$ is orthogonal to $\vec{T}\,'(t)$.

We'll see that $\vec{T}\,'(t)$ is more than just a convenient choice of vector that is orthogonal to $\vec{r}\,'(t)$; rather, it is the "right" choice. Since all we care about is the direction, we define this newly found vector to be a unit vector.
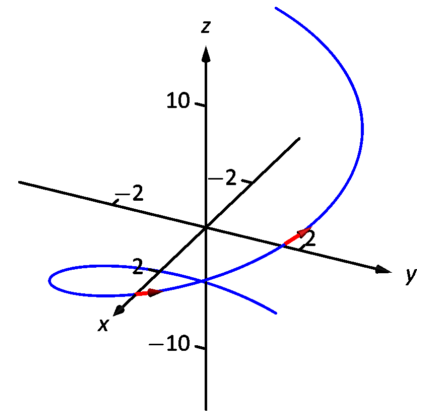
Notes:

Figure 12.19: Plotting unit tangent vectors in Example 1.
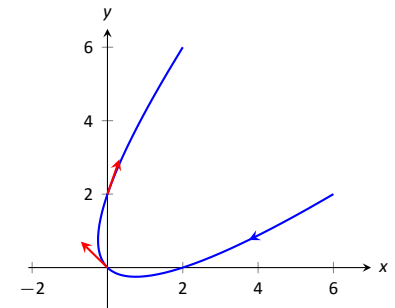

Figure 12.20: Plotting unit tangent vectors in Example 2.

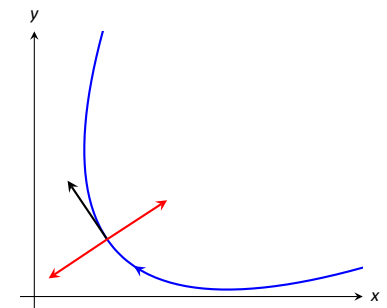Figure 12.21: Given a direction in the plane, there are always two directions orthogonal to it.

Note: $\vec{T}(t)$ is a unit vector, by definition. This *does not* imply that $\vec{T}\,'(t)$ is also a unit vector.

Definition 78   Unit Normal Vector

Let $\vec{r}(t)$ be a vector–valued function where the unit tangent vector, $\vec{T}(t)$, is smooth on an open interval I. The **unit normal vector** $\vec{N}(t)$ is

$$\vec{N}(t) = \frac{1}{\left\| \vec{T}'(t) \right\|} \vec{T}'(t).$$

Example 3   Computing the unit normal vector

Let $\vec{r}(t) = \langle 3\cos t, 3\sin t, 4t \rangle$ as in Example 1. Sketch both $\vec{T}(\pi/2)$ and $\vec{N}(\pi/2)$ with initial points at $\vec{r}(\pi/2)$.

Solution   In Example 1, we found $\vec{T}(t) = \left\langle (-3/5)\sin t, (3/5)\cos t, 4/5 \right\rangle$. Therefore

$$\vec{T}'(t) = \left\langle -\frac{3}{5}\cos t, -\frac{3}{5}\sin t, 0 \right\rangle \quad \text{and} \quad \left\| \vec{T}'(t) \right\| = \frac{3}{5}.$$

Thus

$$\vec{N}(t) = \frac{\vec{T}'(t)}{3/5} = \langle -\cos t, -\sin t, 0 \rangle.$$

We compute $\vec{T}(\pi/2) = \langle -3/5, 0, 4/5 \rangle$ and $\vec{N}(\pi/2) = \langle 0, -1, 0 \rangle$. These are sketched in Figure 12.22.

The previous example was once again "too nice." In general, the expression for $\vec{T}(t)$ contains fractions of square–roots, hence the expression of $\vec{T}'(t)$ is very messy. We demonstrate this in the next example.

Example 4   Computing the unit normal vector

Let $\vec{r}(t) = \langle t^2 - t, t^2 + t \rangle$ as in Example 2. Find $\vec{N}(t)$ and sketch $\vec{r}(t)$ with the unit tangent and normal vectors at $t = -1, 0$ and 1.

Solution   In Example 2, we found

$$\vec{T}(t) = \left\langle \frac{2t - 1}{\sqrt{8t^2 + 2}}, \frac{2t + 1}{\sqrt{8t^2 + 2}} \right\rangle.$$

Finding $\vec{T}'(t)$ requires two applications of the Quotient Rule:

Figure 12.22: Plotting unit tangent and normal vectors in Figure 12.22.

Notes:

$$\vec{T}'(t) = \left\langle \frac{\sqrt{8t^2+2}(2)-(2t-1)\left(\frac{1}{2}(8t^2+2)^{-1/2}(16t)\right)}{8t^2+2}, \right.$$
$$\left. \frac{\sqrt{8t^2+2}(2)-(2t+1)\left(\frac{1}{2}(8t^2+2)^{-1/2}(16t)\right)}{8t^2+2} \right\rangle$$

$$= \left\langle \frac{4(2t+1)}{(8t^2+2)^{3/2}}, \frac{4(1-2t)}{(8t^2+2)^{3/2}} \right\rangle$$

This is not a unit vector; to find $\vec{N}(t)$, we need to divide $\vec{T}'(t)$ by it's magnitude.

$$\|\vec{T}'(t)\| = \sqrt{\frac{16(2t+1)^2}{(8t^2+2)^3} + \frac{16(1-2t)^2}{(8t^2+2)^3}}$$

$$= \sqrt{\frac{16(8t^2+2)}{(8t^2+2)^3}}$$

$$= \frac{4}{8t^2+2}.$$

Finally,

$$\vec{N}(t) = \frac{1}{4/(8t^2+2)} \left\langle \frac{4(2t+1)}{(8t^2+2)^{3/2}}, \frac{4(1-2t)}{(8t^2+2)^{3/2}} \right\rangle$$

$$= \left\langle \frac{2t+1}{\sqrt{8t^2+2}}, -\frac{2t-1}{\sqrt{8t^2+2}} \right\rangle.$$

Because we are normalizing $\vec{T}'(t)$, it is usually easier to scale it first. We see that $\vec{T}'(t)$ is parallel to $\langle 2t+1, 1-2t \rangle$, which has length $\sqrt{(2t+1)^2 + (1-2t)^2} = \sqrt{8t^2+2}$, leading to the same $\vec{N}(t)$.

Using this formula for $\vec{N}(t)$, we compute the unit tangent and normal vectors for $t = -1, 0$ and 1 and sketch them in Figure 12.23.

The final result for $\vec{N}(t)$ in Example 4 is suspiciously similar to $\vec{T}(t)$. There is a clear reason for this. If $\vec{u} = \langle u_1, u_2 \rangle$ is a unit vector in $\mathbb{R}^2$, then the *only* unit vectors orthogonal to $\vec{u}$ are $\langle -u_2, u_1 \rangle$ and $\langle u_2, -u_1 \rangle$. Given $\vec{T}(t)$, we can quickly determine $\vec{N}(t)$ if we know which term to multiply by (-1).

Consider again Figure 12.23, where we have plotted some unit tangent and normal vectors. Note how $\vec{N}(t)$ always points "inside" the curve, or to the concave side of the curve. This is not a coincidence; this is true in general. Knowing the direction that $\vec{r}(t)$ "turns" allows us to quickly find $\vec{N}(t)$.
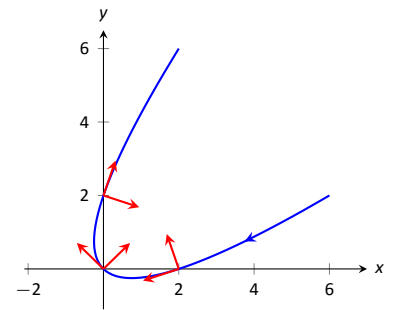

Figure 12.23: Plotting unit tangent and normal vectors in Example 4.

Notes:

> **Theorem 99   Unit Normal Vectors in $\mathbb{R}^2$**
>
> Let $\vec{r}(t)$ be a vector–valued function in $\mathbb{R}^2$ where $\vec{T}'(t)$ is smooth on an open interval I. Let t_0 be in I and $\vec{T}(t_0) = \langle t_1, t_2 \rangle$ Then $\vec{N}(t_0)$ is either
>
> $$\vec{N}(t_0) = \langle -t_2, t_1 \rangle \quad \text{or} \quad \vec{N}(t_0) = \langle t_2, -t_1 \rangle,$$
>
> whichever is the vector that points to the concave side of the graph of $\vec{r}$.

Application to Acceleration

Let $\vec{r}(t)$ be a position function. It is a fact (stated later in Theorem 100) that acceleration, $\vec{a}(t)$, lies in the plane defined by $\vec{T}$ and $\vec{N}$. That is, there are scalars a_T and a_N such that

$$\vec{a}(t) = a_T \vec{T}(t) + a_N \vec{N}(t).$$

The scalar a_T measures "how much" acceleration is in the direction of travel, that is, it measures the component of acceleration that affects the speed. The scalar a_N measures "how much" acceleration is perpendicular to the direction of travel, that is, it measures the component of acceleration that affects the direction of travel.

We can find a_T using the orthogonal projection of $\vec{a}(t)$ onto $\vec{T}(t)$ (review Definition 62 in Section 11.3 if needed). Recalling that since $\vec{T}(t)$ is a unit vector, $\vec{T}(t) \cdot \vec{T}(t) = 1$, so we have

$$\text{proj}_{\vec{T}(t)}\, \vec{a}(t) = \frac{\vec{a}(t) \cdot \vec{T}(t)}{\vec{T}(t) \cdot \vec{T}(t)} \vec{T}(t) = \underbrace{\left(\vec{a}(t) \cdot \vec{T}(t) \right)}_{a_T} \vec{T}(t).$$

Thus the amount of $\vec{a}(t)$ in the direction of $\vec{T}(t)$ is $a_T = \vec{a}(t) \cdot \vec{T}(t)$. The same logic gives $a_N = \vec{a}(t) \cdot \vec{N}(t)$.

While this is a fine way of computing a_T, there are simpler ways of finding a_N (as finding $\vec{N}$ itself can be complicated). The following theorem gives alternate formulas for a_T and a_N.

Note: Keep in mind that both a_T and a_N are functions of t; that is, the scalar changes depending on t. It is convention to drop the "(t)" notation from $a_T(t)$ and simply write a_T.

Notes:

> **Theorem 100**   **Acceleration in the Plane Defined by $\vec{T}$ and $\vec{N}$**
>
> Let $\vec{r}(t)$ be a position function with acceleration $\vec{a}(t)$ and unit tangent and normal vectors $\vec{T}(t)$ and $\vec{N}(t)$. Then $\vec{a}(t)$ lies in the plane defined by $\vec{T}(t)$ and $\vec{N}(t)$; that is, there exists scalars a_T and a_N such that
>
> $$\vec{a}(t) = a_T \vec{T}(t) + a_N \vec{N}(t).$$
>
> Moreover,
>
> $$a_T = \vec{a}(t) \cdot \vec{T}(t) = \frac{d}{dt} \left(\|\vec{v}(t)\| \right)$$
>
> $$a_N = \vec{a}(t) \cdot \vec{N}(t) = \sqrt{\|\vec{a}(t)\|^2 - a_T^2} = \frac{\|\vec{a}(t) \times \vec{v}(t)\|}{\|\vec{v}(t)\|} = \|\vec{v}(t)\| \, \|\vec{T}'(t)\|$$

Note the second formula for a_T: $\dfrac{d}{dt} \left(\|\vec{v}(t)\| \right)$. This measures the rate of change of speed, which again is the amount of acceleration in the direction of travel.

Proof
We see that

$$\vec{a}(t) = \frac{d}{dt} \vec{v}(t) = \frac{d}{dt} \left(\|\vec{v}(t)\| \, \vec{T}(t) \right) = \left(\frac{d}{dt} \|\vec{v}(t)\| \right) \vec{T}(t) + \|\vec{v}(t)\| \, \vec{T}'(t)$$

$$= \left(\frac{d}{dt} \|\vec{v}(t)\| \right) \vec{T}(t) + \|\vec{v}(t)\| \, \|\vec{T}'(t)\| \, \vec{N}(t).$$

Since $\vec{T}(t)$ and $\vec{N}(t)$ are not parallel, this decomposition is unique and the coefficients tell us a_T and a_N.

Because $\|\vec{T}\| = 1$, Theorem 95 tells us that $\vec{T}$ and $\vec{T}' = \|\vec{T}'\| \, \vec{N}$ are orthogonal. This means that

$$\|\vec{a}(t) \times \vec{v}(t)\| = \left\| a_N \vec{N}(t) \times \|\vec{v}(t)\| \, \vec{T}(t) \right\| = a_N \|\vec{v}(t)\|.$$

Also, the Pythagorean theorem tells us that

$$\|\vec{a}(t)\|^2 = \left\| a_T \vec{T}(t) \right\|^2 + \left\| a_N \vec{N}(t) \right\|^2 = a_T^2 + a_N^2. \qquad \square$$

Example 5   **Computing a_T and a_N**
Let $\vec{r}(t) = \langle 3\cos t, 3\sin t, 4t \rangle$ as in Examples 1 and 3. Find a_T and a_N.

Notes:

SOLUTION   The previous examples give $\vec{a}(t) = \langle -3\cos t, -3\sin t, 0\rangle$ and

$$\vec{T}(t) = \left\langle -\frac{3}{5}\sin t, \frac{3}{5}\cos t, \frac{4}{5}\right\rangle \quad \text{and} \quad \vec{N}(t) = \langle -\cos t, -\sin t, 0\rangle.$$

We can find a_T and a_N directly with dot products:

$$a_T = \vec{a}(t)\cdot\vec{T}(t) = \frac{9}{5}\cos t\sin t - \frac{9}{5}\cos t\sin t + 0 = 0.$$
$$a_N = \vec{a}(t)\cdot\vec{N}(t) = 3\cos^2 t + 3\sin^2 t + 0 = 3.$$

Thus $\vec{a}(t) = 0\vec{T}(t) + 3\vec{N}(t) = 3\vec{N}(t)$, which is clearly the case.

What is the practical interpretation of these numbers? $a_T = 0$ means the object is moving at a constant speed, and hence all acceleration comes in the form of direction change.

Example 6   **Computing a_T and a_N**
Let $\vec{r}(t) = \langle t^2 - t, t^2 + t\rangle$ as in Examples 2 and 4. Find a_T and a_N.

SOLUTION   The previous examples give $\vec{a}(t) = \langle 2, 2\rangle$ and

$$\vec{T}(t) = \left\langle \frac{2t-1}{\sqrt{8t^2+2}}, \frac{2t+1}{\sqrt{8t^2+2}}\right\rangle \quad \text{and} \quad \vec{N}(t) = \left\langle \frac{2t+1}{\sqrt{8t^2+2}}, -\frac{2t-1}{\sqrt{8t^2+2}}\right\rangle.$$

While we can compute a_N using $\vec{N}(t)$, we instead demonstrate using another formula from Theorem 100.

$$a_T = \vec{a}(t)\cdot\vec{T}(t) = \frac{4t-2}{\sqrt{8t^2+2}} + \frac{4t+2}{\sqrt{8t^2+2}} = \frac{8t}{\sqrt{8t^2+2}}.$$

$$a_N = \sqrt{\|\vec{a}(t)\|^2 - a_T^2} = \sqrt{8 - \left(\frac{8t}{\sqrt{8t^2+2}}\right)^2} = \frac{4}{\sqrt{8t^2+2}}.$$

When $t = 2$, $a_T = \frac{16}{\sqrt{34}}$ and $a_N = \frac{4}{\sqrt{34}}$. We interpret this to mean that at $t = 2$, the particle is acculturating mostly by increasing speed, not by changing direction. As the path near $t = 2$ is relatively straight, this should make intuitive sense. Figure 12.24 gives a graph of the path for reference.

Contrast this with $t = 0$, where $a_T = 0$ and $a_N = 4/\sqrt{2} = 2\sqrt{2}$. Here the particle's speed is not changing and all acceleration is in the form of direction change.
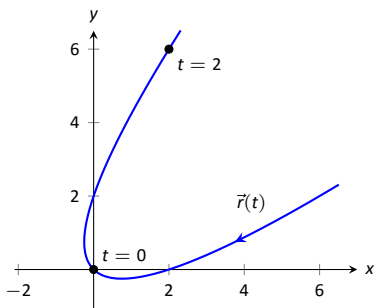

Figure 12.24: Graphing $\vec{r}(t)$ in Example 6.

Notes:

Example 7   **Analyzing projectile motion**

A ball is thrown from a height of 240ft with an initial speed of 64ft/s and an angle of elevation of 30°. Find the position function $\vec{r}(t)$ of the ball and analyze a_T and a_N.

SOLUTION   Using Equation (12.1) of Section 12.3 we form the position function of the ball:

$$\vec{r}(t) = \left\langle \left(64\cos 30°\right)t, -16t^2 + \left(64\sin 30°\right)t + 240 \right\rangle,$$

which we plot in Figure 12.25.

From this we find $\vec{v}(t) = \langle 64\cos 30°, -32t + 64\sin 30° \rangle$ and $\vec{a}(t) = \langle 0, -32 \rangle$. Computing $\vec{T}(t)$ is not difficult, and with some simplification we find

$$\vec{T}(t) = \left\langle \frac{\sqrt{3}}{\sqrt{t^2 - 2t + 4}}, \frac{1-t}{\sqrt{t^2 - 2t + 4}} \right\rangle.$$

With $\vec{a}(t)$ as simple as it is, finding a_T is also simple:

$$a_T = \vec{a}(t) \cdot \vec{T}(t) = \frac{32t - 32}{\sqrt{t^2 - 2t + 4}}.$$

We choose to not find $\vec{N}(t)$ and find a_N through the formula $a_N = \sqrt{\|\vec{a}(t)\|^2 - a_T^2}$:

$$a_N = \sqrt{32^2 - \left(\frac{32t - 32}{\sqrt{t^2 - 2t + 4}}\right)^2} = \frac{32\sqrt{3}}{\sqrt{t^2 - 2t + 4}}.$$

Figure 12.26 gives a table of values of a_T and a_N. When $t = 0$, we see the ball's speed is decreasing; when $t = 1$ the speed of the ball is unchanged. This corresponds to the fact that at $t = 1$ the ball reaches its highest point.

After $t = 1$ we see that a_N is decreasing in value. This is because as the ball falls, it's path becomes straighter and most of the acceleration is in the form of speeding up the ball, and not in changing its direction.

Our understanding of the unit tangent and normal vectors is aiding our understanding of motion. The work in Example 7 gave quantitative analysis of what we intuitively knew.

The next section provides two more important steps towards this analysis. We currently describe position only in terms of time. In everyday life, though, we often describe position in terms of distance ("The gas station is about 2 miles ahead, on the left."). The *arc length parameter* allows us to reference position in terms of distance traveled.

We also intuitively know that some paths are straighter than others – and some are curvier than others, but we lack a measurement of "curviness." The arc length parameter provides a way for us to compute *curvature*, a quantitative measurement of how curvy a curve is.
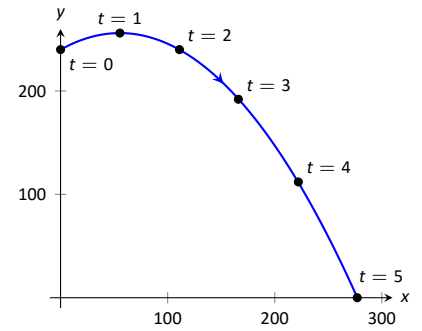
Figure 12.25: Plotting the position of a thrown ball, with 1s increments shown.

t	a_T	a_N
0	−16	27.7
1	0	32
2	16	27.7
3	24.2	20.9
4	27.7	16
5	29.4	12.7

Figure 12.26: A table of values of a_T and a_N in Example 7.

Notes:

Exercises 12.4

Terms and Concepts

1. If $\vec{T}(t)$ is a unit tangent vector, what is $\left\|\vec{T}(t)\right\|$?

2. If $\vec{N}(t)$ is a unit normal vector, what is $\vec{N}(t) \cdot \vec{r}\,'(t)$?

3. The acceleration vector $\vec{a}(t)$ lies in the plane defined by what two vectors?

4. a_T measures how much the acceleration is affecting the _____ of an object.

Problems

In Exercises 5–8, given $\vec{r}(t)$, find $\vec{T}(t)$ and evaluate it at the indicated value of t.

5. $\vec{r}(t) = \langle 2t^2, t^2 - t \rangle, \quad t = 1$

6. $\vec{r}(t) = \langle t, \cos t \rangle, \quad t = \pi/4$

7. $\vec{r}(t) = \langle \cos^3 t, \sin^3 t \rangle, \quad t = \pi/4$

8. $\vec{r}(t) = \langle \cos t, \sin t \rangle, \quad t = \pi$

In Exercises 9–12, find the equation of the line tangent to the curve at the indicated t-value using the unit tangent vector. Note: these are the same problems as in Exercises 5 – 8.

9. $\vec{r}(t) = \langle 2t^2, t^2 - t \rangle, \quad t = 1$

10. $\vec{r}(t) = \langle t, \cos t \rangle, \quad t = \pi/4$

11. $\vec{r}(t) = \langle \cos^3 t, \sin^3 t \rangle, \quad t = \pi/4$

12. $\vec{r}(t) = \langle \cos t, \sin t \rangle, \quad t = \pi$

In Exercises 13–16, find $\vec{N}(t)$ using Definition 78. Confirm the result using Theorem 99.

13. $\vec{r}(t) = \langle 3\cos t, 3\sin t \rangle$

14. $\vec{r}(t) = \langle t, t^2 \rangle$

15. $\vec{r}(t) = \langle \cos t, 2\sin t \rangle$

16. $\vec{r}(t) = \langle e^t, e^{-t} \rangle$

In Exercises 17–20, a position function $\vec{r}(t)$ is given along with its unit tangent vector $\vec{T}(t)$ evaluated at $t = a$, for some value of a.

(a) Confirm that $\vec{T}(a)$ is as stated.

(b) Using a graph of $\vec{r}(t)$ and Theorem 99, find $\vec{N}(a)$.

17. $\vec{r}(t) = \langle 3\cos t, 5\sin t \rangle; \quad \vec{T}(\pi/4) = \left\langle -\dfrac{3}{\sqrt{34}}, \dfrac{5}{\sqrt{34}} \right\rangle.$

18. $\vec{r}(t) = \left\langle t, \dfrac{1}{t^2+1} \right\rangle; \quad \vec{T}(1) = \left\langle \dfrac{2}{\sqrt{5}}, -\dfrac{1}{\sqrt{5}} \right\rangle.$

19. $\vec{r}(t) = (1 + 2\sin t)\langle \cos t, \sin t \rangle; \quad \vec{T}(0) = \left\langle \dfrac{2}{\sqrt{5}}, \dfrac{1}{\sqrt{5}} \right\rangle.$

20. $\vec{r}(t) = \langle \cos^3 t, \sin^3 t \rangle; \quad \vec{T}(\pi/4) = \left\langle -\dfrac{1}{\sqrt{2}}, \dfrac{1}{\sqrt{2}} \right\rangle.$

In Exercises 21–24, find $\vec{N}(t)$.

21. $\vec{r}(t) = \langle 4t, 2\sin t, 2\cos t \rangle$

22. $\vec{r}(t) = \langle 5\cos t, 3\sin t, 4\sin t \rangle$

23. $\vec{r}(t) = \langle a\cos t, a\sin t, bt \rangle; \quad a > 0$

24. $\vec{r}(t) = \langle \cos(at), \sin(at), t \rangle$

In Exercises 25–30, find a_T and a_N given $\vec{r}(t)$. Sketch $\vec{r}(t)$ on the indicated interval, and comment on the relative sizes of a_T and a_N at the indicated t values.

25. $\vec{r}(t) = \langle t, t^2 \rangle$ on $[-1, 1]$; consider $t = 0$ and $t = 1$.

26. $\vec{r}(t) = \langle t, 1/t \rangle$ on $(0, 4]$; consider $t = 1$ and $t = 2$.

27. $\vec{r}(t) = \langle 2\cos t, 2\sin t \rangle$ on $[0, 2\pi]$; consider $t = 0$ and $t = \pi/2$.

28. $\vec{r}(t) = \langle \cos(t^2), \sin(t^2) \rangle$ on $(0, 2\pi]$; consider $t = \sqrt{\pi/2}$ and $t = \sqrt{\pi}$.

29. $\vec{r}(t) = \langle a\cos t, a\sin t, bt \rangle$ on $[0, 2\pi]$, where $a, b > 0$; consider $t = 0$ and $t = \pi/2$.

30. $\vec{r}(t) = \langle 5\cos t, 4\sin t, 3\sin t \rangle$ on $[0, 2\pi]$; consider $t = 0$ and $t = \pi/2$.

12.5   The Arc Length Parameter and Curvature

In normal conversation we describe position in terms of both *time* and *distance*. For instance, imagine driving to visit a friend. If she calls and asks where you are, you might answer "I am 20 minutes from your house," or you might say "I am 10 miles from your house." Both answers provide your friend with a general idea of where you are.

Currently, our vector–valued functions have defined points with a parameter t, which we often take to represent time. Consider Figure 12.27(a), where $\vec{r}(t) = \langle t^2 - t, t^2 + t \rangle$ is graphed and the points corresponding to $t = 0, \ 1$ and 2 are shown. Note how the arc length between $t = 0$ and $t = 1$ is smaller than the arc length between $t = 1$ and $t = 2$; if the parameter t is time and $\vec{r}$ is position, we can say that the particle traveled faster on $[1, 2]$ than on $[0, 1]$.

Now consider Figure 12.27(b), where the same graph is parametrized by a different variable s. Points corresponding to $s = 0$ through $s = 6$ are plotted. The arc length of the graph between each adjacent pair of points is 1. We can view this parameter s as distance; that is, the arc length of the graph from $s = 0$ to $s = 3$ is 3, the arc length from $s = 2$ to $s = 6$ is 4, etc. If one wants to find the point 2.5 units from an initial location (i.e., $s = 0$), one would compute $\vec{r}(2.5)$. This parameter s is very useful, and is called the **arc length parameter**.

How do we find the arc length parameter?

Start with any parametrization of $\vec{r}$. We can compute the arc length of the graph of $\vec{r}$ on the interval $[0, t]$ with

$$\text{arc length } = \int_0^t \|\vec{r}'(u)\| \ du.$$

We can turn this into a function: as t varies, we find the arc length s from 0 to t. This function is

$$s(t) = \int_0^t \|\vec{r}'(u)\| \ du. \tag{12.2}$$

This establishes a relationship between s and t. Knowing this relationship explicitly, we can rewrite $\vec{r}(t)$ as a function of s: $\vec{r}(s)$. We demonstrate this in an example.

Watch the video:
Parametrize a Curve with Respect to Arc Length at
`https://youtu.be/GOR-qialKlE`
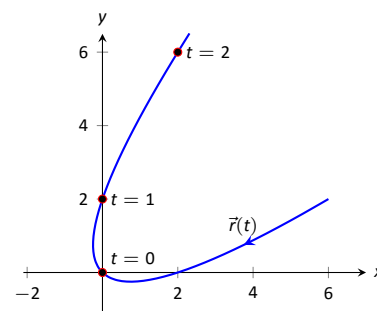
(a)
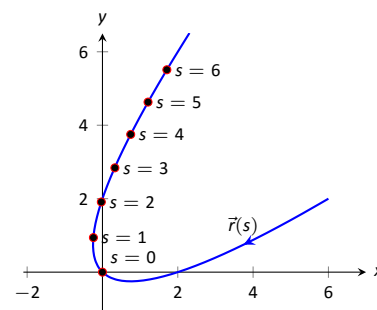
(b)

Figure 12.27: Introducing the arc length parameter.

Notes:

Example 1 **Finding the arc length parameter**

Let $\vec{r}(t) = \langle 3t - 1, 4t + 2 \rangle$. Parametrize $\vec{r}$ with the arc length parameter s.

SOLUTION Using Equation (12.2), we write

$$s(t) = \int_0^t \|\vec{r}\,'(u)\| \ du.$$

We can integrate this, explicitly finding a relationship between s and t:

$$s(t) = \int_0^t \|\vec{r}\,'(u)\| \ du$$
$$= \int_0^t \sqrt{3^2 + 4^2} \ du$$
$$= \int_0^t 5 \ du$$
$$= 5t.$$

Since $s = 5t$, we can write $t = s/5$ and replace t in $\vec{r}(t)$ with $s/5$:

$$\vec{r}(s) = \langle 3(s/5) - 1, 4(s/5) + 2 \rangle = \left\langle \frac{3}{5}s - 1, \frac{4}{5}s + 2 \right\rangle.$$

Clearly, as shown in Figure 12.28, the graph of $\vec{r}$ is a line, where $t = 0$ corresponds to the point $(-1, 2)$. What point on the line is 2 units away from this initial point? We find it with $s(2) = \langle 1/5, 18/5 \rangle$.

Is the point $(1/5, 18/5)$ really 2 units away from $(-1, 2)$? We use the Distance Formula to check:

$$d = \sqrt{\left(\frac{1}{5} - (-1)\right)^2 + \left(\frac{18}{5} - 2\right)^2} = \sqrt{\frac{36}{25} + \frac{64}{25}} = \sqrt{4} = 2.$$

Yes, $s(2)$ is indeed 2 units away, in the direction of travel, from the initial point.

Things worked out very nicely in Example 1; we were able to establish directly that $s = 5t$. Usually, the arc length parameter is much more difficult to describe in terms of t, a result of integrating a square–root. There are a number of things that we can learn about the arc length parameter from Equation (12.2), though, that are incredibly useful.

First, take the derivative of s with respect to t. The Fundamental Theorem of Calculus (see Theorem 38) states that

$$\frac{ds}{dt} = s\,'(t) = \|\vec{r}\,'(t)\|. \tag{12.3}$$

Figure 12.28: Graphing $\vec{r}$ in Example 1 with parameters t and s.

Notes:

Letting t represent time and $\vec{r}(t)$ represent position, we see that the rate of change of s with respect to t is speed; that is, the rate of change of "distance traveled" is speed, which should match our intuition.

The Chain Rule states that

$$\frac{d\vec{r}}{dt} = \frac{d\vec{r}}{ds} \cdot \frac{ds}{dt}$$
$$\vec{r}\,'(t) = \vec{r}\,'(s) \cdot \|\vec{r}\,'(t)\|.$$

Solving for $\vec{r}\,'(s)$, we have

$$\vec{r}\,'(s) = \frac{\vec{r}\,'(t)}{\|\vec{r}\,'(t)\|} = \vec{T}(t), \tag{12.4}$$

where $\vec{T}(t)$ is the unit tangent vector. Equation (12.4) is often misinterpreted, as one is tempted to think it states $\vec{r}\,'(t) = \vec{T}(t)$, but there is a big difference between $\vec{r}\,'(s)$ and $\vec{r}\,'(t)$. The key to take from it is that $\vec{r}\,'(s)$ is a unit vector. In fact, the following theorem states that this characterizes the arc length parameter.

Theorem 101   Arc Length Parameter

Let $\vec{r}(s)$ be a vector–valued function. The parameter s is the arc length parameter if, and only if, $\|\vec{r}\,'(s)\| = 1$.

Curvature

Consider points A and B on the curve graphed in Figure 12.29(a). One can readily argue that the curve curves more sharply at A than at B. It is useful to use a number to describe how sharply the curve bends; that number is the **curvature** of the curve.

We derive this number in the following way. Consider Figure 12.29(b), where unit tangent vectors are graphed around points A and B. Notice how the direction of the unit tangent vector changes quite a bit near A, whereas it does not change as much around B. This leads to an important concept: measuring the rate of change of the unit tangent vector with respect to arc length gives us a measurement of curvature.
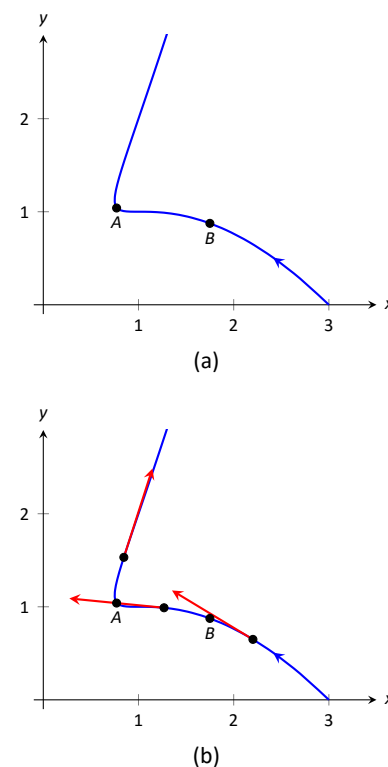

(a)

(b)

Figure 12.29: Establishing the concept of curvature.

Notes:

Definition 79   Curvature

Let $\vec{r}(s)$ be a vector–valued function where s is the arc length parameter. The curvature κ of the graph of $\vec{r}(s)$ is

$$\kappa = \left\|\frac{d\vec{T}}{ds}\right\| = \left\|\vec{T}'(s)\right\|.$$

If $\vec{r}(s)$ is parametrized by the arc length parameter, then

$$\vec{T}(s) = \frac{\vec{r}'(s)}{\|\vec{r}'(s)\|} \quad\text{and}\quad \vec{N}(s) = \frac{\vec{T}'(s)}{\|\vec{T}'(s)\|}.$$

Having defined $\|\vec{T}'(s)\| = \kappa$, we can rewrite the second equation as

$$\vec{T}'(s) = \kappa\vec{N}(s). \tag{12.5}$$

We already knew that $\vec{T}'(s)$ is in the same direction as $\vec{N}(s)$; that is, we can think of $\vec{T}(s)$ as being "pulled" in the direction of $\vec{N}(s)$. How "hard" is it being pulled? By a factor of κ. When the curvature is large, $\vec{T}(s)$ is being "pulled hard" and the direction of $\vec{T}(s)$ changes rapidly. When κ is small, $T(s)$ is not being pulled hard and hence its direction is not changing rapidly.

We use Definition 79 to find the curvature of the line in Example 1.

Example 2   Finding the curvature of a line

Use Definition 79 to find the curvature of $\vec{r}(t) = \langle 3t - 1, 4t + 2\rangle$.

SOLUTION   In Example 1, we found that the arc length parameter was defined by $s = 5t$, so $\vec{r}(s) = \langle 3s/5 - 1, 4s/5 + 2\rangle$ parametrized $\vec{r}$ with the arc length parameter. To find κ, we need to find $\vec{T}'(s)$.

$$\vec{T}(s) = \vec{r}'(s) \quad\text{(recall this is a unit vector)}$$
$$= \langle 3/5, 4/5\rangle.$$

Therefore

$$\vec{T}'(s) = \langle 0, 0\rangle$$

and

$$\kappa = \left\|\vec{T}'(s)\right\| = 0.$$

It probably comes as no surprise that the curvature of a line is 0. (How "curvy" is a line? It is not curvy at all.)

Notes:

While the definition of curvature is a useful mathematical concept, it is nearly impossible to use most of the time; writing $\vec{r}$ in terms of the arc length parameter is generally very hard. Fortunately, there are other methods of calculating this value that are much easier. There is a trade-off: the definition is "easy" to understand though hard to compute, whereas these other formulas are easy to compute though it may be hard to understand why they work.

Theorem 102   Formulas for Curvature

Let C be a smooth curve on an open interval I in the plane or in space.

1. If C is defined by $y = f(x)$, then

$$\kappa = \frac{|f''(x)|}{\left(1 + \left(f'(x)\right)^2\right)^{3/2}}.$$

2. If C is defined as a vector–valued function in the plane, $\vec{r}(t) = \langle x(t), y(t) \rangle$, then

$$\kappa = \frac{|x'y'' - x''y'|}{\left((x')^2 + (y')^2\right)^{3/2}}.$$

3. If C is defined in space by a vector–valued function $\vec{r}(t)$, then

$$\kappa = \frac{\left\|\vec{T}'(t)\right\|}{\left\|\vec{r}'(t)\right\|} = \frac{\left\|\vec{r}'(t) \times \vec{r}''(t)\right\|}{\left\|\vec{r}'(t)\right\|^3} = \frac{\vec{a}(t) \cdot \vec{N}(t)}{\left\|\vec{v}(t)\right\|^2}.$$

Proof

We'll prove statement 3; statements 1 and 2 are applications that we leave to the exercises. By the chain rule and then Equation (12.3),

$$\frac{d\vec{T}(t)}{dt} = \frac{d\vec{T}(s)}{ds}\frac{ds}{dt} = \frac{d\vec{T}(s)}{ds}\left\|\vec{r}'(t)\right\|,$$

so that

$$\kappa = \left\|\frac{d\vec{T}(s)}{ds}\right\| = \left\|\frac{\frac{d\vec{T}(t)}{dt}}{\left\|\vec{r}'(t)\right\|}\right\| = \frac{\left\|\frac{d\vec{T}(t)}{dt}\right\|}{\left\|\vec{r}'(t)\right\|} = \frac{\left\|\vec{T}'(t)\right\|}{\left\|\vec{r}'(t)\right\|}.$$

Notes:

Now, $\vec{r}\,'(t) = \|\vec{r}\,'(t)\| \, \vec{T}(t)$ so that

$$\vec{r}\,''(t) = (\|\vec{r}\,'(t)\|)' \, \vec{T}(t) + \|\vec{r}\,'(t)\| \, \vec{T}\,'(t) \qquad \text{and}$$

$$\vec{r}\,'(t) \times \vec{r}\,''(t) = \|\vec{r}\,'(t)\| \, \vec{T}(t) \times \|\vec{r}\,'(t)\| \, \vec{T}\,'(t) = \|\vec{r}\,'(t)\|^2 \, \vec{T}(t) \times \vec{T}\,'(t)$$

Because $\|\vec{T}\| = 1$, Theorem 95 tells us that $\vec{T}$ and $\vec{T}\,'$ are orthogonal. This means that

$$\|\vec{r}\,'(t) \times \vec{r}\,''(t)\| = \|\vec{r}\,'(t)\|^2 \, \|\vec{T}(t)\| \, \|\vec{T}\,'(t)\| = \|\vec{r}\,'(t)\|^2 \, \|\vec{T}\,'(t)\| \,, \qquad \text{and}$$

$$\kappa = \frac{\|\vec{T}\,'(t)\|}{\|\vec{r}\,'(t)\|} = \frac{\|\vec{r}\,'(t) \times \vec{r}\,''(t)\|}{\|\vec{r}\,'(t)\|^3}.$$

Theorem 100 tells us that $\vec{r}\,''(t) = a_T \vec{T}(t) + (\vec{r}\,''(t) \cdot \vec{N}(t))\vec{N}(t)$. Since $\vec{r}\,'(t) = \vec{v}(t)$ and $T(t)$ are parallel, their cross product is zero and

$$\kappa = \frac{\left\| \vec{v}(t) \times [a_T \vec{T}(t) + (\vec{r}\,''(t) \cdot \vec{N}(t))\vec{N}(t)] \right\|}{\|\vec{v}(t)\|^3} = \frac{\left\| \vec{v}(t) \times (\vec{a}(t) \cdot \vec{N}(t))\vec{N}(t) \right\|}{\|\vec{v}(t)\|^3}.$$

Since $\vec{v}(t)$ and $\vec{N}(t)$ are orthogonal, the norm of their cross product is the product of their norms, and

$$\kappa = \frac{\|\vec{v}(t)\| \, (\vec{a}(t) \cdot \vec{N}(t)) \, \left\| \vec{N}(t) \right\|}{\|\vec{v}(t)\|^3} = \frac{\vec{a}(t) \cdot \vec{N}(t)}{\|\vec{v}(t)\|^2}. \qquad \square$$

We practice using these formulas.

Example 3　　　Finding the curvature of a circle

Find the curvature of a circle with radius r, defined by $\vec{c}(t) = \langle r\cos t, r\sin t \rangle$.

SOLUTION　　　Before we start, we should expect the curvature of a circle to be constant, and not dependent on t. (Why?)

We compute κ using the second part of Theorem 102.

$$\kappa = \frac{|(-r\sin t)(-r\sin t) - (-r\cos t)(r\cos t)|}{\left((-r\sin t)^2 + (r\cos t)^2\right)^{3/2}}$$

$$= \frac{r^2(\sin^2 t + \cos^2 t)}{\left(r^2(\sin^2 t + \cos^2 t)\right)^{3/2}}$$

$$= \frac{r^2}{r^3} = \frac{1}{r}.$$

We have found that a circle with radius r has curvature $\kappa = 1/r$.

Notes:

Example 3 gives a great result. Before this example, if we were told "The curve has a curvature of 5 at point *A*," we would have no idea what this really meant. Is 5 "big" – does is correspond to a really sharp turn, or a not-so-sharp turn? Now we can think of 5 in terms of a circle with radius 1/5. Knowing the units (inches vs. miles, for instance) allows us to determine how sharply the curve is curving.

Let a point *P* on a smooth curve *C* be given, and let κ be the curvature of the curve at *P*. A circle that:

- passes through *P*,

- lies on the concave side of *C*,

- has a common tangent line as *C* at *P* and

- has radius $r = 1/\kappa$ (hence has curvature κ)

is the **osculating circle**, or **circle of curvature**, to *C* at *P*, and *r* is the **radius of curvature**. Figure 12.30 shows the graph of the curve seen earlier in Figure 12.29 and its osculating circles at *A* and *B*. A sharp turn corresponds to a circle with a small radius; a gradual turn corresponds to a circle with a large radius. Being able to think of curvature in terms of the radius of a circle is very useful. (The word "osculating" comes from a Latin word related to kissing; an osculating circle "kisses" the graph at a particular point. Many useful ideas in mathematics have come from studying the osculating circles to a curve.)

Figure 12.30: Illustrating the osculating circles for the curve seen in Figure 12.29.

Example 4   Finding curvature
Find the curvature of the parabola defined by $y = x^2$ at the vertex and at $x = 1$.

SOLUTION   We use the first formula found in Theorem 102.

$$\kappa(x) = \frac{|2|}{\left(1 + (2x)^2\right)^{3/2}}$$
$$= \frac{2}{\left(1 + 4x^2\right)^{3/2}}.$$

At the vertex ($x = 0$), the curvature is $\kappa = 2$. At $x = 1$, the curvature is $\kappa = 2/(5)^{3/2}$. So at $x = 0$, the curvature of $y = x^2$ is that of a circle of radius 1/2; at $x = 1$, the curvature is that of a circle with radius $(5)^{3/2}/2 \approx 5.59$. This is illustrated in Figure 12.31. At $x = 3$, the curvature is 0.009; the graph is nearly straight as the curvature is very close to 0.
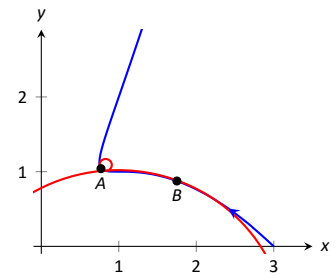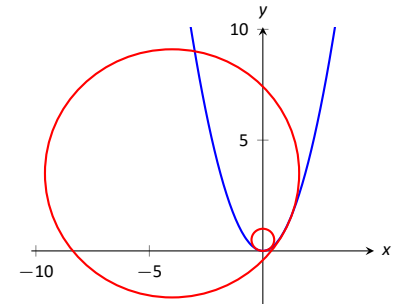

Figure 12.31: Examining the curvature of $y = x^2$.

Notes:

Example 5 **Finding curvature**

Find where the curvature of $\vec{r}(t) = \langle t, t^2, 2t^3 \rangle$ is maximized.

SOLUTION We use the third formula in Theorem 102 as $\vec{r}(t)$ is defined in space. We leave it to the reader to verify that

$$\vec{r}\,'(t) = \langle 1, 2t, 6t^2 \rangle, \quad \vec{r}\,''(t) = \langle 0, 2, 12t \rangle, \quad \text{and} \quad \vec{r}\,'(t) \times \vec{r}\,''(t) = \langle 12t^2, -12t, 2 \rangle.$$

Thus

$$\kappa(t) = \frac{\|\vec{r}\,'(t) \times \vec{r}\,''(t)\|}{\|\vec{r}\,'(t)\|^3}$$

$$= \frac{\|\langle 12t^2, -12t, 2 \rangle\|}{\|\langle 1, 2t, 6t^2 \rangle\|^3}$$

$$= \frac{\sqrt{144t^4 + 144t^2 + 4}}{\left(\sqrt{1 + 4t^2 + 36t^4}\right)^3}$$

While this is not a particularly "nice" formula, it does explicitly tell us what the curvature is at a given t value. To maximize $\kappa(t)$, we should solve $\kappa'(t) = 0$ for t. This is doable, but *very* time consuming. Instead, consider the graph of $\kappa(t)$ as given in Figure 12.32(a). We see that κ is maximized at two t values; using a numerical solver, we find these values are $t \approx \pm 0.189$. In part (b) of the figure we graph $\vec{r}(t)$ and indicate the points where curvature is maximized.

Curvature and Motion

Let $\vec{r}(t)$ be a position function of an object, with velocity $\vec{v}(t) = \vec{r}\,'(t)$ and acceleration $\vec{a}(t) = \vec{r}\,''(t)$. In Section 12.4 we established that acceleration is in the plane formed by $\vec{T}(t)$ and $\vec{N}(t)$, and that we can find scalars a_T and a_N such that

$$\vec{a}(t) = a_\mathrm{T}\vec{T}(t) + a_\mathrm{N}\vec{N}(t).$$

Theorem 100 gives formulas for a_T and a_N:

$$a_\mathrm{T} = \frac{d}{dt}\left(\|\vec{v}(t)\|\right) \quad \text{and} \quad a_\mathrm{N} = \frac{\|\vec{v}(t) \times \vec{a}(t)\|}{\|\vec{v}(t)\|}.$$

We understood that the amount of acceleration in the direction of $\vec{T}$ relates only to how the speed of the object is changing, and that the amount of acceleration in the direction of $\vec{N}$ relates to how the direction of travel of the object is changing. (That is, if the object travels at constant speed, $a_\mathrm{T} = 0$; if the object travels in a constant direction, $a_\mathrm{N} = 0$.)
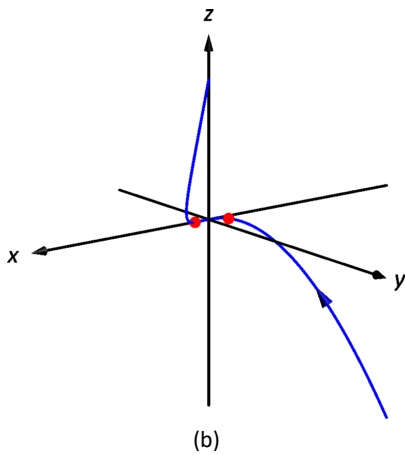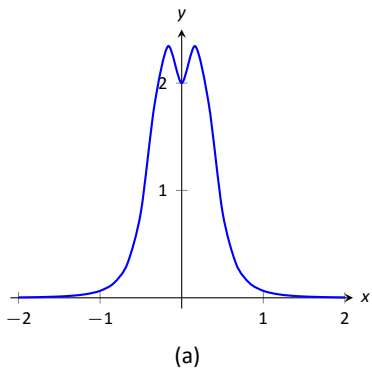

(a)

(b)

Figure 12.32: Understanding the curvature of a curve in space. The top is κ as a function of t; the bottom is $\vec{r}(t)$ as a function of t.

Notes:

In Equation (12.3) at the beginning of this section, we found $s'(t) = \|\vec{v}(t)\|$. We can combine this fact with the above formula for a_T to write

$$a_\mathrm{T} = \frac{d}{dt}\Big(\|\vec{v}(t)\| \Big) = \frac{d}{dt}\big(s'(t)\big) = s''(t).$$

Since $s'(t)$ is speed, $s''(t)$ is the rate at which speed is changing with respect to time. We see once more that the component of acceleration in the direction of travel relates only to speed, not to a change in direction.

Now compare the formula for a_N above to the formula for curvature in Theorem 102:

$$a_\mathrm{N} = \frac{\|\vec{v}(t) \times \vec{a}(t)\|}{\|\vec{v}(t)\|} \quad \text{and} \quad \kappa = \frac{\|\vec{r}'(t) \times \vec{r}''(t)\|}{\|\vec{r}'(t)\|^3} = \frac{\|\vec{v}(t) \times \vec{a}(t)\|}{\|\vec{v}(t)\|^3}.$$

Thus

$$a_\mathrm{N} = \kappa \|\vec{v}(t)\|^2 = \kappa \big(s'(t)\big)^2 \tag{12.6}$$

This last equation shows that the component of acceleration that changes the object's direction is dependent on two things: the curvature of the path and the speed of the object.

Imagine driving a car in a clockwise circle. You will naturally feel a force pushing you towards the door (more accurately, the door is pushing you as the car is turning and you want to travel in a straight line). If you keep the radius of the circle constant but speed up (i.e., increasing $s'(t)$), the door pushes harder against you (a_N has increased). If you keep your speed constant but tighten the turn (i.e., increase κ), once again the door will push harder against you.

Putting our new formulas for a_T and a_N together, we have

$$\vec{a}(t) = s''(t)\vec{T}(t) + \kappa \|\vec{v}(t)\|^2 \, \vec{N}(t).$$

This is not a particularly practical way of finding a_T and a_N, but it reveals some great concepts about how acceleration interacts with speed and the shape of a curve.

Example 6   Curvature and road design
The minimum radius of the curve in a highway cloverleaf is determined by the operating speed, as given in the table in Figure 12.33. For each curve and speed, compute a_N.

Solution   Using Equation (12.6), we can compute the acceleration normal to the curve in each case. We start by converting each speed from "miles per hour" to "feet per second" by multiplying by $5280/3600$.

Operating Speed (mph)	Minimum Radius (ft)
35	310
40	430
45	540

Figure 12.33: Operating speed and minimum radius in highway cloverleaf design.

Notes:

$$35\text{mph}, 310\text{ft} \;\Rightarrow\; 51.33\text{ft/s}, \quad \kappa = 1/310$$

$$a_N = \kappa \, \|\vec{v}(t)\|^2$$

$$= \frac{1}{310}\left(51.33\right)^2$$

$$= 8.50\text{ft/s}^2.$$

$$40\text{mph}, 430\text{ft} \;\Rightarrow\; 58.67\text{ft/s}, \quad \kappa = 1/430$$

$$a_N = \frac{1}{430}\left(58.67\right)^2$$

$$= 8.00\text{ft/s}^2.$$

$$45\text{mph}, 540\text{ft} \;\Rightarrow\; 66\text{ft/s}, \quad \kappa = 1/540$$

$$a_N = \frac{1}{540}\left(66\right)^2$$

$$= 8.07\text{ft/s}^2.$$

Note that each acceleration is similar; this is by design. Considering the classic "Force = mass × acceleration" formula, this acceleration must be kept small in order for the tires of a vehicle to keep a "grip" on the road. If one travels on a turn of radius 310ft at a rate of 50mph, the acceleration is double, at 17.35ft/s^2. If the acceleration is too high, the frictional force created by the tires may not be enough to keep the car from sliding. Civil engineers routinely compute a "safe" design speed, then subtract 5-10mph to create the posted speed limit for additional safety.

We end this chapter with a reflection on what we've covered. We started with vector–valued functions, which may have seemed at the time to be just another way of writing parametric equations. However, we have seen that the vector perspective has given us great insight into the behavior of functions and the study of motion. Vector–valued position functions convey displacement, distance traveled, speed, velocity, acceleration and curvature information, each of which has great importance in science and engineering.

Notes:

Exercises 12.5

Terms and Concepts

1. It is common to describe position in terms of both _____ and/or _____.

2. A measure of the "curviness" of a curve is _____.

3. Give two shapes with constant curvature.

4. Describe in your own words what an "osculating circle" is.

5. Complete the identity: $\vec{T}'(s) = $ _____ $\vec{N}(s)$.

6. Given a position function $\vec{r}(t)$, how are a_T and a_N affected by the curvature?

Problems

In Exercises 7–10, a position function $\vec{r}(t)$ is given, where $t = 0$ corresponds to the initial position. Find the arc length parameter s, and rewrite $\vec{r}(t)$ in terms of s; that is, find $\vec{r}(s)$.

7. $\vec{r}(t) = \langle 2t, t, -2t \rangle$

8. $\vec{r}(t) = \langle 7 \cos t, 7 \sin t \rangle$

9. $\vec{r}(t) = \langle 3 \cos t, 3 \sin t, 2t \rangle$

10. $\vec{r}(t) = \langle 5 \cos t, 13 \sin t, 12 \cos t \rangle$

In Exercises 11–22, a curve C is described along with 2 points on C.

  (a) Using a sketch, determine at which of these points the curvature is greater.

  (b) Find the curvature κ of C, and evaluate κ at each of the 2 given points.

11. C is defined by $y = x^3 - x$; points given at $x = 0$ and $x = 1/2$.

12. C is defined by $y = \dfrac{1}{x^2 + 1}$; points given at $x = 0$ and $x = 2$.

13. C is defined by $y = \cos x$; points given at $x = 0$ and $x = \pi/2$.

14. C is defined by $y = \sqrt{1 - x^2}$ on $(-1, 1)$; points given at $x = 0$ and $x = 1/2$.

15. C is defined by $\vec{r}(t) = \langle \cos t, \sin(2t) \rangle$; points given at $t = 0$ and $t = \pi/4$.

16. C is defined by $\vec{r}(t) = \langle \cos^2 t, \sin t \cos t \rangle$; points given at $t = 0$ and $t = \pi/3$.

17. C is defined by $\vec{r}(t) = \langle t^2 - 1, t^3 - t \rangle$; points given at $t = 0$ and $t = 5$.

18. C is defined by $\vec{r}(t) = \langle \tan t, \sec t \rangle$; points given at $t = 0$ and $t = \pi/6$.

19. C is defined by $\vec{r}(t) = \langle 4t + 2, 3t - 1, 2t + 5 \rangle$; points given at $t = 0$ and $t = 1$.

20. C is defined by $\vec{r}(t) = \langle t^3 - t, t^3 - 4, t^2 - 1 \rangle$; points given at $t = 0$ and $t = 1$.

21. C is defined by $\vec{r}(t) = \langle 3 \cos t, 3 \sin t, 2t \rangle$; points given at $t = 0$ and $t = \pi/2$.

22. C is defined by $\vec{r}(t) = \langle 5 \cos t, 13 \sin t, 12 \cos t \rangle$; points given at $t = 0$ and $t = \pi/2$.

In Exercises 23–26, find the value of x or t where curvature is maximized.

23. $y = \dfrac{1}{6} x^3$

24. $y = \sin x$

25. $\vec{r}(t) = \langle t^2 + 2t, 3t - t^2 \rangle$

26. $\vec{r}(t) = \langle t, 4/t, 3/t \rangle$

In Exercises 27–30, find the radius of curvature at the indicated value.

27. $y = \tan x$, at $x = \pi/4$

28. $y = x^2 + x - 3$, at $x = \pi/4$

29. $\vec{r}(t) = \langle \cos t, \sin(3t) \rangle$, at $t = 0$

30. $\vec{r}(t) = \langle 5 \cos(3t), t \rangle$, at $t = 0$

In Exercises 31–34, find the equation of the osculating circle to the curve at the indicated t-value.

31. $\vec{r}(t) = \langle t, t^2 \rangle$, at $t = 0$

32. $\vec{r}(t) = \langle 3 \cos t, \sin t \rangle$, at $t = 0$

33. $\vec{r}(t) = \langle 3 \cos t, \sin t \rangle$, at $t = \pi/2$

34. $\vec{r}(t) = \langle t^2 - t, t^2 + t \rangle$, at $t = 0$

35. For Theorem 102, use part 3 to prove part 2: If $\vec{r}(t) = \langle x(t), y(t) \rangle$ is a vector–valued function in the plane, then

$$\kappa = \frac{|x'y'' - x''y'|}{\left((x')^2 + (y')^2\right)^{3/2}}.$$

36. For Theorem 102, use part 2 or 3 to prove part 1: If $y = f(x)$, then

$$\kappa = \frac{|f''(x)|}{\left(1 + (f'(x))^2\right)^{3/2}}.$$

13: Functions of Several Variables

A function of the form $y = f(x)$ is a function of a single variable; given a value of x, we can find a value y. Even the vector–valued functions of Chapter 12 are single–variable functions; the input is a single variable though the output is a vector.

There are many situations where a desired quantity is a function of two or more variables. For instance, wind chill is measured by knowing the temperature and wind speed; the volume of a gas can be computed knowing the pressure and temperature of the gas; to compute a baseball player's batting average, one needs to know the number of hits and the number of at–bats.

This chapter studies **multivariable** functions, that is, functions with more than one input.

13.1 Introduction to Multivariable Functions

Definition 80   **Function of Two Variables**

Let D be a subset of $\mathbb{R}^2$. A **function f of two variables** is a rule that assigns each pair (x, y) in D a value $z = f(x, y)$ in $\mathbb{R}$. The set D is the **domain** of f; the set of all outputs of f is the **range**.

Watch the video:
Finding and Sketching the Domain of a Multivariable Function at
`https://youtu.be/q8ictFvAHLk`

Example 1   **Understanding a function of two variables**

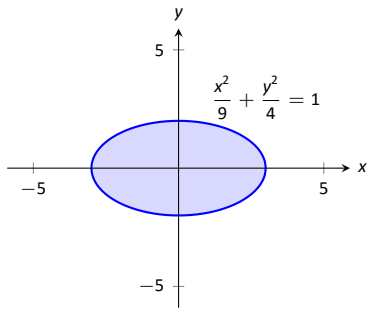Let $z = f(x, y) = x^2 - y$. Evaluate $f(1, 2)$, $f(2, 1)$, and $f(-2, 4)$; find the domain and range of f.

SOLUTION Using the definition $f(x, y) = x^2 - y$, we have:

$$f(1, 2) = 1^2 - 2 = -1$$
$$f(2, 1) = 2^2 - 1 = 3$$
$$f(-2, 4) = (-2)^2 - 4 = 0$$

The domain is not specified, so we take it to be all possible pairs in $\mathbb{R}^2$ for which f is defined. In this example, f is defined for *all* pairs (x, y), so the domain D of f is $\mathbb{R}^2$.

The output of f can be made as large or small as possible; any real number r can be the output. (In fact, given any real number r, $f(0, -r) = r$.) So the range R of f is $\mathbb{R}$.

Example 2   Understanding a function of two variables

Let $f(x, y) = \sqrt{1 - \dfrac{x^2}{9} - \dfrac{y^2}{4}}$. Find the domain and range of f.

SOLUTION The domain is all pairs (x, y) allowable as input in f. Because of the square–root, we need (x, y) such that $0 \leq 1 - \frac{x^2}{9} - \frac{y^2}{4}$:

$$\frac{x^2}{9} + \frac{y^2}{4} \leq 1$$

The above equation describes an ellipse and its interior as shown in Figure 13.1. We can represent the domain D graphically with the figure; in set notation, we can write $D = \{(x, y) | \frac{x^2}{9} + \frac{y^2}{4} \leq 1\}$.

The range is the set of all possible output values. The square–root ensures that all output is ≥ 0. Since the x and y terms are squared, then subtracted, inside the square-root, the largest output value comes at $x = 0, y = 0$: $f(0, 0) = 1$. Thus the range R is the interval $[0, 1]$.

Graphing Functions of Two Variables

The **graph** of a function f of two variables is the set of all points $(x, y, f(x, y))$ where (x, y) is in the domain of f. This creates a **surface** in space.

One can begin sketching a graph by plotting points, but this has limitations. Consider Figure 13.2(a) where 25 points have been plotted of $f(x, y) = \dfrac{1}{x^2 + y^2 + 1}$. More points have been plotted than one would reasonably want to do by hand, yet it is not clear at all what the graph of the function looks like. Technology allows us to plot lots of points, connect adjacent points with lines and add shading

Figure 13.1: Illustrating the domain of $f(x, y)$ in Example 2.

(a)

(b)
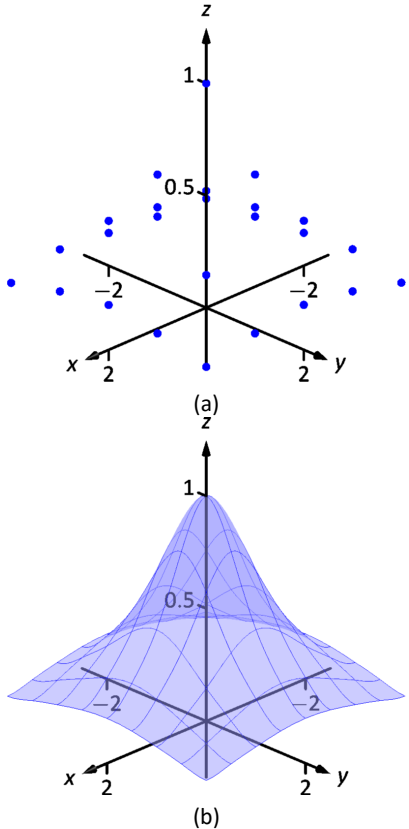
Figure 13.2: Graphing a function of two variables.

Notes:

to create a graph like Figure 13.2(b) which does a far better job of illustrating the behavior of f.

While technology is readily available to help us graph functions of two variables, there is still a paper–and–pencil approach that is useful to understand and master as it, combined with high–quality graphics, gives one great insight into the behavior of a function. This technique is known as sketching **level curves**.

Level Curves

It may be surprising to find that the problem of representing a three dimensional surface on paper is familiar to most people (they just don't realize it). Topographical maps, like the one shown in Figure 13.3, represent the surface of Earth by indicating points with the same elevation with **contour lines**. The elevations marked are equally spaced; in this example, each thin line indicates an elevation change in 50ft increments and each thick line indicates a change of 200ft. When lines are drawn close together, elevation changes rapidly (as one does not have to travel far to rise 50ft). When lines are far apart, such as near "Aspen Campground," elevation changes more gradually as one has to walk farther to rise 50ft.

Given a function $z = f(x, y)$, we can draw a "topographical map" of f by drawing **level curves** (or, contour lines). A level curve at $z = c$ is a curve in the x-y plane such that for all points (x, y) on the curve, $f(x, y) = c$.

When drawing level curves, it helps to evenly space the c values as that gives the best insight to how quickly the "elevation" is changing. Examples will help one understand this concept.

Example 3 **Drawing Level Curves**

Let $f(x, y) = \sqrt{1 - \dfrac{x^2}{9} - \dfrac{y^2}{4}}$. Find the level curves of f for $c = 0, 0.2, 0.4, 0.6, 0.8$ and 1.

SOLUTION Consider first $c = 0$. The level curve for $c = 0$ is the set of all points (x, y) such that $0 = \sqrt{1 - \frac{x^2}{9} - \frac{y^2}{4}}$. Squaring both sides gives us

$$\frac{x^2}{9} + \frac{y^2}{4} = 1,$$

an ellipse centered at $(0, 0)$ with horizontal major axis of length 6 and minor axis of length 4. Thus for any point (x, y) on this curve, $f(x, y) = 0$.
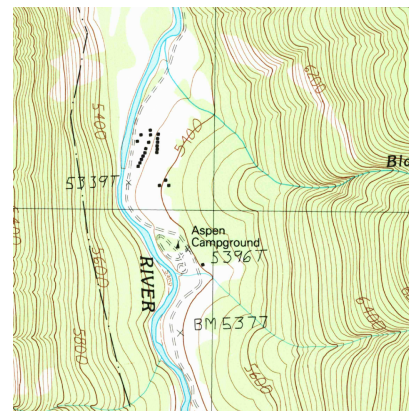

Figure 13.3: A topographical map displays elevation by drawing contour lines, along with the elevation is constant.

Sample taken from the public domain USGS Digital Raster Graphics, http://topmaps.usgs.gove/drg/.

Notes:

Now consider the level curve for $c = 0.2$

$$0.2 = \sqrt{1 - \frac{x^2}{9} - \frac{y^2}{4}}$$

$$0.04 = 1 - \frac{x^2}{9} - \frac{y^2}{4}$$

$$\frac{x^2}{9} + \frac{y^2}{4} = 0.96$$

$$\frac{x^2}{8.64} + \frac{y^2}{3.84} = 1.$$

This is also an ellipse, where $a = \sqrt{8.64} \approx 2.94$ and $b = \sqrt{3.84} \approx 1.96$. In general, for $z = c$, the level curve is:

$$c = \sqrt{1 - \frac{x^2}{9} - \frac{y^2}{4}}$$

$$c^2 = 1 - \frac{x^2}{9} - \frac{y^2}{4}$$

$$\frac{x^2}{9} + \frac{y^2}{4} = 1 - c^2$$

$$\frac{x^2}{9(1 - c^2)} + \frac{y^2}{4(1 - c^2)} = 1,$$

ellipses that are decreasing in size as c increases. A special case is when $c = 1$; there the ellipse is just the point $(0, 0)$.

The level curves are shown in Figure 13.4(a). Note how the level curves for $c = 0$ and $c = 0.2$ are very, very close together: this indicates that f is growing rapidly along those curves.

In Figure 13.4(b), the curves are drawn on a graph of f in space. Note how the elevations are evenly spaced. Near the level curves of $c = 0$ and $c = 0.2$ we can see that f indeed is growing quickly.

Example 4   Analyzing Level Curves

Let $f(x, y) = \dfrac{x + y}{x^2 + y^2 + 1}$. Find the level curves for $z = c$.

SOLUTION   We begin by setting $f(x, y) = c$ for an arbitrary c and seeing if algebraic manipulation of the equation reveals anything significant.

$$\frac{x + y}{x^2 + y^2 + 1} = c$$

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

Notes:

(a)

(b)
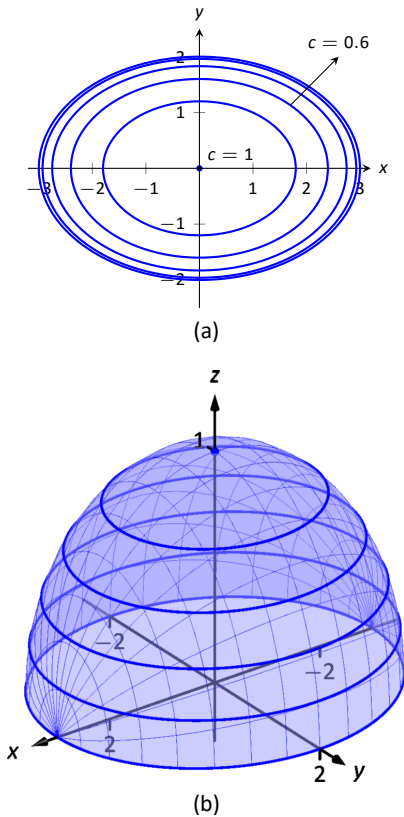
Figure 13.4: Graphing the level curves in Example 3.

We recognize this as a circle, though the center and radius are not yet clear. By completing the square, we can obtain:

$$\left(x - \frac{1}{2c}\right)^2 + \left(y - \frac{1}{2c}\right)^2 = \frac{1}{2c^2} - 1,$$

a circle centered at $\left(1/(2c), 1/(2c)\right)$ with radius $\sqrt{1/(2c^2) - 1}$, where $|c| < 1/\sqrt{2}$. The level curves for $c = \pm 0.2$, ± 0.4 and ± 0.6 are sketched in Figure 13.5(a). To help illustrate "elevation," we use thicker lines for c values near 0, and dashed lines indicate where $c < 0$.

There is one special level curve, when $c = 0$. The level curve in this situation is $x + y = 0$, the line $y = -x$.

In Figure 13.5(b) we see a graph of the surface. Note how the y-axis is pointing away from the viewer to more closely resemble the orientation of the level curves in (a).

Seeing the level curves helps us understand the graph. For instance, the graph does not make it clear that one can "walk" along the line $y = -x$ without elevation change, though the level curve does.
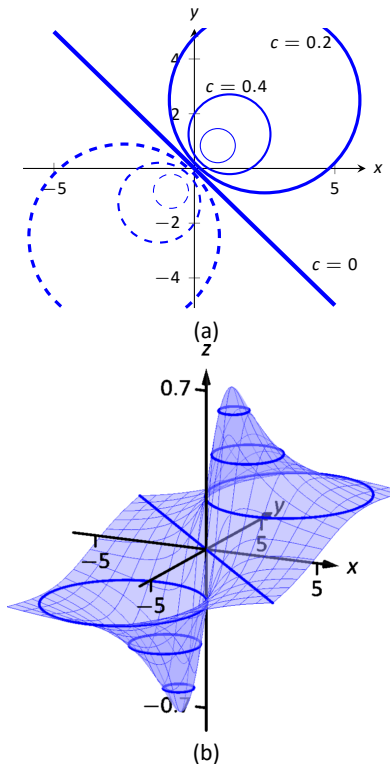
Functions of Three Variables

We extend our study of multivariable functions to functions of three variables. (One can make a function of as many variables as one likes; we limit our study to three variables so that we are able to view the domain without exceeding three dimensions.)

> **Definition 81   Function of Three Variables**
> Let D be a subset of $\mathbb{R}^3$. A **function f of three variables** is a rule that assigns each triple (x, y, z) in D a value $w = f(x, y, z)$ in $\mathbb{R}$. The set D is the **domain** of f; the set of all outputs of f is the **range**.

Note how this definition closely resembles that of Definition 80.

Example 5   Understanding a function of three variables
Let $f(x, y, z) = \dfrac{x^2 + z + 3\sin y}{x + 2y - z}$. Evaluate f at the point $(3, 0, 2)$ and find the domain and range of f.

(a)

(b)

Figure 13.5: Graphing the level curves in Example 4.

Notes:

SOLUTION $f(3, 0, 2) = \dfrac{3^2 + 2 + 3\sin 0}{3 + 2(0) - 2} = 11.$

As the domain of f is not specified, we take it to be the set of all triples (x, y, z) for which $f(x, y, z)$ is defined. As we cannot divide by 0, we find the domain D is

$$D = \{(x, y, z) \mid x + 2y - z \neq 0\}.$$

We recognize that the set of all points in $\mathbb{R}^3$ that *are not* in D form a plane in space that passes through the origin (with normal vector $\langle 1, 2, -1\rangle$).

We determine the range R is $\mathbb{R}$; that is, all real numbers are possible outputs of f. There is no set way of establishing this. Rather, to get numbers near 0 we can let $y = 0$ and choose $z \approx -x^2$. To get numbers of arbitrarily large magnitude, we can let $z \approx x + 2y$.

Level Surfaces

It is very difficult to produce a meaningful graph of a function of three variables. A function of *one* variable is a *curve* drawn in *2* dimensions; a function of *two* variables is a *surface* drawn in *3* dimensions; a function of *three* variables is a *hypersurface* drawn in *4* dimensions.

There are a few techniques one can employ to try to "picture" a graph of three variables. One is an analogue of level curves: **level surfaces**. Given $w = f(x, y, z)$, the level surface at $w = c$ is the surface in space formed by all points (x, y, z) where $f(x, y, z) = c$.

Example 6   **Finding level surfaces**

If a point source S is radiating energy, the intensity I at a given point P in space is inversely proportional to the square of the distance between S and P. That is, when $S = (0, 0, 0)$, $I(x, y, z) = \dfrac{k}{x^2 + y^2 + z^2}$ for some constant k.

Let $k = 1$; find the level surfaces of I.

SOLUTION   We can (mostly) answer this question using "common sense." If energy (say, in the form of light) is emanating from the origin, its intensity will be the same at all points equidistant from the origin. That is, at any point on the surface of a sphere centered at the origin, the intensity should be the same. Therefore, the level surfaces are spheres.

We now find this mathematically. The level surface at $I = c$ is defined by

$$c = \frac{1}{x^2 + y^2 + z^2}.$$

Notes:

taking reciprocals reveals

$$x^2 + y^2 + z^2 = \frac{1}{c}.$$

Given an intensity c, the level surface $I = c$ is a sphere of radius $1/\sqrt{c}$, centered at the origin.

Figure 13.6 gives a table of the radii of the spheres for given c values. Normally one would use equally spaced c values, but these values have been chosen purposefully. At a distance of 0.25 from the point source, the intensity is 16; to move to a point of half that intensity, one just moves out 0.1 to 0.35 – not much at all. To again halve the intensity, one moves 0.15, a little more than before.

Note how each time the intensity if halved, the distance required to move away grows. We conclude that the closer one is to the source, the more rapidly the intensity changes.

In the next section we apply the concepts of limits to functions of two or more variables.

c	r
16.	0.25
8.	0.35
4.	0.5
2.	0.71
1.	1.
0.5	1.41
0.25	2.
0.125	2.83
0.0625	4.

Figure 13.6: A table of c values and the corresponding radius r of the spheres of constant value in Example 6.

Notes:

Exercises 13.1

Terms and Concepts

1. Give two examples (other than those given in the text) of "real world" functions that require more than one input.

2. The graph of a function of two variables is a _____.

3. Most people are familiar with the concept of level curves in the context of _____ maps.

4. T/F: Along a level curve, the output of a function does not change.

5. The analogue of a level curve for functions of three variables is a level _____.

6. What does it mean when level curves are close together? Far apart?

Problems

In Exercises 7–14, give the domain and range of the multivariable function.

7. $f(x, y) = x^2 + y^2 + 2$

8. $f(x, y) = x + 2y$

9. $f(x, y) = x - 2y$

10. $f(x, y) = \dfrac{1}{x + 2y}$

11. $f(x, y) = \dfrac{1}{x^2 + y^2 + 1}$

12. $f(x, y) = \sin x \cos y$

13. $f(x, y) = \sqrt{9 - x^2 - y^2}$

14. $f(x, y) = \dfrac{1}{\sqrt{x^2 + y^2 - 9}}$

In Exercises 15–22, describe in words and sketch the level curves for the function and given c values.

15. $f(x, y) = 3x - 2y; c = -2, 0, 2$

16. $f(x, y) = x^2 - y^2; c = -1, 0, 1$

17. $f(x, y) = x - y^2; c = -2, 0, 2$

18. $f(x, y) = \dfrac{1 - x^2 - y^2}{2y - 2x}; c = -2, 0, 2$

19. $f(x, y) = \dfrac{2x - 2y}{x^2 + y^2 + 1}; c = -1, 0, 1$

20. $f(x, y) = \dfrac{y - x^3 - 1}{x}; c = -3, -1, 0, 1, 3$

21. $f(x, y) = \sqrt{x^2 + 4y^2}; c = 1, 2, 3, 4$

22. $f(x, y) = x^2 + 4y^2; c = 1, 2, 3, 4$

In Exercises 23–26, give the domain and range of the functions of three variables.

23. $f(x, y, z) = \dfrac{x}{x + 2y - 4z}$

24. $f(x, y, z) = \dfrac{1}{1 - x^2 - y^2 - z^2}$

25. $f(x, y, z) = \sqrt{z - x^2 + y^2}$

26. $f(x, y, z) = z^2 \sin x \cos y$

In Exercises 27–30, describe the level surfaces of the given functions of three variables.

27. $f(x, y, z) = x^2 + y^2 + z^2$

28. $f(x, y, z) = z - x^2 + y^2$

29. $f(x, y, z) = \dfrac{x^2 + y^2}{z}$

30. $f(x, y, z) = \dfrac{z}{x - y}$

31. Compare the level curves of Exercises 21 and 22. How are they similar, and how are they different? Each surface is a quadric surface; describe how the level curves are consistent with what we know about each surface.

13.2 Limits and Continuity of Multivariable Functions

We continue with the pattern we have established in this text: after defining a new kind of function, we apply calculus ideas to it. The previous section defined functions of two and three variables; this section investigates what it means for these functions to be "continuous."

We begin with a series of definitions. We are used to "open intervals" such as $(1, 3)$, which represents the set of all x such that $1 < x < 3$, and "closed intervals" such as $[1, 3]$, which represents the set of all x such that $1 \leq x \leq 3$. We need analogous definitions for open and closed sets in the x-y plane.

Definition 82   **Open Disk, Boundary and Interior Points, Open and Closed Sets, Bounded Sets**

An **open disk** B in $\mathbb{R}^2$ centered at (x_0, y_0) with radius r is the set of all points (x, y) such that $\sqrt{(x - x_0)^2 + (y - y_0)^2} < r$.

Let S be a set of points in $\mathbb{R}^2$. A point P in $\mathbb{R}^2$ is a **boundary point** of S if all open disks centered at P contain both points in S and points not in S.

A point P in S is an **interior point** of S if there is an open disk centered at P that contains only points in S.

A set S is **open** if every point in S is an interior point.

A set S is **closed** if it contains all of its boundary points.

A set S is **bounded** if there is an $M > 0$ such that the open disk, centered at the origin with radius M, contains S. A set that is not bounded is **unbounded**.

Figure 13.7 shows several sets in the x-y plane. In each set, point P_1 lies on the boundary of the set as all open disks centered there contain both points in, and not in, the set. In contrast, point P_2 is an interior point for there is an open disk centered there that lies entirely within the set.

The set depicted in Figure 13.7(a) is a closed set as it contains all of its boundary points. The set in (b) is open, for all of its points are interior points (or, equivalently, it does not contain any of its boundary points). The set in (c) is neither open nor closed as it contains some of its boundary points.
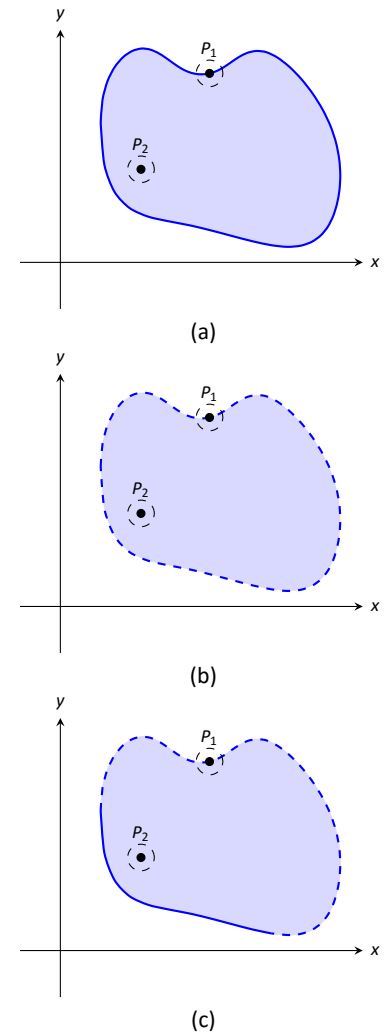
(a)

(b)

(c)

Figure 13.7: Illustrating open and closed sets in the x-y plane.

Notes:

Example 1   **Determining open/closed, bounded/unbounded**

Determine if the domain of the function $f(x,y) = \sqrt{1 - \frac{x^2}{9} - \frac{y^2}{4}}$ is open, closed, or neither, and if it is bounded.

> **SOLUTION**   This domain of this function was found in Example 13.1.2 to be $D = \{(x,y) \mid \frac{x^2}{9} + \frac{y^2}{4} \leq 1\}$, the region *bounded* by the ellipse $\frac{x^2}{9} + \frac{y^2}{4} = 1$. Since the region includes the boundary (indicated by the use of "$\leq$"), the set contains all of its boundary points and hence is closed. The region is bounded as a disk of radius 4, centered at the origin, contains D.

Example 2   **Determining open/closed, bounded/unbounded**

Determine if the domain of $f(x,y) = \frac{1}{x-y}$ is open, closed, or neither.

> **SOLUTION**   As we cannot divide by 0, we find the domain to be $D = \{(x,y) \mid x - y \neq 0\}$. In other words, the domain is the set of all points (x,y) *not* on the line $y = x$.
>
> The domain is sketched in Figure 13.8. Note how we can draw an open disk around any point in the domain that lies entirely inside the domain, and also note how the only boundary points of the domain are the points on the line $y = x$. We conclude the domain is an open set. The set is unbounded.
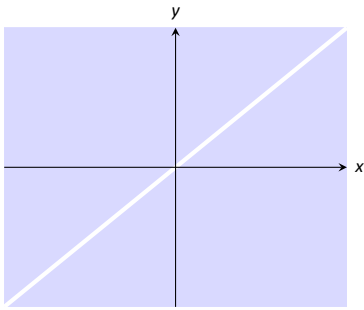

Figure 13.8: Sketching the domain of the function in Example 2.

Limits

Recall a pseudo–definition of the limit of a function of one variable: "$\lim_{x \to c} f(x) = L$" means that if x is "really close" to c, then $f(x)$ is "really close" to L. A similar pseudo–definition holds for functions of two variables. We'll say that

$$\text{"} \lim_{(x,y) \to (x_0, y_0)} f(x,y) = L \text{"}$$

means "if the point (x,y) is really close to the point (x_0, y_0), then $f(x,y)$ is really close to L." The formal definition is given below.

Notes:

> **Definition 83   Limit of a Function of Two Variables**
>
> Let S be an open set containing (x_0, y_0), and let f be a function of two variables defined on S, except possibly at (x_0, y_0). The **limit** of $f(x, y)$ as (x, y) approaches (x_0, y_0) is L, denoted
>
> $$\lim_{(x,y)\to(x_0,y_0)} f(x, y) = L,$$
>
> means that given any $\varepsilon > 0$, there exists $\delta > 0$ such that for all $(x, y) \neq (x_0, y_0)$, if (x, y) is in the open disk centered at (x_0, y_0) with radius δ, then $|f(x, y) - L| < \varepsilon$.

The concept behind Definition 83 is sketched in Figure 13.9. Given $\varepsilon > 0$, find $\delta > 0$ such that if (x, y) is any point in the open disk centered at (x_0, y_0) in the *x-y* plane with radius δ, then $f(x, y)$ should be within ε of L.

Figure 13.9: **Illustrating the definition of a limit.** The open disk in the *x-y* plane has radius δ. Let (x, y) be any point in this disk; $f(x, y)$ is within ε of L.

Computing limits using this definition is rather cumbersome. The following theorem allows us to evaluate limits much more easily.
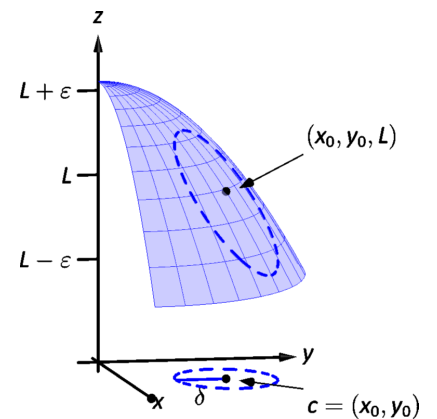
Notes:

> **Theorem 103**   **Basic Limit Properties of Functions of Two Variables**
> Let b, x_0, y_0, L and K be real numbers, let n be a positive integer, and let f and g be functions with the following limits:
>
> $$\lim_{(x,y)\to(x_0,y_0)} f(x,y) = L \quad \text{and} \quad \lim_{(x,y)\to(x_0,y_0)} g(x,y) = K.$$
>
> The following limits hold.
>
> 1. Constants: $\displaystyle\lim_{(x,y)\to(x_0,y_0)} b = b$
>
> 2. Identity $\displaystyle\lim_{(x,y)\to(x_0,y_0)} x = x_0; \quad \lim_{(x,y)\to(x_0,y_0)} y = y_0$
>
> 3. Sums/Differences: $\displaystyle\lim_{(x,y)\to(x_0,y_0)} \big(f(x,y) \pm g(x,y)\big) = L \pm K$
>
> 4. Scalar Multiples: $\displaystyle\lim_{(x,y)\to(x_0,y_0)} b \cdot f(x,y) = bL$
>
> 5. Products: $\displaystyle\lim_{(x,y)\to(x_0,y_0)} f(x,y) \cdot g(x,y) = LK$
>
> 6. Quotients: $\displaystyle\lim_{(x,y)\to(x_0,y_0)} f(x,y)/g(x,y) = L/K, \, (K \neq 0)$
>
> 7. Powers: $\displaystyle\lim_{(x,y)\to(x_0,y_0)} f(x,y)^n = L^n$
>
> 8. Roots: $\displaystyle\lim_{(x,y)\to(x_0,y_0)} \sqrt[n]{f(x,y)} = \sqrt[n]{L}$ (when n is odd
>   or $L \geq 0$)

This theorem can be proved by the same arguments as the analogous results for functions of one variable in Theorem 1. Combined with Theorems 3 and 4 of Section 1.3, this allows us to evaluate many limits.

Example 3   **Evaluating a limit**
Evaluate the following limits:

$$1. \ \lim_{(x,y)\to(1,\pi)} \frac{y}{x} + \cos(xy) \qquad\qquad 2. \ \lim_{(x,y)\to(0,0)} \frac{3xy}{x^2 + y^2}$$

SOLUTION

1. The aforementioned theorems allow us to simply evaluate $y/x + \cos(xy)$ when $x = 1$ and $y = \pi$. If an indeterminate form is returned, we must do

Notes:

more work to evaluate the limit; otherwise, the result is the limit. There-fore

$$\lim_{(x,y)\to(1,\pi)} \frac{y}{x} + \cos(xy) = \frac{\pi}{1} + \cos\pi$$

$$= \pi - 1.$$

2. We attempt to evaluate the limit by substituting 0 in for x and y, but the result is the indeterminate form "$0/0$." To evaluate this limit, we must "do more work," but we have not yet learned what "kind" of work to do. Therefore we cannot yet evaluate this limit.

When dealing with functions of a single variable we also considered one–sided limits and stated

$$\lim_{x\to c} f(x) = L \quad \text{if and only if both} \quad \lim_{x\to c^+} f(x) = L \quad \text{and} \quad \lim_{x\to c^-} f(x) = L.$$

That is, the limit is L if and only if $f(x)$ approaches L when x approaches c from **either** direction, the left or the right.

In the plane, there are infinite directions from which (x, y) might approach (x_0, y_0). In fact, we do not have to restrict ourselves to approaching (x_0, y_0) from a particular direction, but rather we can approach that point along any possible path. It is possible to arrive at different limiting values by approaching (x_0, y_0) along different paths. If this happens, we say that $\lim_{(x,y)\to(x_0,y_0)} f(x, y)$ does not exist (this is analogous to the left and right hand limits of single variable functions not being equal).

Our theorems tell us that we can evaluate most limits quite simply, without worrying about paths. When indeterminate forms arise, the limit may or may not exist. If it does exist, it can be difficult to prove this as we need to show the same limiting value is obtained regardless of the path chosen. The case where the limit does not exist is often easier to deal with, for we can often pick two paths along which the limit is different.

Watch the video:
Showing a Limit Does Not Exist at
`https://youtu.be/q9xIdF33ql8`

Notes:

Example 4   **Showing limits do not exist**

1. Show $\displaystyle\lim_{(x,y)\to(0,0)} \frac{3xy}{x^2 + y^2}$ does not exist by finding the limits along the lines $y = mx$.

2. Show $\displaystyle\lim_{(x,y)\to(0,0)} \frac{\sin(xy)}{x + y}$ does not exist by finding the limit along the path $y = -\sin x$.

SOLUTION

1. Evaluating $\displaystyle\lim_{(x,y)\to(0,0)} \frac{3xy}{x^2 + y^2}$ along the lines $y = mx$ means replace all y's with mx and evaluating the resulting limit:

$$\lim_{(x,mx)\to(0,0)} \frac{3x(mx)}{x^2 + (mx)^2} = \lim_{x\to 0} \frac{3mx^2}{x^2(m^2 + 1)}$$
$$= \lim_{x\to 0} \frac{3m}{m^2 + 1}$$
$$= \frac{3m}{m^2 + 1}.$$

While the limit exists for each choice of m, we get a *different* limit for each choice of m. That is, along different lines we get differing limiting values, meaning *the* limit does not exist.

2. Let $f(x, y) = \frac{\sin(xy)}{x+y}$. We are to show that $\displaystyle\lim_{(x,y)\to(0,0)} f(x, y)$ does not exist by finding the limit along the path $y = -\sin x$. First, however, consider the limits found along the lines $y = mx$ as done above.

$$\lim_{(x,mx)\to(0,0)} \frac{\sin\big(x(mx)\big)}{x + mx} = \lim_{x\to 0} \frac{\sin(mx^2)}{x(m + 1)}$$
$$= \lim_{x\to 0} \frac{\sin(mx^2)}{x} \cdot \frac{1}{m + 1}.$$

By applying L'Hôpital's Rule, we can show this limit is 0 *except* when $m = -1$, that is, along the line $y = -x$. This line is not in the domain of f, so we have found the following fact: along every line $y = mx$ in the domain of f, $\displaystyle\lim_{(x,y)\to(0,0)} f(x, y) = 0$.

Now consider the limit along the path $y = -\sin x$:

$$\lim_{(x,-\sin x)\to(0,0)} \frac{\sin\big(-x\sin x\big)}{x - \sin x} = \lim_{x\to 0} \frac{\sin\big(-x\sin x\big)}{x - \sin x}$$

Notes:

Now apply L'Hôpital's Rule twice:

$$= \lim_{x \to 0} \frac{\cos\left(-x\sin x\right)\left(-\sin x - x\cos x\right)}{1 - \cos x} \quad (\text{``} = 0/0\text{''})$$

$$= \lim_{x \to 0} \frac{-\sin\left(-x\sin x\right)\left(-\sin x - x\cos x\right)^2 + \cos\left(-x\sin x\right)\left(-2\cos x + x\sin x\right)}{\sin x}$$

$$= \text{``}-2/0\text{''} \Rightarrow \text{the limit does not exist.}$$

Step back and consider what we have just discovered. Along any line $y = mx$ in the domain of the $f(x, y)$, the limit is 0. However, along the path $y = -\sin x$, which lies in the domain of $f(x, y)$ for all $x \neq 0$, the limit does not exist. Since the limit is not the same along every path to $(0, 0)$, we say

$$\lim_{(x,y)\to(0,0)} \frac{\sin(xy)}{x + y} \text{ does not exist.}$$

Example 5   **Finding a limit**

Let $f(x, y) = \dfrac{5x^2 y^2}{x^2 + y^2}$. Find $\displaystyle\lim_{(x,y)\to(0,0)} f(x, y)$.

SOLUTION   It is relatively easy to show that along any line $y = mx$, the limit is 0. This is not enough to prove that the limit exists, as demonstrated in the previous example, but it tells us that if the limit does exist then it must be 0.

To prove the limit is 0, we apply Definition 83. Let $\varepsilon > 0$ be given. We want to find $\delta > 0$ such that if $\sqrt{(x - 0)^2 + (y - 0)^2} < \delta$, then $|f(x, y) - 0| < \varepsilon$.

Set $\delta < \sqrt{\varepsilon/5}$. Note that $\left| \dfrac{5y^2}{x^2 + y^2} \right| < 5$ for all $(x, y) \neq (0, 0)$, and that if $\sqrt{x^2 + y^2} < \delta$, then $x^2 < \delta^2$.

Let $\sqrt{(x - 0)^2 + (y - 0)^2} = \sqrt{x^2 + y^2} < \delta$. Consider $|f(x, y) - 0|$:

$$|f(x, y) - 0| = \left| \frac{5x^2 y^2}{x^2 + y^2} - 0 \right|$$

$$= \left| x^2 \cdot \frac{5y^2}{x^2 + y^2} \right|$$

$$< \delta^2 \cdot 5$$

$$< \frac{\varepsilon}{5} \cdot 5$$

$$= \varepsilon.$$

Thus if $\sqrt{(x - 0)^2 + (y - 0)^2} < \delta$ then $|f(x, y) - 0| < \varepsilon$, which is what we wanted to show. Thus $\displaystyle\lim_{(x,y)\to(0,0)} \frac{5x^2 y^2}{x^2 + y^2} = 0$.

Notes:

Continuity

Definition 7 defines what it means for a function of one variable to be continuous. In brief, it meant that the function always equaled its limit. We define continuity for functions of two variables in a similar way as we did for functions of one variable.

Definition 84   Continuous

Let a function $f(x, y)$ be defined on an open disk B containing the point (x_0, y_0).

1. f is **continuous** at (x_0, y_0) if $\lim\limits_{(x,y)\to(x_0,y_0)} f(x, y) = f(x_0, y_0)$.

2. f is **continuous on an open set** S if f is continuous at each points in S. (We say that f is **continuous everywhere** if f is continuous on $\mathbb{R}^2$.)

Example 6   Continuity of a function of two variables

Let $f(x, y) = \begin{cases} \frac{\cos y \sin x}{x} & x \neq 0 \\ \cos y & x = 0 \end{cases}$. Is f continuous at $(0, 0)$? Is f continuous everywhere?

SOLUTION   To determine if f is continuous at $(0, 0)$, we need to compare $\lim\limits_{(x,y)\to(0,0)} f(x, y)$ to $f(0, 0)$.

Applying the definition of f, we see that $f(0, 0) = \cos 0 = 1$.

We now consider the limit $\lim\limits_{(x,y)\to(0,0)} f(x, y)$. Substituting 0 for x and y in $(\cos y \sin x)/x$ returns the indeterminate form "0/0", so we need to do more work to evaluate this limit.

Consider two related limits: $\lim\limits_{(x,y)\to(0,0)} \cos y$ and $\lim\limits_{(x,y)\to(0,0)} \frac{\sin x}{x}$. The first limit does not contain x, and since $\cos y$ is continuous,

$$\lim_{(x,y)\to(0,0)} \cos y = \lim_{y\to 0} \cos y = \cos 0 = 1.$$

The second limit does not contain y. By Theorem 6 we can say

$$\lim_{(x,y)\to(0,0)} \frac{\sin x}{x} = \lim_{x\to 0} \frac{\sin x}{x} = 1.$$

Notes:

Finally, Theorem 103 of this section states that we can combine these two limits as follows:

$$\lim_{(x,y)\to(0,0)} \frac{\cos y \sin x}{x} = \lim_{(x,y)\to(0,0)} (\cos y)\left(\frac{\sin x}{x}\right)$$

$$= \left(\lim_{(x,y)\to(0,0)} \cos y\right)\left(\lim_{(x,y)\to(0,0)} \frac{\sin x}{x}\right)$$

$$= (1)(1)$$

$$= 1.$$

We have found that $\lim_{(x,y)\to(0,0)} \frac{\cos y \sin x}{x} = f(0,0)$, so f is continuous at $(0,0)$.

A similar analysis shows that f is continuous at all points in $\mathbb{R}^2$. As long as $x \neq 0$, we can evaluate the limit directly; when $x = 0$, a similar analysis shows that the limit is $\cos y$. Thus we can say that f is continuous everywhere. A graph of f is given in Figure 13.10. Notice how it has no breaks, jumps, etc.

The following theorems are very similar to Theorems 10 and 11, giving us ways to combine continuous functions to create other continuous functions.

Figure 13.10: A graph of $f(x,y)$ in Example 6.

> **Theorem 104   Properties of Continuous Functions**
> Let f and g be continuous on an open set S, let c be a real number, and let n be a positive integer. The following functions are continuous on S.
>
> 1. Sums/Differences: $f \pm g$
> 2. Constant Multiples: $c \cdot f$
> 3. Products: $f \cdot g$
> 4. Quotients: f/g (as longs as $g \neq 0$ on B)
> 5. Powers: f^n
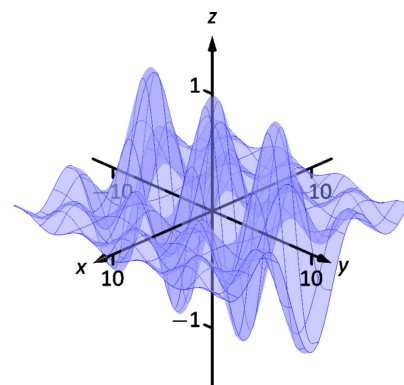> 6. Roots: $\sqrt[n]{f}$ (if $f \geq 0$ on B or n is odd)

Notes:

> **Theorem 105   Continuity of Compositions**
> Let f be continuous on S, where the range of f on S is J, and let g be a single variable function that is continuous on J. Then
>
> $$(g \circ f)(x, y) = g(f(x, y)),$$
>
> is continuous on S.

Example 7   Establishing continuity of a function
Let $f(x, y) = \sin(x^2 \cos y)$. Show f is continuous everywhere.

SOLUTION   We will apply Theorems 10, 104, and 105. Let $f_1(x, y) = x^2$. Since y is not actually used in the function, and polynomials are continuous (by Theorem 10), we conclude f_1 is continuous everywhere. A similar statement can be made about $f_2(x, y) = \cos y$. Part 3 of Theorem 104 states that $f_3 = f_1 \cdot f_2$ is continuous everywhere, and Theorem 105 states the composition of sine with f_3 is continuous: that is, $\sin(f_3) = \sin(x^2 \cos y)$ is continuous everywhere.

Functions of Three Variables

The definitions and theorems given in this section can be extended in a natural way to definitions and theorems about functions of three (or more) variables. We cover the key concepts here; some terms from Definitions 82 and 84 are not redefined but their analogous meanings should be clear to the reader.

Notes:

Definition 85   Open Balls, Limit, Continuous

1. An **open ball** in $\mathbb{R}^3$ centered at (x_0, y_0, z_0) with radius r is the set of all points (x, y, z) such that $\sqrt{(x - x_0)^2 + (y - y_0)^2 + (z - z_0)^2} = r$.

2. Let D be an open set in $\mathbb{R}^3$ containing (x_0, y_0, z_0), and let $f(x, y, z)$ be a function of three variables defined on D, except possibly at (x_0, y_0, z_0). The **limit** of $f(x, y, z)$ as (x, y, z) approaches (x_0, y_0, z_0) is L, denoted

$$\lim_{(x,y,z) \to (x_0, y_0, z_0)} f(x, y, z) = L,$$

means that given any $\varepsilon > 0$, there is a $\delta > 0$ such that for all $(x, y, z) \neq (x_0, y_0, z_0)$, if (x, y, z) is in the open ball centered at (x_0, y_0, z_0) with radius δ, then $|f(x, y, z) - L| < \varepsilon$.

3. Let $f(x, y, z)$ be defined on an open ball B containing (x_0, y_0, z_0). Then f is **continuous** at (x_0, y_0, z_0) if $\lim\limits_{(x,y,z) \to (x_0, y_0, z_0)} f(x, y, z) = f(x_0, y_0, z_0)$.

These definitions can also be extended naturally to apply to functions of four or more variables. Theorems 104 and 105 also applies to function of three or more variables, allowing us to say that the function

$$f(x, y, z) = \frac{e^{x^2 + y}\sqrt{y^2 + z^2 + 3}}{\sin(xyz) + 5}$$

is continuous everywhere.

When considering single variable functions, we studied limits, then continuity, then the derivative. In our current study of multivariable functions, we have studied limits and continuity. In the next section we study derivation, which takes on a slight twist as we are in a multivarible context.

Notes:

Exercises 13.2

Terms and Concepts

1. Describe in your own words the difference between boundary and interior point of a set.
2. Use your own words to describe (informally) what $\lim\limits_{(x,y)\to(1,2)} f(x,y) = 17$ means.
3. Give an example of a closed, bounded set.
4. Give an example of a closed, unbounded set.
5. Give an example of a open, bounded set.
6. Give an example of a open, unbounded set.

Problems

In Exercises 7–10, a set S is given.

(a) Give one boundary point and one interior point, when possible, of S.

(b) State whether S is open, closed, or neither.

(c) State whether S is bounded or unbounded.

7. $S = \left\{ (x,y) \,\middle|\, \dfrac{(x-1)^2}{4} + \dfrac{(y-3)^2}{9} \leq 1 \right\}$

8. $S = \left\{ (x,y) \,\middle|\, y \neq x^2 \right\}$

9. $S = \left\{ (x,y) \,\middle|\, x^2 + y^2 = 1 \right\}$

10. $S = \left\{ (x,y) \,\middle|\, y > \sin x \right\}$

In Exercises 11–14:

(a) Find the domain D of the given function.

(b) State whether D is an open or closed set.

(c) State whether D is bounded or unbounded.

11. $f(x,y) = \sqrt{9 - x^2 - y^2}$

12. $f(x,y) = \sqrt{y - x^2}$

13. $f(x,y) = \dfrac{1}{\sqrt{y - x^2}}$

14. $f(x,y) = \dfrac{x^2 - y^2}{x^2 + y^2}$

In Exercises 15–20, a limit is given. Evaluate the limit along the paths given, then state why these results show the given limit does not exist.

15. $\lim\limits_{(x,y)\to(0,0)} \dfrac{x^2 - y^2}{x^2 + y^2}$

(a) Along the path $y = 0$.

(b) Along the path $x = 0$.

16. $\lim\limits_{(x,y)\to(0,0)} \dfrac{x + y}{x - y}$

(a) Along the path $y = mx$.

17. $\lim\limits_{(x,y)\to(0,0)} \dfrac{xy - y^2}{y^2 + x}$

(a) Along the path $y = mx$.

(b) Along the path $x = 0$.

18. $\lim\limits_{(x,y)\to(0,0)} \dfrac{\sin(x^2)}{y}$

(a) Along the path $y = mx$.

(b) Along the path $y = x^2$.

19. $\lim\limits_{(x,y)\to(1,2)} \dfrac{x + y - 3}{x^2 - 1}$

(a) Along the path $y = 2$.

(b) Along the path $y = x + 1$.

20. $\lim\limits_{(x,y)\to(\pi,\pi/2)} \dfrac{\sin x}{\cos y}$

(a) Along the path $x = \pi$.

(b) Along the path $y = x - \pi/2$.

13.3 Partial Derivatives

Let y be a function of x. We have studied in great detail the derivative of y with respect to x, that is, $\frac{dy}{dx}$, which measures the rate at which y changes with respect to x. Consider now $z = f(x, y)$. It makes sense to want to know how z changes with respect to x and/or y. This section begins our investigation into these rates of change.

Consider the function $z = f(x, y) = x^2 + 2y^2$, as graphed in Figure 13.11(a). By fixing $y = 2$, we focus our attention to all points on the surface where the y-value is 2, shown in both parts (a) and (b) of the figure. These points form a curve in space: $z = f(x, 2) = x^2 + 8$ which is a function of just one variable. We can take the derivative of z with respect to x along this curve and find equations of tangent lines, etc.

The key notion to extract from this example is: by treating y as constant (it does not vary) we can consider how z changes with respect to x. In a similar fashion, we can hold x constant and consider how z changes with respect to y. This is the underlying principle of **partial derivatives**. We state the formal, limit–based definition first, then show how to compute these partial derivatives without directly taking limits.

Definition 86   Partial Derivative

Let $z = f(x, y)$ be a continuous function on an open set S in $\mathbb{R}^2$.

1. The **partial derivative of f with respect to x** is:

$$f_x(x, y) = \lim_{h \to 0} \frac{f(x + h, y) - f(x, y)}{h}.$$

2. The **partial derivative of f with respect to y** is:

$$f_y(x, y) = \lim_{h \to 0} \frac{f(x, y + h) - f(x, y)}{h}.$$

Watch the video:
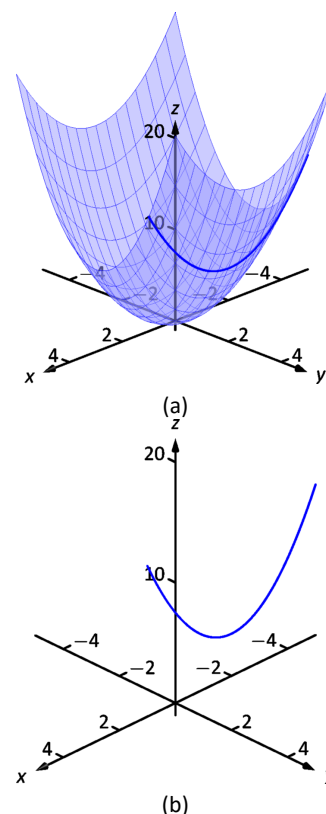Partial Derivatives at
`https://youtu.be/SbfRDBmyAMI`

(a)

(b)

Figure 13.11: By fixing $y = 2$, the surface $f(x, y) = x^2 + 2y^2$ is a curve in space.

Alternate notations for $f_x(x, y)$ include:

$$\frac{\partial}{\partial x}f(x, y), \quad \frac{\partial f}{\partial x}, \quad \frac{\partial z}{\partial x}, \quad \text{and } z_x,$$

with similar notations for $f_y(x, y)$. For ease of notation, $f_x(x, y)$ is often abbreviated f_x.

Notes:

Example 1   **Computing partial derivatives with the limit definition**

Let $f(x, y) = x^2 y + 2x + y^3$. Find $f_x(x, y)$ using the limit definition.

> **SOLUTION**   Using Definition 86, we have:
>
> $$\begin{aligned}
> f_x(x, y) &= \lim_{h \to 0} \frac{f(x + h, y) - f(x, y)}{h} \\
> &= \lim_{h \to 0} \frac{(x + h)^2 y + 2(x + h) + y^3 - (x^2 y + 2x + y^3)}{h} \\
> &= \lim_{h \to 0} \frac{x^2 y + 2xhy + h^2 y + 2x + 2h + y^3 - (x^2 y + 2x + y^3)}{h} \\
> &= \lim_{h \to 0} \frac{2xhy + h^2 y + 2h}{h} \\
> &= \lim_{h \to 0} 2xy + hy + 2 \\
> &= 2xy + 2.
> \end{aligned}$$

We have found $f_x(x, y) = 2xy + 2$.

Example 1 found a partial derivative using the formal, limit–based definition. Using limits is not necessary, though, as we can rely on our previous knowledge of derivatives to compute partial derivatives easily. When computing $f_x(x, y)$, we hold y fixed – it does not vary. Therefore we can compute the derivative with respect to x by treating y as a constant or coefficient.

Just as $\frac{d}{dx}\left(5x^2\right) = 10x$, we compute $\frac{\partial}{\partial x}\left(x^2 y\right) = 2xy$. Here we are treating y as a coefficient.

Just as $\frac{d}{dx}\left(5^3\right) = 0$, we compute $\frac{\partial}{\partial x}\left(y^3\right) = 0$. Here we are treating y as a constant. More examples will help make this clear.

Example 2   **Finding partial derivatives**

Find $f_x(x, y)$ and $f_y(x, y)$ in each of the following.

1. $f(x, y) = x^3 y^2 + 5y^2 - x + 7$

2. $f(x, y) = \cos(xy^2) + \sin x$

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

> **SOLUTION**
>
> 1. We have $f(x, y) = x^3 y^2 + 5y^2 - x + 7$.
>   Begin with $f_x(x, y)$. Keep y fixed, treating it as a constant or coefficient, as appropriate:
> $$f_x(x, y) = 3x^2 y^2 - 1.$$

Notes:

Note how the $5y^2$ and 7 terms go to zero.

To compute $f_y(x, y)$, we hold x fixed:

$$f_y(x, y) = 2x^3y + 10y.$$

Note how the $-x$ and 7 terms go to zero.

2. We have $f(x, y) = \cos(xy^2) + \sin x$.

Begin with $f_x(x, y)$. We need to apply the Chain Rule with the cosine term; y^2 is the coefficient of the x-term inside the cosine function.

$$f_x(x, y) = -\sin(xy^2)(y^2) + \cos x = -y^2 \sin(xy^2) + \cos x.$$

To find $f_y(x, y)$, note that x is the coefficient of the y^2 term inside of the cosine term; also note that since x is fixed, $\sin x$ is also fixed, and we treat it as a constant.

$$f_y(x, y) = -\sin(xy^2)(2xy) = -2xy \sin(xy^2).$$

3. We have $f(x, y) = e^{x^2 y^3} \sqrt{x^2 + 1}$.

Beginning with $f_x(x, y)$, note how we need to apply the Product Rule.

$$f_x(x, y) = e^{x^2 y^3}(2xy^3)\sqrt{x^2 + 1} + e^{x^2 y^3}\frac{1}{2}(x^2 + 1)^{-1/2}(2x)$$

$$= 2xy^3 e^{x^2 y^3}\sqrt{x^2 + 1} + \frac{xe^{x^2 y^3}}{\sqrt{x^2 + 1}}.$$

Note that when finding $f_y(x, y)$ we do not have to apply the Product Rule; since $\sqrt{x^2 + 1}$ does not contain y, we treat it as fixed and hence becomes a coefficient of the $e^{x^2 y^3}$ term.

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

We have shown *how* to compute a partial derivative, but it may still not be clear what a partial derivative *means*. Given $z = f(x, y)$, $f_x(x, y)$ measures the rate at which z changes as only x varies: y is held constant.

Imagine standing in a rolling meadow, then beginning to walk due east. Depending on your location, you might walk up, sharply down, or perhaps not change elevation at all. This is similar to measuring z_x: you are moving only east (in the "x"-direction) and not north/south at all. Going back to your original location, imagine now walking due north (in the "y"-direction). Perhaps walking due north does not change your elevation at all. This is analogous to $z_y = 0$: z does not change with respect to y. We can see that z_x and z_y do not have to be the same, or even similar, as it is easy to imagine circumstances where walking east means you walk downhill, though walking north makes you walk uphill.

The following example helps us visualize this more.

Notes:

Example 3   **Evaluating partial derivatives**

Let $z = f(x, y) = -x^2 - \frac{1}{2}y^2 + xy + 10$. Find $f_x(2, 1)$ and $f_y(2, 1)$ and interpret their meaning.

SOLUTION   We begin by computing $f_x(x, y) = -2x + y$ and $f_y(x, y) = -y + x$. Thus

$$f_x(2, 1) = -3 \quad \text{and} \quad f_y(2, 1) = 1.$$

It is also useful to note that $f(2, 1) = 7.5$. What does each of these numbers mean?

Consider $f_x(2, 1) = -3$, along with Figure 13.12(a). If one "stands" on the surface at the point $(2, 1, 7.5)$ and moves parallel to the x-axis (i.e., only the x-value changes, not the y-value), then the instantaneous rate of change is -3. Increasing the x-value will decrease the z-value; decreasing the x-value will increase the z-value.

Now consider $f_y(2, 1) = 1$, illustrated in Figure 13.12(b). Moving along the curve drawn on the surface, i.e., parallel to the y-axis and not changing the x-values, increases the z-value instantaneously at a rate of 1. Increasing the y-value by 1 would increase the z-value by approximately 1.

Since the magnitude of f_x is greater than the magnitude of f_y at $(2, 1)$, it is "steeper" in the x-direction than in the y-direction.
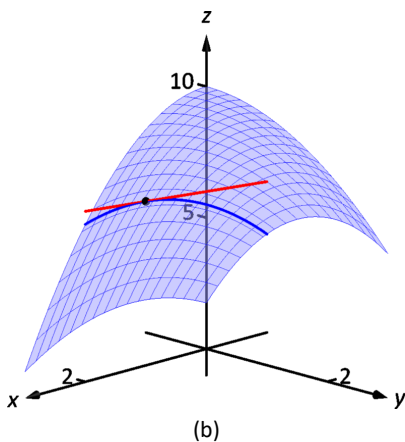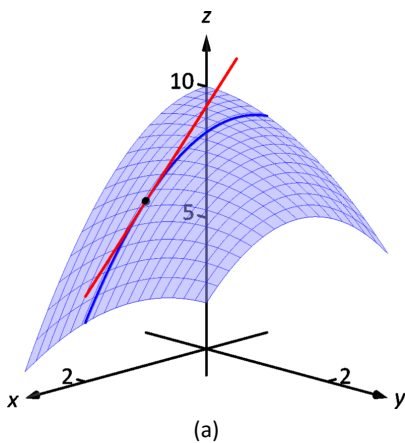

(a)

(b)

Figure 13.12: Illustrating the meaning of partial derivatives.

Second Partial Derivatives

Let $z = f(x, y)$. We have learned to find the partial derivatives $f_x(x, y)$ and $f_y(x, y)$, which are each functions of x and y. Therefore we can take partial derivatives of these, each with respect to x and y. We define these "second partials" along with the notation, give examples, then discuss their meaning.

Notes:

Definition 87   Second Partial Derivative, Mixed Partial Derivative

Let $z = f(x, y)$ be continuous on an open set S.

1. The **second partial derivative of f with respect to x then x** is

$$\frac{\partial}{\partial x}\left(\frac{\partial f}{\partial x}\right) = \frac{\partial^2 f}{\partial x^2} = \left(f_x\right)_x = f_{xx}$$

2. The **second partial derivative of f with respect to x then y** is

$$\frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}\right) = \frac{\partial^2 f}{\partial y \partial x} = \left(f_x\right)_y = f_{xy}$$

Similar definitions hold for $\dfrac{\partial^2 f}{\partial y^2} = f_{yy}$ and $\dfrac{\partial^2 f}{\partial x \partial y} = f_{yx}$.

The second partial derivatives f_{xy} and f_{yx} are **mixed partial derivatives**.

The notation of second partial derivatives gives some insight into the notation of the second derivative of a function of a single variable. If $y = f(x)$, then $f''(x) = \dfrac{d^2 y}{dx^2}$. The "$d^2 y$" portion means "take the derivative of y twice," while "dx^2" means "with respect to x both times." When we only know of functions of a single variable, this latter phrase seems silly: there is only one variable to take the derivative with respect to. Now that we understand functions of multiple variables, we see the importance of specifying which variables we are referring to.

Note: The terms in Definition 87 all depend on limits, so each definition comes with the caveat "where the limit exists."

The way to keep track of the order is to start with the variable nearest to the function. Unfortunately, this means that while $\dfrac{\partial^2 f}{\partial y \partial x} = f_{xy}$ both mean "differentiate with respect to x first", the x and y appear in a different order. Fortunately Theorem 106 will soon tell us that the order doesn't usually matter.

Example 4   Second partial derivatives

For each of the following, find all six first and second partial derivatives. That is, find

$$f_x, \quad f_y, \quad f_{xx}, \quad f_{yy}, \quad f_{xy} \quad \text{and} \quad f_{yx}.$$

1. $f(x, y) = x^3 y^2 + 2xy^3 + \cos x$

2. $f(x, y) = \dfrac{x^3}{y^2}$

SOLUTION   In each, we give f_x and f_y immediately and then spend time deriving the second partial derivatives.

Notes:

1.

$$f(x, y) = x^3 y^2 + 2xy^3 + \cos x$$
$$f_x(x, y) = 3x^2 y^2 + 2y^3 - \sin x$$
$$f_y(x, y) = 2x^3 y + 6xy^2$$
$$f_{xx}(x, y) = \frac{\partial}{\partial x}(f_x) = \frac{\partial}{\partial x}(3x^2 y^2 + 2y^3 - \sin x) = 6xy^2 - \cos x$$
$$f_{yy}(x, y) = \frac{\partial}{\partial y}(f_y) = \frac{\partial}{\partial y}(2x^3 y + 6xy^2) = 2x^3 + 12xy$$
$$f_{xy}(x, y) = \frac{\partial}{\partial y}(f_x) = \frac{\partial}{\partial y}(3x^2 y^2 + 2y^3 - \sin x) = 6x^2 y + 6y^2$$
$$f_{yx}(x, y) = \frac{\partial}{\partial x}(f_y) = \frac{\partial}{\partial x}(2x^3 y + 6xy^2) = 6x^2 y + 6y^2$$

2.

$$f(x, y) = \frac{x^3}{y^2} = x^3 y^{-2}$$
$$f_x(x, y) = \frac{3x^2}{y^2}$$
$$f_y(x, y) = -\frac{2x^3}{y^3}$$
$$f_{xx}(x, y) = \frac{\partial}{\partial x}(f_x) = \frac{\partial}{\partial x}\left(\frac{3x^2}{y^2}\right) = \frac{6x}{y^2}$$
$$f_{yy}(x, y) = \frac{\partial}{\partial y}(f_y) = \frac{\partial}{\partial y}\left(-\frac{2x^3}{y^3}\right) = \frac{6x^3}{y^4}$$
$$f_{xy}(x, y) = \frac{\partial}{\partial y}(f_x) = \frac{\partial}{\partial y}\left(\frac{3x^2}{y^2}\right) = -\frac{6x^2}{y^3}$$
$$f_{yx}(x, y) = \frac{\partial}{\partial x}(f_y) = \frac{\partial}{\partial x}\left(-\frac{2x^3}{y^3}\right) = -\frac{6x^2}{y^3}.$$

Notice how in both of the functions in Example 4, $f_{xy} = f_{yx}$. Due to the complexity of the examples, this likely is not a coincidence. The following theorem states that it is not.

> **Theorem 106   Mixed Partial Derivatives**
> Let f be defined such that f_{xy} and f_{yx} are continuous on an open set S.
> Then for each point (x, y) in S, $f_{xy}(x, y) = f_{yx}(x, y)$.

Finding f_{xy} and f_{yx} independently and comparing the results provides a convenient way of checking our work.

Notes:

Understanding Second Partial Derivatives

Now that we know *how* to find second partials, we investigate *what* they tell us.

Again we refer back to a function $y = f(x)$ of a single variable. The second derivative of f is "the derivative of the derivative," or "the rate of change of the rate of change." The second derivative measures how much the derivative is changing. If $f''(x) < 0$, then the derivative is getting smaller (so the graph of f is concave down); if $f''(x) > 0$, then the derivative is growing, making the graph of f concave up.

Now consider $z = f(x, y)$. Similar statements can be made about f_{xx} and f_{yy} as could be made about $f''(x)$ above. When taking derivatives with respect to x twice, we measure how much f_x changes with respect to x. If $f_{xx}(x, y) < 0$, it means that as x increases, f_x decreases, and the graph of f will be concave down *in the x-direction*. Using the analogy of standing in the rolling meadow used earlier in this section, f_{xx} measures whether one's path is concave up/down when walking due east.

Similarly, f_{yy} measures the concavity in the y-direction. If $f_{yy}(x, y) > 0$, then f_y is increasing with respect to y and the graph of f will be concave up in the y-direction. Appealing to the rolling meadow analogy again, f_{yy} measures whether one's path is concave up/down when walking due north.

We now consider the mixed partials f_{xy} and f_{yx}. The mixed partial f_{xy} measures how much f_x changes with respect to y. Once again using the rolling meadow analogy, f_x measures the slope if one walks due east. Looking east, begin walking *north* (side–stepping). Is the path towards the east getting steeper? If so, $f_{xy} > 0$. Is the path towards the east not changing in steepness? If so, then $f_{xy} = 0$. A similar thing can be said about f_{yx}: consider the steepness of paths heading north while side–stepping to the east.

The following example examines these ideas with concrete numbers and graphs.

Example 5   **Understanding second partial derivatives**
Let $z = x^2 - y^2 + xy$. Evaluate the 6 first and second partial derivatives at $(-1/2, 1/2)$ and interpret what each of these numbers mean.

SOLUTION   We find that:
$f_x(x, y) = 2x + y,$ $f_y(x, y) = -2y + x,$ $f_{xx}(x, y) = 2,$ $f_{yy}(x, y) = -2$ and $f_{xy}(x, y) = f_{yx}(x, y) = 1$. Thus at $(-1/2, 1/2)$ we have

$$f_x(-1/2, 1/2) = -1/2, \qquad f_y(-1/2, 1/2) = -3/2.$$

The slope of the tangent line at $(-1/2, 1/2, -1/4)$ in the direction of x is $-1/2$: if one moves from that point parallel to the x-axis, the instantaneous rate of change will be $-1/2$. The slope of the tangent line at this point in the direction
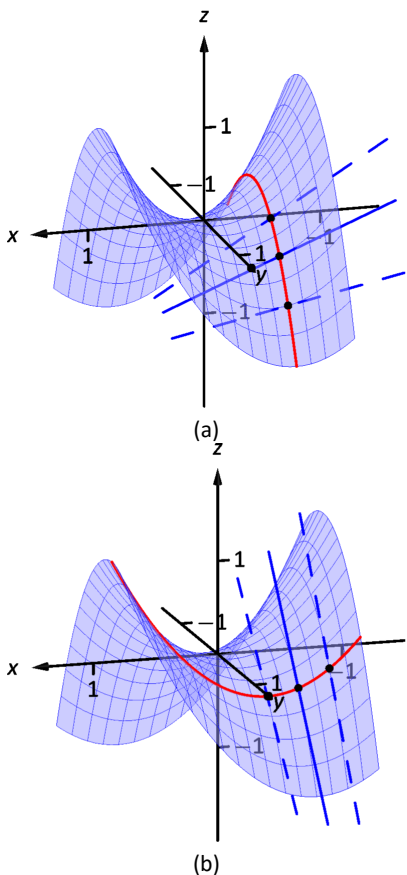

Notes:

of y is $-3/2$: if one moves from this point parallel to the y-axis, the instantaneous rate of change will be $-3/2$. These tangents lines are graphed in Figure 13.13(a) and (b), respectively, where the tangent lines are drawn in a solid line.

Now consider only Figure 13.13(a). Three directed tangent lines are drawn (two are dashed), each in the direction of x; that is, each has a slope determined by f_x. Note how as y increases, the slope of these lines get closer to 0. Since the slopes are all negative, getting closer to 0 means the *slopes are increasing.* The slopes given by f_x are increasing as y increases, meaning f_{xy} must be positive.

Since $f_{xy} = f_{yx}$, we also expect f_y to increase as x increases. Consider Figure 13.13(b) where again three directed tangent lines are drawn, this time each in the direction of y with slopes determined by f_y. As x increases, the slopes become less steep (closer to 0). Since these are negative slopes, this means the slopes are increasing.

Thus far we have a visual understanding of f_x, f_y, and $f_{xy} = f_{yx}$. We now interpret f_{xx} and f_{yy}. In Figure 13.13(a), we see a curve drawn where x is held constant at $x = -1/2$: only y varies. This curve is clearly concave down, corresponding to the fact that $f_{yy} < 0$. In part (b) of the figure, we see a similar curve where y is constant and only x varies. This curve is concave up, corresponding to the fact that $f_{xx} > 0$.

Partial Derivatives and Functions of Three Variables

The concepts underlying partial derivatives can be easily extend to more than two variables. We give some definitions and examples in the case of three variables and trust the reader can extend these definitions to more variables if needed.

(a)

(b)

Figure 13.13: Understanding the second partial derivatives in Example 5.

Definition 88   Partial Derivatives with Three Variables

Let $w = f(x, y, z)$ be a continuous function on an open set S in $\mathbb{R}^3$. The **partial derivative of f with respect to x** is:

$$f_x(x, y, z) = \lim_{h \to 0} \frac{f(x + h, y, z) - f(x, y, z)}{h}.$$

Similar definitions hold for $f_y(x, y, z)$ and $f_z(x, y, z)$.

By taking partial derivatives of partial derivatives, we can find second partial derivatives of f with respect to z then y, for instance, just as before.

Notes:

Example 6   **Partial derivatives of functions of three variables**

For each of the following, find f_x, f_y, f_z, f_{xz}, f_{yz}, and f_{zz}.

1. $f(x, y, z) = x^2 y^3 z^4 + x^2 y^2 + x^3 z^3 + y^4 z^4$

2. $f(x, y, z) = x \sin(yz)$

SOLUTION

1.

$$f_x = 2xy^3 z^4 + 2xy^2 + 3x^2 z^3 \qquad f_y = 3x^2 y^2 z^4 + 2x^2 y + 4y^3 z^4$$
$$f_z = 4x^2 y^3 z^3 + 3x^3 z^2 + 4y^4 z^3 \qquad f_{xz} = 8xy^3 z^3 + 9x^2 z^2$$
$$f_{yz} = 12x^2 y^2 z^3 + 16y^3 z^3 \qquad\quad f_{zz} = 12x^2 y^3 z^2 + 6x^3 z + 12y^4 z^2$$

2.

$$f_x = \sin(yz) \qquad f_y = xz \cos(yz) \qquad\qquad\quad f_z = xy \cos(yz)$$
$$f_{xz} = y \cos(yz) \quad f_{yz} = x \cos(yz) - xyz \sin(yz) \quad f_{zz} = -xy^2 \sin(xy)$$

Higher Order Partial Derivatives

We can continue taking partial derivatives of partial derivatives of partial derivatives of ...; we do not have to stop with second partial derivatives. These higher order partial derivatives do not have a tidy graphical interpretation; nevertheless they are not hard to compute and worthy of some practice.

We do not formally define each higher order derivative, but rather give just a few examples of the notation.

$$f_{xyx}(x, y) = \frac{\partial}{\partial x}\left(\frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}\right)\right) \quad \text{and}$$

$$f_{xyz}(x, y, z) = \frac{\partial}{\partial z}\left(\frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}\right)\right).$$

Example 7   **Higher order partial derivatives**

1. Let $f(x, y) = x^2 y^2 + \sin(xy)$. Find f_{xxy} and f_{yxx}.

2. Let $f(x, y, z) = x^3 e^{xy} + \cos(z)$. Find f_{xyz}.

Notes:

SOLUTION

1. To find f_{xxy}, we first find f_x, then f_{xx}, then f_{xxy}:

$$f_x = 2xy^2 + y\cos(xy) \qquad f_{xx} = 2y^2 - y^2\sin(xy)$$
$$f_{xxy} = 4y - 2y\sin(xy) - xy^2\cos(xy).$$

To find f_{yxx}, we first find f_y, then f_{yx}, then f_{yxx}:

$$f_y = 2x^2y + x\cos(xy) \qquad f_{yx} = 4xy + \cos(xy) - xy\sin(xy)$$
$$f_{yxx} = 4y - y\sin(xy) - \big(y\sin(xy) + xy^2\cos(xy)\big)$$
$$= 4y - 2y\sin(xy) - xy^2\cos(xy).$$

Note how $f_{xxy} = f_{yxx}$.

2. To find f_{xyz}, we find f_x, then f_{xy}, then f_{xyz}:

$$f_x = 3x^2e^{xy} + x^3ye^{xy} \qquad f_{xy} = 3x^3e^{xy} + x^3e^{xy} + x^4ye^{xy} = 4x^3e^{xy} + x^4ye^{xy}$$
$$f_{xyz} = 0.$$

In the previous example we saw that $f_{xxy} = f_{yxx}$; this is not a coincidence. While we do not state this as a formal theorem, as long as each partial derivative is continuous, it does not matter the order in which the partial derivatives are taken. For instance, $f_{xxy} = f_{xyx} = f_{yxx}$.

This can be useful at times. Had we known this, the second part of Example 7 would have been much simpler to compute. Instead of computing f_{xyz} in the x, y then z orders, we could have applied the z, then x then y order (as $f_{xyz} = f_{zxy}$). It is easy to see that $f_z = -\sin z$; then f_{zx} and f_{zxy} are clearly 0 as f_z does not contain an x or y.

We have seen that partial derivatives measure the instantaneous rate of change of a multivariable function with respect to one variable. With $z = f(x, y)$, the partial derivatives f_x and f_y measure the instantaneous rate of change of z when moving parallel to the x- and y-axes, respectively. How do we measure the rate of change at a point when we do not move parallel to one of these axes? What if we move in the direction given by the vector $\langle 2, 1 \rangle$? Can we measure that rate of change? The answer is, of course, yes, we can. This is the topic of Section 13.6. First, we need to define what it means for a function of two variables to be *differentiable.*

Notes:

Exercises 13.3

Terms and Concepts

1. What is the difference between a constant and a coefficient?

2. Given a function $z = f(x, y)$, explain in your own words how to compute f_x.

3. In the expression f_{xy}, which is computed first, f_x or f_y?

4. In the expression $\dfrac{\partial^2 f}{\partial x \partial y}$, which is computed first, f_x or f_y?

Problems

In Exercises 5–8, evaluate $f_x(x, y)$ and $f_y(x, y)$ at the indicated point.

5. $f(x, y) = x^2 y - x + 2y + 3$ at $(1, 2)$

6. $f(x, y) = x^3 - 3x + y^2 - 6y$ at $(-1, 3)$

7. $f(x, y) = \sin y \cos x$ at $(\pi/3, \pi/3)$

8. $f(x, y) = \ln(xy)$ at $(-2, -3)$

In Exercises 9–26, find f_x, f_y, f_{xx}, f_{yy}, f_{xy} and f_{yx}.

9. $f(x, y) = x^2 y + 3x^2 + 4y - 5$

10. $f(x, y) = y^3 + 3xy^2 + 3x^2 y + x^3$

11. $f(x, y) = \dfrac{x}{y}$

12. $f(x, y) = \dfrac{4}{xy}$

13. $f(x, y) = e^{x^2 + y^2}$

14. $f(x, y) = e^{x + 2y}$

15. $f(x, y) = \sin x \cos y$

16. $f(x, y) = (x + y)^3$

17. $f(x, y) = \cos(5xy^3)$

18. $f(x, y) = \sin(5x^2 + 2y^3)$

19. $f(x, y) = \sqrt{4xy^2 + 1}$

20. $f(x, y) = (2x + 5y)\sqrt{y}$

21. $f(x, y) = \dfrac{1}{x^2 + y^2 + 1}$

22. $f(x, y) = 5x - 17y$

23. $f(x, y) = 3x^2 + 1$

24. $f(x, y) = \ln(x^2 + y)$

25. $f(x, y) = \dfrac{\ln x}{4y}$

26. $f(x, y) = 5e^x \sin y + 9$

In Exercises 27–30, form a function $z = f(x, y)$ such that f_x and f_y match those given.

27. $f_x = \sin y + 1$, $\quad f_y = x \cos y$

28. $f_x = x + y$, $\quad f_y = x + y$

29. $f_x = 6xy - 4y^2$, $\quad f_y = 3x^2 - 8xy + 2$

30. $f_x = \dfrac{2x}{x^2 + y^2}$, $\quad f_y = \dfrac{2y}{x^2 + y^2}$

In Exercises 31–34, find f_x, f_y, f_z, f_{yz} and f_{zy}.

31. $f(x, y, z) = x^2 e^{2y - 3z}$

32. $f(x, y, z) = x^3 y^2 + x^3 z + y^2 z$

33. $f(x, y, z) = \dfrac{3x}{7y^2 z}$

34. $f(x, y, z) = \ln(xyz)$

13.4   Differentiability and the Total Differential

We studied **differentials** in Section 4.3, where Definition 20 states that if $y = f(x)$ and f is differentiable, then $dy = f'(x)dx$. One important use of this differential is in Integration by Substitution. Another important application is approximation. Let $\Delta x = dx$ represent a change in x. When dx is small, $dy \approx \Delta y$, the change in y resulting from the change in x. Fundamental in this understanding is this: as dx gets small, the difference between Δy and dy goes to 0. Another way of stating this: as dx goes to 0, the *error* in approximating Δy with dy goes to 0.

We extend this idea to functions of two variables. Let $z = f(x, y)$, and let $\Delta x = dx$ and $\Delta y = dy$ represent changes in x and y, respectively. Let $\Delta z = f(x + dx, y + dy) - f(x, y)$ be the change in z over the change in x and y. Recalling that f_x and f_y give the instantaneous rates of z-change in the x- and y-directions, respectively, we can approximate Δz with $dz = f_x dx + f_y dy$; in words, the total change in z is approximately the change caused by changing x plus the change caused by changing y. In a moment we give an indication of whether or not this approximation is any good. First we give a name to dz.

Definition 89   Total Differential

Let $z = f(x, y)$ be continuous on an open set S. Let dx and dy represent changes in x and y, respectively. Where the partial derivatives f_x and f_y exist, the **total differential of z** is

$$dz = f_x(x, y)dx + f_y(x, y)dy.$$

Watch the video:
Differentials of Functions of Two Variables at
`https://youtu.be/C1Xcj5Xmngc`

Example 1   Finding the total differential
Let $z = x^4 e^{3y}$. Find dz.

SOLUTION   We compute the partial derivatives: $f_x = 4x^3 e^{3y}$ and $f_y = 3x^4 e^{3y}$. Following Definition 89, we have

$$dz = 4x^3 e^{3y} dx + 3x^4 e^{3y} dy.$$

Notes:

We *can* approximate Δz with dz, but as with all approximations, there is error involved. A good approximation is one in which the error is small. At a given point (x_0, y_0), let E_x and E_y be functions of dx and dy such that $E_x dx + E_y dy$ describes this error. Then

$$\Delta z = dz + E_x dx + E_y dy$$
$$= f_x(x_0, y_0)dx + f_y(x_0, y_0)dy + E_x dx + E_y dy.$$

If the approximation of Δz by dz is good, then as dx and dy get small, so does $E_x dx + E_y dy$. The approximation of Δz by dz is even better if, as dx and dy go to 0, so do E_x and E_y. This leads us to our definition of differentiability.

Definition 90   Multivariable Differentiability

Let $z = f(x, y)$ be defined on an open set S containing (x_0, y_0) where $f_x(x_0, y_0)$ and $f_y(x_0, y_0)$ exist. Let dz be the total differential of z at (x_0, y_0), let $\Delta z = f(x_0 + dx, y_0 + dy) - f(x_0, y_0)$, and let E_x and E_y be functions of dx and dy such that

$$\Delta z = dz + E_x dx + E_y dy.$$

1. f is **differentiable at** (x_0, y_0) if

$$\lim_{(dx, dy) \to (0,0)} \| \langle E_x, E_y \rangle \| = 0.$$

2. f is **differentiable on** S if f is differentiable at every point in S. If f is differentiable on $\mathbb{R}^2$, we say that f is **differentiable everywhere**.

Example 2   Showing a function is differentiable
Show $f(x, y) = xy + 3y^2$ is differentiable using Definition 90.

SOLUTION   We begin by finding $f(x + dx, y + dy)$, Δz, f_x and f_y.

$$f(x + dx, y + dy) = (x + dx)(y + dy) + 3(y + dy)^2$$
$$= xy + xdy + ydx + dxdy + 3y^2 + 6ydy + 3dy^2.$$

$\Delta z = f(x + dx, y + dy) - f(x, y)$, so

$$\Delta z = xdy + ydx + dxdy + 6ydy + 3dy^2.$$

Notes:

It is straightforward to compute $f_x = y$ and $f_y = x + 6y$. Consider once more Δz:

$$\Delta z = xdy + ydx + dxdy + 6ydy + 3dy^2 \qquad \text{(now reorder)}$$
$$= ydx + xdy + 6ydy + dxdy + 3dy^2$$
$$= \underbrace{(y)}_{f_x}\,dx + \underbrace{(x + 6y)}_{f_y}\,dy + \underbrace{(dy)}_{E_x}\,dx + \underbrace{(3dy)}_{E_y}\,dy$$
$$= f_xdx + f_ydy + E_xdx + E_ydy.$$

With $E_x = dy$ and $E_y = 3dy$, it is clear that as dx and dy go to 0, E_x and E_y also go to 0. Since this did not depend on a specific point (x_0, y_0), we can say that $f(x, y)$ is differentiable for all pairs (x, y) in $\mathbb{R}^2$, or, equivalently, that f is differentiable everywhere.

Our intuitive understanding of differentiability of functions $y = f(x)$ of one variable was that the graph of f was "smooth." A similar intuitive understanding of functions $z = f(x, y)$ of two variables is that the surface defined by f is also "smooth," not containing cusps, edges, breaks, etc. The following theorem states that differentiable functions are continuous, followed by another theorem that provides a more tangible way of determining whether a great number of functions are differentiable or not.

Theorem 107   Continuity and Differentiability of Multivariable Functions

Let $z = f(x, y)$ be defined on an open set S containing (x_0, y_0). If f is differentiable at (x_0, y_0), then f is continuous at (x_0, y_0).

Theorem 108   Differentiability of Multivariable Functions

Let $z = f(x, y)$ be defined on an open set S containing (x_0, y_0). If f_x and f_y are both continuous on S, then f is differentiable on S.

The theorems assure us that essentially all functions that we see in the course of our studies here are differentiable (and hence continuous) on their natural domains. There is a difference between Definition 90 and Theorem 108, though: it is possible for a function f to be differentiable yet f_x or f_y is *not* continuous. Such strange behavior of functions is a source of delight for many mathematicians. When this happens, we need to use other methods to determine whether or not f is differentiable at that point.

Notes:

Approximating with the Total Differential

By the definition, when *f* is differentiable *dz* is a good approximation for Δz when *dx* and *dy* are small. We give some simple examples of how this is used here.

Example 3   **Approximating with the total differential**
Let $f(x, y) = \sqrt{x} \sin y$. Approximate $f(4.1, 0.2)$.

    SOLUTION   We can approximate $f(4.1, 0.2)$ using $f(4, 0) = 0$. Without calculus, this is the best approximation we could reasonably come up with. The total differential gives us a way of adjusting this initial approximation to hopefully get a more accurate answer.

    We let $\Delta z = f(4.1, 0.2) - f(4, 0)$. The total differential *dz* is approximately equal to Δz, so

$$f(4.1, 0.2) - f(4, 0) \approx dz \quad \Rightarrow \quad f(4.1, 0.2) \approx dz + f(4, 0). \qquad (13.1)$$

To find *dz*, we need f_x and f_y.

$$f_x(x, y) = \frac{\sin y}{2\sqrt{x}} \quad \Rightarrow \qquad f_x(4, 0) = \frac{\sin 0}{2\sqrt{4}} = 0$$

$$f_y(x, y) = \sqrt{x} \cos y \quad \Rightarrow \qquad f_y(4, 0) = \sqrt{4} \cos 0 = 2$$

Approximating 4.1 with 4 gives $dx = 0.1$; approximating 0.2 with 0 gives $dy = 0.2$. Thus

$$dz(4, 0) = f_x(4, 0)(0.1) + f_y(4, 0)(0.2) = 0(0.1) + 2(0.2) = 0.4.$$

Returning to Equation (13.1), we have

$$f(4.1, 0.2) \approx 0.4 + 0 = .4.$$

We, of course, can compute the actual value of $f(4.1, 0.2)$ with a calculator; to 5 places after the decimal, this is 0.40228. Obviously our approximation is quite good.

    The point of the previous example was *not* to develop an approximation method for known functions. After all, we can very easily compute $f(4.1, 0.2)$ using readily available technology. Rather, it serves to illustrate how well this method of approximation works, and to reinforce the following concept:

        "New position = old position + amount of change," so
      "New position ≈ old position + approximate amount of change."

Notes:

In the previous example, we could easily compute $f(4, 0)$ and could approximate the amount of z-change when computing $f(4.1, 0.2)$, letting us approximate the new z-value.

It may be surprising to learn that it is not uncommon to know the values of f, f_x and f_y at a particular point without actually knowing the function f. The total differential gives a good method of approximating f at nearby points.

Example 4   **Approximating an unknown function**
Given that $f(2, -3) = 6$, $f_x(2, -3) = 1.3$ and $f_y(2, -3) = -0.6$, approximate $f(2.1, -3.03)$.

SOLUTION   The total differential approximates how much f changes from the point $(2, -3)$ to the point $(2.1, -3.03)$. With $dx = 0.1$ and $dy = -0.03$, we have

$$dz = f_x(2, -3)dx + f_y(2, -3)dy$$
$$= 1.3(0.1) + (-0.6)(-0.03)$$
$$= 0.148.$$

The change in z is approximately 0.148, so we approximate $f(2.1, -3.03) \approx 6.148$.

Error/Sensitivity Analysis

The total differential gives an approximation of the change in z given small changes in x and y. We can use this to approximate error propagation; that is, if the input is a little off from what it should be, how far from correct will the output be? We demonstrate this in an example.

Example 5   **Sensitivity analysis**
A cylindrical steel storage tank is to be built that is 10ft tall and 4ft across in diameter. It is known that the steel will expand/contract with temperature changes; is the overall volume of the tank more sensitive to changes in the diameter or in the height of the tank?

SOLUTION   A cylindrical solid with height h and radius r has volume $V = \pi r^2 h$. We can view V as a function of two variables, r and h. We can compute partial derivatives of V:

$$\frac{\partial V}{\partial r} = V_r(r, h) = 2\pi rh \qquad \text{and} \qquad \frac{\partial V}{\partial h} = V_h(r, h) = \pi r^2.$$

Notes:

The total differential is $dV = (2\pi rh)dr + (\pi r^2)dh$. When $h = 10$ and $r = 2$, we have $dV = 40\pi dr + 4\pi dh$. Note that the coefficient of dr is 40π; the coefficient of dh is a tenth of that. A small change in radius will be multiplied by 40π, whereas a small change in height will be multiplied by 4π. Thus the volume of the tank is more sensitive to changes in radius than in height.

The previous example showed that the volume of a particular tank was more sensitive to changes in radius than in height. Keep in mind that this analysis only applies to a tank of those dimensions. A tank with a height of 1ft and radius of 5ft would be more sensitive to changes in height than in radius.

One could make a chart of small changes in radius and height and find exact changes in volume given specific changes. While this provides exact numbers, it does not give as much insight as the error analysis using the total differential.

Differentiability of Functions of Three Variables

The definition of differentiability for functions of three variables is very similar to that of functions of two variables. We again start with the total differential.

Definition 91   Total Differential

Let $w = f(x, y, z)$ be continuous on an open set S. Let dx, dy and dz represent changes in x, y and z, respectively. Where the partial derivatives f_x, f_y and f_z exist, the **total differential of w** is

$$dz = f_x(x, y, z)dx + f_y(x, y, z)dy + f_z(x, y, z)dz.$$

This differential can be a good approximation of the change in w when $w = f(x, y, z)$ is **differentiable**.

Notes:

Definition 92   Multivariable Differentiability

Let $w = f(x, y, z)$ be defined on an open set S containing (x_0, y_0, z_0) where $f_x(x_0, y_0, z_0)$, $f_y(x_0, y_0, z_0)$ and $f_z(x_0, y_0, z_0)$ exist. Let dw be the total differential of w at (x_0, y_0, z_0), let $\Delta w = f(x_0 + dx, y_0 + dy, z_0 + dz) - f(x_0, y_0, z_0)$, and let E_x, E_y and E_z be functions of dx, dy and dz such that

$$\Delta w = dw + E_x dx + E_y dy + E_z dz.$$

1. f is **differentiable at** (x_0, y_0, z_0) if, given $\varepsilon > 0$, there is a $\delta > 0$ such that if $\|\langle dx, dy, dz \rangle\| < \delta$, then $\|\langle E_x, E_y, E_z \rangle\| < \varepsilon$.

2. f is **differentiable on** S if f is differentiable at every point in S. If f is differentiable on $\mathbb{R}^3$, we say that f is **differentiable everywhere**.

Just as before, this definition gives a rigorous statement about what it means to be differentiable that is not very intuitive. We follow it with a theorem similar to Theorem 108.

Theorem 109   Continuity and Differentiability of Functions of Three Variables

Let $w = f(x, y, z)$ be defined on an open set S containing (x_0, y_0, z_0).

1. If f is differentiable at (x_0, y_0, z_0), then f is continuous at (x_0, y_0, z_0).

2. If f_x, f_y and f_z are continuous on S, then f is differentiable on B.

This set of definition and theorem extends to functions of any number of variables. The theorem again gives us a simple way of verifying that most functions that we encounter are differentiable on their natural domains.

This section has given us a formal definition of what it means for a functions to be "differentiable," along with a theorem that gives a more accessible understanding. The following sections return to notions prompted by our study of partial derivatives that make use of the fact that most functions we encounter are differentiable.

Notes:

Exercises 13.4

Terms and Concepts

1. T/F: If $f(x, y)$ is differentiable on S, the f is continuous on S.

2. T/F: If f_x and f_y are continuous on S, then f is differentiable on S.

3. T/F: If $z = f(x, y)$ is differentiable, then the change in z over small changes dx and dy in x and y is approximately dz.

4. Finish the sentence: "The new z-value is approximately the old z-value plus the approximate _____."

Problems

In Exercises 5–8, find the total differential dz.

5. $z = x \sin y + x^2$

6. $z = (2x^2 + 3y)^2$

7. $z = 5x - 7y$

8. $z = xe^{x+y}$

In Exercises 9–12, a function $z = f(x, y)$ is given. Give the indicated approximation using the total differential.

9. $f(x, y) = \sqrt{x^2 + y}$. Approximate $f(2.95, 7.1)$ knowing $f(3, 7) = 4$.

10. $f(x, y) = \sin x \cos y$. Approximate $f(0.1, -0.1)$ knowing $f(0, 0) = 0$.

11. $f(x, y) = x^2 y - xy^2$. Approximate $f(2.04, 3.06)$ knowing $f(2, 3) = -6$.

12. $f(x, y) = \ln(x - y)$. Approximate $f(5.1, 3.98)$ knowing $f(5, 4) = 0$.

Exercises 13–16 ask a variety of questions dealing with approximating error and sensitivity analysis.

13. A cylindrical storage tank is to be 2ft tall with a radius of 1ft. Is the volume of the tank more sensitive to changes in the radius or the height?

14. **Projectile Motion:** The x-value of an object moving under the principles of projectile motion is $x(\theta, v_0, t) = (v_0 \cos \theta)t$. A particular projectile is fired with an initial velocity of $v_0 = 250$ft/s and an angle of elevation of $\theta = 60°$. It travels a distance of 375ft in 3 seconds.

    Is the projectile more sensitive to errors in initial speed or angle of elevation?

15. The length ℓ of a long wall is to be approximated. The angle θ, as shown in the diagram (not to scale), is measured to be 85°, and the distance x is measured to be 30′. Assume that the triangle formed is a right triangle.

    Is the measurement of the length of ℓ more sensitive to errors in the measurement of x or in θ?
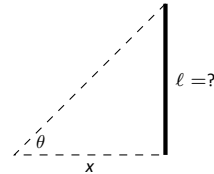
16. It is "common sense" that it is far better to measure a long distance with a long measuring tape rather than a short one. A measured distance D can be viewed as the product of the length ℓ of a measuring tape times the number n of times it was used. For instance, using a 3′ tape 10 times gives a length of 30′. To measure the same distance with a 12′ tape, we would use the tape 2.5 times. (I.e., $30 = 12 \times 2.5$.) Thus $D = n\ell$.

    Suppose each time a measurement is taken with the tape, the recorded distance is within 1/16″ of the actual distance. (I.e., $d\ell = 1/16″ \approx 0.005$ft). Using differentials, show why common sense proves correct in that it is better to use a long tape to measure long distances.

In Exercises 17–18, find the total differential dw.

17. $w = x^2 y z^3$

18. $w = e^x \sin y \ln z$

In Exercises 19–22, use the information provided and the total differential to make the given approximation.

19. $f(3, 1) = 7$, $f_x(3, 1) = 9$, $f_y(3, 1) = -2$. Approximate $f(3.05, 0.9)$.

20. $f(-4, 2) = 13$, $f_x(-4, 2) = 2.6$, $f_y(-4, 2) = 5.1$. Approximate $f(-4.12, 2.07)$.

21. $f(2, 4, 5) = -1$, $f_x(2, 4, 5) = 2$, $f_y(2, 4, 5) = -3$, $f_z(2, 4, 5) = 3.7$. Approximate $f(2.5, 4.1, 4.8)$.

22. $f(3, 3, 3) = 5$, $f_x(3, 3, 3) = 2$, $f_y(3, 3, 3) = 0$, $f_z(3, 3, 3) = -2$. Approximate $f(3.1, 3.1, 3.1)$.

23. Find where the function $z = \sqrt{x^2 + y^2}$ is differentiable.

13.5   The Multivariable Chain Rule

The Chain Rule, as learned in Section 2.5, states that $\dfrac{d}{dx}\Big(f\big(g(x)\big)\Big) = f'\big(g(x)\big)g'(x)$.
If $t = g(x)$, we can express the Chain Rule as

$$\frac{df}{dx} = \frac{df}{dt}\frac{dt}{dx}.$$

In this section we extend the Chain Rule to functions of more than one variable.

Theorem 110   Multivariable Chain Rule, Part I

Let $z = f(x,y)$, $x = g(t)$ and $y = h(t)$, where f, g and h are differentiable functions. Then $z = f(x,y) = f\big(g(t), h(t)\big)$ is a function of t, and

$$\frac{dz}{dt} = \frac{df}{dt} = f_x(x,y)\frac{dx}{dt} + f_y(x,y)\frac{dy}{dt} = \frac{\partial f}{\partial x}\frac{dx}{dt} + \frac{\partial f}{\partial y}\frac{dy}{dt}.$$

Proof

By definition,

$$\frac{df}{dt}(x,y) = \lim_{h\to 0}\frac{f(x(t+h), y(t+h)) - f(x,y)}{h}.$$

Let

$$\begin{aligned}
\Delta f &= f(x(t+h), y(t+h)) - f(x,y),\\
dx &= x(t+h) - x(t), \qquad \text{and}\\
dy &= y(t+h) - y(t).
\end{aligned}$$

Because f is differentiable, Definition 90 gives us functions E_x and E_y so that

$$E_x dx + E_y dy = \Delta f - f_x(x,y)dx - f_y(x,y)dy$$
$$\text{and} \qquad \lim_{\langle dx,dy\rangle \to 0}\langle E_x, E_y\rangle = 0.$$

This means that

$$\begin{aligned}
\frac{df}{dt}(x,y) &= \lim_{h\to 0}\frac{f_x(x,y)dx + f_y(x,y)dy + E_x dx + E_y dy}{h}\\
&= f_x(x,y)\lim_{h\to 0}\frac{dx}{h} + f_y(x,y)\lim_{h\to 0}\frac{dy}{h}\\
&\quad + \lim_{h\to 0}E_x\lim_{h\to 0}\frac{dx}{h} + \lim_{h\to 0}E_y\lim_{h\to 0}\frac{dy}{h}\\
&= f_x(x,y)x'(t) + f_y(x,y)y'(t) + 0x'(t) + 0y'(t). \qquad \square
\end{aligned}$$

Notes:

It is good to understand what the situation of $z = f(x, y)$, $x = g(t)$ and $y = h(t)$ describes. We know that $z = f(x, y)$ describes a surface; we also recognize that $x = g(t)$ and $y = h(t)$ are parametric equations for a curve in the x-y plane. Combining these together, we are describing a curve that lies on the surface described by f. The parametric equations for this curve are $x = g(t)$, $y = h(t)$ and $z = f(g(t), h(t))$.

Consider Figure 13.14 in which a surface is drawn, along with a dashed curve in the x-y plane. Restricting f to just the points on this circle gives the curve shown on the surface. The derivative $\frac{df}{dt}$ gives the instantaneous rate of change of f with respect to t. If we consider an object traveling along this path, $\frac{df}{dt}$ gives the rate at which the object rises/falls.

We now practice applying the Multivariable Chain Rule.

Example 1   **Using the Multivariable Chain Rule**

Let $z = x^2 y + x$, where $x = \sin t$ and $y = e^{5t}$. Find $\dfrac{dz}{dt}$ using the Chain Rule.

     SOLUTION   Following Theorem 110, we find

$$f_x(x, y) = 2xy + 1, \qquad f_y(x, y) = x^2, \qquad \frac{dx}{dt} = \cos t, \qquad \frac{dy}{dt} = 5e^{5t}.$$

Applying the theorem, we have

$$\frac{dz}{dt} = (2xy + 1)\cos t + 5x^2 e^{5t}.$$

This may look odd, as it seems that $\frac{dz}{dt}$ is a function of x, y and t. Since x and y are functions of t, $\frac{dz}{dt}$ is really just a function of t, and we can replace x with $\sin t$ and y with e^{5t}:

$$\frac{dz}{dt} = (2xy + 1)\cos t + 5x^2 e^{5t} = (2\sin(t)e^{5t} + 1)\cos t + 5e^{5t}\sin^2 t.$$

The previous example can make us wonder: if we substituted for x and y at the end to show that $\frac{dz}{dt}$ is really just a function of t, why not substitute before differentiating, showing clearly that z is a function of t?

That is, $z = x^2 y + x = (\sin t)^2 e^{5t} + \sin t$. Applying the Chain and Product Rules, we have

$$\frac{dz}{dt} = 2\sin t \cos t\, e^{5t} + 5\sin^2 t\, e^{5t} + \cos t,$$
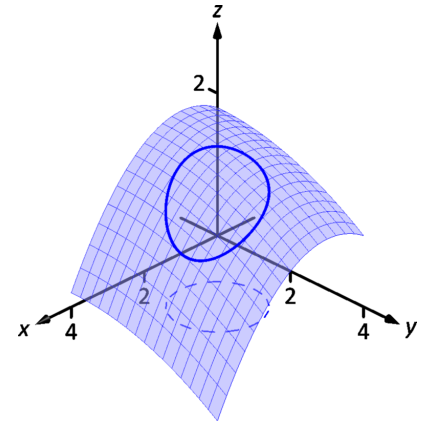
which matches the result from the example.

Figure 13.14: Understanding the application of the Multivariable Chain Rule.

Notes:

This may now make one wonder "What's the point? If we could already find the derivative, why learn another way of finding it?" In some cases, applying this rule makes deriving simpler, but this is hardly the power of the Chain Rule. Rather, in the case where $z = f(x, y)$, $x = g(t)$ and $y = h(t)$, the Chain Rule is extremely powerful when *we do not know what f, g and/or h are*. It may be hard to believe, but often in "the real world" we know rate–of–change information (i.e., information about derivatives) without explicitly knowing the underlying functions. The Chain Rule allows us to combine several rates of change to find another rate of change. The Chain Rule also has theoretic use, giving us insight into the behavior of certain constructions (as we'll see in the next section).

We demonstrate this in the next example.

Example 2   Applying the Multivarible Chain Rule
An object travels along a path on a surface. The exact path and surface are not known, but at time $t = t_0$ it is known that :

$$\frac{\partial z}{\partial x} = 5, \qquad \frac{\partial z}{\partial y} = -2, \qquad \frac{dx}{dt} = 3 \qquad \text{and} \qquad \frac{dy}{dt} = 7.$$

Find $\frac{dz}{dt}$ at time t_0.

SOLUTION   The Multivariable Chain Rule states that

$$\begin{aligned}
\frac{dz}{dt} &= \frac{\partial z}{\partial x}\frac{dx}{dt} + \frac{\partial z}{\partial y}\frac{dy}{dt} \\
&= 5(3) + (-2)(7) \\
&= 1.
\end{aligned}$$

By knowing certain rates–of–change information about the surface and about the path of the particle in the *x-y* plane, we can determine how quickly the object is rising/falling.

We next apply the Chain Rule to solve a max/min problem.

Example 3   Applying the Multivariable Chain Rule
Consider the surface $z = x^2 + y^2 - xy$, a paraboloid, on which a particle moves with x and y coordinates given by $x = \cos t$ and $y = \sin t$. Find $\frac{dz}{dt}$ when $t = 0$, and find where the particle reaches its maximum/minimum z-values.

SOLUTION   It is straightforward to compute

$$f_x(x, y) = 2x - y, \qquad f_y(x, y) = 2y - x, \qquad \frac{dx}{dt} = -\sin t, \qquad \frac{dy}{dt} = \cos t.$$

Notes:

Combining these according to the Chain Rule gives:

$$\frac{dz}{dt} = -(2x - y)\sin t + (2y - x)\cos t.$$

When $t = 0$, $x = 1$ and $y = 0$. Thus $\frac{dz}{dt} = -(2)(0) + (-1)(1) = -1$. When $t = 0$, the particle is moving down, as shown in Figure 13.15.

To find where z-value is maximized/minimized on the particle's path, we set $\frac{dz}{dt} = 0$ and solve for t:

$$\frac{dz}{dt} = 0 = -(2x - y)\sin t + (2y - x)\cos t$$
$$0 = -(2\cos t - \sin t)\sin t + (2\sin t - \cos t)\cos t$$
$$0 = \sin^2 t - \cos^2 t$$
$$\cos^2 t = \sin^2 t$$
$$t = n\frac{\pi}{4} \quad \text{(for odd } n\text{)}$$

Figure 13.15: Plotting the path of a particle on a surface in Example 3.

We can use the First Derivative Test to find that on $[0, 2\pi]$, z has reaches its absolute minimum at $t = \pi/4$ and $5\pi/4$; it reaches its absolute maximum at $t = 3\pi/4$ and $7\pi/4$, as shown in Figure 13.15.

We can extend the Chain Rule to include the situation where z is a function of more than one variable, and each of these variables is also a function of more than one variable. The basic case of this is where $z = f(x, y)$, and x and y are functions of two variables, say s and t.

Notes:

> **Theorem 111   Multivariable Chain Rule, Part II**
>
> 1. Let $z = f(x, y)$, $x = g(s, t)$ and $y = h(s, t)$, where f, g and h are differentiable functions. Then z is a function of s and t, and
>
> - $\dfrac{\partial z}{\partial s} = \dfrac{\partial f}{\partial x} \dfrac{\partial x}{\partial s} + \dfrac{\partial f}{\partial y} \dfrac{\partial y}{\partial s}$,   and
>
> - $\dfrac{\partial z}{\partial t} = \dfrac{\partial f}{\partial x} \dfrac{\partial x}{\partial t} + \dfrac{\partial f}{\partial y} \dfrac{\partial y}{\partial t}$.
>
> 2. Let $z = f(x_1, x_2, \ldots, x_m)$ be a differentiable function of m variables, where each of the x_i is a differentiable function of the variables $t_1, t_2, \ldots, t_n$. Then z is a function of the t_i, and
>
> $$\frac{\partial z}{\partial t_i} = \frac{\partial f}{\partial x_1} \frac{\partial x_1}{\partial t_i} + \frac{\partial f}{\partial x_2} \frac{\partial x_2}{\partial t_i} + \cdots + \frac{\partial f}{\partial x_m} \frac{\partial x_m}{\partial t_i}.$$

The proof of Part II follows quickly from Part I, because $\frac{\partial}{\partial t_i}$ means that we hold the other variables constant and we are back to the one variable case already proved. A helpful way to remember the derivatives is to examine the following chart
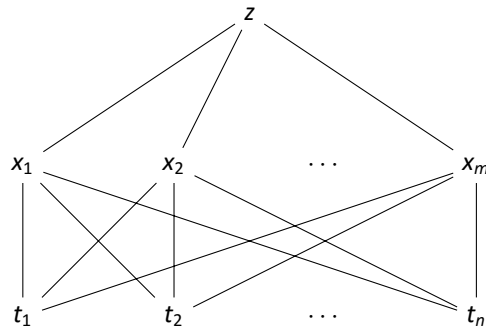
Each possible path from f to the variable t_i contributes a term to the sum, and each line segment in a path contributes a factor to that term.

Notes:

Watch the video:
Generalized Chain Rule — Part 1 at
`https://youtu.be/HOYAO-pOHsg`

Example 4   **Using the Multivarible Chain Rule, Part II**
Let $z = x^2y + x$, $x = s^2 + 3t$ and $y = 2s - t$. Find $\frac{\partial z}{\partial s}$ and $\frac{\partial z}{\partial t}$, and evaluate each when $s = 1$ and $t = 2$.

SOLUTION   Following Theorem 111, we compute the following partial derivatives:

$$\frac{\partial f}{\partial x} = 2xy + 1 \qquad \frac{\partial f}{\partial y} = x^2,$$

$$\frac{\partial x}{\partial s} = 2s \qquad \frac{\partial x}{\partial t} = 3 \qquad \frac{\partial y}{\partial s} = 2 \qquad \frac{\partial y}{\partial t} = -1.$$

Thus

$$\frac{\partial z}{\partial s} = (2xy + 1)(2s) + (x^2)(2) = 4xys + 2s + 2x^2, \quad \text{and}$$

$$\frac{\partial z}{\partial t} = (2xy + 1)(3) + (x^2)(-1) = 6xy - x^2 + 3.$$

When $s = 1$ and $t = 2$, $x = 7$ and $y = 0$, so

$$\frac{\partial z}{\partial s} = 100 \qquad \text{and} \qquad \frac{\partial z}{\partial t} = -46.$$

Example 5   **Using the Multivarible Chain Rule, Part II**
Let $w = xy + z^2$, where $x = t^2e^s$, $y = t\cos s$, and $z = s\sin t$. Find $\frac{\partial w}{\partial t}$ when $s = 0$ and $t = \pi$.

SOLUTION   Following Theorem 111, we compute the following partial derivatives:

$$\frac{\partial f}{\partial x} = y \qquad \frac{\partial f}{\partial y} = x \qquad \frac{\partial f}{\partial z} = 2z,$$

$$\frac{\partial x}{\partial t} = 2te^s \qquad \frac{\partial y}{\partial t} = \cos s \qquad \frac{\partial z}{\partial t} = s\cos t.$$

Notes:

Thus
$$\frac{\partial w}{\partial t} = y(2te^s) + x(\cos s) + 2z(s \cos t).$$

When $s = 0$ and $t = \pi$, we have $x = \pi^2$, $y = \pi$ and $z = 0$. Thus
$$\frac{\partial w}{\partial t} = \pi(2\pi) + \pi^2 = 3\pi^2.$$

Implicit Differentiation

We studied finding $\frac{dy}{dx}$ when y is given as an implicit function of x in detail in Section 2.6. We find here that the Multivariable Chain Rule gives a simpler method of finding $\frac{dy}{dx}$.

For instance, consider the implicit function $x^2y - xy^3 = 3$. We learned to use the following steps to find $\frac{dy}{dx}$:

$$\frac{d}{dx}\left(x^2y - xy^3\right) = \frac{d}{dx}\left(3\right)$$
$$2xy + x^2\frac{dy}{dx} - y^3 - 3xy^2\frac{dy}{dx} = 0$$
$$\frac{dy}{dx} = -\frac{2xy - y^3}{x^2 - 3xy^2}. \tag{13.2}$$

Instead of using this method, consider $z = x^2y - xy^3$. The implicit function above describes the level curve $z = 3$. Considering x and y as functions of x, the Multivariable Chain Rule states that

$$\frac{dz}{dx} = \frac{\partial z}{\partial x}\frac{dx}{dx} + \frac{\partial z}{\partial y}\frac{dy}{dx}. \tag{13.3}$$

Since z is constant (in our example, $z = 3$), $\frac{dz}{dx} = 0$. We also know $\frac{dx}{dx} = 1$. Equation (13.3) becomes

$$0 = \frac{\partial z}{\partial x}(1) + \frac{\partial z}{\partial y}\frac{dy}{dx}$$

$$\frac{dy}{dx} = -\frac{\partial z}{\partial x} \bigg/ \frac{\partial z}{\partial y}$$

$$= -\frac{f_x}{f_y}.$$

Note how our solution for $\frac{dy}{dx}$ in Equation (13.2) is just the partial derivative of z with respect to x, divided by the partial derivative of z with respect to y.

We state the above as a theorem for two and three variables.

Notes:

Theorem 112   Implicit Differentiation

If f is a differentiable function of x and y, where $f(x, y) = c$ defines y as an implicit function of x for some constant c, then

$$\frac{dy}{dx} = -\frac{f_x(x, y)}{f_y(x, y)}.$$

If f is a differentiable function of x, y, and z, where $f(x, y, z) = c$ defines z as an implicit function of x and y for some constant c, then

$$\frac{\partial z}{\partial x} = -\frac{f_x(x, y, z)}{f_z(x, y, z)} \quad \text{and} \quad \frac{\partial z}{\partial y} = -\frac{f_y(x, y, z)}{f_z(x, y, z)}.$$

We practice using Theorem 112 by applying it to a problem from Section 2.6.

Example 6   Implicit Differentiation

Given the implicitly defined function $\sin(x^2 y^2) + y^3 = x + y$, find y'. Note: this is the same problem as given in Example 2.6.4 of Section 2.6.

SOLUTION   Let $f(x, y) = \sin(x^2 y^2) + y^3 - x - y$; the implicitly defined function above is equivalent to $f(x, y) = 0$. We find $\frac{dy}{dx}$ by applying Theorem 112. We find

$$f_x(x, y) = 2xy^2 \cos(x^2 y^2) - 1 \quad \text{and} \quad f_y(x, y) = 2x^2 y \cos(x^2 y^2) + 3y^2 - 1,$$

so

$$\frac{dy}{dx} = -\frac{2xy^2 \cos(x^2 y^2) + 3y^2 - 1}{2x^2 y \cos(x^2 y^2) - 1},$$

which matches our solution from Example 2.6.4.

Notes:

Exercises 13.5

Terms and Concepts

1. Let a level curve of $z = f(x, y)$ be described by $x = g(t)$, $y = h(t)$. Explain why $\frac{dz}{dt} = 0$.

2. Fill in the blank: The single variable Chain Rule states
$$\frac{d}{dx}\Big(f(g(x))\Big) = f'(g(x)) \cdot \underline{\hspace{1cm}}.$$

3. Fill in the blank: The Multivariable Chain Rule states
$$\frac{df}{dt} = \frac{\partial f}{\partial x} \cdot \underline{\hspace{1cm}} + \underline{\hspace{1cm}} \cdot \frac{dy}{dt}.$$

4. If $z = f(x, y)$, where $x = g(t)$ and $y = h(t)$, we can substitute and write z as an explicit function of t.
   T/F: Using the Multivariable Chain Rule to find $\frac{dz}{dt}$ is sometimes easier than first substituting and then taking the derivative.

5. T/F: The Multivariable Chain Rule is only useful when all the related functions are known explicitly.

6. The Multivariable Chain Rule allows us to compute implicit derivatives easily by just computing two _____ derivatives.

Problems

In Exercises 7–12, functions $z = f(x, y)$, $x = g(t)$ and $y = h(t)$ are given.

  (a) Use the Multivariable Chain Rule to compute $\frac{dz}{dt}$.

  (b) Evaluate $\frac{dz}{dt}$ at the indicated t-value.

7. $z = 3x + 4y$, $x = t^2$, $y = 2t$; $t = 1$

8. $z = x^2 - y^2$, $x = t$, $y = t^2 - 1$; $t = 1$

9. $z = 5x + 2y$, $x = 2\cos t + 1$, $y = \sin t - 3$; $t = \pi/4$

10. $z = \dfrac{x}{y^2 + 1}$, $x = \cos t$, $y = \sin t$; $t = \pi/2$

11. $z = x^2 + 2y^2$, $x = \sin t$, $y = 3\sin t$; $t = \pi/4$

12. $z = \cos x \sin y$, $x = \pi t$, $y = 2\pi t + \pi/2$; $t = 3$

In Exercises 13–18, functions $z = f(x, y)$, $x = g(t)$ and $y = h(t)$ are given. Find the values of t where $\frac{dz}{dt} = 0$. Note: these are the same surfaces/curves as found in Exercises 7 – 12.

13. $z = 3x + 4y$, $x = t^2$, $y = 2t$

14. $z = x^2 - y^2$, $x = t$, $y = t^2 - 1$

15. $z = 5x + 2y$, $x = 2\cos t + 1$, $y = \sin t - 3$

16. $z = \dfrac{x}{y^2 + 1}$, $x = \cos t$, $y = \sin t$

17. $z = x^2 + 2y^2$, $x = \sin t$, $y = 3\sin t$

18. $z = \cos x \sin y$, $x = \pi t$, $y = 2\pi t + \pi/2$

In Exercises 19–22, functions $z = f(x, y)$, $x = g(s, t)$ and $y = h(s, t)$ are given.

  (a) Use the Multivariable Chain Rule to compute $\frac{\partial z}{\partial s}$ and $\frac{\partial z}{\partial t}$.

  (b) Evaluate $\frac{\partial z}{\partial s}$ and $\frac{\partial z}{\partial t}$ at the indicated s and t values.

19. $z = x^2 y$, $x = s - t$, $y = 2s + 4t$; $s = 1, t = 0$

20. $z = \cos\left(\pi x + \frac{\pi}{2} y\right)$, $x = st^2$, $y = s^2 t$; $s = 1, t = 1$

21. $z = x^2 + y^2$, $x = s\cos t$, $y = s\sin t$; $s = 2, t = \pi/4$

22. $z = e^{-(x^2 + y^2)}$, $x = t$, $y = st^2$; $s = 1, t = 1$

In Exercises 23–26, find $\frac{dy}{dx}$ using Implicit Differentiation and Theorem 112.

23. $x^2 \tan y = 50$

24. $(3x^2 + 2y^3)^4 = 2$

25. $\dfrac{x^2 + y}{x + y^2} = 17$

26. $\ln(x^2 + xy + y^2) = 1$

In Exercises 27–30, find $\frac{dz}{dt}$, or $\frac{\partial z}{\partial s}$ and $\frac{\partial z}{\partial t}$, using the supplied information.

27. $\dfrac{\partial z}{\partial x} = 2$, $\dfrac{\partial z}{\partial y} = 1$, $\dfrac{dx}{dt} = 4$, $\dfrac{dy}{dt} = -5$

28. $\dfrac{\partial z}{\partial x} = 1$, $\dfrac{\partial z}{\partial y} = -3$, $\dfrac{dx}{dt} = 6$, $\dfrac{dy}{dt} = 2$

29. $\dfrac{\partial z}{\partial x} = -4$, $\dfrac{\partial z}{\partial y} = 9$,
$\dfrac{\partial x}{\partial s} = 5$, $\dfrac{\partial x}{\partial t} = 7$, $\dfrac{\partial y}{\partial s} = -2$, $\dfrac{\partial y}{\partial t} = 6$

30. $\dfrac{\partial z}{\partial x} = 2$, $\dfrac{\partial z}{\partial y} = 1$,
$\dfrac{\partial x}{\partial s} = -2$, $\dfrac{\partial x}{\partial t} = 3$, $\dfrac{\partial y}{\partial s} = 2$, $\dfrac{\partial y}{\partial t} = -1$

13.6   Directional Derivatives

Partial derivatives give us an understanding of how a surface changes when we move in the *x* and *y* directions. We made the comparison to standing in a rolling meadow and heading due east: the amount of rise/fall in doing so is comparable to f_x. Likewise, the rise/fall in moving due north is comparable to f_y. The steeper the slope, the greater in magnitude f_y.

But what if we didn't move due north or east? What if we needed to move northeast and wanted to measure the amount of rise/fall? Partial derivatives alone cannot measure this. This section investigates **directional derivatives**, which do measure this rate of change.

We begin with a definition.

Definition 93   Directional Derivatives

Let $z = f(x, y)$ be continuous on an open set S and let $\vec{u} = \langle u_1, u_2 \rangle$ be a unit vector. For all points (x, y), the **directional derivative of f at (x, y) in the direction of $\vec{u}$** is

$$D_{\vec{u}} f(x, y) = \lim_{h \to 0} \frac{f(x + hu_1, y + hu_2) - f(x, y)}{h}.$$

The partial derivatives f_x and f_y are defined with similar limits, but only *x* or *y* varies with *h*, not both. Here both *x* and *y* vary with a weighted *h*, determined by a particular unit vector $\vec{u}$. This may look a bit intimidating but in reality it is not too difficult to deal with; it often just requires extra algebra. However, the following theorem reduces this algebraic load.

Theorem 113   Directional Derivatives

Let $z = f(x, y)$ be differentiable at (x_0, y_0), and let $\vec{u} = \langle u_1, u_2 \rangle$ be a unit vector. The directional derivative of f at (x_0, y_0) in the direction of $\vec{u}$ is

$$D_{\vec{u}} f(x_0, y_0) = f_x(x_0, y_0) u_1 + f_y(x_0, y_0) u_2.$$

Proof
This is really a quick application of Definition 90. Because f is differentiable at

---

Notes:

(x_0, y_0),

$$D_{\vec{u}}f(x_0, y_0) = \lim_{h \to 0} \frac{f_x(x_0, y_0)hu_1 + f_y(x_0, y_0)hu_2 + E_x(x_0, y_0)hu_1 + E_y(x_0, y_0)hu_2}{h}$$

$$= f_x(x_0, y_0)u_1 + f_y(x_0, y_0)u_2 + \lim_{h \to 0} E_x(x_0, y_0)u_1 + E_y(x_0, y_0)u_2$$

$$= f_x(x_0, y_0)u_1 + f_y(x_0, y_0)u_2. \qquad \Box$$

Watch the video:
Finding the Directional Derivative — Ex 1 at
`https://youtu.be/uCYOXYXgYQo`

Example 1   Computing directional derivatives

Let $z = 14 - x^2 - y^2$ and let $P = (1, 2)$. Find the directional derivative of f, at P, in the following directions:

1. toward the point $Q = (3, 4)$,

2. in the direction of $\langle 2, -1 \rangle$, and

3. toward the origin.

SOLUTION   The surface is plotted in Figure 13.16, where the point $P = (1, 2)$ is indicated in the x, y-plane as well as the point $(1, 2, 9)$ which lies on the surface of f. We find that $f_x(x, y) = -2x$ and $f_x(1, 2) = -2$; $f_y(x, y) = -2y$ and $f_y(1, 2) = -4$.

1. Let $\vec{u}_1$ be the unit vector that points from the point $(1, 2)$ to the point $Q = (3, 4)$, as shown in the figure. The vector $\overrightarrow{PQ} = \langle 2, 2 \rangle$; the unit vector in this direction is $\vec{u}_1 = \langle 1/\sqrt{2}, 1/\sqrt{2} \rangle$. Thus the directional derivative of f at $(1, 2)$ in the direction of $\vec{u}_1$ is

$$D_{\vec{u}_1}f(1, 2) = -2(1/\sqrt{2}) + (-4)(1/\sqrt{2}) = -6/\sqrt{2}.$$

Thus the instantaneous rate of change in moving from the point $(1, 2, 9)$ on the surface in the direction of $\vec{u}_1$ (which points toward the point Q) is $-3\sqrt{2}$. Moving in this direction moves one steeply downward.
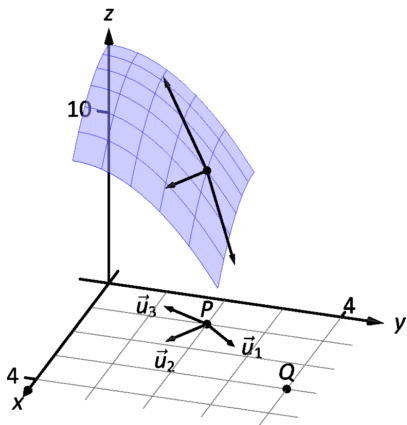
Figure 13.16: Understanding the directional derivative in Example 1.

Notes:

2. We seek the directional derivative in the direction of $\langle 2, -1 \rangle$. The unit vector in this direction is $\vec{u}_2 = \langle 2/\sqrt{5}, -1/\sqrt{5} \rangle$. Thus the directional derivative of f at $(1, 2)$ in the direction of $\vec{u}_2$ is

$$D_{\vec{u}_2} f(1, 2) = -2(2/\sqrt{5}) + (-4)(-1/\sqrt{5}) = 0.$$

Starting on the surface of f at $(1, 2)$ and moving in the direction of $\langle 2, -1 \rangle$ (or $\vec{u}_2$) results in no instantaneous change in z-value. This is analogous to standing on the side of a hill and choosing a direction to walk that does not change the elevation. One neither walks up nor down, rather just "along the side" of the hill.

Finding these directions of "no elevation change" is important.

3. At $P = (1, 2)$, the direction towards the origin is given by the vector $\langle -1, -2 \rangle$; the unit vector in this direction is $\vec{u}_3 = \langle -1/\sqrt{5}, -2/\sqrt{5} \rangle$. The directional derivative of f at P in the direction of the origin is

$$D_{\vec{u}_3} f(1, 2) = -2(-1/\sqrt{5}) + (-4)(-2/\sqrt{5}) = 10/\sqrt{5}.$$

Moving towards the origin means "walking uphill" quite steeply, with an initial slope of $2\sqrt{5}$.

As we study directional derivatives, it will help to make an important connection between the unit vector $\vec{u} = \langle u_1, u_2 \rangle$ that describes the direction and the partial derivatives f_x and f_y. We start with a definition and follow this with a Key Idea.

Definition 94   Gradient

Let $z = f(x, y)$ be differentiable on an open set S that contains the point (x_0, y_0).

1. The **gradient of f** is $\nabla f(x, y) = \langle f_x(x, y), f_y(x, y) \rangle$.

2. The **gradient of f at** (x_0, y_0) **is** $\nabla f(x_0, y_0) = \langle f_x(x_0, y_0), f_y(x_0, y_0) \rangle$.

To simplify notation, we often express the gradient as $\nabla f = \langle f_x, f_y \rangle$. It is often useful to think of the gradient ∇ as an operator:

$$\nabla = \left\langle \frac{\partial}{\partial x}, \frac{\partial}{\partial y} \right\rangle.$$

The operator ∇ only has any meaning when it operates on a function; it doesn't mean anything by itself. But this notation does help us to apply it correctly to find the gradient. The gradient allows us to compute directional derivatives in terms of a dot product.

Note: The symbol "∇" is named "nabla," derived from the Greek name of a Jewish harp. Oddly enough, in mathematics the expression ∇f is pronounced "del f."

Notes:

> **Key Idea 59   The Gradient and Directional Derivatives**
> The directional derivative of $z = f(x, y)$ in the direction of the unit vector $\vec{u}$ is
> $$D_{\vec{u}} f = \nabla f \cdot \vec{u}.$$

The properties of the dot product previously studied allow us to investigate the properties of the directional derivative. Given that the directional derivative gives the instantaneous rate of change of z when moving in the direction of $\vec{u}$, three questions naturally arise:

1. In what direction(s) is the change in z the greatest (i.e., the "steepest up-hill")?

2. In what direction(s) is the change in z the least (i.e., the "steepest down-hill")?

3. In what direction(s) is there no change in z?

Using the key property of the dot product, we have

$$\nabla f \cdot \vec{u} = \|\nabla f\| \ \|\vec{u}\| \cos \theta = \|\nabla f\| \cos \theta, \tag{13.4}$$

where θ is the angle between the gradient and $\vec{u}$. (Since $\vec{u}$ is a unit vector, $\|\vec{u}\| = 1$.) This equation allows us to answer the three questions stated previously.

1. Equation (13.4) is maximized when $\cos \theta = 1$, i.e., when the gradient and $\vec{u}$ have the same direction. We conclude the gradient points in the direction of greatest z change.

2. Equation (13.4) is minimized when $\cos \theta = -1$, i.e., when the gradient and $\vec{u}$ have opposite directions. We conclude the gradient points in the opposite direction of the least z change.

3. Equation (13.4) is 0 when $\cos \theta = 0$, i.e., when the gradient and $\vec{u}$ are orthogonal to each other. We conclude the gradient is orthogonal to directions of no z change.

This result is rather amazing. Once again imagine standing in a rolling meadow and face the direction that leads you steepest uphill. Then the direction that leads steepest downhill is directly behind you, and side–stepping either left or right (i.e., moving perpendicularly to the direction you face) does not change your elevation at all.

Recall that a level curve is defined as a curve in the x-y plane along which the z-values of a function do not change. Let a surface $z = f(x, y)$ be given, and let's

---

Notes:

represent one such level curve as a vector–valued function, $\vec{r}(t) = \langle x(t), y(t)\rangle$. As the output of f does not change along this curve, $f(x(t), y(t)) = c$ for all t, for some constant c.

Since f is constant for all t, $\frac{df}{dt} = 0$. By the Multivariable Chain Rule, we also know

$$\frac{df}{dt} = f_x(x, y)x'(t) + f_y(x, y)y'(t)$$
$$= \langle f_x(x, y), f_y(x, y)\rangle \cdot \langle x'(t), y'(t)\rangle$$
$$= \nabla f \cdot \vec{r}'(t)$$
$$= 0.$$

This last equality states $\nabla f \cdot \vec{r}'(t) = 0$: the gradient is orthogonal to the derivative of $\vec{r}$. Our conclusion: at any point on a surface, the gradient at that point is orthogonal to the level curve that passes through that point.

We restate these ideas in a theorem, then use them in an example.

Theorem 114   The Gradient and Directional Derivatives

Let $z = f(x, y)$ be differentiable on an open set S with gradient ∇f, let $P = (x_0, y_0)$ be a point in S and let $\vec{u}$ be a unit vector.

1. The maximum value of $D_{\vec{u}}f(x_0, y_0)$ is $\|\nabla f(x_0, y_0)\|$; the direction of maximal z increase is $\nabla f(x_0, y_0)$.

2. The minimum value of $D_{\vec{u}}f(x_0, y_0)$ is $-\|\nabla f(x_0, y_0)\|$; the direction of minimal z increase is $-\nabla f(x_0, y_0)$.

3. At P, $\nabla f(x_0, y_0)$ is orthogonal to the level curve passing through $(x_0, y_0, f(x_0, y_0))$.

Example 2   Finding directions of maximal and minimal increase
Let $f(x, y) = \sin x \cos y$ and let $P = (\pi/3, \pi/3)$. Find the directions of maximal/minimal increase, and find a direction where the instantaneous rate of z change is 0.

SOLUTION   We begin by finding the gradient. We see that $f_x = \cos x \cos y$ and $f_y = -\sin x \sin y$, thus

$$\nabla f = \langle \cos x \cos y, -\sin x \sin y\rangle \quad \text{and, at } P, \quad \nabla f\left(\frac{\pi}{3}, \frac{\pi}{3}\right) = \left\langle \frac{1}{4}, -\frac{3}{4}\right\rangle.$$

Notes:

Thus the direction of maximal increase is $\langle 1/4, -3/4 \rangle$. In this direction, the instantaneous rate of z change is $\|\langle 1/4, -3/4 \rangle\| = \sqrt{10}/4$.

Figure 13.17 shows the surface plotted from two different perspectives. In each, the gradient is drawn at P with a dashed line (because of the nature of this surface, the gradient points "into" the surface). Let $\vec{u} = \langle u_1, u_2 \rangle$ be the unit vector in the direction of ∇f at P. Each graph of the figure also contains the vector $\langle u_1, u_2, \|\nabla f\| \rangle$. This vector has a "run" of 1 (because in the x-y plane it moves 1 unit) and a "rise" of $\|\nabla f\|$, hence we can think of it as a vector with slope of $\|\nabla f\|$ in the direction of ∇f, helping us visualize how "steep" the surface is in its steepest direction.

The direction of minimal increase is $\langle -1/4, 3/4 \rangle$; in this direction the instantaneous rate of z change is $-\sqrt{10}/4$.

Any direction orthogonal to ∇f is a direction of no z change. We have two choices: the direction of $\langle 3, 1 \rangle$ and the direction of $\langle -3, -1 \rangle$. The unit vector in the direction of $\langle 3, 1 \rangle$ is shown in each graph of the figure as well. The level curve at $z = \sqrt{3}/4$ is drawn: recall that along this curve the z-values do not change. Since $\langle 3, 1 \rangle$ is a direction of no z-change, this vector is tangent to the level curve at P.

(a)

(b)

Figure 13.17: Graphing the surface and important directions in Example 2.
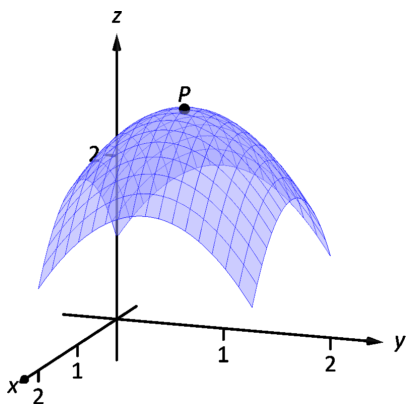
Figure 13.18: At the top of a paraboloid, all directional derivatives are 0.

Example 3   **Understanding when $\nabla f = \vec{0}$**
Let $f(x, y) = -x^2 + 2x - y^2 + 2y + 1$. Find the directional derivative of f in any direction at $P = (1, 1)$.

SOLUTION   We find $\nabla f = \langle -2x + 2, -2y + 2 \rangle$. At P, we have $\nabla f(1, 1) = \langle 0, 0 \rangle$. According to Theorem 114, this is the direction of maximal increase. However, $\langle 0, 0 \rangle$ is directionless; it has no displacement. And regardless of the unit vector $\vec{u}$ chosen, $D_{\vec{u}} f = 0$.

Figure 13.18 helps us understand what this means. We can see that P lies at the top of a paraboloid. In all directions, the instantaneous rate of change is 0.

So what is the direction of maximal increase? It is fine to give an answer of $\vec{0} = \langle 0, 0 \rangle$, as this indicates that all directional derivatives are 0.

The fact that the gradient of a surface always points in the direction of steepest increase/decrease is very useful, as illustrated in the following example.

Example 4   **The flow of water downhill**
Consider the surface given by $f(x, y) = 20 - x^2 - 2y^2$. Water is poured on the surface at $(1, 1/4)$. What path does it take as it flows downhill?

SOLUTION   Let $\vec{r}(t) = \langle x(t), y(t) \rangle$ be the vector–valued function describing the path of the water in the x-y plane; we seek $x(t)$ and $y(t)$. We know

---

Notes:

that water will always flow downhill in the steepest direction; therefore, at any point on its path, it will be moving in the direction of $-\nabla f$. (We ignore the physical effects of momentum on the water.) Thus $\vec{r}\,'(t)$ will be parallel to ∇f, and there is some constant c such that $c\nabla f = \vec{r}\,'(t) = \langle x'(t), y'(t)\rangle$.

We find $\nabla f = \langle -2x, -4y\rangle$ and write $x'(t)$ as $\frac{dx}{dt}$ and $y'(t)$ as $\frac{dy}{dt}$. Then

$$c\nabla f = \langle x'(t), y'(t)\rangle$$

$$\langle -2cx, -4cy\rangle = \left\langle \frac{dx}{dt}, \frac{dy}{dt}\right\rangle.$$

This implies

$$-2cx = \frac{dx}{dt} \quad \text{and} \quad -4cy = \frac{dy}{dt}, \text{ i.e.,}$$

$$c = -\frac{1}{2x}\frac{dx}{dt} \quad \text{and} \quad c = -\frac{1}{4y}\frac{dy}{dt}.$$

As c equals both expressions, we have

$$\frac{1}{2x}\frac{dx}{dt} = \frac{1}{4y}\frac{dy}{dt}.$$

To find an explicit relationship between x and y, we can integrate both sides with respect to t. Recall from our study of differentials that $\frac{dx}{dt}dt = dx$. Thus:

$$\int \frac{1}{2x}\frac{dx}{dt}\,dt = \int \frac{1}{4y}\frac{dy}{dt}\,dt$$

$$\int \frac{1}{2x}\,dx = \int \frac{1}{4y}\,dy$$

$$\frac{1}{2}\ln|x| = \frac{1}{4}\ln|y| + C_1$$

$$2\ln|x| = \ln|y| + C_1$$

$$\ln\left|x^2\right| = \ln|y| + C_1$$

Now raise both sides as a power of e:

$$x^2 = e^{\ln|y| + C_1}$$

$$x^2 = e^{\ln|y|}e^{C_1} \qquad \text{(Note that } e^{C_1} \text{ is just a constant.)}$$

$$x^2 = yC_2$$

$$\frac{1}{C_2}x^2 = y \qquad \text{(Note that } 1/C_2 \text{ is just a constant.)}$$

$$Cx^2 = y.$$

Notes:

As the water started at the point $(1, 1/4)$, we can solve for C:

$$C(1)^2 = \frac{1}{4} \quad \Rightarrow \quad C = \frac{1}{4}.$$

Thus the water follows the curve $y = x^2/4$ in the x-y plane. The surface and the path of the water is graphed in Figure 13.19(a). In part (b) of the figure, the level curves of the surface are plotted in the x-y plane, along with the curve $y = x^2/4$. Notice how the path intersects the level curves at right angles. As the path follows the gradient downhill, this reinforces the fact that the gradient is orthogonal to level curves.

(a)

(b)

Figure 13.19: A graph of the surface described in Example 4 along with the path in the x-y plane with the level curves.

Functions of Three Variables

The concepts of directional derivatives and the gradient are easily extended to three (and more) variables. We combine the concepts behind Definitions 93 and 94 and Theorem 113 into one set of definitions.

Definition 95   Directional Derivatives and Gradient with Three Variables

Let $w = F(x, y, z)$ be differentiable on an open ball B and let $\vec{u}$ be a unit vector in $\mathbb{R}^3$.

1. The **gradient** of F is $\nabla F = \langle F_x, F_y, F_z \rangle$.

2. The **gradient of F at** (x_0, y_0, z_0) **is**

    $$\nabla F(x_0, y_0, z_0) = \langle f_x(x_0, y_0, z_0), f_y(x_0, y_0, z_0), f_z(x_0, y_0, z_0) \rangle.$$

3. The **directional derivative of F in the direction of $\vec{u}$** is

    $$D_{\vec{u}} F = \nabla F \cdot \vec{u}.$$

The same properties of the gradient given in Theorem 114, when f is a function of two variables, hold for F, a function of three variables.

Notes:

> **Theorem 115**   **The Gradient and Directional Derivatives with Three Variables**
>
> Let $w = F(x, y, z)$ be differentiable on an open set S with gradient ∇F, let $P = (x_0, y_0, z_0)$ be a point in S, and let $\vec{u}$ be a unit vector.
>
> 1. The maximum value of $D_{\vec{u}} F(x_0, y_0, z_0)$ is $\|\nabla F(x_0, y_0, z_0)\|$; the direction of maximal increase is $\nabla F(x_0, y_0, z_0)$.
>
> 2. The minimum value of $D_{\vec{u}} F(x_0, y_0, z_0)$ is $-\|\nabla F(x_0, y_0, z_0)\|$; the direction of minimal increase is $-\nabla F(x_0, y_0, z_0)$.
>
> 3. At P, $D_{\vec{u}} F(x_0, y_0, z_0) = 0$ when $\nabla F(x_0, y_0, z_0)$ and $\vec{u}$ are orthogonal.

We interpret the third statement of the theorem as "the gradient is orthogonal to level surfaces," the three–variable analogue to level curves.

Example 5   **Finding directional derivatives with functions of three variables**

If a point source S is radiating energy, the intensity I at a given point P in space is inversely proportional to the square of the distance between S and P. That is, when $S = (0, 0, 0)$, $I(x, y, z) = \dfrac{k}{x^2 + y^2 + z^2}$ for some constant k.

Let $k = 1$, let $\vec{u} = \langle 2/3, 2/3, 1/3 \rangle$ be a unit vector, and let $P = (2, 5, 3)$. Measure distances in inches. Find the directional derivative of I at P in the direction of $\vec{u}$, and find the direction of greatest intensity increase at P.

SOLUTION   We need the gradient ∇I, meaning we need I_x, I_y and I_z. Each partial derivative requires a simple application of the Quotient Rule, giving

$$\nabla I = \left\langle \frac{-2x}{(x^2 + y^2 + z^2)^2}, \frac{-2y}{(x^2 + y^2 + z^2)^2}, \frac{-2z}{(x^2 + y^2 + z^2)^2} \right\rangle$$

$$\nabla I(2, 5, 3) = \left\langle \frac{-4}{1444}, \frac{-10}{1444}, \frac{-6}{1444} \right\rangle$$

$$D_{\vec{u}} I = \nabla I(2, 5, 3) \cdot \vec{u}$$

$$= -\frac{17}{2166}.$$

The directional derivative tells us that moving in the direction of $\vec{u}$ from P results in a slight decrease in intensity. (The intensity is decreasing as $\vec{u}$ moves one farther from the origin than P.)

Notes:

The gradient gives the direction of greatest intensity increase. Notice that

$$\nabla I(2,5,3) = \left\langle \frac{-4}{1444}, \frac{-10}{1444}, \frac{-6}{1444} \right\rangle$$

$$= \frac{2}{1444} \left\langle -2, -5, -3 \right\rangle.$$

That is, the gradient at $(2,5,3)$ is pointing in the direction of $\langle -2, -5, -3 \rangle$, that is, towards the origin. That should make intuitive sense: the greatest increase in intensity is found by moving towards to source of the energy.

The directional derivative allows us to find the instantaneous rate of z change in any direction at a point. We can use these instantaneous rates of change to define lines and planes that are *tangent* to a surface at a point, which is the topic of the next section.

Notes:

Exercises 13.6

Terms and Concepts

1. What is the difference between a directional derivative and a partial derivative?

2. For what $\vec{u}$ is $D_{\vec{u}}f = f_x$?

3. For what $\vec{u}$ is $D_{\vec{u}}f = f_y$?

4. The gradient is _____ to level curves.

5. The gradient points in the direction of _____ increase.

6. It is generally more informative to view the directional derivative not as the result of a limit, but rather as the result of a _____ product.

Problems

In Exercises 7–12, a function $z = f(x, y)$ is given. Find ∇f.

7. $f(x, y) = -x^2y + xy^2 + xy$

8. $f(x, y) = \sin x \cos y$

9. $f(x, y) = \dfrac{1}{x^2 + y^2 + 1}$

10. $f(x, y) = -4x + 3y$

11. $f(x, y) = x^2 + 2y^2 - xy - 7x$

12. $f(x, y) = x^2y^3 - 2x$

In Exercises 13–18, a function $z = f(x, y)$ and a point P are given. Find the directional derivative of f in the indicated directions. Note: these are the same functions as in Exercises 7–12.

13. $f(x, y) = -x^2y + xy^2 + xy$, $P = (2, 1)$
    (a) In the direction of $\vec{v} = \langle 3, 4 \rangle$
    (b) In the direction toward the point $Q = (1, -1)$.

14. $f(x, y) = \sin x \cos y$, $P = \left(\dfrac{\pi}{4}, \dfrac{\pi}{3}\right)$
    (a) In the direction of $\vec{v} = \langle 1, 1 \rangle$.
    (b) In the direction toward the point $Q = (0, 0)$.

15. $f(x, y) = \dfrac{1}{x^2 + y^2 + 1}$, $P = (1, 1)$.
    (a) In the direction of $\vec{v} = \langle 1, -1 \rangle$.
    (b) In the direction toward the point $Q = (-2, -2)$.

16. $f(x, y) = -4x + 3y$, $P = (5, 2)$
    (a) In the direction of $\vec{v} = \langle 3, 1 \rangle$.
    (b) In the direction toward the point $Q = (2, 7)$.

17. $f(x, y) = x^2 + 2y^2 - xy - 7x$, $P = (4, 1)$
    (a) In the direction of $\vec{v} = \langle -2, 5 \rangle$
    (b) In the direction toward the point $Q = (4, 0)$.

18. $f(x, y) = x^2y^3 - 2x$, $P = (1, 1)$
    (a) In the direction of $\vec{v} = \langle 3, 3 \rangle$
    (b) In the direction toward the point $Q = (1, 2)$.

In Exercises 19–24, a function $z = f(x, y)$ and a point P are given.
    (a) Find the direction of maximal increase of f at P.
    (b) What is the maximal value of $D_{\vec{u}}f$ at P?
    (c) Find the direction of minimal increase of f at P.
    (d) Give a direction $\vec{u}$ such that $D_{\vec{u}}f = 0$ at P.

Note: these are the same functions and points as in Exercises 13 – 18.

19. $f(x, y) = -x^2y + xy^2 + xy$, $P = (2, 1)$

20. $f(x, y) = \sin x \cos y$, $P = \left(\dfrac{\pi}{4}, \dfrac{\pi}{3}\right)$

21. $f(x, y) = \dfrac{1}{x^2 + y^2 + 1}$, $P = (1, 1)$.

22. $f(x, y) = -4x + 3y$, $P = (5, 4)$.

23. $f(x, y) = x^2 + 2y^2 - xy - 7x$, $P = (4, 1)$

24. $f(x, y) = x^2y^3 - 2x$, $P = (1, 1)$

In Exercises 25–28, a function $w = F(x, y, z)$, a vector $\vec{v}$ and a point P are given.
    (a) Find $\nabla F(x, y, z)$.
    (b) Find $D_{\vec{u}}F$ at P.

25. $F(x, y, z) = 3x^2z^3 + 4xy - 3z^2$, $\vec{v} = \langle 1, 1, 1 \rangle$, $P = (3, 2, 1)$

26. $F(x, y, z) = \sin(x)\cos(y)e^z$, $\vec{v} = \langle 2, 2, 1 \rangle$, $P = (0, 0, 0)$

27. $F(x, y, z) = x^2y^2 - y^2z^2$, $\vec{v} = \langle -1, 7, 3 \rangle$, $P = (1, 0, -1)$

28. $F(x, y, z) = \dfrac{2}{x^2 + y^2 + z^2}$, $\vec{v} = \langle 1, 1, -2 \rangle$, $P = (1, 1, 1)$

13.7 Tangent Lines, Normal Lines, and Tangent Planes

Derivatives and tangent lines go hand–in–hand. Given $y = f(x)$, the line tangent to the graph of f at $x = x_0$ is the line through $(x_0, f(x_0))$ with slope $f'(x_0)$; that is, the slope of the tangent line is the instantaneous rate of change of f at x_0.

When dealing with functions of two variables, the graph is no longer a curve but a surface. At a given point on the surface, it seems there are many lines that fit our intuition of being "tangent" to the surface.

In Figure 13.20 we see lines that are tangent to curves in space. Since each curve lies on a surface, it makes sense to say that the lines are also tangent to the surface. The next definition formally defines what it means to be "tangent to a surface."
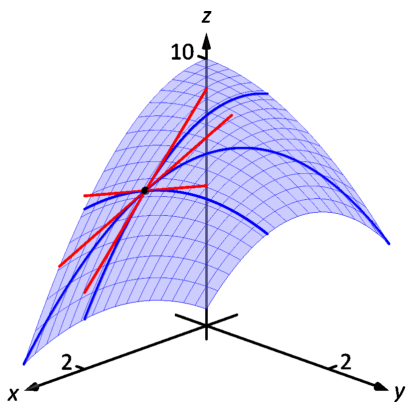
Figure 13.20: Showing various lines tangent to a surface.

> **Definition 96   Directional Tangent Line**
> Let $z = f(x, y)$ be differentiable on an open set S containing (x_0, y_0) and let $\vec{u} = \langle u_1, u_2 \rangle$ be a unit vector. The line $\ell_{\vec{u}}$ through $(x_0, y_0, f(x_0, y_0))$ parallel to $\langle u_1, u_2, D_{\vec{u}} f(x_0, y_0) \rangle$ is the **tangent line to f in the direction of $\vec{u}$ at** (x_0, y_0).

We will also follow the convention that

$$\ell_{\langle 1,0 \rangle} = \ell_x \qquad \text{and} \qquad \ell_{\langle 0,1 \rangle} = \ell_y.$$

It is instructive to consider each of three directions given in the definition in terms of "slope." The direction of ℓ_x is $\langle 1, 0, f_x(x_0, y_0) \rangle$; that is, the "run" is one unit in the x-direction and the "rise" is $f_x(x_0, y_0)$ units in the z-direction. Note how the slope is just the partial derivative with respect to x. A similar statement can be made for ℓ_y. The direction of $\ell_{\vec{u}}$ is $\langle u_1, u_2, D_{\vec{u}} f(x_0, y_0) \rangle$; the "run" is one unit in the $\vec{u}$ direction (where $\vec{u}$ is a unit vector) and the "rise" is the directional derivative of z in that direction.

Definition 96 leads to the following parametric equations of directional tangent lines:

$$\ell_x(t) = \begin{cases} x = x_0 + t \\ y = y_0 \\ z = z_0 + f_x(x_0, y_0)t \end{cases}, \quad \ell_y(t) = \begin{cases} x = x_0 \\ y = y_0 + t \\ z = z_0 + f_y(x_0, y_0)t \end{cases} \quad \text{and} \quad \ell_{\vec{u}}(t) = \begin{cases} x = x_0 + u_1 t \\ y = y_0 + u_2 t \\ z = z_0 + D_{\vec{u}} f(x_0, y_0)t \end{cases}$$

Notes:

Watch the video:
Determining a Unit Normal Vector to a Surface at
`https://youtu.be/DRBNp7SZCvU`

Example 1   **Finding directional tangent lines**

Find the lines tangent to the surface $z = \sin x \cos y$ at $(\pi/2, \pi/2)$ in the x and y directions and also in the direction of $\vec{v} = \langle -1, 1 \rangle$.

SOLUTION   The partial derivatives with respect to x and y are:

$$f_x(x, y) = \cos x \cos y \quad \Rightarrow \quad f_x(\pi/2, \pi/2) = 0$$
$$f_y(x, y) = -\sin x \sin y \quad \Rightarrow \quad f_y(\pi/2, \pi/2) = -1.$$

At $(\pi/2, \pi/2)$, the z-value is 0.

Thus the parametric equations of the line tangent to f at $(\pi/2, \pi/2)$ in the directions of x and y are:

$$\ell_x(t) = \begin{cases} x = \pi/2 + t \\ y = \pi/2 \\ z = 0 \end{cases} \quad \text{and} \quad \ell_y(t) = \begin{cases} x = \pi/2 \\ y = \pi/2 + t \\ z = -t \end{cases}.$$

The two lines are shown with the surface in Figure 13.21(a). To find the equation of the tangent line in the direction of $\vec{v}$, we first find the unit vector in the direction of $\vec{v}$: $\vec{u} = \langle -1/\sqrt{2}, 1/\sqrt{2} \rangle$. The directional derivative at $(\pi/2, \pi, 2)$ in the direction of $\vec{u}$ is

$$D_{\vec{u}}f(\pi/2, \pi, 2) = \langle 0, -1 \rangle \cdot \langle -1/\sqrt{2}, 1/\sqrt{2} \rangle = -1/\sqrt{2}.$$

Thus the directional tangent line is

$$\ell_{\vec{u}}(t) = \begin{cases} x = \pi/2 - t/\sqrt{2} \\ y = \pi/2 + t/\sqrt{2} \\ z = -t/\sqrt{2} \end{cases}.$$

The curve through $(\pi/2, \pi/2, 0)$ in the direction of $\vec{v}$ is shown in Figure 13.21(b) along with $\ell_{\vec{u}}(t)$.
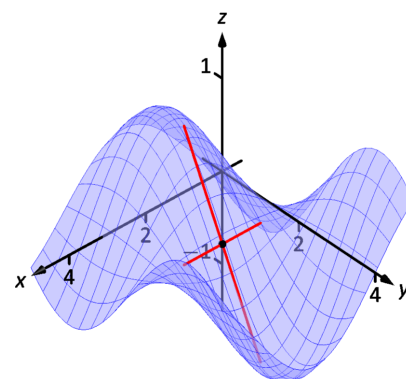
(a)
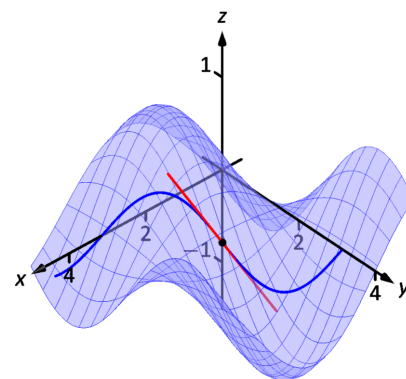
(b)

Figure 13.21: A surface and directional tangent lines in Example 1.

Notes:

Example 2   **Finding directional tangent lines**

Let $f(x, y) = 4xy - x^4 - y^4$. Find the equations of *all* directional tangent lines to f at $(1, 1)$.

SOLUTION   First note that $f(1, 1) = 2$. We need to compute directional derivatives, so we need ∇f. We begin by computing partial derivatives.

$$f_x = 4y - 4x^3 \Rightarrow f_x(1, 1) = 0; \quad f_y = 4x - 4y^3 \Rightarrow f_y(1, 1) = 0.$$

Thus $\nabla f(1, 1) = \langle 0, 0 \rangle$. Let $\vec{u} = \langle u_1, u_2 \rangle$ be any unit vector. The directional derivative of f at $(1, 1)$ will be $D_{\vec{u}}f(1, 1) = \langle 0, 0 \rangle \cdot \langle u_1, u_2 \rangle = 0$. It does not matter what direction we choose; the directional derivative is always 0. Therefore

$$\ell_{\vec{u}}(t) = \begin{cases} x = 1 + u_1 t \\ y = 1 + u_2 t \\ z = 2 \end{cases}.$$

Figure 13.22 shows a graph of f and the point $(1, 1, 2)$. Note that this point comes at the top of a "hill," and therefore every tangent line through this point will have a "slope" of 0.

That is, consider any curve on the surface that goes through this point. Each curve will have a relative maximum at this point, hence its tangent line will have a slope of 0. The following section investigates the points on surfaces where all tangent lines have a slope of 0.
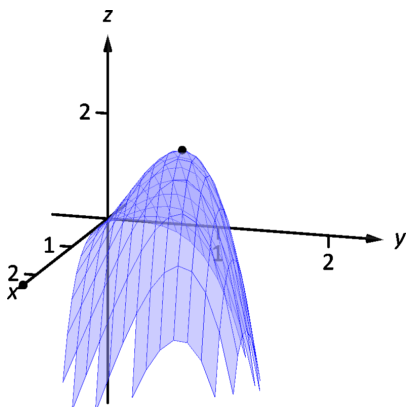
Figure 13.22: Graphing f in Example 2.

Normal Lines

When dealing with a function $y = f(x)$ of one variable, we stated that a line through $(c, f(c))$ was *tangent* to f if the line had a slope of $f'(c)$ and was *normal* (or, *perpendicular, orthogonal*) to f if it had a slope of $-1/f'(c)$. We extend the concept of normal, or orthogonal, to functions of two variables.

Let $z = f(x, y)$ be a differentiable function of two variables. By Definition 96, at (x_0, y_0), $\ell_x(t)$ is a line parallel to the vector $\vec{d}_x = \langle 1, 0, f_x(x_0, y_0) \rangle$ and $\ell_y(t)$ is a line parallel to $\vec{d}_y = \langle 0, 1, f_y(x_0, y_0) \rangle$. Since lines in these directions through $(x_0, y_0, f(x_0, y_0))$ are *tangent* to the surface, a line through this point and orthogonal to these directions would be *orthogonal*, or *normal*, to the surface. We can use this direction to create a normal line.

The direction of the normal line is orthogonal to $\vec{d}_x$ and $\vec{d}_y$, hence the direction is parallel to $\vec{d}_n = \vec{d}_x \times \vec{d}_y$. It turns out this cross product has a very simple form:

$$\vec{d}_x \times \vec{d}_y = \langle 1, 0, f_x \rangle \times \langle 0, 1, f_y \rangle = \langle -f_x, -f_y, 1 \rangle.$$

It is often more convenient to refer to the opposite of this direction, namely $\langle f_x, f_y, -1 \rangle$. This leads to a definition.

Notes:

> **Definition 97   Normal Line**
> Let $z = f(x, y)$ be differentiable on an open set S containing (x_0, y_0).
>
> 1. A nonzero vector parallel to $\vec{n} = \langle f_x(x_0, y_0), f_y(x_0, y_0), -1 \rangle$ is **orthogonal to f at $P = \big(x_0, y_0, f(x_0, y_0)\big)$**.
>
> 2. The line ℓ_n through P with direction parallel to $\vec{n}$ is the **normal line to f at P**.

Thus the parametric equations of the normal line to a surface f at $\big(x_0, y_0, f(x_0, y_0)\big)$ is:

$$\ell_n(t) = \begin{cases} x = x_0 + f_x(x_0, y_0)t \\ y = y_0 + f_y(x_0, y_0)t \\ z = f(x_0, y_0) - t. \end{cases}$$

Example 3   Finding a normal line
Find the equation of the normal line to $z = -x^2 - y^2 + 2$ at $(0, 1)$.

SOLUTION   We find $z_x(x, y) = -2x$ and $z_y(x, y) = -2y$; at $(0, 1)$, we have $z_x = 0$ and $z_y = -2$. We take the direction of the normal line, following Definition 97, to be $\vec{n} = \langle 0, -2, -1 \rangle$. The line with this direction going through the point $(0, 1, 1)$ is

$$\ell_n(t) = \begin{cases} x = 0 \\ y = -2t + 1 \\ z = -t + 1 \end{cases} \quad \text{or} \quad \ell_n(t) = \langle 0, -2, -1 \rangle\, t + \langle 0, 1, 1 \rangle.$$

The surface $z = -x^2 - y^2 + 2$, along with the found normal line, is graphed in Figure 13.23.

The direction of the normal line has many uses, one of which is the definition of the **tangent plane** which we define shortly. Another use is in measuring distances from the surface to a point. Given a point Q in space, it is a general geometric concept to define the distance from Q to the surface as being the length of the shortest line segment $\overline{PQ}$ over all points P on the surface. This, in turn, implies that $\overrightarrow{PQ}$ will be orthogonal to the surface at P. Therefore we can measure the distance from Q to the surface f by finding a point P on the surface such that $\overrightarrow{PQ}$ is parallel to the normal line to f at P.

Example 4   Finding the distance from a point to a surface
Let $f(x, y) = 2 - x^2 - y^2$ and let $Q = (2, 2, 2)$. Find the distance from Q to the surface defined by f.
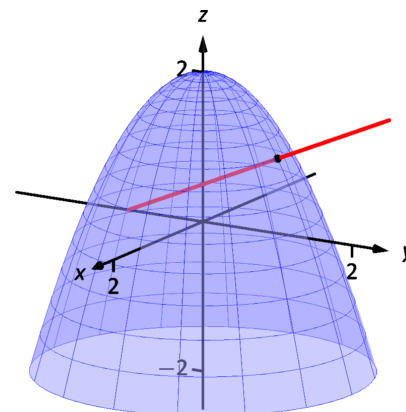
Figure 13.23: Graphing a surface with a normal line from Example 3.

Notes:

SOLUTION   This surface is used in Example 2, so we know that at (x, y), the direction of the normal line will be $\vec{d}_n = \langle -2x, -2y, -1 \rangle$. A point P on the surface will have coordinates $(x, y, 2 - x^2 - y^2)$, so $\overrightarrow{PQ} = \langle 2 - x, 2 - y, x^2 + y^2 \rangle$. To find where $\overrightarrow{PQ}$ is parallel to $\vec{d}_n$, we need to find x, y and c such that $c\overrightarrow{PQ} = \vec{d}_n$.

$$c\overrightarrow{PQ} = \vec{d}_n$$
$$c \langle 2 - x, 2 - y, x^2 + y^2 \rangle = \langle -2x, -2y, -1 \rangle .$$

This implies

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

In each equation, we can solve for c:

$$c = \frac{-2x}{2 - x} = \frac{-2y}{2 - y} = \frac{-1}{x^2 + y^2}.$$

The first two fractions imply $x = y$, and so the last fraction can be rewritten as $c = -1/(2x^2)$. Then

$$\frac{-2x}{2 - x} = \frac{-1}{2x^2}$$
$$-2x(2x^2) = -1(2 - x)$$
$$4x^3 = 2 - x$$
$$4x^3 + x - 2 = 0.$$

This last equation is a cubic, which is not difficult to solve with a numeric solver. We find that $x = 0.689$, hence $P = (0.689, 0.689, 1.051)$. We find the distance from Q to the surface of f is

$$\left\| \overrightarrow{PQ} \right\| = \sqrt{(2 - 0.689)^2 + (2 - 0.689)^2 + (2 - 1.051)^2} = 2.083.$$

We can take the concept of measuring the distance from a point to a surface to find a point Q a particular distance from a surface at a given point P on the surface.

Example 5   **Finding a point a set distance from a surface**
Let $f(x, y) = x - y^2 + 3$. Let $P = \big(2, 1, f(2, 1)\big) = (2, 1, 4)$. Find points Q in space that are 4 units from the surface of f at P. That is, find Q such that $\left\| \overrightarrow{PQ} \right\| = 4$ and $\overrightarrow{PQ}$ is orthogonal to f at P.

Notes:

SOLUTION   We begin by finding partial derivatives:

$$f_x(x, y) = 1 \qquad \Rightarrow \qquad f_x(2, 1) = 1$$
$$f_y(x, y) = -2y \qquad \Rightarrow \qquad f_y(2, 1) = -2$$

The vector $\vec{n} = \langle 1, -2, -1 \rangle$ is orthogonal to f at P. For reasons that will become more clear in a moment, we find the unit vector in the direction of $\vec{n}$:

$$\vec{u} = \frac{\vec{n}}{\|\vec{n}\|} = \frac{\langle 1, -2, -1 \rangle}{\sqrt{6}}.$$

Thus a the normal line to f at P can be written as

$$\ell_n(t) = \langle 2, 1, 4 \rangle + \frac{t}{\sqrt{6}} \langle 1, -2, -1 \rangle.$$

An advantage of this parametrization of the line is that letting $t = t_0$ gives a point on the line that is $|t_0|$ units from P. (This is because the direction of the line is given in terms of a unit vector.) There are thus two points in space 4 units from P:

$$Q_1 = \ell_n(4) \qquad\qquad\qquad Q_2 = \ell_n(-4)$$
$$= \left\langle 2 + \frac{4}{\sqrt{6}}, 1 - \frac{8}{\sqrt{6}}, 4 - \frac{4}{\sqrt{6}} \right\rangle \qquad = \left\langle 2 - \frac{4}{\sqrt{6}}, 1 + \frac{8}{\sqrt{6}}, 4 + \frac{4}{\sqrt{6}} \right\rangle$$

The surface is graphed along with points P, Q_1, Q_2 and a portion of the normal line to f at P.

Tangent Planes

We can use the direction of the normal line to define a plane. With $a = f_x(x_0, y_0)$, $b = f_y(x_0, y_0)$ and $P = (x_0, y_0, f(x_0, y_0))$, the vector $\vec{n} = \langle a, b, -1 \rangle$ is orthogonal to f at P. The plane through P with normal vector $\vec{n}$ is therefore **tangent** to f at P.

Figure 13.24: Graphing the surface in Example 5 along with points 4 units from the surface.

> **Definition 98   Tangent Plane**
>
> Let $z = f(x, y)$ be differentiable on an open set S containing (x_0, y_0), where $a = f_x(x_0, y_0)$, $b = f_y(x_0, y_0)$, $\vec{n} = \langle a, b, -1 \rangle$ and $P = (x_0, y_0, f(x_0, y_0))$.
>
> The plane through P with normal vector $\vec{n}$ is the **tangent plane to f at P**. The standard form of this plane is
>
> $$a(x - x_0) + b(y - y_0) - (z - f(x_0, y_0)) = 0.$$

Notes:

Example 6 **Finding tangent planes**

Find the equation of the tangent plane to $z = -x^2 - y^2 + 2$ at $(0, 1)$.

SOLUTION Note that this is the same surface and point used in Example 3. There we found $\vec{n} = \langle 0, -2, -1 \rangle$ and $P = (0, 1, 1)$. Therefore the equation of the tangent plane is

$$-2(y - 1) - (z - 1) = 0.$$

The surface $z = -x^2 - y^2 + 2$ and tangent plane are graphed in Figure 13.25.

Figure 13.25: Graphing a surface with tangent plane from Example 6.

Example 7 **Using the tangent plane to approximate function values**

The point $(3, -1, 4)$ lies on the surface of an unknown differentiable function f where $f_x(3, -1) = 2$ and $f_y(3, -1) = -1/2$. Find the equation of the tangent plane to f at P, and use this to approximate the value of $f(2.9, -0.8)$.

SOLUTION Knowing the partial derivatives at $(3, -1)$ allows us to form the normal vector to the tangent plane, $\vec{n} = \langle 2, -1/2, -1 \rangle$. Thus the equation of the tangent line to f at P is:

$$2(x-3) - 1/2(y+1) - (z-4) = 0 \quad \Rightarrow \quad z = 2(x-3) - 1/2(y+1) + 4. \quad \text{(13.5)}$$

Just as tangent lines provide excellent approximations of curves near their point of intersection, tangent planes provide excellent approximations of surfaces near their point of intersection. So $f(2.9, -0.8) \approx z(2.9, -0.8) = 3.7$.

This is not a new method of approximation. Compare the right hand expression for z in Equation (13.5) to the total differential:

$$dz = f_x dx + f_y dy \quad \text{and} \quad z = \underbrace{\underbrace{2}_{f_x} \underbrace{(x - 3)}_{dx} + \underbrace{-1/2}_{f_y} \underbrace{(y + 1)}_{dy}}_{dz} + 4.$$

Thus the "new z-value" is the sum of the change in z (i.e., dz) and the old z-value (4). As mentioned when studying the total differential, it is not uncommon to know partial derivative information about a unknown function, and tangent planes are used to give accurate approximations of the function.

The Gradient and Normal Lines, Tangent Planes

The methods developed in this section so far give a straightforward method of finding equations of normal lines and tangent planes for surfaces with explicit

Notes:

equations of the form $z = f(x, y)$. However, they do not handle implicit equations well, such as $x^2 + y^2 + z^2 = 1$. There is a technique that allows us to find vectors orthogonal to these surfaces based on the **gradient**.

Recall that when $z = f(x, y)$, the gradient $\nabla f = \langle f_x, f_y \rangle$ is orthogonal to level curves of f. Theorem 115 part 3 made an analogous statement about the gradient ∇F, where $w = F(x, y, z)$. Given a point (x_0, y_0, z_0), let $c = F(x_0, y_0, z_0)$. Then $F(x, y, z) = c$ is a **level surface** that contains the point (x_0, y_0, z_0) and $\nabla F(x_0, y_0, z_0)$ is orthogonal to this level surface. This direction can be used to find tangent planes and normal lines.

Example 8   Using the gradient to find a tangent plane

Find the equation of the plane tangent to the ellipsoid $\dfrac{x^2}{12} + \dfrac{y^2}{6} + \dfrac{z^2}{4} = 1$ at $P = (1, 2, 1)$.

SOLUTION   We consider the equation of the ellipsoid as a level surface of a function F of three variables, where $F(x, y, z) = \frac{x^2}{12} + \frac{y^2}{6} + \frac{z^2}{4}$. The gradient is:

$$\nabla F(x, y, z) = \langle F_x, F_y, F_z \rangle$$
$$= \left\langle \frac{x}{6}, \frac{y}{3}, \frac{z}{2} \right\rangle.$$

At P, the gradient is $\nabla F(1, 2, 1) = \langle 1/6, 2/3, 1/2 \rangle$. Thus the equation of the plane tangent to the ellipsoid at P is

$$\frac{1}{6}(x - 1) + \frac{2}{3}(y - 2) + \frac{1}{2}(z - 1) = 0.$$

The ellipsoid and tangent plane are graphed in Figure 13.26.

Figure 13.26: An ellipsoid and its tangent plane at a point.

Tangent lines and planes to surfaces have many uses, including the study of instantaneous rates of changes and making approximations. Normal lines also have many uses. In this section we focused on using them to measure distances from a surface. Another interesting application is in computer graphics, where the effects of light on a surface are determined using normal vectors.

The next section investigates another use of partial derivatives: determining relative extrema. When dealing with functions of the form $y = f(x)$, we found relative extrema by finding x where $f'(x) = 0$. We can *start* finding relative extrema of $z = f(x, y)$ by setting f_x and f_y to 0, but it turns out that there is more to consider.

Notes:

Exercises 13.7

Terms and Concepts

1. Explain how the vector $\vec{v} = \langle 1, 0, 3 \rangle$ can be thought of as having a "slope" of 3.

2. Explain how the vector $\vec{v} = \langle 0.6, 0.8, -2 \rangle$ can be thought of as having a "slope" of -2.

3. T/F: Let $z = f(x, y)$ be differentiable at P. If $\vec{n}$ is a normal vector to the tangent plane of f at P, then $\vec{n}$ is orthogonal to f_x and f_y at P.

4. Explain in your own words why we do not refer to *the* tangent line to a surface at a point, but rather to *directional* tangent line*s* to a surface at a point.

Problems

In Exercises 5–8, a function $z = f(x, y)$, a vector $\vec{v}$ and a point P are given. Give the parametric equations of the following directional tangent lines to f at P:

(a) $\ell_x(t)$

(b) $\ell_y(t)$

(c) $\ell_{\vec{u}}(t)$, where $\vec{u}$ is the unit vector in the direction of $\vec{v}$.

5. $f(x, y) = 2x^2y - 4xy^2$, $\vec{v} = \langle 1, 3 \rangle$, $P = (2, 3)$.

6. $f(x, y) = 3 \cos x \sin y$, $\vec{v} = \langle 1, 2 \rangle$, $P = (\pi/3, \pi/6)$.

7. $f(x, y) = 3x - 5y$, $\vec{v} = \langle 1, 1 \rangle$, $P = (4, 2)$.

8. $f(x, y) = x^2 - 2x - y^2 + 4y$, $\vec{v} = \langle 1, 1 \rangle$, $P = (1, 2)$.

In Exercises 9–12, a function $z = f(x, y)$ and a point P are given. Find the equation of the normal line to f at P. Note: these are the same functions as in Exercises 5 – 8.

9. $f(x, y) = 2x^2y - 4xy^2$, $P = (2, 3)$.

10. $f(x, y) = 3 \cos x \sin y$, $P = (\pi/3, \pi/6)$.

11. $f(x, y) = 3x - 5y$, $P = (4, 2)$.

12. $f(x, y) = x^2 - 2x - y^2 + 4y$, $P = (1, 2)$.

In Exercises 13–16, a function $z = f(x, y)$ and a point P are given. Find the two points that are 2 units from the surface f at P. Note: these are the same functions as in Exercises 5 – 8.

13. $f(x, y) = 2x^2y - 4xy^2$, $P = (2, 3)$.

14. $f(x, y) = 3 \cos x \sin y$, $P = (\pi/3, \pi/6)$.

15. $f(x, y) = 3x - 5y$, $P = (4, 2)$.

16. $f(x, y) = x^2 - 2x - y^2 + 4y$, $P = (1, 2)$.

In Exercises 17–20, a function $z = f(x, y)$ and a point P are given. Find the equation of the tangent plane to f at P. Note: these are the same functions as in Exercises 5 – 8.

17. $f(x, y) = 2x^2y - 4xy^2$, $P = (2, 3)$.

18. $f(x, y) = 3 \cos x \sin y$, $P = (\pi/3, \pi/6)$.

19. $f(x, y) = 3x - 5y$, $P = (4, 2)$.

20. $f(x, y) = x^2 - 2x - y^2 + 4y$, $P = (1, 2)$.

In Exercises 21–24, an implicitly defined function of x, y and z is given along with a point P that lies on the surface. Use the gradient ∇F to:

(a) find the equation of the normal line to the surface at P, and

(b) find the equation of the plane tangent to the surface at P.

21. $\dfrac{x^2}{8} + \dfrac{y^2}{4} + \dfrac{z^2}{16} = 1$, at $P = (1, \sqrt{2}, \sqrt{6})$

22. $z^2 - \dfrac{x^2}{4} - \dfrac{y^2}{9} = 0$, at $P = (4, -3, \sqrt{5})$

23. $xy^2 - xz^2 = 0$, at $P = (2, 1, -1)$

24. $\sin(xy) + \cos(yz) = 0$, at $P = (2, \pi/12, 4)$

13.8   Extreme Values

Given a function $z = f(x, y)$, we are often interested in points where z takes on the largest or smallest values. For instance, if z represents a cost function, we would likely want to know what (x, y) values minimize the cost. If z represents the ratio of a volume to surface area, we would likely want to know where z is greatest. This leads to the following definition.

Definition 99   Relative and Absolute Extrema

Let $z = f(x, y)$ be defined on a set S containing the point $P = (x_0, y_0)$.

1. If there is an open disk D containing P such that $f(x_0, y_0) \geq f(x, y)$ for all (x, y) in D, then f has a **relative maximum** at P; if $f(x_0, y_0) \leq f(x, y)$ for all (x, y) in D, then f has a **relative minimum** at P.

2. If $f(x_0, y_0) \geq f(x, y)$ for all (x, y) in S, then f has an **absolute maximum** at P; if $f(x_0, y_0) \leq f(x, y)$ for all (x, y) in S, then f has an **absolute minimum** at P.

3. If f has a relative maximum or minimum at P, then f has a **relative extrema** at P; if f has an absolute maximum or minimum at P, then f has an **absolute extrema** at P.

If f has a relative or absolute maximum at $P = (x_0, y_0)$, it means every curve on the surface of f through P will also have a relative or absolute maximum at P. Recalling what we learned in Section 3.1, the slopes of the tangent lines to these curves at P must be 0 or undefined. Since directional derivatives are computed using f_x and f_y, we are led to the following definition and theorem.

Definition 100   Critical Point

Let $z = f(x, y)$ be continuous on an open set S. A **critical point** $P = (x_0, y_0)$ of f is a point in S such that

- $f_x(x_0, y_0) = 0$ and $f_y(x_0, y_0) = 0$, or

- $f_x(x_0, y_0)$ or $f_y(x_0, y_0)$ is undefined.

Notes:

> **Theorem 116   Critical Points and Relative Extrema**
> Let $z = f(x, y)$ be defined on an open set S containing $P = (x_0, y_0)$. If f has a relative extrema at P, then P is a critical point of f.

Therefore, to find relative extrema, we find the critical points of f and determine which correspond to relative maxima, relative minima, or neither.

> Watch the video:
> Local Maximum and Minimum Values / Function of Two Variables at
> `https://youtu.be/Hm5QnuDjNmY`

The following examples demonstrate this process.

Example 1   Finding critical points and relative extrema
Let $f(x, y) = x^2 + y^2 - xy - x - 2$. Find the relative extrema of f.

SOLUTION   We start by computing the partial derivatives of f:

$$f_x(x, y) = 2x - y - 1 \qquad \text{and} \qquad f_y(x, y) = 2y - x.$$

Each is never undefined. A critical point occurs when f_x and f_y are simultaneously 0, leading us to solve the following system of linear equations:

$$2x - y - 1 = 0 \qquad \text{and} \qquad -x + 2y = 0.$$

This solution to this system is $x = 2/3$, $y = 1/3$. (Check that at $(2/3, 1/3)$, both f_x and f_y are 0.)

The graph in Figure 13.27 shows f along with this critical point. It is clear from the graph that this is a relative minimum; further consideration of the function shows that this is actually the absolute minimum.

Example 2   Finding critical points and relative extrema
Let $f(x, y) = -\sqrt{x^2 + y^2} + 2$. Find the relative extrema of f.
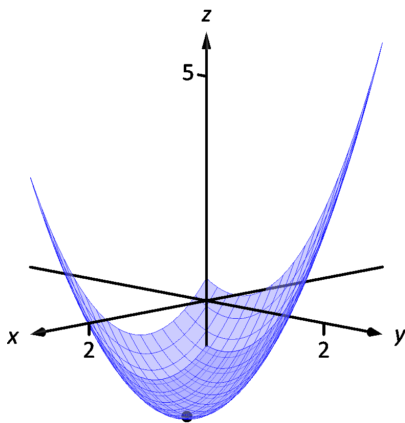
Figure 13.27: The surface in Example 1 with its absolute minimum indicated.

Notes:

SOLUTION   We start by computing the partial derivatives of f:

$$f_x(x, y) = \frac{-x}{\sqrt{x^2 + y^2}} \quad \text{and} \quad f_y(x, y) = \frac{-y}{\sqrt{x^2 + y^2}}.$$

It is clear that $f_x = 0$ when $x = 0$ & $y \neq 0$, and that $f_y = 0$ when $y = 0$ & $x \neq 0$. At $(0, 0)$, both f_x and f_y are *not* 0, but rather undefined. The point $(0, 0)$ is still a critical point, though, because the partial derivatives are undefined. This is the only critical point of f.

The surface of f is graphed in Figure 13.28 along with the point $(0, 0, 2)$. The graph shows that this point is the absolute maximum of f.

In each of the previous two examples, we found a critical point of f and then determined whether or not it was a relative (or absolute) maximum or minimum by graphing. It would be nice to be able to determine whether a critical point corresponded to a max or a min without a graph. Before we develop such a test, we do one more example that sheds more light on the issues our test needs to consider.

Example 3   Finding critical points and relative extrema

Let $f(x, y) = x^3 - 3x - y^2 + 4y$. Find the relative extrema of f.

SOLUTION   Once again we start by finding the partial derivatives of f:

$$f_x(x, y) = 3x^2 - 3 \quad \text{and} \quad f_y(x, y) = -2y + 4.$$

Each is always defined. Setting each equal to 0 and solving for x and y, we find

$$f_x(x, y) = 0 \quad \Rightarrow x = \pm 1$$
$$f_y(x, y) = 0 \quad \Rightarrow y = 2.$$

We have two critical points: $(-1, 2)$ and $(1, 2)$. To determine if they correspond to a relative maximum or minimum, we consider the graph of f in Figure 13.29.

The critical point $(-1, 2)$ clearly corresponds to a relative maximum. However, the critical point at $(1, 2)$ is neither a maximum nor a minimum, displaying a different, interesting characteristic.

If one walks parallel to the y-axis towards this critical point, then this point becomes a relative maximum along this path. But if one walks towards this point parallel to the x-axis, this point becomes a relative minimum along this path. A point that seems to act as both a max and a min is a **saddle point**. A formal definition follows.
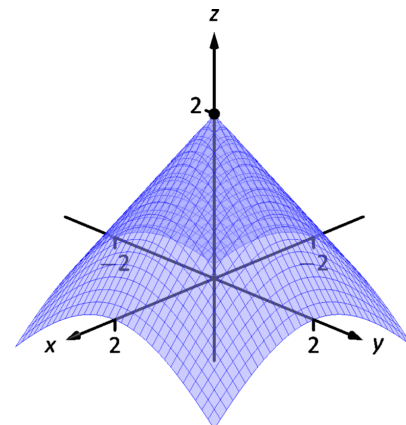
Figure 13.28: The surface in Example 2 with its absolute maximum indicated.
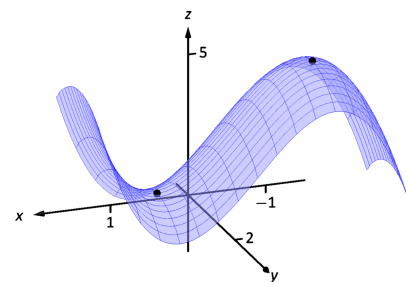
Figure 13.29: The surface in Example 3 with both critical points marked.

Notes:

> **Definition 101   Saddle Point**
> Let $P = (x_0, y_0)$ be in the domain of f where $f_x = 0$ and $f_y = 0$ at P. We say P is a **saddle point** of f if, for every open disk D containing P, there are points (x_1, y_1) and (x_2, y_2) in D such that $f(x_0, y_0) > f(x_1, y_1)$ and $f(x_0, y_0) < f(x_2, y_2)$.

At a saddle point, the instantaneous rate of change in all directions is 0 and there are points nearby with z-values both less than and greater than the z-value of the saddle point.

Before Example 3 we mentioned the need for a test to differentiate between relative maxima and minima. We now recognize that our test also needs to account for saddle points. To do so, we consider the second partial derivatives of f.

Recall that with single variable functions, such as $y = f(x)$, if $f''(c) > 0$, then f is concave up at c, and if $f'(c) = 0$, then f has a relative minimum at $x = c$. (We called this the Second Derivative Test.) Note that at a saddle point, it seems the graph is "both" concave up and concave down, depending on which direction you are considering.

It would be nice if the following were true:

$$
\begin{aligned}
f_{xx} \text{ and } f_{yy} > 0 \quad &\Rightarrow \quad \text{relative minimum} \\
f_{xx} \text{ and } f_{yy} < 0 \quad &\Rightarrow \quad \text{relative maximum} \\
f_{xx} \text{ and } f_{yy} \text{ have opposite signs} \quad &\Rightarrow \quad \text{saddle point.}
\end{aligned}
$$

However, this is not the case. Functions f exist where f_{xx} and f_{yy} are both positive but a saddle point still exists. In such a case, while the concavity in the x-direction is up (i.e., $f_{xx} > 0$) and the concavity in the y-direction is also up (i.e., $f_{yy} > 0$), the concavity switches somewhere in between the x- and y-directions.

To account for this, consider $D = f_{xx}f_{yy} - f_{xy}f_{yx}$. Since f_{xy} and f_{yx} are equal when continuous (refer back to Theorem 106), we can rewrite this as $D = f_{xx}f_{yy} - f_{xy}^2$. Then D can be used to test whether the concavity at a point changes depending on direction. If $D > 0$, the concavity does not switch (i.e., at that point, the graph is concave up or down in all directions). If $D < 0$, the concavity does switch. If $D = 0$, our test fails to determine whether concavity switches or not. We state the use of D in the following theorem.

Notes:

Theorem 117   Second Derivative Test

Let $z = f(x, y)$ be defined on an open set containing a critical point $P = (x_0, y_0)$ where all second order derivatives of f are continuous at P. Define

$$D = f_{xx}(x_0, y_0) f_{yy}(x_0, y_0) - f_{xy}^2(x_0, y_0).$$

1. If $D > 0$ and $f_{xx}(x_0, y_0) > 0$, then P is a relative minimum of f.

2. If $D > 0$ and $f_{xx}(x_0, y_0) < 0$, then P is a relative maximum of f.

3. If $D < 0$, then P is a saddle point of f.

4. If $D = 0$, the test is inconclusive.

Proof

Let $\vec{u} = \langle h, k \rangle$ be a unit vector. Then at the critical point P, $D_{\vec{u}} f = 0$. This means that along the line $h(y - y_0) = k(x - x_0)$, P is a critical point that is a maximum or minimum according to the sign of $D_{\vec{u}}^2 f$. Now,

$$\begin{aligned}
D_{\vec{u}}^2 f &= D_{\vec{u}}(f_x h + f_y k) \\
&= (f_x h)_x h + (f_x h)_y k + (f_y k)_x h + (f_y k) \\
&= f_{xx} h^2 + 2 f_{xy} hk + f_{yy} k^2
\end{aligned}$$

because f_{xy} and f_{yx} are continuous and therefore equal.

Suppose now that $D > 0$. Then we must have $f_{xx} \neq 0$, and we can complete the square to see that

$$D_{\vec{u}}^2 f = f_{xx} \left(h + \frac{f_{xy} k}{f_{xx}} \right)^2 + f_{yy} k^2 - \frac{f_{xy}^2 k^2}{f_{xx}} = f_{xx} \left[\left(h + \frac{f_{xy} k}{f_{xx}} \right)^2 + \frac{Dk^2}{f_{xx}^2} \right].$$

Because we assumed $D > 0$, everything in the brackets is positive, and $D_{\vec{u}}^2 f$ always has the same sign as f_{xx}. This shows parts 1 and 2.

If $D < 0$, our task is easier because we only need to find two different $\vec{u}$ that give $D_{\vec{u}}^2 f$ opposite signs. If $f_{xx} \neq 0$, let $\vec{v} = \langle -f_{xy}, f_{xx} \rangle$ and we can choose

$$\begin{aligned}
\vec{u}_1 &= \frac{1}{\|\vec{v}\|} \vec{v} \\
\vec{u}_2 &= \langle 1, 0 \rangle
\end{aligned} \quad \Rightarrow \quad \begin{aligned}
D_{\vec{u}_1}^2 f &= \frac{1}{\|\vec{v}\|^2} f_{xx} D \\
D_{\vec{u}_2}^2 f &= f_{xx}.
\end{aligned}$$

Notes:

Similarly, if $f_{yy} \neq 0$, let $\vec{v} = \langle f_{yy}, -f_{xy} \rangle$ and we can choose

$$\vec{u}_1 = \frac{1}{\|\vec{v}\|} \vec{v} \qquad \Rightarrow \qquad D^2_{\vec{u}_1} f = \frac{1}{\|\vec{v}\|^2} f_{yy} D$$
$$\vec{u}_2 = \langle 0, 1 \rangle \qquad\qquad D^2_{\vec{u}_2} f = f_{yy}.$$

Finally, if $f_{xx} = f_{yy} = 0$, then $D^2_{\vec{u}} f = 2f_{xy} hk$ has opposite signs for the vectors $\vec{u}_1 = \frac{1}{\sqrt{2}} \langle 1, 1 \rangle$ and $\vec{u}_2 = \frac{1}{\sqrt{2}} \langle 1, -1 \rangle$. $\qquad\square$

We first practice using this test with the function in the previous example, where we visually determined we had a relative maximum and a saddle point.

Example 4   **Using the Second Derivative Test**
Let $f(x, y) = x^3 - 3x - y^2 + 4y$ as in Example 3. Determine whether the function has a relative minimum, maximum, or saddle point at each critical point.

SOLUTION   We determined previously that the critical points of f are $(-1, 2)$ and $(1, 2)$. To use the Second Derivative Test, we must find the second partial derivatives of f:

$$f_{xx} = 6x; \qquad f_{yy} = -2; \qquad f_{xy} = 0.$$

Thus $D(x, y) = -12x$.

At $(-1, 2)$: $D(-1, 2) = 12 > 0$, and $f_{xx}(-1, 2) = -6$. By the Second Derivative Test, f has a relative maximum at $(-1, 2)$.

At $(1, 2)$: $D(1, 2) = -12 < 0$. The Second Derivative Test states that f has a saddle point at $(1, 2)$.

The Second Derivative Test confirmed what we determined visually.

Example 5   **Using the Second Derivative Test**
Find the relative extrema of $f(x, y) = x^2 y + y^2 + xy$.

SOLUTION   We start by finding the first and second partial derivatives of f:

$$f_x = 2xy + y \qquad\qquad f_y = x^2 + 2y + x$$
$$f_{xx} = 2y \qquad\qquad f_{yy} = 2$$
$$f_{xy} = 2x + 1 \qquad\qquad f_{yx} = 2x + 1.$$

We find the critical points by finding where f_x and f_y are simultaneously 0 (they are both never undefined). Setting $f_x = 0$, we have:

$$f_x = 0 \quad \Rightarrow \quad 2xy + y = 0 \quad \Rightarrow \quad y(2x + 1) = 0.$$

Notes:

This implies that for $f_x = 0$, either $y = 0$ or $2x + 1 = 0$.

Assume $y = 0$ then consider $f_y = 0$:

$$f_y = 0$$
$$x^2 + 2y + x = 0, \qquad \text{and since } y = 0, \text{ we have}$$
$$x^2 + x = 0$$
$$x(x + 1) = 0.$$

Thus if $y = 0$, we have either $x = 0$ or $x = -1$, giving two critical points: $(-1, 0)$ and $(0, 0)$.

Going back to f_x, now assume $2x + 1 = 0$, i.e., that $x = -1/2$, then consider $f_y = 0$:

$$f_y = 0$$
$$x^2 + 2y + x = 0, \qquad \text{and since } x = -1/2, \text{ we have}$$
$$1/4 + 2y - 1/2 = 0$$
$$y = 1/8.$$

Thus if $x = -1/2$, $y = 1/8$ giving the critical point $(-1/2, 1/8)$.

With $D = 4y - (2x + 1)^2$, we apply the Second Derivative Test to each critical point.

At $(-1, 0)$, $D < 0$, so $(-1, 0)$ is a saddle point.

At $(0, 0)$, $D < 0$, so $(0, 0)$ is also a saddle point.

At $(-1/2, 1/8)$, $D > 0$ and $f_{xx} > 0$, so $(-1/2, 1/8)$ is a relative minimum.

Figure 13.30 shows a graph of f and the three critical points. Note how this function does not vary much near the critical points – that is, visually it is difficult to determine whether a point is a saddle point or relative minimum (or even a critical point at all!). This is one reason why the Second Derivative Test is so important to have.

Constrained Optimization

When optimizing functions of one variable such as $y = f(x)$, we made use of Theorem 22, the Extreme Value Theorem, that said that over a closed interval I, a continuous function has both a maximum and minimum value. To find these maximum and minimum values, we evaluated f at all critical points in the interval, as well as at the endpoints (the "boundary") of the interval.

A similar theorem and procedure applies to functions of two variables. A continuous function over a closed set also attains a maximum and minimum value (see the following theorem). We can find these values by evaluating the function at the critical values in the set and over the boundary of the set. After formally stating this extreme value theorem, we give examples.
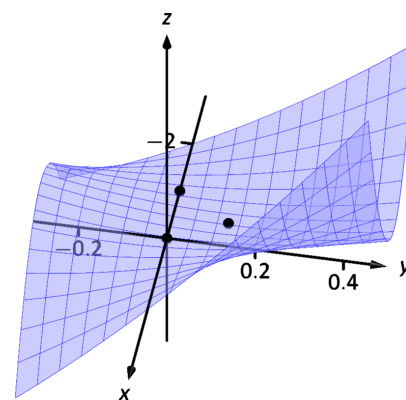
Figure 13.30: Graphing f from Example 5 and its relative extrema.

Notes:

> **Theorem 118**   **Extreme Value Theorem**
> Let $z = f(x, y)$ be a continuous function on a closed, bounded set S.
> Then f has a maximum and minimum value on S.

Example 6   **Finding extrema on a closed set**
Let $f(x, y) = x^2 - y^2 + 5$ and let S be the triangle with vertices $(-1, -2)$, $(0, 1)$ and $(2, -2)$. Find the maximum and minimum values of f on S.

    SOLUTION   It can help to see a graph of f along with the set S. In Figure 13.31(a) the triangle defining S is shown in the x-y plane in a dashed line. Above it is the surface of f; we are only concerned with the portion of f enclosed by the "triangle" on its surface.

    We begin by finding the critical points of f. With $f_x = 2x$ and $f_y = -2y$, we find only one critical point, at $(0, 0)$.

    We now find the maximum and minimum values that f attains along the boundary of S, that is, along the edges of the triangle. In Figure 13.31(b) we see the triangle sketched in the plane with the equations of the lines forming its edges labeled.

    Start with the bottom edge, along the line $y = -2$. If y is -2, then on the surface, we are considering points $f(x, -2)$; that is, our function reduces to $f(x, -2) = x^2 - (-2)^2 + 5 = x^2 + 1 = f_1(x)$. We want to maximize/minimize $f_1(x) = x^2 + 1$ on the interval $[-1, 2]$. To do so, we evaluate $f_1(x)$ at its critical points and at the endpoints. The critical points of f_1 are found by setting its derivative equal to 0:

$$f_1'(x) = 0 \quad \Rightarrow \quad x = 0,$$

so that we will need to evaluate f at the points $(-1, -2)$, $(0, -2)$, and $(2, -2)$.

    We need to do this process twice more, for the other two edges of the triangle.

    Along the left edge, along the line $y = 3x + 1$, we substitute $3x + 1$ in for y in $f(x, y)$:

$$f(x, y) = f(x, 3x + 1) = x^2 - (3x + 1)^2 + 5 = -8x^2 - 6x + 4 = f_2(x).$$

We want the maximum and minimum values of f_2 on the interval $[-1, 0]$, so we evaluate f_2 at its critical points and the endpoints of the interval. We find the critical points:

$$f_2'(x) = -16x - 6 = 0 \quad \Rightarrow \quad x = -3/8,$$

(a)

(b)
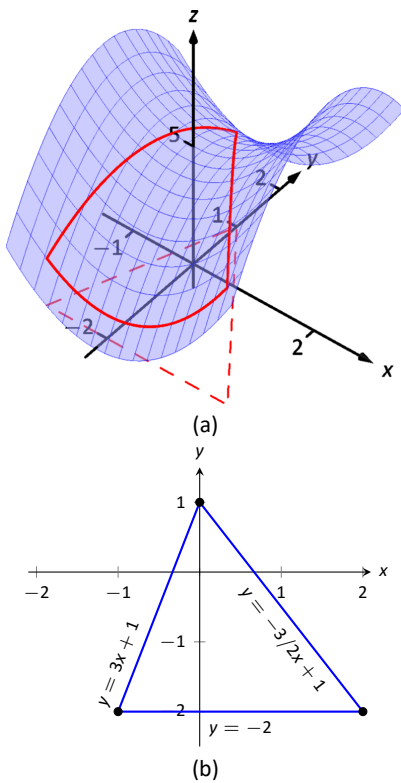
Figure 13.31: Plotting the surface of f along with the restricted domain S.

Notes:

so that we will need to evaluate f at the points $(-1, -2)$, $(-\frac{3}{8}, -\frac{1}{8})$, and $(0, 1)$.

Finally, we evaluate f along the right edge of the triangle, where $y = -3/2x + 1$.

$$f(x, y) = f(x, -3/2x + 1) = x^2 - (-3/2x + 1)^2 + 5 = -\frac{5}{4}x^2 + 3x + 4 = f_3(x).$$

The critical points of $f_3(x)$ are:

$$f'_3(x) = 0 \quad \Rightarrow \quad x = 6/5,$$

so that we will need to evaluate f at the points $(0, 1)$, $(\frac{6}{5}, -\frac{4}{5})$, and $(2, -2)$.

We now evaluate f at a total of 7 different places, all shown in Figure 13.32.

$$f(-1, -2) = 2, \quad f(0, -2) = 1, \quad f(2, -2) = 5,$$
$$f(-\frac{3}{8}, -\frac{1}{8}) = \frac{41}{8}, \quad f(0, 1) = 4, \quad \text{and} \quad f(\frac{6}{5}, -\frac{4}{5}) = \frac{29}{5}.$$

Of all the z-values found, the maximum is $\frac{29}{5}$, found at $(\frac{6}{5}, -\frac{4}{5})$; the minimum is 1, found at $(0, -2)$.

This portion of the text is entitled "Constrained Optimization" because we want to optimize a function (i.e., find its maximum and/or minimum values) subject to a *constraint* – some limit to what values the function can attain. In the previous example, we constrained ourselves by considering a function only within the boundary of a triangle. This was largely arbitrary; the function and the boundary were chosen just as an example, with no real "meaning" behind the function or the chosen constraint.

However, solving constrained optimization problems is a very important topic in applied mathematics. The techniques developed here are the basis for solving larger problems, where more than two variables are involved.

We illustrate the technique once more with a classic problem.

Example 7   **Constrained Optimization**
The U.S. Postal Service states that the girth+length of Standard Post Package must not exceed 130". Given a rectangular box, the "length" is the longest side, and the "girth" is twice the width+height.

Given a rectangular box where the width and height are equal, what are the dimensions of the box that give the maximum volume subject to the constraint of the size of a Standard Post Package?

Solution   Let w, h, and ℓ denote the width, height, and length of a rectangular box; we assume here that $w = h$. The girth is then $2(w + h) = 4w$. The volume of the box is $V(w, \ell) = wh\ell = w^2\ell$. We wish to maximize this
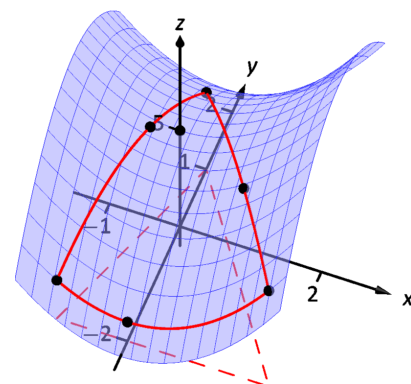
Figure 13.32: The surface of f along with important points along the boundary of S and the interior.

Notes:

volume subject to the constraint $4w + \ell \leq 130$, or $\ell \leq 130 - 4w$. (Common sense also indicates that $\ell > 0, w > 0$, so that we don't need to check the boundary where either is zero.)

We begin by finding the critical values of V. We find that $V_w = 2w\ell$ and $V_\ell = w^2$; these are simultaneously 0 only at $(0,0)$. This gives a volume of 0, so we can ignore this critical point.

We now consider the volume along the constraint $\ell = 130 - 4w$. Along this line, we have:

$$V(w, \ell) = V(w, 130 - 4w) = w^2(130 - 4w) = 130w^2 - 4w^3 = V_1(w).$$

The constraint is applicable on the w-interval $[0, 32.5]$ as indicated in the figure. Thus we want to maximize V_1 on $[0, 32.5]$.

Finding the critical values of V_1, we take the derivative and set it equal to 0:

$$V_1'(w) = 260w - 12w^2 = 0 \quad \Rightarrow \quad w(260 - 12w) = 0 \quad \Rightarrow \quad w = 0, \frac{260}{12} \approx 21.67.$$

We found two critical values: when $w = 0$ and when $w = 21.67$. We again ignore the $w = 0$ solution; the maximum volume, subject to the constraint, comes at $w = h = 21.67$, $\ell = 130 - 4(21.6) = 43.33$. This gives a volume of $V(21.67, 43.33) \approx 19{,}408\text{in}^3$.

The volume function $V(w, \ell)$ is shown in Figure 13.33 along with the constraint $\ell = 130 - 4w$. As done previously, the constraint is drawn dashed in the x-y plane and also along the surface of the function. The point where the volume is maximized is indicated.

It is hard to overemphasize the importance of optimization. In "the real world," we routinely seek to make *something* better. By expressing the *something* as a mathematical function, "making *something* better" means "optimize *some function.*"

The techniques shown here are only the beginning of an incredibly important field. Many functions that we seek to optimize are incredibly complex, making the step of "find the gradient and set it equal to $\vec{0}$" highly nontrivial. Mastery of the principles here are key to being able to tackle these more complicated problems.

Figure 13.33: Graphing the volume of a box with girth $4w$ and length ℓ, subject to a size constraint.

Notes:

Exercises 13.8

Terms and Concepts

1. T/F: Theorem 116 states that if f has a critical point at P, then f has a relative extrema at P.

2. T/F: A point P is a critical point of f if f_x and f_y are both 0 at P.

3. T/F: A point P is a critical point of f if f_x or f_y are undefined at P.

4. Explain what it means to "solve a constrained optimization" problem.

Problems

In Exercises 5–14, find the critical points of the given function. Use the Second Derivative Test to determine if each critical point corresponds to a relative maximum, minimum, or saddle point.

5. $f(x, y) = \frac{1}{2}x^2 + 2y^2 - 8y + 4x$

6. $f(x, y) = x^2 + 4x + y^2 - 9y + 3xy$

7. $f(x, y) = x^2 + 3y^2 - 6y + 4xy$

8. $f(x, y) = \dfrac{1}{x^2 + y^2 + 1}$

9. $f(x, y) = x^2 + y^3 - 3y + 1$

10. $f(x, y) = \frac{1}{3}x^3 - x + \frac{1}{3}y^3 - 4y$

11. $f(x, y) = x^2 y^2$

12. $f(x, y) = x^4 - 2x^2 + y^3 - 27y - 15$

13. $f(x, y) = \sqrt{16 - (x - 3)^2 - y^2}$

14. $f(x, y) = \sqrt{x^2 + y^2}$

In Exercises 15–18, find the absolute maximum and minimum of the function subject to the given constraint.

15. $f(x, y) = x^2 + y^2 + y + 1$, constrained to the triangle with vertices $(0, 1)$, $(-1, -1)$ and $(1, -1)$.

16. $f(x, y) = 5x - 7y$, constrained to the region bounded by $y = x^2$ and $y = 1$.

17. $f(x, y) = x^2 + 2x + y^2 + 2y$, constrained to the region bounded by the circle $x^2 + y^2 = 4$.

18. $f(x, y) = 3y - 2x^2$, constrained to the region bounded by the parabola $y = x^2 + x - 1$ and the line $y = x$.

13.9   Lagrange Multipliers

In the previous section, we were concerned with finding maxima and minima of functions without any constraints on the variables (other than being in the domain of the function). We ended by discussing what we would do if there were constraints on the variables. The following example illustrates a simple case of this type of problem.

Example 1   **Maximizing an Area**

For a rectangle whose perimeter is 20 m, find the dimensions that will maximize the area.

Solution   The area A of a rectangle with width x and height y is $A = xy$. The perimeter P of the rectangle is then given by the formula $P = 2x + 2y$. Since we are given that the perimeter $P = 20$, this problem can be stated as:

$$\text{Maximize } f(x, y) = xy \text{ subject to } 2x + 2y = 20$$

The reader is probably familiar with a simple method, using single-variable calculus, for solving this problem. Since we must have $2x + 2y = 20$, then we can solve for, say, y in terms of x using that equation. This gives $y = 10 - x$, which we then substitute into f to get $f(x, y) = xy = x(10 - x) = 10x - x^2$. This is now a function of x alone, so we now just have to maximize the function $f(x) = 10x - x^2$ on the interval $[0, 10]$. Since $f'(x) = 10 - 2x = 0 \Rightarrow x = 5$ and $f''(5) = -2 < 0$, then the Second Derivative Test tells us that $x = 5$ is a local maximum for f, and hence $x = 5$ must be the global maximum on the interval $[0, 10]$ (since $f = 0$ at the endpoints of the interval). So since $y = 10 - x = 5$, then the maximum area occurs for a rectangle whose width and height both are 5 m.

Note: Joseph Louis Lagrange (1736–1813) was a French mathematician and astronomer.

Notice in the above example that the ease of the solution depended on being able to solve for one variable in terms of the other in the equation $2x + 2y = 20$. But what if that were not possible (which is often the case)? In this section we will use a general method, called the *Lagrange multiplier method*, for solving *constrained optimization* problems:

$$\text{Maximize (or minimize) } f(x, y) \text{ subject to } g(x, y) = c$$

for some constant c. The equation $g(x, y) = c$ is called the *constraint equation*, and we say that x and y are *constrained* by $g(x, y) = c$. Points (x, y) which are maxima or minima of $f(x, y)$ with the condition that they satisfy the constraint equation $g(x, y) = c$ are called *constrained maximum* or *constrained minimum* points, respectively. Similar definitions hold for functions of three variables.

Notes:

The previous section optimized a function on a set *S*. In this section, "subject to $g(x, y) = c$" is the same as saying that the set *S* is given by $\{(x, y)|g(x, y) = c\}$. The Lagrange multiplier method for solving such problems can now be stated:

Theorem 119   Lagrange Multipliers

Let $f(x, y)$ and $g(x, y)$ be functions with continuous partial derivatives of all orders, and suppose that c is a scalar constant such that $\nabla g(x, y) \neq \vec{0}$ for all (x, y) that satisfy the equation $g(x, y) = c$. Then to solve the constrained optimization problem

$$\text{Maximize (or minimize) } f(x, y) \text{ subject to } g(x, y) = c,$$

find the points (x, y) that solve the equation $\nabla f(x, y) = \lambda \nabla g(x, y)$ for some constant λ (the number λ is called the *Lagrange multiplier*). If there is a constrained maximum or minimum, then it must be at such a point.

A rigorous proof of the above theorem is well beyond the scope of this text. Note that the theorem only gives a *necessary* condition for a point to be a constrained maximum or minimum. Whether a point (x, y) that satisfies $\nabla f(x, y) = \lambda \nabla g(x, y)$ for some λ actually *is* a constrained maximum or minimum can *sometimes* be determined by the nature of the problem itself. For instance, in Example 1 it was clear that there had to be a global maximum.

So how can you tell when a point that satisfies the condition in Theorem 119 really is a constrained maximum or minimum? The answer is that it depends on the constraint function $g(x, y)$, together with any implicit constraints. It can be shown that if the constraint equation $g(x, y) = c$ (plus any hidden constraints) describes a bounded set B in $\mathbb{R}^2$, then the constrained maximum or minimum of $f(x, y)$ will occur either at a point (x, y) satisfying $\nabla f(x, y) = \lambda \nabla g(x, y)$ or at a "boundary" point of the set B.

Watch the video:
LaGrange Multipliers at
`https://youtu.be/ry9cgNx1QV8`

In Example 1 the constraint equation $2x + 2y = 20$ describes a line in $\mathbb{R}^2$, which by itself is not bounded. However, there are "hidden" constraints, due to

---

Notes:

the nature of the problem, namely $0 \leq x, y \leq 10$, which cause that line to be restricted to a *line segment* in $\mathbb{R}^2$ (including the endpoints of that line segment), which *is* bounded.

Example 2   **Maximizing an Area**
For a rectangle whose perimeter is 20 m, use the Lagrange multiplier method to find the dimensions that will maximize the area.

   SOLUTION   As we saw in Example 1, with *x* and *y* representing the width and height, respectively, of the rectangle, this problem can be stated as:

$$\text{Maximize } f(x, y) = xy \text{ subject to } g(x, y) = 2x + 2y = 20.$$

Then solving the equation $\nabla f(x, y) = \lambda \nabla g(x, y)$ for some λ means solving the equations $\dfrac{\partial f}{\partial x} = \lambda \dfrac{\partial g}{\partial x}$ and $\dfrac{\partial f}{\partial y} = \lambda \dfrac{\partial g}{\partial y}$, namely:

$$y = 2\lambda,$$
$$x = 2\lambda$$

The general idea is to solve for λ in both equations, then set those expressions equal (since they both equal λ) to solve for *x* and *y*. Doing this we get

$$\frac{y}{2} = \lambda = \frac{x}{2} \quad \Rightarrow \quad x = y,$$

so now substitute either of the expressions for *x* or *y* into the constraint equation to solve for *x* and *y*:

$$20 = g(x, y) = 2x + 2y = 2x + 2x = 4x \quad \Rightarrow \quad x = 5 \quad \Rightarrow \quad y = 5$$

There must be a maximum area, since the minimum area is 0 and $f(5, 5) = 25 > 0$, so the point $(5, 5)$ that we found (called a *constrained critical point*) must be the constrained maximum. Therefore, the maximum area occurs for a rectangle whose width and height both are 5 m.

Example 3   **Extreme Values on a Circle**
Find the points on the circle $x^2 + y^2 = 80$ which are closest to and farthest from the point $(1, 2)$.

   SOLUTION   The distance *d* from any point (x, y) to the point $(1, 2)$ is

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

Notes:

and minimizing the distance is equivalent to minimizing the square of the distance. Thus the problem can be stated as:

Maximize (and minimize) $f(x, y) = (x-1)^2 + (y-2)^2$ subject to $g(x, y) = x^2 + y^2 = 80$.

Solving $\nabla f(x, y) = \lambda \nabla g(x, y)$ means solving the following equations:

$$2(x - 1) = 2\lambda x,$$
$$2(y - 2) = 2\lambda y$$

Note that $x \neq 0$ since otherwise we would get $-2 = 0$ in the first equation. Similarly, $y \neq 0$. So we can solve both equations for λ as follows:

$$\frac{x - 1}{x} = \lambda = \frac{y - 2}{y} \quad \Rightarrow \quad xy - y = xy - 2x \quad \Rightarrow \quad y = 2x$$

Substituting this into $g(x, y) = x^2 + y^2 = 80$ yields $5x^2 = 80$, so $x = \pm 4$. So the two constrained critical points are $(4, 8)$ and $(-4, -8)$. Since $f(4, 8) = 45$ and $f(-4, -8) = 125$, and since there must be points on the circle closest to and farthest from $(1, 2)$, then it must be the case that $(4, 8)$ is the point on the circle closest to $(1, 2)$ and $(-4, -8)$ is the farthest from $(1, 2)$ (see Figure 13.34).

Notice that since the constraint equation $x^2 + y^2 = 80$ describes a circle, which is a bounded set in $\mathbb{R}^2$, then we were guaranteed that the constrained critical points we found were indeed the constrained maximum and minimum.
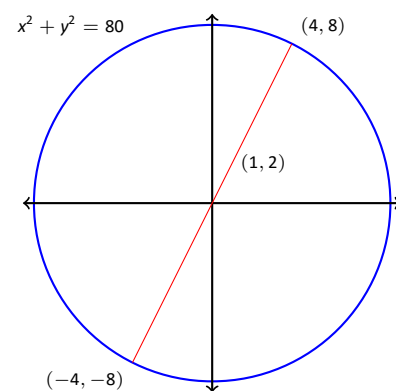
Figure 13.34: The circle in Example 3.

The Lagrange multiplier method can be extended to functions of three variables.

Example 4 **Maximizing a Function of Three Variables**
Maximize (and minimize) $f(x, y, z) = x + z$ subject to $g(x, y, z) = x^2 + y^2 + z^2 = 1$.

SOLUTION Solve the equation $\nabla f(x, y, z) = \lambda \nabla g(x, y, z)$:

$$1 = 2\lambda x$$
$$0 = 2\lambda y$$
$$1 = 2\lambda z$$

The first equation implies $\lambda \neq 0$ (otherwise we would have $1 = 0$), so we can divide by λ in the second equation to get $y = 0$ and we can divide by λ in the first and third equations to get $x = \frac{1}{2\lambda} = z$. Substituting these expressions into the constraint equation $g(x, y, z) = x^2 + y^2 + z^2 = 1$ yields the constrained critical points $\left(\frac{1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}}\right)$ and $\left(\frac{-1}{\sqrt{2}}, 0, \frac{-1}{\sqrt{2}}\right)$. Since $f\left(\frac{1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}}\right) > f\left(\frac{-1}{\sqrt{2}}, 0, \frac{-1}{\sqrt{2}}\right)$,

Notes:

and since the constraint equation $x^2 + y^2 + z^2 = 1$ describes a sphere (which is bounded) in $\mathbb{R}^3$, then $\left(\frac{1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}}\right)$ is the constrained maximum point and $\left(\frac{-1}{\sqrt{2}}, 0, \frac{-1}{\sqrt{2}}\right)$ is the constrained minimum point.

Two Constraints

When we have two constraints, we can still use Lagrange multipliers once we've made a slight modification. The optimization problem

Maximize (or minimize) $f(x, y, z)$ subject to $g(x, y, z) = c_1$ and $h(x, y, z) = c_2$

is satisfied when $\nabla f(x, y, z) = \lambda \nabla g(x, y, z) + \mu \nabla h(x, y, z)$.

Example 5   Optimizing with Two Constraints
The plane $x - y + z = 2$ intersects the cylinder $x^2 + y^2 = 4$ in an ellipse. Find the points on the ellipse closest to and farthest from the origin.

SOLUTION   We can optimize the distance $\sqrt{x^2 + y^2 + z^2}$ by optimizing the function $f(x, y, z) = x^2 + y^2 + z^2$, which has a simpler derivative. Let $g(x, y, z) = x - y + z$ be the plane constraint, and $h(x, y, z) = x^2 + y^2$ be the cylinder constraint. We see that

$$\nabla f(x, y, z) = \langle 2x, 2y, 2z \rangle$$
$$\nabla g(x, y, z) = \langle 1, -1, 1 \rangle$$
$$\nabla h(x, y, z) = \langle 2x, 2y, 0 \rangle .$$

The equation $\nabla f = \lambda \nabla g + \mu \nabla h$ means that

$$2x = \lambda + 2\mu x$$
$$2y = -\lambda + 2\mu y$$
$$2z = \lambda.$$

Adding the first two equations tells us that $x + y = \mu(x + y)$, so that $\mu = 1$ or $x = -y$. If $\mu = 1$, then $\lambda = z = 0$, and the constraint equations become

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

Substituting $x = y + 2$ into $x^2 + y^2 = 4$ tells us that $(y + 2)^2 + y^2 = 4$, which simplifies to $2y(y + 2) = 0$. This means that we need to look at the points

---

Notes:

$(2, 0, 0)$ and $(0, -2, 0)$, which are both distance 2 from the origin. If $x = -y$, then the constraint equations become

$$2x + z = 2$$
$$2x^2 = 4$$

and we need to look at the points $(\pm\sqrt{2}, \mp\sqrt{2}, 2 \mp 2\sqrt{2})$. These have distance $\sqrt{2 + 2 + (2 \mp 2\sqrt{2})^2} = \sqrt{16 \mp 8\sqrt{2}}$, which are both greater than 2. Therefore, the closest points are $(2, 0, 0)$ and $(0, -2, 0)$, while the furthest point is $(-\sqrt{2}, \sqrt{2}, 2 + \sqrt{2})$.

Finally, note that solving the equation $\nabla f(x, y) = \lambda \nabla g(x, y)$ means having to solve a system of two (possibly nonlinear) equations in three unknowns, which as we have seen before, may not be possible to do. And the 3-variable case can get even more complicated. All of this somewhat restricts the usefulness of Lagrange's method to relatively simple functions. Luckily there are many numerical methods for solving constrained optimization problems, though we will not discuss them here.

Notes:

Exercises 13.9

Problems

1. Find the constrained maxima and minima of $f(x, y) = 2x + y$ given that $x^2 + y^2 = 4$.
2. Find the constrained maxima and minima of $f(x, y) = xy$ given that $x^2 + 3y^2 = 6$.
3. Find the points on the circle $x^2 + y^2 = 100$ which are closest to and farthest from the point $(2, 3)$.
4. Find the constrained maxima and minima of $f(x, y, z) = x + y^2 + 2z$ given that $4x^2 + 9y^2 - 36z^2 = 36$.
5. Find the volume of the largest rectangular parallelepiped that can be inscribed in the ellipsoid

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1.$$

6. Find the minimum surface area of a box that holds 2 cubic meters.
7. The *girth* of a box is the perimeter of a cross section perpendicular to its length. The US post office will accept packages whose combined length and girth are at most 130 inches. Find the dimensions of the largest volume box that will be accepted.

8. Using Lagrange multipliers, find the shortest distance from the point (x_0, y_0, z_0) to the plane $ax + by + cz = d$. (See also Key Idea 54.)
9. Find all points on the surface $xz - y^2 + 1 = 0$ that are closest to the origin.
10. Find the three positive numbers whose sum is 60 and whose product is as large as possible.
11. Find all points on the plane $x + y + z = 5$ in the first octant at which $f(x, y, z) = x^2 y z^2$ has a maximum value.
12. Find the points on the surface $z^2 - xy = 5$ that are closest to the origin.
13. Find the maximum and minimum points of $f(x, y) = xy + \sqrt{9 - x^2 - y^2}$ when $x^2 + y^2 \leq 9$.
14. Find three real numbers whose sum is 12 and the sum of whose squares is a small as possible.
15. Find the maximum volume of a rectangular box inscribed in the unit sphere.
16. The plane $x + y - z = 1$ intersects the cylinder $x^2 + y^2 = 1$ in an ellipse. Find the points on the ellipse closest to and farthest from the origin.

14: Multiple Integration

The previous chapter introduced multivariable functions and we applied concepts of differential calculus to these functions. We learned how we can view a function of two variables as a surface in space, and learned how partial derivatives convey information about how the surface is changing in any direction.

In this chapter we apply techniques of integral calculus to multivariable functions. In Chapter 5 we learned how the definite integral of a single variable function gave us "area under the curve." In this chapter we will see that integration applied to a multivariable function gives us "volume under a surface." And just as we learned applications of integration beyond finding areas, we will find applications of integration in this chapter beyond finding volume.

14.1 Iterated Integrals and Area

In Chapter 13 we found that it was useful to differentiate functions of several variables with respect to one variable, while treating all the other variables as constants or coefficients. We can integrate functions of several variables in a similar way. For instance, if we are told that $f_x(x, y) = 2xy$, we can treat y as staying constant and integrate to obtain $f(x, y)$:

$$f(x, y) = \int f_x(x, y)\, dx$$

$$= \int 2xy\, dx$$

$$= x^2 y + C.$$

Make a careful note about the constant of integration, C. This "constant" is something with a derivative of 0 with respect to x, so it could be any expression that contains only constants and functions of y. For instance, if $f(x, y) = x^2 y + \sin y + y^3 + 17$, then $f_x(x, y) = 2xy$. To signify that C is actually a function of y, we write:

$$f(x, y) = \int f_x(x, y)\, dx = x^2 y + C(y).$$

Using this process we can even evaluate definite integrals.

Example 1   **Integrating functions of more than one variable**

Evaluate the integral $\int_1^{2y} 2xy\, dx$.

SOLUTION   We find the indefinite integral as before, then apply the Fundamental Theorem of Calculus to evaluate the definite integral:

$$\int_1^{2y} 2xy \, dx = x^2y \Big|_1^{2y}$$
$$= (2y)^2y - (1)^2y$$
$$= 4y^3 - y.$$

We can also integrate with respect to y. In general,

$$\int_{h_1(y)}^{h_2(y)} f_x(x, y) \, dx = f(x, y) \Big|_{h_1(y)}^{h_2(y)} = f\big(h_2(y), y\big) - f\big(h_1(y), y\big),$$

and

$$\int_{g_1(x)}^{g_2(x)} f_y(x, y) \, dy = f(x, y) \Big|_{g_1(x)}^{g_2(x)} = f\big(x, g_2(x)\big) - f\big(x, g_1(x)\big).$$

Note that when integrating by x, the bounds do not depend on x, and the result is no longer a function of x. When integrating by y, the bounds do not depend on y, and the result is no longer a function of y. Another example will help us understand this.

Example 2   **Integrating functions of more than one variable**

Evaluate $\int_1^x \left(5x^3y^{-3} + 6y^2\right) dy$.

SOLUTION   We consider x as staying constant and integrate with respect to y:

$$\int_1^x \left(5x^3y^{-3} + 6y^2\right) dy = \left(\frac{5x^3y^{-2}}{-2} + \frac{6y^3}{3}\right)\Big|_1^x$$
$$= \left(-\frac{5}{2}x^3x^{-2} + 2x^3\right) - \left(-\frac{5}{2}x^3 + 2\right)$$
$$= \frac{9}{2}x^3 - \frac{5}{2}x - 2.$$

Note how the bounds of the integral are from $y = 1$ to $y = x$ and that the final answer is a function of x.

In the previous example, we integrated a function with respect to y and ended up with a function of x. We can integrate this as well. This process is known as **iterated integration**, or **multiple integration.**

Notes:

Watch the video:
Double Integrals — Basic Idea and Examples at
`https://youtu.be/DYsv6L-VcsQ`

Example 3 **Integrating an integral**

Evaluate $\displaystyle\int_1^2 \left(\int_1^x \left(5x^3 y^{-3} + 6y^2 \right) dy \right) dx$.

SOLUTION We follow a standard "order of operations" and perform the operations inside parentheses first (which is the integral evaluated in Example 2.)

$$\int_1^2 \left(\int_1^x \left(5x^3 y^{-3} + 6y^2 \right) dy \right) dx = \int_1^2 \left(\left[\frac{5x^3 y^{-2}}{-2} + \frac{6y^3}{3} \right] \Big|_1^x \right) dx$$

$$= \int_1^2 \left(\frac{9}{2}x^3 - \frac{5}{2}x - 2 \right) dx$$

$$= \left(\frac{9}{8}x^4 - \frac{5}{4}x^2 - 2x \right) \Big|_1^2$$

$$= \frac{89}{8}.$$

Note how the bounds of x were $x = 1$ to $x = 2$ and the final result was a number.

The previous example showed how we could perform something called an iterated integral; we do not yet know *why* we would be interested in doing so nor what the result, such as the number $89/8$, *means*. Before we investigate these questions, we offer some definitions.

Notes:

> **Definition 102   Iterated Integration**
>
> **Iterated integration** is the process of repeatedly integrating the results of previous integrations. Integrating one integral is denoted as follows.
>
> Let a, b, c and d be numbers and let $g_1(x)$, $g_2(x)$, $h_1(y)$ and $h_2(y)$ be functions of x and y, respectively. Then:
>
> 1. $\displaystyle \int_c^d \int_{h_1(y)}^{h_2(y)} f(x,y)\, dx\, dy = \int_c^d \left(\int_{h_1(y)}^{h_2(y)} f(x,y)\, dx \right) dy.$
>
> 2. $\displaystyle \int_a^b \int_{g_1(x)}^{g_2(x)} f(x,y)\, dy\, dx = \int_a^b \left(\int_{g_1(x)}^{g_2(x)} f(x,y)\, dy \right) dx.$

Again make note of the bounds of these iterated integrals.

With $\displaystyle \int_c^d \int_{h_1(y)}^{h_2(y)} f(x,y)\, dx\, dy$, x varies from $h_1(y)$ to $h_2(y)$, whereas y varies from c to d. That is, the bounds of x are *curves*, the curves $x = h_1(y)$ and $x = h_2(y)$, whereas the bounds of y are *constants*, $y = c$ and $y = d$. It is useful to remember that after integrating with respect to a variable, that variable is no longer present.

We now begin to investigate *why* we are interested in iterated integrals and *what* they mean.

Area of a plane region

Consider the plane region R bounded by $a \leq x \leq b$ and $g_1(x) \leq y \leq g_2(x)$, shown in Figure 14.1. We learned in Section 6.1 that the area of R is given by

$$\int_a^b \left(g_2(x) - g_1(x) \right) dx.$$

We can view the expression $\left(g_2(x) - g_1(x) \right)$ as

$$\left(g_2(x) - g_1(x) \right) = \int_{g_1(x)}^{g_2(x)} 1\, dy = \int_{g_1(x)}^{g_2(x)} dy,$$

meaning we can express the area of R as an iterated integral:

$$\text{area of } R = \int_a^b \left(g_2(x) - g_1(x) \right) dx = \int_a^b \left(\int_{g_1(x)}^{g_2(x)} dy \right) dx = \int_a^b \int_{g_1(x)}^{g_2(x)} dy\, dx.$$
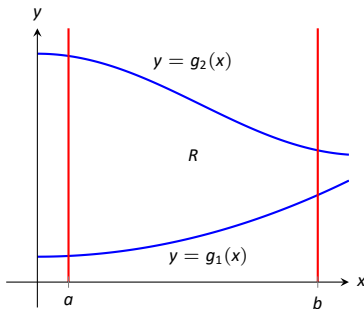
Figure 14.1: Calculating the area of a plane region R with an iterated integral.

Notes:

In short: a certain iterated integral can be viewed as giving the area of a plane region.

A region R could also be defined by $c \leq y \leq d$ and $h_1(y) \leq x \leq h_2(y)$, as shown in Figure 14.2. Using a process similar to that above, we have

$$\text{the area of } R = \int_c^d \int_{h_1(y)}^{h_2(y)} dx\, dy.$$

We state this formally in a theorem.

Theorem 120   Area of a plane region

1. Let R be a plane region bounded by $a \leq x \leq b$ and $g_1(x) \leq y \leq g_2(x)$, where g_1 and g_2 are continuous functions on $[a, b]$. The area A of R is

$$A = \int_a^b \int_{g_1(x)}^{g_2(x)} dy\, dx.$$

2. Let R be a plane region bounded by $c \leq y \leq d$ and $h_1(y) \leq x \leq h_2(y)$, where h_1 and h_2 are continuous functions on $[c, d]$. The area A of R is

$$A = \int_c^d \int_{h_1(y)}^{h_2(y)} dx\, dy.$$
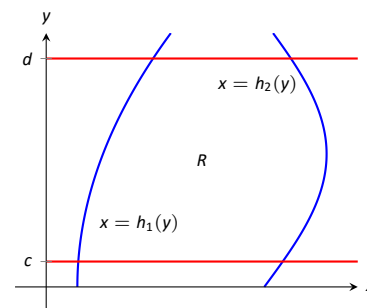
Figure 14.2: Calculating the area of a plane region R with an iterated integral.

The following examples should help us understand this theorem.

Example 4   Area of a rectangle
Find the area A of the rectangle with corners $(-1, 1)$ and $(3, 3)$, as shown in Figure 14.3.

SOLUTION   Multiple integration is obviously overkill in this situation, but we proceed to establish its use.

The region R is bounded by $x = -1$, $x = 3$, $y = 1$ and $y = 3$. Choosing to integrate with respect to y first, we have

$$A = \int_{-1}^3 \int_1^3 1\, dy\, dx = \int_{-1}^3 \left(y \Big|_1^3 \right) dx = \int_{-1}^3 2\, dx = 2x \Big|_{-1}^3 = 8.$$

We could also integrate with respect to x first, giving:

$$A = \int_1^3 \int_{-1}^3 1\, dx\, dy = \int_1^3 \left(x \Big|_{-1}^3 \right) dy = \int_1^3 4\, dy = 4y \Big|_1^3 = 8.$$
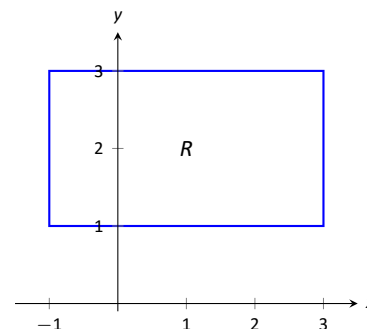
Figure 14.3: Calculating the area of a rectangle with an iterated integral in Example 4.

Notes:

Clearly there are simpler ways to find this area, but it is interesting to note that this method works.

Example 5   Area of a triangle

Find the area A of the triangle with vertices at $(1,1)$, $(3,1)$ and $(5,5)$, as shown in Figure 14.4.

SOLUTION   The triangle is bounded by the lines as shown in the figure. Choosing to integrate with respect to x first gives that x is bounded by $x = y$ to $x = \frac{y+5}{2}$, while y is bounded by $y = 1$ to $y = 5$. (Recall that since x-values increase from left to right, the leftmost curve, $x = y$, is the lower bound and the rightmost curve, $x = (y+5)/2$, is the upper bound.) The area is

$$A = \int_1^5 \int_y^{\frac{y+5}{2}} dx\, dy$$

$$= \int_1^5 \left(x \Big|_y^{\frac{y+5}{2}} \right) dy$$

$$= \int_1^5 \left(-\frac{1}{2}y + \frac{5}{2} \right) dy$$

$$= \left(-\frac{1}{4}y^2 + \frac{5}{2}y \right) \Big|_1^5$$

$$= 4.$$

We can also find the area by integrating with respect to y first. In this situation, though, we have two functions that act as the lower bound for the region R, $y = 1$ and $y = 2x - 5$. This requires us to use two iterated integrals. Note how the x-bounds are different for each integral:

$$
\begin{aligned}
A &= \int_1^3 \int_1^x 1\, dy\, dx &&+ &&\int_3^5 \int_{2x-5}^x 1\, dy\, dx \\
&= \int_1^3 (y)\Big|_1^x dx &&+ &&\int_3^5 (y)\Big|_{2x-5}^x dx \\
&= \int_1^3 (x-1)\, dx &&+ &&\int_3^5 (-x+5)\, dx \\
&= 2 &&+ &&2 \\
&= 4.
\end{aligned}
$$

As expected, we get the same answer both ways. This equality will also be justified by Theorem 122 in the next section.

Figure 14.4: Calculating the area of a triangle with iterated integrals in Example 5.

Notes:

Example 6 **Area of a plane region**

Find the area of the region enclosed by $y = 2x$ and $y = x^2$, as shown in Figure 14.5.

SOLUTION Once again we'll find the area of the region using both orders of integration.

Using $dy\,dx$:

$$\int_0^2 \int_{x^2}^{2x} 1\,dy\,dx = \int_0^2 (2x - x^2)\,dx = \left(x^2 - \frac{1}{3}x^3\right)\Big|_0^2 = \frac{4}{3}.$$

Using $dx\,dy$:

$$\int_0^4 \int_{y/2}^{\sqrt{y}} 1\,dx\,dy = \int_0^4 (\sqrt{y} - y/2)\,dy = \left(\frac{2}{3}y^{3/2} - \frac{1}{4}y^2\right)\Big|_0^4 = \frac{4}{3}.$$
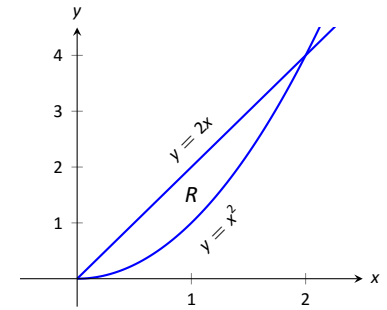
Figure 14.5: Calculating the area of a plane region with iterated integrals in Example 6.

Changing Order of Integration

In each of the previous examples, we have been given a region R and found the bounds needed to find the area of R using both orders of integration. We integrated using both orders of integration to demonstrate their equality.

We now approach the skill of describing a region using both orders of integration from a different perspective. Instead of starting with a region and creating iterated integrals, we will start with an iterated integral and rewrite it in the other integration order. To do so, we'll need to understand the region over which we are integrating.

The simplest of all cases is when both integrals are bound by constants. The region described by these bounds is a rectangle (see Example 4), and so:

$$\int_a^b \int_c^d 1\,dy\,dx = \int_c^d \int_a^b 1\,dx\,dy.$$

When the inner integral's bounds are not constants, it is generally very useful to sketch the bounds to determine what the region we are integrating over looks like. From the sketch we can then rewrite the integral with the other order of integration.

Examples will help us develop this skill.

Example 7 **Changing the order of integration**

Rewrite the iterated integral $\displaystyle\int_0^6 \int_0^{x/3} 1\,dy\,dx$ with the order of integration $dx\,dy$.

Notes:

857
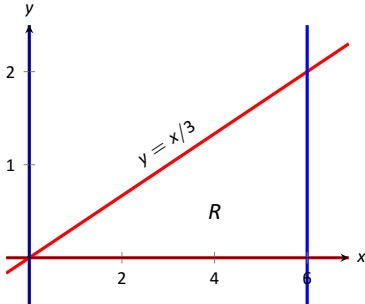
Figure 14.6: Sketching the region R described by the iterated integral in Example 7.
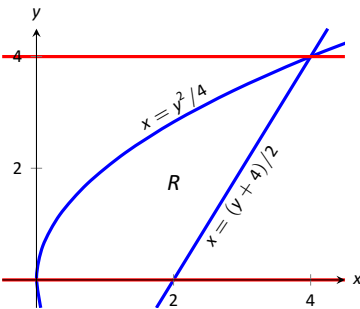
Figure 14.7: Drawing the region determined by the bounds of integration in Example 8.

SOLUTION   We need to use the bounds of integration to determine the region we are integrating over.

The bounds tell us that y is bounded by 0 and $x/3$; x is bounded by 0 and 6. We plot these four curves: $y = 0$, $y = x/3$, $x = 0$ and $x = 6$ to find the region described by the bounds. Figure 14.6 shows these curves, indicating that R is a triangle.

To change the order of integration, we need to consider the curves that bound the x-values. We see that the lower bound is $x = 3y$ and the upper bound is $x = 6$. The bounds on y are 0 to 2. Thus we can rewrite the integral as

$$\int_0^2 \int_{3y}^6 1 \, dx \, dy.$$

Example 8   **Changing the order of integration**

Change the order of integration of $\int_0^4 \int_{y^2/4}^{(y+4)/2} 1 \, dx \, dy$.

SOLUTION   We sketch the region described by the bounds to help us change the integration order. We see x is bounded below and above (i.e., to the left and right) by $x = y^2/4$ and $x = (y + 4)/2$ respectively, and y is bounded between 0 and 4. Graphing the previous curves, we find the region R to be that shown in Figure 14.7.

To change the order of integration, we need to establish curves that bound y. The figure makes it clear that there are two lower bounds for y: $y = 0$ on $0 \le x \le 2$, and $y = 2x - 4$ on $2 \le x \le 4$. Thus we need two double integrals. The upper bound for each is $y = 2\sqrt{x}$. Thus we have

$$\int_0^4 \int_{y^2/4}^{(y+4)/2} 1 \, dx \, dy = \int_0^2 \int_0^{2\sqrt{x}} 1 \, dy \, dx + \int_2^4 \int_{2x-4}^{2\sqrt{x}} 1 \, dy \, dx.$$

This section has introduced a new concept, the iterated integral. We developed one application for iterated integration: area between curves. However, this is not new, for we already know how to find areas bounded by curves.

In the next section we apply iterated integration to solve problems we currently do not know how to handle. The "real" goal of this section was not to learn a new way of computing area. Rather, our goal was to learn how to define a region in the plane using the bounds of an iterated integral. That skill is very important in the following sections.

Notes:

Exercises 14.1

Terms and Concepts

1. When integrating $f_x(x, y)$ with respect to x, the constant of integration C is really which: $C(x)$ or $C(y)$? What does this mean?

2. Integrating an integral is called _____ _____.

3. When evaluating an iterated integral, we integrate from _____ to _____, then from _____ to _____.

4. One understanding of an iterated integral is that
$$\int_a^b \int_{g_1(x)}^{g_2(x)} dy\, dx$$ gives the _____ of a plane region.

Problems

In Exercises 5–10, evaluate the integral and subsequent iterated integral.

5.

(a) $\displaystyle \int_2^5 \left(6x^2 + 4xy - 3y^2\right) dy$

(b) $\displaystyle \int_{-3}^{-2} \int_2^5 \left(6x^2 + 4xy - 3y^2\right) dy\, dx$

6.

(a) $\displaystyle \int_0^\pi \left(2x \cos y + \sin x\right) dx$

(b) $\displaystyle \int_0^{\pi/2} \int_0^\pi \left(2x \cos y + \sin x\right) dx\, dy$

7.

(a) $\displaystyle \int_1^x \left(x^2 y - y + 2\right) dy$

(b) $\displaystyle \int_0^2 \int_1^x \left(x^2 y - y + 2\right) dy\, dx$

8.

(a) $\displaystyle \int_y^{y^2} (x - y)\, dx$

(b) $\displaystyle \int_{-1}^1 \int_y^{y^2} (x - y)\, dx\, dy$

9.

(a) $\displaystyle \int_0^y \left(\cos x \sin y\right) dx$

(b) $\displaystyle \int_0^\pi \int_0^y \left(\cos x \sin y\right) dx\, dy$

10.

(a) $\displaystyle \int_0^x \left(\frac{1}{1 + x^2}\right) dy$

(b) $\displaystyle \int_1^2 \int_0^x \left(\frac{1}{1 + x^2}\right) dy\, dx$

In Exercises 11–16, a graph of a planar region R is given. Give the iterated integrals, with both orders of integration $dy\, dx$ and $dx\, dy$, that give the area of R. Evaluate one of the iterated integrals to find the area.

11.

12.

13.

14.
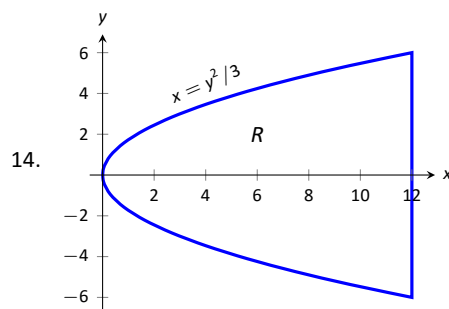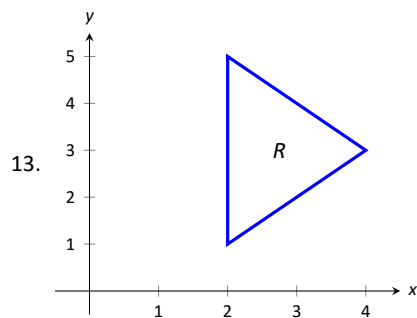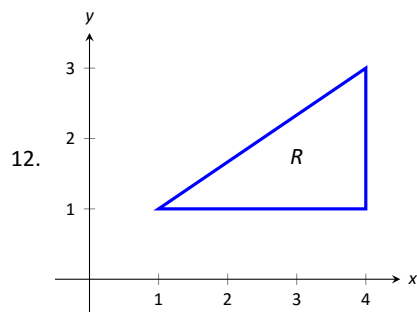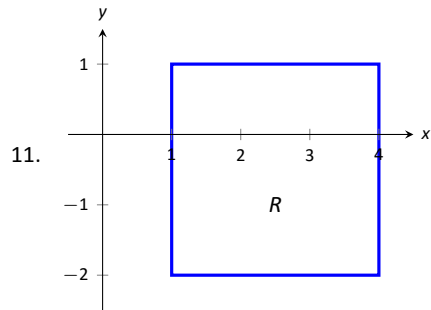
15.
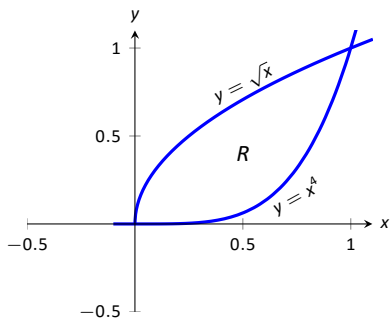
16.
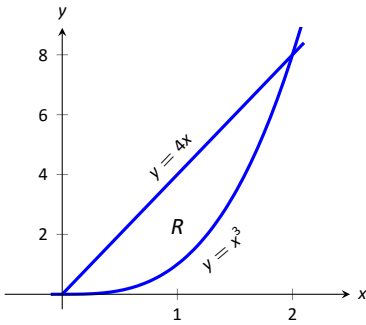
In Exercises 17–22, iterated integrals are given that compute the area of a region R in the x-y plane. Sketch the region R, and give the iterated integral(s) that give the area of R with the opposite order of integration.

17. $\displaystyle\int_{-2}^{2}\int_{0}^{4-x^2} dy\, dx$

18. $\displaystyle\int_{0}^{1}\int_{5-5x}^{5-5x^2} dy\, dx$

19. $\displaystyle\int_{-2}^{2}\int_{0}^{2\sqrt{4-y^2}} dx\, dy$

20. $\displaystyle\int_{-3}^{3}\int_{-\sqrt{9-x^2}}^{\sqrt{9-x^2}} dy\, dx$

21. $\displaystyle\int_{0}^{1}\int_{-\sqrt{y}}^{\sqrt{y}} dx\, dy + \int_{1}^{4}\int_{y-2}^{\sqrt{y}} dx\, dy$

22. $\displaystyle\int_{-1}^{1}\int_{(x-1)/2}^{(1-x)/2} dy\, dx$

14.2 Double Integration and Volume

The definite integral of f over $[a, b]$, $\int_a^b f(x)\ dx$, was introduced as "the signed area under the curve." We approximated the value of this area by first subdividing $[a, b]$ into n subintervals, where the ith subinterval has length Δx_i, and letting c_i be any value in the ith subinterval. We formed rectangles that approximated part of the region under the curve with width Δx_i, height $f(c_i)$, and hence with area $f(c_i)\Delta x_i$. Summing all the rectangle's areas gave an approximation of the definite integral, and Theorem 36 stated that

$$\int_a^b f(x)\ dx = \lim_{\|\Delta x\| \to 0} \sum f(c_i)\Delta x_i,$$

connecting the area under the curve with sums of the areas of rectangles.

We use a similar approach in this section to find volume under a surface.

Let R be a closed, bounded region in the x-y plane and let $z = f(x, y)$ be a continuous function defined on R. We wish to find the signed volume under the surface of f over R. (We use the term "signed volume" to denote that space above the x-y plane, under f, will have a positive volume; space above f and under the x-y plane will have a "negative" volume, similar to the notion of signed area used before.)

We start by partitioning R into n rectangular subregions as shown in Figure 14.8(a). For simplicity's sake, we let all widths be Δx and all heights be Δy. Note that the sum of the areas of the rectangles is not equal to the area of R, but rather is a close approximation. Arbitrarily number the rectangles 1 through n, and pick a point (x_i, y_i) in the ith subregion.

The volume of the rectangular solid whose base is the ith subregion and whose height is $f(x_i, y_i)$ is $V_i = f(x_i, y_i)\Delta x\Delta y$. Such a solid is shown in Figure 14.8(b). Note how this rectangular solid only approximates the true volume under the surface; part of the solid is above the surface and part is below.

For each subregion R_i used to approximate R, create the rectangular solid with base area $\Delta x\Delta y$ and height $f(x_i, y_i)$. The sum of all rectangular solids is

$$\sum_{i=1}^{n} f(x_i, y_i)\Delta x\Delta y.$$

This approximates the signed volume under f over R. As we have done before, to get a better approximation we can use more rectangles to approximate the region R.

In general, each rectangle could have a different width Δx_j and height Δy_k, giving the ith rectangle an area $\Delta A_i = \Delta x_j \Delta y_k$ and the ith rectangular solid a

Notes:

(a)

(b)
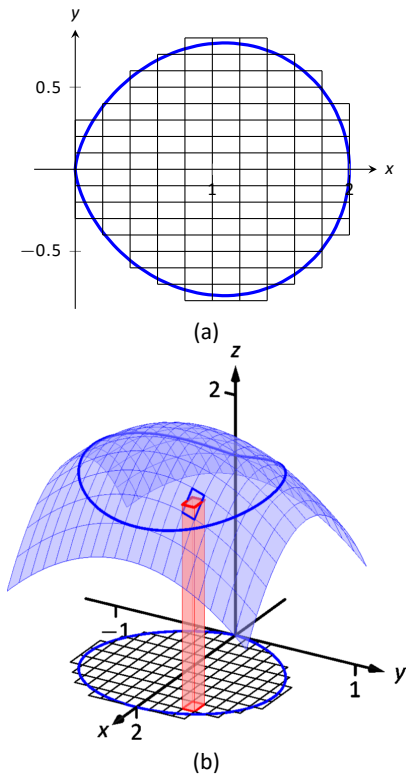
Figure 14.8: Developing a method for finding signed volume under a surface.

volume of $f(x_i, y_i)\Delta A_i$. Let $\|\Delta A\|$ denote the length of the longest diagonal of all rectangles in the subdivision of R; $\|\Delta A\| \to 0$ means each rectangle's width and height are both approaching 0. If f is a continuous function, as $\|\Delta A\|$ shrinks (and hence $n \to \infty$) the summation $\sum_{i=1}^{n} f(x_i, y_i)\Delta A_i$ approximates the signed volume better and better. This leads to a definition.

Note: Recall that the integration symbol "$\int$" is an "elongated S," representing the word "sum." We interpreted $\int_a^b f(x)\, dx$ as "take the *sum* of the areas of rectangles over the interval $[a, b]$." The double integral uses two integration symbols to represent a "double sum." When adding up the volumes of rectangular solids over a partition of a region R, as done in Figure 14.8, one could first add up the volumes across each row (one type of sum), then add these totals together (another sum), as in

$$\sum_{j=1}^{n} \sum_{i=1}^{m} f(x_i, y_j)\Delta x_i \Delta y_j.$$

One can rewrite this as

$$\sum_{j=1}^{n} \left(\sum_{i=1}^{m} f(x_i, y_j)\Delta x_i \right) \Delta y_j.$$

The summation inside the parenthesis indicates the sum of heights $\times$ widths, which gives an area; multiplying these areas by the thickness Δy_j gives a volume. The illustration in Figure 14.9 relates to this understanding.

> **Definition 103   Double Integral, Signed Volume**
>
> Let $z = f(x, y)$ be a continuous function defined over a closed region R in the *x-y* plane. The **signed volume** V under f over R is denoted by the **double integral**
>
> $$V = \iint_R f(x, y)\, dA.$$
>
> Alternate notations for the double integral are
>
> $$\iint_R f(x, y)\, dA = \iint_R f(x, y)\, dx\, dy = \iint_R f(x, y)\, dy\, dx.$$

The definition above does not state how to find the signed volume, though the notation offers a hint. We need the next two theorems to evaluate double integrals to find volume.

> **Theorem 121   Double Integrals and Signed Volume**
>
> Let $z = f(x, y)$ be a continuous function defined over a closed region R in the *x-y* plane. Then the signed volume V under f over R is
>
> $$V = \iint_R f(x, y)\, dA = \lim_{\|\Delta A\| \to 0} \sum_{i=1}^{n} f(x_i, y_i)\Delta A_i.$$

This theorem states that we can find the exact signed volume using a limit of sums. The partition of the region R is not specified, so any partitioning where the diagonal of each rectangle shrinks to 0 results in the same answer.

This does not offer a very satisfying way of computing volume, though. Our experience has shown that evaluating the limits of sums can be tedious. We seek a more direct method.

Recall Theorem 45 in Section 6.2. This stated that if $A(x)$ gives the cross-sectional area of a solid at x, then $\int_a^b A(x)\, dx$ gave the volume of that solid over $[a, b]$.

Notes:

Consider Figure 14.9, where a surface $z = f(x, y)$ is drawn over a region R. Fixing a particular x value, we can consider the area under f over R where x has that fixed value. That area can be found with a definite integral, namely

$$A(x) = \int_{g_1(x)}^{g_2(x)} f(x, y) \, dy.$$

Remember that though the integrand contains x, we are viewing x as fixed. Also note that the bounds of integration are functions of x: the bounds depend on the value of x.

As $A(x)$ is a cross-sectional area function, we can find the signed volume V under f by integrating it:

$$V = \int_a^b A(x) \, dx = \int_a^b \left(\int_{g_1(x)}^{g_2(x)} f(x, y) \, dy \right) dx = \int_a^b \int_{g_1(x)}^{g_2(x)} f(x, y) \, dy \, dx.$$

This gives a concrete method for finding signed volume under a surface. We could do a similar procedure where we started with y fixed, resulting in a iterated integral with the order of integration $dx \, dy$. The following theorem states that both methods give the same result, which is the value of the double integral. It is such an important theorem it has a name associated with it.

Figure 14.9: Finding volume under a surface by sweeping out a cross–sectional area.

Theorem 122   Fubini's Theorem

Let R be a closed, bounded region in the x-y plane and let $z = f(x, y)$ be a continuous function on R.

1. If R is bounded by $a \leq x \leq b$ and $g_1(x) \leq y \leq g_2(x)$, where g_1 and g_2 are continuous functions on $[a, b]$, then

$$\iint_R f(x, y) \, dA = \int_a^b \int_{g_1(x)}^{g_2(x)} f(x, y) \, dy \, dx.$$

2. If R is bounded by $c \leq y \leq d$ and $h_1(y) \leq x \leq h_2(y)$, where h_1 and h_2 are continuous functions on $[c, d]$, then

$$\iint_R f(x, y) \, dA = \int_c^d \int_{h_1(y)}^{h_2(y)} f(x, y) \, dx \, dy.$$

Note that the bounds of integration follow a "curve to curve, point to point" pattern. In fact, one of the main points of the previous section is developing the skill of describing a region R with the bounds of an iterated integral. Once this

---

Notes:

skill is developed, we can use double integrals to compute many quantities, not just signed volume under a surface.

<table>
<tr><td>Watch the video:
Ex: Evaluate a Double Integral to Determine Volume (Basic) at
<code>https://youtu.be/NG2UcXdwzfk</code></td></tr>
</table>

Example 1 **Evaluating a double integral**

Let $f(x, y) = xy + e^y$. Find the signed volume under f on the region R, which is the rectangle with corners $(3, 1)$ and $(4, 2)$ pictured in Figure 14.10, using Fubini's Theorem and both orders of integration.

SOLUTION We wish to evaluate $\iint_R \left(xy + e^y \right) dA$. As R is a rectangle, the bounds are easily described as $3 \leq x \leq 4$ and $1 \leq y \leq 2$.

Using the order *dy dx*:

$$\iint_R \left(xy + e^y \right) dA = \int_3^4 \int_1^2 \left(xy + e^y \right) dy\, dx$$

$$= \int_3^4 \left(\left[\frac{1}{2}xy^2 + e^y \right] \Big|_1^2 \right) dx$$

$$= \int_3^4 \left(\frac{3}{2}x + e^2 - e \right) dx$$

$$= \left(\frac{3}{4}x^2 + (e^2 - e)x \right) \Big|_3^4$$

$$= \frac{21}{4} + e^2 - e.$$

Figure 14.10: Finding the signed volume under a surface in Example 1.

Notes:

Now we check the validity of Fubini's Theorem by using the order $dx\,dy$:

$$\iint_R (xy + e^y)\,dA = \int_1^2 \int_3^4 (xy + e^y)\,dx\,dy$$

$$= \int_1^2 \left(\left[\frac{1}{2}x^2y + xe^y\right]\Big|_3^4 \right) dy$$

$$= \int_1^2 \left(\frac{7}{2}y + e^y\right) dy$$

$$= \left(\frac{7}{4}y^2 + e^y\right)\Big|_1^2$$

$$= \frac{21}{4} + e^2 - e.$$

Both orders of integration return the same result, as expected.

Example 2   Evaluating a double integral
Evaluate $\iint_R \left(3xy - x^2 - y^2 + 6\right) dA$, where R is the triangle bounded by $x = 0$, $y = 0$ and $x/2 + y = 1$, as shown in Figure 14.11.

SOLUTION   While it is not specified which order we are to use, we will evaluate the double integral using both orders to help drive home the point that it does not matter which order we use.
Using the order $dy\,dx$: The bounds on y go from "curve to curve," i.e., $0 \leq y \leq 1 - x/2$, and the bounds on x go from "point to point," i.e., $0 \leq x \leq 2$.

$$\iint_R (3xy - x^2 - y^2 + 6)\,dA = \int_0^2 \int_0^{-\frac{x}{2}+1} (3xy - x^2 - y^2 + 6)\,dy\,dx$$

$$= \int_0^2 \left(\frac{3}{2}xy^2 - x^2y - \frac{1}{3}y^3 + 6y\right)\Big|_0^{-\frac{x}{2}+1} dx$$

$$= \int_0^2 \left(\frac{11}{12}x^3 - \frac{11}{4}x^2 - x + \frac{17}{3}\right) dx$$

$$= \left(\frac{11}{48}x^4 - \frac{11}{12}x^3 - \frac{1}{2}x^2 + \frac{17}{3}x\right)\Big|_0^2$$

$$= \frac{17}{3} = 5.\overline{6}.$$

Now lets consider the order $dx\,dy$. Here x goes from "curve to curve," $0 \leq$
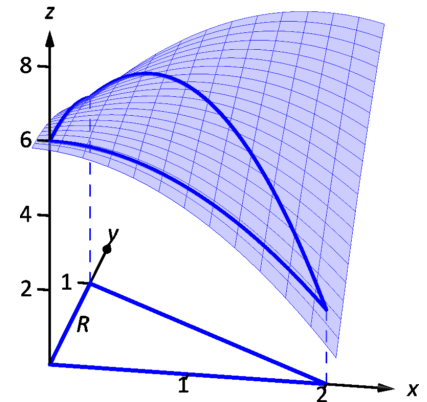
Figure 14.11: Finding the signed volume under the surface in Example 2.

Notes:

$x \leq 2 - 2y$, and y goes from "point to point," $0 \leq y \leq 1$:

$$\iint_R \left(3xy - x^2 - y^2 + 6\right) dA = \int_0^1 \int_0^{2-2y} \left(3xy - x^2 - y^2 + 6\right) dx\, dy$$

$$= \int_0^1 \left(\frac{3}{2}x^2y - \frac{1}{3}x^3 - xy^2 + 6x\right)\Bigg|_0^{2-2y} dy$$

$$= \int_0^1 \left(\frac{32}{3}y^3 - 22y^2 + 2y + \frac{28}{3}\right) dy$$

$$= \left(\frac{8}{3}y^4 - \frac{22}{3}y^3 + y^2 + \frac{28}{3}y\right)\Bigg|_0^1$$

$$= \frac{17}{3} = 5.\overline{6}.$$

We obtained the same result using both orders of integration.

Note how in these two examples that the bounds of integration depend only on R; the bounds of integration have nothing to do with $f(x, y)$. This is an important concept, so we include it as a Key Idea.

> **Key Idea 60**   **Double Integration Bounds**
> When evaluating $\iint_R f(x, y)\, dA$ using an iterated integral, the bounds of integration depend only on R. The surface f does not determine the bounds of integration.

Before doing another example, we give some properties of double integrals. Each should make sense if we view them in the context of finding signed volume under a surface, over a region.

Notes:

Theorem 123   Properties of Double Integrals

Let f and g be continuous functions over a closed, bounded plane region R, and let c be a constant.

1. $\displaystyle\iint_R cf(x, y)\, dA = c \iint_R f(x, y)\, dA.$

2. $\displaystyle\iint_R \big(f(x, y) \pm g(x, y)\big)\, dA = \iint_R f(x, y)\, dA \pm \iint_R g(x, y)\, dA$

3. If $f(x, y) \geq 0$ on R, then $\displaystyle\iint_R f(x, y)\, dA \geq 0.$

4. If $f(x, y) \geq g(x, y)$ on R, then $\displaystyle\iint_R f(x, y)\, dA \geq \iint_R g(x, y)\, dA.$

5. Let R be the union of two nonoverlapping regions, $R = R_1 \bigcup R_2$ (see Figure 14.12). Then
$$\iint_R f(x, y)\, dA = \iint_{R_1} f(x, y)\, dA + \iint_{R_2} f(x, y)\, dA.$$
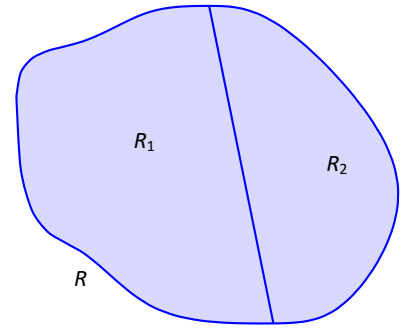
Figure 14.12: R is the union of two nonoverlapping regions, R_1 and R_2.

Example 3   Evaluating a double integral

Let $f(x, y) = \sin x \cos y$ and R be the triangle with vertices $(-1, 0)$, $(1, 0)$ and $(0, 1)$ (see Figure 14.13). Evaluate the double integral $\iint_R f(x, y)\, dA$.

Solution   If we attempt to integrate using an iterated integral with the order $dy\, dx$, note how there are two upper bounds on R meaning we'll need to use two iterated integrals. We would need to split the triangle into two regions along the y-axis, then use Theorem 123, part 5.

Instead, let's use the order $dx\, dy$. The curves bounding x are $y - 1 \leq x \leq 1 - y$; the bounds on y are $0 \leq y \leq 1$. This gives us:

$$
\begin{aligned}
\iint_R f(x, y)\, dA &= \int_0^1 \int_{y-1}^{1-y} \sin x \cos y\, dx\, dy \\
&= \int_0^1 \left(-\cos x \cos y \right)\Big|_{y-1}^{1-y} dy \\
&= \int_0^1 \cos y \Big(-\cos(1 - y) + \cos(y - 1) \Big) dy.
\end{aligned}
$$

Recall that the cosine function is an even function; that is, $\cos x = \cos(-x)$. Therefore, from the last integral above, we have $\cos(y - 1) = \cos(1 - y)$. Thus
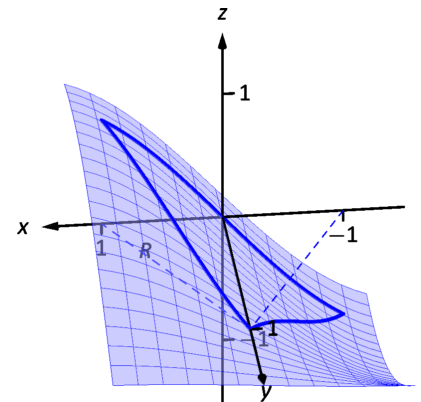
Figure 14.13: Finding the signed volume under a surface in Example 3.

Notes:

the integrand simplifies to 0, and we have

$$\iint_R f(x,y)\,dA = \int_0^1 0\,dy$$
$$= 0.$$

It turns out that over R, there is just as much volume above the x-y plane as below (look again at Figure 14.13), giving a final signed volume of 0.

Example 4 **Evaluating a double integral**
Evaluate $\iint_R (4-y)\,dA$, where R is the region bounded by the parabolas $y^2 = 4x$ and $x^2 = 4y$, graphed in Figure 14.14.

SOLUTION Graphing each curve can help us find their points of intersection. Solving analytically, the second equation tells us that $y = x^2/4$. Substituting this value in for y in the first equation gives us $x^4/16 = 4x$. Solving for x:

$$\frac{x^4}{16} = 4x$$
$$x^4 - 64x = 0$$
$$x(x^3 - 64) = 0$$
$$x = 0,\ 4.$$

Thus we've found analytically what was easy to approximate graphically: the regions intersect at $(0,0)$ and $(4,4)$, as shown in Figure 14.14.

We now choose an order of integration: $dy\,dx$ or $dx\,dy$? Either order works; since the integrand does not contain x, choosing $dx\,dy$ might be simpler – at least, the first integral is very simple.

Thus we have the following "curve to curve, point to point" bounds: $y^2/4 \le x \le 2\sqrt{y}$, and $0 \le y \le 4$.

$$\iint_R (4-y)\,dA = \int_0^4 \int_{y^2/4}^{2\sqrt{y}} (4-y)\,dx\,dy$$
$$= \int_0^4 \left(x(4-y)\right)\Big|_{y^2/4}^{2\sqrt{y}}\,dy$$
$$= \int_0^4 \left((2\sqrt{y} - \frac{y^2}{4})(4-y)\right)\,dy = \int_0^4 \left(\frac{y^3}{4} - y^2 - 2y^{3/2} + 8y^{1/2}\right)\,dy$$
$$= \left(\frac{y^4}{16} - \frac{y^3}{3} - \frac{4y^{5/2}}{5} + \frac{16y^{3/2}}{3}\right)\Big|_0^4$$
$$= \frac{176}{15} = 11.7\overline{3}.$$

Notes:

Figure 14.14: Finding the volume under the surface in Example 4.

The signed volume under the surface f is about 11.7 cubic units.

In the previous section we practiced changing the order of integration of a given iterated integral, where the region R was not explicitly given. Changing the bounds of an integral is more than just an test of understanding. Rather, there are cases where integrating in one order is really hard, if not impossible, whereas integrating with the other order is feasible.

Example 5   Changing the order of integration

Rewrite the iterated integral $\int_0^3 \int_y^3 e^{-x^2}\, dx\, dy$ with the order $dy\, dx$. Comment on the feasibility to evaluate each integral.

SOLUTION   Once again we make a sketch of the region over which we are integrating to facilitate changing the order. The bounds on x are from $x = y$ to $x = 3$; the bounds on y are from $y = 0$ to $y = 3$. These curves are sketched in Figure 14.15, enclosing the region R.

To change the bounds, note that the curves bounding y are $y = 0$ up to $y = x$; the triangle is enclosed between $x = 0$ and $x = 3$. Thus the new bounds of integration are $0 \le y \le x$ and $0 \le x \le 3$, giving the iterated integral $\int_0^3 \int_0^x e^{-x^2}\, dy\, dx$.

How easy is it to evaluate each iterated integral? Consider the order of integrating $dx\, dy$, as given in the original problem. The first indefinite integral we need to evaluate is $\int e^{-x^2}\, dx$; we have stated before (see Section 8.7) that this integral cannot be evaluated in terms of elementary functions. We are stuck.

Changing the order of integration makes a big difference here. In the second iterated integral, we are faced with $\int e^{-x^2}\, dy$; integrating with respect to y gives us $ye^{-x^2} + C$, and the first definite integral evaluates to

$$\int_0^x e^{-x^2}\, dy = xe^{-x^2}.$$

Thus

$$\int_0^3 \int_0^x e^{-x^2}\, dy\, dx = \int_0^3 \left(xe^{-x^2}\right) dx.$$

This last integral is easy to evaluate with substitution, giving a final answer of $\frac{1}{2}(1 - e^{-9})$. Figure 14.16 shows the surface over R.

In short, evaluating one iterated integral is impossible; the other iterated integral is relatively simple.
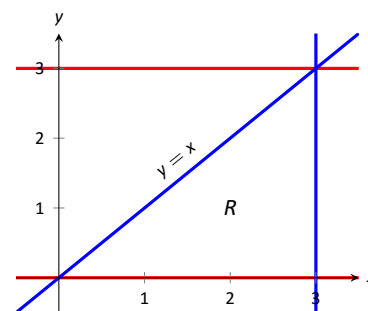
Figure 14.15: Determining the region R determined by the bounds of integration in Example 5.
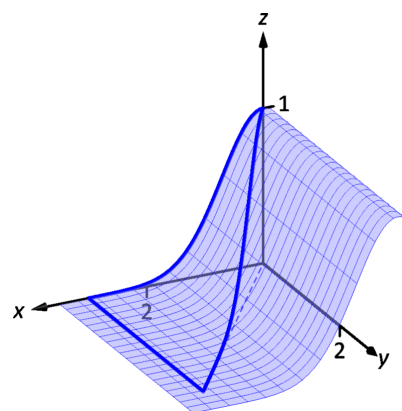
Figure 14.16: Showing the surface f defined in Example 5 over its region R.

Notes:

Definition 25 defines the average value of a single–variable function $f(x)$ on the interval $[a, b]$ as

$$\text{average value of } f(x) \text{ on } [a, b] = \frac{1}{b-a} \int_a^b f(x) \, dx;$$

that is, it is the "area under f over an interval divided by the length of the interval." We make an analogous statement here: the average value of $z = f(x, y)$ over a region R is the volume under f over R divided by the area of R.

Definition 104   The Average Value of f on R

Let $z = f(x, y)$ be a continuous function defined over a closed region R in the x-y plane. The **average value of f on R** is

$$\text{average value of } f \text{ on } R = \frac{\displaystyle\iint_R f(x, y) \, dA}{\displaystyle\iint_R dA}.$$

Example 6   Finding average value of a function over a region R

Find the average value of $f(x, y) = 4 - y$ over the region R, which is bounded by the parabolas $y^2 = 4x$ and $x^2 = 4y$. Note: this is the same function and region as used in Example 4.

SOLUTION   In Example 4 we found

$$\iint_R f(x, y) \, dA = \int_0^4 \int_{y^2/4}^{2\sqrt{y}} (4 - y) \, dx \, dy = \frac{176}{15}.$$

We find the area of R by computing $\iint_R dA$:

$$\iint_R dA = \int_0^4 \int_{y^2/4}^{2\sqrt{y}} dx \, dy = \frac{16}{3}.$$

Dividing the volume under the surface by the area gives the average value:

$$\text{average value of } f \text{ on } R = \frac{176/15}{16/3} = \frac{11}{5} = 2.2.$$

While the surface, as shown in Figure 14.17, covers z-values from $z = 0$ to $z = 4$, the "average" z-value on R is 2.2.
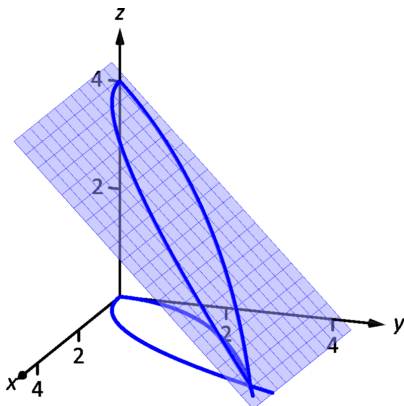
Notes:

Figure 14.17: Finding the average value of f in Example 6.

The previous section introduced the iterated integral in the context of finding the area of plane regions. This section has extended our understanding of iterated integrals; now we see they can be used to find the signed volume under a surface.

This new understanding allows us to revisit what we did in the previous section. Given a region R in the plane, we computed $\iint_R 1\, dA$; again, our understanding at the time was that we were finding the area of R. However, we can now view the function $z = 1$ as a surface, a flat surface with constant z-value of 1. The double integral $\iint_R 1\, dA$ finds the volume, under $z = 1$, over R, as shown in Figure 14.18. Basic geometry tells us that if the base of a general right cylinder has area A, its volume is $A \cdot h$, where h is the height. In our case, the height is 1. We were "actually" computing the volume of a solid, though we interpreted the number as an area.

The next section extends our abilities to find "volumes under surfaces." Currently, some integrals are hard to compute because either the region R we are integrating over is hard to define with rectangular curves, or the integrand itself is hard to deal with. Some of these problems can be solved by converting everything into polar coordinates.
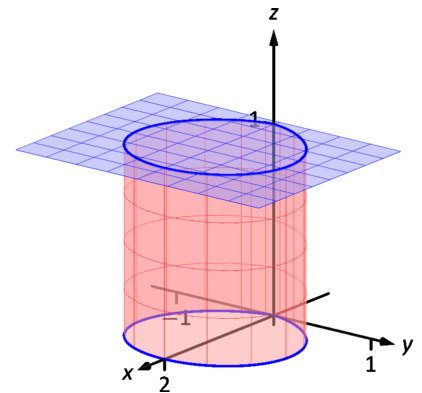
Figure 14.18: Showing how an iterated integral used to find area also finds a certain volume.

Notes:

Exercises 14.2

Terms and Concepts

1. An integral can be interpreted as giving the signed area over an interval; a double integral can be interpreted as giving the signed _____ over a region.

2. Explain why the following statement is false: "Fubini's Theorem states that $\int_a^b \int_{g_1(x)}^{g_2(x)} f(x, y) \; dy \; dx = \int_a^b \int_{g_1(y)}^{g_2(y)} f(x, y) \; dx \; dy$."

3. Explain why if $f(x, y) > 0$ over a region R, then $\iint_R f(x, y) \; dA > 0$.

4. If $\iint_R f(x, y) \; dA = \iint_R g(x, y) \; dA$, does this imply $f(x, y) = g(x, y)$?

Problems

In Exercises 5–10,

(a) Evaluate the given iterated integral, and

(b) rewrite the integral using the other order of integration.

5. $\int_1^2 \int_{-1}^1 \left(\frac{x}{y} + 3 \right) \; dx \; dy$

6. $\int_{-\pi/2}^{\pi/2} \int_0^\pi (\sin x \cos y) \; dx \; dy$

7. $\int_0^4 \int_0^{-x/2+2} \left(3x^2 - y + 2 \right) \; dy \; dx$

8. $\int_1^3 \int_y^3 \left(x^2 y - xy^2 \right) \; dx \; dy$

9. $\int_0^1 \int_{-\sqrt{1-y}}^{\sqrt{1-y}} (x + y + 2) \; dx \; dy$

10. $\int_0^9 \int_{y/3}^{\sqrt{y}} \left(xy^2 \right) \; dx \; dy$

In Exercises 11–18:

(a) Sketch the region R given by the problem.

(b) Set up the iterated integrals, in both orders, that evaluate the given double integral for the described region R.

(c) Evaluate one of the iterated integrals to find the signed volume under the surface $z = f(x, y)$ over the region R.

11. $\iint_R x^2 y \; dA$, where R is bounded by $y = \sqrt{x}$ and $y = x^2$.

12. $\iint_R x^2 y \; dA$, where R is bounded by $y = \sqrt[3]{x}$ and $y = x^3$.

13. $\iint_R x^2 - y^2 \; dA$, where R is the rectangle with corners $(-1, -1)$, $(1, -1)$, $(1, 1)$ and $(-1, 1)$.

14. $\iint_R ye^x \; dA$, where R is bounded by $x = 0$, $x = y^2$ and $y = 1$.

15. $\iint_R \left(6 - 3x - 2y \right) \; dA$, where R is bounded by $x = 0$, $y = 0$ and $3x + 2y = 6$.

16. $\iint_R e^y \; dA$, where R is bounded by $y = \ln x$ and $y = \frac{1}{e-1}(x-1)$.

17. $\iint_R \left(x^3 y - x \right) \; dA$, where R is the half disk $x^2 + y^2 \le 9$ in the first and second quadrants.

18. $\iint_R \left(4 - 3y \right) \; dA$, where R is bounded by $y = 0$, $y = x/e$ and $y = \ln x$.

In Exercises 19–22, state why it is difficult/impossible to integrate the iterated integral in the given order of integration. Change the order of integration and evaluate the new iterated integral.

19. $\int_0^4 \int_{y/2}^2 e^{x^2} \; dx \; dy$

20. $\int_0^{\sqrt{\pi/2}} \int_x^{\sqrt{\pi/2}} \cos \left(y^2 \right) \; dy \; dx$

21. $\int_0^1 \int_y^1 \frac{2y}{x^2 + y^2} \; dx \; dy$

22. $\int_{-1}^1 \int_1^2 \frac{x \tan^2 y}{1 + \ln y} \; dy \; dx$

In Exercises 23–26, find the average value of f over the region R. Notice how these functions and regions are related to the iterated integrals given in Exercises 5 – 8.

23. $f(x, y) = \frac{x}{y} + 3$; R is the rectangle with opposite corners $(-1, 1)$ and $(1, 2)$.

24. $f(x, y) = \sin x \cos y$; R is bounded by $x = 0$, $x = \pi$, $y = -\pi/2$ and $y = \pi/2$.

25. $f(x, y) = 3x^2 - y + 2$; R is bounded by the lines $y = 0$, $y = 2 - x/2$ and $x = 0$.

26. $f(x, y) = x^2 y - xy^2$; R is bounded by $y = x$, $y = 1$ and $x = 3$.

14.3 Double Integration with Polar Coordinates

We have used iterated integrals to evaluate double integrals, which give the signed volume under a surface, $z = f(x, y)$, over a region R of the x-y plane. The integrand is simply $f(x, y)$, and the bounds of the integrals are determined by the region R.

Some regions R are easy to describe using rectangular coordinates – that is, with equations of the form $y = f(x)$, $x = a$, etc. However, some regions are easier to handle if we represent their boundaries with polar equations of the form $r = f(\theta)$, $\theta = \alpha$, etc.

The basic form of the double integral is $\iint_R f(x, y)\, dA$. We interpret this integral as follows: over the region R, sum up lots of products of heights (given by $f(x_i, y_i)$) and areas (given by ΔA_i). That is, dA represents "a little bit of area." In rectangular coordinates, we can describe a small rectangle as having area $dx\, dy$ or $dy\, dx$ – the area of a rectangle is simply length$\times$width – a small change in x times a small change in y. Thus we replace dA in the double integral with $dx\, dy$ or $dy\, dx$.

Now consider representing a region R with polar coordinates. Consider Figure 14.19(a). Let R be the region in the first quadrant bounded by the curve. We can approximate this region using the natural shape of polar coordinates: portions of sectors of circles. In the figure, one such region is shaded, shown again in part (b) of the figure.

As the area of a sector of a circle with radius r, subtended by an angle θ, is $A = \frac{1}{2}r^2\theta$, we can find the area of the shaded region. The whole sector has area $\frac{1}{2}r_2^2\Delta\theta$, whereas the smaller, unshaded sector has area $\frac{1}{2}r_1^2\Delta\theta$. The area of the shaded region is the difference of these areas:

$$\Delta A_i = \frac{1}{2}r_2^2\Delta\theta - \frac{1}{2}r_1^2\Delta\theta = \frac{1}{2}\left(r_2^2 - r_1^2\right)\left(\Delta\theta\right) = \frac{r_2 + r_1}{2}\left(r_2 - r_1\right)\Delta\theta.$$

Note that $(r_2 + r_1)/2$ is just the average of the two radii.

To approximate the region R, we use many such subregions; doing so shrinks the difference $r_2 - r_1$ between radii to 0 and shrinks the change in angle $\Delta\theta$ also to 0. We represent these infinitesimal changes in radius and angle as dr and $d\theta$, respectively. Finally, as dr is small, $r_2 \approx r_1$, and so $(r_2 + r_1)/2 \approx r_1$. Thus, when dr and $d\theta$ are small,

$$\Delta A_i \approx r_i\, dr\, d\theta.$$

Taking a limit, where the number of subregions goes to infinity and both $r_2 - r_1$ and $\Delta\theta$ go to 0, we get

$$dA = r\, dr\, d\theta.$$

So to evaluate $\iint_R f(x, y)\, dA$, replace dA with $r\, dr\, d\theta$. Convert the function $z = f(x, y)$ to a function with polar coordinates with the substitutions $x = r\cos\theta$,

(a)

(b)
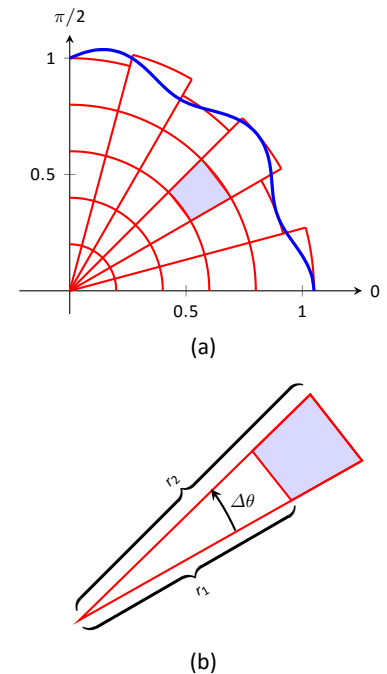
Figure 14.19: Approximating a region R with portions of sectors of circles.

Notes:

$y = r \sin \theta$. Finally, find bounds $g_1(\theta) \le r \le g_2(\theta)$ and $\alpha \le \theta \le \beta$ that describe R. This is the key principle of this section, so we restate it here as a Key Idea.

Key Idea 61   Evaluating Double Integrals with Polar Coordinates

Let R be a plane region bounded by the polar equations $\alpha \le \theta \le \beta$ and $g_1(\theta) \le r \le g_2(\theta)$. Then

$$\iint_R f(x, y) \, dA = \int_\alpha^\beta \int_{g_1(\theta)}^{g_2(\theta)} f\big(r \cos \theta, r \sin \theta\big) \, r \, dr \, d\theta.$$

Watch the video:
Double Integral Using Polar Coordinates — Part 1 of 3 at
`https://youtu.be/sQM-80j4Ecg`

Examples will help us understand this Key Idea.

Example 1   Evaluating a double integral with polar coordinates
Find the signed volume under the plane $z = 4 - x - 2y$ over the disk with equation $x^2 + y^2 \le 1$.

SOLUTION   The bounds of the integral are determined solely by the region R over which we are integrating. The surface and region are shown in Figure 14.20. In this case, it is a circle with equation $x^2 + y^2 = 1$. We need to find polar bounds for this region. It may help to review Section 10.4; bounds for this circle are $0 \le r \le 1$ and $0 \le \theta \le 2\pi$.

We replace $f(x, y)$ with $f(r \cos \theta, r \sin \theta)$. That means we make the following substitutions:

$$4 - x - 2y \quad \Rightarrow \quad 4 - r \cos \theta - 2r \sin \theta.$$

Finally, we replace dA in the double integral with $r \, dr \, d\theta$. This gives the final
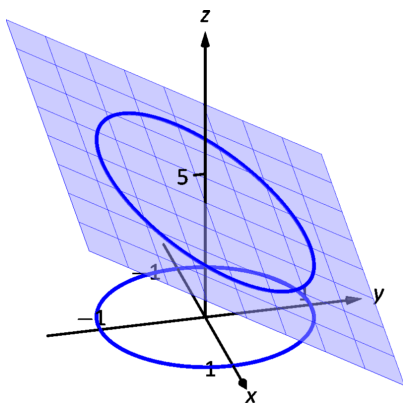
Figure 14.20: Evaluating a double integral with polar coordinates in Example 1.

Notes:

iterated integral, which we evaluate:

$$
\begin{aligned}
\iint_R f(x,y)\, dA &= \int_0^{2\pi} \int_0^1 \left(4 - r\cos\theta - 2r\sin\theta\right) r \, dr \, d\theta \\
&= \int_0^{2\pi} \int_0^1 \left(4r - r^2(\cos\theta - 2\sin\theta)\right) dr \, d\theta \\
&= \int_0^{2\pi} \left(2r^2 - \frac{1}{3}r^3(\cos\theta - 2\sin\theta)\right)\Big|_0^1 \, d\theta \\
&= \int_0^{2\pi} \left(2 - \frac{1}{3}(\cos\theta - 2\sin\theta)\right) d\theta \\
&= \left(2\theta - \frac{1}{3}(\sin\theta + 2\cos\theta)\right)\Big|_0^{2\pi} \\
&= 4\pi.
\end{aligned}
$$

Example 2   Evaluating a double integral with polar coordinates
Find the volume under the paraboloid $z = 4 - (x-2)^2 - y^2$ over the region bounded by the circles $(x-1)^2 + y^2 = 1$ and $(x-2)^2 + y^2 = 4$.

(a)

SOLUTION   At first glance, this seems like a very hard volume to compute as the region R (shown in Figure 14.21(a)) has a hole in it, cutting out a strange portion of the surface, as shown in part (b) of the figure. However, by describing R in terms of polar equations, the volume is not very difficult to compute. It is straightforward to show that the circle $(x-1)^2 + y^2 = 1$ has polar equation $r = 2\cos\theta$, and that the circle $(x-2)^2 + y^2 = 4$ has polar equation $r = 4\cos\theta$. Each of these circles is traced out on the interval $0 \le \theta \le \pi$. The bounds on r are $2\cos\theta \le r \le 4\cos\theta$.

Replacing x with $r\cos\theta$ in the integrand, along with replacing y with $r\sin\theta$, prepares us to evaluate the double integral $\iint_R f(x,y)\, dA$:

(b)
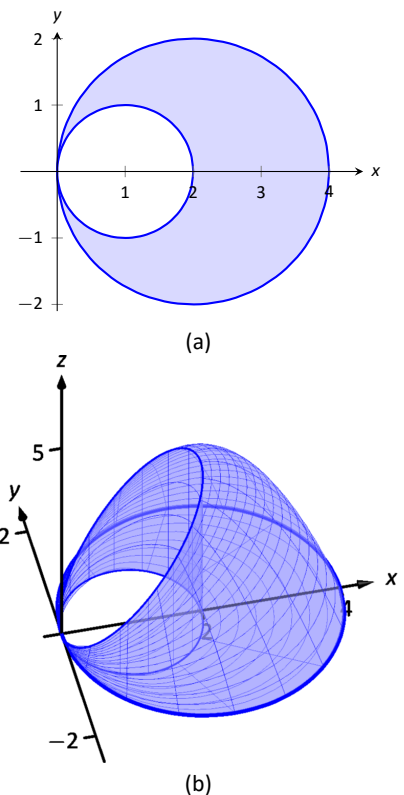
Figure 14.21: Showing the region R and surface used in Example 2.

Notes:

875

$$\iint_R f(x, y) \, dA = \int_0^\pi \int_{2\cos\theta}^{4\cos\theta} \left(4 - \left(r\cos\theta - 2 \right)^2 - \left(r\sin\theta \right)^2 \right) r \, dr \, d\theta$$

$$= \int_0^\pi \int_{2\cos\theta}^{4\cos\theta} \left(-r^3 + 4r^2\cos\theta \right) dr \, d\theta$$

$$= \int_0^\pi \left(-\frac{1}{4}r^4 + \frac{4}{3}r^3\cos\theta \right) \Bigg|_{2\cos\theta}^{4\cos\theta} d\theta$$

$$= \int_0^\pi \left(\left[-\frac{1}{4}(256\cos^4\theta) + \frac{4}{3}(64\cos^4\theta) \right] - \left[-\frac{1}{4}(16\cos^4\theta) + \frac{4}{3}(8\cos^4\theta) \right] \right) d\theta$$

$$= \int_0^\pi \frac{44}{3}\cos^4\theta \, d\theta.$$

To integrate $\cos^4\theta$, rewrite it as $\cos^2\theta\cos^2\theta$ and employ the half–angle formula twice:

$$\cos^4\theta = \cos^2\theta\cos^2\theta$$

$$= \frac{1}{2}\left(1 + \cos(2\theta) \right)\frac{1}{2}\left(1 + \cos(2\theta) \right)$$

$$= \frac{1}{4}\left(1 + 2\cos(2\theta) + \cos^2(2\theta) \right)$$

$$= \frac{1}{4}\left(1 + 2\cos(2\theta) + \frac{1}{2}\left(1 + \cos(4\theta) \right) \right)$$

$$= \frac{3}{8} + \frac{1}{2}\cos(2\theta) + \frac{1}{8}\cos(4\theta).$$

Picking up from where we left off above, we have

$$\iint_R f(x, y) \, dA = \int_0^\pi \frac{44}{3}\cos^4\theta \, d\theta$$

$$= \int_0^\pi \frac{44}{3}\left(\frac{3}{8} + \frac{1}{2}\cos(2\theta) + \frac{1}{8}\cos(4\theta) \right) d\theta$$

$$= \frac{44}{3}\left(\frac{3}{8}\theta + \frac{1}{4}\sin(2\theta) + \frac{1}{32}\sin(4\theta) \right) \Bigg|_0^\pi$$

$$= \frac{11}{2}\pi.$$

While this example was not trivial, the double integral would have been *much* harder to evaluate had we used rectangular coordinates.

Notes:

Example 3   **Evaluating a double integral with polar coordinates**

Find the volume under the surface $f(x, y) = \dfrac{1}{x^2 + y^2 + 1}$ over the sector of the circle with radius a centered at the origin in the first quadrant, as shown in Figure 14.22.

    Solution   The region R we are integrating over is a circle with radius a, restricted to the first quadrant. Thus, in polar, the bounds on R are $0 \le r \le a$, $0 \le \theta \le \pi/2$. The integrand is rewritten in polar as

$$\frac{1}{x^2 + y^2 + 1} \Rightarrow \frac{1}{r^2 \cos^2\theta + r^2 \sin^2\theta + 1} = \frac{1}{r^2 + 1}.$$

We find the volume as follows:

$$\iint_R f(x, y)\, dA = \int_0^{\pi/2} \int_0^a \frac{r}{r^2 + 1}\, dr\, d\theta$$

$$= \int_0^{\pi/2} \frac{1}{2}\left(\ln |r^2 + 1| \right)\Big|_0^a \, d\theta$$

$$= \int_0^{\pi/2} \frac{1}{2} \ln(a^2 + 1)\, d\theta$$

$$= \left(\frac{1}{2} \ln(a^2 + 1)\theta \right)\Big|_0^{\pi/2}$$

$$= \frac{\pi}{4} \ln(a^2 + 1).$$

Figure 14.22 shows that f shrinks to near 0 very quickly. Regardless, as a grows, so does the volume, without bound.

Figure 14.22: The surface and region R used in Example 3.

Note: Previous work has shown that there is finite *area* under $\frac{1}{x^2+1}$ over the entire x-axis. However, Example 3 shows that there is infinite *volume* under $\frac{1}{x^2+y^2+1}$ over the entire x-y plane.

Example 4   **Finding the volume of a sphere**

Find the volume of a sphere with radius a.

    Solution   The sphere of radius a, centered at the origin, has equation $x^2 + y^2 + z^2 = a^2$; solving for z, we have $z = \sqrt{a^2 - x^2 - y^2}$. This gives the upper half of a sphere. We wish to find the volume under this top half, then double it to find the total volume.

    The region we need to integrate over is the circle of radius a, centered at the origin. Polar bounds for this equation are $0 \le r \le a$, $0 \le \theta \le 2\pi$.

    All together, the volume of a sphere with radius a is:

$$2 \iint_R \sqrt{a^2 - x^2 - y^2}\, dA = 2 \int_0^{2\pi} \int_0^a \sqrt{a^2 - (r\cos\theta)^2 - (r\sin\theta)^2}\, r\, dr\, d\theta$$

$$= 2 \int_0^{2\pi} \int_0^a r\sqrt{a^2 - r^2}\, dr\, d\theta.$$

Notes:

We can evaluate this inner integral with substitution. With $u = a^2 - r^2$, $du = -2r\,dr$. The new bounds of integration are $u(0) = a^2$ to $u(a) = 0$. Thus we have:

$$= \int_0^{2\pi} \int_{a^2}^0 \left(-u^{1/2} \right) du\,d\theta$$

$$= \int_0^{2\pi} \left(-\frac{2}{3}u^{3/2} \right)\Big|_{a^2}^0 d\theta$$

$$= \int_0^{2\pi} \left(\frac{2}{3}a^3 \right) d\theta$$

$$= \left(\frac{2}{3}a^3\theta \right)\Big|_0^{2\pi}$$

$$= \frac{4}{3}\pi a^3.$$

Generally, the formula for the volume of a sphere with radius r is given as $4/3\pi r^3$; we have justified this formula with our calculation.
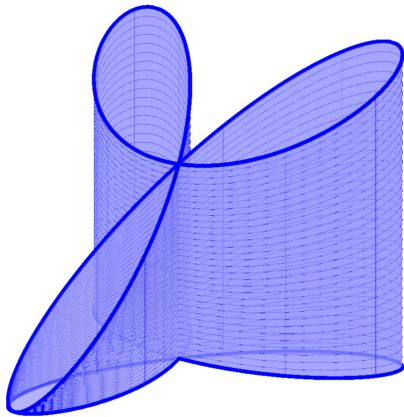
Figure 14.23: Visualizing the solid used in Example 5.

Example 5   Finding the volume of a solid
A sculptor wants to make a solid bronze cast of the solid shown in Figure 14.23, where the base of the solid has boundary, in polar coordinates, $r = \cos(3\theta)$, and the top is defined by the plane $z = 1 - x + 0.1y$. Find the volume of the solid.

   SOLUTION   From the outset, we should recognize that knowing *how to set up* this problem is probably more important than knowing *how to compute the integrals.* The iterated integral to come is not "hard" to evaluate, though it is long, requiring lots of algebra. Once the proper iterated integral is determined, one can use readily–available technology to help compute the final answer.
   The region R that we are integrating over is bound by $0 \leq r \leq \cos(3\theta)$, for $0 \leq \theta \leq \pi$ (note that this rose curve is traced out on the interval $[0, \pi]$, not $[0, 2\pi]$). This gives us our bounds of integration. The integrand is $z = 1 - x + 0.1y$; converting to polar, we have that the volume V is:

$$V = \iint_R f(x, y)\,dA = \int_0^\pi \int_0^{\cos(3\theta)} \left(1 - r\cos\theta + 0.1r\sin\theta \right) r\,dr\,d\theta.$$

Distributing the r, the inner integral is easy to evaluate, leading to

$$\int_0^\pi \left(\frac{1}{2}\cos^2(3\theta) - \frac{1}{3}\cos^3(3\theta)\cos\theta + \frac{0.1}{3}\cos^3(3\theta)\sin\theta \right) d\theta.$$

Notes:

This integral takes time to compute by hand; it is rather long and cumbersome. The powers of cosine need to be reduced, and products like $\cos(3\theta)\cos\theta$ need to be turned to sums using the Product To Sum formulas in the back cover of this text.

We rewrite $\frac{1}{2}\cos^2(3\theta)$ as $\frac{1}{4}(1+\cos(6\theta))$. We can also rewrite $\frac{1}{3}\cos^3(3\theta)\cos\theta$ as:

$$\frac{1}{3}\cos^3(3\theta)\cos\theta = \frac{1}{3}\cos^2(3\theta)\cos(3\theta)\cos\theta = \frac{1}{3}\frac{1+\cos(6\theta)}{2}\big(\cos(4\theta)+\cos(2\theta)\big).$$

This last expression still needs simplification, but eventually all terms can be reduced to the form $a\cos(m\theta)$ or $a\sin(m\theta)$ for various values of a and m.

We forgo the algebra and recommend the reader employ technology, such as WolframAlpha®, to compute the numeric answer. Such technology gives:

$$\int_0^\pi \int_0^{\cos(3\theta)} \big(1 - r\cos\theta + 0.1r\sin\theta\big)r\,dr\,d\theta = \frac{\pi}{4}\text{ units}^3.$$

Since the units were not specified, we leave the result as almost 0.8 cubic units (meters, feet, etc.).

We have used iterated integrals to find areas of plane regions and volumes under surfaces. Just as a single integral can be used to compute much more than "area under the curve," iterated integrals can be used to compute much more than we have thus far seen. The next two sections show two, among many, applications of iterated integrals.

Notes:

Exercises 14.3

Terms and Concepts

1. When evaluating $\iint_R f(x, y)\, dA$ using polar coordinates, $f(x, y)$ is replaced with _____ and dA is replaced with _____.

2. Why would one be interested in evaluating a double integral with polar coordinates?

Problems

In Exercises 3–10, a function $f(x, y)$ is given and a region R of the x-y plane is described. Set up and evaluate $\iint_R f(x, y)\, dA$ using polar coordinates.

3. $f(x, y) = 3x - y + 4$; R is the region enclosed by the circle $x^2 + y^2 = 1$.

4. $f(x, y) = 4x + 4y$; R is the region enclosed by the circle $x^2 + y^2 = 4$.

5. $f(x, y) = 8 - y$; R is the region enclosed by the circles with polar equations $r = \cos\theta$ and $r = 3\cos\theta$.

6. $f(x, y) = 4$; R is the region enclosed by the petal of the rose curve $r = \sin(2\theta)$ in the first quadrant.

7. $f(x, y) = \ln\left(x^2 + y^2\right)$; R is the annulus enclosed by the circles $x^2 + y^2 = 1$ and $x^2 + y^2 = 4$.

8. $f(x, y) = 1 - x^2 - y^2$; R is the region enclosed by the circle $x^2 + y^2 = 1$.

9. $f(x, y) = x^2 - y^2$; R is the region enclosed by the circle $x^2 + y^2 = 36$ in the first and fourth quadrants.

10. $f(x, y) = (x - y)/(x + y)$; R is the region enclosed by the lines $y = x$, $y = 0$ and the circle $x^2 + y^2 = 1$ in the first quadrant.

In Exercises 11–14, an iterated integral in rectangular coordinates is given. Rewrite the integral using polar coordinates and evaluate the new double integral.

11. $\displaystyle\int_0^5 \int_{-\sqrt{25-x^2}}^{\sqrt{25-x^2}} \sqrt{x^2 + y^2}\, dy\, dx$

12. $\displaystyle\int_{-4}^{4} \int_{-\sqrt{16-y^2}}^{0} \left(2y - x\right) dx\, dy$

13. $\displaystyle\int_0^2 \int_y^{\sqrt{8-y^2}} \left(x + y\right) dx\, dy$

14. $\displaystyle\int_{-2}^{-1} \int_0^{\sqrt{4-x^2}} \left(x+5\right) dy\, dx + \int_{-1}^{1} \int_{\sqrt{1-x^2}}^{\sqrt{4-x^2}} \left(x+5\right) dy\, dx +$
$\displaystyle\int_1^2 \int_0^{\sqrt{4-x^2}} \left(x + 5\right) dy\, dx$

Hint: draw the region of each integral carefully and see how they all connect.

In Exercises 15–16, special double integrals are presented that are especially well suited for evaluation in polar coordinates.

15. Consider $\displaystyle\iint_R e^{-(x^2+y^2)}\, dA$.

   (a) Why is this integral difficult to evaluate in rectangular coordinates, regardless of the region R?

   (b) Let R be the region bounded by the circle of radius a centered at the origin. Evaluate the double integral using polar coordinates.

   (c) Take the limit of your answer from (b), as $a \to \infty$. What does this imply about the volume under the surface of $e^{-(x^2+y^2)}$ over the entire x-y plane?

16. The surface of a right circular cone with height h and base radius a can be described by the equation $f(x, y) = h - h\sqrt{\dfrac{x^2}{a^2} + \dfrac{y^2}{a^2}}$, where the tip of the cone lies at $(0, 0, h)$ and the circular base lies in the x-y plane, centered at the origin.

   Confirm that the volume of a right circular cone with height h and base radius a is $V = \dfrac{1}{3}\pi a^2 h$ by evaluating $\displaystyle\iint_R f(x, y)\, dA$ in polar coordinates.

14.4 Center of Mass

We have used iterated integrals to find areas of plane regions and signed volumes under surfaces. A brief recap of these uses will be useful in this section as we apply iterated integrals to compute the **mass** and **center of mass** of planar regions.

To find the area of a planar region, we evaluated the double integral $\iint_R dA$. That is, summing up the areas of lots of little subregions of R gave us the total area. Informally, we think of $\iint_R dA$ as meaning "sum up lots of little areas over R."

To find the signed volume under a surface, we evaluated the double integral $\iint_R f(x, y)\, dA$. Recall that the "dA" is not just a "bookend" at the end of an integral; rather, it is multiplied by $f(x, y)$. We regard $f(x, y)$ as giving a height, and dA still giving an area: $f(x, y)\, dA$ gives a volume. Thus, informally, $\iint_R f(x, y)\, dA$ means "sum up lots of little volumes over R."

We now extend these ideas to other contexts.

Mass and Weight

Consider a thin sheet of material with constant thickness and finite area. Mathematicians (and physicists and engineers) call such a sheet a **lamina**. So consider a lamina, as shown in Figure 14.24(a), with the shape of some planar region R, as shown in part (b).

We can write a simple double integral that represents the mass of the lamina: $\iint_R dm$, where "dm" means "a little mass." That is, the double integral states the total mass of the lamina can be found by "summing up lots of little masses over R."

To evaluate this double integral, partition R into n subregions as we have done in the past. The ith subregion has area ΔA_i. A fundamental property of mass is that "mass=density×area." If the lamina has a constant density δ, then the mass of this i^{th} subregion is $\Delta m_i = \delta \Delta A_i$. That is, we can compute a small amount of mass by multiplying a small amount of area by the density.

If density is variable, with density function $\delta = \delta(x, y)$, then we can approximate the mass of the ith subregion of R by multiplying ΔA_i by $\delta(x_i, y_i)$, where (x_i, y_i) is a point in that subregion. That is, for a small enough subregion of R, the density across that region is almost constant.

The total mass M of the lamina is approximately the sum of approximate masses of subregions:

$$M \approx \sum_{i=1}^{n} \Delta m_i = \sum_{i=1}^{n} \delta(x_i, y_i)\Delta A_i.$$

(a)

(b)
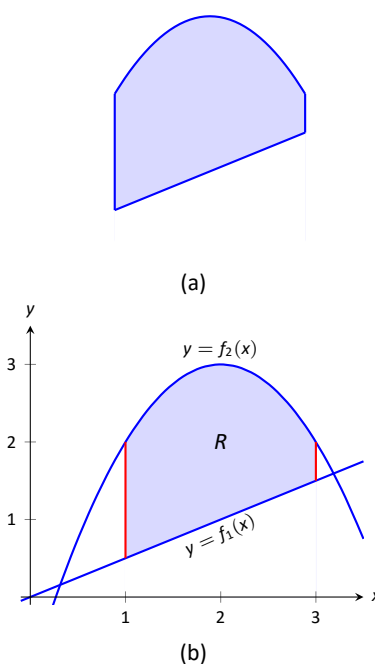
Figure 14.24: Illustrating the concept of a lamina.

Note: *Mass* and *weight* are different measures. Since they are scalar multiples of each other, it is often easy to treat them as the same measure. In this section we effectively treat them as the same, as our technique for finding mass is the same as for finding weight. The density functions used will simply have different units.

Notes:

Taking the limit as the size of the subregions shrinks to 0 gives us the actual mass; that is, integrating $\delta(x, y)$ over R gives the mass of the lamina.

> **Definition 105**   **Mass of a Lamina with Vairable Density**
>
> Let $\delta(x, y)$ be a continuous density function of a lamina corresponding to a plane region R. The mass M of the lamina is
>
> $$\text{mass } M = \iint_R dm = \iint_R \delta(x, y) \, dA.$$

Watch the video:
Center of Mass for a Rectangle of Variable Density at
`https://youtu.be/5CmgNCjRVFE`

Example 1   **Finding the mass of a lamina with constant density**
Find the mass of a square lamina, with side length 1, with a density of $\delta = 3\text{g/cm}^2$.

    SOLUTION   We represent the lamina with a square region in the plane as shown in Figure 14.25. As the density is constant, it does not matter where we place the square.

Following Definition 105, the mass M of the lamina is

$$M = \iint_R 3 \, dA = \int_0^1 \int_0^1 3 \, dx \, dy = 3 \int_0^1 \int_0^1 dx \, dy = 3\text{g}.$$

    This is all very straightforward; note that all we really did was find the area of the lamina and multiply it by the constant density of 3g/cm^2.

Example 2   **Finding the mass of a lamina with variable density**
Find the mass of a square lamina, represented by the unit square with lower lefthand corner at the origin (see Figure 14.25), with variable density $\delta(x, y) = (x + y + 2)\text{g/cm}^2$.
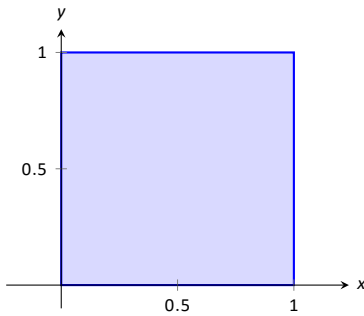
Figure 14.25: A region R representing a lamina in Example 1.

Notes:

Solution The variable density δ, in this example, is very uniform, giving a density of 3 in the center of the square and changing linearly. A graph of $\delta(x, y)$ can be seen in Figure 14.26; notice how "same amount" of density is above $z = 3$ as below. We'll comment on the significance of this momentarily.

The mass M is found by integrating $\delta(x, y)$ over R. The order of integration is not important; we choose $dx\,dy$ arbitrarily. Thus:

$$M = \iint_R (x + y + 2)\, dA = \int_0^1 \int_0^1 (x + y + 2)\, dx\, dy$$

$$= \int_0^1 \left(\frac{1}{2}x^2 + x(y + 2) \right)\Big|_0^1 dy$$

$$= \int_0^1 \left(\frac{5}{2} + y \right) dy$$

$$= \left(\frac{5}{2}y + \frac{1}{2}y^2 \right)\Big|_0^1$$

$$= 3\text{ g.}$$

It turns out that since since the density of the lamina is so uniformly distributed "above and below" $z = 3$ that the mass of the lamina is the same as if it had a constant density of 3. The density functions in Examples 1 and 2 are graphed in Figure 14.26, which illustrates this concept.

Example 3 **Finding the weight of a lamina with variable density**
Find the weight of the lamina represented by the circle with radius 2ft, centered at the origin, with density function $\delta(x, y) = (x^2 + y^2 + 1)\text{lb/ft}^2$. Compare this to the weight of the same lamina with density $\delta(x, y) = (2\sqrt{x^2 + y^2} + 1)\text{lb/ft}^2$.

Figure 14.26: Graphing the density functions in Examples 1 and 2.

Solution A direct application of Definition 105 states that the weight of the lamina is $\iint_R \delta(x, y)\, dA$. Since our lamina is in the shape of a circle, it makes sense to approach the double integral using polar coordinates.

The density function $\delta(x, y) = x^2 + y^2 + 1$ becomes $\delta(r, \theta) = (r\cos\theta)^2 + (r\sin\theta)^2 + 1 = r^2 + 1$. The circle is bounded by $0 \leq r \leq 2$ and $0 \leq \theta \leq 2\pi$.

Notes:

Thus the weight W is:

$$W = \int_0^{2\pi} \int_0^2 (r^2 + 1)r \, dr \, d\theta$$

$$= \int_0^{2\pi} \left(\frac{1}{4}r^4 + \frac{1}{2}r^2 \right) \Big|_0^2 \, d\theta$$

$$= \int_0^{2\pi} (6) \, d\theta$$

$$= 12\pi \text{lb}.$$

Now compare this with the density function $\delta(x, y) = 2\sqrt{x^2 + y^2} + 1$. Converting this to polar coordinates gives $\delta(r, \theta) = 2\sqrt{(r\cos\theta)^2 + (r\sin\theta)^2} + 1 = 2r + 1$. Thus the weight W is:

$$W = \int_0^{2\pi} \int_0^2 (2r + 1)r \, dr \, d\theta$$

$$= \int_0^{2\pi} \left(\frac{2}{3}r^3 + \frac{1}{2}r^2 \right) \Big|_0^2 \, d\theta$$

$$= \int_0^{2\pi} \left(\frac{22}{3} \right) \, d\theta$$

$$= \frac{44}{3}\pi \text{lb}.$$

One would expect different density functions to return different weights, as we have here. The density functions were chosen, though, to be similar: each gives a density of 1 at the origin and a density of 5 at the outside edge of the circle, as seen in Figure 14.27.

Notice how $x^2 + y^2 + 1 \leq 2\sqrt{x^2 + y^2} + 1$ over the circle; this results in less weight.

Plotting the density functions can be useful as our understanding of mass can be related to our understanding of "volume under a surface." We interpreted $\iint_R f(x, y) \, dA$ as giving the volume under f over R; we can understand $\iint_R \delta(x, y) \, dA$ in the same way. The "volume" under δ over R is actually mass; by compressing the "volume" under δ onto the x-y plane, we get "more mass" in some areas than others – i.e., areas of greater density.

Knowing the mass of a lamina is one of several important measures. Another is the **center of mass**, which we discuss next.

Notes:
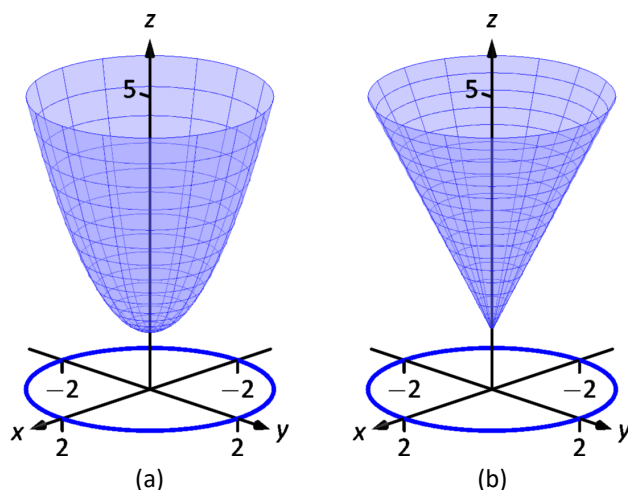
Figure 14.27: Graphing the density functions in Example 3. In (a) is the density function $\delta(x,y) = x^2 + y^2 + 1$; in (b) is $\delta(x,y) = 2\sqrt{x^2 + y^2} + 1$.

Center of Mass

Consider a disk of radius 1 with uniform density. It is common knowledge that the disk will balance on a point if the point is placed at the center of the disk. What if the disk does not have a uniform density? Through trial-and-error, we should still be able to find a spot on the disk at which the disk will balance on a point. This balance point is referred to as the **center of mass**, or **center of gravity**. It is as though all the mass is "centered" there. In fact, if the disk has a mass of 3kg, the disk will behave physically as though it were a point-mass of 3kg located at its center of mass. For instance, the disk will naturally spin with an axis through its center of mass (which is why it is important to "balance" the tires of your car: if they are "out of balance", their center of mass will be outside of the axle and it will shake terribly).

We find the center of mass based on the principle of a **weighted average**. Consider a college class in which your homework average is 90%, your test average is 73%, and your final exam grade is an 85%. Experience tells us that our final grade *is not* the *average* of these three grades: that is, it is not:

$$\frac{0.9 + 0.73 + 0.85}{3} \approx 0.837 = 83.7\%.$$

That is, you are probably not pulling a B in the course. Rather, your grades are *weighted*. Let's say the homework is worth 10% of the grade, tests are 60% and

Notes:

the exam is 30%. Then your final grade is:

$$(0.1)(0.9) + (0.6)(0.73) + (0.3)(0.85) = 0.783 = 78.3\%.$$

Each grade is multiplied by a **weight**.

In general, given values $x_1, x_2, \ldots, x_n$ and weights $w_1, w_2, \ldots, w_n$, the weighted average of the n values is

$$\sum_{i=1}^{n} w_i x_i \Big/ \sum_{i=1}^{n} w_i.$$

In the grading example above, the sum of the weights 0.1, 0.6 and 0.3 is 1, so we don't see the division by the sum of weights in that instance.

How this relates to center of mass is given in the following theorem.

Theorem 124   Center of Mass of Discrete Linear System

Let point masses $m_1, m_2, \ldots, m_n$ be distributed along the x-axis at locations $x_1, x_2, \ldots, x_n$, respectively. The center of mass $\bar{x}$ of the system is located at

$$\bar{x} = \sum_{i=1}^{n} m_i x_i \Big/ \sum_{i=1}^{n} m_i.$$

Example 4   Finding the center of mass of a discrete linear system

1. Point masses of 2g are located at $x = -1$, $x = 2$ and $x = 3$ are connected by a thin rod of negligible weight. Find the center of mass of the system.

2. Point masses of 10g, 2g and 1g are located at $x = -1$, $x = 2$ and $x = 3$, respectively, are connected by a thin rod of negligible weight. Find the center of mass of the system.

SOLUTION

1. Following Theorem 124, we compute the center of mass as:

$$\bar{x} = \frac{2(-1) + 2(2) + 2(3)}{2 + 2 + 2} = \frac{4}{3}.$$

So the system would balance on a point placed at $x = 4/3$, as illustrated in Figure 14.28(a).

(a)

(b)
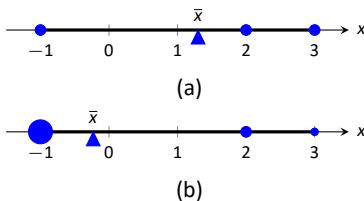
Figure 14.28: Illustrating point masses along a thin rod and the center of mass.

Notes:

2. Again following Theorem 124, we find:

$$\bar{x} = \frac{10(-1) + 2(2) + 1(3)}{10 + 2 + 1} = \frac{-3}{13}.$$

Placing a large weight at the left hand side of the system moves the center of mass left, as shown in Figure 14.28(b).

In a discrete system (i.e., mass is located at individual points, not along a continuum) we find the center of mass by dividing the mass into a **moment** of the system. In general, a moment is a weighted measure of distance from a particular point or line. In the case described by Theorem 124, we are finding a weighted measure of distances from the y-axis, so we refer to this as **the moment about the y-axis**, represented by M_y. Letting M be the total mass of the system, we have $\bar{x} = M_y/M$.

We can extend the concept of the center of mass of discrete points along a line to the center of mass of discrete points in the plane rather easily. To do so, we define some terms then give a theorem.

Definition 106   Moments about the x- and y- Axes.

Let point masses m_1, $m_2, \ldots, m_n$ be located at points (x_1, y_1), $(x_2, y_2) \ldots, (x_n, y_n)$, respectively, in the x-y plane.

1. The **moment about the y-axis**, M_y, is $M_y = \displaystyle\sum_{i=1}^{n} m_i x_i$.

2. The **moment about the x-axis**, M_x, is $M_x = \displaystyle\sum_{i=1}^{n} m_i y_i$.

One can think that these definitions are "backwards" as M_y sums up "x" distances. But remember, "x" distances are measurements of distance from the y-axis, hence defining the moment about the y-axis.

We now define the center of mass of discrete points in the plane.

Notes:

Figure 14.29: Illustrating the center of mass of a discrete planar system in Example 5.

> **Theorem 125   Center of Mass of Discrete Planar System**
>
> Let point masses m_1, $m_2, \ldots, m_n$ be located at points (x_1, y_1), $(x_2, y_2) \ldots, (x_n, y_n)$, respectively, in the x-y plane, and let $M = \sum_{i=1}^{n} m_i$.
>
> The center of mass of the system is at $(\bar{x}, \bar{y})$, where
>
> $$\bar{x} = \frac{M_y}{M} \quad \text{and} \quad \bar{y} = \frac{M_x}{M}.$$

Example 5   Finding the center of mass of a discrete planar system

Let point masses of 1kg, 2kg and 5kg be located at points $(2, 0)$, $(1, 1)$ and $(3, 1)$, respectively, and are connected by thin rods of negligible weight. Find the center of mass of the system.

SOLUTION   We follow Theorem 125 and Definition 106 to find M, M_x and M_y: first, $M = 1 + 2 + 5 = 8$kg. Next, we see that

$$M_x = \sum_{i=1}^{n} m_i y_i \qquad\qquad M_y = \sum_{i=1}^{n} m_i x_i$$
$$= 1(0) + 2(1) + 5(1) \qquad\qquad = 1(2) + 2(1) + 5(3)$$
$$= 7. \qquad\qquad\qquad = 19.$$

Thus the center of mass is $(\bar{x}, \bar{y}) = \left(\dfrac{M_y}{M}, \dfrac{M_x}{M} \right) = \left(\dfrac{19}{8}, \dfrac{7}{8} \right) = (2.375, 0.875)$, illustrated in Figure 14.29.

We finally arrive at our true goal of this section: finding the center of mass of a lamina with variable density. While the above measurement of center of mass is interesting, it does not directly answer more realistic situations where we need to find the center of mass of a contiguous region. However, understanding the discrete case allows us to approximate the center of mass of a planar lamina; using calculus, we can refine the approximation to an exact value.

We begin by representing a planar lamina with a region R in the x-y plane with density function $\delta(x, y)$. Partition R into n subdivisions, each with area ΔA_i. As done before, we can approximate the mass of the ith subregion with $\delta(x_i, y_i) \Delta A_i$, where (x_i, y_i) is a point inside the ith subregion. We can approximate the moment of this subregion about the y-axis with $x_i \delta(x_i, y_i) \Delta A_i$ – that is, by multiplying the approximate mass of the region by its approximate distance

Notes:

from the *y*-axis. Similarly, we can approximate the moment about the *x*-axis with $y_i\delta(x_i, y_i)\Delta A_i$. By summing over all subregions, we have:

$$\text{mass: } M \approx \sum_{i=1}^{n} \delta(x_i, y_i)\Delta A_i \quad \text{(as seen before)}$$

$$\text{moment about the } x\text{-axis: } M_x \approx \sum_{i=1}^{n} y_i\delta(x_i, y_i)\Delta A_i$$

$$\text{moment about the } y\text{-axis: } M_y \approx \sum_{i=1}^{n} x_i\delta(x_i, y_i)\Delta A_i$$

By taking limits, where size of each subregion shrinks to 0 in both the *x* and *y* directions, we arrive at the double integrals given in the following theorem.

Theorem 126   Center of Mass of a Planar Lamina, Moments

Let a planar lamina be represented by a region *R* in the *x*-*y* plane with density function $\delta(x, y)$.

1. mass: $M = \displaystyle\iint_R \delta(x, y)\, dA$

2. moment about the *x*-axis: $M_x = \displaystyle\iint_R y\delta(x, y)\, dA$

3. moment about the *y*-axis: $M_y = \displaystyle\iint_R x\delta(x, y)\, dA$

4. The center of mass of the lamina is

$$(\bar{x}, \bar{y}) = \left(\frac{M_y}{M}, \frac{M_x}{M} \right).$$

We start our practice of finding centers of mass by revisiting some of the lamina used previously in this section when finding mass. We will just set up the integrals needed to compute M, M_x and M_y and leave the details of the integration to the reader.

Example 6   Finding the center of mass of a lamina
Find the center mass of a square lamina, with side length 1, with a density of $\delta = 3\text{g/cm}^2$. (Note: this is the lamina from Example 1.)

Notes:
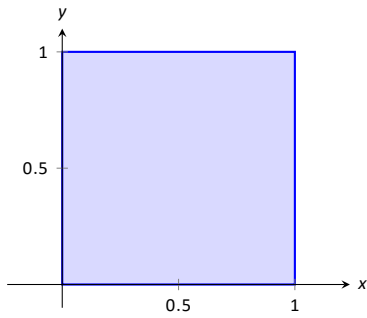
Figure 14.30: A region R representing a lamina in Examples 1 and 6.

SOLUTION   We represent the lamina with a square region in the plane as shown in Figure 14.30 as done previously.

Following Theorem 126, we find M, M_x and M_y:

$$M = \iint_R 3 \, dA = \int_0^1 \int_0^1 3 \, dx \, dy = 3g.$$

$$M_x = \iint_R 3y \, dA = \int_0^1 \int_0^1 3y \, dx \, dy = 3/2 = 1.5.$$

$$M_y = \iint_R 3x \, dA = \int_0^1 \int_0^1 3x \, dx \, dy = 3/2 = 1.5.$$

Thus the center of mass is $(\bar{x}, \bar{y}) = \left(\dfrac{M_y}{M}, \dfrac{M_x}{M} \right) = (1.5/3, 1.5/3) = (0.5, 0.5)$.
This is what we should have expected: the center of mass of a square with constant density is the center of the square.

Example 7   **Finding the center of mass of a lamina**
Find the center of mass of a square lamina, represented by the unit square with lower lefthand corner at the origin (see Figure 14.30), with variable density $\delta(x, y) = (x + y + 2)$g/cm^2. (Note: this is the lamina from Example 2.)

SOLUTION   We follow Theorem 126, to find M, M_x and M_y:

$$M = \iint_R (x + y + 2) \, dA = \int_0^1 \int_0^1 (x + y + 2) \, dx \, dy = 3g.$$

$$M_x = \iint_R y(x + y + 2) \, dA = \int_0^1 \int_0^1 y(x + y + 2) \, dx \, dy = \frac{19}{12}.$$

$$M_y = \iint_R x(x + y + 2) \, dA = \int_0^1 \int_0^1 x(x + y + 2) \, dx \, dy = \frac{19}{12}.$$

Thus the center of mass is $(\bar{x}, \bar{y}) = \left(\dfrac{M_y}{M}, \dfrac{M_x}{M} \right) = \left(\dfrac{19}{36}, \dfrac{19}{36} \right)$. While the mass of this lamina is the same as the lamina in the previous example, the greater density found with greater x and y values pulls the center of mass from the center slightly towards the upper righthand corner.

Example 8   **Finding the center of mass of a lamina**
Find the center of mass of the lamina represented by the circle with radius 2ft, centered at the origin, with density function $\delta(x, y) = (x^2 + y^2 + 1)$lb/ft^2. (Note: this is one of the lamina used in Example 3.)

Notes:

Solution As done in Example 3, it is best to describe R using polar coordinates. Thus when we compute M_y, we will integrate not $x\delta(x, y) = x(x^2 + y^2 + 1)$, but rather $\left(r\cos\theta\right)\delta(r\cos\theta, r\sin\theta) = \left(r\cos\theta\right)\left(r^2 + 1\right)$. We compute M, M_x and M_y:

$$M = \int_0^{2\pi}\int_0^2 (r^2 + 1)r\,dr\,d\theta = 12\pi\text{lb}.$$

$$M_x = \int_0^{2\pi}\int_0^2 (r\sin\theta)(r^2 + 1)r\,dr\,d\theta = 0.$$

$$M_y = \int_0^{2\pi}\int_0^2 (r\cos\theta)(r^2 + 1)r\,dr\,d\theta = 0.$$

Since R and the density of R are both symmetric about the x and y axes, it should come as no big surprise that the moments about each axis is 0. Thus the center of mass is $(\bar{x}, \bar{y}) = (0, 0)$.

Example 9 **Finding the center of mass of a lamina**
Find the center of mass of the lamina represented by the region R shown in Figure 14.31, half an annulus with outer radius 6 and inner radius 5, with constant density 2lb/ft^2.

Solution Once again it will be useful to represent R in polar coordinates. Using the description of R and/or the illustration, we see that R is bounded by $5 \leq r \leq 6$ and $0 \leq \theta \leq \pi$. As the lamina is symmetric about the y-axis, we should expect $M_y = 0$. We compute M, M_x and M_y:

$$M = \int_0^{\pi}\int_5^6 (2)r\,dr\,d\theta = 11\pi\text{lb}.$$

$$M_x = \int_0^{\pi}\int_5^6 (r\sin\theta)(2)r\,dr\,d\theta = \frac{364}{3}.$$

$$M_y = \int_0^{\pi}\int_5^6 (r\cos\theta)(2)r\,dr\,d\theta = 0.$$

Thus the center of mass is $(\bar{x}, \bar{y}) = \left(0, \frac{364}{33\pi}\right) \approx (0, 3.51)$. The center of mass is indicated in Figure 14.31; note how it lies outside of R.

This section has shown us another use for iterated integrals beyond finding area or signed volume under the curve. While there are many uses for iterated integrals, we give one more application in the following section: computing surface area.
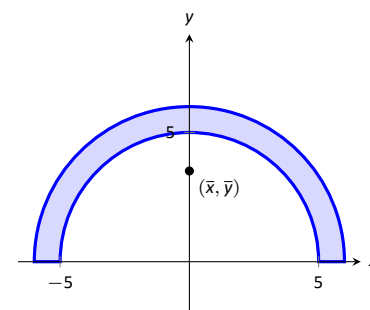
Figure 14.31: Illustrating the region R in Example 9.

Notes:

Exercises 14.4

Terms and Concepts

1. Why is it easy to use "mass" and "weight" interchangeably, even though they are different measures?

2. Given a point (x, y), the value of x is a measure of distance from the _____-axis.

3. We can think of $\iint_R \, dm$ as meaning "sum up lots of _____"

4. What is a "discrete planar system?"

5. Why does M_x use $\iint_R y\delta(x, y) \, dA$ instead of $\iint_R x\delta(x, y) \, dA$; that is, why do we use "y" and not "x"?

6. Describe a situation where the center of mass of a lamina does not lie within the region of the lamina itself.

Problems

In Exercises 7–10, point masses are given along a line or in the plane. Find the center of mass $\bar{x}$ or $(\bar{x}, \bar{y})$, as appropriate. (All masses are in grams and distances are in cm.)

7. $m_1 = 4$ at $x = 1$; $\quad m_2 = 3$ at $x = 3$; $\quad m_3 = 5$ at $x = 10$

8. $m_1 = 2$ at $x = -3$; $\quad m_2 = 2$ at $x = -1$;
   $m_3 = 3$ at $x = 0$; $\quad m_4 = 3$ at $x = 7$

9. $m_1 = 2$ at $(-2, -2)$; $\quad m_2 = 2$ at $(2, -2)$;
   $m_3 = 20$ at $(0, 4)$

10. $m_1 = 1$ at $(-1, -1)$; $\quad m_2 = 2$ at $(-1, 1)$;
    $m_3 = 2$ at $(1, 1)$; $\quad m_4 = 1$ at $(1, -1)$

In Exercises 11–18, find the mass/weight of the lamina described by the region R in the plane and its density function $\delta(x, y)$.

11. R is the rectangle with corners $(1, -3)$, $(1, 2)$, $(7, 2)$ and $(7, -3)$; $\delta(x, y) = 5$g/cm^2

12. R is the rectangle with corners $(1, -3)$, $(1, 2)$, $(7, 2)$ and $(7, -3)$; $\delta(x, y) = (x + y^2)$g/cm^2

13. R is the triangle with corners $(-1, 0)$, $(1, 0)$, and $(0, 1)$; $\delta(x, y) = 2$lb/in^2

14. R is the triangle with corners $(0, 0)$, $(1, 0)$, and $(0, 1)$; $\delta(x, y) = (x^2 + y^2 + 1)$lb/in^2

15. R is the circle centered at the origin with radius 2; $\delta(x, y) = (x + y + 4)$kg/m^2

16. R is the circle sector bounded by $x^2 + y^2 = 25$ in the first quadrant; $\delta(x, y) = (\sqrt{x^2 + y^2} + 1)$kg/m^2

17. R is the annulus in the first and second quadrants bounded by $x^2 + y^2 = 9$ and $x^2 + y^2 = 36$; $\delta(x, y) = 4$lb/ft^2

18. R is the annulus in the first and second quadrants bounded by $x^2 + y^2 = 9$ and $x^2 + y^2 = 36$; $\delta(x, y) = \sqrt{x^2 + y^2}$lb/ft^2

In Exercises 19–26, find the center of mass of the lamina described by the region R in the plane and its density function $\delta(x, y)$.
Note: these are the same lamina as in Exercises 11 – 18.

19. R is the rectangle with corners $(1, -3)$, $(1, 2)$, $(7, 2)$ and $(7, -3)$; $\delta(x, y) = 5$g/cm^2

20. R is the rectangle with corners $(1, -3)$, $(1, 2)$, $(7, 2)$ and $(7, -3)$; $\delta(x, y) = (x + y^2)$g/cm^2

21. R is the triangle with corners $(-1, 0)$, $(1, 0)$, and $(0, 1)$; $\delta(x, y) = 2$lb/in^2

22. R is the triangle with corners $(0, 0)$, $(1, 0)$, and $(0, 1)$; $\delta(x, y) = (x^2 + y^2 + 1)$lb/in^2

23. R is the circle centered at the origin with radius 2; $\delta(x, y) = (x + y + 4)$kg/m^2

24. R is the circle sector bounded by $x^2 + y^2 = 25$ in the first quadrant; $\delta(x, y) = (\sqrt{x^2 + y^2} + 1)$kg/m^2

25. R is the annulus in the first and second quadrants bounded by $x^2 + y^2 = 9$ and $x^2 + y^2 = 36$; $\delta(x, y) = 4$lb/ft^2

26. R is the annulus in the first and second quadrants bounded by $x^2 + y^2 = 9$ and $x^2 + y^2 = 36$; $\delta(x, y) = \sqrt{x^2 + y^2}$lb/ft^2

The *moment of inertia I* is a measure of the tendency of a lamina to resist rotating about an axis or continue to rotate about an axis. I_x is the moment of inertia about the x-axis, I_y is the moment of inertia about the y-axis, and I_O is the moment of inertia about the origin. These are computed as follows:

- $I_x = \displaystyle\iint_R y^2 \, dm$

- $I_y = \displaystyle\iint_R x^2 \, dm$

- $I_O = \displaystyle\iint_R \left(x^2 + y^2\right) \, dm$

In Exercises 27–30, a lamina corresponding to a planar region R is given with a mass of 16 units. For each, compute I_x, I_y and I_O.

27. R is the 4×4 square with corners at $(-2, -2)$ and $(2, 2)$ with density $\delta(x, y) = 1$.

28. R is the 8×2 rectangle with corners at $(-4, -1)$ and $(4, 1)$ with density $\delta(x, y) = 1$.

29. R is the 4×2 rectangle with corners at $(-2, -1)$ and $(2, 1)$ with density $\delta(x, y) = 2$.

30. R is the circle with radius 2 centered at the origin with density $\delta(x, y) = 4/\pi$.

14.5   Surface Area

In Section 10.1 we used definite integrals to compute the arc length of plane curves of the form $y = f(x)$. We later extended these ideas to compute the arc length of plane curves defined by parametric or polar equations.

The natural extension of the concept of "arc length over an interval" to surfaces is "surface area over a region."

Consider the surface $z = f(x, y)$ over a region R in the x-y plane, shown in Figure 14.32(a). Because of the domed shape of the surface, the surface area will be greater than that of the area of the region R. We can find this area using the same basic technique we have used over and over: we'll make an approximation, then using limits, we'll refine the approximation to the exact value.

As done to find the volume under a surface or the mass of a lamina, we subdivide R into n subregions. Here we subdivide R into rectangles, as shown in the figure. One such subregion is outlined in the figure, where the rectangle has dimensions Δx_i and Δy_i, along with its corresponding region on the surface.

In part (b) of the figure, we zoom in on this portion of the surface. When Δx_i and Δy_i are small, the function is approximated well by the tangent plane at any point (x_i, y_i) in this subregion, which is graphed in part (b). In fact, the tangent plane approximates the function so well that in this figure, it is virtually indistinguishable from the surface itself! Therefore we can approximate the surface area S_i of this region of the surface with the area T_i of the corresponding portion of the tangent plane.

This portion of the tangent plane is a parallelogram, defined by sides $\vec{u}$ and $\vec{v}$, as shown. One of the applications of the cross product from Section 11.4 is that the area of this parallelogram is $\|\vec{u} \times \vec{v}\|$. Once we can determine $\vec{u}$ and $\vec{v}$, we can determine the area.

$\vec{u}$ is tangent to the surface in the direction of x, therefore, from Section 13.7, $\vec{u}$ is parallel to $\langle 1, 0, f_x(x_i, y_i) \rangle$. The x-displacement of $\vec{u}$ is Δx_i, so we know that $\vec{u} = \Delta x_i \langle 1, 0, f_x(x_i, y_i) \rangle$. Similar logic shows that $\vec{v} = \Delta y_i \langle 0, 1, f_y(x_i, y_i) \rangle$. Thus:

$$\text{surface area } S_i \approx \text{area of } T_i$$
$$= \|\vec{u} \times \vec{v}\|$$
$$= \|\Delta x_i \langle 1, 0, f_x(x_i, y_i) \rangle \times \Delta y_i \langle 0, 1, f_y(x_i, y_i) \rangle\|$$
$$= \sqrt{1 + f_x(x_i, y_i)^2 + f_y(x_i, y_i)^2}\, \Delta x_i \Delta y_i.$$

Note that $\Delta x_i \Delta y_i = \Delta A_i$, the area of the ith subregion.

Summing up all n of the approximations to the surface area gives

$$\text{surface area over } R \approx \sum_{i=1}^{n} \sqrt{1 + f_x(x_i, y_i)^2 + f_y(x_i, y_i)^2}\, \Delta A_i.$$

(a)

(b)
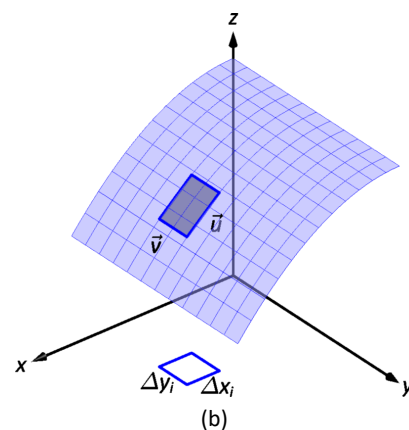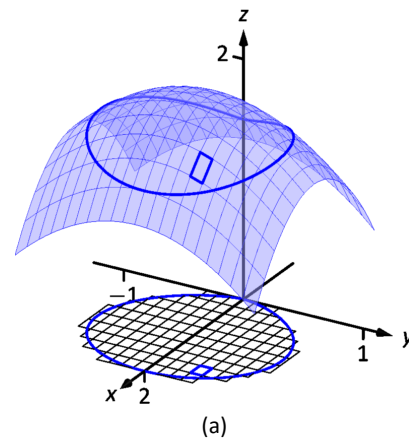
Figure 14.32: Developing a method of computing surface area.

Notes:

Once again take a limit as all of the Δx_i and Δy_i shrink to 0; this leads to a double integral.

Note: As before, we think of "$\iint_R dS$" as meaning "sum up lots of little surface areas over R."

The concept of surface area is *defined* here, for while we already have a notion of the area of a region in the *plane*, we did not yet have a solid grasp of what "the area of a surface in *space*" means.

> **Definition 107   Surface Area**
> Let $z = f(x, y)$ where f_x and f_y are continuous over a closed, bounded region R. The surface area S over R is
>
> $$S = \iint_R dS$$
> $$= \iint_R \sqrt{1 + f_x(x, y)^2 + f_y(x, y)^2}\, dA.$$

Watch the video:
Surface area of $z = (x^2 + y^2)^{1/2}$ at
`https://youtu.be/ricG1_x6xCo`

We test this definition by using it to compute surface areas of known surfaces. We start with a triangle.

Example 1   Finding the surface area of a plane over a triangle
Let $f(x, y) = 4 - x - 2y$, and let R be the region in the plane bounded by $x = 0$, $y = 0$ and $y = 2 - x/2$, as shown in Figure 14.33. Find the surface area of f over R.

SOLUTION   We follow Definition 107. We start by noting that $f_x(x, y) = -1$ and $f_y(x, y) = -2$. To define R, we use bounds $0 \le y \le 2 - x/2$ and $0 \le x \le 4$. Therefore

$$S = \iint_R dS$$
$$= \int_0^4 \int_0^{2-x/2} \sqrt{1 + (-1)^2 + (-2)^2}\, dy\, dx$$
$$= \int_0^4 \sqrt{6}\left(2 - \frac{x}{2}\right) dx$$
$$= 4\sqrt{6}.$$
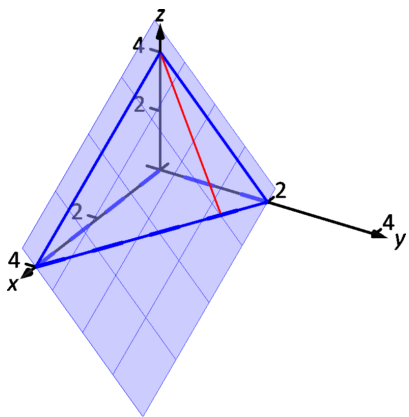
Figure 14.33: Finding the area of a triangle in space in Example 1.

Notes:

Because the surface is a triangle, we can figure out the area using geometry. Considering the base of the triangle to be the side in the x-y plane, we find the length of the base to be $\sqrt{20}$. We can find the height using our knowledge of vectors: let $\vec{u}$ be the side in the x-z plane and let $\vec{v}$ be the side in the x-y plane. The height is then $\|\vec{u} - \text{proj}_{\vec{v}}\,\vec{u}\| = 4\sqrt{6/5}$. Geometry states that the area is thus

$$\frac{1}{2} \cdot 4\sqrt{6/5} \cdot \sqrt{20} = 4\sqrt{6}.$$

We affirm the validity of our formula.

It is "common knowledge" that the surface area of a sphere of radius r is $4\pi r^2$. We confirm this in the following example, which involves using our formula with polar coordinates.

Example 2   **The surface area of a sphere.**
Find the surface area of the sphere with radius a centered at the origin, whose top hemisphere has equation $f(x, y) = \sqrt{a^2 - x^2 - y^2}$.

SOLUTION   We start by computing partial derivatives and find

$$f_x(x, y) = \frac{-x}{\sqrt{a^2 - x^2 - y^2}} \quad \text{and} \quad f_y(x, y) = \frac{-y}{\sqrt{a^2 - x^2 - y^2}}.$$

As our function f only defines the top upper hemisphere of the sphere, we double our surface area result to get the total area:

$$S = 2 \iint_R \sqrt{1 + f_x(x, y)^2 + f_y(x, y)^2}\; dA$$

$$= 2 \iint_R \sqrt{1 + \frac{x^2 + y^2}{a^2 - x^2 - y^2}}\; dA.$$

The region R that we are integrating over is the disk, centered at the origin, with radius a: $x^2 + y^2 \leq a^2$. Because of this region, we are likely to have greater success with our integration by converting to polar coordinates. Using the substitutions $x = r\cos\theta$, $y = r\sin\theta$, $dA = r\,dr\,d\theta$ and bounds $0 \leq \theta \leq 2\pi$ and $0 \leq r \leq a$, we have:

$$S = 2 \int_0^{2\pi} \int_0^a \sqrt{1 + \frac{r^2 \cos^2\theta + r^2 \sin^2\theta}{a^2 - r^2 \cos^2\theta - r^2 \sin^2\theta}}\; r\,dr\,d\theta$$

$$= 2 \int_0^{2\pi} \int_0^a r\sqrt{1 + \frac{r^2}{a^2 - r^2}}\; dr\,d\theta$$

$$= 2 \int_0^{2\pi} \int_0^a r\sqrt{\frac{a^2}{a^2 - r^2}}\; dr\,d\theta. \tag{14.1}$$

Note: The inner integral in Equation (14.1) is an improper integral, as the integrand of $\int_0^a r\sqrt{\dfrac{a^2}{a^2 - r^2}}\; dr$ is not defined at $r = a$. To properly evaluate this integral, one must use the techniques of Section 8.6.

The reason this need arises is that the function $f(x, y) = \sqrt{a^2 - x^2 - y^2}$ fails the requirements of Definition 107, as f_x and f_y are not continuous on the boundary of the circle $x^2 + y^2 = a^2$.

The computation of the surface area is still valid. The definition makes stronger requirements than necessary in part to avoid the use of improper integration, as when f_x and/or f_y are not continuous, the resulting improper integral may not converge. Since the improper integral does converge in this example, the surface area is accurately computed.

Notes:

Apply substitution $u = a^2 - r^2$ and integrate the inner integral, giving

$$= 2 \int_0^{2\pi} a^2 \, d\theta$$

$$= 4\pi a^2.$$

Our work confirms our previous formula.

Example 3　　　**Finding the surface area of a cone**

The general formula for a right cone with height h and base radius a, as shown in Figure 14.34, is

$$f(x, y) = h - \frac{h}{a}\sqrt{x^2 + y^2}.$$

Find the surface area of this cone.

SOLUTION　　　We begin by computing partial derivatives.

$$f_x(x, y) = -\frac{xh}{a\sqrt{x^2 + y^2}} \quad \text{and} \quad f_y(x, y) = -\frac{yh}{a\sqrt{x^2 + y^2}}.$$

Since we are integrating over the disk $x^2 + y^2 \leq a^2$, we again use polar coordinates. Using the standard substitutions, our integrand becomes

$$\sqrt{1 + \left(\frac{hr\cos\theta}{a\sqrt{r^2}}\right)^2 + \left(\frac{hr\sin\theta}{a\sqrt{r^2}}\right)^2}.$$

This may look intimidating at first, but there are lots of simple simplifications to be done. It amazingly reduces to just

$$\sqrt{1 + \frac{h^2}{a^2}} = \frac{1}{a}\sqrt{a^2 + h^2}.$$

Our polar bounds are $0 \leq \theta \leq 2\pi$ and $0 \leq r \leq a$. Thus

$$S = \int_0^{2\pi} \int_0^a r\frac{1}{a}\sqrt{a^2 + h^2} \, dr \, d\theta$$

$$= \int_0^{2\pi} \left(\frac{1}{2}r^2\frac{1}{a}\sqrt{a^2 + h^2}\right)\Big|_0^a \, d\theta$$

$$= \int_0^{2\pi} \frac{1}{2}a\sqrt{a^2 + h^2} \, d\theta$$

$$= \pi a\sqrt{a^2 + h^2}.$$

This matches the formula found in the back of this text.

Notes:

Figure 14.34: Finding the surface area of a cone in Example 3.

Note: Note that once again f_x and f_y are not continuous on the domain of f, as both are undefined at $(0, 0)$. (A similar problem occurred in the previous example.) Once again the resulting improper integral converges and the computation of the surface area is valid.

Example 4   Finding surface area over a region

Find the area of the surface $f(x, y) = x^2 - 3y + 3$ over the region R bounded by $-x \leq y \leq x$, $0 \leq x \leq 4$, as pictured in Figure 14.35.

SOLUTION   It is straightforward to compute $f_x(x, y) = 2x$ and $f_y(x, y) = -3$. Thus the surface area is described by the double integral

$$\iint_R \sqrt{1 + (2x)^2 + (-3)^2}\, dA = \iint_R \sqrt{10 + 4x^2}\, dA.$$

As with integrals describing arc length, double integrals describing surface area are in general hard to evaluate directly because of the square–root. This particular integral can be easily evaluated, though, with judicious choice of our order of integration.

Integrating with order $dx\, dy$ requires us to evaluate $\int \sqrt{10 + 4x^2}\, dx$. This can be done, though it involves Integration By Parts and $\sinh^{-1} x$. Integrating with order $dy\, dx$ has as its first integral $\int \sqrt{10 + 4x^2}\, dy$, which is easy to evaluate: it is simply $y\sqrt{10 + 4x^2} + C$. So we proceed with the order $dy\, dx$; the bounds are already given in the statement of the problem.

$$\iint_R \sqrt{10 + 4x^2}\, dA = \int_0^4 \int_{-x}^{x} \sqrt{10 + 4x^2}\, dy\, dx$$

$$= \int_0^4 \left(y\sqrt{10 + 4x^2} \right)\Big|_{-x}^{x} dx$$

$$= \int_0^4 \left(2x\sqrt{10 + 4x^2} \right) dx.$$

Apply substitution with $u = 10 + 4x^2$:

$$= \left(\frac{1}{6} (10 + 4x^2)^{3/2} \right)\Big|_0^4$$

$$= \frac{1}{3} \left(37\sqrt{74} - 5\sqrt{10} \right) \text{ units}^2.$$

So while the region R over which we integrate has an area of 16 units^2, the surface has a much greater area as its z-values change dramatically over R.

In practice, technology helps greatly in the evaluation of such integrals. High powered computer algebra systems can compute integrals that are difficult, or at least time consuming, by hand, and can at the least produce very accurate approximations with numerical methods. In general, just knowing *how* to set up the proper integrals brings one very close to being able to compute the needed
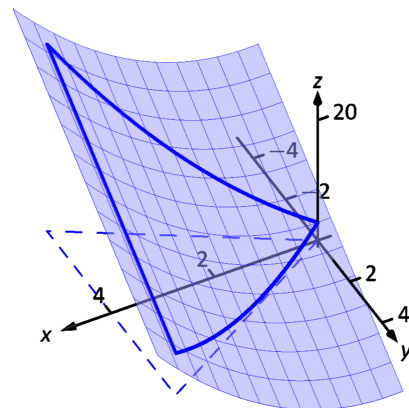
Figure 14.35: Graphing the surface in Example 4.

Notes:

value. Most of the work is actually done in just describing the region R in terms of polar or rectangular coordinates. Once this is done, technology can usually provide a good answer.

We have learned how to integrate integrals; that is, we have learned to evaluate double integrals. In the next section, we learn how to integrate double integrals – that is, we learn to evaluate *triple integrals*, along with learning some uses for this operation.

Notes:

Exercises 14.5

Terms and Concepts

1. "Surface area" is analogous to what previously studied concept?

2. To approximate the area of a small portion of a surface, we computed the area of its _____ plane.

3. We interpret $\iint_R dS$ as "sum up lots of little _____ _____."

4. Why is it important to know how to set up a double integral to compute surface area, even if the resulting integral is hard to evaluate?

5. Why do $z = f(x, y)$ and $z = g(x, y) = f(x, y) + h$, for some real number h, have the same surface area over a region R?

6. Let $z = f(x, y)$ and $z = g(x, y) = 2f(x, y)$. Why is the surface area of g over a region R not twice the surface area of f over R?

Problems

In Exercises 7–19, *set up* the iterated integral that computes the surface area of the given surface over the region R.

7. $f(x, y) = \sin x \cos y$; R is the rectangle with bounds $0 \le x \le 2\pi$, $0 \le y \le 2\pi$.
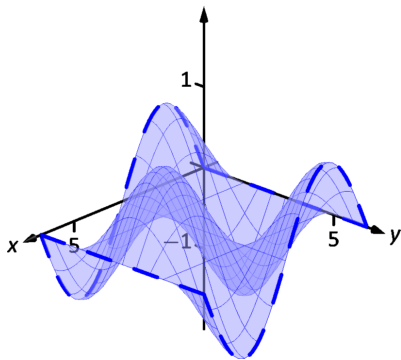
8. $f(x, y) = \dfrac{1}{x^2 + y^2 + 1}$; R is the disk $x^2 + y^2 \le 9$.
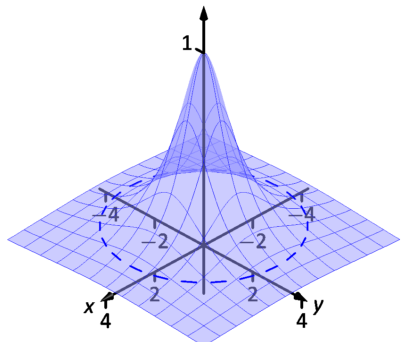
9. $f(x, y) = x^2 - y^2$; R is the rectangle with opposite corners $(-1, -1)$ and $(1, 1)$.
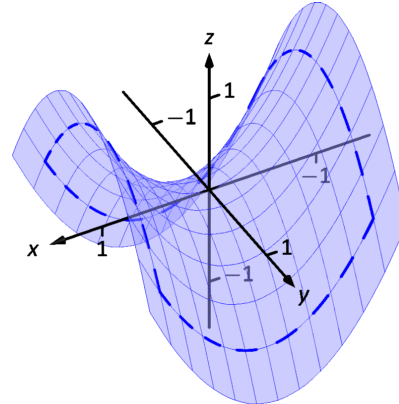
10. $f(x, y) = \dfrac{1}{e^{x^2} + 1}$; R is the rectangle bounded by $-5 \le x \le 5$ and $0 \le y \le 1$.
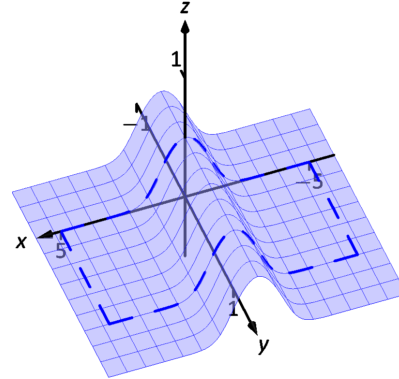
In Exercises 11–19, find the area of the given surface over the region R.

11. $f(x, y) = 3x - 7y + 2$; R is the rectangle with opposite corners $(-1, 0)$ and $(1, 3)$.

12. $f(x, y) = 2x + 2y + 2$; R is the triangle with corners $(0, 0)$, $(1, 0)$ and $(0, 1)$.

13. $f(x, y) = x^2 + y^2 + 10$; R is the disk $x^2 + y^2 \le 16$.

14. $f(x, y) = -2x + 4y^2 + 7$ over R, the triangle bounded by $y = -x$, $y = x$, $0 \le y \le 1$.

15. $f(x, y) = x^2 + y$ over R, the triangle bounded by $y = 2x$, $y = 0$ and $x = 2$.

16. $f(x, y) = \frac{2}{3}x^{3/2} + 2y^{3/2}$ over R, the rectangle with opposite corners $(0, 0)$ and $(1, 1)$.

17. $f(x, y) = 10 - 2\sqrt{x^2 + y^2}$ over R, the disk $x^2 + y^2 \le 25$. (This is the cone with height 10 and base radius 5; be sure to compare you result with the known formula.)

18. Find the surface area of the sphere with radius 5 by doubling the surface area of $f(x, y) = \sqrt{25 - x^2 - y^2}$ over R, the disk $x^2 + y^2 \le 25$. (Be sure to compare you result with the known formula.)

19. Find the surface area of the ellipse formed by restricting the plane $f(x, y) = cx + dy + h$ to the region R, the disk $x^2 + y^2 \le 1$, where c, d and h are some constants. Your answer should be given in terms of c and d; why does the value of h not matter?

14.6 Volume Between Surfaces and Triple Integration

We learned in Section 14.2 how to compute the signed volume V under a surface $z = f(x, y)$ over a region R: $V = \iint_R f(x, y)\, dA$. It follows naturally that if $f(x, y) \geq g(x, y)$ on R, then the **volume between $f(x, y)$ and $g(x, y)$ on R** is

$$V = \iint_R f(x, y)\, dA - \iint_R g(x, y)\, dA = \iint_R \big(f(x, y) - g(x, y)\big)\, dA.$$

Theorem 127   Volume Between Surfaces

Let f and g be continuous functions on a closed, bounded region R, where $f(x, y) \geq g(x, y)$ for all (x, y) in R. The volume V between f and g over R is

$$V = \iint_R \big(f(x, y) - g(x, y)\big)\, dA.$$

Example 1   Finding volume between surfaces

Find the volume of the space region bounded by the planes $z = 3x + y - 4$ and $z = 8 - 3x - 2y$ in the 1st octant. In Figure 14.36(a) the planes are drawn; in (b), only the defined region is given.

SOLUTION   We need to determine the region R over which we will integrate. To do so, we need to determine where the planes intersect. They have common z-values when $3x + y - 4 = 8 - 3x - 2y$. Applying a little algebra, we have:

$$3x + y - 4 = 8 - 3x - 2y$$
$$6x + 3y = 12$$
$$2x + y = 4$$

The planes intersect along the line $2x + y = 4$. Therefore the region R is bounded by $x = 0$, $y = 0$, and $y = 4 - 2x$; we can convert these bounds to integration bounds of $0 \leq x \leq 2$, $0 \leq y \leq 4 - 2x$. Thus

$$V = \iint_R \big(8 - 3x - 2y - (3x + y - 4)\big)\, dA$$
$$= \int_0^2 \int_0^{4-2x} \big(12 - 6x - 3y\big)\, dy\, dx$$
$$= 16\,\text{units}^3.$$

The volume between the surfaces is 16 cubic units.

(a)

(b)
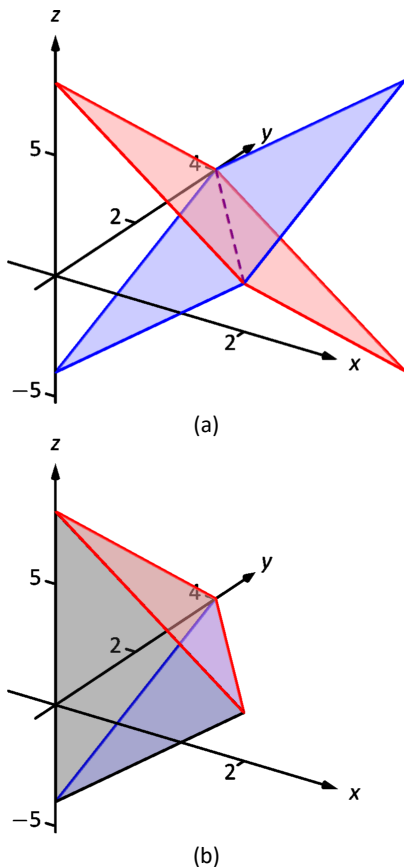
Figure 14.36: Finding the volume between the planes given in Example 1.

Notes:

In the preceding example, we found the volume by evaluating the integral

$$\int_0^2 \int_0^{4-2x} \big(8 - 3x - 2y - (3x + y - 4)\big)\, dy\, dx.$$

Note how we can rewrite the integrand as an integral, much as we did in Section 14.1:

$$8 - 3x - 2y - (3x + y - 4) = \int_{3x+y-4}^{8-3x-2y} dz.$$

Thus we can rewrite the double integral that finds volume as

$$\int_0^2 \int_0^{4-2x} \big(8 - 3x - 2y - (3x+y-4)\big)\, dy\, dx = \int_0^2 \int_0^{4-2x} \left(\int_{3x+y-4}^{8-3x-2y} dz\right) dy\, dx.$$

This no longer looks like a "double integral," but more like a "triple integral." Just as our first introduction to double integrals was in the context of finding the area of a plane region, our introduction into triple integrals will be in the context of finding the volume of a space region.

To formally find the volume of a closed, bounded region D in space, such as the one shown in Figure 14.37(a), we start with an approximation. Break D into n rectangular solids; the solids near the boundary of D may possibly not include portions of D and/or include extra space. In Figure 14.37(b), we zoom in on a portion of the boundary of D to show a rectangular solid that contains space not in D; as this is an approximation of the volume, this is acceptable and this error will be reduced as we shrink the size of our solids.

The volume ΔV_i of the ith solid D_i is $\Delta V_i = \Delta x_i \Delta y_i \Delta z_i$, where Δx_i, Δy_i and Δz_i give the dimensions of the rectangular solid in the x, y and z directions, respectively. By summing up the volumes of all n solids, we get an approximation of the volume V of D:

$$V \approx \sum_{i=1}^n \Delta V_i = \sum_{i=1}^n \Delta x_i \Delta y_i \Delta z_i.$$

Let $\|\Delta D\|$ represent the length of the longest diagonal of rectangular solids in the subdivision of D. As $\|\Delta D\| \to 0$, the volume of each solid goes to 0, as do each of Δx_i, Δy_i and Δz_i, for all i. Our calculus experience tells us that taking a limit as $\|\Delta D\| \to 0$ turns our approximation of V into an exact calculation of V. Before we state this result in a theorem, we use a definition to define some terms.

Notes:

(a)

(b)
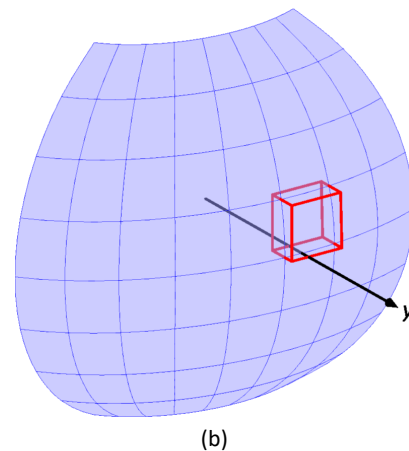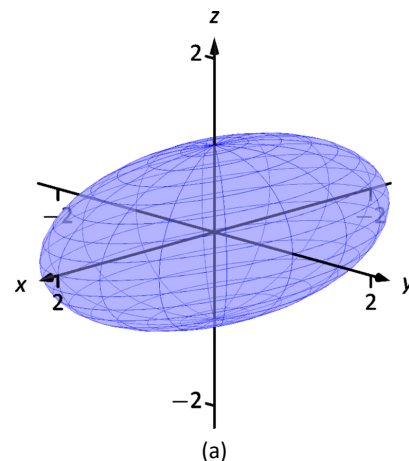
Figure 14.37: Approximating the volume of a region D in space.

Definition 108   Triple Integrals, Iterated Integration (Part I)

Let D be a closed, bounded region in space. Let a and b be real numbers, let $g_1(x)$ and $g_2(x)$ be continuous functions of x, and let $f_1(x, y)$ and $f_2(x, y)$ be continuous functions of x and y.

1. The volume V of D is denoted by a **triple integral**, $V = \iiint_D dV$.

2. The iterated integral $\displaystyle\int_a^b \int_{g_1(x)}^{g_2(x)} \int_{f_1(x,y)}^{f_2(x,y)} dz\, dy\, dx$ is evaluated as

$$\int_a^b \int_{g_1(x)}^{g_2(x)} \int_{f_1(x,y)}^{f_2(x,y)} dz\, dy\, dx = \int_a^b \int_{g_1(x)}^{g_2(x)} \left(\int_{f_1(x,y)}^{f_2(x,y)} dz \right) dy\, dx.$$

Evaluating the above iterated integral is **triple integration.**

Our informal understanding of the notation $\iiint_D dV$ is "sum up lots of little volumes over D," analogous to our understanding of $\iint_R dA$ and $\iint_R dm$. We now state the major theorem of this section.

Theorem 128   Triple Integration (Part I)

Let D be a closed, bounded region in space and let ΔD be any subdivision of D into n rectangular solids, where the ith subregion D_i has dimensions $\Delta x_i \times \Delta y_i \times \Delta z_i$ and volume ΔV_i.

1. The volume V of D is

$$V = \iiint_D dV = \lim_{\|\Delta D\| \to 0} \sum_{i=1}^n \Delta V_i = \lim_{\|\Delta D\| \to 0} \sum_{i=1}^n \Delta x_i \Delta y_i \Delta z_i.$$

2. If D is defined as the region bounded by the planes $x = a$ and $x = b$, the cylinders $y = g_1(x)$ and $y = g_2(x)$, and the surfaces $z = f_1(x, y)$ and $z = f_2(x, y)$, where $a < b$, $g_1(x) \le g_2(x)$ and $f_1(x, y) \le f_2(x, y)$ on D, then

$$\iiint_D dV = \int_a^b \int_{g_1(x)}^{g_2(x)} \int_{f_1(x,y)}^{f_2(x,y)} dz\, dy\, dx.$$

3. V can be determined using iterated integration with other orders of integration (there are 6 total), as long as D is defined by the region enclosed by a pair of planes, a pair of cylinders, and a pair of surfaces.

Notes:

We evaluated the area of a plane region R by iterated integration, where the bounds were "from curve to curve, then from point to point." Theorem 128 allows us to find the volume of a space region with an iterated integral with bounds "from surface to surface, then from curve to curve, then from point to point." In the iterated integral

$$\int_a^b \int_{g_1(x)}^{g_2(x)} \int_{f_1(x,y)}^{f_2(x,y)} dz\, dy\, dx,$$

the bounds $a \le x \le b$ and $g_1(x) \le y \le g_2(x)$ define a region R in the x-y plane over which the region D exists in space. However, these bounds are also defining surfaces in space; $x = a$ is a plane and $y = g_1(x)$ is a cylinder. The combination of these 6 surfaces enclose, and define, D.

Watch the video:
Triple Integrals at
`https://youtu.be/zFy-OpajEtA`

Examples will help us understand triple integration, including integrating with various orders of integration.

Example 2   Finding the volume of a space region with triple integration

Find the volume of the space region in the 1st octant bounded by the plane $z = 2 - y/3 - 2x/3$, shown in Figure 14.38(a), using the order of integration $dz\, dy\, dx$. Set up the triple integrals that give the volume in the other 5 orders of integration.

SOLUTION   Starting with the order of integration $dz\, dy\, dx$, we need to first find bounds on z. The region D is bounded below by the plane $z = 0$ (because we are restricted to the first octant) and above by $z = 2 - y/3 - 2x/3$; $0 \le z \le 2 - y/3 - 2x/3$.

To find the bounds on y and x, we "collapse" the region onto the x-y plane, giving the triangle shown in Figure 14.38(b). (We know the equation of the line $y = 6 - 2x$ in two ways. First, by setting $z = 0$, we have $0 = 2 - y/3 - 2x/3 \Rightarrow y = 6 - 2x$. Secondly, we know this is going to be a straight line between the points $(3, 0)$ and $(0, 6)$ in the x-y plane.)

Notes:

(a)

(b)
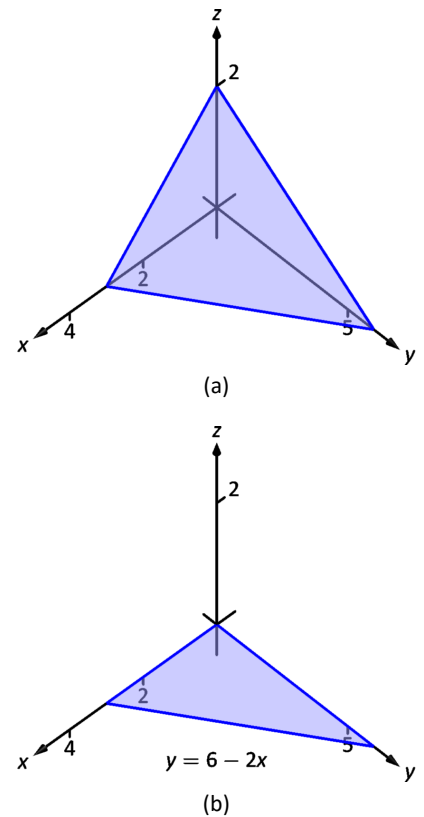
Figure 14.38: The region D used in Example 2 in (a); in (b), the region found by collapsing D onto the x-y plane.

We define that region R, in the integration order of $dy\,dx$, with bounds $0 \le y \le 6 - 2x$ and $0 \le x \le 3$. Thus the volume V of the region D is:

$$V = \iiint_D dV$$

$$= \int_0^3 \int_0^{6-2x} \int_0^{2-\frac{1}{3}y-\frac{2}{3}x} dz\,dy\,dz$$

$$= \int_0^3 \int_0^{6-2x} \left(\int_0^{2-\frac{1}{3}y-\frac{2}{3}x} dz \right) dy\,dz$$

$$= \int_0^3 \int_0^{6-2x} z \Big|_0^{2-\frac{1}{3}y-\frac{2}{3}x} dy\,dz$$

$$= \int_0^3 \int_0^{6-2x} \left(2 - \frac{1}{3}y - \frac{2}{3}x \right) dy\,dz.$$

From this step on, we are evaluating a double integral as done many times before. We skip these steps and give the final volume,

$$= 6 \text{units}^3.$$

The order $dz\,dx\,dy$:

Now consider the volume using the order of integration $dz\,dx\,dy$. The bounds on z are the same as before, $0 \le z \le 2 - y/3 - 2x/3$. Collapsing the space region on the x-y plane as shown in Figure 14.38(b), we now describe this triangle with the order of integration $dx\,dy$. This gives bounds $0 \le x \le 3 - y/2$ and $0 \le y \le 6$. Thus the volume is given by the triple integral

$$V = \int_0^6 \int_0^{3-\frac{1}{2}y} \int_0^{2-\frac{1}{3}y-\frac{2}{3}x} dz\,dx\,dy.$$

The order $dx\,dy\,dz$:

Following our "surface to surface..." strategy, we need to determine the *x-surfaces* that bound our space region. To do so, approach the region "from behind," in the direction of increasing x. The first surface we hit as we enter the region is the y-z plane, defined by $x = 0$. We come out of the region at the plane $z = 2 - y/3 - 2x/3$; solving for x, we have $x = 3 - y/2 - 3z/2$. Thus the bounds on x are: $0 \le x \le 3 - y/2 - 3z/2$.

Now collapse the space region onto the y-z plane, as shown in Figure 14.39(a). (Again, we find the equation of the line $z = 2 - y/3$ by setting $x = 0$ in the equation $x = 3 - y/2 - 3z/2$.) We need to find bounds on this region with the order

Notes:

dy dz. The *curves* that bound *y* are $y = 0$ and $y = 6 - 3z$; the *points* that bound *z* are 0 and 2. Thus the triple integral giving volume is:

$$0 \leq x \leq 3 - y/2 - 3z/2$$
$$0 \leq y \leq 6 - 3z \qquad \Rightarrow \qquad \int_0^2 \int_0^{6-3z} \int_0^{3-y/2-3z/2} dx \, dy \, dz.$$
$$0 \leq z \leq 2$$

The order *dx dz dy*:

The *x*-bounds are the same as the order above. We now consider the triangle in Figure 14.39(a) and describe it with the order *dz dy*: $0 \leq z \leq 2 - y/3$ and $0 \leq y \leq 6$. Thus the volume is given by:

$$0 \leq x \leq 3 - y/2 - 3z/2$$
$$0 \leq z \leq 2 - y/3 \qquad \Rightarrow \qquad \int_0^6 \int_0^{2-y/3} \int_0^{3-y/2-3z/2} dx \, dz \, dy.$$
$$0 \leq y \leq 6$$

The order *dy dz dx*:

We now need to determine the *y*-surfaces that determine our region. Approaching the space region from "behind" and moving in the direction of increasing *y*, we first enter the region at $y = 0$, and exit along the plane $z = 2 - y/3 - 2x/3$. Solving for *y*, this plane has equation $y = 6 - 2x - 3z$. Thus *y* has bounds $0 \leq y \leq 6 - 2x - 3z$.

Now collapse the region onto the *x-z* plane, as shown in Figure 14.39(b). The curves bounding this triangle are $z = 0$ and $z = 2 - 2x/3$; *x* is bounded by the points $x = 0$ to $x = 3$. Thus the triple integral giving volume is:

$$0 \leq y \leq 6 - 2x - 3z$$
$$0 \leq z \leq 2 - 2x/3 \qquad \Rightarrow \qquad \int_0^3 \int_0^{2-2x/3} \int_0^{6-2x-3z} dy \, dz \, dx.$$
$$0 \leq x \leq 3$$

The order *dy dx dz*:

The *y*-bounds are the same as in the order above. We now determine the bounds of the triangle in Figure 14.39(b) using the order *dy dx dz*. We see *x* is bounded by $x = 0$ and $x = 3 - 3z/2$; *z* is bounded between $z = 0$ and $z = 2$. This leads to the triple integral:

$$0 \leq y \leq 6 - 2x - 3z$$
$$0 \leq x \leq 3 - 3z/2 \qquad \Rightarrow \qquad \int_0^2 \int_0^{3-3z/2} \int_0^{6-2x-3z} dy \, dx \, dz.$$
$$0 \leq z \leq 2$$
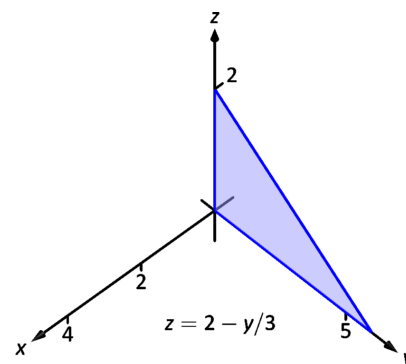
(a)
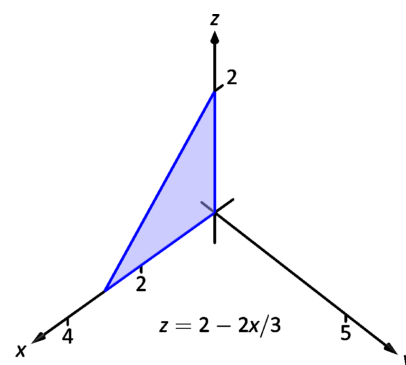
(b)

Figure 14.39: The region *D* in Example 2 is collapsed onto the *y-z* plane in (a); in (b), the region is collapsed onto the *x-z* plane.

Notes:

(a)

(b)

Figure 14.40: Finding the projections of the curve of intersection in Example 3.

This problem was long, but hopefully useful, demonstrating how to determine bounds with every order of integration to describe the region *D*. In practice, we only need 1, but being able to do them all gives us flexibility to choose the order that suits us best.

In the previous example, we collapsed the surface into the *x-y*, *x-z*, and *y-z* planes as we determined the "curve to curve, point to point" bounds of integration. Since the surface was a triangular portion of a plane, this collapsing, or *projecting*, was simple: the *projection* of a straight line in space onto a coordinate plane is a line.

The following example shows us how to do this when dealing with more complicated surfaces and curves.

Example 3   Finding the projection of a curve in space onto the coordinate planes

Consider the surfaces $z = 3 - x^2 - y^2$ and $z = 2y$, as shown in Figure 14.40(a). The curve of their intersection is shown, along with the projection of this curve into the coordinate planes, shown dashed. Find the equations of the projections into the coordinate planes.

SOLUTION   The two surfaces are $z = 3 - x^2 - y^2$ and $z = 2y$. To find where they intersect, it is natural to set them equal to each other: $3 - x^2 - y^2 = 2y$. This is an implicit function of *x* and *y* that gives all points (x, y) in the *x-y* plane where the *z* values of the two surfaces are equal.

We can rewrite this implicit function by completing the square:

$$3 - x^2 - y^2 = 2y \quad \Rightarrow \quad y^2 + 2y + x^2 = 3 \quad \Rightarrow \quad (y+1)^2 + x^2 = 4.$$

Thus in the *x-y* plane the projection of the intersection is a circle with radius 2, centered at $(0, -1)$.

To project onto the *x-z* plane, we do a similar procedure: find the *x* and *z* values where the *y* values on the surface are the same. We start by solving the equation of each surface for *y*. In this particular case, it works well to actually solve for y^2:

$z = 3 - x^2 - y^2 \quad \Rightarrow \quad y^2 = 3 - x^2 - z$
$z = 2y \quad \Rightarrow \quad y^2 = z^2/4.$

Thus we have (after again completing the square):

$$3 - x^2 - z = z^2/4 \quad \Rightarrow \quad \frac{(z+2)^2}{16} + \frac{x^2}{4} = 1,$$

and ellipse centered at $(0, -2)$ in the *x-z* plane with a major axis of length 8 and a minor axis of length 4.
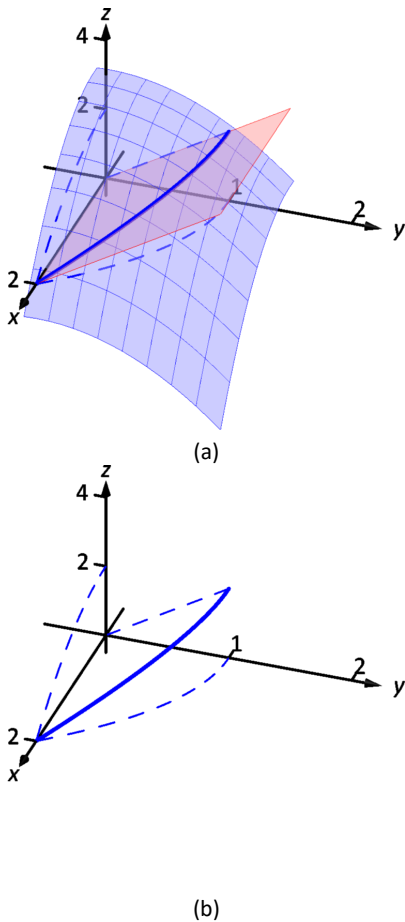
Notes:

Finally, to project the curve of intersection into the *y-z* plane, we solve equation for *x*. Since $z = 2y$ is a cylinder that lacks the variable *x*, it becomes our equation of the projection in the *y-z* plane.

All three projections are shown in Figure 14.40(b).

Example 4   **Finding the volume of a space region with triple integration**
Set up the triple integrals that find the volume of the space region *D* bounded by the surfaces $x^2 + y^2 = 1$, $z = 0$ and $z = -y$, as shown in Figure 14.41(a), with the orders of integration *dz dy dx*, *dy dx dz* and *dx dz dy*.

SOLUTION   The order *dz dy dx*:

The region *D* is bounded below by the plane $z = 0$ and above by the plane $z = -y$. The cylinder $x^2 + y^2 = 1$ does not offer any bounds in the *z*-direction, as that surface is parallel to the *z*-axis. Thus $0 \le z \le -y$.

Collapsing the region into the *x-y* plane, we get part of the region bounded by the circle with equation $x^2 + y^2 = 1$ as shown in Figure 14.41(b). As a function of *x*, this half circle has equation $y = -\sqrt{1 - x^2}$. Thus *y* is bounded below by $-\sqrt{1 - x^2}$ and above by $y = 0$: $-\sqrt{1 - x^2} \le y \le 0$. The *x* bounds of the half circle are $-1 \le x \le 1$. All together, the bounds of integration and triple integral are as follows:

$$\begin{array}{c} 0 \le z \le -y \\ -\sqrt{1 - x^2} \le y \le 0 \\ -1 \le x \le 1 \end{array} \quad \Rightarrow \quad \int_{-1}^{1} \int_{-\sqrt{1-x^2}}^{0} \int_{0}^{-y} dz\, dy\, dx.$$

We evaluate this triple integral:

$$\int_{-1}^{1} \int_{-\sqrt{1-x^2}}^{0} \int_{0}^{-y} dz\, dy\, dx = \int_{-1}^{1} \int_{-\sqrt{1-x^2}}^{0} (-y)\, dy\, dx$$

$$= \int_{-1}^{1} \left(-\frac{1}{2}y^2 \right) \Big|_{-\sqrt{1-x^2}}^{0} dx$$

$$= \int_{-1}^{1} \frac{1}{2}(1 - x^2)\, dx$$

$$= \left(\frac{1}{2} \left(x - \frac{1}{3}x^3 \right) \right) \Big|_{-1}^{1}$$

$$= \frac{2}{3} \text{units}^3.$$

With the order *dy dx dz*:
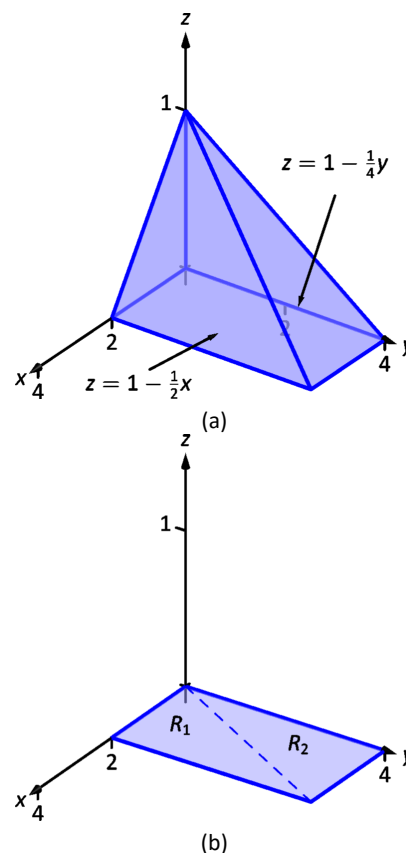

$z = -y$

$x^2 + y^2 = 1$

(a)

(b)
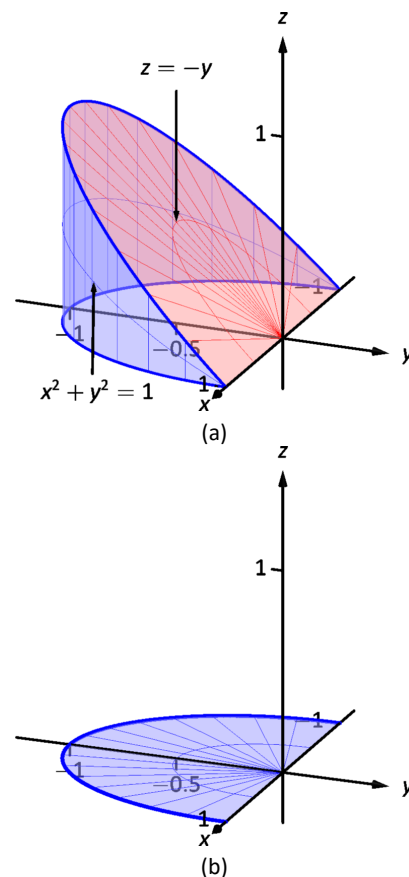
Figure 14.41: The region *D* in Example 4 is shown in (a); in (b), it is collapsed onto the *x-y* plane.
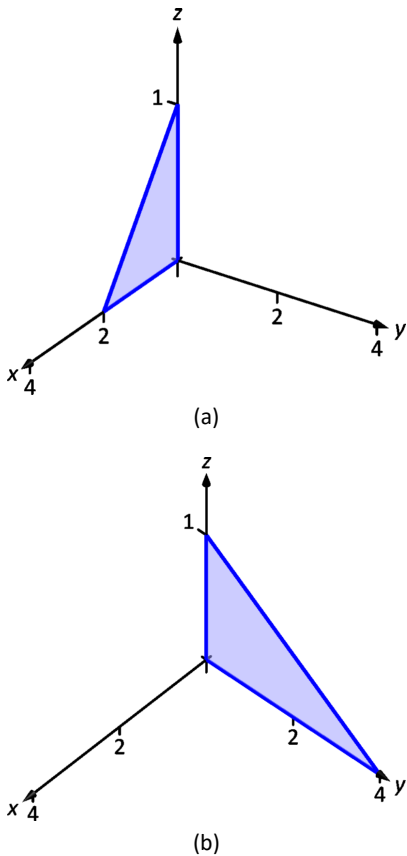

Notes:

(a)

(b)
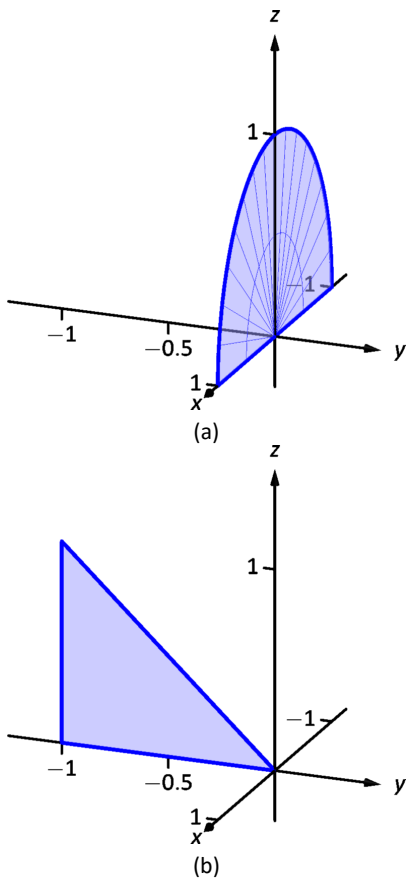
Figure 14.42: The region D in Example 4 is shown collapsed onto the x-z plane in (a); in (b), it is collapsed onto the y-z plane.

The region is bounded "below" in the y-direction by the surface $x^2 + y^2 = 1 \Rightarrow y = -\sqrt{1 - x^2}$ and "above" by the surface $y = -z$. Thus the y bounds are $-\sqrt{1 - x^2} \leq y \leq -z$.

Collapsing the region onto the x-z plane gives the region shown in Figure 14.42(a); this half disk is bounded by a circle with equation $x^2 + z^2 = 1$. (We find this curve by solving each surface for y^2, then setting them equal to each other. We have $y^2 = 1 - x^2$ and $y = -z \Rightarrow y^2 = z^2$. Thus $x^2 + z^2 = 1$.) It is bounded below by $x = -\sqrt{1 - z^2}$ and above by $x = \sqrt{1 - z^2}$, where z is bounded by $0 \leq z \leq 1$. All together, we have:

$$
\begin{aligned}
-\sqrt{1 - x^2} &\leq y \leq -z \\
-\sqrt{1 - z^2} &\leq x \leq \sqrt{1 - z^2} \\
0 &\leq z \leq 1
\end{aligned}
\qquad \Rightarrow \qquad
\int_0^1 \int_{-\sqrt{1-z^2}}^{\sqrt{1-z^2}} \int_{-\sqrt{1-x^2}}^{-z} dy\, dx\, dz.
$$

With the order $dx\, dz\, dy$:

D is bounded below by the surface $x = -\sqrt{1 - y^2}$ and above by $\sqrt{1 - y^2}$. We then collapse the region onto the y-z plane and get the triangle shown in Figure 14.42(b). (The hypotenuse is the line $z = -y$, just as the plane.) Thus z is bounded by $0 \leq z \leq -y$ and y is bounded by $-1 \leq y \leq 0$. This gives:

$$
\begin{aligned}
-\sqrt{1 - y^2} &\leq x \leq \sqrt{1 - y^2} \\
0 &\leq z \leq -y \\
-1 &\leq y \leq 0
\end{aligned}
\qquad \Rightarrow \qquad
\int_{-1}^0 \int_0^{-y} \int_{-\sqrt{1-y^2}}^{\sqrt{1-y^2}} dx\, dz\, dy.
$$

The following theorem states two things that should make "common sense" to us. First, using the triple integral to find volume of a region D should always return a positive number; we are computing *volume* here, not *signed volume*. Secondly, to compute the volume of a "complicated" region, we could break it up into subregions and compute the volumes of each subregion separately, summing them later to find the total volume.

Theorem 129   Properties of Triple Integrals

Let D be a closed, bounded region in space, and let D_1 and D_2 be non-overlapping regions such that $D = D_1 \bigcup D_2$.

1. $\displaystyle\iiint_D dV \geq 0$

2. $\displaystyle\iiint_D dV = \iiint_{D_1} dV + \iiint_{D_2} dV.$

Notes:

We use this latter property in the next example.

Example 5   **Finding the volume of a space region with triple integration**
Find the volume of the space region D bounded by the coordinate planes, $z = 1 - x/2$ and $z = 1 - y/4$, as shown in Figure 14.43(a). Set up the triple integrals that find the volume of D in all 6 orders of integration.

SOLUTION   Following the bounds–determining strategy of "surface to surface, curve to curve, and point to point," we can see that the most difficult orders of integration are the two in which we integrate with respect to z first, for there are two "upper" surfaces that bound D in the z-direction. So we start by noting that we have

$$0 \le z \le 1 - \frac{1}{2}x \quad \text{and} \quad 0 \le z \le 1 - \frac{1}{4}y.$$

We now collapse the region D onto the x-y axis, as shown in Figure 14.43(b). The boundary of D, the line from $(0, 0, 1)$ to $(2, 4, 0)$, is shown in part (b) of the figure as a dashed line; it has equation $y = 2x$. (We can recognize this in two ways: one, in collapsing the line from $(0, 0, 1)$ to $(2, 4, 0)$ onto the x-y plane, we simply ignore the z-values, meaning the line now goes from $(0, 0)$ to $(2, 4)$. Secondly, the two surfaces meet where $z = 1 - x/2$ is equal to $z = 1 - y/4$: thus $1 - x/2 = 1 - y/4 \Rightarrow y = 2x$.)
We use the second property of Theorem 129 to state that

$$\iiint_D dV = \iiint_{D_1} dV + \iiint_{D_2} dV,$$

where D_1 and D_2 are the space regions above the plane regions R_1 and R_2, respectively. Thus we can say

$$\iiint_D dV = \iint_{R_1} \left(\int_0^{1-x/2} dz \right) dA + \iint_{R_2} \left(\int_0^{1-y/4} dz \right) dA.$$

All that is left is to determine bounds of R_1 and R_2, depending on whether we are integrating with order $dx\,dy$ or $dy\,dx$. We give the final integrals here, leaving it to the reader to confirm these results.

dz dy dx:
$$\begin{array}{ll} 0 \le z \le 1 - x/2 & \quad 0 \le z \le 1 - y/4 \\ 0 \le y \le 2x & \quad 2x \le y \le 4 \\ 0 \le x \le 2 & \quad 0 \le x \le 2 \end{array}$$

$$\iiint_D dV = \int_0^2 \int_0^{2x} \int_0^{1-x/2} dz\,dy\,dx + \int_0^2 \int_{2x}^4 \int_0^{1-y/4} dz\,dy\,dx$$

Figure 14.43: The region D in Example 5 is shown in (a); in (b), it is collapsed onto the x-y plane.

Notes:

(a)

(b)

Figure 14.44: The region D in Example 5 is shown collapsed onto the x-z plane in (a); in (b), it is collapsed onto the y-z plane.

$dz\ dx\ dy$:

$$0 \le z \le 1 - x/2 \qquad\qquad 0 \le z \le 1 - y/4$$
$$y/2 \le x \le 2 \qquad\qquad\quad 0 \le x \le y/2$$
$$0 \le y \le 4 \qquad\qquad\qquad 0 \le y \le 4$$

$$\iiint_D dV = \int_0^4 \int_{y/2}^2 \int_0^{1-x/2} dz\ dx\ dy + \int_0^4 \int_0^{y/2} \int_0^{1-y/4} dz\ dx\ dy$$

The remaining four orders of integration do not require a sum of triple integrals. In Figure 14.44 we show D collapsed onto the other two coordinate planes. Using these graphs, we give the final orders of integration here, again leaving it to the reader to confirm these results.

$dy\ dx\ dz$:

$$0 \le y \le 4 - 4z$$
$$0 \le x \le 2 - 2z \qquad \Rightarrow \qquad \int_0^1 \int_0^{2-2z} \int_0^{4-4z} dy\ dx\ dz$$
$$0 \le z \le 1$$

$dy\ dz\ dx$:

$$0 \le y \le 4 - 4z$$
$$0 \le z \le 1 - x/2 \qquad \Rightarrow \qquad \int_0^2 \int_0^{1-x/2} \int_0^{4-4z} dy\ dx\ dz$$
$$0 \le x \le 2$$

$dx\ dy\ dz$:

$$0 \le x \le 2 - 2z$$
$$0 \le y \le 4 - 4z \qquad \Rightarrow \qquad \int_0^1 \int_0^{4-4z} \int_0^{2-2z} dx\ dy\ dz$$
$$0 \le z \le 1$$

$dx\ dz\ dy$:

$$0 \le x \le 2 - 2z$$
$$0 \le z \le 1 - y/4 \qquad \Rightarrow \qquad \int_0^4 \int_0^{1-y/4} \int_0^{2-2z} dx\ dz\ dy$$
$$0 \le y \le 4$$

We give one more example of finding the volume of a space region.

Example 6   **Finding the volume of a space region**
Set up a triple integral that gives the volume of the space region D bounded by $z = 2x^2 + 2$ and $z = 6 - 2x^2 - y^2$. These surfaces are plotted in Figure 14.45(a) and (b), respectively; the region D is shown in part (c) of the figure.

Notes:
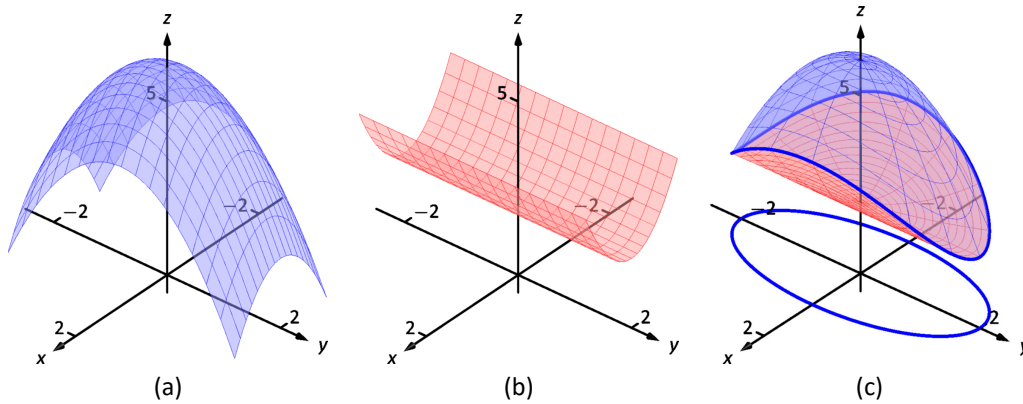
Figure 14.45: The region D is bounded by the surfaces shown in (a) and (b); D is shown in (c).

SOLUTION   The main point of this example is this: integrating with respect to z first is rather straightforward; integrating with respect to x first is not.

The order *dz dy dx*:

The bounds on z are clearly $2x^2 + 2 \leq z \leq 6 - 2x^2 - y^2$. Collapsing D onto the *x-y* plane gives the ellipse shown in Figure 14.45(c). The equation of this ellipse is found by setting the two surfaces equal to each other:

$$2x^2 + 2 = 6 - 2x^2 - y^2 \quad \Rightarrow \quad 4x^2 + y^2 = 4 \quad \Rightarrow \quad x^2 + \frac{y^2}{4} = 1.$$

We can describe this ellipse with the bounds

$$-\sqrt{4 - 4x^2} \leq y \leq \sqrt{4 - 4x^2} \quad \text{and} \quad -1 \leq x \leq 1.$$

Thus we find volume as

$$\begin{array}{c} 2x^2 + 2 \leq z \leq 6 - 2x^2 - y^2 \\ -\sqrt{4 - 4x^2} \leq y \leq \sqrt{4 - 4x^2} \\ -1 \leq x \leq 1 \end{array} \quad \Rightarrow \quad \int_{-1}^{1} \int_{-\sqrt{4-4x^2}}^{\sqrt{4-4x^2}} \int_{2x^2+2}^{6-2x^2-y^2} dz \, dy \, dx$$

The order *dy dz dx*:

Integrating with respect to y is not too difficult. Since the surface $z = 2x^2 + 2$ is a cylinder whose directrix is the *y*-axis, it does not create a border for y. The paraboloid $z = 6 - 2x^2 - y^2$ does; solving for y, we get the bounds

$$-\sqrt{6 - 2x^2 - z} \leq y \leq \sqrt{6 - 2x^2 - z}.$$

Notes:
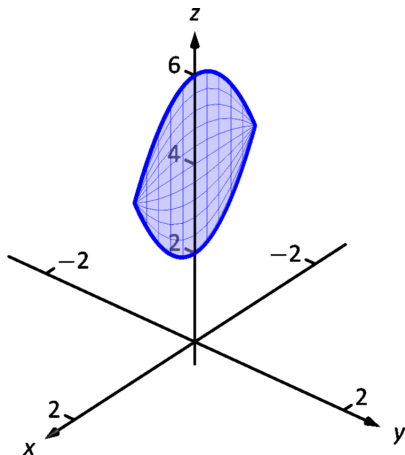
(a)
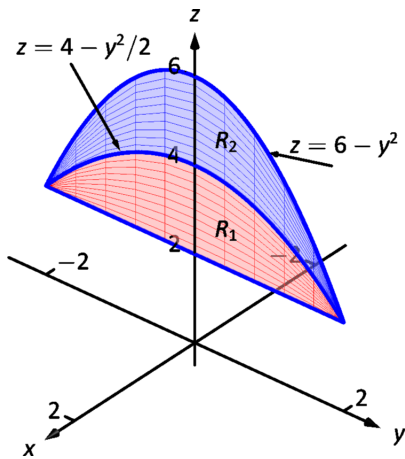
(b)

Figure 14.46: The region D in Example 6 is collapsed onto the x-z plane in (a); in (b), it is collapsed onto the y-z plane.

Collapsing D onto the x-z axes gives the region shown in Figure 14.46(a); the lower curve is from the cylinder, with equation $z = 2x^2 + 2$. The upper curve is from the paraboloid; with $y = 0$, the curve is $z = 6 - 2x^2$. Thus bounds on z are $2x^2 + 2 \leq z \leq 6 - 2x^2$; the bounds on x are $-1 \leq x \leq 1$. Thus we have:

$$-\sqrt{6 - 2x^2 - z} \leq y \leq \sqrt{6 - 2x^2 - z}$$
$$2x^2 + 2 \leq z \leq 6 - 2x^2 \quad \Rightarrow \quad \int_{-1}^{1} \int_{2x^2+2}^{6-2x^2} \int_{-\sqrt{6-2x^2-z}}^{\sqrt{6-2x^2-z}} dy\, dz\, dx.$$
$$-1 \leq x \leq 1$$

The order $dx\, dz\, dy$:

This order takes more effort as D must be split into two subregions. The two surfaces create two sets of upper/lower bounds in terms of x; the cylinder creates bounds

$$-\sqrt{z/2 - 1} \leq x \leq \sqrt{z/2 - 1}$$

for region D_1 and the paraboloid creates bounds

$$-\sqrt{3 - y^2/2 - z^2/2} \leq x \leq \sqrt{3 - y^2/2 - z^2/2}$$

for region D_2.

Collapsing D onto the y-z axes gives the regions shown in Figure 14.46(b). We find the equation of the curve $z = 4 - y^2/2$ by noting that the equation of the ellipse seen in Figure 14.45(c) has equation

$$x^2 + y^2/4 = 1 \quad \Rightarrow \quad x = \sqrt{1 - y^2/4}.$$

Substitute this expression for x in either surface equation, $z = 6 - 2x^2 - y^2$ or $z = 2x^2 + 2$. In both cases, we find

$$z = 4 - \frac{1}{2}y^2.$$

Region R_1, corresponding to D_1, has bounds

$$2 \leq z \leq 4 - y^2/2, \quad -2 \leq y \leq 2$$

and region R_2, corresponding to D_2, has bounds

$$4 - y^2/2 \leq z \leq 6 - y^2, \quad -2 \leq y \leq 2.$$

Thus the volume of D is given by:

$$\int_{-2}^{2} \int_{2}^{4-y^2/2} \int_{-\sqrt{z/2-1}}^{\sqrt{z/2-1}} dx\, dz\, dy + \int_{-2}^{2} \int_{4-y^2/2}^{6-y^2} \int_{-\sqrt{3-y^2/2-z^2/2}}^{\sqrt{3-y^2/2-z^2/2}} dx\, dz\, dy.$$

Notes:

If all one wanted to do in Example 6 was find the volume of the region D, one would have likely stopped at the first integration setup (with order *dz dy dx*) and computed the volume from there. However, we included the other two methods (1) to show that it could be done, "messy" or not, and (2) because sometimes we "have" to use a less desirable order of integration in order to actually integrate.

Triple Integration and Functions of Three Variables

There are uses for triple integration beyond merely finding volume, just as there are uses for integration beyond "area under the curve." These uses start with understanding how to integrate functions of three variables, which is effectively no different than integrating functions of two variables. This leads us to a definition, followed by an example.

Definition 109   Iterated Integration, (Part II)

Let D be a closed, bounded region in space, over which $g_1(x)$, $g_2(x)$, $f_1(x, y)$, $f_2(x, y)$ and $h(x, y, z)$ are all continuous, and let a and b be real numbers.

The **iterated integral** $\displaystyle\int_a^b \int_{g_1(x)}^{g_2(x)} \int_{f_1(x,y)}^{f_2(x,y)} h(x, y, z)\, dz\, dy\, dx$ is evaluated as

$$\int_a^b \int_{g_1(x)}^{g_2(x)} \int_{f_1(x,y)}^{f_2(x,y)} h(x, y, z)\, dz\, dy\, dx = \int_a^b \int_{g_1(x)}^{g_2(x)} \left(\int_{f_1(x,y)}^{f_2(x,y)} h(x, y, z)\, dz \right) dy\, dx.$$

Example 7   Evaluating a triple integral of a function of three variables

Evaluate $\displaystyle\int_0^1 \int_{x^2}^x \int_{x^2-y}^{2x+3y} \left(xy + 2xz \right) dz\, dy\, dx$.

SOLUTION   We evaluate this integral according to Definition 109.

Notes:

$$\int_0^1 \int_{x^2}^x \int_{x^2-y}^{2x+3y} \left(xy + 2xz\right) dz\, dy\, dx$$

$$= \int_0^1 \int_{x^2}^x \left(\int_{x^2-y}^{2x+3y} \left(xy + 2xz\right) dz\right) dy\, dx$$

$$= \int_0^1 \int_{x^2}^x \left(\left(xyz + xz^2\right)\Big|_{x^2-y}^{2x+3y}\right) dy\, dx$$

$$= \int_0^1 \int_{x^2}^x \left(xy(2x+3y) + x(2x+3y)^2 - \left(xy(x^2-y) + x(x^2-y)^2\right)\right) dy\, dx$$

$$= \int_0^1 \int_{x^2}^x \left(-x^5 + x^3y + 4x^3 + 14x^2y + 12xy^2\right) dy\, dx.$$

We continue as we have in the past, showing fewer steps.

$$= \int_0^1 \left(-\frac{7}{2}x^7 - 8x^6 - \frac{7}{2}x^5 + 15x^4\right) dx$$

$$= \frac{281}{336}.$$

We now know *how* to evaluate a triple integral of a function of three variables; we do not yet understand what it *means*. We build up this understanding in a way very similar to how we have understood integration and double integration.

Let $h(x, y, z)$ be a continuous function of three variables, defined over some space region D. We can partition D into n rectangular–solid subregions, each with dimensions $\Delta x_i \times \Delta y_i \times \Delta z_i$. Let (x_i, y_i, z_i) be some point in the ith subregion, and consider the product $h(x_i, y_i, z_i)\Delta x_i\Delta y_i\Delta z_i$. It is the product of a function value (that's the $h(x_i, y_i, z_i)$ part) and a small volume ΔV_i (that's the $\Delta x_i\Delta y_i\Delta z_i$ part). One of the simplest understanding of this type of product is when h describes the density of an object, for then $h \times$ volume $=$ mass.

We can sum up all n products over D. Again letting $\|\Delta D\|$ represent the length of the longest diagonal of the n rectangular solids in the partition, we can take the limit of the sums of products as $\|\Delta D\| \to 0$. That is, we can find

$$S = \lim_{\|\Delta D\| \to 0} \sum_{i=1}^n h(x_i, y_i, z_i)\Delta V_i = \lim_{\|\Delta D\| \to 0} \sum_{i=1}^n h(x_i, y_i, z_i)\Delta x_i\Delta y_i\Delta z_i.$$

While this limit has lots of interpretations depending on the function h, in the case where h describes density, S is the total mass of the object described by the region D.

Notes:

We now use the above limit to define the **triple integral**, give a theorem that relates triple integrals to iterated iteration, followed by the application of triple integrals to find the centers of mass of solid objects.

Definition 110   Triple Integral

Let $w = h(x, y, z)$ be a continuous function over a closed, bounded space region D, and let ΔD be any partition of D into n rectangular solids with volume ΔV_i. The **triple integral of h over D** is

$$\iiint_D h(x, y, z)\, dV = \lim_{\|\Delta D\| \to 0} \sum_{i=1}^{n} h(x_i, y_i, z_i) \Delta V_i.$$

The following theorem assures us that the above limit exists for continuous functions h and gives us a method of evaluating the limit.

Theorem 130   Triple Integration (Part II)

Let $w = h(x, y, z)$ be a continuous function over a closed, bounded space region D, and let ΔD be any partition of D into n rectangular solids with volume V_i.

1. The limit $\displaystyle \lim_{\|\Delta D\| \to 0} \sum_{i=1}^{n} h(x_i, y_i, z_i) \Delta V_i$ exists.

2. If D is defined as the region bounded by the planes $x = a$ and $x = b$, the cylinders $y = g_1(x)$ and $y = g_2(x)$, and the surfaces $z = f_1(x, y)$ and $z = f_2(x, y)$, where $a < b$, $g_1(x) \le g_2(x)$ and $f_1(x, y) \le f_2(x, y)$ on D, then

$$\iiint_D h(x, y, z)\, dV = \int_a^b \int_{g_1(x)}^{g_2(x)} \int_{f_1(x,y)}^{f_2(x,y)} h(x, y, z)\, dz\, dy\, dx.$$

We now apply triple integration to find the centers of mass of solid objects.

Mass and Center of Mass

One may wish to review Section 14.4 for a reminder of the relevant terms and concepts.

Notes:

Definition 111   Mass, Center of Mass of Solids

Let a solid be represented by a region D in space with variable density function $\delta(x, y, z)$.

1. The **mass** of the object is $M = \iiint_D dm = \iiint_D \delta(x, y, z)\, dV$.

2. The **moment about the x-y plane** is $M_{xy} = \iiint_D z\delta(x, y, z)\, dV$.

3. The **moment about the x-z plane** is $M_{xz} = \iiint_D y\delta(x, y, z)\, dV$.

4. The **moment about the y-z plane** is $M_{yz} = \iiint_D x\delta(x, y, z)\, dV$.

5. The **center of mass** of the object is

$$\left(\bar{x}, \bar{y}, \bar{z}\right) = \left(\frac{M_{yz}}{M}, \frac{M_{xz}}{M}, \frac{M_{xy}}{M}\right).$$

Example 8   Finding the center of mass of a solid

Find the mass and center of mass of the solid represented by the space region bounded by the coordinate planes and $z = 2 - y/3 - 2x/3$, shown in Figure 14.47, with constant density $\delta(x, y, z) = 3\text{g/cm}^3$. (Note: this space region was used in Example 2.)

SOLUTION   We apply Definition 111. In Example 2, we found bounds for the order of integration $dz\,dy\,dx$ to be $0 \le z \le 2 - y/3 - 2x/3$, $0 \le y \le 6 - 2x$ and $0 \le x \le 3$. We find the mass of the object:

$$M = \iiint_D \delta(x, y, z)\, dV$$
$$= \int_0^3 \int_0^{6-2x} \int_0^{2-y/3-2x/3} (3)\, dz\, dy\, dx$$
$$= 3 \int_0^3 \int_0^{6-2x} \int_0^{2-y/3-2x/3} dz\, dy\, dx$$
$$= 3(6) = 18\text{g}.$$

The evaluation of the triple integral is done in Example 2, so we skipped those steps above. Note how the mass of an object with constant density is simply "density × volume."

Figure 14.47: Finding the center of mass of this solid in Example 8.

Notes:

We now find the moments about the planes.

$$M_{xy} = \iiint_D 3z \, dV$$

$$= \int_0^3 \int_0^{6-2x} \int_0^{2-y/3-2x/3} \left(3z\right) dz \, dy \, dx$$

$$= \int_0^3 \int_0^{6-2x} \frac{3}{2}\left(2 - y/3 - 2x/3\right)^2 dy \, dx$$

$$= \int_0^3 -\frac{4}{9}(x-3)^3 \, dx$$

$$= 9.$$

We omit the steps of integrating to find the other moments.

$$M_{yz} = \iiint_D 3x \, dV = \frac{27}{2}.$$

$$M_{xz} = \iiint_D 3y \, dV = 27.$$

The center of mass is

$$(\bar{x}, \bar{y}, \bar{z}) = \left(\frac{27/2}{18}, \frac{27}{18}, \frac{9}{18}\right) = \left(0.75, 1.5, 0.5\right).$$

Example 9   **Finding the center of mass of a solid**
Find the center of mass of the solid represented by the region bounded by the planes $z = 0$ and $z = -y$ and the cylinder $x^2 + y^2 = 1$, shown in Figure 14.48, with density function $\delta(x, y, z) = 10 + x^2 + 5y - 5z$. (Note: this space region was used in Example 4.)

     Solution   As we start, consider the density function. It is symmetric about the y-z plane, and the farther one moves from this plane, the denser the object is. The symmetry indicates that $\bar{x}$ should be 0.

     As one moves away from the origin in the y or z directions, the object becomes less dense, though there is more volume in these regions.

     Though none of the integrals needed to compute the center of mass are particularly hard, they do require a number of steps. We emphasize here the importance of knowing how to set up the proper integrals; in complex situations we can appeal to technology for a good approximation, if not the exact answer. We use the order of integration $dz \, dy \, dx$, using the bounds found in Example 4.
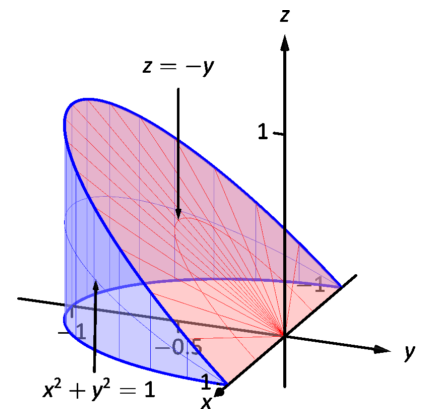
Figure 14.48: Finding the center of mass of this solid in Example 9.

Notes:

(As these are the same for all four triple integrals, we explicitly show the bounds only for *M*.)

$$M = \iiint_D \left(10 + x^2 + 5y - 5z\right) dV$$

$$= \int_{-1}^{1} \int_{-\sqrt{1-x^2}}^{0} \int_{0}^{-y} \left(10 + x^2 + 5y - 5z\right) dV$$

$$= \frac{64}{5} - \frac{15\pi}{16} \approx 3.855.$$

$$M_{yz} = \iiint_D x\left(10 + x^2 + 5y - 5z\right) dV = 0.$$

$$M_{xz} = \iiint_D y\left(10 + x^2 + 5y - 5z\right) dV = 2 - \frac{61\pi}{48} \approx -1.99.$$

$$M_{xy} = \iiint_D z\left(10 + x^2 + 5y - 5z\right) dV = \frac{61\pi}{96} - \frac{10}{9} \approx 0.885.$$

Note how $M_{yz} = 0$, as expected. The center of mass is

$$\left(\bar{x}, \bar{y}, \bar{z}\right) = \left(0, \frac{-1.99}{3.855}, \frac{0.885}{3.855}\right) \approx \left(0, -0.516, 0.230\right).$$

As stated before, there are many uses for triple integration beyond finding volume. When $h(x, y, z)$ describes a rate of change function over some space region *D*, then $\iiint_D h(x, y, z)\, dV$ gives the total change over *D*. Our one specific example of this was computing mass; a density function is simply a "rate of mass change per volume" function. Integrating density gives total mass.

While knowing *how to integrate* is important, it is arguably much more important to know *how to set up* integrals. It takes skill to create a formula that describes a desired quantity; modern technology is very useful in evaluating these formulas quickly and accurately.

Notes:

Exercises 14.6

Terms and Concepts

1. The strategy for establishing bounds for triple integrals is "_____ to _____, _____ to _____ and _____ to _____."

2. Give an informal interpretation of what "$\iiint_D dV$" means.

3. Give two uses of triple integration.

4. If an object has a constant density δ and a volume V, what is its mass?

Problems

In Exercises 5–8, two surfaces $f_1(x, y)$ and $f_2(x, y)$ and a region R in the x, y plane are given. Set up and evaluate the double integral that finds the volume between these surfaces over R.

5. $f_1(x, y) = 8 - x^2 - y^2, f_2(x, y) = 2x + y$;
   R is the square with corners $(-1, -1)$ and $(1, 1)$.

6. $f_1(x, y) = x^2 + y^2, f_2(x, y) = -x^2 - y^2$;
   R is the square with corners $(0, 0)$ and $(2, 3)$.

7. $f_1(x, y) = \sin x \cos y, f_2(x, y) = \cos x \sin y + 2$;
   R is the triangle with corners $(0, 0)$, $(\pi, 0)$ and (π, π).

8. $f_1(x, y) = 2x^2 + 2y^2 + 3, f_2(x, y) = 6 - x^2 - y^2$;
   R is the disk $x^2 + y^2 \le 1$.

In Exercises 9–16, a domain D is described by its bounding surfaces, along with a graph. Set up the triple integrals that give the volume of D in all 6 orders of integration, and find the volume of D by evaluating the indicated triple integral.

9. D is bounded by the coordinate planes and $z = 2 - 2x/3 - 2y$.
   Evaluate the triple integral with order $dz\, dy\, dx$.

$z = 2 - \frac{2}{3}x - 2y$
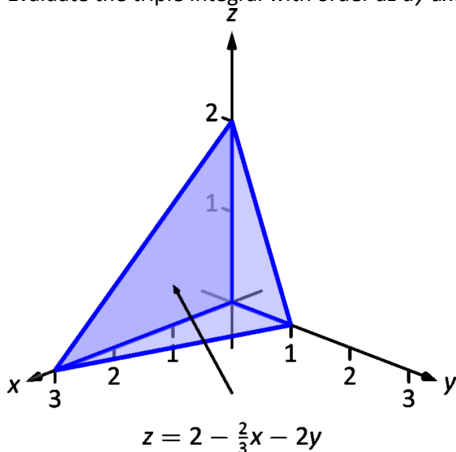
10. D is bounded by the planes $y = 0$, $y = 2$, $x = 1$, $z = 0$ and $z = (3 - x)/2$.
    Evaluate the triple integral with order $dx\, dy\, dz$.

$z = \frac{1}{2}(3 - x)$
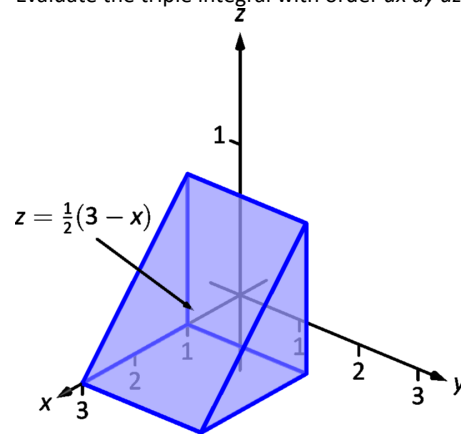
11. D is bounded by the planes $x = 0$, $x = 2$, $z = -y$ and by $z = y^2/2$.
    Evaluate the triple integral with the order $dy\, dz\, dx$.

$z = \frac{1}{2}y^2$

$z = -y$
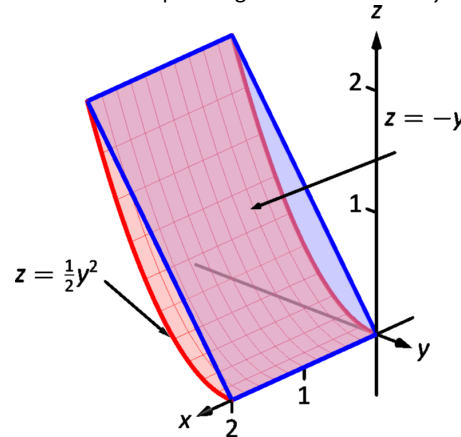
12. D is bounded by the planes $z = 0$, $y = 9$, $x = 0$ and by $z = \sqrt{y^2 - 9x^2}$.
    Do not evaluate any triple integral.

$z = \sqrt{y^2 - 9x^2}$
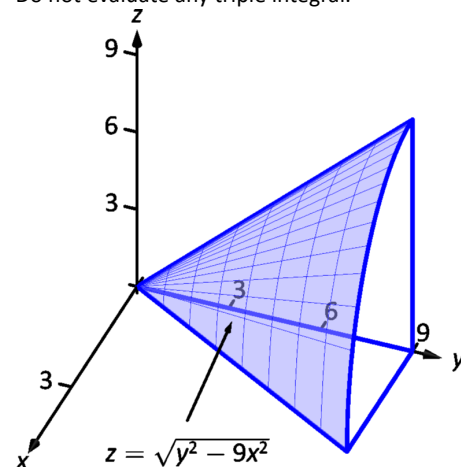
13. *D* is bounded by the planes $x = 2$, $y = 1$, $z = 0$ and $z = 2x + 4y - 4$.
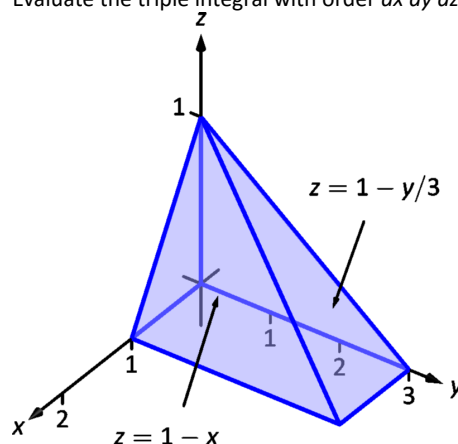Evaluate the triple integral with the order *dx dy dz*.
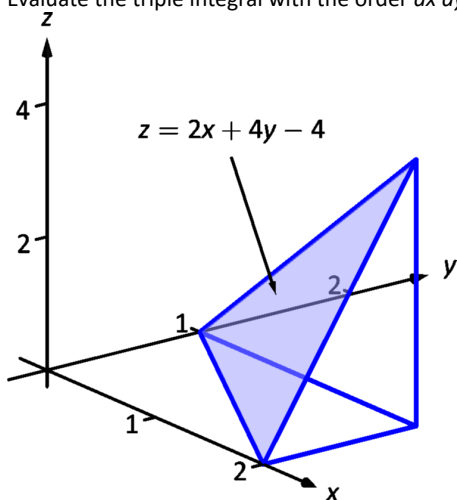
14. *D* is bounded by the plane $z = 2y$ and by $y = 4 - x^2$.
Evaluate the triple integral with the order *dz dy dx*.
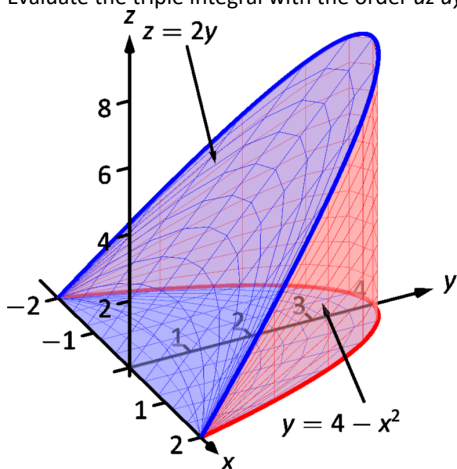
15. *D* is bounded by the coordinate planes and by $y = 1 - x^2$ and $y = 1 - z^2$.
Do not evaluate any triple integral. Which order is easier to evaluate: *dz dy dx* or *dy dz dx*? Explain why.
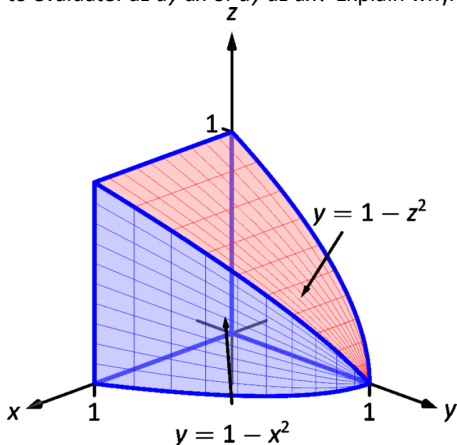
16. *D* is bounded by the coordinate planes and by $z = 1 - y/3$ and $z = 1 - x$.
Evaluate the triple integral with order *dx dy dz*.

In Exercises 17–20, evaluate the triple integral.

17. $\displaystyle\int_{-\pi/2}^{\pi/2} \int_0^\pi \int_0^\pi \left(\cos x \sin y \sin z \right) \, dz \, dy \, dx$

18. $\displaystyle\int_0^1 \int_0^x \int_0^{x+y} \left(x + y + z \right) \, dz \, dy \, dx$

19. $\displaystyle\int_0^\pi \int_0^1 \int_0^z \left(\sin(yz) \right) \, dx \, dy \, dz$

20. $\displaystyle\int_\pi^{\pi^2} \int_x^{x^3} \int_{-y^2}^{y^2} \left(z \frac{x^2 y + y^2 x}{e^{x^2 + y^2}} \right) \, dz \, dy \, dx$

In Exercises 21–24, find the center of mass of the solid represented by the indicated space region *D* with density function $\delta(x, y, z)$.

21. *D* is bounded by the coordinate planes and $z = 2 - 2x/3 - 2y$; $\quad \delta(x, y, z) = 10$g/cm^3.
(Note: this is the same region as used in Exercise 9.)

22. *D* is bounded by the planes $y = 0$, $y = 2$, $x = 1$, $z = 0$ and $z = (3 - x)/2$; $\quad \delta(x, y, z) = 2$g/cm^3.
(Note: this is the same region as used in Exercise 10.)

23. *D* is bounded by the planes $x = 2$, $y = 1$, $z = 0$ and $z = 2x + 4y - 4$; $\quad \delta(x, y, z) = x^2$lb/in^3.
(Note: this is the same region as used in Exercise 13.)

24. *D* is bounded by the plane $z = 2y$ and by $y = 4 - x^2$. $\delta(x, y, z) = y^2$lb/in^3.
(Note: this is the same region as used in Exercise 14.)

920

14.7   Change of Variables in Multiple Integrals

Given the difficulty of evaluating multiple integrals, the reader may be wondering if it is possible to simplify those integrals using a suitable substitution for the variables. The answer is yes, though it is a bit more complicated than the substitution method which you learned in single-variable calculus.

Recall that if you are given, for example, the definite integral

$$\int_1^2 x^3 \sqrt{x^2 - 1} \, dx,$$

then you would make the substitution

$$u = x^2 - 1 \Rightarrow x^2 = u + 1$$
$$du = 2x \, dx$$

which changes the limits of integration

$$x = 1 \Rightarrow u = 0$$
$$x = 2 \Rightarrow u = 3$$

so that we get

$$\begin{aligned}
\int_1^2 x^3 \sqrt{x^2 - 1} \, dx &= \int_1^2 \frac{1}{2} x^2 \cdot 2x \sqrt{x^2 - 1} \, dx \\
&= \int_0^3 \frac{1}{2} (u + 1) \sqrt{u} \, du \\
&= \frac{1}{2} \int_0^3 \left(u^{3/2} + u^{1/2} \right) \, du \\
&= \frac{14 \sqrt{3}}{5}.
\end{aligned}$$

Let us take a different look at what happened when we did that substitution, which will give some motivation for how substitution works in multiple integrals. First, we let $u = x^2 - 1$. On the interval of integration $[1, 2]$, the function $x \mapsto x^2 - 1$ is strictly increasing (and maps $[1, 2]$ onto $[0, 3]$) and hence has an inverse function (defined on the interval $[0, 3]$). That is, on $[0, 3]$ we can define x as a function of u, namely

$$x = g(u) = \sqrt{u + 1}.$$

Then substituting that expression for x into the function $f(x) = x^3 \sqrt{x^2 - 1}$ gives

$$f(x) = f(g(u)) = (u + 1)^{3/2} \sqrt{u},$$

Notes:

and we see that

$$\frac{dx}{du} = g'(u) \Rightarrow dx = g'(u)\, du$$

$$dx = \frac{1}{2}(u+1)^{-1/2}\, du,$$

so since

$$g(0) = 1 \Rightarrow 0 = g^{-1}(1)$$
$$g(3) = 2 \Rightarrow 3 = g^{-1}(2)$$

then performing the substitution as we did earlier gives

$$\int_1^2 f(x)\, dx = \int_1^2 x^3 \sqrt{x^2 - 1}\, dx$$

$$= \int_0^3 \frac{1}{2}(u+1)\sqrt{u}\, du, \text{ which can be written as}$$

$$= \int_0^3 (u+1)^{3/2} \sqrt{u} \cdot \frac{1}{2}(u+1)^{-1/2}\, du, \text{ which means}$$

$$\int_1^2 f(x)\, dx = \int_{g^{-1}(1)}^{g^{-1}(2)} f(g(u))\, g'(u)\, du.$$

In general, if $x = g(u)$ is a one-to-one, differentiable function from an interval $[c, d]$ (which you can think of as being on the "u-axis") onto an interval $[a, b]$ (on the x-axis), which means that $g'(u) \neq 0$ on the interval (c, d), so that $a = g(c)$ and $b = g(d)$, then $c = g^{-1}(a)$ and $d = g^{-1}(b)$, and

$$\int_a^b f(x)\, dx = \int_{g^{-1}(a)}^{g^{-1}(b)} f(g(u))\, g'(u)\, du.$$

This is called the *change of variable* formula for integrals of single-variable functions, and it is what you were implicitly using when doing integration by substitution. This formula turns out to be a special case of a more general formula which can be used to evaluate multiple integrals. We will state the formulas for double and triple integrals involving real-valued functions of two and three variables, respectively. We will assume that all the functions involved are continuously differentiable and that the regions and solids involved all have "reasonable" boundaries. The proof of the following theorem is beyond the scope of the text.

Notes:

Theorem 131   Change of Variables Formula for Multiple Integrals
Let $x = x(u, v)$ and $y = y(u, v)$ define a one-to-one mapping of a region R' in the uv-plane onto a region R in the xy-plane such that the determinant

$$J(u, v) = \begin{vmatrix} \dfrac{\partial x}{\partial u} & \dfrac{\partial x}{\partial v} \\[2mm] \dfrac{\partial y}{\partial u} & \dfrac{\partial y}{\partial v} \end{vmatrix} \qquad (14.2)$$

is never 0 in R'. Then

$$\iint_R f(x, y)\, dA(x, y) = \iint_{R'} f(x(u, v), y(u, v))\, |J(u, v)|\, dA(u, v). \quad (14.3)$$

We use the notation $dA(x, y)$ and $dA(u, v)$ to denote the area element in the (x, y) and (u, v) coordinates, respectively.
Similarly, if $x = x(u, v, w)$, $y = y(u, v, w)$ and $z = z(u, v, w)$ define a one-to-one mapping of a solid S' in uvw-space onto a solid S in xyz-space such that the determinant

$$J(u, v, w) = \begin{vmatrix} \dfrac{\partial x}{\partial u} & \dfrac{\partial x}{\partial v} & \dfrac{\partial x}{\partial w} \\[2mm] \dfrac{\partial y}{\partial u} & \dfrac{\partial y}{\partial v} & \dfrac{\partial y}{\partial w} \\[2mm] \dfrac{\partial z}{\partial u} & \dfrac{\partial z}{\partial v} & \dfrac{\partial z}{\partial w} \end{vmatrix} \qquad (14.4)$$

is never 0 in S', then

$$\iiint_S f(x, y, z)\, dV(x, y, z) =$$

$$\iiint_{S'} f(x(u, v, w), y(u, v, w), z(u, v, w))\, |J(u, v, w)|\, dV(u, v, w). \quad (14.5)$$

The determinant $J(u, v)$ in Equation (14.2) is called the **Jacobian** of x and y with respect to u and v, and is sometimes written as

$$J(u, v) = \frac{\partial(x, y)}{\partial(u, v)}.$$

Similarly, the Jacobian $J(u, v, w)$ of three variables is sometimes written as

$$J(u, v, w) = \frac{\partial(x, y, z)}{\partial(u, v, w)}.$$

Notes:

Notice that Equation (14.3) is saying that $dA(x, y) = |J(u, v)| \, dA(u, v)$, which you can think of as a two-variable version of the relation $dx = g'(u) \, du$ in the single-variable case.

<div style="border:1px solid; padding:1em;">

Watch the video:

Jacobian at

`https://youtu.be/Bw5yEqwMjQU`
</div>

The following example shows how the change of variables formula is used.

Example 1   **Change of Variables**

Evaluate $\displaystyle\iint_R e^{\frac{x-y}{x+y}} \, dA$, where $R = \{(x, y) : x \geq 0, y \geq 0, x + y \leq 1\}$.

SOLUTION   First, note that evaluating this double integral *without* using substitution is probably impossible, at least in a closed form. By looking at the numerator and denominator of the exponent of e, we will try the substitution $u = x - y$ and $v = x + y$. To use the change of variables Equation (14.3), we need to write both x and y in terms of u and v. So solving for x and y gives $x = \frac{1}{2}(u + v)$ and $y = \frac{1}{2}(v - u)$. In Figure 14.49 below, we see how the mapping $x = x(u, v) = \frac{1}{2}(u + v)$, $y = y(u, v) = \frac{1}{2}(v - u)$ maps the region R' onto R in a one-to-one manner.
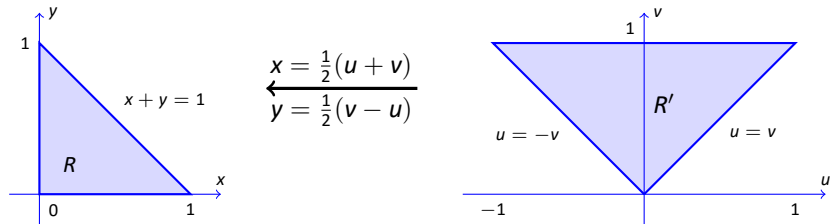
Figure 14.49: The regions R and R'

Now we see that

$$J(u, v) = \begin{vmatrix} \dfrac{\partial x}{\partial u} & \dfrac{\partial x}{\partial v} \\[2mm] \dfrac{\partial y}{\partial u} & \dfrac{\partial y}{\partial v} \end{vmatrix} = \begin{vmatrix} \dfrac{1}{2} & \dfrac{1}{2} \\[2mm] -\dfrac{1}{2} & \dfrac{1}{2} \end{vmatrix} = \dfrac{1}{2} \Rightarrow |J(u, v)| = \left| \dfrac{1}{2} \right| = \dfrac{1}{2},$$

Notes:

so using horizontal slices in R', we have

$$
\begin{aligned}
\iint_R e^{\frac{x-y}{x+y}}\, dA &= \iint_{R'} f(x(u,v), y(u,v))\; |J(u,v)|\; dA \\
&= \int_0^1 \int_{-v}^{v} e^{\frac{u}{v}} \frac{1}{2}\, du\, dv \\
&= \int_0^1 \left(\frac{v}{2} e^{\frac{u}{v}} \Big|_{u=-v}^{u=v} \right) dv \\
&= \int_0^1 \frac{v}{2}(e - e^{-1})\, dv \\
&= \frac{v^2}{4}(e - e^{-1}) \Big|_0^1 = \frac{1}{4}\left(e - \frac{1}{e}\right) = \frac{e^2 - 1}{4e}.
\end{aligned}
$$

The change of variables formula can be used to evaluate double integrals in polar coordinates. Letting

$$ x = x(r, \theta) = r\cos\theta \quad \text{and} \quad y = y(r, \theta) = r\sin\theta, $$

we have

$$
J(u,v) = \begin{vmatrix} \dfrac{\partial x}{\partial r} & \dfrac{\partial x}{\partial \theta} \\[2mm] \dfrac{\partial y}{\partial r} & \dfrac{\partial y}{\partial \theta} \end{vmatrix} = \begin{vmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{vmatrix} = r\cos^2\theta + r\sin^2\theta = r \Rightarrow |J(u,v)| = |r| = r,
$$

which verifies Key Idea 61.

In a similar fashion, it can be shown (see Exercises 5 and 6) that triple integrals in cylindrical and spherical coordinates take the following forms:

Key Idea 62   Triple Integral in Cylindrical Coordinates

$$ \iiint_S f(x, y, z)\, dx\, dy\, dz = \iiint_{S'} f(r\cos\theta, r\sin\theta, z)\, r\, dr\, d\theta\, dz, \qquad (14.6) $$

where the mapping $x = r\cos\theta$, $y = r\sin\theta$, $z = z$ maps the solid S' in $r\theta z$-space onto the solid S in xyz-space in a one-to-one manner.

Notes:

> **Key Idea 63**   **Triple Integral in Spherical Coordinates**
>
> $$\iiint_S f(x, y, z)\, dx\, dy\, dz =$$
>
> $$\iiint_{S'} f(\rho \sin \phi \, \cos \theta, \rho \sin \phi \, \sin \theta, \rho \cos \phi)\, \rho^2 \, \sin \phi \, d\rho \, d\phi \, d\theta, \quad (14.7)$$
>
> where the mapping $x = \rho \sin \phi \, \cos \theta$, $y = \rho \sin \phi \, \sin \theta$, $z = \rho \cos \phi$ maps the solid S' in $\rho\phi\theta$-space onto the solid S in xyz-space in a one-to-one manner.

Example 2   **Finding the Volume of a Sphere**

For $a > 0$, find the volume V inside the sphere $S = x^2 + y^2 + z^2 = a^2$.

    SOLUTION   We see that S is the set $\rho = a$ in spherical coordinates, so

$$V = \iiint_S 1 \, dV = \int_0^{2\pi} \int_0^{\pi} \int_0^a 1\, \rho^2 \, \sin \phi \, d\rho \, d\phi \, d\theta$$

$$= \int_0^{2\pi} \int_0^{\pi} \left(\frac{\rho^3}{3} \bigg|_{\rho=0}^{\rho=a} \right) \sin \phi \, d\phi \, d\theta = \int_0^{2\pi} \int_0^{\pi} \frac{a^3}{3} \sin \phi \, d\phi \, d\theta$$

$$= \int_0^{2\pi} \left(-\frac{a^3}{3} \cos \phi \bigg|_{\phi=0}^{\phi=\pi} \right) d\theta = \int_0^{2\pi} \frac{2a^3}{3} \, d\theta = \frac{4\pi a^3}{3}.$$

    This chapter investigated the natural follow–on to partial derivatives: iterated integration. We learned how to use the bounds of a double integral to describe a region in the plane using both rectangular and polar coordinates, then later expanded to use the bounds of a triple integral to describe a region in space. We used double integrals to find volumes under surfaces, surface area, and the center of mass of lamina; we used triple integrals as an alternate method of finding volumes of space regions and also to find the center of mass of a region in space.

    Integration does not stop here. We could continue to iterate our integrals, next investigating "quadruple integrals" whose bounds describe a region in 4–dimensional space (which are very hard to visualize). We can also look back to "regular" integration where we found the area under a curve in the plane. A natural analogue to this is finding the "area under a curve," where the curve is in space, not in a plane. These are just two of many avenues to explore under the heading of "integration."

Notes:

Exercises 14.7

Problems

1. Find the volume V inside the paraboloid $z = x^2 + y^2$ for $0 \le z \le 4$.

2. Find the volume V inside the cone $z = \sqrt{x^2 + y^2}$ for $0 \le z \le 3$.

3. Find the volume V of the solid inside both $x^2 + y^2 + z^2 = 4$ and $x^2 + y^2 = 1$.

4. Find the volume V inside both the sphere $x^2 + y^2 + z^2 = 1$ and the cone $z = \sqrt{x^2 + y^2}$.

5. Prove Equation (14.6).

6. Prove Equation (14.7).

7. Evaluate $\displaystyle\iint_R \sin\left(\frac{x+y}{2}\right) \cos\left(\frac{x-y}{2}\right) dA$, where R is the triangle with vertices $(0,0)$, $(2,0)$ and $(1,1)$. (*Hint: Use the change of variables* $u = (x+y)/2$, $v = (x-y)/2$.)

8. Find the volume of the solid bounded by $z = x^2 + y^2$ and $z^2 = 4(x^2 + y^2)$.

9. Find the volume inside the elliptic cylinder $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$ for $0 \le z \le 2$.

10. Show that the volume inside the ellipsoid $\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1$ is $\frac{4\pi abc}{3}$. (*Hint: Use the change of variables* $x = au$, $y = bv$, $z = cw$, *then consider Example 2.*)

11. Show that the *Beta function*, defined by

$$B(x, y) = \int_0^1 t^{x-1}(1-t)^{y-1}\, dt, \quad \text{for } x > 0, y > 0,$$

satisfies the relation $B(y, x) = B(x, y)$ for $x > 0, y > 0$.

12. Using the substitution $t = u/(u+1)$, show that the Beta function can be written as

$$B(x, y) = \int_0^\infty \frac{u^{x-1}}{(u+1)^{x+y}}\, du, \quad \text{for } x > 0, y > 0.$$

15: Line and Surface Integrals

15.1 Line Integrals

In single-variable calculus you learned how to integrate a real-valued function $f(x)$ over an interval $[a, b]$ in $\mathbb{R}^1$. This integral (usually called a *Riemann integral*) can be thought of as an integral over a *path* in $\mathbb{R}^1$, since an interval (or collection of intervals) is really the only kind of "path" in $\mathbb{R}^1$. You may also recall that if $f(x)$ represented the force applied along the x-axis to an object at position x in $[a, b]$, then the *work W* done in moving that object from position $x = a$ to $x = b$ was defined as the integral:

$$W = \int_a^b f(x)\, dx$$

In this section, we will see how to define the integral of a function (either real-valued or vector-valued) of two variables over a general path (i.e. a curve) in $\mathbb{R}^2$. This definition will be motivated by the physical notion of work. We will begin with real-valued functions of two variables.

In physics, the intuitive idea of work is that

$$\text{Work} = \text{Force} \times \text{Distance}.$$

Suppose that we want to find the total amount W of work done in moving an object along a curve C in $\mathbb{R}^2$ with a smooth parametrization $x = x(t), y = y(t)$, $a \le t \le b$, with a force $f(x, y)$ which varies with the position (x, y) of the object and is applied in the direction of motion along C (see Figure 15.1 below).
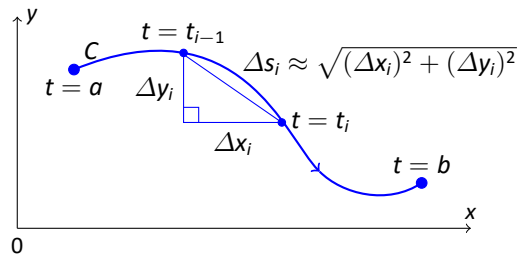
Figure 15.1: Curve $C : x = x(t), y = y(t)$ for t in $[a, b]$

We will assume for now that the function $f(x, y)$ is continuous and real-valued, so we only consider the magnitude of the force. Partition the interval $[a, b]$ as follows:

$$a = t_0 < t_1 < t_2 < \cdots < t_{n-1} < t_n = b, \text{ for some integer } n \ge 2$$

As we can see from Figure 15.1, over a typical subinterval $[t_{i-1}, t_i]$ the distance Δs_i traveled along the curve is approximately $\sqrt{(\Delta x_i)^2 + (\Delta y_i)^2}$, by the Pythagorean

Theorem. Thus, if the subinterval is small enough then the work done in moving the object along that piece of the curve is approximately

$$\text{Force} \times \text{Distance} \approx f(x_i^*, y_i^*) \sqrt{(\Delta x_i)^2 + (\Delta y_i)^2},$$

where $(x_i^*, y_i^*) = (x(t_i^*), y(t_i^*))$ for some t_i^* in $[t_{i-1}, t_i]$, and so

$$W \approx \sum_{i=1}^{n} f(x_i^*, y_i^*) \sqrt{(\Delta x_i)^2 + (\Delta y_i)^2}$$

is approximately the total amount of work done over the entire curve. But since

$$\sqrt{(\Delta x_i)^2 + (\Delta y_i)^2} = \sqrt{\left(\frac{\Delta x_i}{\Delta t_i}\right)^2 + \left(\frac{\Delta y_i}{\Delta t_i}\right)^2}\, \Delta t_i,$$

where $\Delta t_i = t_i - t_{i-1}$, then

$$W \approx \sum_{i=1}^{n} f(x_i^*, y_i^*) \sqrt{\left(\frac{\Delta x_i}{\Delta t_i}\right)^2 + \left(\frac{\Delta y_i}{\Delta t_i}\right)^2}\, \Delta t_i.$$

Taking the limit of that sum as the length of the largest subinterval goes to 0, the sum over all subintervals becomes the integral from $t = a$ to $t = b$, $\frac{\Delta x_i}{\Delta t_i}$ and $\frac{\Delta y_i}{\Delta t_i}$ become $x'(t)$ and $y'(t)$, respectively, and $f(x_i^*, y_i^*)$ becomes $f(x(t), y(t))$, so that

$$W = \int_a^b f(x(t), y(t)) \sqrt{x'(t)^2 + y'(t)^2}\, dt.$$

The integral on the right side of the above equation gives us our idea of how to define, for *any* real-valued function $f(x, y)$, the integral of $f(x, y)$ along the curve C, called a *line integral*:

Definition 112   Line Integral of a Real Valued Function
For a real-valued function $f(x, y)$ and a curve C in $\mathbb{R}^2$, parametrized by $x = x(t)$, $y = y(t)$, $a \le t \le b$, the **line integral of $f(x, y)$ along C with respect to arc length s** is

$$\int_C f(x, y)\, ds = \int_a^b f(x(t), y(t)) \sqrt{x'(t)^2 + y'(t)^2}\, dt.$$

The symbol ds is the differential of the arc length function

$$s = s(t) = \int_a^t \sqrt{x'(u)^2 + y'(u)^2}\, du,$$

Notes:

which you may recognize from Section 10.3 as the length of the curve C over the interval $[a, t]$, for all t in $[a, b]$. That is,

$$ds = s'(t)\, dt = \sqrt{x'(t)^2 + y'(t)^2}\ dt,$$

by the Fundamental Theorem of Calculus.

For a general real-valued function $f(x, y)$, what does the line integral $\int_C f(x, y)\, ds$ represent? The preceding discussion of ds gives us a clue. You can think of differentials as infinitesimal lengths. So if you think of $f(x, y)$ as the height of a picket fence along C, then $f(x, y)\, ds$ can be thought of as approximately the area of a section of that fence over some infinitesimally small section of the curve, and thus the line integral $\int_C f(x, y)\, ds$ is the total area of that picket fence (see Figure 15.2).
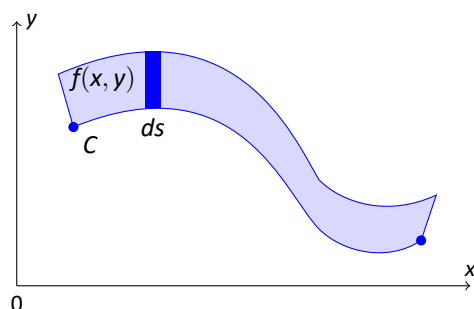
Figure 15.2: Area of shaded rectangle $=$ height $\times$ width $\approx f(x, y)\, ds$

Watch the video:
Line Integrals — Evaluating a Line Integral at
`https://youtu.be/fjEvsinvtnw`

Example 1   Using the Line Integral
Use a line integral to show that the lateral surface area A of a right circular cylinder of radius r and height h is $2\pi rh$.

SOLUTION   We will use the right circular cylinder with base circle C given by $x^2 + y^2 = r^2$ and with height h in the positive z direction (see Figure 15.3). Parametrize C as follows:

$$x = x(t) = r\cos t, \quad y = y(t) = r\sin t, \quad 0 \le t \le 2\pi$$
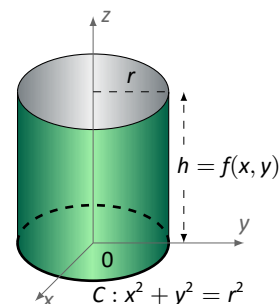
Figure 15.3: Figure for Example 1

Notes:

Let $f(x, y) = h$ for all (x, y). Then

$$A = \int_C f(x, y)\, ds = \int_a^b f(x(t), y(t))\, \sqrt{x'(t)^2 + y'(t)^2}\, dt$$

$$= \int_0^{2\pi} h\sqrt{(-r \sin t)^2 + (r \cos t)^2}\, dt$$

$$= h \int_0^{2\pi} r\sqrt{\sin^2 t + \cos^2 t}\, dt$$

$$= rh \int_0^{2\pi} 1\, dt = 2\pi rh.$$

Note in Example 1 that if we had traversed the circle C twice, i.e. let t vary from 0 to 4π, then we would have gotten an area of $4\pi rh$, i.e. twice the desired area, even though the *curve* itself is still the same (namely, a circle of radius r). Also, notice that we traversed the circle in the counter-clockwise direction. If we had gone in the clockwise direction, using the parametrization

$$x = x(t) = r\cos(2\pi - t), \quad y = y(t) = r\sin(2\pi - t), \quad 0 \le t \le 2\pi, \quad (15.1)$$

then it is easy to verify (see Exercise 12) that the value of the line integral is unchanged.

In general, it can be shown (see Exercise 15) that reversing the direction in which a curve C is traversed leaves $\int_C f(x, y)\, ds$ unchanged, for any $f(x, y)$. If a curve C has a parametrization $x = x(t)$, $y = y(t)$, $a \le t \le b$, then denote by $-C$ the same curve as C but traversed in the opposite direction. Then $-C$ is parametrized by

$$x = x(a + b - t), \quad y = y(a + b - t), \quad a \le t \le b, \quad (15.2)$$

and we have

$$\int_C f(x, y)\, ds = \int_{-C} f(x, y)\, ds.$$

Notice that our definition of the line integral was with respect to the arc length parameter s. We can also define

$$\int_C f(x, y)\, dx = \int_a^b f(x(t), y(t))\, x'(t)\, dt$$

as the *line integral of $f(x, y)$ along C with respect to x*, and

$$\int_C f(x, y)\, dy = \int_a^b f(x(t), y(t))\, y'(t)\, dt$$

Notes:

as the *line integral of f(x, y) along C with respect to y.*

In the derivation of the formula for a line integral, we used the idea of work as force multiplied by distance. However, we know that force is actually a *vector.* So it would be helpful to develop a vector form for a line integral. For this, suppose that we have a function $\vec{f}(x, y)$ defined on $\mathbb{R}^2$ by

$$\vec{f}(x, y) = P(x, y)\,\vec{\imath} + Q(x, y)\,\vec{\jmath}$$

for some continuous real-valued functions $P(x, y)$ and $Q(x, y)$ on $\mathbb{R}^2$. Such a function $\vec{f}$ is called a **vector field** on $\mathbb{R}^2$. It is defined at *points* in $\mathbb{R}^2$, and its values are *vectors* in $\mathbb{R}^2$. For a curve C with a smooth parametrization $x = x(t)$, $y = y(t)$, $a \le t \le b$, let

$$\vec{r}(t) = x(t)\,\vec{\imath} + y(t)\,\vec{\jmath}$$

be the position vector for a point $(x(t), y(t))$ on C. Then $\vec{r}''(t) = x'(t)\,\vec{\imath} + y'(t)\,\vec{\jmath}$ and so

$$\int_C P(x, y)\,dx + \int_C Q(x, y)\,dy = \int_a^b P(x(t), y(t))\,x'(t)\,dt + \int_a^b Q(x(t), y(t))\,y'(t)\,dt$$

$$= \int_a^b (P(x(t), y(t))\,x'(t) + Q(x(t), y(t))\,y'(t))\,dt$$

$$= \int_a^b \vec{f}(x(t), y(t)) \cdot \vec{r}'(t)\,dt$$

by definition of $\vec{f}(x, y)$. Notice that the function $\vec{f}(x(t), y(t)) \cdot \vec{r}'(t)$ is a *real-valued* function on $[a, b]$, so the last integral on the right looks somewhat similar to our earlier definition of a line integral. This leads us to the following definition:

Definition 113   Line Integral of a Vector Valued Function
For a vector field $\vec{f}(x, y) = P(x, y)\,\vec{\imath} + Q(x, y)\,\vec{\jmath}$ and a curve C with a smooth parametrization $x = x(t)$, $y = y(t)$, $a \le t \le b$, the **line integral of f along C** is

$$\int_C \vec{f} \cdot d\vec{r} = \int_C P(x, y)\,dx + \int_C Q(x, y)\,dy \qquad (15.3)$$

$$= \int_a^b \vec{f}(x(t), y(t)) \cdot \vec{r}'(t)\,dt, \qquad (15.4)$$

where $\vec{r}(t) = x(t)\,\vec{\imath} + y(t)\,\vec{\jmath}$ is the position vector for points on C.

Notes:

We use the notation $d\vec{r} = \vec{r}'(t)\,dt = dx\,\vec{\imath} + dy\,\vec{\jmath}$ to denote the **differential** of the vector-valued function $\vec{r}$. The line integral in Definition 113 is often called a *line integral of a vector field* to distinguish it from the line integral in Definition 112 which is called a *line integral of a scalar field*. For convenience we will often write

$$\int_C P(x, y)\,dx + \int_C Q(x, y)\,dy = \int_C P(x, y)\,dx + Q(x, y)\,dy,$$

Note: We defined total differential in Definition 89 in Section 13.4 as $dF = \frac{\partial F}{\partial x}\,dx + \frac{\partial F}{\partial y}\,dy$.

where it is understood that the line integral along C is being applied to both P and Q. The quantity $P(x, y)\,dx + Q(x, y)\,dy$ is known as a **differential form**. A differential form $P(x, y)\,dx + Q(x, y)\,dy$ is called **exact** if it equals the total differential dF for some function $F(x, y)$.

Recall that if the points on a curve C have position vector $\vec{r}(t) = x(t)\,\vec{\imath} + y(t)\,\vec{\jmath}$, then $\vec{r}'(t)$ is a tangent vector to C at the point $(x(t), y(t))$ in the direction of increasing t (which we call the *direction of C*). Since C is a smooth curve, then $\vec{r}'(t) \neq \vec{0}$ on $[a, b]$ and hence

$$\vec{T}(t) = \frac{\vec{r}'(t)}{\|\vec{r}'(t)\|}$$

is the unit tangent vector to C at $(x(t), y(t))$. Putting Definitions 112 and 113 together we get the following theorem:

> **Theorem 132   Line Integrals and Tangent Vectors**
> For a vector field $\vec{f}(x, y) = P(x, y)\,\vec{\imath} + Q(x, y)\,\vec{\jmath}$ and a curve C with a smooth parametrization $x = x(t)$, $y = y(t)$, $a \leq t \leq b$ and position vector $\vec{r}(t) = x(t)\,\vec{\imath} + y(t)\,\vec{\jmath}$,
>
> $$\int_C \vec{f} \cdot d\vec{r} = \int_C \vec{f} \cdot \vec{T}\,ds,$$
>
> where $\vec{T}(t) = \frac{\vec{r}'(t)}{\|\vec{r}'(t)\|}$ is the unit tangent vector to C at $(x(t), y(t))$.

If the vector field $\vec{f}(x, y)$ represents the force moving an object along a curve C, then the work W done by this force is

$$W = \int_C \vec{f} \cdot \vec{T}\,ds = \int_C \vec{f} \cdot d\vec{r}.$$

Notes:

Example 2   Evaluating Line Integrals
Evaluate $\int_C (x^2 + y^2)\, dx + 2xy\, dy$, where:

1. $C : x = t, \quad y = 2t, \quad 0 \le t \le 1$

2. $C : x = t, \quad y = 2t^2, \quad 0 \le t \le 1$

SOLUTION   Figure 15.4 shows both curves.

1. Since $x'(t) = 1$ and $y'(t) = 2$, then

$$\int_C (x^2 + y^2)\, dx + 2xy\, dy = \int_0^1 \left((x(t)^2 + y(t)^2)x'(t) + 2x(t)y(t)\, y'(t) \right)\, dt$$

$$= \int_0^1 \left((t^2 + 4t^2)(1) + 2t(2t)(2) \right)\, dt$$

$$= \int_0^1 13t^2\, dt$$

$$= \frac{13t^3}{3} \Big|_0^1 = \frac{13}{3}$$

Figure 15.4: Figure for Example 2

2. Since $x'(t) = 1$ and $y'(t) = 4t$, then

$$\int_C (x^2 + y^2)\, dx + 2xy\, dy = \int_0^1 \left((x(t)^2 + y(t)^2)x'(t) + 2x(t)y(t)\, y'(t) \right)\, dt$$

$$= \int_0^1 \left((t^2 + 4t^4)(1) + 2t(2t^2)(4t) \right)\, dt$$

$$= \int_0^1 (t^2 + 20t^4)\, dt$$

$$= \frac{t^3}{3} + 4t^5 \Big|_0^1 = \frac{1}{3} + 4 = \frac{13}{3}$$

So in both cases, if the vector field $\vec{f}(x, y) = (x^2 + y^2)\,\vec{\imath} + 2xy\,\vec{\jmath}$ represents the force moving an object from $(0,0)$ to $(1,2)$ along the given curve C, then the work done is $\frac{13}{3}$. This may lead you to think that work (and more generally, the line integral of a vector field) is independent of the path taken. However, as we will see in the next section, this is not always the case.

Notes:

Although we defined line integrals over a single smooth curve, if C is a *piecewise smooth curve*, that is

$$C = C_1 \cup C_2 \cup \ldots \cup C_n$$

is the union of smooth curves $C_1, \ldots, C_n$, then we can define

$$\int_C \vec{f} \cdot d\vec{r} = \int_{C_1} \vec{f} \cdot d\vec{r}_1 + \int_{C_2} \vec{f} \cdot d\vec{r}_2 + \cdots + \int_{C_n} \vec{f} \cdot d\vec{r}_n$$

where each $\vec{r}_i$ is the position vector of the curve C_i.

Example 3   A Piecewise Smooth Line Integral
Evaluate $\int_C (x^2 + y^2)\, dx + 2xy\, dy$, where C is the polygonal path from $(0,0)$ to $(0,2)$ to $(1,2)$.

SOLUTION   Write $C = C_1 \cup C_2$, where C_1 is the curve given by $x = 0$, $y = t$, $0 \leq t \leq 2$ and C_2 is the curve given by $x = t$, $y = 2$, $0 \leq t \leq 1$ (see Figure 15.5). Then

$$\int_C (x^2 + y^2)\, dx + 2xy\, dy$$

$$= \int_{C_1} (x^2 + y^2)\, dx + 2xy\, dy$$

$$+ \int_{C_2} (x^2 + y^2)\, dx + 2xy\, dy$$

$$= \int_0^2 \left((0^2 + t^2)(0) + 2(0)t(1) \right)\, dt + \int_0^1 \left((t^2 + 4)(1) + 2t(2)(0) \right)\, dt$$

$$= \int_0^2 0\, dt + \int_0^1 (t^2 + 4)\, dt$$

$$= \frac{t^3}{3} + 4t \Big|_0^1 = \frac{1}{3} + 4 = \frac{13}{3}$$
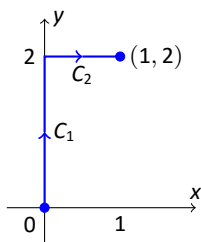
Figure 15.5: The Figure for Example 3

Line integral notation varies quite a bit. For example, in physics it is common to see the notation $\int_a^b \vec{f} \cdot d\vec{l}$, where it is understood that the limits of integration a and b are for the underlying parameter t of the curve, and the letter **l** signifies length. Also, the formulation $\int_C \vec{f} \cdot \vec{T}\, ds$ from Theorem 132 is often preferred in physics since it emphasizes the idea of integrating the tangential component $\vec{f} \cdot \vec{T}$ of $\vec{f}$ in the direction of $\vec{T}$ (i.e., in the direction of C), which is a useful physical interpretation of line integrals.

Notes:

Exercises 15.1

Problems

In Exercises 1–4, calculate $\int_C f(x, y)\,ds$ for the given function $f(x, y)$ and curve C.

1. $f(x, y) = xy$; $C : x = \cos t, y = \sin t, 0 \le t \le \pi/2$

2. $f(x, y) = \dfrac{x}{x^2 + 1}$; $C : x = t, y = 0, 0 \le t \le 1$

3. $f(x, y) = 2x + y$; C: polygonal path from $(0, 0)$ to $(3, 0)$ to $(3, 2)$

4. $f(x, y) = x + y^2$; C: path from $(2, 0)$ counterclockwise along the circle $x^2 + y^2 = 4$ to the point $(-2, 0)$ and then back to $(2, 0)$ along the x-axis

5. Use a line integral to find the lateral surface area of the part of the cylinder $x^2 + y^2 = 4$ below the plane $x + 2y + z = 6$ and above the xy-plane.

In Exercises 6–11, calculate $\int_C \vec{f} \cdot d\vec{r}$ for the given vector field $\vec{f}(x, y)$ and curve C.

6. $\vec{f}(x, y) = \vec{\imath} - \vec{\jmath}$; $C : x = 3t, y = 2t, 0 \le t \le 1$

7. $\vec{f}(x, y) = y\,\vec{\imath} - x\,\vec{\jmath}$; $C : x = \cos t, y = \sin t, 0 \le t \le 2\pi$

8. $\vec{f}(x, y) = x\,\vec{\imath} + y\,\vec{\jmath}$; $C : x = \cos t, y = \sin t, 0 \le t \le 2\pi$

9. $\vec{f}(x, y) = (x^2 - y)\,\vec{\imath} + (x - y^2)\,\vec{\jmath}$; $C : x = \cos t, y = \sin t, 0 \le t \le 2\pi$

10. $\vec{f}(x, y) = xy^2\,\vec{\imath} + xy^3\,\vec{\jmath}$; $C :$ the polygonal path from $(0, 0)$ to $(1, 0)$ to $(0, 1)$ to $(0, 0)$

11. $\vec{f}(x, y) = (x^2 + y^2)\,\vec{\imath}$; $C : x = 2 + \cos t, y = \sin t, 0 \le t \le 2\pi$

12. Verify that the value of the line integral in Example 1 is unchanged when using the parametrization of the circle C given in Equation (15.1).

13. Show that if $\vec{f} \perp \vec{r}\,'(t)$ at each point $\vec{r}(t)$ along a smooth curve C, then $\int_C \vec{f} \cdot d\vec{r} = 0$.

14. Show that if $\vec{f}$ points in the same direction as $\vec{r}\,'(t)$ at each point $\vec{r}(t)$ along a smooth curve C, then $\int_C \vec{f} \cdot d\vec{r} = \int_C \left\| \vec{f} \right\| ds$.

15. Prove that $\int_C f(x, y)\,ds = \int_{-C} f(x, y)\,ds$. (*Hint: Use Equation (15.2).*)

16. Let C be a smooth curve with arc length L, and suppose that $\vec{f}(x, y) = P(x, y)\,\vec{\imath} + Q(x, y)\,\vec{\jmath}$ is a vector field such that $\left\| \vec{f}(x, y) \right\| \le M$ for all (x, y) on C. Show that $\left| \int_C \vec{f} \cdot d\vec{r} \right| \le ML$. (*Hint: Recall that* $\left| \int_a^b g(x)\,dx \right| \le \int_a^b |g(x)|\,dx$ *for Riemann integrals.*)

17. Prove that the Riemann integral $\int_a^b f(x)\,dx$ is a special case of a line integral.

15.2   Properties of Line Integrals

We know from the previous section that for line integrals of real-valued functions (scalar fields), reversing the direction in which the integral is taken along a curve does not change the value of the line integral:

$$\int_C f(x,y)\, ds = \int_{-C} f(x,y)\, ds$$

For line integrals of vector fields, however, the value does change. To see this, let $\vec{f}(x,y) = P(x,y)\,\vec{\imath} + Q(x,y)\,\vec{\jmath}$ be a vector field, with P and Q continuously differentiable functions. Let C be a smooth curve parametrized by $x = x(t)$, $y = y(t)$, $a \le t \le b$, with position vector $\vec{r}(t) = x(t)\,\vec{\imath} + y(t)\,\vec{\jmath}$ (we will usually abbreviate this by saying that $C : \vec{r}(t) = x(t)\,\vec{\imath} + y(t)\,\vec{\jmath}$ is a smooth curve). We know that the curve $-C$ traversed in the opposite direction is parametrized by $x = x(a+b-t)$, $y = y(a+b-t)$, $a \le t \le b$. Then

$$\int_{-C} P(x,y)\, dx = \int_a^b P(x(a+b-t), y(a+b-t)) \frac{d}{dt}(x(a+b-t))\, dt$$

$$= \int_a^b P(x(a+b-t), y(a+b-t))\,(-x'(a+b-t))\, dt \text{ (by the Chain Rule)}$$

$$= \int_b^a P(x(u), y(u))\,(-x'(u))\,(-du) \text{ (by letting } u = a+b-t)$$

$$= \int_b^a P(x(u), y(u))\, x'(u)\, du$$

$$= -\int_a^b P(x(u), y(u))\, x'(u)\, du, \text{ since } \int_b^a = -\int_a^b, \text{ so}$$

$$\int_{-C} P(x,y)\, dx = -\int_C P(x,y)\, dx$$

since we are just using a different letter (u) for the line integral along C. A similar argument shows that

$$\int_{-C} Q(x,y)\, dy = -\int_C Q(x,y)\, dy,$$

Notes:

and hence

$$\int_{-C} \vec{f} \cdot d\vec{r} = \int_{-C} P(x, y)\, dx + \int_{-C} Q(x, y)\, dy$$

$$= - \int_{C} P(x, y)\, dx + - \int_{C} Q(x, y)\, dy$$

$$= - \left(\int_{C} P(x, y)\, dx + \int_{C} Q(x, y)\, dy \right)$$

$$\int_{-C} \vec{f} \cdot d\vec{r} = - \int_{C} \vec{f} \cdot d\vec{r}.$$

The above formula can be interpreted in terms of the work done by a force $\vec{f}(x, y)$ (treated as a vector) moving an object along a curve C: the total work performed moving the object along C from its initial point to its terminal point, and then back to the initial point moving backwards along the same path, is zero. This is because when force is considered as a vector, direction is accounted for.

The preceding discussion shows the importance of always taking the *direction* of the curve into account when using line integrals of vector fields. For this reason, the curves in line integrals are sometimes referred to as *directed curves* or *oriented curves*.

Recall that our definition of a line integral required that we have *a* parametrization $x = x(t)$, $y = y(t)$, $a \leq t \leq b$ for the curve C. But as we know, any curve has infinitely many parameterizations. So could we get a different value for a line integral using some other parametrization of C, say, $x = \tilde{x}(u)$, $y = \tilde{y}(u)$, $c \leq u \leq d$? If so, this would mean that our definition is not well-defined. Luckily, it turns out that the value of a line integral of a vector field is unchanged as long as the direction of the curve C is preserved by whatever parametrization is chosen:

> **Theorem 133   Line Integral is Independent of Parameterization**
> Let $\vec{f}(x, y) = P(x, y)\,\vec{i} + Q(x, y)\,\vec{j}$ be a vector field, and let C be a smooth curve parametrized by $x = x(t)$, $y = y(t)$, $a \leq t \leq b$. Suppose that $t = \alpha(u)$ for $c \leq u \leq d$, such that $a = \alpha(c)$, $b = \alpha(d)$, and $\alpha'(u) > 0$ on the open interval (c, d) (i.e., $\alpha(u)$ is strictly increasing on $[c, d]$). Then $\int_C \vec{f} \cdot d\vec{r}$ has the same value for the parameterizations $x = x(t)$, $y = y(t)$, $a \leq t \leq b$ and $x = \tilde{x}(u) = x(\alpha(u))$, $y = \tilde{y}(u) = y(\alpha(u))$, $c \leq u \leq d$.

Proof
Since $\alpha(u)$ is strictly increasing and maps $[c, d]$ onto $[a, b]$, then we know that

Notes:

$t = \alpha(u)$ has an inverse function $u = \alpha^{-1}(t)$ defined on $[a, b]$ such that $c = \alpha^{-1}(a)$, $d = \alpha^{-1}(b)$, and $\frac{du}{dt} = \frac{1}{\alpha'(u)}$. Also, $dt = \alpha'(u)\, du$, and by the Chain Rule

$$\tilde{x}'(u) = \frac{d\tilde{x}}{du} = \frac{d}{du}(x(\alpha(u))) = \frac{dx}{dt}\frac{dt}{du} = x'(t)\,\alpha'(u) \quad \Rightarrow \quad x'(t) = \frac{\tilde{x}'(u)}{\alpha'(u)}$$

so making the substitution $t = \alpha(u)$ gives

$$\int_a^b P(x(t), y(t))\, x'(t)\, dt = \int_{\alpha^{-1}(a)}^{\alpha^{-1}(b)} P(x(\alpha(u)), y(\alpha(u))) \frac{\tilde{x}'(u)}{\alpha'(u)}\, (\alpha'(u)\, du)$$

$$= \int_c^d P(\tilde{x}(u), \tilde{y}(u))\, \tilde{x}'(u)\, du,$$

which shows that $\int_C P(x, y)\, dx$ has the same value for both parameterizations. A similar argument shows that $\int_C Q(x, y)\, dy$ has the same value for both parameterizations, and hence $\int_C \vec{f} \cdot d\vec{r}$ has the same value. $\qquad\qquad \square$

Notice that the condition $\alpha'(u) > 0$ in Theorem 133 means that the two parameterizations move along C in the same direction. That was *not* the case with the "reverse" parametrization for $-C$: for $u = a+b-t$ we have $t = \alpha(u) = a + b - u \Rightarrow \alpha'(u) = -1 < 0$.

Example 1   **Re-evaluating a Line Integral**
Evaluate the line integral $\int_C (x^2 + y^2)\, dx + 2xy\, dy$ from Example 15.1.2 in Section 15.1, along the curve $C : x = t, y = 2t^2, 0 \le t \le 1$, where $t = \sin u$ for $0 \le u \le \pi/2$.

SOLUTION   First, we notice that $0 = \sin 0$, $1 = \sin(\pi/2)$, and $\frac{dt}{du} = \cos u > 0$ on $(0, \pi/2)$. So by Theorem 133 we know that if C is parametrized by

$$x = \sin u, \quad y = 2\sin^2 u, \quad 0 \le u \le \pi/2$$

then $\int_C (x^2 + y^2)\, dx + 2xy\, dy$ should have the same value as we found in Exam-

Notes:

ple 15.1.2, namely $\frac{13}{3}$. And we can indeed verify this:

$$\int_C (x^2 + y^2)\, dx + 2xy\, dy$$

$$= \int_0^{\pi/2} \left((\sin^2 u + (2\sin^2 u)^2)\cos u + 2(\sin u)(2\sin^2 u)4\sin u\, \cos u \right)\, du$$

$$= \int_0^{\pi/2} \left(\sin^2 u + 20\sin^4 u \right)\cos u\, du$$

$$= \frac{\sin^3 u}{3} + 4\sin^5 u\, \Big|_0^{\pi/2}$$

$$= \frac{1}{3} + 4 = \frac{13}{3}$$

In other words, the line integral is unchanged whether t or u is the parameter for C.

By a **closed curve**, we mean a curve C whose initial point and terminal point are the same, i.e. for C: $x = x(t)$, $y = y(t)$, $a \le t \le b$, we have $(x(a), y(a)) = (x(b), y(b))$.

(a) Closed   (b) Not Closed
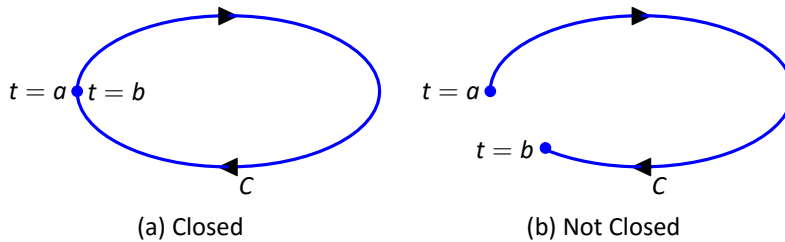
Figure 15.6: Closed vs nonclosed curves

A **simple closed curve** is a closed curve which does not intersect itself. Note that any closed curve can be regarded as a union of simple closed curves (think of the loops in a figure eight). We use the special notation

$$\oint_C f(x, y)\, ds \quad \text{and} \quad \oint_C \vec{f} \cdot d\vec{r}$$

to denote line integrals of scalar and vector fields, respectively, along closed curves. In some older texts you may see the notation $\oint$ or $\oint$ to indicate a line integral traversing a closed curve in a counterclockwise or clockwise direction, respectively.

So far, the examples we have seen of line integrals (e.g., Example 15.1.2) have had the same value for different curves joining the initial point to the terminal

Notes:

point. That is, the line integral has been independent of the path joining the two points. As we mentioned before, this is not always the case. The following theorem gives a necessary and sufficient condition for this *path independence*:

Theorem 134   Path Independence of Line Integrals

In a region R, the line integral $\int_C \vec{f} \cdot d\vec{r}$ is independent of the path between any two points in R if and only if $\oint_C \vec{f} \cdot d\vec{r} = 0$ for every closed curve C which is contained in R.
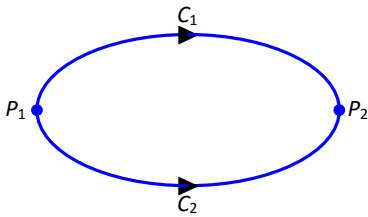
Figure 15.7: The idea of proving Theorem 134.

Proof

Suppose that $\oint_C \vec{f} \cdot d\vec{r} = 0$ for every closed curve C which is contained in R. Let P_1 and P_2 be two distinct points in R. Let C_1 be a curve in R going from P_1 to P_2, and let C_2 be another curve in R going from P_1 to P_2, as in Figure 15.7.

Then $C = C_1 \cup -C_2$ is a closed curve in R (from P_1 to P_1), and so $\oint_C \vec{f} \cdot d\vec{r} = 0$. Thus,

$$0 = \oint_C \vec{f} \cdot d\vec{r}$$

$$= \int_{C_1} \vec{f} \cdot d\vec{r} + \int_{-C_2} \vec{f} \cdot d\vec{r}$$

$$= \int_{C_1} \vec{f} \cdot d\vec{r} - \int_{C_2} \vec{f} \cdot d\vec{r}, \text{ and so}$$

$\int_{C_1} \vec{f} \cdot d\vec{r} = \int_{C_2} \vec{f} \cdot d\vec{r}$. This proves path independence.

Conversely, suppose that the line integral $\int_C \vec{f} \cdot d\vec{r}$ is independent of the path between any two points in R. Let C be a closed curve contained in R. Let P_1 and P_2 be two distinct points on C. Let C_1 be a part of the curve C that goes from P_1 to P_2, and let C_2 be the remaining part of C that goes from P_1 to P_2, again as in

---

Notes:

Figure 15.7. Then by path independence we have

$$\int_{C_1} \vec{f} \cdot d\vec{r} = \int_{C_2} \vec{f} \cdot d\vec{r}$$

$$\int_{C_1} \vec{f} \cdot d\vec{r} - \int_{C_2} \vec{f} \cdot d\vec{r} = 0$$

$$\int_{C_1} \vec{f} \cdot d\vec{r} + \int_{-C_2} \vec{f} \cdot d\vec{r} = 0, \text{ so}$$

$$\oint_C \vec{f} \cdot d\vec{r} = 0$$

since $C = C_1 \cup -C_2$. □

Clearly, the above theorem does not give a practical way to determine path independence, since it is impossible to check the line integrals around all possible closed curves in a region. What it mostly does is give an idea of the way in which line integrals behave, and how seemingly unrelated line integrals can be related (in this case, a specific line integral between two points and *all* line integrals around closed curves). We will now prove the following *sufficient* condition for path independence of line integrals:

Theorem 135   Fundamental Theorem of Line Integrals
Let $\vec{f}(x, y) = P(x, y)\,\vec{i} + Q(x, y)\,\vec{j}$ be a vector field in some region R without holes, with P and Q continuously differentiable functions on R. Let C be a smooth curve in R parametrized by $x = x(t), y = y(t), a \leq t \leq b$. Suppose that there is a real-valued function $F(x, y)$ such that $\nabla F = \vec{f}$ on R. Then

$$\int_C \vec{f} \cdot d\vec{r} = F(B) - F(A),$$

where $A = (x(a), y(a))$ and $B = (x(b), y(b))$ are the endpoints of C. Thus, the line integral is independent of the path between its endpoints, since it depends only on the values of F at those endpoints.

Proof

Notes:

943

By definition of $\int_C \vec{f} \cdot d\vec{r}$, we have

$$\int_C \vec{f} \cdot d\vec{r} = \int_a^b \left(P(x(t), y(t)) \, x'(t) + Q(x(t), y(t)) \, y'(t) \right) dt$$

$$= \int_a^b \left(\frac{\partial F}{\partial x} \frac{dx}{dt} + \frac{\partial F}{\partial y} \frac{dy}{dt} \right) dt \text{ (since } \nabla F = \vec{f} \Rightarrow \frac{\partial F}{\partial x} = P \text{ and } \frac{\partial F}{\partial y} = Q)$$

$$= \int_a^b F'(x(t), y(t)) \, dt \text{ (by Theorem 110)}$$

$$= F(x(t), y(t)) \, |_a^b = F(B) - F(A)$$

by the Fundamental Theorem of Calculus.　　　　　　　　　　　　□

Theorem 135 can be thought of as the line integral version of the Fundamental Theorem of Calculus. A real-valued function $F(x, y)$ such that $\nabla F(x, y) = \vec{f}(x, y)$ is called a **potential** for $\vec{f}$. A **conservative** vector field is one which has a potential.

> Watch the video:
> Ex 1: Fundamental Theorem of Line Integrals —
> Given Vector Field in a Plane at
> `https://youtu.be/62oBGKSjYiY`

Example 2　　　　　Using the Fundamental Theorem of Line Integrals

Recall from Examples 15.1.2 and 15.1.3 in Section 15.1 that the line integral $\int_C (x^2 + y^2) \, dx + 2xy \, dy$ was found to have the value $\frac{13}{3}$ for three different curves C going from the point $(0, 0)$ to the point $(1, 2)$. Use Theorem 135 to show that this line integral is indeed path independent.

SOLUTION　　　　We need to find a real-valued function $F(x, y)$ such that

$$\frac{\partial F}{\partial x} = x^2 + y^2 \quad \text{and} \quad \frac{\partial F}{\partial y} = 2xy.$$

Suppose that $\frac{\partial F}{\partial x} = x^2 + y^2$, Then we must have $F(x, y) = \frac{1}{3}x^3 + xy^2 + g(y)$ for some function $g(y)$. So $\frac{\partial F}{\partial y} = 2xy + g'(y)$ satisfies the condition $\frac{\partial F}{\partial y} = 2xy$ if $g'(y) = 0$, i.e., $g(y) = K$, where K is a constant. Since any choice for K will do (why?), we pick $K = 0$. Thus, a potential $F(x, y)$ for $\vec{f}(x, y) = (x^2 + y^2)\,\vec{\imath} + 2xy\,\vec{\jmath}$ exists, namely

$$F(x, y) = \frac{1}{3}x^3 + xy^2.$$

Notes:

Hence the line integral $\int_C (x^2 + y^2)\,dx + 2xy\,dy$ is path independent.

Note that we can also verify that the value of the line integral of $\vec{f}$ along any curve C going from $(0,0)$ to $(1,2)$ will always be $\frac{13}{3}$, since by Theorem 135

$$\int_C \vec{f} \cdot d\vec{r} = F(1,2) - F(0,0) = \frac{1}{3}(1)^3 + (1)(2)^2 - (0+0) = \frac{1}{3} + 4 = \frac{13}{3}.$$

A consequence of Theorem 135 in the special case where C is a closed curve, so that the endpoints A and B are the same point, is the following:

> **Theorem 136   Closed Line Integrals of Conservative Fields**
>
> If a vector field $\vec{f}$ has a potential in a region R without holes, then $\oint_C \vec{f} \cdot d\vec{r} = 0$ for any closed curve C in R (i.e., $\oint_C \nabla F \cdot d\vec{r} = 0$ for any real-valued function $F(x,y)$).

Example 3   Calculating a Closed Line Integral of a Conservative Field

Evaluate $\oint_C x\,dx + y\,dy$ for $C : x = 2\cos t,\ y = 3\sin t,\ 0 \le t \le 2\pi$.

> **SOLUTION**   The vector field $\vec{f}(x,y) = x\,\vec{\imath} + y\,\vec{\jmath}$ has a potential $F(x,y)$:
>
> $$\frac{\partial F}{\partial x} = x \Rightarrow F(x,y) = \frac{1}{2}x^2 + g(y),\text{ so}$$
>
> $$\frac{\partial F}{\partial y} = y \Rightarrow g'(y) = y \Rightarrow g(y) = \frac{1}{2}y^2 + K$$
>
> for any constant K, so $F(x,y) = \frac{1}{2}x^2 + \frac{1}{2}y^2$ is a potential for $\vec{f}(x,y)$. Thus,
>
> $$\oint_C x\,dx + y\,dy = \oint_C \vec{f} \cdot d\vec{r} = 0$$
>
> by Theorem 136, since the curve C is closed (it is the ellipse $\frac{x^2}{4} + \frac{y^2}{9} = 1$).

Notes:

Exercises 15.2

Problems

1. Evaluate $\oint_C (x^2 + y^2)\,dx + 2xy\,dy$ for $C : x = \cos t, y = \sin t$, $0 \le t \le 2\pi$.

2. Evaluate $\int_C (x^2 + y^2)\,dx + 2xy\,dy$ for $C : x = \cos t, y = \sin t$, $0 \le t \le \pi$.

3. Is there a potential $F(x, y)$ for $\vec{f}(x, y) = y\,\vec{\imath} - x\,\vec{\jmath}$? If so, find one.

4. Is there a potential $F(x, y)$ for $\vec{f}(x, y) = x\,\vec{\imath} - y\,\vec{\jmath}$? If so, find one.

5. Is there a potential $F(x, y)$ for $\vec{f}(x, y) = xy^2\,\vec{\imath} + x^3y\,\vec{\jmath}$? If so, find one.

6. Let $\vec{f}(x, y)$ and $\vec{g}(x, y)$ be vector fields, let a and b be constants, and let C be a curve in $\mathbb{R}^2$. Show that

$$\int_C (a\vec{f} \pm b\vec{g}) \cdot d\vec{r} = a \int_C \vec{f} \cdot d\vec{r} \pm b \int_C \vec{g} \cdot d\vec{r}.$$

7. Let C be a curve whose arc length is L. Show that $\int_C 1\,ds = L$.

8. Let $f(x, y)$ and $g(x, y)$ be continuously differentiable real-valued functions in a region R. Show that

$$\oint_C (f \nabla g) \cdot d\vec{r} = - \oint_C (g \nabla f) \cdot d\vec{r}$$

for any closed curve C in R.

9. Let $\vec{f}(x, y) = \frac{-y}{x^2+y^2}\,\vec{\imath} + \frac{x}{x^2+y^2}\,\vec{\jmath}$ for all $(x, y) \ne (0, 0)$, and $C : x = \cos t, y = \sin t, 0 \le t \le 2\pi$.

   (a) Show that $\vec{f} = \nabla F$, for $F(x, y) = \tan^{-1}(y/x)$.

   (b) Show that $\oint_C \vec{f} \cdot d\vec{r} = 2\pi$. Does this contradict Theorem 136? Explain.

10. Let $g(x)$ and $h(y)$ be differentiable functions, and let $\vec{f}(x, y) = h(y)\,\vec{\imath} + g(x)\,\vec{\jmath}$. Is it possible for $\vec{f}$ to have a potential $F(x, y)$? If so, find an example. You may assume that F would be smooth. (*Hint: Consider the mixed partial derivatives of F.*)

15.3 Green's Theorem

We will now see a way of evaluating the line integral of a *smooth* vector field around a simple closed curve. A vector field $\vec{f}(x, y) = P(x, y)\,\vec{\imath} + Q(x, y)\,\vec{\jmath}$ is **smooth** if its component functions $P(x, y)$ and $Q(x, y)$ are smooth. We will use *Green's Theorem* (sometimes called *Green's Theorem in the plane*) to relate the *line* integral around a closed curve with a *double* integral over the region inside the curve:

Theorem 137   Green's Theorem

Let R be a region in $\mathbb{R}^2$ whose boundary is a simple closed curve C which is piecewise smooth. Let $\vec{f}(x, y) = P(x, y)\,\vec{\imath} + Q(x, y)\,\vec{\jmath}$ be a smooth vector field defined on both R and C. Then

$$\oint_C \vec{f} \cdot d\vec{r} = \iint_R \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} \right) dA, \qquad (15.5)$$

where C is traversed so that R is always on the left side of C.

Proof

We will prove the theorem in the case for a *simple* region R, that is, where the boundary curve C can be written as $C = C_1 \cup C_2$ in two distinct ways:

$$C_1 = \text{the curve } y = y_1(x) \text{ from the point } X_1 \text{ to the point } X_2 \qquad (15.6)$$
$$C_2 = \text{the curve } y = y_2(x) \text{ from the point } X_2 \text{ to the point } X_1, \qquad (15.7)$$

where X_1 and X_2 are the points on C farthest to the left and right, respectively; and

$$C_1 = \text{the curve } x = x_1(y) \text{ from the point } Y_2 \text{ to the point } Y_1 \qquad (15.8)$$
$$C_2 = \text{the curve } x = x_2(y) \text{ from the point } Y_1 \text{ to the point } Y_2, \qquad (15.9)$$

where Y_1 and Y_2 are the lowest and highest points, respectively, on C. See Figure 15.8.

Notes:
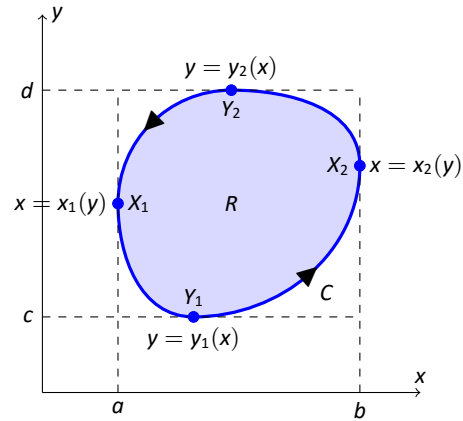
Figure 15.8: Figure for Green's Theorem on a simple region.

Integrate $P(x, y)$ around C using the representation $C = C_1 \cup C_2$ given by Equations (15.6) and (15.7). Since $y = y_1(x)$ along C_1 (as x goes from a to b) and $y = y_2(x)$ along C_2 (as x goes from b to a), as we see from Figure 15.8, then we have

$$\oint_C P(x, y)\, dx = \int_{C_1} P(x, y)\, dx + \int_{C_2} P(x, y)\, dx$$

$$= \int_a^b P(x, y_1(x))\, dx + \int_b^a P(x, y_2(x))\, dx$$

$$= \int_a^b P(x, y_1(x))\, dx - \int_a^b P(x, y_2(x))\, dx$$

$$= -\int_a^b \left(P(x, y_2(x)) - P(x, y_1(x)) \right) dx$$

$$= -\int_a^b \left(P(x, y) \Big|_{y=y_1(x)}^{y=y_2(x)} \right) dx$$

$$= -\int_a^b \int_{y_1(x)}^{y_2(x)} \frac{\partial P(x, y)}{\partial y}\, dy\, dx \text{ (by the Fundamental Theorem of Calculus)}$$

$$= -\iint_R \frac{\partial P}{\partial y}\, dA.$$

Likewise, integrate $Q(x, y)$ around C using the representation $C = C_1 \cup C_2$ given by Equations (15.8) and (15.9). Since $x = x_1(y)$ along C_1 (as y goes from d to c) and $x = x_2(y)$ along C_2 (as y goes from c to d), as we see from Figure 15.8, then

Notes:

we have

$$\oint_C Q(x, y)\, dy = \int_{C_1} Q(x, y)\, dy + \int_{C_2} Q(x, y)\, dy$$

$$= \int_d^c Q(x_1(y), y)\, dy + \int_c^d Q(x_2(y), y)\, dy$$

$$= -\int_c^d Q(x_1(y), y)\, dy + \int_c^d Q(x_2(y), y)\, dy$$

$$= \int_c^d \left(Q(x_2(y), y) - Q(x_1(y), y) \right) dy$$

$$= \int_c^d \left(Q(x, y) \Big|_{x=x_1(y)}^{x=x_2(y)} \right) dy$$

$$= \int_c^d \int_{x_1(y)}^{x_2(y)} \frac{\partial Q(x, y)}{\partial x}\, dx\, dy \text{ (by the Fundamental Theorem of Calculus)}$$

$$= \iint_R \frac{\partial Q}{\partial x}\, dA.$$

Putting this together, we have

$$\oint_C \vec{f} \cdot d\vec{r} = \oint_C P(x, y)\, dx + \oint_C Q(x, y)\, dy$$

$$= -\iint_R \frac{\partial P}{\partial y}\, dA + \iint_R \frac{\partial Q}{\partial x}\, dA$$

$$= \iint_R \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} \right) dA. \qquad \square$$

Though we proved Green's Theorem only for a simple region R, the theorem can also be proved for more general regions (say, a union of simple regions).

Watch the video:
Green's Theorem at
`https://youtu.be/a_zdFvYXX_c`

Notes:

Example 1   **Using Green's Theorem**

Evaluate $\oint_C (x^2 + y^2)\, dx + 2xy\, dy$, where C is the boundary (traversed counter-clockwise) of the region $R = \{(x, y) : 0 \le x \le 1, 2x^2 \le y \le 2x\}$.

SOLUTION R is the shaded region in Figure 15.9. By Green's Theorem, for $P(x, y) = x^2 + y^2$ and $Q(x, y) = 2xy$, we have

$$\oint_C (x^2 + y^2)\, dx + 2xy\, dy = \iint_R \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} \right) dA$$

$$= \iint_R (2y - 2y)\, dA = \iint_R 0\, dA = 0.$$

We actually already knew that the answer was zero. Recall from Example 15.2.2 in Section 15.2 that the vector field $\vec{f}(x, y) = (x^2 + y^2)\,\vec{\imath} + 2xy\,\vec{\jmath}$ has a potential function $F(x, y) = \frac{1}{3}x^3 + xy^2$, and so $\oint_C \vec{f} \cdot d\vec{r} = 0$ by Theorem 136.

Figure 15.9: Figure for Example 1

Example 2   **Green's Theorem with a Hole**

Let $\vec{f}(x, y) = P(x, y)\,\vec{\imath} + Q(x, y)\,\vec{\jmath}$, where

$$P(x, y) = \frac{-y}{x^2 + y^2} \quad \text{and} \quad Q(x, y) = \frac{x}{x^2 + y^2},$$

and let $R = \{(x, y) : 0 < x^2 + y^2 \le 1\}$. For the boundary curve $C : x^2 + y^2 = 1$, traversed counterclockwise, it was shown in Exercise 9(b) in Section 15.2 that $\oint_C \vec{f} \cdot d\vec{r} = 2\pi$. But

$$\frac{\partial Q}{\partial x} = \frac{y^2 - x^2}{(x^2 + y^2)^2} = \frac{\partial P}{\partial y} \Rightarrow \iint_R \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} \right) dA = \iint_R 0\, dA = 0.$$

This would seem to contradict Green's Theorem. However, note that R is not the *entire* region enclosed by C, since the point $(0, 0)$ is not contained in R. That is, R has a "hole" at the origin, so Green's Theorem does not apply.

If we modify the region R to be the *annulus* $R = \{(x, y) : 1/4 \le x^2 + y^2 \le 1\}$ (see Figure 15.10), and take the "boundary" C of R to be $C = C_1 \cup C_2$, where C_1 is the unit circle $x^2 + y^2 = 1$ traversed counterclockwise and C_2 is the circle $x^2 + y^2 = 1/4$ traversed *clockwise*, then it can be shown (see Exercise 8) that

$$\oint_C \vec{f} \cdot d\vec{r} = 0.$$

We would still have $\iint_R \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} \right) dA = 0$, so for this R we would have

$$\oint_C \vec{f} \cdot d\vec{r} = \iint_R \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} \right) dA,$$
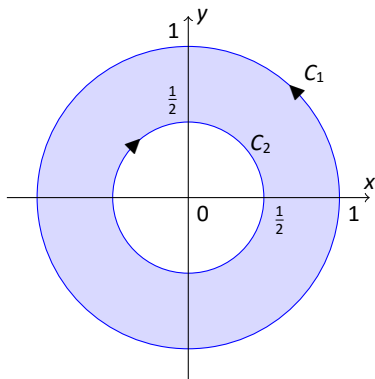
Figure 15.10: The annulus R

Notes:

which shows that Green's Theorem holds for the annular region R.

It turns out that Green's Theorem can be extended to *multiply connected* regions, that is, regions like the annulus in Example 2, which have one or more *regions* cut out from the interior, as opposed to discrete *points* being cut out. For such regions, the "outer" boundary and the "inner" boundaries are traversed so that R is always on the left side.

Region R with one hole   Region R with two holes
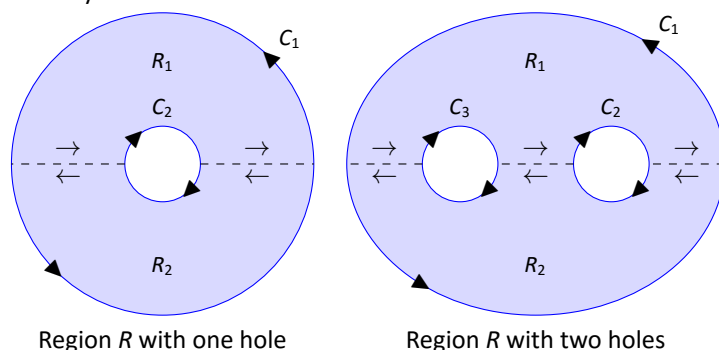
Figure 15.11: Multiply connected regions

The intuitive idea for why Green's Theorem holds for multiply connected regions is shown in Figure 15.11 above. The idea is to cut "slits" between the boundaries of a multiply connected region R so that R is divided into subregions which do *not* have any "holes". For example, in Figure 15.11(a) the region R is the union of the regions R_1 and R_2, which are divided by the slits indicated by the dashed lines. Those slits are part of the boundary of both R_1 and R_2, and we traverse then in the manner indicated by the arrows. Notice that along each slit the boundary of R_1 is traversed in the opposite direction as that of R_2, which means that the line integrals of $\vec{f}$ along those slits add to 0. Since R_1 and R_2 do not have holes in them, then Green's Theorem holds in each subregion, so that

$$\oint_{\substack{\text{bdy} \\ \text{of } R_1}} \vec{f} \cdot d\vec{r} = \iint_{R_1} \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} \right) dA \quad \text{and} \quad \oint_{\substack{\text{bdy} \\ \text{of } R_2}} \vec{f} \cdot d\vec{r} = \iint_{R_2} \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} \right) dA.$$

But since the line integrals along the slits are opposite each other, we have

$$\oint_{C_1 \cup C_2} \vec{f} \cdot d\vec{r} = \oint_{\substack{\text{bdy} \\ \text{of } R_1}} \vec{f} \cdot d\vec{r} + \oint_{\substack{\text{bdy} \\ \text{of } R_2}} \vec{f} \cdot d\vec{r},$$

and so

$$\oint_{C_1 \cup C_2} \vec{f} \cdot d\vec{r} = \iint_{R_1} \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} \right) dA + \iint_{R_2} \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} \right) dA = \iint_{R} \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} \right) dA,$$

Notes:

which shows that Green's Theorem holds in the region R. A similar argument shows that the theorem holds in the region with two holes shown in Figure 15.11(b).

We know from Theorem 136 that when a smooth vector field $\vec{f}(x, y) = P(x, y)\,\vec{\imath} + Q(x, y)\,\vec{\jmath}$ on a region R (whose boundary is a piecewise smooth, simple closed curve C) has a potential in R, then $\oint_C \vec{f} \cdot d\vec{r} = 0$. And if the potential $F(x, y)$ is smooth in R, then $\frac{\partial F}{\partial x} = P$ and $\frac{\partial F}{\partial y} = Q$, and so we know that

$$\frac{\partial^2 F}{\partial y\,\partial x} = \frac{\partial^2 F}{\partial x\,\partial y} \;\Rightarrow\; \frac{\partial P}{\partial y} = \frac{\partial Q}{\partial x} \;\text{ in } R.$$

Conversely, if $\frac{\partial P}{\partial y} = \frac{\partial Q}{\partial x}$ in R then

$$\oint_C \vec{f} \cdot d\vec{r} = \iint_R \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} \right) dA = \iint_R 0\,dA = 0.$$

For a **simply connected** region R (i.e. a region with no holes), the following can be shown:

Theorem 138   Equivalence of Path Independence

The following statements are equivalent for a simply connected region R in $\mathbb{R}^2$:

1. $\vec{f}(x, y) = P(x, y)\,\vec{\imath} + Q(x, y)\,\vec{\jmath}$ has a smooth potential $F(x, y)$ in R

2. $\displaystyle\int_C \vec{f} \cdot d\vec{r}$ is independent of the path for any curve C in R

3. $\displaystyle\oint_C \vec{f} \cdot d\vec{r} = 0$ for every simple closed curve C in R

4. $\dfrac{\partial P}{\partial y} = \dfrac{\partial Q}{\partial x}$ in R (in this case, the differential form $P\,dx + Q\,dy$ is exact)

Notes:

Exercises 15.3

Problems

In Exercises 1–4, use Green's Theorem to evaluate the given line integral around the curve C, traversed counterclockwise.

1. $\oint_C (x^2 - y^2)\, dx + 2xy\, dy$; C is the boundary of $R = \{ (x, y) : 0 \leq x \leq 1, 2x^2 \leq y \leq 2x \}$

2. $\oint_C x^2 y\, dx + 2xy\, dy$; C is the boundary of $R = \{ (x, y) : 0 \leq x \leq 1, x^2 \leq y \leq x \}$

3. $\oint_C 2y\, dx - 3x\, dy$; C is the circle $x^2 + y^2 = 1$

4. $\oint_C (e^{x^2} + y^2)\, dx + (e^{y^2} + x^2)\, dy$; C is the boundary of the triangle with vertices $(0, 0)$, $(4, 0)$ and $(0, 4)$

5. Is there a potential $F(x, y)$ for $\vec{f}(x, y) = (y^2 + 3x^2)\, \vec{\imath} + 2xy\, \vec{\jmath}$? If so, find one.

6. Is there a potential $F(x, y)$ for $\vec{f}(x, y) = (x^3 \cos(xy) + 2x \sin(xy))\, \vec{\imath} + x^2 y \cos(xy)\, \vec{\jmath}$? If so, find one.

7. Is there a potential $F(x, y)$ for $\vec{f}(x, y) = (8xy + 3)\, \vec{\imath} + 4(x^2 + y)\, \vec{\jmath}$? If so, find one.

8. Show that for any constants a, b and any closed simple curve C, $\oint_C a\, dx + b\, dy = 0$.

9. For the vector field $\vec{f}$ as in Example 2, show directly that $\oint_C \vec{f} \cdot d\vec{r} = 0$, where C is the boundary of the annulus $R = \{ (x, y) : 1/4 \leq x^2 + y^2 \leq 1 \}$ traversed so that R is always on the left.

10. Evaluate $\oint_C e^x \sin y\, dx + (y^3 + e^x \cos y)\, dy$, where C is the boundary of the rectangle with vertices $(1, -1)$, $(1, 1)$, $(-1, 1)$ and $(-1, -1)$, traversed counterclockwise.

11. For a region R bounded by a simple closed curve C, show that the area A of R is

$$A = -\oint_C y\, dx = \oint_C x\, dy = \frac{1}{2} \oint_C x\, dy - y\, dx,$$

where C is traversed so that R is always on the left. (*Hint: Use Green's Theorem and the fact that $A = \iint_R 1\, dA$.*)

15.4   Surface Integrals and the Divergence Theorem

In Section 15.1 we learned how to integrate along a curve. We will now learn how to perform integration over a *surface* in $\mathbb{R}^3$, such as a sphere or a paraboloid. Recall from Section 12.1 how we identified points (x, y, z) on a curve C in $\mathbb{R}^3$, parametrized by $x = x(t)$, $y = y(t)$, $z = z(t)$, $a \leq t \leq b$, with the terminal points of the position vector

$$\vec{r}(t) = x(t)\vec{\imath} + y(t)\vec{\jmath} + z(t)\vec{k} \qquad \text{for } t \text{ in } [a, b].$$

The idea behind a parametrization of a curve is that it "transforms" a subset of $\mathbb{R}^1$ (normally an interval $[a, b]$) into a curve in $\mathbb{R}^2$ or $\mathbb{R}^3$ (see Figure 15.12).
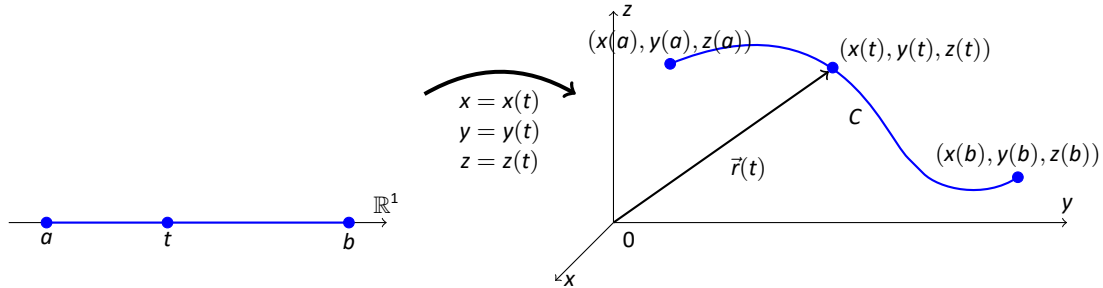
Figure 15.12: Parametrization of a curve C in $\mathbb{R}^3$

Similar to how we used a parametrization of a curve to define the line integral along the curve, we will use a parametrization of a surface to define a *surface integral*. We will use *two* variables, u and v, to parametrize a surface Σ in $\mathbb{R}^3$: $x = x(u, v)$, $y = y(u, v)$, $z = z(u, v)$, for (u, v) in some region R in $\mathbb{R}^2$ (see Figure 15.13).
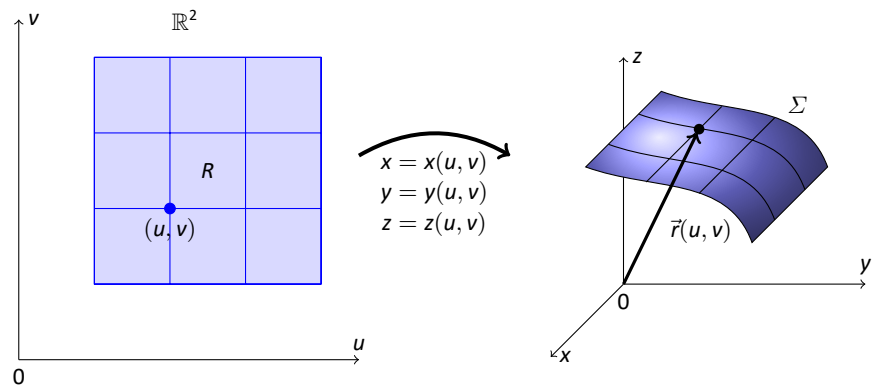
Figure 15.13: Parametrization of a surface Σ in $\mathbb{R}^3$

Notes:

In this case, the position vector of a point on the surface Σ is given by the vector-valued function

$$\vec{r}(u, v) = x(u, v)\vec{\imath} + y(u, v)\vec{\jmath} + z(u, v)\vec{k} \text{ for } (u, v) \text{ in } R.$$

Since $\vec{r}(u, v)$ is a function of two variables, define the partial derivatives $\frac{\partial \vec{r}}{\partial u}$ and $\frac{\partial \vec{r}}{\partial v}$ for (u, v) in R by

$$\frac{\partial \vec{r}}{\partial u}(u, v) = \frac{\partial x}{\partial u}(u, v)\vec{\imath} + \frac{\partial y}{\partial u}(u, v)\vec{\jmath} + \frac{\partial z}{\partial u}(u, v)\vec{k}, \text{ and}$$
$$\frac{\partial \vec{r}}{\partial v}(u, v) = \frac{\partial x}{\partial v}(u, v)\vec{\imath} + \frac{\partial y}{\partial v}(u, v)\vec{\jmath} + \frac{\partial z}{\partial v}(u, v)\vec{k}.$$

The parametrization of Σ can be thought of as "transforming" a region in $\mathbb{R}^2$ (in the *uv*-plane) into a 2-dimensional surface in $\mathbb{R}^3$. This parametrization of the surface is sometimes called a *patch*, based on the idea of "patching" the region *R* onto Σ in the grid-like manner shown in Figure 15.13.

In fact, those gridlines in *R* lead us to how we will define a surface integral over Σ. Along the vertical gridlines in *R*, the variable *u* is constant. So those lines get mapped to curves on Σ, and the variable *u* is constant along the position vector $\vec{r}(u, v)$. Thus, the tangent vector to those curves at a point (u, v) is $\frac{\partial \vec{r}}{\partial v}$. Similarly, the horizontal gridlines in *R* get mapped to curves on Σ whose tangent vectors are $\frac{\partial \vec{r}}{\partial u}$.

Now take a point (u, v) in *R* as, say, the lower left corner of one of the rectangular grid sections in *R*, as shown in Figure 15.13. Suppose that this rectangle has a small width and height of Δu and Δv, respectively. The corner points of that rectangle are (u, v), $(u + \Delta u, v)$, $(u + \Delta u, v + \Delta v)$ and $(u, v + \Delta v)$. So the area of that rectangle is $A = \Delta u\, \Delta v$. Then that rectangle gets mapped by the parametrization onto some section of the surface Σ which, for Δu and Δv small enough, will have a surface area (call it $d\sigma$) that is very close to the area of the parallelogram which has adjacent sides $\vec{r}(u + \Delta u, v) - \vec{r}(u, v)$ (corresponding to the line segment from (u, v) to $(u + \Delta u, v)$ in *R*) and $\vec{r}(u, v + \Delta v) - \vec{r}(u, v)$ (corresponding to the line segment from (u, v) to $(u, v + \Delta v)$ in *R*). By combining our usual notion of a partial derivative (see Definition 86 in Section 13.3) with that of the derivative of a vector-valued function (see Definition 73 in Section 12.2) applied to a function of two variables, we have

$$\frac{\partial \vec{r}}{\partial u} \approx \frac{\vec{r}(u + \Delta u, v) - \vec{r}(u, v)}{\Delta u}, \text{ and}$$
$$\frac{\partial \vec{r}}{\partial v} \approx \frac{\vec{r}(u, v + \Delta v) - \vec{r}(u, v)}{\Delta v},$$

Notes:

and so the surface area element $d\sigma$ is approximately

$$\|(\vec{r}(u + \Delta u, v) - \vec{r}(u, v)) \times (\vec{r}(u, v + \Delta v) - \vec{r}(u, v))\|$$

$$\approx \left\|(\Delta u \frac{\partial \vec{r}}{\partial u}) \times (\Delta v \frac{\partial \vec{r}}{\partial v})\right\| = \left\|\frac{\partial \vec{r}}{\partial u} \times \frac{\partial \vec{r}}{\partial v}\right\| \Delta u\, \Delta v$$

by Equation (11.5). Thus, the total surface area S of Σ is approximately the sum of all the quantities $\left\|\frac{\partial \vec{r}}{\partial u} \times \frac{\partial \vec{r}}{\partial v}\right\| \Delta u\, \Delta v$, summed over the rectangles in R. Taking the limit of that sum as the diagonal of the largest rectangle goes to 0 gives

$$S = \iint_R \left\|\frac{\partial \vec{r}}{\partial u} \times \frac{\partial \vec{r}}{\partial v}\right\| du\, dv.$$

We will write the double integral on the right using the special notation

$$\iint_\Sigma d\sigma = \iint_R \left\|\frac{\partial \vec{r}}{\partial u} \times \frac{\partial \vec{r}}{\partial v}\right\| du\, dv.$$

This is a special case of a *surface integral* over the surface Σ, where the surface area element $d\sigma$ can be thought of as $1\, d\sigma$. Replacing 1 by a general real-valued function $f(x, y, z)$ defined in $\mathbb{R}^3$, we have the following:

> **Definition 114　　Scalar Surface Integral**
> Let Σ be a surface in $\mathbb{R}^3$ parametrized by $x = x(u, v)$, $y = y(u, v)$, $z = z(u, v)$, for (u, v) in some region R in $\mathbb{R}^2$. Let $\vec{r}(u, v) = x(u, v)\vec{\imath} + y(u, v)\vec{\jmath} + z(u, v)\vec{k}$ be the position vector for any point on Σ, and let $f(x, y, z)$ be a real-valued function defined on some subset of $\mathbb{R}^3$ that contains Σ. The **surface integral** of $f(x, y, z)$ over Σ is
>
> $$\iint_\Sigma f(x, y, z)\, d\sigma = \iint_R f(x(u, v), y(u, v), z(u, v)) \left\|\frac{\partial \vec{r}}{\partial u} \times \frac{\partial \vec{r}}{\partial v}\right\| du\, dv.$$
>
> In particular, the surface area S of Σ is
>
> $$S = \iint_\Sigma 1\, d\sigma.$$

Sometimes, the notation $\oiint_\Sigma f(x, y, z)\, d\sigma$ is used instead of $\iint_\Sigma f(x, y, z)\, d\sigma$ when Σ is a closed surface. Especially in physics texts, it is common to see simply $\oint_\Sigma$ instead of $\oiint_\Sigma$.

Notes:

Example 1   Computing a Surface Integral

A *torus T* is a surface obtained by revolving a circle of radius *a* in the *yz*-plane around the *z*-axis, where the circle's center is at a distance *b* from the z-axis ($0 < a < b$), as in Figure 15.14. Find the surface area of *T*.

Circle in the *yz*-plane   Torus *T*
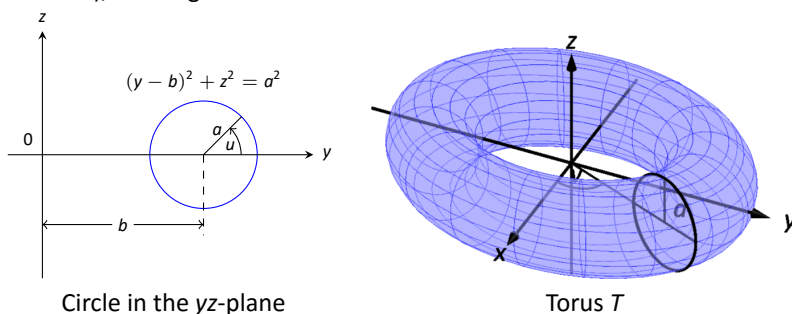
Figure 15.14: Creating a torus

SOLUTION   For any point on the circle, the line segment from the center of the circle to that point makes an angle *u* with the *y*-axis in the positive *y* direction (see Figure 15.14(a)). And as the circle revolves around the *z*-axis, the line segment from the origin to the center of that circle sweeps out an angle *v* with the positive *x*-axis (see Figure 15.14(b)). Thus, the torus can be parametrized as:

$$x = (b + a\cos u)\cos v, \quad y = (b + a\cos u)\sin v, \quad z = a\sin u,$$

where $0 \le u \le 2\pi$, and $0 \le v \le 2\pi$. So for the position vector

$$\vec{r}(u,v) = x(u,v)\vec{i} + y(u,v)\vec{j} + z(u,v)\vec{k}$$
$$= (b + a\cos u)\cos v\,\vec{i} + (b + a\cos u)\sin v\,\vec{j} + a\sin u\,\vec{k}$$

we see that

$$\frac{\partial \vec{r}}{\partial u} = -a\sin u\,\cos v\,\vec{i} - a\sin u\,\sin v\,\vec{j} + a\cos u\,\vec{k}$$

$$\frac{\partial \vec{r}}{\partial v} = -(b + a\cos u)\sin v\,\vec{i} + (b + a\cos u)\cos v\,\vec{j} + 0\vec{k},$$

Notes:

and so computing the cross product gives

$$\frac{\partial \vec{r}}{\partial u} \times \frac{\partial \vec{r}}{\partial v} =$$

$$-a(b+a\cos u)\cos v\,\cos u\,\vec{\imath}-a(b+a\cos u)\sin v\,\cos u\,\vec{\jmath}-a(b+a\cos u)\sin u\,\vec{k},$$

which has magnitude

$$\left\|\frac{\partial \vec{r}}{\partial u} \times \frac{\partial \vec{r}}{\partial v}\right\| = a(b + a\cos u).$$

Thus, the surface area of T is

$$S = \iint_{\Sigma} 1\, d\sigma$$

$$= \int_0^{2\pi} \int_0^{2\pi} \left\|\frac{\partial \vec{r}}{\partial u} \times \frac{\partial \vec{r}}{\partial v}\right\|\, du\, dv$$

$$= \int_0^{2\pi} \int_0^{2\pi} a(b + a\cos u)\, du\, dv$$

$$= \int_0^{2\pi} \left(abu + a^2 \sin u \,\bigg|_{u=0}^{u=2\pi}\right) dv$$

$$= \int_0^{2\pi} 2\pi ab\, dv$$

$$= 4\pi^2 ab$$

Since $\frac{\partial \vec{r}}{\partial u}$ and $\frac{\partial \vec{r}}{\partial v}$ are tangent to the surface Σ (i.e. lie in the tangent plane to Σ at each point on Σ), then their cross product $\frac{\partial \vec{r}}{\partial u} \times \frac{\partial \vec{r}}{\partial v}$ is perpendicular to the tangent plane to the surface at each point of Σ. Thus,

$$\iint_{\Sigma} f(x, y, z)\, d\sigma = \iint_{R} f(x(u, v), y(u, v), z(u, v)) \, \|\vec{n}\| \, d\sigma,$$

where $\vec{n} = \frac{\partial \vec{r}}{\partial u} \times \frac{\partial \vec{r}}{\partial v}$. We say that $\vec{n}$ is a **normal vector** to Σ.

Recall that normal vectors to a plane can point in two opposite directions. By an **outward unit normal vector** to a surface Σ, we will mean the unit vector that is normal to Σ and points away from the "top" (or "outer" part) of the surface. This is a hazy definition, but the picture in Figure 15.15 gives a better idea of what outward normal vectors look like, in the case of a sphere. With this idea in mind, we make the following definition of a surface integral of a 3-dimensional *vector field* over a surface:
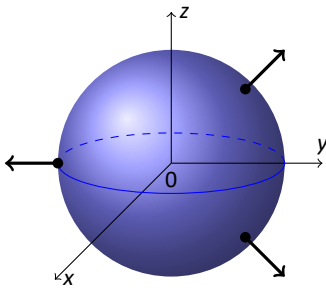
Figure 15.15: Outward Unit Normal Vectors

Notes:

> **Definition 115   Vector Surface Integral**
>
> Let Σ be a surface in $\mathbb{R}^3$ and let $\vec{f}(x, y, z) = f_1(x, y, z)\vec{\imath} + f_2(x, y, z)\vec{\jmath} + f_3(x, y, z)\vec{k}$ be a vector field defined on some subset of $\mathbb{R}^3$ that contains Σ. The **surface integral** of $\vec{f}$ over Σ is
>
> $$\iint_{\Sigma} \vec{f} \cdot d\vec{\sigma} = \iint_{R} \vec{f} \cdot \vec{n}\, d\sigma, \qquad (15.10)$$
>
> where, at any point on Σ, $\vec{n}$ is the outward unit normal vector to Σ.

Sometimes, the notation $\oiint_{\Sigma} \vec{f} \cdot d\vec{\sigma}$ is used instead of $\iint_{\Sigma} \vec{f} \cdot d\vec{\sigma}$ when Σ is a closed surface. Especially in physics texts, it is common to see simply $\oint_{\Sigma}$ instead of $\oiint_{\Sigma}$.

Note in the above definition that the dot product inside the integral on the right is a real-valued function, and hence we can use Definition 114 to evaluate the integral.

Example 2   Evaluating a Surface Integral

Evaluate the surface integral $\iint_{\Sigma} \vec{f} \cdot d\vec{\sigma}$, where $\vec{f}(x, y, z) = yz\vec{\imath} + xz\vec{\jmath} + xy\vec{k}$ and Σ is the part of the plane $x + y + z = 1$ with $x \geq 0$, $y \geq 0$, and $z \geq 0$, with the outward unit normal $\vec{n}$ pointing in the positive z direction (see Figure 15.16).

SOLUTION   Since the vector $\vec{v} = (1, 1, 1)$ is normal to the plane $x + y + z = 1$ (why?), then dividing $\vec{v}$ by its length yields the outward unit normal vector $\vec{n} = \left(\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}} \right)$. We now need to parametrize Σ. As we can see from Figure 15.16, projecting Σ onto the xy-plane yields a triangular region $R = \{ (x, y) : 0 \leq x \leq 1, 0 \leq y \leq 1 - x \}$. Thus, using (u, v) instead of (x, y), we see that

$$x = u, y = v, z = 1 - (u + v), \text{ for } 0 \leq u \leq 1, 0 \leq v \leq 1 - u$$

is a parametrization of Σ over R (since $z = 1 - (x + y)$ on Σ). So on Σ,

$$\vec{f} \cdot \vec{n} = (yz, xz, xy) \cdot \left(\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}} \right) = \frac{1}{\sqrt{3}}(yz + xz + xy)$$

$$= \frac{1}{\sqrt{3}}((x + y)z + xy) = \frac{1}{\sqrt{3}}((u + v)(1 - (u + v)) + uv)$$

$$= \frac{1}{\sqrt{3}}((u + v) - (u + v)^2 + uv)$$

for (u, v) in R, and for $\vec{r}(u, v) = x(u, v)\vec{\imath} + y(u, v)\vec{\jmath} + z(u, v)\vec{k} = u\vec{\imath} + v\vec{\jmath} + (1 - (u + v))\vec{k}$ we have

$$\frac{\partial \vec{r}}{\partial u} \times \frac{\partial \vec{r}}{\partial v} = (1, 0, -1) \times (0, 1, -1) = (1, 1, 1) \quad \Rightarrow \quad \left\| \frac{\partial \vec{r}}{\partial u} \times \frac{\partial \vec{r}}{\partial v} \right\| = \sqrt{3}.$$

Figure 15.16: The surface in Example 2

Notes:

Thus, integrating over R using vertical slices (e.g. as indicated by the dashed line in Figure 15.16) gives

$$
\iint_{\Sigma} \vec{f} \cdot d\vec{\sigma} = \iint_{\Sigma} \vec{f} \cdot \vec{n} \, d\sigma
$$

$$
= \iint_{R} (\vec{f}(x(u,v), y(u,v), z(u,v)) \cdot \vec{n}) \left\| \frac{\partial \vec{r}}{\partial u} \times \frac{\partial \vec{r}}{\partial v} \right\| dv \, du
$$

$$
= \int_{0}^{1} \int_{0}^{1-u} \frac{1}{\sqrt{3}} ((u+v) - (u+v)^2 + uv) \sqrt{3} \, dv \, du
$$

$$
= \int_{0}^{1} \left(\frac{(u+v)^2}{2} - \frac{(u+v)^3}{3} + \frac{uv^2}{2} \Big|_{v=0}^{v=1-u} \right) du
$$

$$
= \int_{0}^{1} \left(\frac{1}{6} + \frac{u}{2} - \frac{3u^2}{2} + \frac{5u^3}{6} \right) du
$$

$$
= \frac{u}{6} + \frac{u^2}{4} - \frac{u^3}{2} + \frac{5u^4}{24} \Big|_{0}^{1} = \frac{1}{8}.
$$

Notice that we divided $\vec{n}$ by its norm, and later multiplied by the same factor with $\left\| \frac{\partial \vec{r}}{\partial u} \times \frac{\partial \vec{r}}{\partial v} \right\|$. This is generally the case:

$$
\vec{n} = \frac{\frac{\partial \vec{r}}{\partial u} \times \frac{\partial \vec{r}}{\partial v}}{\left\| \frac{\partial \vec{r}}{\partial u} \times \frac{\partial \vec{r}}{\partial v} \right\|}
$$

$$
\sigma = \left\| \frac{\partial \vec{r}}{\partial u} \times \frac{\partial \vec{r}}{\partial v} \right\| dv \, du
$$

$$
\vec{n} d\sigma = \frac{\partial \vec{r}}{\partial u} \times \frac{\partial \vec{r}}{\partial v} \, dv \, du,
$$

which simplifies the calculation.

Computing surface integrals can often be tedious, especially when the formula for the outward unit normal vector at each point of Σ changes. The following theorem provides an easier way in the case when Σ is a **closed surface**, that is, when Σ encloses a bounded solid in $\mathbb{R}^3$. For example, spheres, cubes, and ellipsoids are closed surfaces, but planes and paraboloids are not.

Notes:

> **Theorem 139   Divergence Theorem**
> Let Σ be a closed surface in $\mathbb{R}^3$ which bounds a solid S, and let $\vec{f}(x, y, z) = f_1(x, y, z)\vec{\imath} + f_2(x, y, z)\vec{\jmath} + f_3(x, y, z)\vec{k}$ be a vector field defined on some subset of $\mathbb{R}^3$ that contains Σ. Then
>
> $$\iint_{\Sigma} \vec{f} \cdot d\vec{\sigma} = \iiint_S \operatorname{div} \vec{f} dV, \qquad (15.11)$$
>
> where
>
> $$\operatorname{div} \vec{f} = \frac{\partial f_1}{\partial x} + \frac{\partial f_2}{\partial y} + \frac{\partial f_3}{\partial z} \qquad (15.12)$$
>
> is called the **divergence** of $\vec{f}$.

The proof of the Divergence Theorem is very similar to the proof of Green's Theorem, i.e. it is first proved for the simple case when the solid S is bounded above by one surface, bounded below by another surface, and bounded laterally by one or more surfaces. The proof can then be extended to more general solids.

In Definition 95 in Section 13.6, we defined the operator ∇ as a *vector* in $\mathbb{R}^3$, namely

$$\nabla = \frac{\partial}{\partial x}\vec{\imath} + \frac{\partial}{\partial y}\vec{\jmath} + \frac{\partial}{\partial z}\vec{k}.$$

This definition allows us to write $\operatorname{div} \vec{f} = \nabla \cdot \vec{f}$.

Example 3   Using the Divergence Theorem
Evaluate $\iint_{\Sigma} \vec{f} \cdot d\vec{\sigma}$, where $\vec{f}(x, y, z) = x\vec{\imath} + y\vec{\jmath} + z\vec{k}$ and Σ is the unit sphere $x^2 + y^2 + z^2 = 1$.

SOLUTION   We see that $\operatorname{div} \vec{f} = 1 + 1 + 1 = 3$, so

$$\iint_{\Sigma} \vec{f} \cdot d\vec{\sigma} = \iiint_S \operatorname{div} \vec{f} dV = \iiint_S 3 dV$$

$$= 3 \iiint_S 1 dV = 3 \operatorname{vol}(S) = 3 \cdot \frac{4\pi(1)^3}{3} = 4\pi.$$

In physical applications, the surface integral $\iint_{\Sigma} \vec{f} \cdot d\vec{\sigma}$ is often referred to as the **flux** of $\vec{f}$ through the surface Σ. For example, if $\vec{f}$ represents the velocity field of a fluid, then the flux is the net quantity of fluid to flow through the surface Σ per unit time. A positive flux means there is a net flow *out* of the surface (i.e. in the direction of the outward unit normal vector $\vec{n}$), while a negative flux indicates a net flow inward (in the direction of $-\vec{n}$).

Notes:

The term divergence comes from interpreting div $\vec{f}$ as a measure of how much a vector field "diverges" from a point. This is best seen by using another definition of div $\vec{f}$ which is equivalent to the definition given by Equation (15.12). Namely, for a point (x, y, z) in $\mathbb{R}^3$,

$$\text{div}\,\vec{f}(x, y, z) = \lim_{V \to 0} \frac{1}{V} \iint_{\Sigma} \vec{f} \cdot d\vec{\sigma}, \tag{15.13}$$

where V is the volume enclosed by a closed surface Σ around the point (x, y, z). In the limit, $V \to 0$ means that we take smaller and smaller closed surfaces around (x, y, z), which means that the volumes they enclose are going to zero. It can be shown that this limit is independent of the shapes of those surfaces. Notice that the limit being taken is of the ratio of the flux through a surface to the volume enclosed by that surface, which gives a rough measure of the flow "leaving" a point, as we mentioned. Vector fields which have zero divergence are often called *solenoidal* fields.

The following theorem is a simple consequence of Equation (15.13).

Theorem 140   Zero Flux

If the flux of a vector field $\vec{f}$ is zero through every closed surface containing a given point, then div $\vec{f} = 0$ at that point.

Proof
By Equation (15.13), at the given point (x, y, z) we have

$\text{div}\,\vec{f}(x, y, z) = \displaystyle\lim_{V \to 0} \frac{1}{V} \iint_{\Sigma} \vec{f} \cdot d\vec{\sigma}$ for closed surfaces Σ containing (x, y, z), so

$\qquad = \displaystyle\lim_{V \to 0} \frac{1}{V} (0)$ by our assumption that the flux through each Σ is zero, so

$\qquad = \displaystyle\lim_{V \to 0} 0$

$\qquad = 0.$ $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad$ $\square$

This section and the previous introduced four new types of integrals, which we gather in Key Idea 64.

Notes:

Key Idea 64   Integrating Parameterized Curves and Surfaces

	$\vec{c} : \mathbb{R} \to \mathbb{R}^3$ parameterizes a curve	$\vec{r} : \mathbb{R}^2 \to \mathbb{R}^3$ parameterizes a surface
$f : \mathbb{R}^3 \to \mathbb{R}$	scalar line integral: $\int_{\vec{c}(D)} f\, ds$ $\quad = \int_D f(\vec{c}(t))\, \|\vec{c}\,'(t)\|\, dt$	scalar surface integral: $\iint_{\vec{r}(D)} f\, d\sigma$ $\quad = \iint_D f(\vec{r}(u,v))\, \|\vec{r}_u \times \vec{r}_v\|\, du\, dv$
$\vec{F} : \mathbb{R}^3 \to \mathbb{R}^3$	vector line integral: $\int_{\vec{c}(D)} \vec{F} \cdot d\vec{r}$ $\quad = \int_D \vec{F}(\vec{c}(t)) \cdot \vec{c}\,'(t)\, dt$	vector surface integral: $\iint_{\vec{r}(D)} \vec{F} \cdot d\vec{\sigma}$ $\quad = \iint_D \vec{F}(\vec{r}(u,v)) \cdot (\vec{r}_u \times \vec{r}_v)\, du\, dv$

If we are in the bottom left entry, where $\vec{c} : \mathbb{R} \to \mathbb{R}^3$ and $\vec{F} : \mathbb{R}^3 \to \mathbb{R}^3$, we may be able to use the Fundamental Theorem for Gradient Vector Fields: $\int_C \nabla\phi \cdot d\vec{s} = \phi(Q) - \phi(P)$.

Notes:

Exercises 15.4

Problems

In Exercises 1–4, use the Divergence Theorem to evaluate the surface integral $\iint_\Sigma \vec{f} \cdot d\sigma$ of the given vector field $\vec{f}(x, y, z)$ over the surface Σ.

1. $\vec{f}(x, y, z) = x\vec{\imath} + 2y\vec{\jmath} + 3z\vec{k}$, $\Sigma : x^2 + y^2 + z^2 = 9$

2. $\vec{f}(x, y, z) = x\vec{\imath} + y\vec{\jmath} + z\vec{k}$, Σ : boundary of the solid cube $S = \{(x, y, z) : 0 \le x, y, z \le 1\}$

3. $\vec{f}(x, y, z) = x^3\vec{\imath} + y^3\vec{\jmath} + z^3\vec{k}$, $\Sigma : x^2 + y^2 + z^2 = 1$

4. $\vec{f}(x, y, z) = 2\vec{\imath} + 3\vec{\jmath} + 5\vec{k}$, $\Sigma : x^2 + y^2 + z^2 = 1$

5. Show that the flux of any constant vector field through any closed surface is zero.

6. Evaluate the surface integral from Exercise 2 *without* using the Divergence Theorem, i.e. using only Definition 114, as in Example 2. Note that there will be a different outward unit normal vector to each of the six faces of the cube.

7. Evaluate the surface integral $\iint_\Sigma \vec{f} \cdot d\sigma$, where $\vec{f}(x, y, z) = x^2\vec{\imath} + xy\vec{\jmath} + z\vec{k}$ and Σ is the part of the plane $6x + 3y + 2z = 6$ with $x \ge 0$, $y \ge 0$, and $z \ge 0$, with the outward unit normal $\vec{n}$ pointing in the positive z direction.

8. Use a surface integral to show that the surface area of a sphere of radius r is $4\pi r^2$. (*Hint: Use spherical coordinates to parametrize the sphere.*)

9. Use a surface integral to show that the surface area of a right circular cone of radius R and height h is $\pi R\sqrt{h^2 + R^2}$. (*Hint: Use the parametrization $x = r\cos\theta$, $y = r\sin\theta$, $z = \frac{h}{R}r$, for $0 \le r \le R$ and $0 \le \theta \le 2\pi$.*)

10. The ellipsoid $\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1$ can be parametrized using *ellipsoidal coordinates*

$$x = a\sin\phi\,\cos\theta, y = b\sin\phi\,\sin\theta, z = c\cos\phi,$$

for $0 \le \theta \le 2\pi$ and $0 \le \phi \le \pi$. Show that the surface area S of the ellipsoid is

$$S = \int_0^\pi \int_0^{2\pi} \sin\phi\sqrt{a^2 b^2 \cos^2\phi + c^2(a^2\sin^2\theta + b^2\cos^2\theta)\sin^2\phi}\, d\theta\, d\phi.$$

(Note: The above double integral can not be evaluated by elementary means. For specific values of a, b and c it can be evaluated using numerical methods. An alternative is to express the surface area in terms of *elliptic integrals*.)

11. Use Definition 114 to prove that the surface area S over a region R in $\mathbb{R}^2$ of a surface $z = f(x, y)$ is given by the formula

$$S = \iint_R \sqrt{1 + \left(\frac{\partial f}{\partial x}\right)^2 + \left(\frac{\partial f}{\partial y}\right)^2}\, dA.$$

(*Hint: Think of the parametrization of the surface.*)

15.5 Stokes' Theorem

So far the only types of line integrals which we have discussed are those along curves in $\mathbb{R}^2$. But the definitions and properties which were covered in Sections 15.1 and 15.2 can easily be extended to include functions of three variables, so that we can now discuss line integrals along curves in $\mathbb{R}^3$.

Definition 116   **Scalar Line Integral**

For a real-valued function $f(x, y, z)$ and a curve C in $\mathbb{R}^3$, parametrized by $x = x(t)$, $y = y(t)$, $z = z(t)$, $a \le t \le b$, the **line integral of $f(x, y, z)$ along C with respect to arc length s** is

$$\int_C f(x, y, z)\, ds = \int_a^b f(x(t), y(t), z(t))\, \sqrt{x'(t)^2 + y'(t)^2 + z'(t)^2}\ dt.$$

The **line integral of $f(x, y, z)$ along C with respect to x** is

$$\int_C f(x, y, z)\, dx = \int_a^b f(x(t), y(t), z(t))\, x'(t)\, dt.$$

The **line integral of $f(x, y, z)$ along C with respect to y** is

$$\int_C f(x, y, z)\, dy = \int_a^b f(x(t), y(t), z(t))\, y'(t)\, dt.$$

The **line integral of $f(x, y, z)$ along C with respect to z** is

$$\int_C f(x, y, z)\, dz = \int_a^b f(x(t), y(t), z(t))\, z'(t)\, dt.$$

Similar to the two-variable case, if $f(x, y, z) \ge 0$ then the line integral $\int_C f(x, y, z)\, ds$ can be thought of as the total area of the "picket fence" of height $f(x, y, z)$ at each point along the curve C in $\mathbb{R}^3$.

Vector fields in $\mathbb{R}^3$ are defined in a similar fashion to those in $\mathbb{R}^2$, which allows us to define the line integral of a vector field along a curve in $\mathbb{R}^3$.

Notes:

Definition 117   Vector Line Integral
For a vector field $\vec{f}(x, y, z) = P(x, y, z)\,\vec{\imath} + Q(x, y, z)\,\vec{\jmath} + R(x, y, z)\,\vec{k}$ and a curve C in $\mathbb{R}^3$ with a smooth parametrization $x = x(t)$, $y = y(t)$, $z = z(t)$, $a \leq t \leq b$, the **line integral of f along** C is

$$\int_C \vec{f} \cdot d\vec{r} = \int_C P(x, y, z)\,dx + \int_C Q(x, y, z)\,dy + \int_C R(x, y, z)\,dz$$

$$= \int_a^b \vec{f}(x(t), y(t), z(t)) \cdot \vec{r}'(t)\,dt,$$

where $\vec{r}(t) = x(t)\,\vec{\imath} + y(t)\,\vec{\jmath} + z(t)\,\vec{k}$ is the position vector for points on C.

Similar to the two-variable case, if $\vec{f}(x, y, z)$ represents the force applied to an object at a point (x, y, z) then the line integral $\int_C f \cdot d\vec{r}$ represents the work done by that force in moving the object along the curve C in $\mathbb{R}^3$.

Some of the most important results we will need for line integrals in $\mathbb{R}^3$ are stated below without proof (the proofs are similar to their two-variable equivalents).

Theorem 141   Vector Line Integral Along a Curve
For a vector field $\vec{f}(x, y, z) = P(x, y, z)\,\vec{\imath} + Q(x, y, z)\,\vec{\jmath} + R(x, y, z)\,\vec{k}$ and a curve C with a smooth parametrization $x = x(t)$, $y = y(t)$, $z = z(t)$, $a \leq t \leq b$ and position vector $\vec{r}(t) = x(t)\,\vec{\imath} + y(t)\,\vec{\jmath} + z(t)\,\vec{k}$,

$$\int_C \vec{f} \cdot d\vec{r} = \int_C \vec{f} \cdot \vec{T}\,ds,$$

where $\vec{T}(t) = \dfrac{\vec{r}'(t)}{\|\vec{r}'(t)\|}$ is the unit tangent vector to C at $(x(t), y(t), z(t))$.

Notes:

Theorem 142   Fundamental Theorem of Line Integrals in Three Dimensions

Let $\vec{f}(x,y,z) = P(x,y,z)\,\vec{\imath} + Q(x,y,z)\,\vec{\jmath} + R(x,y,z)\,\vec{k}$ be a vector field in some solid S, with P, Q and R continuously differentiable functions on S. Let C be a smooth curve in S parametrized by $x = x(t)$, $y = y(t)$, $z = z(t)$, $a \leq t \leq b$. Suppose that there is a real-valued function $F(x,y,z)$ such that $\nabla F = \vec{f}$ on S. Then

$$\int_C \vec{f} \cdot d\vec{r} = F(B) - F(A), \qquad (15.14)$$

where $A = (x(a), y(a), z(a))$ and $B = (x(b), y(b), z(b))$ are the endpoints of C.

Theorem 143   Zero Line Integral

If a vector field $\vec{f}$ has a potential in a solid S without holes, then $\oint_C \vec{f} \cdot d\vec{r} = 0$ for any closed curve C in S (i.e., $\oint_C \nabla F \cdot d\vec{r} = 0$ for any real-valued function $F(x,y,z)$).

Watch the video:
Fundamental Theorem for Line Integrals at
`https://youtu.be/xVhHow_usMQ`

Example 1   Evaluating a Line Integral

Let $f(x,y,z) = z$ and let C be the curve in $\mathbb{R}^3$ parametrized by

$$x = t\sin t, \quad y = t\cos t, \quad z = t, \quad 0 \leq t \leq 8\pi.$$

Evaluate $\int_C f(x,y,z)\,ds$. (Note: C is called a *conical helix*. See Figure 15.17).

Notes:
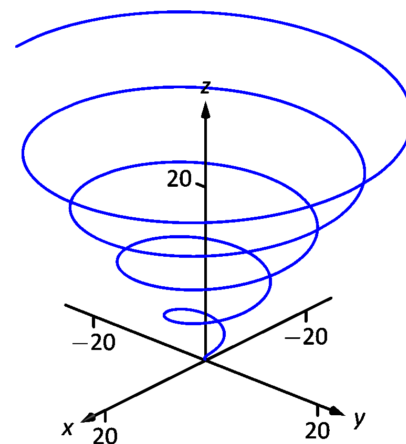
Figure 15.17: Conical helix C for Example 1.

SOLUTION Since $x'(t) = \sin t + t \cos t$, $y'(t) = \cos t - t \sin t$, and $z'(t) = 1$, we have

$$
\begin{aligned}
x'(t)^2 + y'(t)^2 + z'(t)^2 &= (\sin^2 t + 2t \sin t \cos t + t^2 \cos^2 t) \\
&\quad + (\cos^2 t - 2t \sin t \cos t + t^2 \sin^2 t) + 1 \\
&= t^2(\sin^2 t + \cos^2 t) + \sin^2 t + \cos^2 t + 1 \\
&= t^2 + 2,
\end{aligned}
$$

so since $f(x(t), y(t), z(t)) = z(t) = t$ along the curve C, then

$$
\begin{aligned}
\int_C f(x, y, z)\, ds &= \int_0^{8\pi} f(x(t), y(t), z(t))\, \sqrt{x'(t)^2 + y'(t)^2 + z'(t)^2}\, dt \\
&= \int_0^{8\pi} t\, \sqrt{t^2 + 2}\, dt \\
&= \left(\frac{1}{3}(t^2 + 2)^{3/2} \right) \Bigg|_0^{8\pi} = \frac{1}{3}\left((64\pi^2 + 2)^{3/2} - 2\sqrt{2} \right).
\end{aligned}
$$

Example 2   **Evaluating a Line Integral**

Let $\vec{f}(x, y, z) = x\,\vec{\imath} + y\,\vec{\jmath} + 2z\,\vec{k}$ be a vector field in $\mathbb{R}^3$. Using the same curve C from Example 1, evaluate $\int_C f \cdot d\vec{r}$.

SOLUTION It is easy to see that $F(x, y, z) = \frac{x^2}{2} + \frac{y^2}{2} + z^2$ is a potential for $\vec{f}(x, y, z)$ (i.e., $\nabla F = \vec{f}$). So by Theorem 142 we know that

$$
\int_C \vec{f} \cdot d\vec{r} = F(B) - F(A), \text{ where } A = (x(0), y(0), z(0)) \text{ and } B = (x(8\pi), y(8\pi), z(8\pi)), \text{ so}
$$

$$
\begin{aligned}
&= F(8\pi \sin 8\pi, 8\pi \cos 8\pi, 8\pi) - F(0 \sin 0, 0 \cos 0, 0) \\
&= F(0, 8\pi, 8\pi) - F(0, 0, 0) \\
&= 0 + \frac{(8\pi)^2}{2} + (8\pi)^2 - (0 + 0 + 0) = 96\pi^2.
\end{aligned}
$$

We will now discuss a generalization of Green's Theorem in $\mathbb{R}^2$ to *orientable* surfaces in $\mathbb{R}^3$, called *Stokes' Theorem*. A surface Σ in $\mathbb{R}^3$ is **orientable** if there is a continuous vector field $\vec{n}$ in $\mathbb{R}^3$ such that $\vec{n}$ is nonzero and normal to Σ (i.e. perpendicular to the tangent plane) at each point of Σ. We say that such an $\vec{n}$ is a *normal vector field*.

For example, the unit sphere $x^2 + y^2 + z^2 = 1$ is orientable, since the continuous vector field $\vec{n}(x, y, z) = x\,\vec{\imath} + y\,\vec{\jmath} + z\,\vec{k}$ is nonzero and normal to the sphere at
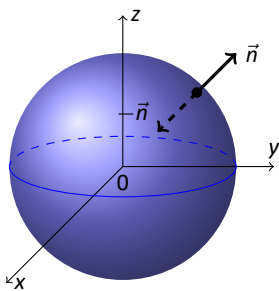
Figure 15.18: An orientation of the sphere.

Notes:

each point. In fact, $-\vec{n}(x, y, z)$ is another normal vector field (see Figure 15.18). We see in this case that $\vec{n}(x, y, z)$ is what we have called an outward normal vector, and $-\vec{n}(x, y, z)$ is an *inward* normal vector. These "outward" and "inward" normal vector fields on the sphere correspond to an "outer" and "inner" side, respectively, of the sphere. That is, we say that the sphere is a *two-sided* surface. Roughly, "two-sided" means "orientable". Other examples of two-sided, and hence orientable, surfaces are cylinders, paraboloids, ellipsoids, and planes.

You may be wondering what kind of surface would *not* have two sides. An example is the **Möbius strip**, which is constructed by taking a thin rectangle and connecting its ends at the opposite corners, resulting in a "twisted" strip (see Figure 15.19).

Connect *A* to *A* and *B* to *B* along the ends   Not orientable
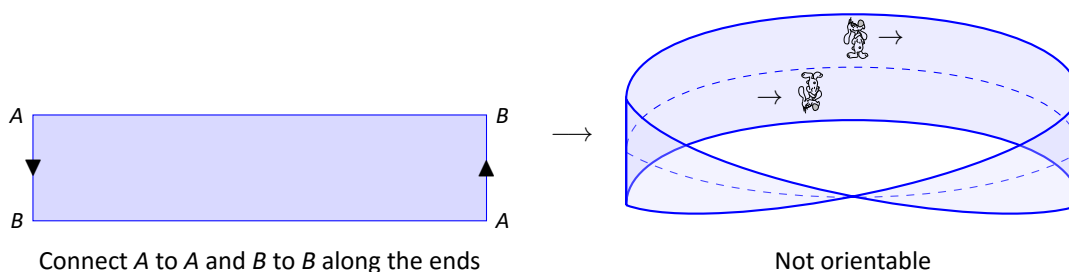
Figure 15.19: Möbius strip

If you imagine walking along a line down the center of the Möbius strip, as in Figure 15.19(b), then you arrive back at the same place from which you started but upside down! That is, your *orientation* changed even though your motion was continuous along that center line. Informally, thinking of your vertical direction as a normal vector field along the strip, there is a discontinuity at your starting point (and, in fact, at every point) since your vertical direction takes two different values there. The Möbius strip has only *one side*, and hence is nonorientable.

For an orientable surface Σ which has a boundary curve *C*, pick a unit normal vector $\vec{n}$ such that if you walked along *C* with your head pointing in the direction of $\vec{n}$, then the surface would be on your left. We say in this situation that $\vec{n}$ is a *positive unit normal vector* and that *C* is traversed $\vec{n}$-*positively*. We can now state Stokes' Theorem:

Notes:

<div style="border: 1px solid green;">

Theorem 144   Stokes' Theorem

Let Σ be an orientable surface in $\mathbb{R}^3$ whose boundary is a simple closed curve C, and let $\vec{f}(x, y, z) = P(x, y, z)\vec{\imath} + Q(x, y, z)\vec{\jmath} + R(x, y, z)\vec{k}$ be a smooth vector field defined on some subset of $\mathbb{R}^3$ that contains Σ. Then

$$\oint_C \vec{f} \cdot d\vec{r} = \iint_\Sigma (\text{curl}\,\vec{f}) \cdot \vec{n}\,d\sigma, \qquad (15.15)$$

where

$$\text{curl}\,\vec{f} = \left(\frac{\partial R}{\partial y} - \frac{\partial Q}{\partial z} \right)\vec{\imath} + \left(\frac{\partial P}{\partial z} - \frac{\partial R}{\partial x} \right)\vec{\jmath} + \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} \right)\vec{k}, \quad (15.16)$$

$\vec{n}$ is a positive unit normal vector over Σ, and C is traversed $\vec{n}$-positively.

</div>

The formula for $\text{curl}\,\vec{f}$ is unfortunately complicated. If we recall that we have defined the operator ∇ as a *vector* in $\mathbb{R}^3$ by

$$\nabla = \frac{\partial}{\partial x}\,\vec{\imath} + \frac{\partial}{\partial y}\,\vec{\jmath} + \frac{\partial}{\partial z}\,\vec{k},$$

then we can write $\text{curl}\,\vec{f} = \nabla \times \vec{f}$.

Proof

As the general case is beyond the scope of this text, we will prove the theorem only for the special case where Σ is the graph of $z = z(x, y)$ for some smooth real-valued function $z(x, y)$, with (x, y) varying over a region D in $\mathbb{R}^2$. Projecting Σ onto the xy-plane, we see that the closed curve C (the boundary curve of Σ) projects onto a closed curve C_D which is the boundary curve of D (see Figure 15.20). Assuming that C has a smooth parametrization, its projection C_D in the xy-plane also has a smooth parametrization, say

$$C_D : x = x(t), y = y(t), a \le t \le b,$$

and so C can be parametrized (in $\mathbb{R}^3$) as

$$C : x = x(t), y = y(t), z = z(x(t), y(t)), a \le t \le b,$$

since the curve C is part of the surface $z = z(x, y)$. Now, by the Chain Rule (Theorem 110 in Section 13.5), for $z = z(x(t), y(t))$ as a function of t, we know that

$$z'(t) = \frac{\partial z}{\partial x}\,x'(t) + \frac{\partial z}{\partial y}\,y'(t),$$
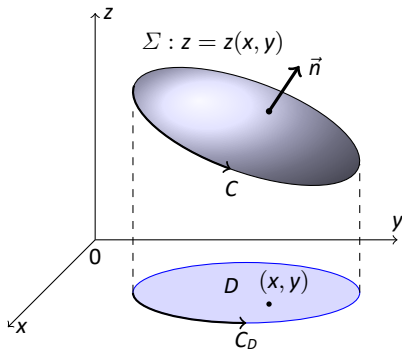
Figure 15.20: A particular case of Stokes' Theorem

Notes:

and so

$$
\oint_C \vec{f} \cdot d\vec{r} = \int_C P(x, y, z)\, dx + Q(x, y, z)\, dy + R(x, y, z)\, dz
$$

$$
= \int_a^b \left(P\, x'(t) + Q\, y'(t) + R\left(\frac{\partial z}{\partial x} x'(t) + \frac{\partial z}{\partial y} y'(t) \right) \right) dt
$$

$$
= \int_a^b \left(\left(P + R\frac{\partial z}{\partial x} \right) x'(t) + \left(Q + R\frac{\partial z}{\partial y} \right) y'(t) \right) dt
$$

$$
= \int_{C_D} \tilde{P}(x, y)\, dx + \tilde{Q}(x, y)\, dy,
$$

where

$$
\tilde{P}(x, y) = P(x, y, z(x, y)) + R(x, y, z(x, y)) \frac{\partial z}{\partial x}(x, y), \text{ and}
$$

$$
\tilde{Q}(x, y) = Q(x, y, z(x, y)) + R(x, y, z(x, y)) \frac{\partial z}{\partial y}(x, y)
$$

for (x, y) in D. Thus, by Green's Theorem applied to the region D, we have

$$
\oint_C \vec{f} \cdot d\vec{r} = \iint_D \left(\frac{\partial \tilde{Q}}{\partial x} - \frac{\partial \tilde{P}}{\partial y} \right) dA. \tag{15.17}
$$

Thus,

$$
\frac{\partial \tilde{Q}}{\partial x} = \frac{\partial}{\partial x} \left(Q(x, y, z(x, y)) + R(x, y, z(x, y)) \frac{\partial z}{\partial y}(x, y) \right), \text{ so by the Product Rule we get}
$$

$$
= \frac{\partial}{\partial x} \left(Q(x, y, z(x, y)) \right) + \left(\frac{\partial}{\partial x} R(x, y, z(x, y)) \right) \frac{\partial z}{\partial y}(x, y) + R(x, y, z(x, y)) \frac{\partial}{\partial x} \left(\frac{\partial z}{\partial y}(x, y) \right).
$$

Now, by Theorem 111, we have

$$
\frac{\partial}{\partial x} \left(Q(x, y, z(x, y)) \right) = \frac{\partial Q}{\partial x} \frac{\partial x}{\partial x} + \frac{\partial Q}{\partial y} \frac{\partial y}{\partial x} + \frac{\partial Q}{\partial z} \frac{\partial z}{\partial x}
$$

$$
= \frac{\partial Q}{\partial x} \cdot 1 + \frac{\partial Q}{\partial y} \cdot 0 + \frac{\partial Q}{\partial z} \frac{\partial z}{\partial x}
$$

$$
= \frac{\partial Q}{\partial x} + \frac{\partial Q}{\partial z} \frac{\partial z}{\partial x}.
$$

Similarly,

$$
\frac{\partial}{\partial x} \left(R(x, y, z(x, y)) \right) = \frac{\partial R}{\partial x} + \frac{\partial R}{\partial z} \frac{\partial z}{\partial x}.
$$

Notes:

Thus,

$$\frac{\partial \tilde{Q}}{\partial x} = \frac{\partial Q}{\partial x} + \frac{\partial Q}{\partial z}\frac{\partial z}{\partial x} + \left(\frac{\partial R}{\partial x} + \frac{\partial R}{\partial z}\frac{\partial z}{\partial x}\right)\frac{\partial z}{\partial y} + R(x,y,z(x,y))\frac{\partial^2 z}{\partial x\,\partial y}$$

$$= \frac{\partial Q}{\partial x} + \frac{\partial Q}{\partial z}\frac{\partial z}{\partial x} + \frac{\partial R}{\partial x}\frac{\partial z}{\partial y} + \frac{\partial R}{\partial z}\frac{\partial z}{\partial x}\frac{\partial z}{\partial y} + R\frac{\partial^2 z}{\partial x\,\partial y}.$$

In a similar fashion, we can calculate

$$\frac{\partial \tilde{P}}{\partial y} = \frac{\partial P}{\partial y} + \frac{\partial P}{\partial z}\frac{\partial z}{\partial y} + \frac{\partial R}{\partial y}\frac{\partial z}{\partial x} + \frac{\partial R}{\partial z}\frac{\partial z}{\partial y}\frac{\partial z}{\partial x} + R\frac{\partial^2 z}{\partial y\,\partial x}.$$

So subtracting gives

$$\frac{\partial \tilde{Q}}{\partial x} - \frac{\partial \tilde{P}}{\partial y} = \left(\frac{\partial Q}{\partial z} - \frac{\partial R}{\partial y}\right)\frac{\partial z}{\partial x} + \left(\frac{\partial R}{\partial x} - \frac{\partial P}{\partial z}\right)\frac{\partial z}{\partial y} + \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right) \quad (15.18)$$

since $\frac{\partial^2 z}{\partial x\,\partial y} = \frac{\partial^2 z}{\partial y\,\partial x}$ by the smoothness of $z = z(x,y)$. Hence, by Equation (15.17),

$$\oint_C \vec{f}\cdot d\vec{r} = \iint_D \left(-\left(\frac{\partial R}{\partial y} - \frac{\partial Q}{\partial z}\right)\frac{\partial z}{\partial x} - \left(\frac{\partial P}{\partial z} - \frac{\partial R}{\partial x}\right)\frac{\partial z}{\partial y} + \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right)\right)\,dA$$
$$(15.19)$$

after factoring out a -1 from the terms in the first two products in Equation (15.18).

Now, recall from Section 13.7 that the vector $\vec{n} = -\frac{\partial z}{\partial x}\,\vec{\imath} - \frac{\partial z}{\partial y}\,\vec{\jmath} + \vec{k}$ is normal to the tangent plane to the surface $z = z(x,y)$ at each point of Σ. Thus,

$$\vec{n} = \frac{\vec{n}}{\|\vec{n}\|} = \frac{-\frac{\partial z}{\partial x}\,\vec{\imath} - \frac{\partial z}{\partial y}\,\vec{\jmath} + \vec{k}}{\sqrt{1 + \left(\frac{\partial z}{\partial x}\right)^2 + \left(\frac{\partial z}{\partial y}\right)^2}}$$

is in fact a positive unit normal vector to Σ (see Figure 15.20). Hence, using the parametrization $\vec{r}(x,y) = x\,\vec{\imath} + y\,\vec{\jmath} + z(x,y)\,\vec{k}$, for (x,y) in D, of the surface Σ, we have $\frac{\partial \vec{r}}{\partial x} = \vec{\imath} + \frac{\partial z}{\partial x}\,\vec{k}$ and $\frac{\partial \vec{r}}{\partial y} = \vec{\jmath} + \frac{\partial z}{\partial y}\,\vec{k}$, and so $\left\|\frac{\partial \vec{r}}{\partial x} \times \frac{\partial \vec{r}}{\partial y}\right\| = \sqrt{1 + \left(\frac{\partial z}{\partial x}\right)^2 + \left(\frac{\partial z}{\partial y}\right)^2}$. So we see that using Equation (15.16) for $\operatorname{curl}\vec{f}$, we have

$$\iint_\Sigma (\operatorname{curl}\vec{f})\cdot\vec{n}\,d\sigma$$

$$= \iint_D (\operatorname{curl}\vec{f})\cdot\vec{n}\left\|\frac{\partial \vec{r}}{\partial x} \times \frac{\partial \vec{r}}{\partial y}\right\|\,dA$$

$$= \iint_D \left(\left(\frac{\partial R}{\partial y} - \frac{\partial Q}{\partial z}\right)\vec{\imath} + \left(\frac{\partial P}{\partial z} - \frac{\partial R}{\partial x}\right)\vec{\jmath} + \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right)\vec{k}\right)\cdot\left(-\frac{\partial z}{\partial x}\vec{\imath} - \frac{\partial z}{\partial y}\vec{\jmath} + \vec{k}\right)\,dA$$

$$= \iint_D \left(-\left(\frac{\partial R}{\partial y} - \frac{\partial Q}{\partial z}\right)\frac{\partial z}{\partial x} - \left(\frac{\partial P}{\partial z} - \frac{\partial R}{\partial x}\right)\frac{\partial z}{\partial y} + \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right)\right)\,dA,$$

Notes:

which, upon comparing to Equation (15.19), proves the Theorem.   □

Note: The condition in Stokes' Theorem that the surface Σ have a (continuously varying) positive unit normal vector $\vec{n}$ and a boundary curve C traversed $\vec{n}$-positively can be expressed more precisely as follows: if $\vec{r}(t)$ is the position vector for C and $\vec{T}(t) = \vec{r}'(t)/\|\vec{r}'(t)\|$ is the unit tangent vector to C, then the vectors $\vec{T}, \vec{n}, \vec{T} \times \vec{n}$ form a right-handed system.

Also, it should be noted that Stokes' Theorem holds even when the boundary curve C is piecewise smooth.

Example 3   Verifying Stokes' Theorem

Verify Stokes' Theorem for $\vec{f}(x, y, z) = z\,\vec{\imath} + x\,\vec{\jmath} + y\,\vec{k}$ when Σ is the paraboloid $z = x^2 + y^2$ such that $z \le 1$ (see Figure 15.21).

   SOLUTION   The positive unit normal vector to the surface $z = z(x, y) = x^2 + y^2$ is

$$\vec{n} = \frac{-\frac{\partial z}{\partial x}\,\vec{\imath} - \frac{\partial z}{\partial y}\,\vec{\jmath} + \vec{k}}{\sqrt{1 + \left(\frac{\partial z}{\partial x}\right)^2 + \left(\frac{\partial z}{\partial y}\right)^2}} = \frac{-2x\,\vec{\imath} - 2y\,\vec{\jmath} + \vec{k}}{\sqrt{1 + 4x^2 + 4y^2}},$$

and $\operatorname{curl}\vec{f} = (1 - 0)\,\vec{\imath} + (1 - 0)\,\vec{\jmath} + (1 - 0)\,\vec{k} = \vec{\imath} + \vec{\jmath} + \vec{k}$, so

$$(\operatorname{curl}\vec{f}) \cdot \vec{n} = (-2x - 2y + 1)/\sqrt{1 + 4x^2 + 4y^2}.$$

Since Σ can be parametrized as $\vec{r}(x, y) = x\,\vec{\imath} + y\,\vec{\jmath} + (x^2 + y^2)\,\vec{k}$ for (x, y) in the region $D = \{\,(x, y) : x^2 + y^2 \le 1\,\}$, then

$$\iint_{\Sigma} (\operatorname{curl}\vec{f}) \cdot \vec{n}\, d\sigma = \iint_{D} (\operatorname{curl}\vec{f}) \cdot \vec{n} \left\| \frac{\partial \vec{r}}{\partial x} \times \frac{\partial \vec{r}}{\partial y} \right\| dA$$

$$= \iint_{D} \frac{-2x - 2y + 1}{\sqrt{1 + 4x^2 + 4y^2}} \sqrt{1 + 4x^2 + 4y^2}\, dA$$

$$= \iint_{D} (-2x - 2y + 1)\, dA, \text{ so switching to polar coordinates gives}$$

$$= \int_{0}^{2\pi} \int_{0}^{1} (-2r\cos\theta - 2r\sin\theta + 1)r\, dr\, d\theta$$

$$= \int_{0}^{2\pi} \int_{0}^{1} (-2r^2\cos\theta - 2r^2\sin\theta + r)\, dr\, d\theta$$

$$= \int_{0}^{2\pi} \left(-\tfrac{2r^3}{3}\cos\theta - \tfrac{2r^3}{3}\sin\theta + \tfrac{r^2}{2} \Big|_{r=0}^{r=1} \right) d\theta$$

$$= \int_{0}^{2\pi} \left(-\tfrac{2}{3}\cos\theta - \tfrac{2}{3}\sin\theta + \tfrac{1}{2} \right) d\theta$$

$$= -\tfrac{2}{3}\sin\theta + \tfrac{2}{3}\cos\theta + \tfrac{1}{2}\theta \Big|_{0}^{2\pi} = \pi.$$
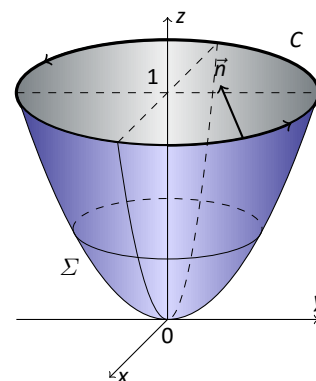
Figure 15.21: $z = x^2 + y^2$ for Example 3.

Notes:

The boundary curve C is the unit circle $x^2 + y^2 = 1$ laying in the plane $z = 1$ (see Figure 15.21), which can be parametrized as $x = \cos t$, $y = \sin t$, $z = 1$ for $0 \le t \le 2\pi$. So

$$\oint_C \vec{f} \cdot d\vec{r} = \int_0^{2\pi} ((1)(-\sin t) + (\cos t)(\cos t) + (\sin t)(0))\, dt$$

$$= \int_0^{2\pi} \left(-\sin t + \frac{1 + \cos 2t}{2}\right) dt \quad \left(\text{here we used } \cos^2 t = \frac{1 + \cos 2t}{2}\right)$$

$$= \cos t + \frac{t}{2} + \frac{\sin 2t}{4} \bigg|_0^{2\pi} = \pi.$$

So we see that $\oint_C \vec{f} \cdot d\vec{r} = \iint_\Sigma (\text{curl}\,\vec{f}) \cdot \vec{n}\, d\sigma$, as predicted by Stokes' Theorem.

The line integral in the preceding example was far simpler to calculate than the surface integral, but this will not always be the case.

Example 4   **Using Stokes' Theorem**

Let Σ be the elliptic paraboloid $z = \frac{x^2}{4} + \frac{y^2}{9}$ for $z \le 1$, and let C be its boundary curve. Calculate $\oint_C \vec{f} \cdot d\vec{r}$ for $\vec{f}(x, y, z) = (9xz + 2y)\vec{\imath} + (2x + y^2)\vec{\jmath} + (-2y^2 + 2z)\vec{k}$, where C is traversed counterclockwise.

SOLUTION   The surface is similar to the one in Example 3, except now the boundary curve C is the ellipse $\frac{x^2}{4} + \frac{y^2}{9} = 1$ laying in the plane $z = 1$. In this case, using Stokes' Theorem is easier than computing the line integral directly. As in Example 3, at each point $(x, y, z(x, y))$ on the surface $z = z(x, y) = \frac{x^2}{4} + \frac{y^2}{9}$ the vector

$$\vec{n} = \frac{-\frac{\partial z}{\partial x}\,\vec{\imath} - \frac{\partial z}{\partial y}\,\vec{\jmath} + \vec{k}}{\sqrt{1 + \left(\frac{\partial z}{\partial x}\right)^2 + \left(\frac{\partial z}{\partial y}\right)^2}} = \frac{-\frac{x}{2}\,\vec{\imath} - \frac{2y}{9}\,\vec{\jmath} + \vec{k}}{\sqrt{1 + \frac{x^2}{4} + \frac{4y^2}{9}}},$$

is a positive unit normal vector to Σ. And calculating the curl of $\vec{f}$ gives

$$\text{curl}\,\vec{f} = (-4y - 0)\vec{\imath} + (9x - 0)\vec{\jmath} + (2 - 2)\vec{k} = -4y\,\vec{\imath} + 9x\,\vec{\jmath} + 0\,\vec{k},$$

so

$$(\text{curl}\,\vec{f}) \cdot \vec{n} = \frac{(-4y)(-\frac{x}{2}) + (9x)(-\frac{2y}{9}) + (0)(1)}{\sqrt{1 + \frac{x^2}{4} + \frac{4y^2}{9}}} = \frac{2xy - 2xy + 0}{\sqrt{1 + \frac{x^2}{4} + \frac{4y^2}{9}}} = 0,$$

and so by Stokes' Theorem

$$\oint_C \vec{f} \cdot d\vec{r} = \iint_\Sigma (\text{curl}\,\vec{f}) \cdot \vec{n}\, d\sigma = \iint_\Sigma 0\, d\sigma = 0.$$

Notes:

In physical applications, for a simple closed curve C the line integral $\oint_C \vec{f} \cdot d\vec{r}$ is often called the **circulation** of $\vec{f}$ around C. For example, if $\vec{E}$ represents the electrostatic field due to a point charge, then it turns out that curl $\vec{E} = \vec{0}$, which means that the circulation $\oint_C \vec{E} \cdot d\vec{r} = 0$ by Stokes' Theorem. Vector fields which have zero curl are often called *irrotational* fields.

In fact, the term curl was created by the 19th century Scottish physicist James Clerk Maxwell in his study of electromagnetism, where it is used extensively. In physics, the curl is interpreted as a measure of *circulation density*. This is best seen by using another definition of curl $\vec{f}$ which is equivalent to the definition given by Equation (15.16). Namely, for a point (x, y, z) in $\mathbb{R}^3$,

$$\vec{n} \cdot (\operatorname{curl}\vec{f})(x, y, z) = \lim_{S \to 0} \frac{1}{S} \oint_C \vec{f} \cdot d\vec{r},$$

where S is the surface area of a surface Σ containing the point (x, y, z) and with a simple closed boundary curve C and positive unit normal vector $\vec{n}$ at (x, y, z). In the limit, think of the curve C shrinking to the point (x, y, z), which causes Σ, the surface it bounds, to have smaller and smaller surface area. That ratio of circulation to surface area in the limit is what makes the curl a rough measure of circulation density (i.e., circulation per unit area).
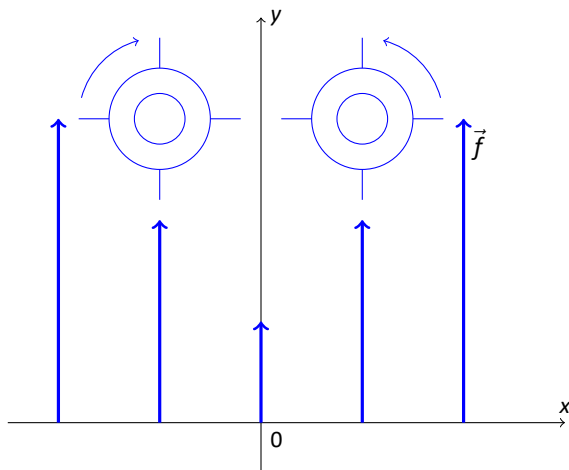
Figure 15.22: Curl and rotation

An idea of how the curl of a vector field is related to rotation is shown in Figure 15.22. Suppose we have a vector field $\vec{f}(x, y, z)$ which is always parallel to the *xy*-plane at each point (x, y, z) and that the vectors grow larger the further the point (x, y, z) is from the *y*-axis. For example, $\vec{f}(x, y, z) = (1 + x^2)\,\vec{j}$. Think of the vector field as representing the flow of water, and imagine dropping two wheels with paddles into that water flow, as in Figure 15.22. Since the

Notes:

flow is stronger (i.e., the magnitude of $\vec{f}$ is larger) as you move away from the y-axis, then such a wheel would rotate counterclockwise if it were dropped to the right of the y-axis, and it would rotate clockwise if it were dropped to the left of the y-axis. In both cases the curl would be nonzero (curl $\vec{f}(x, y, z) = 2x \vec{k}$ in our example) and would obey the right-hand rule, that is, curl $\vec{f}(x, y, z)$ points in the direction of your thumb as you cup your right hand in the direction of the rotation of the wheel. So the curl points outward (in the positive z-direction) if $x > 0$ and points inward (in the negative z-direction) if $x < 0$. Notice that if all the vectors had the same direction *and* the same magnitude, then the wheels would not rotate and hence there would be no curl (which is why such fields are called *irrotational*, meaning no rotation).

Finally, by Stokes' Theorem, we know that if C is a simple closed curve in some solid region S in $\mathbb{R}^3$ and if $\vec{f}(x, y, z)$ is a smooth vector field such that curl $\vec{f} = \vec{0}$ in S, then

$$\oint_C \vec{f} \cdot d\vec{r} = \iint_\Sigma (\text{curl}\,\vec{f}) \cdot \vec{n}\, d\sigma = \iint_\Sigma \vec{0} \cdot \vec{n}\, d\sigma = \iint_\Sigma 0 \, d\sigma = 0,$$

where Σ is any orientable surface inside S whose boundary is C (such a surface is sometimes called a *capping surface* for C). So similar to the two-variable case, we have a three-dimensional version of a result from Section 15.3, for solid regions in $\mathbb{R}^3$ which are **simply connected** (i.e. regions having no holes):

Theorem 145   Irrotational Equivalences

The following statements are equivalent for a simply connected solid region S in $\mathbb{R}^3$:

1. $\vec{f}(x, y, z) = P(x, y, z)\,\vec{i} + Q(x, y, z)\,\vec{j} + R(x, y, z)\,\vec{k}$ has a smooth potential $F(x, y, z)$ in S

2. $\displaystyle\int_C \vec{f} \cdot d\vec{r}$ is independent of the path for any curve C in S

3. $\displaystyle\oint_C \vec{f} \cdot d\vec{r} = 0$ for every simple closed curve C in S

4. $\dfrac{\partial R}{\partial y} = \dfrac{\partial Q}{\partial z}, \dfrac{\partial P}{\partial z} = \dfrac{\partial R}{\partial x}$, and $\dfrac{\partial Q}{\partial x} = \dfrac{\partial P}{\partial y}$ in S (i.e., curl $\vec{f} = \vec{0}$ in S, or the differential form $P\,dx + Q\,dy + R\,dz$ is exact)

Example 5   Determining Irrotation

Determine if the vector field $\vec{f}(x, y, z) = xyz\,\vec{i} + xz\,\vec{j} + xy\,\vec{k}$ has a potential in $\mathbb{R}^3$.

Notes:

SOLUTION   Since $\mathbb{R}^3$ is simply connected, we just need to check whether $\text{curl}\,\vec{f} = \vec{0}$ throughout $\mathbb{R}^3$, that is,

$$\frac{\partial R}{\partial y} = \frac{\partial Q}{\partial z}, \quad \frac{\partial P}{\partial z} = \frac{\partial R}{\partial x}, \quad \text{and} \quad \frac{\partial Q}{\partial x} = \frac{\partial P}{\partial y}$$

throughout $\mathbb{R}^3$, where $P(x, y, z) = xyz$, $Q(x, y, z) = xz$, and $R(x, y, z) = xy$. But we see that

$$\frac{\partial P}{\partial z} = xy, \frac{\partial R}{\partial x} = y \quad \Rightarrow \quad \frac{\partial P}{\partial z} \neq \frac{\partial R}{\partial x} \text{ for some } (x, y, z) \text{ in } \mathbb{R}^3.$$

Thus, $\vec{f}(x, y, z)$ does not have a potential in $\mathbb{R}^3$.

The theorems in this chapter all relate an integral over a domain to an integral over the boundary of the domain. This means that we need to pay special attention to the orientation of the domain. We'll occasionally use the symbol ∂ to indicate the boundary of something (so the boundary of D would be ∂D). This is the same symbol as a partial derivative, so you'll have to look at the context to figure out which definition is being used.

We also have four new theorems about these types of integrals. These are similar to the Fundamental Theorem of Calculus, so we'll put all five together into Key Idea 65.

Key Idea 65   Fundamental Theorems Relating Integrals and Domains

Theorem	Equation	Orientation
Fundamental Theorem of Calculus	$f(b) - f(a) = \int_a^b f'(x)dx$	
Fundamental Theorem of Gradient Fields	$\phi(Q) - \phi(P) = \int_C \nabla\phi \cdot ds$	C goes from P to Q
Green's Theorem	$\oint_{\partial D} Pdx + Qdy = \iint_D Q_x - P_y dA$	∂D oriented counterclockwise
Stokes' Theorem	$\int_{\partial S} \vec{F} \cdot d\vec{r} = \iint_S \text{curl}\,\vec{F} \cdot \vec{n}d\sigma$	∂S oriented with S to the left
Divergence Theorem	$\iint_{\partial W} \vec{F} \cdot d\vec{\sigma} = \iiint_W \text{div}\,\vec{F}dV$	∂W oriented outwards

---

Notes:

Exercises 15.5

Problems

In Exercises 1–3, calculate $\int_C f(x, y, z)\, ds$ for the given function $f(x, y, z)$ and curve C.

1. $f(x, y, z) = z;\quad C : x = \cos t,\, y = \sin t,\, z = t,\, 0 \le t \le 2\pi$

2. $f(x, y, z) = \dfrac{x}{y} + y + 2yz;\quad C : x = t^2,\, y = t,\, z = 1,$
   $1 \le t \le 2$

3. $f(x, y, z) = z^2;\quad C : x = t \sin t,\, y = t \cos t,\, z = \frac{2\sqrt{2}}{3} t^{3/2},$
   $0 \le t \le 1$

In Exercises 4–9, calculate $\int_C \vec{f} \cdot d\vec{r}$ for the given vector field $\vec{f}(x, y, z)$ and curve C.

4. $\vec{f}(x, y, z) = \vec{\imath} - \vec{\jmath} + \vec{k};\quad C : x = 3t,\, y = 2t,\, z = t,\, 0 \le t \le 1$

5. $\vec{f}(x, y, z) = y\vec{\imath} - x\vec{\jmath} + z\vec{k};\quad C : x = \cos t,\, y = \sin t,\, z = t,$
   $0 \le t \le 2\pi$

6. $\vec{f}(x, y, z) = x\vec{\imath} + y\vec{\jmath} + z\vec{k};\quad C : x = \cos t,\, y = \sin t,\, z = 2,$
   $0 \le t \le 2\pi$

7. $\vec{f}(x, y, z) = (y - 2z)\vec{\imath} + xy\vec{\jmath} + (2xz + y)\vec{k};\quad C : x = t,$
   $y = 2t,\, z = t^2 - 1,\, 0 \le t \le 1$

8. $\vec{f}(x, y, z) = yz\vec{\imath} + xz\vec{\jmath} + xy\vec{k};\quad C$: the polygonal path from $(0, 0, 0)$ to $(1, 0, 0)$ to $(1, 2, 0)$

9. $\vec{f}(x, y, z) = xy\vec{\imath} + (z - x)\vec{\jmath} + 2yz\vec{k};\quad C$: the polygonal path from $(0, 0, 0)$ to $(1, 0, 0)$ to $(1, 2, 0)$ to $(1, 2, -2)$

In Exercises 10–13, state whether or not the vector field $\vec{f}(x, y, z)$ has a potential in $\mathbb{R}^3$ (you do not need to find the potential itself).

10. $\vec{f}(x, y, z) = y\vec{\imath} - x\vec{\jmath} + z\vec{k}$

11. $\vec{f}(x, y, z) = a\vec{\imath} + b\vec{\jmath} + c\vec{k}$ (a, b, c constant)

12. $\vec{f}(x, y, z) = (x + y)\vec{\imath} + x\vec{\jmath} + z^2\vec{k}$

13. $\vec{f}(x, y, z) = xy\vec{\imath} - (x - yz^2)\vec{\jmath} + y^2z\vec{k}$

In Exercises 14–15, verify Stokes' Theorem for the given vector field $\vec{f}(x, y, z)$ and surface Σ.

14. $\vec{f}(x, y, z) = 2y\vec{\imath} - x\vec{\jmath} + z\vec{k};\quad \Sigma : x^2 + y^2 + z^2 = 1,\, z \ge 0$

15. $\vec{f}(x, y, z) = xy\vec{\imath} + xz\vec{\jmath} + yz\vec{k};\quad \Sigma : z = x^2 + y^2,\, z \le 1$

16. Construct a Möbius strip from a piece of paper, then draw a line down its center (like the dotted line in Figure 15.19(b)). Cut the Möbius strip along that center line completely around the strip. How many surfaces does this result in? How would you describe them? Are they orientable?

17. Use a computer algebra system to plot the Möbius strip parametrized as:
    $$\vec{r}(u, v) = \left\langle \cos u \left(1 + v \cos \tfrac{u}{2}\right), \sin u \left(1 + v \cos \tfrac{u}{2}\right), v \sin \tfrac{u}{2} \right\rangle,$$
    where $0 \le u \le 2\pi,\, -\tfrac{1}{2} \le v \le \tfrac{1}{2}$

18. Let Σ be a closed surface and $\vec{f}(x, y, z)$ a smooth vector field. Show that $\displaystyle\iint_\Sigma (\text{curl}\,\vec{f}) \cdot \vec{n}\, d\sigma = 0$. (*Hint: Split Σ in half.*)

19. Show that Green's Theorem is a special case of Stokes' Theorem.

15.6   Gradient, Divergence, Curl and Laplacian

In this final section we will establish some relationships between the gradient, divergence and curl, and we will also introduce a new quantity called the *Laplacian*. We will then show how to write these quantities in cylindrical and spherical coordinates.

For a real-valued function $f(x, y, z)$ on $\mathbb{R}^3$, the gradient $\nabla f(x, y, z)$ is a vector-valued function on $\mathbb{R}^3$, that is, its value at a point (x, y, z) is the vector

$$\nabla f(x, y, z) = \left\langle \frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \frac{\partial f}{\partial z} \right\rangle = \frac{\partial f}{\partial x}\,\vec{\imath} + \frac{\partial f}{\partial y}\,\vec{\jmath} + \frac{\partial f}{\partial z}\,\vec{k}$$

in $\mathbb{R}^3$, where each of the partial derivatives is evaluated at the point (x, y, z). So in this way, you can think of the *symbol* ∇ as being "applied" to a real-valued function f to produce a vector ∇f.

In Section 15.4 and Section 15.5, we noted that the divergence and curl can also be expressed in terms of the symbol ∇ by thinking of ∇ as a *vector* in $\mathbb{R}^3$, namely

$$\nabla = \frac{\partial}{\partial x}\,\vec{\imath} + \frac{\partial}{\partial y}\,\vec{\jmath} + \frac{\partial}{\partial z}\,\vec{k}. \tag{15.20}$$

Here, the symbols $\frac{\partial}{\partial x}$, $\frac{\partial}{\partial y}$ and $\frac{\partial}{\partial z}$ are to be thought of as "partial derivative operators" that will get "applied" to a real-valued function, say $f(x, y, z)$, to produce the partial derivatives $\frac{\partial f}{\partial x}$, $\frac{\partial f}{\partial y}$ and $\frac{\partial f}{\partial z}$. For instance, $\frac{\partial}{\partial x}$ "applied" to $f(x, y, z)$ produces $\frac{\partial f}{\partial x}$.

Is ∇ *really* a vector? Strictly speaking, no, since $\frac{\partial}{\partial x}$, $\frac{\partial}{\partial y}$ and $\frac{\partial}{\partial z}$ are not actual numbers. But it helps to *think* of ∇ as a vector, especially with the divergence and curl, as we will soon see. The process of "applying" $\frac{\partial}{\partial x}$, $\frac{\partial}{\partial y}$, $\frac{\partial}{\partial z}$ to a real-valued function $f(x, y, z)$ is normally thought of as *multiplying* the quantities:

$$\left(\frac{\partial}{\partial x} \right) (f) = \frac{\partial f}{\partial x}, \quad \left(\frac{\partial}{\partial y} \right) (f) = \frac{\partial f}{\partial y}, \quad \left(\frac{\partial}{\partial z} \right) (f) = \frac{\partial f}{\partial z}$$

For this reason, ∇ is often referred to as the "del operator", since it "operates" on functions.

For example, it is often convenient to write the divergence $\operatorname{div} \vec{f}$ as $\nabla \cdot \vec{f}$, since for a vector field $\vec{f}(x, y, z) = f_1(x, y, z)\vec{\imath} + f_2(x, y, z)\vec{\jmath} + f_3(x, y, z)\vec{k}$, the dot product

---

Notes:

of $\vec{f}$ with ∇ (thought of as a vector) makes sense:

$$\nabla \cdot \vec{f} = \left(\frac{\partial}{\partial x} \, \vec{\imath} + \frac{\partial}{\partial y} \, \vec{\jmath} + \frac{\partial}{\partial z} \, \vec{k} \right) \cdot \left(f_1(x,y,z)\vec{\imath} + f_2(x,y,z)\vec{\jmath} + f_3(x,y,z)\vec{k} \right)$$

$$= \left(\frac{\partial}{\partial x} \right) (f_1) + \left(\frac{\partial}{\partial y} \right) (f_2) + \left(\frac{\partial}{\partial z} \right) (f_3)$$

$$= \frac{\partial f_1}{\partial x} + \frac{\partial f_2}{\partial y} + \frac{\partial f_3}{\partial z}$$

$$= \operatorname{div} \vec{f} \tag{15.21}$$

We can also write $\operatorname{curl} \vec{f}$ in terms of ∇, namely as $\nabla \times \vec{f}$, since for a vector field $\vec{f}(x,y,z) = P(x,y,z)\vec{\imath} + Q(x,y,z)\vec{\jmath} + R(x,y,z)\vec{k}$, we have:

$$\nabla \times \vec{f} = \begin{vmatrix} \vec{\imath} & \vec{\jmath} & \vec{k} \\ \dfrac{\partial}{\partial x} & \dfrac{\partial}{\partial y} & \dfrac{\partial}{\partial z} \\ P(x,y,z) & Q(x,y,z) & R(x,y,z) \end{vmatrix}$$

$$= \left(\frac{\partial R}{\partial y} - \frac{\partial Q}{\partial z} \right) \vec{\imath} - \left(\frac{\partial R}{\partial x} - \frac{\partial P}{\partial z} \right) \vec{\jmath} + \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} \right) \vec{k}$$

$$= \left(\frac{\partial R}{\partial y} - \frac{\partial Q}{\partial z} \right) \vec{\imath} + \left(\frac{\partial P}{\partial z} - \frac{\partial R}{\partial x} \right) \vec{\jmath} + \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} \right) \vec{k}$$

$$= \operatorname{curl} \vec{f} \tag{15.22}$$

For a real-valued function $f(x,y,z)$, the gradient $\nabla f(x,y,z) = \frac{\partial f}{\partial x} \, \vec{\imath} + \frac{\partial f}{\partial y} \, \vec{\jmath} + \frac{\partial f}{\partial z} \, \vec{k}$ is a vector field, so we can take its divergence:

$$\operatorname{div} \nabla f = \nabla \cdot \nabla f$$

$$= \left(\frac{\partial}{\partial x} \, \vec{\imath} + \frac{\partial}{\partial y} \, \vec{\jmath} + \frac{\partial}{\partial z} \, \vec{k} \right) \cdot \left(\frac{\partial f}{\partial x} \, \vec{\imath} + \frac{\partial f}{\partial y} \, \vec{\jmath} + \frac{\partial f}{\partial z} \, \vec{k} \right)$$

$$= \frac{\partial}{\partial x} \left(\frac{\partial f}{\partial x} \right) + \frac{\partial}{\partial y} \left(\frac{\partial f}{\partial y} \right) + \frac{\partial}{\partial z} \left(\frac{\partial f}{\partial z} \right)$$

$$= \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} + \frac{\partial^2 f}{\partial z^2}$$

Note that this is a real-valued function, to which we will give a special name:

Notes:

> **Definition 118   Laplacian**
> For a real-valued function $f(x, y, z)$, the **Laplacian** of f, denoted by Δf, is given by
> $$\Delta f(x, y, z) = \nabla \cdot \nabla f = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} + \frac{\partial^2 f}{\partial z^2}.$$

Often the notation $\nabla^2 f$ is used for the Laplacian instead of Δf, using the convention $\nabla^2 = \nabla \cdot \nabla$.

Watch the video:
Laplacian intuition at
`https://youtu.be/EW08rD-GFh0`

Example 1   Computing a Laplacian
Let $\vec{r}(x, y, z) = x\,\vec{\imath} + y\,\vec{\jmath} + z\,\vec{k}$ be the position vector field on $\mathbb{R}^3$. Then $\|\vec{r}(x, y, z)\|^2 = \vec{r} \cdot \vec{r} = x^2 + y^2 + z^2$ is a real-valued function. Find

1. the gradient of $\|\vec{r}\|^2$

2. the divergence of $\vec{r}$

3. the curl of $\vec{r}$

4. the Laplacian of $\|\vec{r}\|^2$

SOLUTION

1. $\nabla \|\vec{r}\|^2 = 2x\,\vec{\imath} + 2y\,\vec{\jmath} + 2z\,\vec{k} = 2\,\vec{r}$

2. $\nabla \cdot \vec{r} = \frac{\partial}{\partial x}(x) + \frac{\partial}{\partial y}(y) + \frac{\partial}{\partial z}(z) = 1 + 1 + 1 = 3$

3.
$$\nabla \times \vec{r} = \begin{vmatrix} \vec{\imath} & \vec{\jmath} & \vec{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ x & y & z \end{vmatrix} = (0 - 0)\,\vec{\imath} - (0 - 0)\,\vec{\jmath} + (0 - 0)\,\vec{k} = \vec{0}$$

Notes:

4. $\Delta \left\| \vec{r} \right\|^2 = \frac{\partial^2}{\partial x^2}(x^2 + y^2 + z^2) + \frac{\partial^2}{\partial y^2}(x^2 + y^2 + z^2) + \frac{\partial^2}{\partial z^2}(x^2 + y^2 + z^2) = 2 + 2 + 2 = 6$.

Note that we could have calculated $\Delta \left\| \vec{r} \right\|^2$ another way, using the ∇ notation along with the first two parts:

$$\Delta \left\| \vec{r} \right\|^2 = \nabla \cdot \nabla \left\| \vec{r} \right\|^2 = \nabla \cdot 2\vec{r} = 2\,\nabla \cdot \vec{r} = 2(3) = 6$$

Notice that in Example 1 if we take the curl of the gradient of $\left\| \vec{r} \right\|^2$ we get

$$\nabla \times (\nabla \left\| \vec{r} \right\|^2) = \nabla \times 2\vec{r} = 2\,\nabla \times \vec{r} = 2\vec{0} = \vec{0}.$$

The following theorem shows that this will be the case in general:

Theorem 146   Gradient is Irrotational
For any smooth real-valued function $f(x, y, z)$, $\nabla \times (\nabla f) = \vec{0}$.

Proof
We see by the smoothness of f that

$\nabla \times (\nabla f)$

$$= \begin{vmatrix} \vec{\imath} & \vec{\jmath} & \vec{k} \\ \dfrac{\partial}{\partial x} & \dfrac{\partial}{\partial y} & \dfrac{\partial}{\partial z} \\ \dfrac{\partial f}{\partial x} & \dfrac{\partial f}{\partial y} & \dfrac{\partial f}{\partial z} \end{vmatrix}$$

$$= \left(\frac{\partial^2 f}{\partial y\, \partial z} - \frac{\partial^2 f}{\partial z\, \partial y} \right) \vec{\imath} - \left(\frac{\partial^2 f}{\partial x\, \partial z} - \frac{\partial^2 f}{\partial z\, \partial x} \right) \vec{\jmath} + \left(\frac{\partial^2 f}{\partial x\, \partial y} - \frac{\partial^2 f}{\partial y\, \partial x} \right) \vec{k}$$

$$= \vec{0},$$

since the mixed partial derivatives in each component are equal.   □

Another way of stating Theorem 146 is that potentials are irrotational.

Theorem 147   Potentials are Irrotational
If a vector field $\vec{f}(x, y, z)$ has a potential, then $\text{curl}\,\vec{f} = \vec{0}$.

Notes:

Also, notice that in Example 1 if we take the divergence of the curl of $\vec{r}$ we trivially get

$$\nabla \cdot (\nabla \times \vec{r}) = \nabla \cdot \vec{0} = 0.$$

The following theorem shows that this will be the case in general:

Theorem 148   Curl is Divergence Free
For any smooth vector field $\vec{f}(x, y, z)$, $\nabla \cdot (\nabla \times \vec{f}) = 0$.

The proof is straightforward and left as Exercise 24.

Theorem 149   Flux of a Curl Through a Closed Surface
The flux of the curl of a smooth vector field $\vec{f}(x, y, z)$ through any closed surface is zero.

Proof
Let Σ be a closed surface which bounds a solid S. The flux of $\nabla \times \vec{f}$ through Σ is

$$\iint_{\Sigma} (\nabla \times \vec{f}) \cdot d\sigma = \iiint_{S} \nabla \cdot (\nabla \times \vec{f}) dV \quad \text{(by the Divergence Theorem)}$$

$$= \iiint_{S} 0 \, dV \quad \text{(by Theorem 148)}$$

$$= 0. \qquad \qquad \qquad \square$$

There is another method for proving Theorem 146 which can be useful, and is often used in physics. Namely, if the surface integral $\iint_{\Sigma} f(x, y, z) \, d\sigma = 0$ for *all* surfaces Σ in some solid region (usually all of $\mathbb{R}^3$), then we must have $f(x, y, z) = 0$ throughout that region. The proof is not trivial, and physicists do not usually bother to prove it. But the result is true, and can also be applied to double and triple integrals.

For instance, to prove Theorem 146, assume that $f(x, y, z)$ is a smooth real-valued function on $\mathbb{R}^3$. Let C be a simple closed curve in $\mathbb{R}^3$ and let Σ be any capping surface for C (i.e., Σ is orientable and its boundary is C). Since ∇f is a vector field, then

$$\iint_{\Sigma} (\nabla \times (\nabla f)) \cdot \vec{n} \, d\sigma = \oint_{C} \nabla f \cdot d\vec{r} \quad \text{by Stokes' Theorem, so}$$

$$= 0 \quad \text{by Theorem 143.}$$

Notes:

Since the choice of Σ was arbitrary, then we must have $(\nabla \times (\nabla f)) \cdot \vec{n} = 0$ throughout $\mathbb{R}^3$, where $\vec{n}$ is any unit vector. Using $\vec{i}, \vec{j}$ and $\vec{k}$ in place of $\vec{n}$, we see that we must have $\nabla \times (\nabla f) = \vec{0}$ in $\mathbb{R}^3$, which completes the proof.

Example 2   **Maxwell's Equation**

A system of electric charges has a *charge density* $\rho(x, y, z)$ and produces an elec-trostatic field $\vec{E}(x, y, z)$ at points (x, y, z) in space. *Gauss' Law* states that

$$\iint_{\Sigma} \vec{E} \cdot d\sigma = 4\pi \iiint_{S} \rho dV$$

for any closed surface Σ which encloses the charges, with S being the solid re-gion enclosed by Σ. Show that $\nabla \cdot \vec{E} = 4\pi\rho$. This is one of *Maxwell's Equations*.

SOLUTION   By the Divergence Theorem, we have

$$\iiint_{S} \nabla \cdot \vec{E} dV = \iint_{\Sigma} \vec{E} \cdot d\sigma$$
$$= 4\pi \iiint_{S} \rho dV \quad \text{by Gauss' Law}$$

Combining the integrals gives

$$\iiint_{S} (\nabla \cdot \vec{E} - 4\pi\rho) dV = 0, \quad \text{so}$$
$$\nabla \cdot \vec{E} - 4\pi\rho = 0 \quad \text{since } \Sigma \text{ and hence } S \text{ was arbitrary, so}$$
$$\nabla \cdot \vec{E} = 4\pi\rho.$$

Often (especially in physics) it is convenient to use other coordinate systems when dealing with quantities such as the gradient, divergence, curl and Lapla-cian. We will present the formulas for these in cylindrical and spherical coordi-nates.

Recall from Section 11.7 that a point (x, y, z) can be represented in cylindrical coordinates (r, θ, z), where $x = r\cos\theta, y = r\sin\theta, z = z$. At each point (r, θ, z), let $\vec{e}_r, \vec{e}_\theta, \vec{e}_z$ be unit vectors in the direction of increasing r, θ, z, respectively (see Figure 15.23(a)). Then $\vec{e}_r, \vec{e}_\theta, \vec{e}_z$ form an orthonormal set of vectors. Note, by the right-hand rule, that $\vec{e}_z \times \vec{e}_r = \vec{e}_\theta$.

Notes:
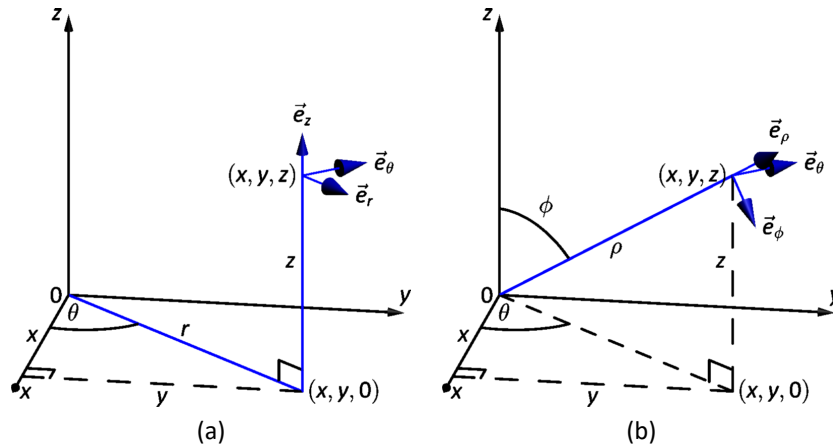
Figure 15.23: Orthonormal vectors in cylindrical (a) and spherical (b) coordinates

Similarly, a point (x, y, z) can be represented in spherical coordinates (ρ, θ, ϕ), where $x = \rho \sin \phi \cos \theta$, $y = \rho \sin \phi \sin \theta$, $z = \rho \cos \phi$. At each point (ρ, θ, ϕ), let $\vec{e}_\rho, \vec{e}_\theta, \vec{e}_\phi$ be unit vectors in the direction of increasing ρ, θ, ϕ, respectively (see Figure 15.23(b)). Then the vectors $\vec{e}_\rho, \vec{e}_\theta, \vec{e}_\phi$ are orthonormal. By the right-hand rule, we see that $\vec{e}_\theta \times \vec{e}_\rho = \vec{e}_\phi$.

We can now summarize the expressions for the gradient, divergence, curl and Laplacian in Cartesian, cylindrical and spherical coordinates in the following tables:

Key Idea 66   Vector Calculus in Cartesian Coordinates
For the scalar function F or vector field $\vec{f} = f_1 \, \vec{\imath} + f_2 \, \vec{\jmath} + f_3 \, \vec{k}$,

$$\text{gradient} : \quad \nabla F = \frac{\partial F}{\partial x} \, \vec{\imath} + \frac{\partial F}{\partial y} \, \vec{\jmath} + \frac{\partial F}{\partial z} \, \vec{k}$$

$$\text{divergence} : \quad \nabla \cdot \vec{f} = \frac{\partial f_1}{\partial x} + \frac{\partial f_2}{\partial y} + \frac{\partial f_3}{\partial z}$$

$$\text{curl} : \nabla \times \vec{f} = \left(\frac{\partial f_3}{\partial y} - \frac{\partial f_2}{\partial z} \right) \vec{\imath} + \left(\frac{\partial f_1}{\partial z} - \frac{\partial f_3}{\partial x} \right) \vec{\jmath} + \left(\frac{\partial f_2}{\partial x} - \frac{\partial f_1}{\partial y} \right) \vec{k}$$

$$\text{Laplacian} : \quad \Delta F = \frac{\partial^2 F}{\partial x^2} + \frac{\partial^2 F}{\partial y^2} + \frac{\partial^2 F}{\partial z^2}$$

Notes:

Key Idea 67   **Vector Calculus in Cylindrical Coordinates**

For the scalar function F or vector field $\vec{f} = f_r\,\vec{e}_r + f_\theta\,\vec{e}_\theta + f_z\,\vec{e}_z$,

gradient : $\nabla F = \dfrac{\partial F}{\partial r}\,\vec{e}_r + \dfrac{1}{r}\dfrac{\partial F}{\partial \theta}\,\vec{e}_\theta + \dfrac{\partial F}{\partial z}\,\vec{e}_z$

divergence : $\nabla \cdot \vec{f} = \dfrac{1}{r}\dfrac{\partial}{\partial r}(rf_r) + \dfrac{1}{r}\dfrac{\partial f_\theta}{\partial \theta} + \dfrac{\partial f_z}{\partial z}$

curl : $\nabla \times \vec{f} = \left(\dfrac{1}{r}\dfrac{\partial f_z}{\partial \theta} - \dfrac{\partial f_\theta}{\partial z}\right)\vec{e}_r + \left(\dfrac{\partial f_r}{\partial z} - \dfrac{\partial f_z}{\partial r}\right)\vec{e}_\theta + \dfrac{1}{r}\left(\dfrac{\partial}{\partial r}(rf_\theta) - \dfrac{\partial f_r}{\partial \theta}\right)\vec{e}_z$

Laplacian : $\Delta F = \dfrac{1}{r}\dfrac{\partial}{\partial r}\left(r\dfrac{\partial F}{\partial r}\right) + \dfrac{1}{r^2}\dfrac{\partial^2 F}{\partial \theta^2} + \dfrac{\partial^2 F}{\partial z^2}$

Key Idea 68   **Vector Calculus in Spherical Coordinates**

For the scalar function F or vector field $\vec{f} = f_\rho\,\vec{e}_\rho + f_\theta\,\vec{e}_\theta + f_\phi\,\vec{e}_\phi$

gradient : $\nabla F = \dfrac{\partial F}{\partial \rho}\,\vec{e}_\rho + \dfrac{1}{\rho \sin\phi}\dfrac{\partial F}{\partial \theta}\,\vec{e}_\theta + \dfrac{1}{\rho}\dfrac{\partial F}{\partial \phi}\,\vec{e}_\phi$

divergence : $\nabla \cdot \vec{f} = \dfrac{1}{\rho^2}\dfrac{\partial}{\partial \rho}(\rho^2 f_\rho) + \dfrac{1}{\rho \sin\phi}\dfrac{\partial f_\theta}{\partial \theta} + \dfrac{1}{\rho \sin\phi}\dfrac{\partial}{\partial \phi}(\sin\phi f_\phi)$

curl : $\nabla \times \vec{f} = \dfrac{1}{\rho \sin\phi}\left(\dfrac{\partial}{\partial \phi}(\sin\phi f_\theta) - \dfrac{\partial f_\phi}{\partial \theta}\right)\vec{e}_\rho + \dfrac{1}{\rho}\left(\dfrac{\partial}{\partial \rho}(\rho f_\phi) - \dfrac{\partial f_\rho}{\partial \phi}\right)\vec{e}_\theta$

$\qquad + \left(\dfrac{1}{\rho \sin\phi}\dfrac{\partial f_\rho}{\partial \theta} - \dfrac{1}{\rho}\dfrac{\partial}{\partial \rho}(\rho f_\theta)\right)\vec{e}_\phi$

Laplacian : $\Delta F = \dfrac{1}{\rho^2}\dfrac{\partial}{\partial \rho}\left(\rho^2\dfrac{\partial F}{\partial \rho}\right) + \dfrac{1}{\rho^2 \sin^2\phi}\dfrac{\partial^2 F}{\partial \theta^2} + \dfrac{1}{\rho^2 \sin\phi}\dfrac{\partial}{\partial \phi}\left(\sin\phi\dfrac{\partial F}{\partial \phi}\right)$

The derivation of the above formulas for cylindrical and spherical coordinates is straightforward but extremely tedious. The basic idea is to take the Cartesian equivalent of the quantity in question and to substitute into that formula using the appropriate coordinate transformation. As an example, we will derive the formula for the gradient in spherical coordinates.

Example 3   **Spherical Coordinates Gradient**

Show that the gradient of a real-valued function $F(\rho, \theta, \phi)$ in spherical coordinates is:

$$\nabla F = \frac{\partial F}{\partial \rho}\,\vec{e}_\rho + \frac{1}{\rho \sin\phi}\frac{\partial F}{\partial \theta}\,\vec{e}_\theta + \frac{1}{\rho}\frac{\partial F}{\partial \phi}\,\vec{e}_\phi$$

Notes:

SOLUTION   The idea is that in the Cartesian gradient formula $\nabla F(x, y, z) = \frac{\partial F}{\partial x}\vec{\imath} + \frac{\partial F}{\partial y}\vec{\jmath} + \frac{\partial F}{\partial z}\vec{k}$, we want to put the Cartesian basis vectors $\vec{\imath}, \vec{\jmath}, \vec{k}$ in terms of the spherical coordinate basis vectors $\vec{e}_\rho, \vec{e}_\theta, \vec{e}_\phi$ and functions of ρ, θ and ϕ. Then put the partial derivatives $\frac{\partial F}{\partial x}, \frac{\partial F}{\partial y}, \frac{\partial F}{\partial z}$ in terms of $\frac{\partial F}{\partial \rho}, \frac{\partial F}{\partial \theta}, \frac{\partial F}{\partial \phi}$ and functions of ρ, θ and ϕ.

Our first step is to get formulas for $\vec{e}_\rho, \vec{e}_\theta, \vec{e}_\phi$ in terms of $\vec{\imath}, \vec{\jmath}, \vec{k}$.

We can see from Figure 15.23(b) that the unit vector $\vec{e}_\rho$ in the ρ direction at a general point (ρ, θ, ϕ) is $\vec{e}_\rho = \frac{\vec{r}}{\|\vec{r}\|}$, where $\vec{r} = x\vec{\imath} + y\vec{\jmath} + z\vec{k}$ is the position vector of the point in Cartesian coordinates. Thus,

$$\vec{e}_\rho = \frac{\vec{r}}{\|\vec{r}\|} = \frac{x\vec{\imath} + y\vec{\jmath} + z\vec{k}}{\sqrt{x^2 + y^2 + z^2}},$$

so using $x = \rho \sin\phi \cos\theta, y = \rho \sin\phi \sin\theta, z = \rho \cos\phi$, and $\rho = \sqrt{x^2 + y^2 + z^2}$, we get:

$$\vec{e}_\rho = \sin\phi\,\cos\theta\,\vec{\imath} + \sin\phi\,\sin\theta\,\vec{\jmath} + \cos\phi\,\vec{k} \qquad (15.23)$$

Now, since the angle θ is measured in the *xy*-plane, then the unit vector $\vec{e}_\theta$ in the θ direction must be parallel to the *xy*-plane. That is, $\vec{e}_\theta$ is of the form $a\vec{\imath} + b\vec{\jmath} + 0\vec{k}$. To figure out what a and b are, note that since $\vec{e}_\theta \perp \vec{e}_\rho$, then in particular $\vec{e}_\theta \perp \vec{e}_\rho$ when $\vec{e}_\rho$ is in the *xy*-plane. That occurs when the angle ϕ is $\pi/2$. Putting $\phi = \pi/2$ into the formula for $\vec{e}_\rho$ gives $\vec{e}_\rho = \cos\theta\,\vec{\imath} + \sin\theta\,\vec{\jmath} + 0\vec{k}$, and we see that a vector perpendicular to that is $-\sin\theta\,\vec{\imath} + \cos\theta\,\vec{\jmath} + 0\vec{k}$. Since this vector is also a unit vector and points in the (positive) θ direction, it must be $\vec{e}_\theta$:

$$\vec{e}_\theta = -\sin\theta\,\vec{\imath} + \cos\theta\,\vec{\jmath} + 0\vec{k} \qquad (15.24)$$

Lastly, since $\vec{e}_\phi = \vec{e}_\theta \times \vec{e}_\rho$, we get:

$$\vec{e}_\phi = \cos\phi\,\cos\theta\,\vec{\imath} + \cos\phi\,\sin\theta\,\vec{\jmath} - \sin\phi\,\vec{k} \qquad (15.25)$$

Now that we have formulas for the three spherical unit vectors, our next step is to solve those for $\vec{\imath}, \vec{\jmath}, \vec{k}$ in terms of $\vec{e}_\rho, \vec{e}_\theta, \vec{e}_\phi$.

This comes down to solving a system of three equations in three unknowns. There are many ways of doing this, but we will do it by combining the formulas for $\vec{e}_\rho$ and $\vec{e}_\phi$ to eliminate $\vec{k}$, which will give us an equation involving just $\vec{\imath}$ and $\vec{\jmath}$. This, with the formula for $\vec{e}_\theta$, will then leave us with a system of two equations in two unknowns ($\vec{\imath}$ and $\vec{\jmath}$), which we will use to solve first for $\vec{\jmath}$ then for $\vec{\imath}$. Lastly, we will solve for $\vec{k}$.

First, note that

$$\sin\phi\,\vec{e}_\rho + \cos\phi\,\vec{e}_\phi = \cos\theta\,\vec{\imath} + \sin\theta\,\vec{\jmath}$$

Notes:

so that

$$\sin\theta\,(\sin\phi\,\vec{e}_\rho + \cos\phi\,\vec{e}_\phi) + \cos\theta\,\vec{e}_\theta = (\sin^2\theta + \cos^2\theta)\vec{j} = \vec{j},$$

and so:

$$\vec{j} = \sin\phi\,\sin\theta\,\vec{e}_\rho + \cos\theta\,\vec{e}_\theta + \cos\phi\,\sin\theta\,\vec{e}_\phi \qquad (15.26)$$

Likewise, we see that

$$\cos\theta\,(\sin\phi\,\vec{e}_\rho + \cos\phi\,\vec{e}_\phi) - \sin\theta\,\vec{e}_\theta = (\cos^2\theta + \sin^2\theta)\vec{i} = \vec{i},$$

and so:

$$\vec{i} = \sin\phi\,\cos\theta\,\vec{e}_\rho - \sin\theta\,\vec{e}_\theta + \cos\phi\,\cos\theta\,\vec{e}_\phi \qquad (15.27)$$

Lastly, we see that:

$$\vec{k} = \cos\phi\,\vec{e}_\rho - \sin\phi\,\vec{e}_\phi \qquad (15.28)$$

Now that we have formulas for the three Cartesian unit vectors, our next step is to get formulas for $\frac{\partial F}{\partial \rho}$, $\frac{\partial F}{\partial \theta}$, $\frac{\partial F}{\partial \phi}$ in terms of $\frac{\partial F}{\partial x}$, $\frac{\partial F}{\partial y}$, $\frac{\partial F}{\partial z}$.

By the Chain Rule, we have

$$\frac{\partial F}{\partial \rho} = \frac{\partial F}{\partial x}\frac{\partial x}{\partial \rho} + \frac{\partial F}{\partial y}\frac{\partial y}{\partial \rho} + \frac{\partial F}{\partial z}\frac{\partial z}{\partial \rho},$$

$$\frac{\partial F}{\partial \theta} = \frac{\partial F}{\partial x}\frac{\partial x}{\partial \theta} + \frac{\partial F}{\partial y}\frac{\partial y}{\partial \theta} + \frac{\partial F}{\partial z}\frac{\partial z}{\partial \theta},$$

$$\frac{\partial F}{\partial \phi} = \frac{\partial F}{\partial x}\frac{\partial x}{\partial \phi} + \frac{\partial F}{\partial y}\frac{\partial y}{\partial \phi} + \frac{\partial F}{\partial z}\frac{\partial z}{\partial \phi},$$

which yields:

$$\frac{\partial F}{\partial \rho} = \sin\phi\,\cos\theta\,\frac{\partial F}{\partial x} + \sin\phi\,\sin\theta\,\frac{\partial F}{\partial y} + \cos\phi\,\frac{\partial F}{\partial z} \qquad (15.29)$$

$$\frac{\partial F}{\partial \theta} = -\rho\sin\phi\,\sin\theta\,\frac{\partial F}{\partial x} + \rho\sin\phi\,\cos\theta\,\frac{\partial F}{\partial y} \qquad (15.30)$$

$$\frac{\partial F}{\partial \phi} = \rho\cos\phi\,\cos\theta\,\frac{\partial F}{\partial x} + \rho\cos\phi\,\sin\theta\,\frac{\partial F}{\partial y} - \rho\sin\phi\,\frac{\partial F}{\partial z} \qquad (15.31)$$

Our next step is to invert the previous relation to solve for $\frac{\partial F}{\partial x}$, $\frac{\partial F}{\partial y}$, $\frac{\partial F}{\partial z}$ in terms of $\frac{\partial F}{\partial \rho}$, $\frac{\partial F}{\partial \theta}$, $\frac{\partial F}{\partial \phi}$.

Again, this involves solving a system of three equations in three unknowns.

Notes:

Using a similar process of elimination as before, we get:

$$\frac{\partial F}{\partial x} = \frac{1}{\rho \sin \phi} \left(\rho \sin^2 \phi \, \cos \theta \, \frac{\partial F}{\partial \rho} - \sin \theta \, \frac{\partial F}{\partial \theta} + \sin \phi \, \cos \phi \, \cos \theta \, \frac{\partial F}{\partial \phi} \right)$$
$$(15.32)$$

$$\frac{\partial F}{\partial y} = \frac{1}{\rho \sin \phi} \left(\rho \sin^2 \phi \, \sin \theta \, \frac{\partial F}{\partial \rho} + \cos \theta \, \frac{\partial F}{\partial \theta} + \sin \phi \, \cos \phi \, \sin \theta \, \frac{\partial F}{\partial \phi} \right) \quad (15.33)$$

$$\frac{\partial F}{\partial z} = \frac{1}{\rho} \left(\rho \cos \phi \, \frac{\partial F}{\partial \rho} - \sin \phi \, \frac{\partial F}{\partial \phi} \right) \tag{15.34}$$

Finally, we substitute the Equations (15.26), (15.27), (15.28), and (15.34) into the Cartesian gradient formula $\nabla F(x, y, z) = \frac{\partial F}{\partial x} \vec{\imath} + \frac{\partial F}{\partial y} \vec{\jmath} + \frac{\partial F}{\partial z} \vec{k}$.

Doing this last step is perhaps the most tedious, since it involves simplifying $3 \times 3 + 3 \times 3 + 2 \times 2 = 22$ terms! Namely,

$$\begin{aligned}
\nabla F = {} & \frac{1}{\rho \sin \phi} \left(\rho \sin^2 \phi \, \cos \theta \, \frac{\partial F}{\partial \rho} - \sin \theta \, \frac{\partial F}{\partial \theta} + \sin \phi \, \cos \phi \, \cos \theta \, \frac{\partial F}{\partial \phi} \right) (\sin \phi \, \cos \theta \, \vec{e}_\rho - \sin \theta \, \vec{e}_\theta \\
& + \cos \phi \, \cos \theta \, \vec{e}_\phi) \\
& + \frac{1}{\rho \sin \phi} \left(\rho \sin^2 \phi \, \sin \theta \, \frac{\partial F}{\partial \rho} + \cos \theta \, \frac{\partial F}{\partial \theta} + \sin \phi \, \cos \phi \, \sin \theta \, \frac{\partial F}{\partial \phi} \right) (\sin \phi \, \sin \theta \, \vec{e}_\rho + \cos \theta \, \vec{e}_\theta \\
& + \cos \phi \, \sin \theta \, \vec{e}_\phi) \\
& + \frac{1}{\rho} \left(\rho \cos \phi \, \frac{\partial F}{\partial \rho} - \sin \phi \, \frac{\partial F}{\partial \phi} \right) (\cos \phi \, \vec{e}_\rho - \sin \phi \, \vec{e}_\phi),
\end{aligned}$$

which we see has 8 terms involving $\vec{e}_\rho$, 6 terms involving $\vec{e}_\theta$, and 8 terms involving $\vec{e}_\phi$. But the algebra is straightforward and yields the desired result:

$$\nabla F = \frac{\partial F}{\partial \rho} \vec{e}_\rho + \frac{1}{\rho \sin \phi} \frac{\partial F}{\partial \theta} \vec{e}_\theta + \frac{1}{\rho} \frac{\partial F}{\partial \phi} \vec{e}_\phi$$

Example 4   Practicing in Spherical Coordinates

In Example 1 we showed that $\nabla \|\vec{r}\|^2 = 2\vec{r}$ and $\Delta \|\vec{r}\|^2 = 6$, where $\vec{r}(x, y, z) = x\vec{\imath} + y\vec{\jmath} + z\vec{k}$ in Cartesian coordinates. Verify that we get the same answers if we switch to spherical coordinates.

SOLUTION   Since $\|\vec{r}\|^2 = x^2 + y^2 + z^2 = \rho^2$ in spherical coordinates, let $F(\rho, \theta, \phi) = \rho^2$ (so that $F(\rho, \theta, \phi) = \|\vec{r}\|^2$). The gradient of F in spherical

Notes:

coordinates is

$$\nabla F = \frac{\partial F}{\partial \rho}\, \vec{e}_\rho + \frac{1}{\rho \sin \phi}\, \frac{\partial F}{\partial \theta}\, \vec{e}_\theta + \frac{1}{\rho}\, \frac{\partial F}{\partial \phi}\, \vec{e}_\phi$$

$$= 2\rho\, \vec{e}_\rho + \frac{1}{\rho \sin \phi}\, (0)\, \vec{e}_\theta + \frac{1}{\rho}\, (0)\, \vec{e}_\phi$$

$$= 2\rho\, \vec{e}_\rho = 2\rho\, \frac{\vec{r}}{\|\vec{r}\|}, \ \text{as we showed earlier, so}$$

$$= 2\rho\, \frac{\vec{r}}{\rho} = 2\, \vec{r}, \ \text{as expected. And the Laplacian is}$$

$$\Delta F = \frac{1}{\rho^2}\, \frac{\partial}{\partial \rho}\left(\rho^2\, \frac{\partial F}{\partial \rho} \right) + \frac{1}{\rho^2 \sin^2 \phi}\, \frac{\partial^2 F}{\partial \theta^2} + \frac{1}{\rho^2 \sin \phi}\, \frac{\partial}{\partial \phi}\left(\sin \phi\, \frac{\partial F}{\partial \phi} \right)$$

$$= \frac{1}{\rho^2}\, \frac{\partial}{\partial \rho}(\rho^2\, 2\rho) + \frac{1}{\rho^2 \sin \phi}\, (0) + \frac{1}{\rho^2 \sin \phi}\, \frac{\partial}{\partial \phi}\, (\sin \phi\, (0))$$

$$= \frac{1}{\rho^2}\, \frac{\partial}{\partial \rho}(2\rho^3) + 0 + 0$$

$$= \frac{1}{\rho^2}\, (6\rho^2) = 6, \ \text{as expected.}$$

Notes:

Exercises 15.6

Problems

In Exercises 1–6, find the Laplacian of the function $f(x, y, z)$ in Cartesian coordinates.

1. $f(x, y, z) = x + y + z$

2. $f(x, y, z) = x^5$

3. $f(x, y, z) = (x^2 + y^2 + z^2)^{3/2}$

4. $f(x, y, z) = e^{x+y+z}$

5. $f(x, y, z) = x^3 + y^3 + z^3$

6. $f(x, y, z) = e^{-x^2-y^2-z^2}$

7. Find the Laplacian of the function in Exercise 3 in spherical coordinates.

8. Find the Laplacian of the function in Exercise 6 in spherical coordinates.

9. Let $f(x, y, z) = \dfrac{z}{x^2 + y^2}$ in Cartesian coordinates. Find ∇f in cylindrical coordinates.

10. For $\vec{f}(r, \theta, z) = r\,\vec{e}_r + z\,\sin\theta\,\vec{e}_\theta + rz\,\vec{e}_z$ in cylindrical coordinates, find $\operatorname{div}\vec{f}$ and $\operatorname{curl}\vec{f}$.

11. For $\vec{f}(\rho, \theta, \phi) = \vec{e}_\rho + \rho\,\cos\theta\,\vec{e}_\theta + \rho\,\vec{e}_\phi$ in spherical coordinates, find $\operatorname{div}\vec{f}$ and $\operatorname{curl}\vec{f}$.

In Exercises 12–23, prove the given formula ($r = \|\vec{r}\|$ is the length of the position vector field $\vec{r}(x, y, z) = x\,\vec{\imath} + y\,\vec{\jmath} + z\,\vec{k}$).

12. $\nabla(1/r) = -\vec{r}/r^3$

13. $\Delta(1/r) = 0$

14. $\nabla \cdot (\vec{r}/r^3) = 0$

15. $\nabla(\ln r) = \vec{r}/r^2$

16. $\operatorname{div}(\vec{F} + \vec{G}) = \operatorname{div}\vec{F} + \operatorname{div}\vec{G}$

17. $\operatorname{curl}(\vec{F} + \vec{G}) = \operatorname{curl}\vec{F} + \operatorname{curl}\vec{G}$

18. $\operatorname{div}(f\vec{F}) = f\operatorname{div}\vec{F} + \vec{F} \cdot \nabla f$

19. $\operatorname{div}(\vec{F} \times \vec{G}) = \vec{G} \cdot \operatorname{curl}\vec{F} - \vec{F} \cdot \operatorname{curl}\vec{G}$

20. $\operatorname{div}(\nabla f \nabla g) = 0$

21. $\operatorname{curl}(f\vec{F}) = f\operatorname{curl}\vec{F} + (\nabla f) \times \vec{F}$

22. $\operatorname{curl}(\operatorname{curl}\vec{F}) = \nabla(\operatorname{div}\vec{F}) - \Delta\vec{F}$

23. $\Delta(fg) = f\,\Delta g + g\,\Delta f + 2(\nabla f \cdot \nabla g)$

24. Prove Theorem 148.

25. Derive the gradient formula in cylindrical coordinates: $\nabla F = \dfrac{\partial F}{\partial r}\vec{e}_r + \dfrac{1}{r}\dfrac{\partial F}{\partial \theta}\vec{e}_\theta + \dfrac{\partial F}{\partial z}\vec{e}_z$

26. Use $\vec{f} = u\,\nabla v$ in the Divergence Theorem to prove:

    (a) *Green's first identity:* $\displaystyle\iiint_S (u\,\Delta v + (\nabla u)\cdot(\nabla v))\,dV = \iint_\Sigma (u\,\nabla v)\cdot d\sigma$

    (b) *Green's second identity:* $\displaystyle\iiint_S (u\,\Delta v - v\,\Delta u)\,dV = \iint_\Sigma (u\,\nabla v - v\,\nabla u)\cdot d\sigma$

27. Suppose that $\Delta u = 0$ (i.e. u is *harmonic*) over $\mathbb{R}^3$. Define the *normal derivative* $\dfrac{\partial u}{\partial n}$ of u over a closed surface Σ with outward unit normal vector $\vec{n}$ by $\dfrac{\partial u}{\partial n} = D_{\vec{n}}u = \vec{n}\cdot\nabla u$. Show that $\displaystyle\iint_\Sigma \dfrac{\partial u}{\partial n}\,d\sigma = 0$. (*Hint: Use Green's second identity.*)

Notes:

A: Solutions To Selected Problems

Chapter 1

Exercises 1.0

1. $(-\infty, \infty)$

2. $[-7, \infty)$

3. $(-\infty, -1] \cup [7, \infty)$

4. $(-\infty, \infty)$

5. $(-\infty, 2) \cup (2, \infty)$

6. $(-\infty, \infty)$

7. $(-\infty, \infty)$

8. $(-\infty, -2) \cup (-2, 2) \cup (2, \infty)$

9. $(-\infty, 0) \cup (0, \infty)$

10. $(-\infty, \infty)$

11.

12.

13.

14.
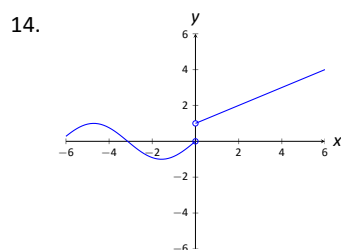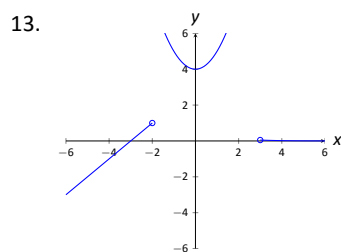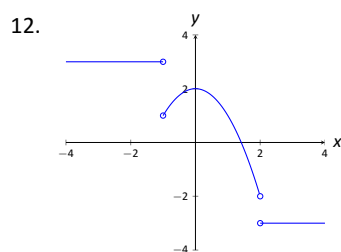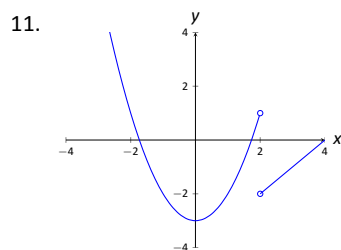
15.
   (a) 14
   (b) 11
   (c) $3a^2 - 2a + 6$
   (d) $3(x + h)^2 - 2(x + h) + 6$
   (e) $\dfrac{h(3h + 6x - 2)}{h}$

16.
   (a) $\sqrt{2}$
   (b) undefined
   (c) $\sqrt{t - 2}$
   (d) $\sqrt{x + h - 2}$
   (e) $\dfrac{\sqrt{x + h - 2} - \sqrt{x - 2}}{h}$
   $= \dfrac{h}{h(\sqrt{x + h - 2} + \sqrt{x - 2})}$

17.
   (a) -1
   (b) $\dfrac{1}{9}$
   (c) $\dfrac{1}{t + 3}$
   (d) $\dfrac{1}{x + h}$
   (e) $\dfrac{\frac{1}{x+h} - \frac{1}{x}}{h} = -\dfrac{h}{hx(x + h)}$

18. $\{2.12\} \cup [2.13, 2.14) \cup (2.15, \infty)$

19. $(-\infty, 5.677) \cup [5.678, 5.679) \cup (5.679, \infty)$

20. $(0.3, 0.8]$

21. $(0.1, 0.2]$

Exercises 1.1

1. Answers will vary.

3. F

5. Answers will vary.

7. -5

9. 2

11. Limit does not exist.

13. 7

15. Limit does not exist.

17.

h	$\frac{f(a+h)-f(a)}{h}$
-0.1	9
-0.01	9
0.01	9
0.1	9

The limit seems to be exactly 9.

19.

h	$\frac{f(a+h)-f(a)}{h}$
−0.1	−0.114943
−0.01	−0.111483
0.01	−0.110742
0.1	−0.107527

The limit is approx. −0.11.

21.

h	$\frac{f(a+h)-f(a)}{h}$
−0.1	0.202027
−0.01	0.2002
0.01	0.1998
0.1	0.198026

The limit is approx. 0.2.

23.

h	$\frac{f(a+h)-f(a)}{h}$
−0.1	−0.0499583
−0.01	−0.00499996
0.01	0.00499996
0.1	0.0499583

The limit is approx. 0.

Exercises 1.2

1. ε should be given first, and the restriction $|x - a| < \delta$ implies $|f(x) - K| < \varepsilon$, not the other way around.

3. T

5. $\delta \leq 0.45$

7. Given $\varepsilon > 0$, choose $\delta = \varepsilon$.

$$|x - 5| < \delta = \varepsilon$$
$$|5 - x| < \varepsilon$$
$$|3 - x - (-2)| < \varepsilon.$$

Thus $\lim_{x \to 5} 3 - x = -2.$

9. Given $\varepsilon > 0$, let $\delta = \dfrac{\varepsilon}{2}$. Then:

$$|x - 3| < \delta = \frac{\varepsilon}{2}$$
$$2\,|x - 3| < \frac{\varepsilon}{2} \cdot 2$$
$$|-2x + 6| < \varepsilon$$
$$|5 - 2x + 1| < \varepsilon$$

Thus $\lim_{x \to 3} 5 - 2x = -1.$

11. Given $\varepsilon > 0$, let $\delta = \frac{\varepsilon}{10}$. Then:

$$|x - 4| < \delta = \frac{\varepsilon}{10}$$
$$|x - 4| < \frac{\varepsilon}{x + 5}$$
$$|x - 4| \cdot |x + 5| < \frac{\varepsilon}{x + 5} \cdot |x + 5|$$

Assuming x is near 4, $x + 5$ is positive and we can drop the absolute value signs on the right.

$$|x - 4| \cdot |x + 5| < \frac{\varepsilon}{x + 5} \cdot (x + 5)$$
$$|x^2 + x - 20| < \varepsilon$$
$$|(x^2 + x - 5) - 15| < \varepsilon.$$

Thus, $\lim_{x \to 4} x^2 + x - 5 = 15.$

13. Let $\varepsilon > 0$ be given. We wish to find $\delta > 0$ such that when $|x - 2| < \delta$, $|f(x) - 5| < \varepsilon$. However, since $f(x) = 5$, a constant function, the latter inequality is simply $|5 - 5| < \varepsilon$, which is always true. Thus we can choose any δ we like; we arbitrarily choose $\delta = \varepsilon$.

Exercises 1.3

1. Answers will vary.

3. Answers will vary.

5. As x is near 1, both f and g are near 0, but f is approximately twice the size of g. (I.e., $f(x) \approx 2g(x)$.)

7. 6

9. Limit does not exist.

11. Not possible to know.

13. −45

15. −1

17. π

19. $-0.000000015 \approx 0$

21. Limit does not exist

23. 2

25. $\frac{\pi^2 + 3\pi + 5}{5\pi^2 - 2\pi - 3} \approx 0.6064$

27. −8

29. 10

31. −3/2

33. 1/6

35. −1/9

37. −8

39. 0

41. 1

43. 0

45. 5/8

47. $\pi/180$

49. 5/7

Exercises 1.4

1. The function approaches different values from the left and right; the function grows without bound; the function oscillates.

3. F

5.
 (a) 2
 (b) 2
 (c) 2
 (d) 1
 (e) As f is not defined for $x < 0$, this limit is not defined.
 (f) 1

7.
    (a) 2

    (b) 0

    (c) Does not exist.

    (d) 1

9.
    (a) 4

    (b) -4

    (c) Does not exist.

    (d) 0

11.
    (a) $a - 1$

    (b) a

    (c) Does not exist.

    (d) a

13. 0

15. DNE

17.
    (a) -1

    (b) 0

    (c) Does not exist.

    (d) 0

19.
    (a) -1

    (b) 0

    (c) Does not exist.

    (d) 0

21.
    (a) 2

    (b) 0

    (c) Does not exist

    (d) 1

23.
    (a) c

    (b) c

    (c) c

    (d) c

25.
    (a) -1

    (b) 1

    (c) Does not exist.

    (d) Undefined

27. Answers will vary.

29. Answers will vary.

31. $-3/5$

33. $\frac{1}{2\sqrt{3}}$

35. -1.63

Exercises 1.5

1. F

3. F

5. T

7. Answers will vary.

9.
    (a) ∞

    (b) ∞

11.
    (a) 1

    (b) 0

    (c) $1/2$

    (d) $1/2$

13.
    (a) Limit does not exist

    (b) Limit does not exist

15. Tables will vary.

    (a)

x	$f(x)$
2.9	-15.1224
2.99	-159.12
2.999	-1599.12

    It seems $\lim_{x\to 3^-} f(x) = -\infty$.

    (b)

x	$f(x)$
3.1	16.8824
3.01	160.88
3.001	1600.88

    It seems $\lim_{x\to 3^+} f(x) = \infty$.

    (c) It seems $\lim_{x\to 3} f(x)$ does not exist.

17. Tables will vary.

    (a)

x	$f(x)$
2.9	132.857
2.99	12124.4

    It seems $\lim_{x\to 3^-} f(x) = \infty$.

    (b)

x	$f(x)$
3.1	108.039
3.01	11876.4

    It seems $\lim_{x\to 3^+} f(x) = \infty$.

    (c) It seems $\lim_{x\to 3} f(x) = \infty$.

19. Horizontal asymptote at $y = 2$; vertical asymptotes at $x = -5, 4$.

21. Horizontal asymptote at $y = 0$; vertical asymptotes at $x = -1, 0$.

23. No horizontal or vertical asymptotes.

25. $y = 2$

27. ∞

29. $-\infty$

31. $-2/3$

33. $-\sqrt{10}/2$

35. Let $\varepsilon > 0$ be given. We wish to find $\delta > 0$ such that when $|x - 1| < \delta$, $|f(x) - 3| < \varepsilon$.
Scratch-Work: Consider $|f(x) - 3| < \varepsilon$, keeping in mind we want to make a statement about $|x - 1|$:

$$|f(x) - 3| < \varepsilon$$
$$|5x - 2 - 3| < \varepsilon$$
$$|5x - 5| < \varepsilon$$
$$5\,|x - 1| < \varepsilon$$
$$|x - 1| < \frac{\varepsilon}{5}$$

suggesting $\delta = \frac{\varepsilon}{5}$.
Proof: Given $\varepsilon > 0$, let $\delta = \dfrac{\varepsilon}{5}$. Then:

$$|x - 1| < \delta$$
$$|x - 1| < \frac{\varepsilon}{5}$$
$$5\,|x - 1| < \frac{\varepsilon}{5} \cdot 5$$
$$|5x - 5| < \varepsilon$$
$$|5x - 2 - 3| < \varepsilon$$

Thus $\lim\limits_{x \to 1} 5x - 2 = 3$.

37. 1

Exercises 1.6

1. Answers will vary.

3. A root of a function f is a value c such that $f(c) = 0$.

5. F

7. T

9. F

11. No; $\lim\limits_{x \to 1} f(x) = 2$, while $f(1) = 1$.

13. No; $f(1)$ does not exist.

15. Yes

17.
 (a) No; $\lim\limits_{x \to -2} f(x) \neq f(-2)$
 (b) Yes
 (c) No; $f(2)$ is not defined.

19.
 (a) Yes
 (b) No; the left and right hand limits at 1 are not equal.

21.
 (a) Yes
 (b) No. $\lim_{x \to 8} f(x) = 16/5 \neq f(8) = 5$.

23. $(-\infty, -2] \cup [2, \infty)$

25. $(-\infty, -\sqrt{6}] \cup [\sqrt{6}, \infty)$

27. $(-\infty, \infty)$

29. $(0, \infty)$

31. $(-\infty, 0]$

33. $(-\infty, -4) \cup (-4, 2) \cup (2, 5) \cup (5, \infty)$

35. Yes. The only "questionable" place is at $x = 3$, but the left and right limits agree.

37. Yes, by the Intermediate Value Theorem. In fact, we can be more specific and state such a value c exists in $(0, 2)$, not just in $(-3, 7)$.

39. We cannot say; the Intermediate Value Theorem only applies to continuous functions. As we do know know if h is continuous, we cannot say.

41. $a = -1$ and $\frac{4}{3}$

43. $a = -1$

45. Answers will vary.

47. Answers will vary.

49. Use the Bisection Method with an appropriate interval.

51. Use the Bisection Method with an appropriate interval.

53.

x	$f(x)$
-0.81	-2.34129
-0.801	-2.33413
-0.79	-2.32542
-0.799	-2.33254

The top two lines give an approximation of the limit from the left: -2.33. The bottom two lines give an approximation from the right: -2.33 as well.

Chapter 2

Exercises 2.0

1. $80x^{12}y^{17}$

2. $\dfrac{a}{16b^7}$

3. $\dfrac{-x^3}{16y^{22}z^{35}}$

4. $x^2 y^4 z^5 \sqrt[4]{z} = x^2 y^4 z^{21/4}$

5. $\dfrac{5(x-1)}{3x^{\frac{1}{3}}}$

6. $\dfrac{-5x+4}{2x^{\frac{1}{2}}(x+4)^2}$

7. $6x(3x^2 + 2)^3(x^2 - 5)^2(7x^2 - 18)$

8. (a) 8   (b) 44   (c) $x^2 - 6x + 8$ (d) $x^2 + 2x - 4$

9. (a) $-\frac{1}{3}$ (b) undefined   (c) $\dfrac{1}{\sqrt{x-2} - 5}$ (d) $\sqrt{\dfrac{1}{x-5} - 2}$

10. (a) Possible solution: $f(x) = \frac{5}{x}$ and $g(x) = x + 4$
(b) Possible solution: $f(x) = |x|$ and $g(x) = 4 - x^2$
(c) Possible solution: $f(x) = \sqrt{x - 5}$ and $g(x) = (x + 2)^2$

11. (a) Possible solution: $f(x) = \sqrt[3]{x}$, $g(x) = x^2$, and $h(x) = 2x + 1$
(b) Possible solution: $f(x) = 2x + 1$, $g(x) = \sqrt[3]{x}$, and $h(x) = x^2$

Exercises 2.1

1. T

3. Answers will vary.

5. Answers will vary.

7. (a) $f'(x) = 2$, (b) $y = 2x$

9. (a) $g'(x) = 2x$, (b) $y = -4x - 4$

11. (a) $f''(x) = 6x - 1$, (b) $y = -7x + 1$

13. (a) $r'(x) = \frac{-1}{x^2}$, (b) $y = -\frac{x}{4} - 1$

15. (a) $f'(x) = \dfrac{-1}{(s-2)^2}$, (b) $y = -x + 4$

17. $f(x) = x^4, c = 3$

19. $f(x) = \cos x, c = -\pi$.

21. $y = .248x + 1.006$

23. $y = 7.77(x - 2) + e^2$, or $y = 7.77(x - 2) + 7.39$.

25.

    (a) Approximations will vary; they should match (c) closely.

    (b) $f'(x) = 2x$

    (c) At $(-1, 0)$, slope is -2. At $(0, -1)$, slope is 0. At $(2, 3)$, slope is 4.

27.

29.
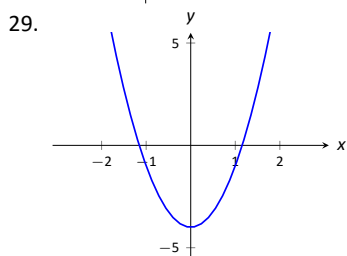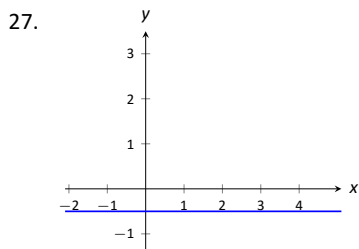
31.

    (a) Approximately on $(-1.5, 1.5)$.

    (b) Approximately on $(-\infty, -1.5) \cup (1.5, \infty)$.

    (c) Approximately at $x = \pm 1.5$.

    (d) On $(-\infty, -1) \cup (0, 1)$.

    (e) On $(-1, 0) \cup (1, \infty)$.

    (f) At $x = \pm 1$.

33. Approximately 0.54.

35.

    (a) 1

    (b) 3

    (c) Does not exist

    (d) $(-\infty, -3) \cup (3, \infty)$

Exercises 2.2

1. Velocity

3. Linear functions.

5. -17

7. $f(10.1)$ is likely most accurate, as accuracy is lost the farther from $x = 10$ we go.

9. 6

11. ft/s^2

13.

    (a) thousands of dollars per car

    (b) It is likely that $P(0) < 0$. That is, negative profit for not producing any cars.

15. $f(x) = g'(x)$

17. $g(x) = f'(x)$

19. $f(6) = 1, f'(6) = -\frac{3}{4}$

21. Answers vary. Possible solution
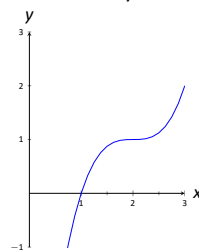
23.

25. $f'(x) = 10x$

27. $f'(\pi) \approx 0$.

Exercises 2.3

1. Power Rule.

3. One answer is $f(x) = 10e^x$.

5. $f(x), g(x), h(x),$ and $m(x)$

7. One possible answer is $f(x) = 17x - 205$.

9. $f'(x)$ is a velocity function, and $f''(x)$ is acceleration.

11. $f'(x) = 14x - 5$

13. $m'(t) = 45t^4 - \frac{3}{8}t^2 + 3$

15. $f'(r) = 6e^r$

17. $f'(x) = \frac{2}{x} - 1$

19. $h'(t) = e^t - \cos t + \sin t$

21. $f'(t) = 0$

23. $g'(x) = 24x^2 - 120x + 150$

25. $f'(x) = 18x - 12$

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

29.

31. $n = -3, 2$

33. d is f, c is f', b is f'', and a is f'''

35. $g'(x) = -2\sin x$ $g''(x) = -2\cos x$ $g'''(x) = 2\sin x$ $g^{(4)}(x) = 2\cos x$

37. $p'(\theta) = 4\theta^3 - 3\theta^2$ $p''(\theta) = 12\theta^2 - 6\theta$ $p'''(\theta) = 24\theta - 6$ $p^{(4)}(\theta) = 24$

39. $f'(x) = f''(x) = f'''(x) = f^{(4)}(x) = 0$

41.

    (a) $v(t) = 5e^x - 5$, $a(t) = 5e^x$

    (b) $a(2) = 5e^2$ ft/s^2

    (c) $v(t) = 0$ at $t = 0$ sec, $a(0) = 5$ in/s^2

43. Tangent line: $y = t + 4$

45. Tangent line: $y = 4$

47. Tangent line: $y = 2x + 3$

49. The tangent line to $f(x) = x^4$ at $x = 3$ is $y = 108(x - 3) + 81$; thus $(3.01)^4 \approx y(3.01) = 108(.01) + 81 = 82.08$.

Exercises 2.4

1. F

3. T

5. F

7.

$$\frac{d}{dx}(\cot x) = \frac{d}{dx}\left(\frac{\cos x}{\sin x}\right)$$
$$= \frac{\sin x(-\sin x) - (\cos x)(\cos x)}{(\sin x)^2}$$
$$= \frac{-[(\sin x)^2 + (\cos x)^2]}{(\sin x)^2}$$
$$= \frac{-1}{(\sin x)^2} = -\csc^2 x$$

9.

$$\frac{d}{dx}(\csc x) = \frac{d}{dx}\left(\frac{1}{\sin x}\right)$$
$$= \frac{\sin x \cdot 0 - 1 \cdot (\cos x)}{(\sin x)^2}$$
$$= \frac{-\cos x}{(\sin x)^2} = -\csc x \cot x$$

11.

    (a) $g'(x) = 4x(5x^3) + 2x^2(15x^2)$

    (b) $g'(x) = 50x^4$

    (c) They are equal.

13.

    (a) $f'(x) = 2x(3 - x^3) + (x^2 + 5)(-3x^2)$

    (b) $f'(x) = -5x^4 - 15x^2 + 6x$

    (c) They are equal.

15.

    (a) $g'(x) = \frac{2x^2(3x^2 - 4x) - (x^3 - 2x^2)(4x)}{4x^4}$

    (b) $g'(x) = 1/2$

    (c) They are equal.

17.

    (a) $f'(t) = \frac{(t+1)(2t) - (t^2 - 1)(1)}{(t+1)^2}$

    (b) $f(t) = t - 1$ when $t \neq -1$, so $f'(t) = 1$.

    (c) They are equal.

19. $f'(t) = \frac{-2}{t^3}(\csc t - 4) + \frac{1}{t^2}(-\csc t \cot t)$

21. $F'(y) = \frac{8}{3}y^{5/3} + 15y^{2/3} = \frac{\sqrt[3]{y^2}(8y + 45)}{3}$

23. $y' = \dfrac{4 - x}{2\sqrt{x}(x + 4)^2}$

25. $g'(t) = \frac{(\cos t - 2t^2)(5t^4) - (t^5)(-\sin t - 4t)}{(\cos t - 2t^2)^2}$

27. $h'(t) = 14t + 6$

29. $f'(x) = -\dfrac{1}{x^2} + \dfrac{5}{2x^3\sqrt{x}} = \dfrac{-2x\sqrt{x} + 5}{2x^3\sqrt{x}}$

31. $g'(x) = -\frac{1 + 2x + 3x^2}{(1 + x + x^2 + x^3)^2}$

33. $f'(x) = 7$

35. $f'(x) = \frac{\sin^2(x) + \cos^2(x) + 3\cos(x)}{(\cos(x) + 3)^2}$

37. $g'(t) = 12t^2 e^t + 4t^3 e^t - \cos^2 t + \sin^2 t$

39. $F'(x) = (8x - 1)(x^2 + 4x + 7)(3x^2) + (8x - 1)(2x + 4)(x^3 - 5) + (8)(x^2 + 4x + 7)(x^3 - 5)$

41. $f'(x) = 2xe^x \tan x = x^2 e^x \tan x + x^2 e^x \sec^2 x$

43. $y = 2x + 2$

45. $y = 4$

47. $x = 3/2$

49. $f'(x)$ is never 0.

51. $f''(x) = 2\cos x - x\sin x$

53. $f''(x) = \cot^2 x \csc x + \csc^3 x$

55. 1

57. -4

59. $-\dfrac{1}{25}$

61. (a) $-\frac{7}{2}$ (b) $\frac{1}{8}$ (c) $-\frac{9}{2}$ (d) $\frac{15}{2}$

Exercises 2.5

1. T

3. F

5. T

7. $f'(x) = 10(4x^3 - x)^9 \cdot (12x^2 - 1) = (120x^2 - 10)(4x^3 - x)^9$

9. $g'(\theta) = 3(\sin\theta + \cos\theta)^2(\cos\theta - \sin\theta)$

11. $f'(x) = 4\left(x + \frac{1}{x}\right)^3\left(1 - \frac{1}{x^2}\right)$

13. $f'(x) = -3\sin(3x)$

15. $h'(t) = 8\sin^3(2t)\cos(2t)$

17. $g'(x) = 2(\tan x \sec^2 x - x\sec^2(x^2))$

19. $f'(x) = -\tan x$

21. $f'(x) = 2/x$

23. $r'(x) = \dfrac{-6(x-1)}{x^3\sqrt{4x-3}}$

25. $h'(x) = 200(2x+1)^9[(2x+1)10+1]^9$

27. $F'(x) = 2(2x+1)(2x+3)^2(24x^2+26x+3)$

29. $g'(t) = 5\cos(t^2+3t)\cos(5t-7) - (2t+3)\sin(t^2+3t)\sin(5t-7)$

31. $a'(t) = 7t^2 e^{\tan(t^2)}(2t^2\sec^2(t^2)+3)$

33. $k'(x) = -\sin(x\sin x^3)(3x^3\cos x^3 + \sin x^3)$

35. 90

37. (a) 12   (b) 2.5   (c) 9   (d) 35

39. $y = 15(t-1)+1$

41. $y = -5e(t+1)+e$

43. In both cases the derivative is the same: k/x.

45. Let $h(x) = x^{-1}$. Then $\dfrac{d}{dx}\dfrac{f(x)}{g(x)} = \dfrac{d}{dx}\left[f(x)\cdot h(g(x))\right] = \dfrac{d}{dx}[f(x)]\cdot h(g(x)) + f(x)\cdot\dfrac{d}{dx}[h(g(x))] = f'(x)\cdot h(g(x)) + f(x)\cdot h'(g(x))\cdot g'(x) = f'(x)[g(x)]^{-1} - f(x)[g(x)]^{-2}g'(x) = \dfrac{f'(x)g(x) - f(x)g'(x)}{[g(x)]^2}$

47.

    (a) $^\circ$ F/mph

    (b) The sign would be negative; when the wind is blowing at 10 mph, any increase in wind speed will make it feel colder, i.e., a lower number on the Fahrenheit scale.

Exercises 2.6

1. Answers will vary.

3. T

5. $\dfrac{dy}{dx} = \dfrac{-4x^3}{2y+1}$

7. $\dfrac{dy}{dx} = \sin x\sec y$

9. $\dfrac{dy}{dx} = \dfrac{y}{x}$

11. $-\dfrac{x}{y^2}$

13. $\dfrac{x^2+2xy^2-y}{2x^2y-x+y^2}$

15. $-\dfrac{x}{y}$

17. $\dfrac{e^x(x+1)}{e^y(y+1)}$

19. $\dfrac{y-4xy\sqrt{xy}}{2x^2\sqrt{xy}-x}$

21. $\dfrac{dy}{dx} = -\dfrac{y+2x}{2y+x}$

23.

    (a) $y = 0$

    (b) $y = -1.859(x-0.1)+0.281$

25.

    (a) $y = 4$

    (b) $y = 0.93(x-2)+\sqrt[4]{108}$

27.

    (a) $y = -\dfrac{1}{\sqrt{3}}\left(x-\dfrac{7}{2}\right)+\dfrac{6+3\sqrt{3}}{2}$

    (b) $y = \sqrt{3}\left(x-\dfrac{4+3\sqrt{3}}{2}\right)+\dfrac{3}{2}$

29. $\dfrac{d^2y}{dx^2} = \dfrac{(2y+1)(-12x^2)+4x^3\left(2\frac{-4x^3}{2y+1}\right)}{(2y+1)^2}$

31. $\dfrac{d^2y}{dx^2} = \dfrac{\cos x\cos y + \sin^2 x\tan y}{\cos^2 y}$

Chapter 3

Exercises 3.1

1. Answers will vary.

3. Answers will vary.

5. F

7. A: abs. min; B: none; C: abs. max; D: none; E: none

9. $f'(0) = 0$

11. $f'(\pi/2) = 0$; $f'(3\pi/2) = 0$

13. $f'(0) = 0$

15. $f'(2)$ is not defined; $f'(6) = 0$

17. min: $(5, -134.5)$
    max: $(0, 3)$

19. min: $(0, 0)$ and $(\pm 2, 0)$
    max: $(\pm 2\sqrt{2/3}, 16\sqrt{3}/9)$

21. min: $(0, 0)$
    max: $(5, 5/6)$

23. min: $(0, 0)$ and $(\pi, 0)$
    max: $(3\pi/4, \frac{\sqrt{2}e^{3\pi/4}}{2})$

25. min: $(2, 2^{2/3}-2)$
    max: $(8/27, 4/27)$

27.

    (a) x^3-x, x^3, and x^3+x have 2, 1, and 0 critical numbers respectively. Because the derivative is a quadratic with at most 2 roots, a cubic cannot have 3 or more critical numbers.

    (b) A cubic can only have 2 or 0 extreme values.

29. $\dfrac{dy}{dx} = \dfrac{y(y-2x)}{x(x-2y)}$

31. $3x^2+1$

Exercises 3.2

1. Answers will vary.

3. Any c in $[-1, 1]$ is valid.

5. $c = -1/2$

7. Rolle's Thm. does not apply.

9. Rolle's Thm. does not apply.

11. $c = 0$

13. $c = 3/\sqrt{2}$

15. The Mean Value Theorem does not apply.

17. $c = \pm\sec^{-1}(2/\sqrt{\pi})$

19. $c = \frac{5 \pm 7\sqrt{7}}{6}$

21. No. Otherwise, with c given by the Mean Value Theorem, $\frac{4 - -1}{2 - 0} = f'(c) \leq 2$, a contradiction.

23. If f has more than 3 real roots, then Rolle's Theorem implies f' is a quadratic with more than 2 real roots.

25. $2pc + q = f'(c) = \frac{f(b) - f(a)}{b - a} = \frac{pb^2 + qb + r - pa^2 - qa - r}{b - a} = \frac{p(b^2 - a^2) + q(b - a)}{b - a} = p(b + a) + q$ implies that $c = \frac{a+b}{2}$.

27. They are the odd, integer valued multiples of $\pi/2$ (such as $0, \pm\pi/2, \pm 3\pi/2, \pm 5\pi/2$, etc.)

Exercises 3.3

1. Answers will vary.

3. Answers will vary.

5. Increasing

7. decreasing on $(0, \frac{\pi}{6})$; $(\frac{\pi}{2}, \frac{5\pi 6}{})$; $(\frac{3\pi}{2}, 2\pi)$,
   increasing on $(\frac{\pi}{6}, \frac{\pi}{2})$; $(\frac{5\pi}{6}, \frac{3\pi}{2})$;
   local maxima when $x = \frac{\pi}{2}, \frac{3\pi}{2}$,
   local minima when $x = \frac{\pi}{6}, \frac{5\pi}{6}$.

9. decreasing on $(-1, 1)$,
   increasing on $(-\infty, -1)$; $(1, \infty)$;
   local maxima when $x = -1$,
   local minima when $x = 1$.

11. Graph and verify.

13. Graph and verify.

15. Graph and verify.

17. Graph and verify.

19. domain=$(-\infty, \infty)$;
    c.p. at $c = -2, 0$;
    increasing on $(-\infty, -2)$; $(0, \infty)$;
    decreasing on $(-2, 0)$;
    rel. min at $x = 0$;
    rel. max at $x = -2$.

21. domain=$(-\infty, \infty)$
    c.p. at $c = 1$;
    increasing on $(-\infty, \infty)$

23. domain=$(-\infty, -1) \cup (-1, 1) \cup (1, \infty)$;
    c.p. at $c = 0$;
    decreasing on $(-\infty, -1)$; $(-1, 0)$;
    increasing on $(0, 1)$; $(1, \infty)$;
    rel. min at $x = 0$.

25. domain=$(-\infty, 0) \cup (0, \infty)$;
    c.p. at $c = 2, 6$;
    decreasing on $(-\infty, 0)$; $(0, 2)$; $(6, \infty)$;
    increasing on $(2, 6)$;
    rel. min at $x = 2$;
    rel. max at $x = 6$.

27. domain = $(-\infty, \infty)$;
    c.p. at $c = -1, 1$;
    decreasing on $(-1, 1)$;
    increasing on $(-\infty, -1)$; $(1, \infty)$;
    rel. min at $x = 1$;
    rel. max at $x = -1$

29. domain=$(-\infty, \infty)$;
    c.p. at $c = \frac{\pi}{2}, \frac{3\pi}{2}$;
    decreasing on $(0, \frac{\pi}{2})$; $(\frac{3\pi}{2}, 2\pi)$;
    increasing on $(\frac{\pi}{2}, \frac{3\pi}{2})$;
    rel. min at $x = \frac{\pi}{2}$;
    rel. max at $x = \frac{3\pi}{2}$

31. domain=$(-\infty, \infty)$;
    c.p. at $c = -1, 0, 1$;
    decreasing on $(-\infty, 0)$;
    increasing on $(0, \infty)$;
    rel. min at $x = 0$

33. domain=$[0, 2\pi]$;
    c.p. at $c = 0, \pi, 2\pi$;
    decreasing on $(0, \pi)$;
    increasing on $(\pi, 2\pi)$;
    rel. min at $x = \pi$

35. domain=$(-\infty, \infty)$;
    c.p. at $c = 0, 1$;
    decreasing on $(-\infty, 0)$; $(1, \infty)$;
    increasing on $(0, 1)$;
    rel. min at $x = 0$;
    rel. max at $x = 1$

37. domain=$[0, 2\pi]$;
    c.p. at $c = 0, \pi/2, \pi, 3\pi/2, 2\pi$;
    decreasing on $(\pi/2, 3\pi/2)$;
    increasing on $(0, \pi/2)$; $(3\pi/2, 2\pi)$;
    rel. max at $x = \pi/2$;
    rel. min at $x = 3\pi/2$

39. $c = 1/2$

Exercises 3.4

1. Answers will vary.

3. Yes; Answers will vary.

5. concave up on $(-2, 2)$;
   concave down on $(-\infty, -2)$; $(2, \infty)$;
   inflection points when $x = \pm 2$

7. concave up on $(-\infty, -1)$; $(1, \infty)$;
   concave down on $(-1, 1)$;
   inflection points when $x = \pm 1$

9. Graph and verify.

11. Graph and verify.

13. Graph and verify.

15. Graph and verify.

17. Graph and verify.

19.
    (a) Possible points of inflection: none
    (b) concave up on $(-\infty, \infty)$
    (c) min: $x = 1$
    (d) f' has no maximal or minimal value.

21.
    (a) Possible points of inflection: $x = 0$
    (b) concave down on $(-\infty, 0)$; concave up on $(0, \infty)$
    (c) max: $x = -1/\sqrt{3}$, min: $x = 1/\sqrt{3}$
    (d) f' has a minimal value at $x = 0$

23.

    (a) Possible points of inflection: $x = -2/3, 0$

    (b) concave down on $(-2/3, 0)$;
        concave up on $(-\infty, -2/3)$, $(0, \infty)$

    (c) min: $x = 1$

    (d) f' has a relative min at: $x = 0$,
        relative max at: $x = -2/3$

25.

    (a) Possible points of inflection: $x = 1$

    (b) concave up on $(-\infty, \infty)$

    (c) min: $x = 1$

    (d) f' has no relative extrema

27.

    (a) Possible points of inflection: $x = 0, \pm 1$

    (b) concave down on $(-\infty, -1)$, $(0, 1)$;
        concave up on $(-1, 0)$, $(1, \infty)$

    (c) critical values: $x = -1, 1$, no max/min

    (d) f' has a relative max at $x = 0$

29.

    (a) Possible points of inflection: $x = -2 \pm \sqrt{2}$

    (b) concave down on $(-2 - \sqrt{2}, -2 + \sqrt{2})$;
        concave up on $(-\infty, -2 - \sqrt{2})$, $(-2 + \sqrt{2}, \infty)$

    (c) max: $x = -2$, min: $x = 0$

    (d) f' has a relative max at $x = -2 - \sqrt{2}$,
        relative min at $x = -2 + \sqrt{2}$

31.

    (a) Possible points of inflection: $x = \pm 1/\sqrt{2}$

    (b) concave down on $(-1/\sqrt{2}, 1/\sqrt{2})$;
        concave up on $(-\infty, -1/\sqrt{2})$, $(1/\sqrt{2}, \infty)$

    (c) max: $x = 0$

    (d) f' has a relative max at $x = -1/\sqrt{2}$,
        a relative min at $x = 1/\sqrt{2}$

33.

    (a) Possible points of inflection: $x = \pi/6, 5\pi/6, 3\pi/2$

    (b) concave down on $(0, \pi/6)$, $(5\pi/6, 2\pi)$;
        concave up on $(\pi/6, 5\pi/6)$

    (c) max: $x = 3\pi/2$, min: $x = 3\pi/2$

    (d) f' has a relative max at $x = 5\pi/6$,
        f' has a relative min at $x = \pi/6$

35.

    (a) Possible points of inflection: $x = \pm 1/\sqrt{3}$

    (b) concave down on $(-1/\sqrt{3}, 1/\sqrt{3})$;
        concave up on $(-\infty, -1/\sqrt{3})$; $(1/\sqrt{3}, \infty)$

    (c) min: $x = \pm 1$, max: $x = 0$

    (d) f' has a relative max at $x = -1/\sqrt{3}$,
        f' has a relative min at $x = 1/\sqrt{3}$

37.

    (a) Possible points of inflection: $x = 81/256$

    (b) concave up on $(0, 81/256)$;
        concave down on $(81/256, \infty)$

    (c) min: $x = \frac{1}{16}$

    (d) f' has a relative max at $x = 81/256$

Exercises 3.5

1. Answers will vary.

3. T

5. concave up on $(-\infty, -1)$; $(1, \infty)$
   concave down on $(-1, 1)$
   inflection points when $x = \pm 1$
   increasing on $(-2, 0)$; $(2, \infty)$
   decreasing on $(-\infty, -2)$; $(0, 2)$
   relative maximum when $x = 0$
   relative minima when $x = \pm 2$

7. A good sketch will include the x and y intercepts and draw the appropriate line.

9. Use technology to verify sketch.

11. Use technology to verify sketch.

13. Use technology to verify sketch.

15. Use technology to verify sketch.

17. Use technology to verify sketch.

19. Use technology to verify sketch.

21. Use technology to verify sketch.

23. Use technology to verify sketch.

25. Use technology to verify sketch.

27. Use technology to verify sketch.

29. Use technology to verify sketch.

31. Use technology to verify sketch.

33. Use technology to verify sketch.

35. Use technology to verify sketch.

37. Use technology to verify sketch.

39. Use technology to verify sketch.

41. Use technology to verify sketch.

43. Use technology to verify sketch.

45. Use technology to verify sketch.

47. various possibilities

49. various possibilities

51. various possibilities

53. Critical point: $x = 0$ Points of inflection: $\pm b/\sqrt{3}$

55. Critical point: $x = (a + b)/2$ Points of inflection: none

Chapter 4

Exercises 4.1

1. T

3. 3 ft/min

5.

   (a) $5/(2\pi) \approx 0.796$cm/s

   (b) $1/(40\pi) \approx 0.00796$ cm/s

   (c) $1/(4000\pi) \approx 0.0000796$ cm/s

7.

   (a) 64.44 mph

   (b) 78.89 mph

9. Due to the height of the plane, the gun does not have to rotate very fast.

   (a) 0.073 rad/s

   (b) 3.66 rad/s (about 1/2 revolution/sec)

   (c) In the limit, rate goes to 7.33 rad/s (more than 1 revolution/sec)

11.

   (a) 30.59 ft/min

   (b) 36.1 ft/min

   (c) 301 ft/min

   (d) The boat no longer floats as usual, but is being pulled up by the winch (assuming it has the power to do so).

13.

   (a) 0.63 ft/sec

   (b) 1.6 ft/sec

  About 52 ft.

15.

   (a) The balloon is 105ft in the air.

   (b) The balloon is rising at a rate of 17.45ft/min. (Hint: convert all angles to radians.)

Exercises 4.2

1. T

3. 2500; the two numbers are each 50.

5. There is no maximum sum; the fundamental equation has only 1 critical value that corresponds to a minimum.

7. Area = 1/4, with sides of length $1/\sqrt{2}$.

9. The radius should be about 3.84cm and the height should be $2r = 7.67$cm. No, this is not the size of the standard can.

11. The height and width should be 18 and the length should be 36, giving a volume of $11,664$in^3.

13. $5 - 10/\sqrt{39} \approx 3.4$ miles should be run underground, giving a minimum cost of $374,899.96.

15. The dog should run about 19 feet along the shore before starting to swim.

17. The largest area is 2 formed by a square with sides of length $\sqrt{2}$.

19. A length of 2 in and height of 2.5 will give a cost of 60 ¢.

Exercises 4.3

1. T

3. F

5. Answers will vary.

7. Use $y = x^2$; $dy = 2x \cdot dx$ with $x = 6$ and $dx = -0.07$. Thus $dy = -0.84$; knowing $6^2 = 36$, we have $5.93^2 \approx 35.16$.

9. Use $y = x^3$; $dy = 3x^2 \cdot dx$ with $x = 7$ and $dx = -0.2$. Thus $dy = -29.4$; knowing $7^3 = 343$, we have $6.8^3 \approx 313.6$.

11. Use $y = \sqrt{x}$; $dy = 1/(2\sqrt{x}) \cdot dx$ with $x = 25$ and $dx = -1$. Thus $dy = -0.1$; knowing $\sqrt{25} = 5$, we have $\sqrt{24} \approx 4.9$.

13. Use $y = \sqrt[3]{x}$; $dy = 1/(3\sqrt[3]{x^2}) \cdot dx$ with $x = 8$ and $dx = 0.5$. Thus $dy = 1/24 \approx 1/25 = 0.04$; knowing $\sqrt[3]{8} = 2$, we have $\sqrt[3]{8.5} \approx 2.04$.

15. Use $y = \cos x$; $dy = -\sin x \cdot dx$ with $x = \pi/2 \approx 1.57$ and $dx \approx -0.07$. Thus $dy = 0.07$; knowing $\cos \pi/2 = 0$, we have $\cos 1.5 \approx 0.07$.

17. $dy = (2x + 3)dx$

19. $dy = \frac{-2}{4x^3}dx$

21. $dy = \left(2xe^{3x} + 3x^2 e^{3x}\right)dx$

23. $dy = \frac{2(\tan x+1) - 2x\sec^2 x}{(\tan x+1)^2}dx$

25. $dy = (e^x \sin x + e^x \cos x)dx$

27. $dy = \frac{1}{(x+2)^2}dx$

29. $dy = (\ln x)dx$

31. $1 - 6x$

33. $1 - \frac{x}{2}$

35. $2^{2/3} + \frac{x}{2^{4/3}}$

37. $dV = \pm 0.157$

39. $\pm 15\pi/8 \approx \pm 5.89$in^2

41.

   (a) 297.8 feet

   (b) ± 62.3 ft

   (c) $\pm 20.9\%$

43.

   (a) 298.9 feet

   (b) ± 8.67 ft

   (c) $\pm 2.9\%$

45. 1%

Exercises 4.4

1. F

3. $x_0 = 1.5$, $x_1 = 1.5709148$, $x_2 = 1.5707963$, $x_3 = 1.5707963$, $x_4 = 1.5707963$, $x_5 = 1.5707963$

5. $x_0 = 0$, $x_1 = 2$, $x_2 = 1.2$, $x_3 = 1.0117647$, $x_4 = 1.0000458$, $x_5 = 1$

7. $x_0 = 2$, $x_1 = 0.6137056389$, $x_2 = 0.9133412072$, $x_3 = 0.9961317034$, $x_4 = 0.9999925085$, $x_5 = 1$

9. roots are: $x = -3.714$, $x = -0.857$, $x = 1$ and $x = 1.571$

11. roots are: $x = -2.165$, $x = 0$, $x = 0.525$ and $x = 1.813$

13. $x = -0.637$, $x = 1.410$

15. $x = \pm 4.493$, $x = 0$

17. The approximations alternate between $x = 1$, $x = 2$ and $x = 3$.

19. $f(x) = x^2 - 24$ and $x_0 = 5$ yield $x_1 = \frac{49}{10} = 4.9$ and $x_2 = \frac{4801/980}{\approx} 4.898980$.

21. $f(x) = x^3 - 8.5$ and $x_0 = 2$ yield $x_1 = \frac{49}{24} \approx 2.0416667$ and $x_2 \approx 2.0408279$.

Chapter 5

Exercises 5.1

1. Answers will vary.

3. Answers will vary.

5. Answers will vary.

7. velocity

9. $x^9/9 + C$

11. $t + C$

13. $-3/(t) + C$

15. $\tan\theta + C$

17. $\sec x - \csc x + C$

19. $4/3 t^3 + 6t^2 + 9t + C$

21. $x^6/6 + C$

23. $-x^{-3} + C$

25. $\dfrac{2}{9}x^{9/2} + C$

27. $5x + \frac{2}{9}x^3 + \frac{3}{16}x^4 + C$

29. $\frac{2}{3}u^3 + \frac{9}{2}u^2 + 4u + C$

31. $2\sqrt{x} + x + \frac{2}{3}x\sqrt{x} + C$

33. $\theta + \tan\theta + C$

35. $-\cot t - t + C$

37. $8\sqrt{u} + 4u\sqrt{u} + C$

39. $6t^{1/3} + \frac{3}{4}t^{4/3} + C$

41.
 (a) $x > 0$
 (b) $1/x$
 (c) $x < 0$
 (d) $1/x$
 (e) $\ln|x| + C$. Explanations will vary.

43. $5e^x + 5$

45. $\tan x + 4$

47. $\frac{7x^3}{6} - \frac{9x}{2} + \frac{40}{3}$

49. $\theta - \sin(\theta) - \pi + 4$

51. $x^{-2} + 1$

53. $s(t) = 2t^{3/2}$.

55.
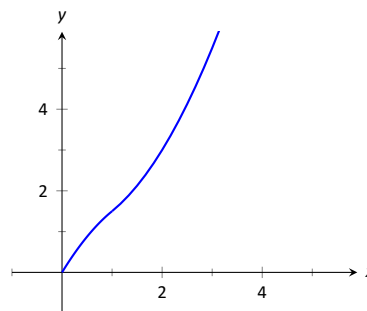
Other antiderivatives are vertical shifts of this one.

57. No answer provided.

Exercises 5.2

1. Answers will vary.

3. 0

5.
 (a) 3
 (b) 4
 (c) 3
 (d) 0
 (e) -4
 (f) 9

7.
 (a) 4
 (b) 2
 (c) 4
 (d) 2
 (e) 1
 (f) 2

9.
 (a) π
 (b) π
 (c) 2π
 (d) 10π

11.
 (a) $4/\pi$
 (b) $-4/\pi$
 (c) 0
 (d) $2/\pi$
 (e) $4/\pi$
 (f) $8/\pi$

13.
 (a) $40/3$
 (b) $26/3$
 (c) $8/3$
 (d) $38/3$

15.

    (a) 3ft/s

    (b) 9.5ft

    (c) 9.5ft

17.

    (a) 96ft/s

    (b) 6 seconds

    (c) 6 seconds

    (d) Never; the maximum height is 208ft.

19. 5

21. Answers can vary; one solution is $a = -2$, $b = 7$

23. -7

25. Answers can vary; one solution is $a = -11$, $b = 18$

27. $-\cos x - \sin x + \tan x + C$

29. $\ln |x| + \csc x + C$

Exercises 5.3

1. limits

3. Rectangles.

5. $2^2 + 3^2 + 4^2 = 29$

7. $0 - 1 + 0 + 1 + 0 = 0$

9. $-1 + 2 - 3 + 4 - 5 + 6 = 3$

11. $1 + 1 + 1 + 1 + 1 + 1 = 6$

13. Answers may vary; $\sum_{i=0}^{8}(i^2 - 1)$

15. Answers may vary; $\sum_{i=0}^{4}(-1)^i e^i$

17. 1045

19. -8525

21. 5050

23. 155

25. 24

27. $\int_{0}^{\pi} \frac{\sin x}{1 + x}\,dx$

29. $\int_{2}^{7} 5x^3 - 4x + 7\,dx$

31. $\lim\limits_{n \to \infty} \frac{3}{n} \sum\limits_{i=1}^{n} 4 - 2\left(2 + \frac{3i}{n}\right)$

33. $\lim\limits_{n \to \infty} \frac{\pi}{n} \sum\limits_{i=1}^{n} \frac{\sin^3(-\pi/2 + \pi i/n)}{2 + \cos(-\pi/2 + \pi i/n)}$

35. $59/8$

37. 8.16986

39. $496/315 \approx 1.5746$

41.

    (a) Exact expressions will vary; $2 + 4/n^2$.

    (b) 51/25, 5001/2500, 500001/250000

    (c) 2

43.

    (a) Exact expressions will vary; $20/3 - 96/(3n) + 64/(3n^2)$.

    (b) 92/25, 3968/625, 103667/15625

    (c) $20/3$

45.

    (a) Exact expressions will vary; $-1/12(1 - 1/n^2)$.

    (b) $-33/400$, $-3333/40000$, $-333333/4000000$

    (c) $-1/12$

47.

    (a) $(5\,\text{s})((0+6+14+23+30+36)\,\text{mph}) = 545\frac{\text{mi}\,\text{s}}{\text{hr}} \times \frac{1\,\text{hr}}{3600\,\text{s}} \times 5280\,\text{ft}1\,\text{mi} = 799$ ft

    (b) $(5\,\text{s})((6+14+23+30+36+40)\,\text{mph}) = 585\frac{\text{mi}\,\text{s}}{\text{hr}} \times \frac{1\,\text{hr}}{3600\,\text{s}} \times 5280\,\text{ft}1\,\text{mi} = 858$ ft

49. Let f and M be as given.

$$\int_{a}^{b} f(x)\,dx = \lim_{n \to \infty} \sum_{i=1}^{n} f(c_i)\Delta x \quad \text{T36.2}$$

$$\le \lim_{n \to \infty} \sum_{i=1}^{n} M\Delta x$$

$$= \int_{a}^{b} M\,dx \quad \text{T36.2}$$

$$= M(b - a)$$

51. $F(x) = 7 \ln |x| + 14$

53. $G(t) = \sin t - \cos t - 78$

Exercises 5.4

1. Answers will vary.

3. T

5. 20

7. 0

9. 1

11. $23/2$

13. $e^3 - e$

15. 4

17. $\ln 2$

19. $1/4$

21. 15

23. 2

25. $-\frac{9}{2}$

27. $\frac{6\pi}{7}$

29. $\sqrt{2}$

31. 36

33. $\frac{69}{4}$

35. Explanations will vary. A sketch will help.

37. $c = \pm 2/\sqrt{3}$

39. $c = 64/9 \approx 7.1$

41. $2/pi$

43. $16/3$

45. $1/(e-1)$

47. (a) 400ft; (b) 850ft

49. (a) 128/5ft; (b) same

51. 50ft/s

53. 0ft/s

55. $F'(x) = -3x^{11}$

57. $F'(x) = e^x \sin(e^x) - \frac{1}{x}\sin(\ln x)$

59. $F'(x) = [\cos^3(\sin x) + 3\tan^3(\sin x)]\cos x$

61. $F'(x) = -2\tan x \sec^2 x[\ln(\tan^2 x) + e^{\tan^4 x - 7}]$

Exercises 5.5

1. Chain Rule.

3. $\frac{1}{8}(x^3 - 5)^8 + C$

5. $\frac{1}{18}(x^2 + 1)^9 + C$

7. $\frac{1}{2}\ln|2x + 7| + C$

9. $\frac{2}{3}(x+3)^{3/2} - 6(x+3)^{1/2} + C = \frac{2}{3}(x-6)\sqrt{x+3} + C$

11. $2e^{\sqrt{x}} + C$

13. $-\frac{1}{2x^2} - \frac{1}{x} + C$

15. $\frac{\sin^3(x)}{3} + C$

17. $-\tan(4-x) + C$

19. $\frac{\tan^3(x)}{3} + C$

21. The key is to rewrite $\cot x$ as $\cos x/\sin x$, and let $u = \sin x$.

23. $\frac{1}{3}e^{3x-1} + C$

25. $\frac{1}{2}e^{(x-1)^2} + C$

27. $\frac{e^{-3x}}{3} - e^{-x} + C$

29. $\frac{(\ln x)^3}{3} + C$

31. $\frac{1}{2}\ln\left(\ln\left(x^2\right)\right) + C$

33. $\frac{1}{45}(5x^3 + 5x^2 + 2)^9 + C$

35. $-\frac{1}{3}\cot\left(x^3 + 1\right) + C$

37. $\ln|x-5| + C$

39. $\ln\left|x^2 + 7x + 3\right| + C$

41. $3\sqrt{x^2 - 2x - 6} + C$

43. $2\sin\sqrt{x} + C$

45. $\frac{1}{10}(2x+3)^{5/2} - \frac{1}{2}(2x+3)^{3/2} + C$

47. $\frac{1}{2}(x^2+1)^2 - 2(x^2+1) + \ln(x^2+1) + C$

49. $-3\cos(\frac{x}{3}) + C$

51. $\frac{2}{3}\sin(x^{3/2} + 1) + C$

53. 352/15

55. 1/5

57. $e - 1$

59. $\ln\left(\frac{4}{1+e}\right)$

61. $\frac{2}{3}$

Chapter 6

Exercises 6.1

1. T

3. Answers will vary.

5. $16/3$

7. π

9. $2\sqrt{2}$

11. $4/3$

13. 8

15. $37/12$

17. $\frac{8}{3}$

19. 9

21. 1

23. $9/2$

25. $1/12(9 - 2\sqrt{2}) \approx 0.514$

27. 5

29. $133/20$

Exercises 6.2

1. T

3. Recall that "dx" does not just "sit there;" it is multiplied by $A(x)$ and represents the thickness of a small slice of the solid. Therefore dx has units of in, giving $A(x)\,dx$ the units of in^3.

5. $175\pi/3$ units^3

7. $\dfrac{768\pi}{7}$

9. $35\pi/3$ units^3

11. $\dfrac{96\pi}{5}$

13.
   (a) $512\pi/15$
   (b) $256\pi/5$
   (c) $832\pi/15$
   (d) $128\pi/3$

15.
   (a) $104\pi/15$
   (b) $64\pi/15$
   (c) $32\pi/5$

17.

    (a) $\frac{\pi^2}{8} + \frac{\pi}{4}$

    (b) $\frac{3\pi^2}{8} + \frac{\pi}{4} - \pi\sqrt{2}$

    (c) $\frac{\pi^2}{8} + \frac{\pi}{4} + \pi\sqrt{2}$

19. The cross–sections of this cone are the same as the cone in item 18. Thus they have the same volume of $250\pi/3$ units^3.

21. Orient the solid so that the *x*-axis is parallel to long side of the base. All cross–sections are trapezoids (at the far left, the trapezoid is a square; at the far right, the trapezoid has a top length of 0, making it a triangle). The area of the trapezoid at *x* is $A(x) = 1/2(-1/2x + 5 + 5)(5) = -5/4x + 25$. The volume is 187.5 units^3.

Exercises 6.3

1. T

3. F

5. $9\pi/2$ units^3

7. $\dfrac{96\pi}{5}$

9. $48\pi\sqrt{3}/5$ units^3

11. $\dfrac{768\pi}{7}$

13.

    (a) $4\pi/5$

    (b) $8\pi/15$

    (c) $\pi/2$

    (d) $5\pi/6$

15.

    (a) $4\pi/3$

    (b) $\pi/3$

    (c) $4\pi/3$

    (d) $2\pi/3$

17.

    (a) $16\pi/3$

    (b) $8\pi/3$

    (c) 8π

    (d) 8π

19. (a) Disk: $\pi \int_1^2 (\sqrt[3]{y-1})^2\, dy = \frac{3\pi}{5}$
Shell: $2\pi \int_0^1 x(2 - (x^3 + 1))\, dx = \frac{3\pi}{5}$
(b) Disk: $\pi \int_0^1 \left[2 - (x^3 + 1)\right]^2\, dx = \frac{9\pi}{14}$
Shell: $2\pi \int_1^2 (2 - y)\sqrt[3]{y-1}\, dy = \frac{9\pi}{14}$.

Exercises 6.4

1. In SI units, it is one Joule, i.e., one Newton–meter, or kg·m/s^2·m. In Imperial Units, it is ft–lb.

3. Smaller.

5.

    (a) 2450 J

    (b) 1568 J

7. 735 J

9. 11,100 ft–lb

11. 125 ft–lb

13. 12.5 ft–lb

15. 0.2625 = 21/80 J

17. 45 ft–lb

19. 953, 284 J

21. 192,767 ft–lb. Note that the tank is oriented horizontally. Let the origin be the center of one of the circular ends of the tank. Since the radius is 3.75 ft, the fluid is being pumped to $y = 4.75$; thus the distance the gas travels is $h(y) = 4.75 - y$. A differential element of water is a rectangle, with length 20 and width $2\sqrt{3.75^2 - y^2}$. Thus the force required to move that slab of gas is $F(y) = 40 \cdot 45.93 \cdot \sqrt{3.75^2 - y^2}dy$. Total work is $\int_{-3.75}^{3.75} 40 \cdot 45.93 \cdot (4.75 - y)\sqrt{3.75^2 - y^2}\, dy$. This can be evaluated without actual integration; split the integral into $\int_{-3.75}^{3.75} 40 \cdot 45.93 \cdot (4.75)\sqrt{3.75^2 - y^2}\, dy + \int_{-3.75}^{3.75} 40 \cdot 45.93 \cdot (-y)\sqrt{3.75^2 - y^2}\, dy$. The first integral can be evaluated as measuring half the area of a circle; the latter integral can be shown to be 0 without much difficulty. (Use substitution and realize the bounds are both 0.)

23.

    (a) approx. 577,000 J

    (b) approx. 399,000 J

    (c) approx. 110,000 J (By volume, half of the water is between the base of the cone and a height of 3.9685 m. If one rounds this to 4 m, the work is approx 104,000 J.)

25. 617,400 J

Exercises 6.5

1. Answers will vary.

3. 499.2 lb

5. 6739.2 lb

7. 3920.7 lb

9. 2496 lb

11. 602.59 lb

13.

    (a) 2340 lb

    (b) 5625 lb

15.

    (a) 1597.44 lb

    (b) 3840 lb

17.

    (a) 56.42 lb

    (b) 135.62 lb

19. 5.1 ft

Chapter 7

1. F

3. The point $(10, 1)$ lies on the graph of $y = f^{-1}(x)$ (assuming f is invertible).

5. Compose $f(g(x))$ and $g(f(x))$ to confirm that each equals x.

7. Compose $f(g(x))$ and $g(f(x))$ to confirm that each equals x.

9. $[-4, 0]$ or $[0, 4]$

11. $(-\infty, 3]$ or $[3, \infty)$

13. 0

15. $-1/5$

17. $1/\sqrt{2}$

19. $5/4$

Exercises 7.2

1. The point $(10, 1)$ lies on the graph of $y = f^{-1}(x)$ (assuming f is invertible) and $(f^{-1})'(10) = 1/5$.

3. $(f^{-1})'(7) = \frac{1}{f'(3)} = 1/4$

5. $(f^{-1})'(8) = \frac{1}{f'(1)} = 1/6$

7. $h'(t) = \frac{2}{\sqrt{1-4t^2}}$

9. $g'(x) = \frac{2}{1+4x^2}$

11. $g'(t) = \cos^{-1}(t)\cos(t) - \frac{\sin(t)}{\sqrt{1-t^2}}$

13. $g'(x) = \frac{1}{\sqrt{x}(2x+2)}$

15.

   (a) $f(x) = x$, so $f'(x) = 1$

   (b) $f'(x) = \cos(\sin^{-1}x)\frac{1}{\sqrt{1-x^2}} = 1$.

17.

   (a) $f(x) = x$, so $f'(x) = 1$

   (b) $f'(x) = \frac{1}{1+\tan^2 x}\sec^2 x = 1$

19. $y = \sqrt{2}(x - \sqrt{2}/2) + \pi/4$

21. $\sqrt{91} \approx 9.54$ feet

23. $2\pi/9$

25. $\frac{1}{8}\tan^{-1}(x^4/2) + C$

27. $\sin^{-1}\left(\frac{x-1}{2}\right) + C$

Exercises 7.3

1. $(-\infty, \infty)$

3. $(-\infty, 0) \cup (0, \infty)$

5. $f'(t) = 3t^2 e^{t^3 - 1}$

7. $f'(x) = \dfrac{1 - x\ln 5 \ln x}{x 5^x \ln 5}$

9. $f'(x) = 1$

11. $h'(r) = \dfrac{3^r \ln 3}{1 + 3^{2r}}$

13. $f'(t) = \frac{e^t}{t} + \ln te^t$

15. $\dfrac{24}{\ln 5}$

17. $\frac{1}{2}e^{x^2 - 1} + C$

19. $\frac{1}{2}\sin^2(e^x) + C$

21. $\dfrac{\ln 245}{\ln 3} - 1$

23. $\ln\left|\tan^{-1} x\right| + C$

25. $y' = (1 + x)^{1/x}\left(\frac{1}{x(x+1)} - \frac{\ln(1+x)}{x^2}\right)$
Tangent line: $y = (1 - 2\ln 2)(x - 1) + 2$

27. $y' = \frac{x^x}{x+1}\left(\ln x + 1 - \frac{1}{x+1}\right)$
Tangent line: $y = (1/4)(x - 1) + 1/2$

29. $y' = \frac{x+1}{x+2}\left(\frac{1}{x+1} - \frac{1}{x+2}\right)$
Tangent line: $y = 1/9(x - 1) + 2/3$

31. $y' = x^{e^x - 1}e^x(1 + x\ln x)$
Tangent line: $y = ex - e + 1$

Exercises 7.4

1. Because $\cosh x$ is always positive.

3. $\coth^2 x - \operatorname{csch}^2 x = \left(\dfrac{e^x + e^{-x}}{e^x - e^{-x}}\right)^2 - \left(\dfrac{2}{e^x - e^{-x}}\right)^2$

   $= \dfrac{(e^{2x} + 2 + e^{-2x}) - (4)}{e^{2x} - 2 + e^{-2x}}$

   $= \dfrac{e^{2x} - 2 + e^{-2x}}{e^{2x} - 2 + e^{-2x}}$

   $= 1$

5. $\cosh^2 x = \left(\dfrac{e^x + e^{-x}}{2}\right)^2$

   $= \dfrac{e^{2x} + 2 + e^{-2x}}{4}$

   $= \dfrac{1}{2}\dfrac{(e^{2x} + e^{-2x}) + 2}{2}$

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

   $= \dfrac{\cosh 2x + 1}{2}.$

7. $\dfrac{d}{dx}[\operatorname{sech} x] = \dfrac{d}{dx}\left[\dfrac{2}{e^x + e^{-x}}\right]$

   $= \dfrac{-2(e^x - e^{-x})}{(e^x + e^{-x})^2}$

   $= -\dfrac{2(e^x - e^{-x})}{(e^x + e^{-x})(e^x + e^{-x})}$

   $= -\dfrac{2}{e^x + e^{-x}} \cdot \dfrac{e^x - e^{-x}}{e^x + e^{-x}}$

   $= -\operatorname{sech} x \tanh x$

9. $\displaystyle\int \tanh x\, dx = \int \dfrac{\sinh x}{\cosh x}\, dx$

   Let $u = \cosh x$; $du = (\sinh x)dx$

   $= \displaystyle\int \dfrac{1}{u}\, du$

   $= \ln|u| + C$

   $= \ln(\cosh x) + C.$

11. $2\sinh 2x$

13. $\coth x$

15. $x \cosh x$

17. $\dfrac{3}{\sqrt{9x^2+1}}$

19. $\dfrac{1}{1-(x+5)^2}$

21. $\sec x$

23. $y = 3/4(x - \ln 2) + 5/4$

25. $y = x$

27. $\frac{1}{2}\ln(\cosh(2x)) + C$

29. $\frac{1}{2}\sinh^2 x + C$ or $1/2\cosh^2 x + C$

31. $\cosh^{-1}(x^2/2) + C = \ln(x^2 + \sqrt{x^4 - 4}) + C$

33. $\tan^{-1}(e^x) + C$

35. 0

Exercises 7.5

1. $0/0, \infty/\infty, 0 \cdot \infty, \infty - \infty, 0^0, 1^\infty, \infty^0$

3. F

5. derivatives; limits

7. Answers will vary.

9. $-5/3$

11. $-\sqrt{2}/2$

13. 0

15. a/b

17. $1/2$

19. 0

21. ∞

23. 0

25. -2

27. 0

29. 0

31. ∞

33. ∞

35. 0

37. 1

39. 1

41. 1

43. 1

45. 1

47. 2

49. $-\infty$

51. 0

Chapter 8

1. T

3. $\sin x - x \cos x + C$

5. $-x^2 \cos x + 2x \sin x + 2 \cos x + C$

7. $1/2 e^{x^2} + C$

9. $-\frac{1}{2}xe^{-2x} - \frac{e^{-2x}}{4} + C$

11. $1/5 e^{2x}(\sin x + 2 \cos x) + C$

13. $1/10 e^{5x}(\sin(5x) + \cos(5x)) + C$

15. $\sqrt{1 - x^2} + x \sin^{-1}(x) + C$

17. $\frac{1}{2}x^2 \tan^{-1}(x) - \frac{x}{2} + \frac{1}{2}\tan^{-1}(x) + C$

19. $\frac{1}{2}x^2 \ln|x| - \frac{x^2}{4} + C$

21. $-\frac{x^2}{4} + \frac{1}{2}x^2 \ln|x - 1| - \frac{x}{2} - \frac{1}{2}\ln|x - 1| + C$

23. $\frac{1}{3}x^3 \ln|x| - \frac{x^3}{9} + C$

25. $2x + x(\ln|x + 1|) + (\ln|x + 1|)^2 - 2x \ln|x + 1| - 2\ln|x + 1| + 2 + C$

27. $\ln|\sin(x)| - x \cot(x) + C$

29. $\frac{1}{3}(x^2 - 2)^{3/2} + C$

31. $x \sec x - \ln|\sec x + \tan x| + C$

33. $x \sinh(x) - \cosh(x) + C$

35. $x \sinh^{-1} x - \sqrt{x^2 + 1} + C$

37. $1/2x(\sin(\ln x) - \cos(\ln x)) + C$

39. $\frac{1}{2}x \ln|x| - \frac{x}{2} + C$

41. $1/2x^2 + C$

43. π

45. 0

47. $1/2$

49. $\dfrac{3}{4e^2} - \dfrac{5}{4e^4}$

51. $1/5 \left(e^\pi + e^{-\pi}\right)$

Exercises 8.2

1. F

3. F

5. $\frac{3}{4}x + \frac{1}{2}\sin 2x + C$

7. $\frac{1}{6}\cos^6 x - \frac{1}{4}\cos^4 x + C$

9. $\frac{1}{2}\cos^2 x - \ln|\cos x| + C$

11. $\left(\frac{2}{7}\cos^3 x - \frac{2}{3}\cos x\right)\sqrt{\cos x} + C$

13. $\frac{1}{2}\left(\frac{1}{4}\sin(4x) - \frac{1}{10}\sin(10x)\right) + C$

15. $\frac{1}{2}\left(\sin(x) + \frac{1}{3}\sin(3x)\right) + C$

17. $\tan x - x + C$

19. $\dfrac{\tan^6(x)}{6} + \dfrac{\tan^4(x)}{4} + C$

21. $\dfrac{\sec^5(x)}{5} - \dfrac{\sec^3(x)}{3} + C$

23. $\frac{1}{3}\tan^3 x - \tan x + x + C$

25. $\frac{1}{2}(\sec x \tan x - \ln|\sec x + \tan x|) + C$

27. $\ln|\csc x - \cot x| + C$

29. $-\frac{1}{2}\cot^2 x + \ln|\csc x| + C$

31. $\frac{2}{5}$

33. $32/315$

35. $2/3$

37. $16/15$

39. $\frac{1}{2} - \ln\sqrt{2}$

Exercises 8.3

1. backwards

3.
   (a) $\tan^2\theta + 1 = \sec^2\theta$
   (b) $9\sec^2\theta$.

5. $\frac{1}{2}\left(x\sqrt{x^2+1} + \ln\left|\sqrt{x^2+1}+x\right|\right) + C$

7. $x\sqrt{x^2+1/4} + \frac{1}{4}\ln\left|2\sqrt{x^2+1/4}+2x\right| + C$
   $= \frac{1}{2}x\sqrt{4x^2+1} + \frac{1}{4}\ln\left|\sqrt{4x^2+1}+2x\right| + C$

9. $4\left(\frac{1}{2}x\sqrt{x^2-1/16} - \frac{1}{32}\ln\left|4x+4\sqrt{x^2-1/16}\right|\right) + C = \frac{1}{2}x\sqrt{16x^2-1} - \frac{1}{8}\ln\left|4x+\sqrt{16x^2-1}\right| + C$

11. $3\sin^{-1}\left(\frac{x}{\sqrt{7}}\right) + C$ (Trig. Subst. is not needed)

13. $\sqrt{x^2-11} - \sqrt{11}\sec^{-1}(x/\sqrt{11}) + C$

15. $-\frac{1}{\sqrt{x^2+9}} + C$ (Trig. Subst. is not needed)

17. $\frac{1}{18}\frac{x+2}{x^2+4x+13} + \frac{1}{54}\tan^{-1}\left(\frac{x+2}{2}\right) + C$

19. $\frac{1}{7}\left(-\frac{\sqrt{5-x^2}}{x} - \sin^{-1}(x/\sqrt{5})\right) + C$

21. $\pi/2$

23. $2\sqrt{2} + 2\ln(1+\sqrt{2})$

25. $9\sin^{-1}(1/3) + \sqrt{8}$ Note: the new bounds of integration are $\sin^{-1}(-1/3) < \theta < \sin^{-1}(1/3)$. The final answer comes with recognizing that $\sin^{-1}(-1/3) = -\sin^{-1}(1/3)$ and that $\cos\left(\sin^{-1}(1/3)\right) = \cos\left(\sin^{-1}(-1/3)\right) = \sqrt{8}/3$.

Exercises 8.4

1. rational

3. $\frac{A}{x} + \frac{B}{x-3}$

5. $\frac{A}{x-\sqrt{7}} + \frac{B}{x+\sqrt{7}}$

7. $3\ln|x-2| + 4\ln|x+5| + C$

9. $\frac{1}{3}(\ln|x+2| - \ln|x-2|) + C$

11. $-\frac{4}{x+8} - 3\ln|x+8| + C$

13. $-\ln|2x-3| + 5\ln|x-1| + 2\ln|x+3| + C$

15. $x + \ln|x-1| - \ln|x+2| + C$

17. $2x + C$

19. $\frac{1}{x} + \frac{1}{2}\ln\left|\frac{x-1}{x+1}\right| + C$

21. $\ln|3x^2+5x-1| + 2\ln|x+1| + C$

23. $\ln|x| - \frac{1}{2}\ln(x^2+1) - \tan^{-1}x - \frac{1}{2(x^2+1)} + C$

25. $\frac{1}{2}\left(3\ln|x^2+2x+17| - 4\ln|x-7| + \tan^{-1}\left(\frac{x+1}{4}\right)\right) + C$

27. $-\frac{1}{4}\ln(x^2+3) + \frac{1}{4}\ln(x^2+1) + C = \frac{1}{4}\ln\frac{x^2+1}{x^2+3} + C$

29. $3\left(\ln|x^2-2x+11| + \ln|x-9|\right) + 3\sqrt{\frac{2}{5}}\tan^{-1}\left(\frac{x-1}{\sqrt{10}}\right) + C$

31. $\frac{1}{16}\tan^{-1}(x/2) + \frac{1}{32}\ln|x-2| + \frac{1}{32}\ln|x+2| + C$

33. $\ln(2000/243) \approx 2.108$

35. $-\pi/4 + \tan^{-1}3 - \ln(11/9) \approx 0.263$

Exercises 8.5

1. $\frac{x^2}{2}\sin^{-1}x - \frac{1}{4}\sin^{-1}x + \frac{x}{4}\sqrt{1-x^2} + C$

3. $18\ln|x-2| - 9\ln|x-1| - 5\ln|x-3| + C$

5. $\frac{x}{25}\sqrt{x^2+25} + C$

7. $2\ln|x-1| - \ln|x| - \frac{x}{(x-1)^2} + C$

9. $\frac{1}{2}e^{x^2}(x^2-1) + C$

11. $\frac{1}{13}e^{2x}(2\sin 3x - 3\cos 3x) + C$

13. $-\sqrt{4-x^2} + C$

15. $2\tan^{-1}\sqrt{x} + C$

17. $\frac{1}{27}\left[6x\sin 3x - (9x^2-2)\cos 3x\right] + C$

19. $\frac{2}{3}(1+e^x)^{3/2} + C$

21. $\frac{1}{3}\tan^3 x + C$

23. $-\frac{1}{4}(8-x^3)^{4/3} + C$

25. $\frac{1}{10}(3-2x)^{5/2} - \frac{1}{2}(3-2x)^{3/2} + C$

27. $\frac{2}{5}x^{5/2} - \frac{8}{3}x^{3/2} + 6x^{1/2}C$

29. $\frac{11}{2}\ln|x+5| - \frac{15}{2}\ln|x+7| + C$

31. $e^{\tan x} + C$

33. $-\frac{1}{5}\cot^5 x - \frac{1}{3}\cot^3 x - \cot x - x + C$

35. $\frac{1}{3}x^3 - \frac{1}{4}\tanh 4x + C$

37. $3\sin^{-1}\left(\frac{x+5}{6}\right) + C$

39. $-\frac{1}{7}\cos 7x + C$

41. $x^3\sin x + 3x^2\cos x - 6x\sin x - 6\cos x + \sin x + C$

43. $24x - \frac{10}{3}\ln|\sin 3x| - \frac{1}{3}\cot 3x + C$

45. $-2\sqrt{1+\cos x} + C$

47. $\ln(x^2+4) - \frac{3}{2}\tan^{-1}\frac{x}{2} + \frac{7}{\sqrt{5}}\tan^{-1}\frac{x}{\sqrt{5}} + C$

49. $\frac{2}{5}x^{5/2}\ln x - \frac{4}{25}x^{5/2} + C$

51. $\frac{e^x}{x+1} + C$

Exercises 8.6

1. The interval of integration is finite, and the integrand is continuous on that interval.

3. converges; could also state < 10.

5. $p > 1$

7. $e^5/2$

9. $1/3$

11. $1/\ln 2$

13. diverges

15. 1

17. diverges

19. diverges

21. $2\sqrt{3}$

23. diverges

25. diverges

27. 1

29. 0

31. $-1/4$

33. -1

35. diverges

37. $1/2$

39. converges; Limit Comparison Test with $1/x^{3/2}$.

41. converges; Direct Comparison Test with xe^{-x}.

43. converges; Direct Comparison Test with xe^{-x}.

45. diverges; Direct Comparison Test with $x/(x^2 + \cos x)$.

47. converges; Limit Comparison Test with $1/e^x$.

Exercises 8.7

1. F

3. They are superseded by the Trapezoidal Rule; it takes an equal amount of work and is generally more accurate.

5.
 (a) $3/4$
 (b) $2/3$
 (c) $2/3$

7.
 (a) $\frac{1}{4}(1 + \sqrt{2})\pi \approx 1.896$
 (b) $\frac{1}{6}(1 + 2\sqrt{2})\pi \approx 2.005$
 (c) 2

9.
 (a) 38.5781
 (b) $147/4 \approx 36.75$
 (c) $147/4 \approx 36.75$

11.
 (a) 0
 (b) 0
 (c) 0

13. Trapezoidal Rule: 0.9006
 Simpson's Rule: 0.90452

15. Trapezoidal Rule: 13.9604
 Simpson's Rule: 13.9066

17. Trapezoidal Rule: 1.1703
 Simpson's Rule: 1.1873

19. Trapezoidal Rule: 1.0803
 Simpson's Rule: 1.077

21.
 (a) $n = 161$ (using $\max\left(f''(x)\right) = 1$)
 (b) $n = 12$ (using $\max\left(f^{(4)}(x)\right) = 1$)

23.
 (a) $n = 1004$ (using $\max\left(f''(x)\right) = 39$)
 (b) $n = 62$ (using $\max\left(f^{(4)}(x)\right) = 800$)

25.
 (a) Area is 30.8667 cm^2.
 (b) Area is $308,667$ yd^2.

27. Let $f(x) = a(x-x_1)^2 + b(x-x_1) + c$, so that $f(x_1) = c = y_1$, $f(x_1 + \Delta x) = a\Delta x^2 + b\Delta x + c = y_2$, and $f(x_1 + 2\Delta x) = 4a\Delta x^2 + 2b\Delta x + c = y3$. Therefore, $a = \frac{y_1 - 2y_2 + y_3}{2(\Delta x)^2}$ and $b = \frac{4y_2 - y_3 - 3y_1}{2\Delta x}$, and $\int_{x_1}^{x_1 + 2\Delta x} a(x-x_1)^2 + b(x-x_1) + cdx = \frac{a(2\Delta x)^3}{3} + \frac{b(2\Delta x)^2}{2} + c(2\Delta x) = \frac{4(y_1 - 2y_2 + y_3)\Delta x}{3} + (4y_2 - y_3 - 3y_1)\Delta x + 2y_1\Delta x = \frac{\Delta x}{3}(4y_1 - 8y_2 + 4y_3 + 12y_2 - 3y_3 - 9y_1 + 6y_1) = \frac{\Delta x}{3}(y_1 + 4y_2 + y_3).$

Chapter 9

Exercises 9.1

1. Answers will vary.

3. Answers will vary.

5. $2, \frac{8}{3}, \frac{8}{3}, \frac{32}{15}, \frac{64}{45}$

7. $\frac{1}{3}, 2, \frac{81}{5}, \frac{512}{3}, \frac{15625}{7}$

9. $a_n = 3n + 1$

11. $a_n = 10 \cdot 2^{n-1}$

13. $1/7$

15. 0

17. diverges

19. converges to 0

21. converges to 0

23. diverges

25. converges to e

27. converges to 5

29. diverges

31. converges to 0

33. converges to 0

35. converges to ln 2

37. converges to 0

39. converges to 1

41. bounded

43. bounded above

45. monotonically increasing for $n \geq 3$

47. monotonically decreasing for $n \geq 3$

49. monotonically decreasing

51.
   (a) Left to reader
   (b) $a_n = 1/3^n$ and $b_n = 1/2^n$

Exercises 9.2

1. Answers will vary.

3. One sequence is the sequence of terms $\{a_n\}$. The other is the sequence of nth partial sums, $\{S_n\} = \{\sum_{i=1}^{n} a_i\}$.

5. F

7.
   (a) $1, \frac{5}{4}, \frac{49}{36}, \frac{205}{144}, \frac{5269}{3600}$
   (b) Plot omitted

9.
   (a) $1, 3, 6, 10, 15$
   (b) Plot omitted

11.
   (a) $\frac{1}{3}, \frac{4}{9}, \frac{13}{27}, \frac{40}{81}, \frac{121}{243}$
   (b) Plot omitted

13.
   (a) $0.1, 0.11, 0.111, 0.1111, 0.11111$
   (b) Plot omitted

15. $\lim\limits_{n\to\infty} a_n = 3$; by Theorem 67 the series diverges.

17. $\lim\limits_{n\to\infty} a_n = \infty$; by Theorem 67 the series diverges.

19. $\lim\limits_{n\to\infty} a_n = \infty$; by Theorem 67 the series diverges.

21. Diverges

23. $\lim\limits_{n\to\infty} a_n = 1/2$; by Theorem 67 the series diverges.

25. Diverges by the Test for Divergence

27. Diverges by Theorem 67

29. Converges

31.
   (a) $S_n = \left(\frac{n(n+1)}{2}\right)^2$
   (b) Diverges

33.
   (a) $S_n = 5\frac{1-1/2^n}{1/2}$
   (b) Converges to 10.

35.
   (a) $S_n = \frac{1-(-1/3)^n}{4/3}$
   (b) Converges to 3/4.

37.
   (a) With partial fractions, $a_n = \frac{3}{2}\left(\frac{1}{n} - \frac{1}{n+2}\right)$.
       Thus $S_n = \frac{3}{2}\left(\frac{3}{2} - \frac{1}{n+1} - \frac{1}{n+2}\right)$.
   (b) Converges to 9/4

39.
   (a) $S_n = \ln\left(1/(n+1)\right)$
   (b) Diverges (to $-\infty$).

41.
   (a) $a_n = \frac{1}{n(n+3)}$; using partial fractions, the resulting telescoping sum reduces to $S_n = \frac{1}{3}\left(1 + \frac{1}{2} + \frac{1}{3} - \frac{1}{n+1} - \frac{1}{n+2} - \frac{1}{n+3}\right)$
   (b) Converges to 11/18.

43.
   (a) With partial fractions, $a_n = \frac{1}{2}\left(\frac{1}{n-1} - \frac{1}{n+1}\right)$. Thus $S_n = \frac{1}{2}\left(3/2 - \frac{1}{n} - \frac{1}{n+1}\right)$.
   (b) Converges to 3/4.

45.
   (a) $S_n = 1 + \frac{1}{2} - \frac{1}{n+1} - \frac{1}{n+2} + \frac{\frac{5}{4}(1-(\frac{1}{4})^n)}{1-\frac{1}{4}}$
   (b) Converges to $\frac{19}{6}$.

47. $(-5, -1)$

49. $(-3, -1)$

Exercises 9.3

1. continuous, positive and decreasing

3. Converges

5. Diverges

7. Converges

9. Converges

11. $p > 1$

13. $p > 1$

Exercises 9.4

1. $\sum_{n=0}^{\infty} b_n$ converges; we cannot conclude anything about $\sum_{n=0}^{\infty} c_n$

3. Converges; compare to $\sum_{n=1}^{\infty} \dfrac{1}{4^n}$, as $1/(4^n + n^2 - n) \le 1/4^n$ for all $n \ge 1$.

5. Converges; compare to $\sum_{n=1}^{\infty} \dfrac{1}{n!}$, as $1/(n! + n) \le 1/n!$ for all $n \ge 1$.

7. Converges; compare to $\sum_{n=1}^{\infty} \dfrac{1}{n^2}$.

9. Diverges; compare to $\sum_{n=1}^{\infty} \dfrac{\ln n}{n}$.

11. Diverges; compare to $\sum_{n=1}^{\infty} \dfrac{1}{n}$.

13. Diverges; compare to $\sum_{n=1}^{\infty} \dfrac{1}{n}$:

$$\frac{1}{n} = \frac{n^2}{n^3} < \frac{n^2 + n + 1}{n^3} < \frac{n^2 + n + 1}{n^3 - 5},$$

for all $n \ge 1$.

15. Diverges; compare to $\sum_{n=1}^{\infty} \dfrac{1}{n}$. Just as $\lim_{n \to 0} \dfrac{\sin n}{n} = 1$,

$\lim_{n \to \infty} \dfrac{\sin(1/n)}{1/n} = 1$.

17. Converges; compare to $\sum_{n=1}^{\infty} \dfrac{1}{n^2}$.

19. Converges by Comparison Test with $\sum \dfrac{1}{4^n}$

21. Diverges; compare to $\sum_{n=1}^{\infty} \dfrac{1}{n^{1/2}}$.

23. Converges; compare to $\sum_{n=1}^{\infty} \dfrac{1}{n^{3/2}}$.

25.
    (a) Converges; use Direct Comparison Test as $\dfrac{a_n}{n} < n$.
    (b) Converges; since original series converges, we know $\lim_{n \to \infty} a_n = 0$. Thus for large n, $a_n a_{n+1} < a_n$.
    (c) Converges; similar logic to part (b) so $(a_n)^2 < a_n$.
    (d) May converge; certainly $n a_n > a_n$ but that does not mean it does not converge.
    (e) Does not converge, using logic from (b) and n^{th} Term Test.

27. Converges; Integral Test, p-Series Test, Direct & Limit Comparison Tests can all be used.

29. Converges; the Direct Comparison Test can be used with sequence $1/(n - 1)!$.

31. Diverges; the nth Term Test can be used, along with the Limit Comparison Test (compare with $1/10$).

33. Converges; the Direct Comparison Test can be used with sequence $1/\sqrt{n}$.

Exercises 9.5

1. The signs of the terms do not alternate; in the given series, some terms are negative and the others positive, but they do not necessarily alternate.

3. Many examples exist; one common example is $a_n = (-1)^n/n$.

5.
    (a) converges
    (b) converges (p-Series)
    (c) absolute

7.
    (a) diverges (limit of terms is not 0)
    (b) diverges
    (c) n/a; diverges

9.
    (a) converges
    (b) diverges (Limit Comparison Test with $1/n$)
    (c) conditional

11.
    (a) diverges (limit of terms is not 0)
    (b) diverges
    (c) n/a; diverges

13.
    (a) diverges (terms oscillate between ± 1)
    (b) diverges
    (c) n/a; diverges

15.
    (a) converges
    (b) converges (Geometric Series with $r = 2/3$)
    (c) absolute

17.
    (a) converges
    (b) converges (Ratio Test)
    (c) absolute

19.
    (a) converges
    (b) diverges (p-Series Test with $p = 1/2$)
    (c) conditional

21. $S_5 = -1.1906$; $S_6 = -0.6767$;
$$-1.1906 \le \sum_{n=1}^{\infty} \frac{(-1)^n}{\ln(n + 1)} \le -0.6767$$

23. $S_6 = 0.3681$; $S_7 = 0.3679$;
$$0.3681 \le \sum_{n=0}^{\infty} \frac{(-1)^n}{n!} \le 0.3679$$

25. $n = 5$

27. Using the theorem, we find $n = 499$ guarantees the sum is within 0.001 of $\pi/4$. (Convergence is actually faster, as the sum is within ε of $\pi/24$ when $n \geq 249$.)

Exercises 9.6

1. algebraic, or polynomial.

3. Integral Test, Limit Comparison Test, and Root Test

5. Converges

7. Converges

9. The Ratio Test is inconclusive; the p-Series Test states it diverges.

11. Converges

13. Converges; note the summation can be rewritten as $\displaystyle\sum_{n=1}^{\infty} \frac{2^n n!}{3^n n!}$, from which the Ratio Test can be applied.

15. Diverges

17. Converges

19. Converges

21. Diverges

23. Diverges. The Root Test is inconclusive, but the nth-Term Test shows divergence. (The terms of the sequence approach e^2, not 0, as $n \to \infty$.)

25. Converges

Exercises 9.7

1. Diverges

3. Diverges

5. Diverges

7. Absolutely converges

9. Conditionally converges

11. Diverges

13. Absolutely converges

15. Absolutely converges

17. Absolutely converges

19. Conditionally converges

21. Absolutely converges

23. Absolutely converges

25. Diverges

27. Diverges

29. Absolutely converges

31. Diverges

33. Absolutely converges

35. Diverges

37. Absolutely converges

Exercises 9.8

1. 1

3. 5

5. $1 + 2x + 4x^2 + 8x^3 + 16x^4$

7. $1 + x + \frac{x^2}{2} + \frac{x^3}{6} + \frac{x^4}{24}$

9.
  (a) $R = \infty$
  (b) $(-\infty, \infty)$

11.
  (a) $R = 1$
  (b) $(2, 4]$

13.
  (a) $R = 2$
  (b) $(-2, 2)$

15.
  (a) $R = 1/5$
  (b) $(4/5, 6/5)$

17.
  (a) $R = 1$
  (b) $(-1, 1)$

19.
  (a) $R = \infty$
  (b) $(-\infty, \infty)$

21.
  (a) $R = 1$
  (b) $[-1, 1]$

23.
  (a) $R = 0$
  (b) $x = 0$

25.
  (a) $R = \frac{1}{9}$
  (b) $x = [\frac{5}{9}, \frac{7}{9}]$

27.
  (a) $R = \infty$
  (b) $x = (-\infty, \infty)$

29. $\displaystyle\sum_{n=0}^{\infty} 8^n x^{n+1}, R = 1/8$

31. $\displaystyle\sum_{n=0}^{\infty} 6(-7)^n x^{4n}, R = 1/\sqrt[4]{7}$

33. $\displaystyle\sum_{n=0}^{\infty} \frac{3 \cdot 2^n x^{n/3+2}}{5^{n+1}}, R = 125/8$

35. $\ln 3 - \displaystyle\sum_{n=1}^{\infty} 3^{-n} x^n / n, R = 3$

37. $\sum_{n=0}^{\infty} \frac{2}{2n+1} x^{2n+1}, R = 1$

39. $\sum_{n=0}^{\infty} (-1)^n \frac{x^{6n+5}}{2n+1}; R = 1$

41. $\sum_{n=0}^{\infty} \frac{(n+2)(n+1)}{2^{n+4}} x^{n+3}; R = 2$

Exercises 9.9

1. The Maclaurin polynomial is a special case of Taylor polynomials. Taylor polynomials are centered at a specific x-value; when that x-value is 0, it is a Maclaurin polynomial.

3. $p_2(x) = 6 + 3x - 4x^2$.

5. $p_3(x) = 1 - x + \frac{1}{2}x^3 - \frac{1}{6}x^3$

7. $p_8(x) = x + x^2 + \frac{1}{2}x^3 + \frac{1}{6}x^4 + \frac{1}{24}x^5$

9. $p_4(x) = \frac{2x^4}{3} + \frac{4x^3}{3} + 2x^2 + 2x + 1$

11. $p_4(x) = x^4 - x^3 + x^2 - x + 1$

13. $p_4(x) = 1 + \frac{1}{2}(-1+x) - \frac{1}{8}(-1+x)^2 + \frac{1}{16}(-1+x)^3 - \frac{5}{128}(-1+x)^4$

15. $p_6(x) = \frac{1}{\sqrt{2}} - \frac{-\frac{\pi}{4}+x}{\sqrt{2}} - \frac{\left(-\frac{\pi}{4}+x\right)^2}{2\sqrt{2}} + \frac{\left(-\frac{\pi}{4}+x\right)^3}{6\sqrt{2}} + \frac{\left(-\frac{\pi}{4}+x\right)^4}{24\sqrt{2}} - \frac{\left(-\frac{\pi}{4}+x\right)^5}{120\sqrt{2}} - \frac{\left(-\frac{\pi}{4}+x\right)^6}{720\sqrt{2}}$

17. $p_5(x) = \frac{1}{2} - \frac{x-2}{4} + \frac{1}{8}(x-2)^2 - \frac{1}{16}(x-2)^3 + \frac{1}{32}(x-2)^4 - \frac{1}{64}(x-2)^5$

19. $p_3(x) = \frac{1}{2} + \frac{1+x}{2} + \frac{1}{4}(1+x)^2$

21. $p_3(x) = x - \frac{x^3}{6}; p_3(0.1) = 0.09983.$ Error is bounded by $\pm\frac{1}{4!} \cdot 0.1^4 \approx \pm 0.000004167.$

23. $p_2(x) = 3 + \frac{1}{6}(-9+x) - \frac{1}{216}(-9+x)^2; p_2(10) = 3.16204.$ The third derivative of $f(x) = \sqrt{x}$ is bounded on $(8, 11)$ by $0.003.$ Error is bounded by $\pm\frac{0.003}{3!} \cdot 1^3 = \pm 0.0005.$

25. The nth derivative of $f(x) = e^x$ is bounded by 3 on intervals containing 0 and 1. Thus $|R_n(1)| \leq \frac{3}{(n+1)!} 1^{(n+1)}.$ When $n = 7$, this is less than 0.0001.

27. The nth derivative of $f(x) = \cos x$ is bounded by 1 on intervals containing 0 and $\pi/3$. Thus $|R_n(\pi/3)| \leq \frac{1}{(n+1)!}(\pi/3)^{(n+1)}.$ When $n = 7$, this is less than 0.0001. Since the Maclaurin polynomial of $\cos x$ only uses even powers, we can actually just use $n = 6$.

29. The nth term is $\frac{1}{n!}x^n$.

31. The nth term is x^n.

33. The nth term is $(-1)^n \frac{(x-1)^n}{n}$.

35. $3 + 15x + \frac{75}{2}x^2 + \frac{375}{6}x^3 + \frac{1875}{24}x^4$

Exercises 9.10

1. A Taylor polynomial is a **polynomial**, containing a finite number of terms. A Taylor series is a **series**, the summation of an infinite number of terms.

3. All derivatives of e^x are e^x which evaluate to 1 at $x = 0$. The Taylor series starts $1 + x + \frac{1}{2}x^2 + \frac{1}{3!}x^3 + \frac{1}{4!}x^4 + \cdots$; the Taylor series is $\sum_{n=0}^{\infty} \frac{x^n}{n!}$

5. The nth derivative of $1/(1-x)$ is $f^{(n)}(x) = (n)!/(1-x)^{n+1}$, which evaluates to $n!$ at $x = 0$. The Taylor series starts $1 + x + x^2 + x^3 + \cdots$; the Taylor series is $\sum_{n=0}^{\infty} x^n$

7. The Taylor series starts $0 - (x-\pi/2) + 0x^2 + \frac{1}{6}(x-\pi/2)^3 + 0x^4 - \frac{1}{120}(x-\pi/2)^5$; the Taylor series is $\sum_{n=0}^{\infty} (-1)^{n+1} \frac{(x-\pi/2)^{2n+1}}{(2n+1)!}$

9. $f^{(n)}(x) = (-1)^n e^{-x}$; at $x = 0, f^{(n)}(0) = -1$ when n is odd and $f^{(n)}(0) = 1$ when n is even. The Taylor series starts $1 - x + \frac{1}{2}x^2 - \frac{1}{3!}x^3 + \cdots$; the Taylor series is $\sum_{n=0}^{\infty} (-1)^n \frac{x^n}{n!}.$

11. $f^{(n)}(x) = (-1)^{n+1} \frac{n!}{(x+1)^{n+1}}$; at $x = 1, f^{(n)}(1) = (-1)^{n+1} \frac{n!}{2^{n+1}}$ The Taylor series starts $\frac{1}{2} + \frac{1}{4}(x-1) - \frac{1}{8}(x-1)^2 + \frac{1}{16}(x-1)^3 \cdots$; the Taylor series is $\sum_{n=0}^{\infty} (-1)^{n+1} \frac{(x-1)^n}{2^{n+1}}.$

13. Given a value x, the magnitude of the error term $R_n(x)$ is bounded by

$$|R_n(x)| \leq \frac{\max \left| f^{(n+1)}(z) \right|}{(n+1)!} \left| x^{(n+1)} \right|,$$

where z is between 0 and x.
If $x > 0$, then $z < x$ and $f^{(n+1)}(z) = e^z < e^x$. If $x < 0$, then $x < z < 0$ and $f^{(n+1)}(z) = e^z < 1$. So given a fixed x value, let $M = \max\{e^x, 1\}; f^{(n)}(z) < M$. This allows us to state

$$|R_n(x)| \leq \frac{M}{(n+1)!} \left| x^{(n+1)} \right|.$$

For any x, $\lim_{n\to\infty} \frac{M}{(n+1)!} \left| x^{(n+1)} \right| = 0$. Thus by the Squeeze Theorem, we conclude that $\lim_{n\to\infty} R_n(x) = 0$ for all x, and hence

$$e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!} \quad \text{for all } x.$$

15. Given a value x, the magnitude of the error term $R_n(x)$ is bounded by

$$|R_n(x)| \leq \frac{\max \left| f^{(n+1)}(z) \right|}{(n+1)!} \left| x^{(n+1)} \right|,$$

where z is between 0 and x. Since $\left| f^{(n+1)}(z) \right| = \frac{n!}{(z+1)^{n+1}}$,

$$|R_n(x)| \leq \frac{1}{n+1} \left(\frac{|x|}{\min z + 1} \right)^{n+1}.$$

If $0 < x < 1$, then $0 < z < x$ and $f^{(n+1)}(z) = \frac{n!}{(z+1)^{n+1}} < n!$. Thus

$$|R_n(x)| \leq \frac{n!}{(n+1)!} \left| x^{(n+1)} \right| = \frac{x^{n+1}}{n+1}.$$

For a fixed $x < 1$,

$$\lim_{n \to \infty} \frac{x^{n+1}}{n+1} = 0.$$

17. Given $\cos x = \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n}}{(2n)!}$,

$$\cos(-x) = \sum_{n=0}^{\infty} (-1)^n \frac{(-x)^{2n}}{(2n)!} = \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n}}{(2n)!} = \cos x,$$

as all powers in the series are even.

19. Given $\sin x = \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n+1}}{(2n+1)!}$,

$$\frac{d}{dx}(\sin x) = \frac{d}{dx} \left(\sum_{n=0}^{\infty} (-1)^n \frac{x^{2n+1}}{(2n+1)!} \right)$$

$$= \sum_{n=0}^{\infty} (-1)^n \frac{(2n+1)x^{2n}}{(2n+1)!} = \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n}}{(2n)!} = \cos x.$$

(The summation still starts at $n = 0$ as there was no constant term in the expansion of $\sin x$).

21. $1 + \dfrac{x}{2} - \dfrac{x^2}{8} + \dfrac{x^3}{16} - \dfrac{5x^4}{128}$

23. $1 + \dfrac{x}{3} - \dfrac{x^2}{9} + \dfrac{5x^3}{81} - \dfrac{10x^4}{243}$

25. $\sum_{n=0}^{\infty} (-1)^n \dfrac{(x^2)^{2n}}{(2n)!} = \sum_{n=0}^{\infty} (-1)^n \dfrac{x^{4n}}{(2n)!}.$

27. $\sum_{n=0}^{\infty} (-1)^n \dfrac{(2x+3)^{2n+1}}{(2n+1)!}.$

29. $x + x^2 + \dfrac{x^3}{3} - \dfrac{x^5}{30}$

31. $\displaystyle\int_0^{\sqrt{\pi}} \sin\left(x^2\right) dx \approx \int_0^{\sqrt{\pi}} \left(x^2 - \dfrac{x^6}{6} + \dfrac{x^{10}}{120} - \dfrac{x^{14}}{5040} \right) dx$
$= 0.8877$

Chapter 10

Exercises 10.0

1. $y = \frac{1}{2}(x-3)^2 + \frac{3}{2}$

2. $y = \frac{-1}{12}(x+1)^2 - 1$

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

4. $x = y^2$

5. $y = -\frac{1}{4}(x-1)^2 + 2$

6. $x = -\frac{1}{12}y^2$

7. $y = 4x^2$

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

9.

10.
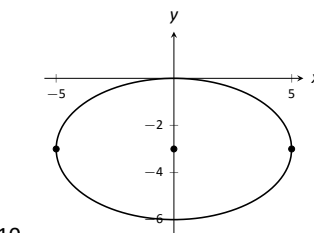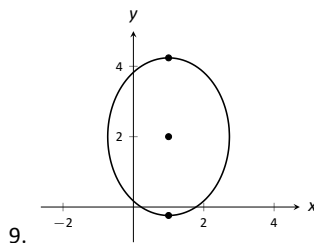
11. $\dfrac{(x+1)^2}{9} + \dfrac{(y-2)^2}{4} = 1$

12. $\dfrac{(x-1)^2}{1/4} + \dfrac{y^2}{9} = 1$

13. $\dfrac{(x-1)^2}{2} + (y-2)^2 = 1$

14. $\dfrac{x^2}{3} + \dfrac{y^2}{5} = 1$

15. $\dfrac{x^2}{4} + \dfrac{(y-3)^2}{6} = 1$

16. $\dfrac{(x-2)^2}{4} + \dfrac{(y-2)^2}{4} = 1$

17. $x^2 - \dfrac{y^2}{3} = 1$

18. $y^2 - \dfrac{x^2}{24} = 1$

19. $\dfrac{(y-3)^2}{4} - \dfrac{(x-1)^2}{9} = 1$

20. $\dfrac{(x-1)^2}{9} - \dfrac{(y-3)^2}{4} = 1$

21.

22.
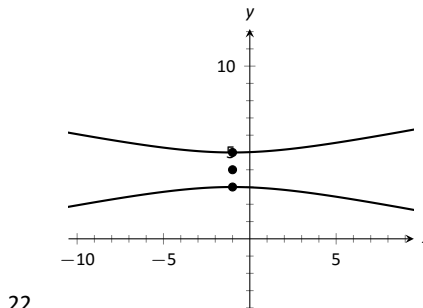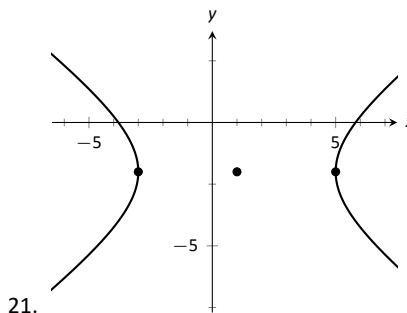
23. $\frac{x^2}{4} - \frac{y^2}{3} = 1$

24. $\frac{x^2}{3} - \frac{(y-1)^2}{9} = 1$

25. $(y-2)^2 - \frac{x^2}{10} = 1$

26. $4y^2 - \frac{x^2}{4} = 1$

Exercises 10.1

1. T

3. $\sqrt{2}$

5. $4/3$

7. $109/2$

9. $12/5$

11. $-\ln(2 - \sqrt{3}) \approx 1.31696$

13. $\int_0^1 \sqrt{1 + 4x^2}\, dx$

15. $\int_0^1 \sqrt{1 + \frac{1}{4x}}\, dx$

17. $\int_{-1}^1 \sqrt{1 + \frac{x^2}{1 - x^2}}\, dx$

19. $\int_1^2 \sqrt{1 + \frac{1}{x^4}}\, dx$

21. 1.4790

23. Simpson's Rule fails, as it requires one to divide by 0. However, recognize the answer should be the same as for $y = x^2$; why?

25. Simpson's Rule fails.

27. 1.4058

29. $2\pi \int_0^1 2x\sqrt{5}\, dx = 2\pi\sqrt{5}$

31. $2\pi \int_0^1 x^3 \sqrt{1 + 9x^4}\, dx = \pi/27(10\sqrt{10} - 1)$

33. $2\pi \int_0^1 \sqrt{1 - x^2}\sqrt{1 + x/(1 - x^2)}\, dx = 4\pi$

Exercises 10.2

1. T

3. rectangular

5.

7.

9.

11.

13.

15.

17.
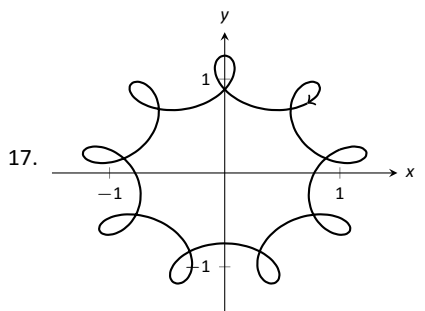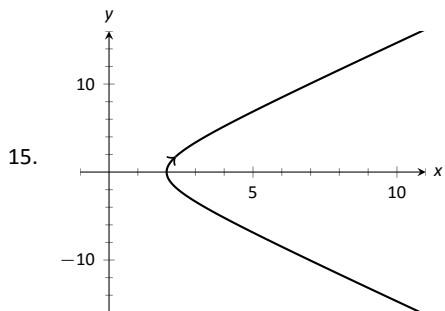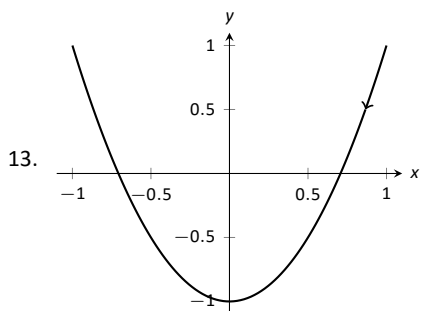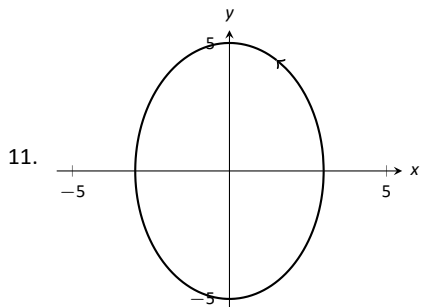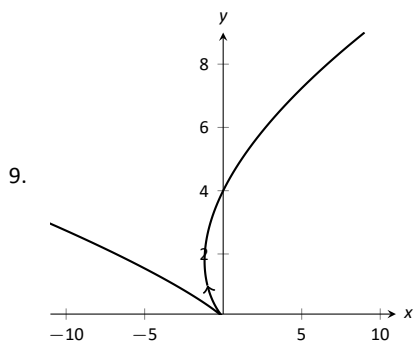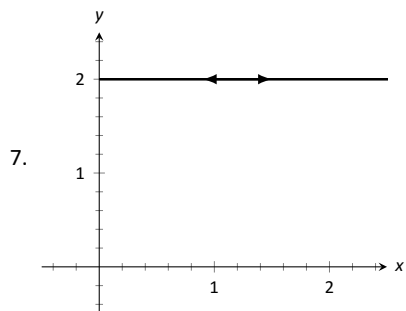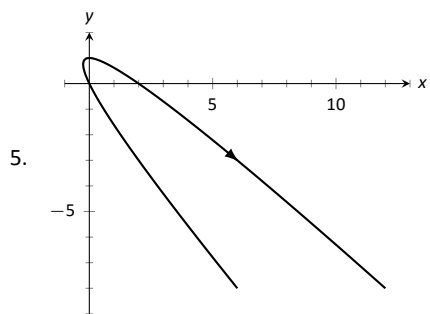
19.

    (a) Traces a circle of radius 1 counterclockwise once.

    (b) Traces a circle of radius 1 counterclockwise over 6 times.

    (c) Traces a circle of radius 1 clockwise infinite times.

    (d) Traces an arc of a circle of radius 1, from an angle of -1 radians to 1 radian, twice.

21. Possible Answer: $x = \frac{5+t^2}{4}, y = t$

23. Possible Answer: $x = \frac{5}{4}t + \frac{11}{4}, y = t, [-3, 1]$

25. Possible Answer: $x = t, y = t^2 + 2t, (-\infty, -1]$

27. $x = (t+11)/6, y = (t^2-97)/12$. At $t = 1, x = 2, y = -8$. $y' = 6x - 11$; when $x = 2, y' = 1$.

29. $x = \cos^{-1} t, y = \sqrt{1-t^2}$. At $t = 1, x = 0, y = 0$. $y' = \cos x$; when $x = 0, y' = 1$.

31. Possible answers:

    (a) $x = \sin t, y = \cos t, [\pi/2, 5\pi/2]$

    (b) $x = \cos t, y = \sin t, [0, 2\pi]$

    (c) $x = \sin t, y = \cos t, [\pi/2, 9\pi/2]$

    (d) $x = \cos t, y = \sin t, [0, 4\pi]$

33. $x = 4t, y = -16t^2 + 64t$

35. $x = 10t, y = -16t^2 + 320t$

37. $x = 3\cos(2\pi t) + 1, y = 3\sin(2\pi t) + 1$; other answers possible

39. $x = 5\cos t, y = \sqrt{24}\sin t$; other answers possible

41. $x = 2\tan t, y = \pm 6\sec t$; other answers possible

43. $x^2 - y^2 = 1$

45. $y = x^{3/2}$

47. $y = x^3 - 3$

49. $y^2 - x^2 = 1$

51. $x = 1 - 2y^2$

53. $x^2 + y^2 = r^2$; circle centered at $(0, 0)$ with radius r.

55. $\frac{(x-h)^2}{a^2} - \frac{(y-k)^2}{b^2} = 1$; hyperbola centered at (h, k) with horizontal transverse axis and asymptotes with slope b/a. The parametric equations only give half of the hyperbola. When $a > 0$, the right half; when $a < 0$, the left half.

57. $t = -1, 2$

59. $t = \pi/6, \pi/2, 5\pi/6$

61. $t = 2$

63. $t = \ldots 0, 2\pi, 4\pi, \ldots$

Exercises 10.3

1. F

3. F

5.

    (a) $\frac{dy}{dx} = 2t$

    (b) Tangent line: $y = 2(x - 1) + 1$; normal line: $y = -1/2(x - 1) + 1$

7.

    (a) $\frac{dy}{dx} = \frac{2t+1}{2t-1}$

    (b) Tangent line: $y = 3x+2$; normal line: $y = -1/3x+2$

9.

    (a) $\frac{dy}{dx} = \csc t$

    (b) $t = \pi/4$: Tangent line: $y = \sqrt{2}(x-\sqrt{2})+1$; normal line: $y = -1/\sqrt{2}(x - \sqrt{2}) + 1$

11.

    (a) $\frac{dy}{dx} = \frac{\cos t \sin(2t)+\sin t \cos(2t)}{-\sin t \sin(2t)+2\cos t \cos(2t)}$

    (b) Tangent line: $y = x-\sqrt{2}$; normal line: $y = -x-\sqrt{2}$

13. $t = 0$

15. $t = -1/2$

17. The graph does not have a horizontal tangent line.

19. The solution is non-trivial; use identities $\sin(2t) = 2\sin t \cos t$ and $\cos(2t) = \cos^2 t - \sin^2 t$ to rewrite $g'(t) = 2\sin t(2\cos^2 t - \sin^2 t)$. On $[0, 2\pi]$, $\sin t = 0$ when $t = 0, \pi, 2\pi$, and $2\cos^2 t - \sin^2 t = 0$ when $t = \tan^{-1}(\sqrt{2}), \pi \pm \tan^{-1}(\sqrt{2}), 2\pi - \tan^{-1}(\sqrt{2})$.

21. $t_0 = 0$; $\lim_{t\to 0}\frac{dy}{dx} = 0$.

23. $t_0 = 1$; $\lim_{t\to 1}\frac{dy}{dx} = \infty$.

25. $\frac{d^2y}{dx^2} = 2$; always concave up

27. $\frac{d^2y}{dx^2} = -\frac{4}{(2t-1)^3}$; concave up on $(-\infty, 1/2)$; concave down on $(1/2, \infty)$.

29. $\frac{d^2y}{dx^2} = -\cot^3 t$; concave up on $(-\infty, 0)$; concave down on $(0, \infty)$.

31. $\frac{d^2y}{dx^2} = \frac{4(13+3\cos(4t))}{(\cos t + 3\cos(3t))^3}$, obtained with a computer algebra system; concave up on $\left(-\tan^{-1}(\frac{1}{\sqrt{2}}), \tan^{-1}(\frac{1}{\sqrt{2}})\right)$, concave down on $\left(-\frac{\pi}{2}, -\tan^{-1}(\frac{1}{\sqrt{2}})\right)$; $\left(\tan^{-1}(\frac{1}{\sqrt{2}}), \frac{\pi}{2}\right)$

33. $L = 6\pi$

35. $L = 2\sqrt{34}$

37. 2π

39. $-\frac{\sqrt{10}}{3} + \ln(3 + \sqrt{10}) + \sqrt{2} - \ln(1 + \sqrt{2})$

41. $L \approx 2.4416$ (actual value: $L = 2.42211$)

43. $L \approx 4.19216$ (actual value: $L = 4.18308$)

45. The answer is 16π for both (of course), but the integrals are different.

47. $\frac{6\pi a^2}{5}$

49. $\frac{24\pi(949\sqrt{26}+1)}{5}$

51. $SA \approx 1.36751$ (actual value $SA = 1.36707$)

Exercises 10.4

1. Answers will vary.

3. T

5.
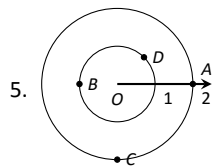
7. $A(2.5, \pi/4)$ and $A(-2.5, 5\pi/4)$;
   $B(-1, 5\pi/6)$ and $B(1, 11\pi/6)$;
   $C(3, 4\pi/3)$ and $C(-3, \pi/3)$;
   $D(1.5, 2\pi/3)$ and $D(-1.5, 5\pi/3)$

9. $A = (\sqrt{2}, \sqrt{2})$;
   $B = (\sqrt{2}, -\sqrt{2})$;
   $C = (\sqrt{5}, -0.46)$;
   $D = (\sqrt{5}, 2.68)$

11.

13.

15.

17.

19.

21.

23.

25.

27.

29.

31.
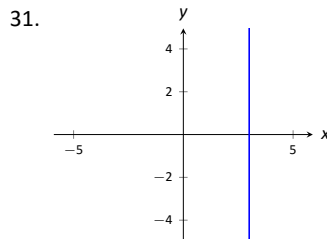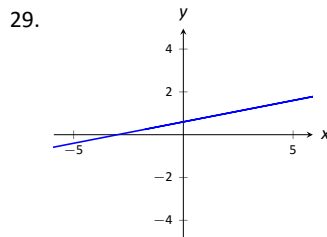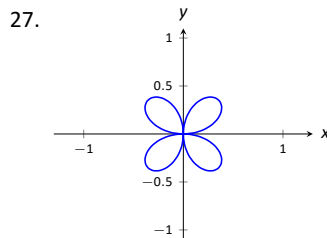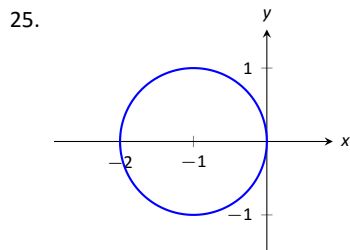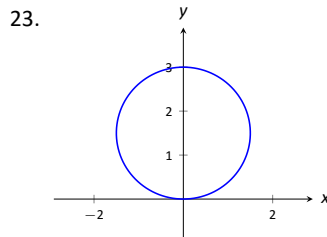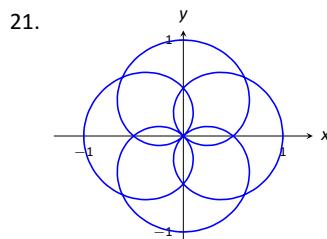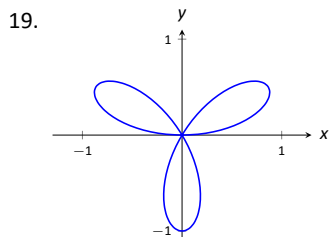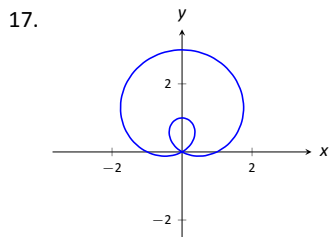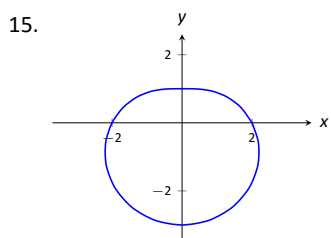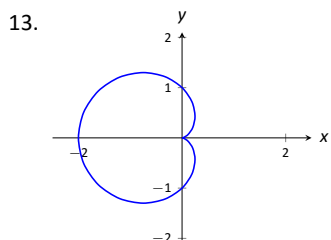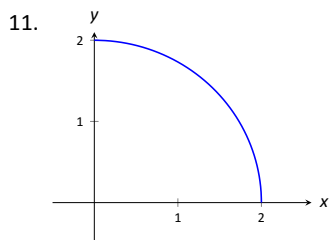
33. $(x - 1)^2 + y^2 = 1$

35. $x^2 + (y - \frac{3}{2})^2 = \frac{9}{4}$

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

39. $x = 3$

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

43. $y = x/\sqrt{3}$

45. $r = 7/(\sin\theta - 4\cos\theta)$

47. $r = 5\csc\theta$

49. $r = 1/\sqrt[3]{\cos^2 \theta \sin \theta}$

51. $r = -2 \cos \theta$

53. $P(1, 0), P(0, \pi/2) = P(0, \pi/4), P(-1/2, \pi/3)$

55. $P(\sqrt{3}/2, \pi/3) = P(-\sqrt{3}/2, 4\pi/3), P(\sqrt{3}/2, 2\pi/3) = P(-\sqrt{3}/2, 5\pi/3), P(0, \pi/2)$

57. $P(3/2, \pi/3), P(3/2, -\pi/3)$

59. $P(0, 0) = P(0, 3\pi/2), P(1 + \sqrt{2}/2, 3\pi/4), P(1 - \sqrt{2}/2, 7\pi/4)$

61. Answers will vary.

Exercises 10.5

1. Using $x = r \cos \theta$ and $y = r \sin \theta$, we can write $x = f(\theta) \cos \theta, y = f(\theta) \sin \theta$.

3.
   (a) $\frac{dy}{dx} = -\cot \theta$
   (b) tangent line: $y = -(x - \sqrt{2}/2) + \sqrt{2}/2$; normal line: $y = x$

5.
   (a) $\frac{dy}{dx} = \frac{\cos \theta (1 + 2 \sin \theta)}{\cos^2 \theta - \sin \theta (1 + \sin \theta)}$
   (b) tangent line: $x = 3\sqrt{3}/4$; normal line: $y = 3/4$

7.
   (a) $\frac{dy}{dx} = \frac{\theta \cos \theta + \sin \theta}{\cos \theta - \theta \sin \theta}$
   (b) tangent line: $y = -2/\pi x + \pi/2$; normal line: $y = \pi/2x + \pi/2$

9.
   (a) $\frac{dy}{dx} = \frac{4 \sin(\theta) \cos(4\theta) + \sin(4\theta) \cos(\theta)}{4 \cos(\theta) \cos(4\theta) - \sin(\theta) \sin(4\theta)}$
   (b) tangent line: $y = 5\sqrt{3}(x + \sqrt{3}/4) - 3/4$; normal line: $y = -1/5\sqrt{3}(x + \sqrt{3}/4) - 3/4$

11. horizontal: $\theta = \pi/2, 3\pi/2$; vertical: $\theta = 0, \pi, 2\pi$

13. horizontal: $\theta = \tan^{-1}(1/\sqrt{5}), \pi/2$, $\pi - \tan^{-1}(1/\sqrt{5}), \pi + \tan^{-1}(1/\sqrt{5}), 3\pi/2, 2\pi - \tan^{-1}(1/\sqrt{5})$; vertical: $\theta = 0, \tan^{-1}(\sqrt{5}), \pi - \tan^{-1}(\sqrt{5}), \pi, \pi + \tan^{-1}(\sqrt{5}), 2\pi - \tan^{-1}(\sqrt{5})$

15. In polar: $\theta = 0 \cong \theta = \pi$
    In rectangular: $y = 0$

17. In polar: $\theta = \frac{\pi}{4}$ and $\theta = -\frac{\pi}{4}$
    In rectangular: $y = x$ and $y = -x$.

19. area = $\frac{8\pi}{3} + 4\sqrt{3}$

21. area = $\pi/12$

23. area = $\pi - 3\sqrt{3}/2$

25. area = $\pi + 3\sqrt{3}$

27. area = $\int_{\pi/12}^{\pi/3} \frac{1}{2} \sin^2(3\theta) \, d\theta - \int_{\pi/12}^{\pi/6} \frac{1}{2} \cos^2(3\theta) \, d\theta = \frac{1}{12} + \frac{\pi}{24}$

29. area = $\int_0^{\pi/3} \frac{1}{2}(1 - \cos \theta)^2 \, d\theta + \int_{\pi/3}^{\pi/2} \frac{1}{2}(\cos \theta)^2 \, d\theta = \frac{7\pi}{24} - \frac{\sqrt{3}}{2} \approx 0.0503$

31. 4π

33. $L \approx 2.2592$; (actual value $L = 2.22748$)

35. $SA = 16\pi$

37. $SA = 32\pi/5$

39. $SA = 36\pi$

Chapter 11

Exercises 11.1

1. right hand

3. curve (a parabola); surface (a cylinder)

5. a hyperboloid of two sheets

7. $\|\overline{AB}\| = \sqrt{6}; \|\overline{BC}\| = \sqrt{17}; \|\overline{AC}\| = \sqrt{11}$. Yes, it is a right triangle as $\|\overline{AB}\|^2 + \|\overline{AC}\|^2 = \|\overline{BC}\|^2$.

9. Center at $(4, -1, 0)$; radius = 3

11. Interior of a sphere with radius 1 centered at the origin.

13. The first octant of space along with its adjacent quarter planes; all points (x, y, z) where each of x, y and z are positive or zero. (Analogous to the first quadrant in the plane.)

15.
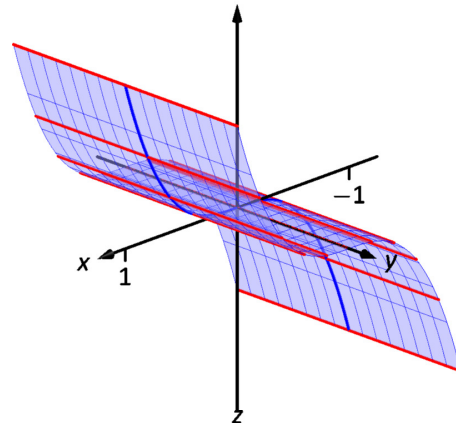
17.
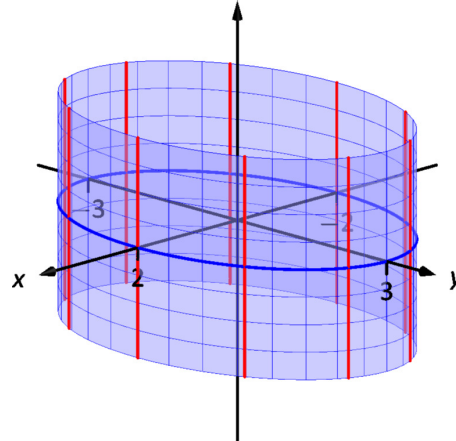
19. $y^2 + z^2 = x^4$

21. $z = (\sqrt{x^2 + y^2})^2 = x^2 + y^2$

23. (a) $x = y^2 + \frac{z^2}{9}$

25. (b) $x^2 + \dfrac{y^2}{9} + \dfrac{z^2}{4} = 1$

27.
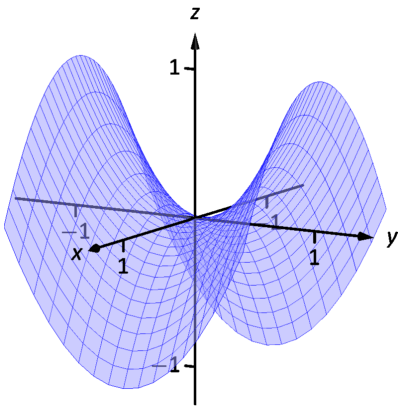
29.
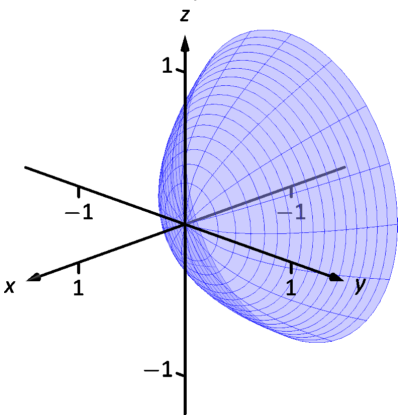
31.
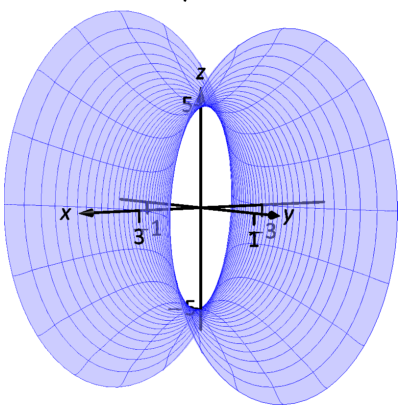
Exercises 11.2

1. Answers will vary.

3. A vector with magnitude 1.

5. It stretches the vector by a factor of 2, and points it in the opposite direction.

7. $\vec{PQ} = \langle 4, -4 \rangle = 4\vec{\imath} - 4\vec{\jmath}$

9. $\vec{PQ} = \langle 2, 2, 0 \rangle = 2\vec{\imath} + 2\vec{\jmath}$

11.

    (a) $\vec{u} + \vec{v} = \langle 3, 2, 1 \rangle$; $\vec{u} - \vec{v} = \langle -1, 0, -3 \rangle$; $\pi\vec{u} - \sqrt{2}\vec{v} = \langle \pi - 2\sqrt{2}, \pi - \sqrt{2}, -\pi - 2\sqrt{2} \rangle$.

    (c) $\vec{x} = \langle -1, 0, -3 \rangle$.

13.
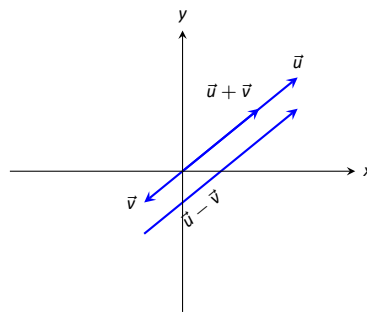
Sketch of $\vec{u} - \vec{v}$ shifted for clarity.

15.
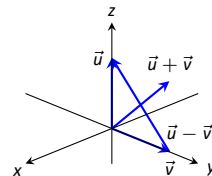
17. $\|\vec{u}\| = \sqrt{17}$, $\|\vec{v}\| = \sqrt{3}$, $\|\vec{u} + \vec{v}\| = \sqrt{14}$, $\|\vec{u} - \vec{v}\| = \sqrt{26}$

19. $\|\vec{u}\| = 7$, $\|\vec{v}\| = 35$, $\|\vec{u} + \vec{v}\| = 42$, $\|\vec{u} - \vec{v}\| = 28$

21. $\vec{u} = \langle 3/\sqrt{58}, 7/\sqrt{58} \rangle$

23. $\vec{u} = \langle 1/3, -2/3, 2/3 \rangle$

25. $\vec{u} = \langle \cos 50°, \sin 50° \rangle \approx \langle 0.643, 0.766 \rangle$.

27.

$$\|\vec{u}\| = \sqrt{\sin^2 \theta \cos^2 \varphi + \sin^2 \theta \sin^2 \varphi + \cos^2 \theta}$$
$$= \sqrt{\sin^2 \theta (\cos^2 \varphi + \sin^2 \varphi) + \cos^2 \theta}$$
$$= \sqrt{\sin^2 \theta + \cos^2 \theta}$$
$$= 1.$$

29. The force on each chain is 100lb.

31. The force on each chain is 50lb.

33. $\theta = 5.71°$; the weight is lifted 0.005 ft (about 1/16th of an inch).

35. $\theta = 84.29°$; the weight is lifted 9 ft.

Exercises 11.3

1. Scalar

3. By considering the sign of the dot product of the two vectors. If the dot product is positive, the angle is acute; if the dot product is negative, the angle is obtuse.

5. -22

7. 3

9. not defined

11. Answers will vary.

13. $\theta = 0.3218 \approx 18.43°$

15. $\theta = \pi/4 = 45°$

17. Answers will vary; two possible answers are $\langle -7, 4 \rangle$ and $\langle 14, -8 \rangle$.

19. Answers will vary; two possible answers are $\langle 1, 0, -1 \rangle$ and $\langle 4, 5, -9 \rangle$.

21. $\text{proj}_{\vec{v}}\,\vec{u} = \langle -1/2, 3/2 \rangle$.

23. $\text{proj}_{\vec{v}}\,\vec{u} = \langle -1/2, -1/2 \rangle$.

25. $\text{proj}_{\vec{v}}\,\vec{u} = \langle 1, 2, 3 \rangle$.

27. $\vec{u} = \langle -1/2, 3/2 \rangle + \langle 3/2, 1/2 \rangle$.

29. $\vec{u} = \langle -1/2, -1/2 \rangle + \langle -5/2, 5/2 \rangle$.

31. $\vec{u} = \langle 1, 2, 3 \rangle + \langle 0, 3, -2 \rangle$.

33. 1.96lb

35. 141.42ft–lb

37. 500ft–lb

39. 500ft–lb

Exercises 11.4

1. vector

3. "Perpendicular" is one answer.

5. Torque

7. $\vec{u} \times \vec{v} = \langle 12, -15, 3 \rangle$

9. $\vec{u} \times \vec{v} = \langle -5, -31, 27 \rangle$

11. $\vec{u} \times \vec{v} = \langle 0, -2, 0 \rangle$

13. $\vec{i} \times \vec{j} = \vec{k}$

15. $\vec{j} \times \vec{k} = \vec{i}$

17. Answers will vary.

19. 21

21. 5

23. $\sqrt{230}$

25. 6

27. $3\sqrt{30}$

29. 5/2

31. $8\sqrt{7/2}$

33. 15

35. $\pm \frac{1}{\sqrt{21}} \langle -2, 1, 4 \rangle$

37. any unit vector orthogonal to $\vec{u}$ works (such as $\frac{1}{\sqrt{2}} \langle 1, 0, -1 \rangle$).

39. $43.75\sqrt{3} \approx 75.78$ft–lb

41. 11.58ft–lb

43. With $\vec{u} = \langle u_1, u_2, u_3 \rangle$, we have

$$\vec{u} \times \vec{u} = \langle u_2 u_3 - u_3 u_2, -(u_1 u_3 - u_3 u_1), u_1 u_2 - u_2 u_1 \rangle$$
$$= \langle 0, 0, 0 \rangle$$
$$= \vec{0}.$$

Exercises 11.5

1. A point on the line and the direction of the line.

3. parallel, skew

5. vector: $\ell(t) = \langle 2, -4, 1 \rangle + t \langle 9, 2, 5 \rangle$
parametric: $x = 2 + 9t, y = -4 + 2t, z = 1 + 5t$
symmetric: $(x-2)/9 = (y+4)/2 = (z-1)/5$

7. Answers can vary: vector: $\ell(t) = \langle 2, 1, 5 \rangle + t \langle 5, -3, -1 \rangle$
parametric: $x = 2 + 5t, y = 1 - 3t, z = 5 - t$
symmetric: $(x-2)/5 = -(y-1)/3 = -(z-5)$

9. Answers can vary; here the direction is given by $\vec{d_1} \times \vec{d_2}$:
vector: $\ell(t) = \langle 0, 1, 2 \rangle + t \langle -10, 43, 9 \rangle$
parametric: $x = -10t, y = 1 + 43t, z = 2 + 9t$
symmetric: $-x/10 = (y-1)/43 = (z-2)/9$

11. Answers can vary; here the direction is given by $\vec{d_1} \times \vec{d_2}$:
vector: $\ell(t) = \langle 7, 2, -1 \rangle + t \langle 1, -1, 2 \rangle$
parametric: $x = 7 + t, y = 2 - t, z = -1 + 2t$
symmetric: $x - 7 = 2 - y = (z+1)/2$

13. vector: $\ell(t) = \langle 1, 1 \rangle + t \langle 2, 3 \rangle$
parametric: $x = 1 + 2t, y = 1 + 3t$
symmetric: $(x-1)/2 = (y-1)/3$

15. parallel

17. intersecting; $\vec{\ell_1}(3) = \vec{\ell_2}(4) = \langle 9, -5, 13 \rangle$

19. skew

21. same

23. $\sqrt{41}/3$

25. $5\sqrt{2}/2$

27. $3/\sqrt{2}$

29. Since both P and Q are on the line, $\overrightarrow{PQ}$ is parallel to $\vec{d}$. Thus $\overrightarrow{PQ} \times \vec{d} = \vec{0}$, giving a distance of 0.

31.

(a) The distance formula cannot be used because since $\vec{d_1}$ and $\vec{d_2}$ are parallel, $\vec{c}$ is $\vec{0}$ and we cannot divide by $\left\| \vec{0} \right\|$.

(b) Since $\vec{d_1}$ and $\vec{d_2}$ are parallel, $\overrightarrow{P_1 P_2}$ lies in the plane formed by the two lines. Thus $\overrightarrow{P_1 P_2} \times \vec{d_2}$ is orthogonal to this plane, and $\vec{c} = (\overrightarrow{P_1 P_2} \times \vec{d_2}) \times \vec{d_2}$ is parallel to the plane, but still orthogonal to both $\vec{d_1}$ and $\vec{d_2}$. We desire the length of the projection of $\overrightarrow{P_1 P_2}$ onto $\vec{c}$, which is what the formula provides.

(c) Since the lines are parallel, one can measure the distance between the lines at any location on either line (just as to find the distance between straight railroad tracks, one can use a measuring tape anywhere along the track, not just at one specific place.) Let $P = P_1$ and $Q = P_2$ as given by the equations of the lines, and apply the formula for distance between a point and a line.

Exercises 11.6

1. A point in the plane and a normal vector (i.e., a direction orthogonal to the plane).

3. Answers will vary.

5. Answers will vary.

7. Standard form: $3(x-2) - (y-3) + 7(z-4) = 0$
general form: $3x - y + 7z = 31$

9. Answers may vary;
Standard form: $8(x-1) + 4(y-2) - 4(z-3) = 0$
general form: $8x + 4y - 4z = 4$

11. Answers may vary;
Standard form: $-7(x-2) + 2(y-1) + (z-2) = 0$
general form: $-7x + 2y + z = -10$

13. Answers may vary;
Standard form: $2(x-1) - (y-1) = 0$
general form: $2x - y = 1$

15. Answers may vary;
 Standard form: $2(x - 2) - (y + 6) - 4(z - 1) = 0$
 general form: $2x - y - 4z = 6$

17. Answers may vary;
 Standard form: $(x - 5) + (y - 7) + (z - 3) = 0$
 general form: $x + y + z = 15$

19. Answers may vary;
 Standard form: $3(x + 4) + 8(y - 7) - 10(z - 2) = 0$
 general form: $3x + 8y - 10z = 24$

21. Answers may vary:
 $$\ell = \begin{cases} x = 14t \\ y = -1 - 10t \\ z = 2 - 8t \end{cases}$$

23. $(-3, -7, -5)$

25. No point of intersection; the plane and line are parallel.

27. $\sqrt{5/7}$

29. $1/\sqrt{3}$

31. If P is any point in the plane, and Q is also in the plane, then $\overrightarrow{PQ}$ lies parallel to the plane and is orthogonal to $\vec{n}$, the normal vector. Thus $\vec{n} \cdot \overrightarrow{PQ} = 0$, giving the distance as 0.

Exercises 11.7

1. (a) $(4, \frac{\pi}{3}, -1)$; (b) $(\sqrt{17}, \frac{\pi}{3}, 1.816)$

3. (a) $(2\sqrt{7}, \frac{11\pi}{6}, 0)$; (b) $(2\sqrt{7}, \frac{11\pi}{6}, \frac{\pi}{2})$

5. (a) $r^2 + z^2 = 25$; (b) $\rho = 5$

7. (a) $r^2 + 9z^2 = 36$; (b) $\rho^2(1 + 8\cos^2 \phi) = 36$

9.

11.

13.

Chapter 12

Exercises 12.1

1. parametric equations

3. displacement

7.

9.

11.

13.

15.
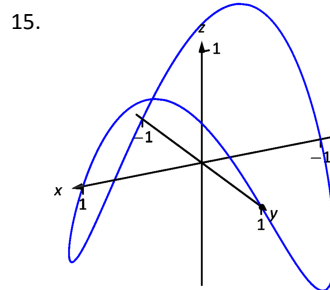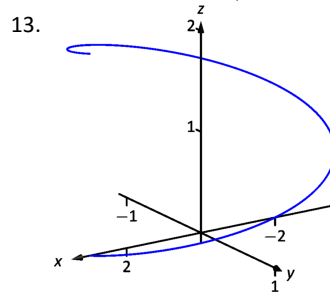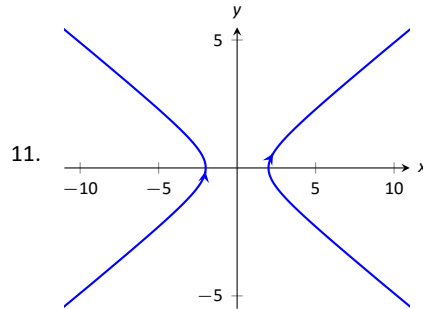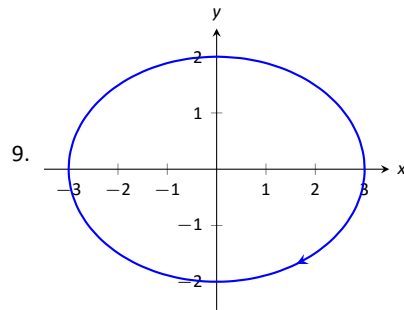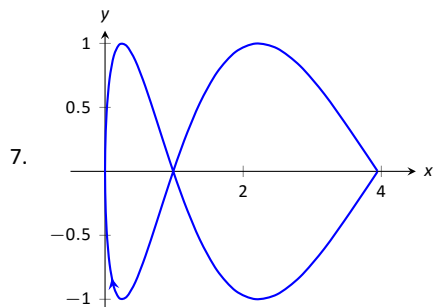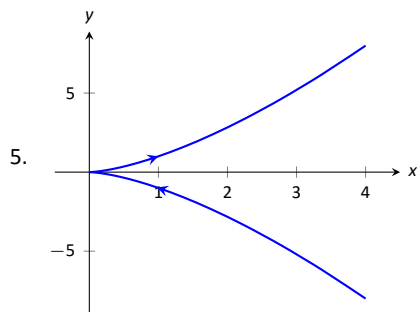
17. $\|\vec{r}(t)\| = \sqrt{25\cos^2 t + 9 \sin^2 t}$.

19. $\|\vec{r}(t)\| = \sqrt{\cos^2 t + t^2 + t^4}$.

21. Answers may vary; three solutions are
 $\vec{r}(t) = \langle 3 \sin t + 5, 3 \cos t + 5 \rangle$,
 $\vec{r}(t) = \langle -3 \cos t + 5, 3 \sin t + 5 \rangle$ and
 $\vec{r}(t) = \langle 3 \cos t + 5, -3 \sin t + 5 \rangle$.

23. Answers may vary, though most direct solutions are
 $\vec{r}(t) = \langle -3 \cos t + 3, 2 \sin t - 2 \rangle$,
 $\vec{r}(t) = \langle 3 \cos t + 3, -2 \sin t - 2 \rangle$ and
 $\vec{r}(t) = \langle 3 \sin t + 3, 2 \cos t - 2 \rangle$.

25. Answers may vary, though most direct solutions are
 $\vec{r}(t) = \langle t, -1/2(t - 1) + 5 \rangle$,
 $\vec{r}(t) = \langle t + 1, -1/2t + 5 \rangle$,
 $\vec{r}(t) = \langle -2t + 1, t + 5 \rangle$ and
 $\vec{r}(t) = \langle 2t + 1, -t + 5 \rangle$.

27. Answers may vary, though most direct solution is
 $\vec{r}(t) = \langle 3 \cos(4\pi t), 3 \sin(4\pi t), 3t \rangle$.

29. $\langle 1, 1 \rangle$

31. $\langle 1, 2, 7 \rangle$

Exercises 12.2

1. component

3. It is difficult to identify the points on the graphs of $\vec{r}(t)$ and $\vec{r}'(t)$ that correspond to each other.

5. $\langle e^3, 0 \rangle$

7. $\langle 2t, 1, 0 \rangle$

9. $(0, \infty)$

11. $\vec{r}'(t) = \langle -1/t^2, 5/(3t+1)^2, \sec^2 t \rangle$

13. $\vec{r}'(t) = \langle 2t, 1 \rangle \cdot \langle \sin t, 2t + 5 \rangle + \langle t^2 + 1, t - 1 \rangle \cdot \langle \cos t, 2 \rangle = (t^2 + 1) \cos t + 2t \sin t + 4t + 3$

15.

$\vec{r}'(t) = \langle 2t + 1, 2t - 1 \rangle$

17.

$\vec{r}'(t) = \langle 2t, 3t^2 - 1 \rangle$
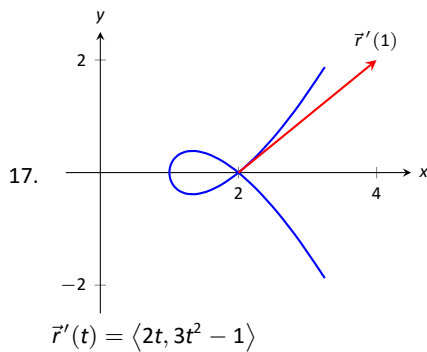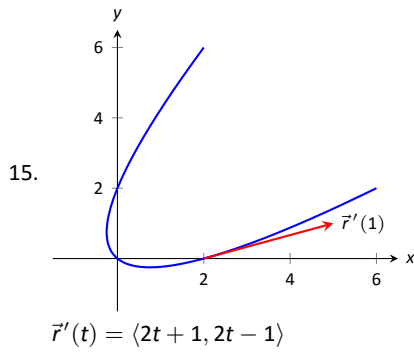
19. $\ell(t) = \langle 2, 0 \rangle + t \langle 3, 1 \rangle$

21. $\ell(t) = \langle -3, 0, \pi \rangle + t \langle 0, -3, 1 \rangle$

23. $t = 2n\pi$, where n is an integer; so $t = \ldots -4\pi, -2\pi, 0, 2\pi, 4\pi, \ldots$

25. $\vec{r}(t)$ is not smooth at $t = 3\pi/4 + n\pi$, where n is an integer

27. Both derivatives return $\langle 5t^4, 4t^3 - 3t^2, 3t^2 \rangle$.

29. Both derivatives return $\langle 2t - e^t - 1, \cos t - 3t^2, (t^2 + 2t)e^t - (t - 1) \cos t - \sin t \rangle$.

31. $\langle \tan^{-1} t, \tan t \rangle + \vec{C}$

33. $\langle 4, -4 \rangle$

35. $\vec{r}(t) = \langle \ln|t + 1| + 1, -\ln|\cos t| + 2 \rangle$

37. $\vec{r}(t) = \langle -\cos t + 1, t - \sin t, e^t - t - 1 \rangle$

39. 10π

41. $\sqrt{2}(1 - e^{-1})$

Exercises 12.3

1. Velocity is a vector, indicating an objects direction of travel and its rate of distance change (i.e., its speed). Speed is a scalar.

3. The average velocity is found by dividing the displacement by the time traveled – it is a vector. The average speed is found by dividing the distance traveled by the time traveled – it is a scalar.

5. One example is traveling at a constant speed s in a circle, ending at the starting position. Since the displacement is $\vec{0}$, the average velocity is $\vec{0}$, hence $\left\| \vec{0} \right\| = 0$. But traveling at constant speed s means the average speed is also $s > 0$.

7. $\vec{v}(t) = \langle 2, 5, 0 \rangle$, $\vec{a}(t) = \langle 0, 0, 0 \rangle$

9. $\vec{v}(t) = \langle -\sin t, \cos t \rangle$, $\vec{a}(t) = \langle -\cos t, -\sin t \rangle$

11. $\vec{v}(t) = \langle 1, \cos t \rangle$, $\vec{a}(t) = \langle 0, -\sin t \rangle$
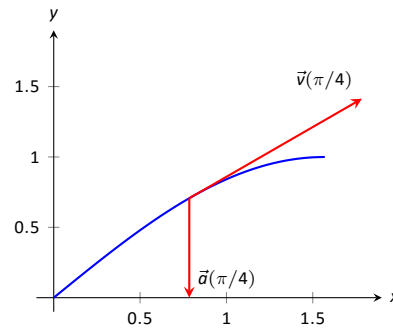
13. $\vec{v}(t) = \langle 2t + 1, -2t + 2 \rangle$, $\vec{a}(t) = \langle 2, -2 \rangle$
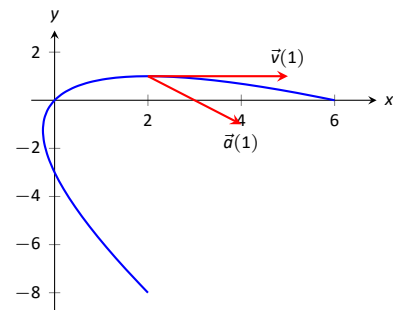
15. $\|\vec{v}(t)\| = \sqrt{4t^2 + 1}$.
Min at $t = 0$; Max at $t = \pm 1$.

17. $\|\vec{v}(t)\| = 5$.
Speed is constant, so there is no difference between min/max

19. $\|\vec{v}(t)\| = |\sec t| \sqrt{\tan^2 t + \sec^2 t}$.
min: $t = 0$; max: $t = \pi/4$

21. $\|\vec{v}(t)\| = 13$.
speed is constant, so there is no difference between min/max

23. $\|\vec{v}(t)\| = \sqrt{4t^2 + 1 + t^2/(1 - t^2)}$.
min: $t = 0$; max: there is no max; speed approaches ∞ as $t \to \pm 1$

25.

(a) $\vec{r}_1(1) = \langle 1, 1 \rangle; \vec{r}_2(1) = \langle 1, 1 \rangle$

(b) $\vec{v}_1(1) = \langle 1, 2 \rangle; \|\vec{v}_1(1)\| = \sqrt{5}; \vec{a}_1(1) = \langle 0, 2 \rangle$
$\vec{v}_2(1) = \langle 2, 4 \rangle; \|\vec{v}_2(1)\| = 2\sqrt{5}; \vec{a}_2(1) = \langle 2, 12 \rangle$

27.

  (a) $\vec{r}_1(2) = \langle 6, 4 \rangle; \vec{r}_2(2) = \langle 6, 4 \rangle$

  (b) $\vec{v}_1(2) = \langle 3, 2 \rangle; \|\vec{v}_1(2)\| = \sqrt{13}; \vec{a}_1(2) = \langle 0, 0 \rangle$
     $\vec{v}_2(2) = \langle 6, 4 \rangle; \|\vec{v}_2(2)\| = 2\sqrt{13}; \vec{a}_2(2) = \langle 0, 0 \rangle$

29. $\vec{v}(t) \quad = \quad \langle 2t + 1, 3t + 2 \rangle, \quad \vec{r}(t) \quad =$
   $\langle t^2 + t + 5, 3t^2/2 + 2t - 2 \rangle$

31. $\vec{v}(t) = \langle \sin t, \cos t \rangle, \vec{r}(t) = \langle 1 - \cos t, \sin t \rangle$

33. Displacement: $\langle 0, 0, 6\pi \rangle$; distance traveled: $2\sqrt{13}\pi \approx$ 22.65ft; average velocity: $\langle 0, 0, 3 \rangle$; average speed: $\sqrt{13} \approx$ 3.61ft/s

35. Displacement: $\langle 0, 0 \rangle$; distance traveled: $2\pi \approx 6.28$ft; average velocity: $\langle 0, 0 \rangle$; average speed: 1ft/s

37. At t-values of $\sin^{-1}(9/30)/(4\pi) + n/2 \approx 0.024 + n/2$ seconds, where n is an integer.

39.

  (a) Holding the crossbow at an angle of 0.013 radians, $\approx 0.745°$ will hit the target 0.4s later. (Another solution exists, with an angle of $89°$, landing 18.75s later, but this is impractical.)

  (b) In the .4 seconds the arrow travels, a deer, traveling at 20mph or 29.33ft/s, can travel 11.7ft. So she needs to lead the deer by 11.7ft.

41. The position function is $\vec{r}(t) = \langle 220t, -16t^2 + 1000 \rangle$. The y-component is 0 when $t = 7.9$; $\vec{r}(7.9) = \langle 1739.25, 0 \rangle$, meaning the box will travel about 1740ft horizontally before it lands.

Exercises 12.4

1. 1

3. $\vec{T}(t)$ and $\vec{N}(t)$.

5. $\vec{T}(t) = \left\langle \dfrac{4t}{\sqrt{20t^2 - 4t + 1}}, \dfrac{2t - 1}{\sqrt{20t^2 - 4t + 1}} \right\rangle$;
  $\vec{T}(1) = \langle 4/\sqrt{17}, 1/\sqrt{17} \rangle$

7. $\vec{T}(t) = \dfrac{\cos t \sin t}{\sqrt{\cos^2 t \sin^2 t}} \langle -\cos t, \sin t \rangle$.   (Be careful; this cannot be simplified as just $\langle -\cos t, \sin t \rangle$ as $\sqrt{\cos^2 t \sin^2 t} \neq \cos t \sin t$, but rather $|\cos t \sin t|$.)
  $\vec{T}(\pi/4) = \langle -\sqrt{2}/2, \sqrt{2}/2 \rangle$

9. $\ell(t) = \langle 2, 0 \rangle + t \langle 4/\sqrt{17}, 1/\sqrt{17} \rangle$; in parametric form,
  $\ell(t) = \begin{cases} x = 2 + 4t/\sqrt{17} \\ y = t/\sqrt{17} \end{cases}$

11. $\ell(t) = \langle \sqrt{2}/4, \sqrt{2}/4 \rangle + t \langle -\sqrt{2}/2, \sqrt{2}/2 \rangle$; in parametric form,
  $\ell(t) = \begin{cases} x = \sqrt{2}/4 - \sqrt{2}t/2 \\ y = \sqrt{2}/4 + \sqrt{2}t/2 \end{cases}$

13. $\vec{T}(t) = \langle -\sin t, \cos t \rangle; \vec{N}(t) = \langle -\cos t, -\sin t \rangle$

15. $\vec{T}(t) = \left\langle -\dfrac{\sin t}{\sqrt{4\cos^2 t + \sin^2 t}}, \dfrac{2\cos t}{\sqrt{4\cos^2 t + \sin^2 t}} \right\rangle$;
  $\vec{N}(t) = \left\langle -\dfrac{2\cos t}{\sqrt{4\cos^2 t + \sin^2 t}}, -\dfrac{\sin t}{\sqrt{4\cos^2 t + \sin^2 t}} \right\rangle$

17.

  (a) Be sure to show work

  (b) $\vec{N}(\pi/4) = \langle -5/\sqrt{34}, -3/\sqrt{34} \rangle$

19.

  (a) Be sure to show work

  (b) $\vec{N}(0) = \left\langle -\dfrac{1}{\sqrt{5}}, \dfrac{2}{\sqrt{5}} \right\rangle$

21. $\vec{T}(t) = \dfrac{1}{\sqrt{5}} \langle 2, \cos t, -\sin t \rangle; \vec{N}(t) = \langle 0, -\sin t, -\cos t \rangle$

23. $\vec{T}(t) = \dfrac{1}{\sqrt{a^2 + b^2}} \langle -a \sin t, a \cos t, b \rangle$;
  $\vec{N}(t) = \langle -\cos t, -\sin t, 0 \rangle$

25. $a_{\mathsf{T}} = \dfrac{4t}{\sqrt{1 + 4t^2}}$ and $a_{\mathsf{N}} = \sqrt{4 - \dfrac{16t^2}{1 + 4t^2}}$
  At $t = 0$, $a_{\mathsf{T}} = 0$ and $a_{\mathsf{N}} = 2$;
  At $t = 1$, $a_{\mathsf{T}} = 4/\sqrt{5}$ and $a_{\mathsf{N}} = 2/\sqrt{5}$.
  At $t = 0$, all acceleration comes in the form of changing the direction of velocity and not the speed; at $t = 1$, more acceleration comes in changing the speed than in changing direction.

27. $a_{\mathsf{T}} = 0$ and $a_{\mathsf{N}} = 2$
  At $t = 0$, $a_{\mathsf{T}} = 0$ and $a_{\mathsf{N}} = 2$;
  At $t = \pi/2$, $a_{\mathsf{T}} = 0$ and $a_{\mathsf{N}} = 2$.
  The object moves at constant speed, so all acceleration comes from changing direction, hence $a_{\mathsf{T}} = 0$. $\vec{a}(t)$ is always parallel to $\vec{N}(t)$, but twice as long, hence $a_{\mathsf{N}} = 2$.

29. $a_{\mathsf{T}} = 0$ and $a_{\mathsf{N}} = a$
  At $t = 0$, $a_{\mathsf{T}} = 0$ and $a_{\mathsf{N}} = a$;
  At $t = \pi/2$, $a_{\mathsf{T}} = 0$ and $a_{\mathsf{N}} = a$.
  The object moves at constant speed, meaning that a_{T} is always 0. The object "rises" along the z-axis at a constant rate, so all acceleration comes in the form of changing direction circling the z-axis. The greater the radius of this circle the greater the acceleration, hence $a_{\mathsf{N}} = a$.

Exercises 12.5

1. time and/or distance

3. Answers may include lines, circles, helixes

5. κ

7. $s = 3t$, so $\vec{r}(s) = \langle 2s/3, s/3, -2s/3 \rangle$

9. $s = \sqrt{13}t$, so
  $\vec{r}(s) = \langle 3\cos(s/\sqrt{13}), 3\sin(s/\sqrt{13}), 2s/\sqrt{13} \rangle$

11. $\kappa = \dfrac{|6x|}{\left(1 + (3x^2 - 1)^2\right)^{3/2}}$;
  $\kappa(0) = 0, \kappa(1/2) = \dfrac{192}{17\sqrt{17}} \approx 2.74$.

13. $\kappa = \dfrac{|\cos x|}{\left(1 + \sin^2 x\right)^{3/2}}$;
  $\kappa(0) = 1, \kappa(\pi/2) = 0$

15. $\kappa = \dfrac{|2\cos t \cos(2t) + 4\sin t \sin(2t)|}{\left(4\cos^2(2t) + \sin^2 t\right)^{3/2}}$;
  $\kappa(0) = 1/4, \kappa(\pi/4) = 8$

17. $\kappa = \dfrac{|6t^2 + 2|}{\left(4t^2 + (3t^2 - 1)^2\right)^{3/2}}$;
  $\kappa(0) = 2, \kappa(5) = \dfrac{19}{1394\sqrt{1394}} \approx 0.0004$

19. $\kappa = 0$;
  $\kappa(0) = 0, \kappa(1) = 0$

21. $\kappa = \dfrac{3}{13}$;
  $\kappa(0) = 3/13, \kappa(\pi/2) = 3/13$

23. maximized at $x = \pm\dfrac{\sqrt{2}}{\sqrt[4]{5}}$

25. maximized at $t = 1/4$

27. radius of curvature is $5\sqrt{5}/4$.

29. radius of curvature is 9.

31. $x^2 + (y - 1/2)^2 = 1/4$, or
$\vec{c}(t) = \langle 1/2 \cos t, 1/2 \sin t + 1/2 \rangle$

33. $x^2 + (y + 8)^2 = 81$, or $\vec{c}(t) = \langle 9 \cos t, 9 \sin t - 8 \rangle$

35. Let $\vec{r}(t) = \langle x(t), y(t), 0 \rangle$ and apply the second formula of part 3.

Chapter 13

Exercises 13.1

1. Answers will vary.

3. topographical

5. surface

7. domain: $\mathbb{R}^2$
range: $z \geq 2$

9. domain: $\mathbb{R}^2$
range: $\mathbb{R}$

11. domain: $\mathbb{R}^2$
range: $0 < z \leq 1$

13. domain: $\{(x, y) \mid x^2 + y^2 \leq 9\}$, i.e., the domain is the circle and interior of a circle centered at the origin with radius 3.
range: $0 \leq z \leq 3$

15. Level curves are lines $y = (3/2)x - c/2$.
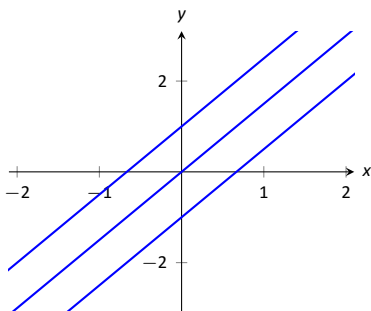
17. Level curves are parabolas $x = y^2 + c$.
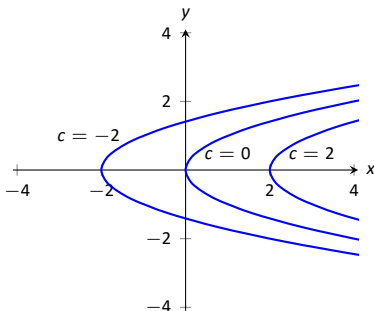
19. Level curves are circles, centered at $(1/c, -1/c)$ with radius $\sqrt{2/c^2 - 1}$. When $c = 0$, the level curve is the line $y = x$.
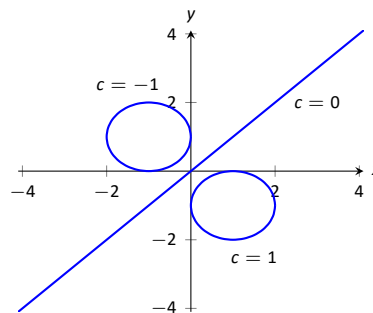
21. Level curves are ellipses of the form $\frac{x^2}{c^2} + \frac{y^2}{c^2/4} = 1$, i.e., $a = c$ and $b = c/2$.
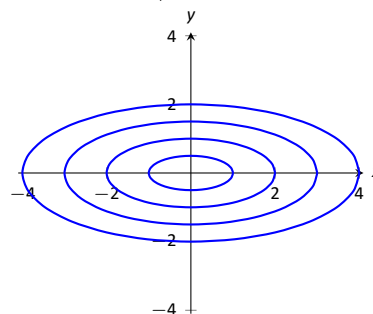
23. domain: $x + 2y - 4z \neq 0$; the set of points in $\mathbb{R}^3$ NOT in the domain form a plane through the origin.
range: $\mathbb{R}$

25. domain: $z \geq x^2 - y^2$; the set of points in $\mathbb{R}^3$ above (and including) the hyperbolic paraboloid $z = x^2 - y^2$.
range: $[0, \infty)$

27. The level surfaces are spheres, centered at the origin, with radius $\sqrt{c}$.

29. The level surfaces are paraboloids of the form $z = \frac{x^2}{c} + \frac{y^2}{c}$; the larger c, the "wider" the paraboloid.

31. The level curves for each surface are similar; for $z = \sqrt{x^2 + 4y^2}$ the level curves are ellipses of the form $\frac{x^2}{c^2} + \frac{y^2}{c^2/4} = 1$, i.e., $a = c$ and $b = c/2$; whereas for $z = x^2 + 4y^2$ the level curves are ellipses of the form $\frac{x^2}{c} + \frac{y^2}{c/4} = 1$, i.e., $a = \sqrt{c}$ and $b = \sqrt{c}/2$. The first set of ellipses are spaced evenly apart, meaning the function grows at a constant rate; the second set of ellipses are more closely spaced together as c grows, meaning the function grows faster and faster as c increases.
The function $z = \sqrt{x^2 + 4y^2}$ can be rewritten as $z^2 = x^2 + 4y^2$, an elliptic cone; the function $z = x^2 + 4y^2$ is a paraboloid, each matching the description above.

Exercises 13.2

1. Answers will vary.

3. Answers will vary.
One possible answer: $\{(x, y) | x^2 + y^2 \leq 1\}$

5. Answers will vary.
One possible answer: $\{(x, y) | x^2 + y^2 < 1\}$

7.

  (a) Answers will vary.
      interior point: $(1, 3)$
      boundary point: $(3, 3)$

  (b) S is a closed set

  (c) S is bounded

9.

  (a) Answers will vary.
      interior point: none
      boundary point: $(0, -1)$

  (b) S is a closed set, consisting only of boundary points

  (c) S is bounded

11.

  (a) $D = \{(x, y) \mid 9 - x^2 - y^2 \geq 0\}$.

  (b) D is a closed set.

  (c) D is bounded.

13.

  (a) $D = \{(x, y) \mid y > x^2\}$.

  (b) D is an open set.

  (c) D is unbounded.

15.

  (a) Along $y = 0$, the limit is 1.

  (b) Along $x = 0$, the limit is -1.

  Since the above limits are not equal, the limit does not exist.

17.

  (a) Along $y = mx$, the limit is 0.

  (b) Along $x = 0$, the limit is -1.

  Since the above limits are not equal, the limit does not exist.

19.

  (a) Along $y = 2$, the limit is:

$$\lim_{(x,y)\to(1,2)} \frac{x+y-3}{x^2-1} = \lim_{x\to1} \frac{x-1}{x^2-1}$$

$$= \lim_{x\to1} \frac{1}{x+1}$$

$$= 1/2.$$

  (b) Along $y = x + 1$, the limit is:

$$\lim_{(x,y)\to(1,2)} \frac{x+y-3}{x^2-1} = \lim_{x\to1} \frac{2(x-1)}{x^2-1}$$

$$= \lim_{x\to1} \frac{2}{x+1}$$

$$= 1.$$

Since the limits along the lines $y = 2$ and $y = x + 1$ differ, the overall limit does not exist.

Exercises 13.3

1. A constant is a number that is added or subtracted in an expression; a coefficient is a number that is being multiplied by a nonconstant function.

3. f_x

5. $f_x = 2xy - 1, f_y = x^2 + 2$
   $f_x(1, 2) = 3, f_y(1, 2) = 3$

7. $f_x = -\sin x \sin y, f_y = \cos x \cos y$
   $f_x(\pi/3, \pi/3) = -3/4, f_y(\pi/3, \pi/3) = 1/4$

9. $f_x = 2xy + 6x, f_y = x^2 + 4$
   $f_{xx} = 2y + 6, f_{yy} = 0$
   $f_{xy} = 2x, f_{yx} = 2x$

11. $f_x = 1/y, f_y = -x/y^2$
    $f_{xx} = 0, f_{yy} = 2x/y^3$
    $f_{xy} = -1/y^2, f_{yx} = -1/y^2$

13. $f_x = 2xe^{x^2+y^2}, f_y = 2ye^{x^2+y^2}$
    $f_{xx} = 2e^{x^2+y^2} + 4x^2 e^{x^2+y^2}, f_{yy} = 2e^{x^2+y^2} + 4y^2 e^{x^2+y^2}$
    $f_{xy} = 4xye^{x^2+y^2}, f_{yx} = 4xye^{x^2+y^2}$

15. $f_x = \cos x \cos y, f_y = -\sin x \sin y$
    $f_{xx} = -\sin x \cos y, f_{yy} = -\sin x \cos y$
    $f_{xy} = -\sin y \cos x, f_{yx} = -\sin y \cos x$

17. $f_x = -5y^3 \sin(5xy^3), f_y = -15xy^2 \sin(5xy^3)$
    $f_{xx} = -25y^6 \cos(5xy^3), f_{yy} = -225x^2y^4 \cos(5xy^3) - 30xy \sin(5xy^3)$
    $f_{xy} = -75xy^5 \cos(5xy^3) - 15y^2 \sin(5xy^3), f_{yx} = -75xy^5 \cos(5xy^3) - 15y^2 \sin(5xy^3)$

19. $f_x = \frac{2y^2}{\sqrt{4xy^2+1}}, f_y = \frac{4xy}{\sqrt{4xy^2+1}}$
    $f_{xx} = -\frac{4y^4}{\sqrt{4xy^2+1}^3}, f_{yy} = -\frac{16x^2y^2}{\sqrt{4xy^2+1}^3} + \frac{4x}{\sqrt{4xy^2+1}}$
    $f_{xy} = -\frac{8xy^3}{\sqrt{4xy^2+1}^3} + \frac{4y}{\sqrt{4xy^2+1}}, f_{yx} = -\frac{8xy^3}{\sqrt{4xy^2+1}^3} + \frac{4y}{\sqrt{4xy^2+1}}$

21. $f_x = -\frac{2x}{(x^2+y^2+1)^2}, f_y = -\frac{2y}{(x^2+y^2+1)^2}$
    $f_{xx} = \frac{8x^2}{(x^2+y^2+1)^3} - \frac{2}{(x^2+y^2+1)^2}, f_{yy} = \frac{8y^2}{(x^2+y^2+1)^3} - \frac{2}{(x^2+y^2+1)^2}$
    $f_{xy} = \frac{8xy}{(x^2+y^2+1)^3}, f_{yx} = \frac{8xy}{(x^2+y^2+1)^3}$

23. $f_x = 6x, f_y = 0$
    $f_{xx} = 6, f_{yy} = 0$
    $f_{xy} = 0, f_{yx} = 0$

25. $f_x = \frac{1}{4xy}, f_y = -\frac{\ln x}{4y^2}$
    $f_{xx} = -\frac{1}{4x^2y}, f_{yy} = \frac{\ln x}{2y^3}$
    $f_{xy} = -\frac{1}{4xy^2}, f_{yx} = -\frac{1}{4xy^2}$

27. $f(x, y) = x \sin y + x + C$, where C is any constant.

29. $f(x, y) = 3x^2y - 4xy^2 + 2y + C$, where C is any constant.

31. $f_x = 2xe^{2y-3z}, f_y = 2x^2 e^{2y-3z}, f_z = -3x^2 e^{2y-3z}$
    $f_{yz} = -6x^2 e^{2y-3z}, f_{zy} = -6x^2 e^{2y-3z}$

33. $f_x = \frac{3}{7y^2z}, f_y = -\frac{6x}{7y^3z}, f_z = -\frac{3x}{7y^2z^2}$
    $f_{yz} = \frac{6x}{7y^3z^2}, f_{zy} = \frac{6x}{7y^3z^2}$

Exercises 13.4

1. T

3. T

5. $dz = (\sin y + 2x)dx + (x\cos y)dy$

7. $dz = 5dx - 7dy$

9. $dz = \frac{x}{\sqrt{x^2+y}}dx + \frac{1}{2\sqrt{x^2+y}}dy$, with $dx = -0.05$ and $dy = .1$.
At $(3,7)$, $dz = 3/4(-0.05) + 1/8(.1) = -0.025$, so $f(2.95, 7.1) \approx -0.025 + 4 = 3.975$.

11. $dz = (2xy - y^2)dx + (x^2 - 2xy)dy$, with $dx = 0.04$ and $dy = 0.06$. At $(2,3)$, $dz = 3(0.04) + (-8)(0.06) = -0.36$, so $f(2.04, 3.06) \approx -0.36 - 6 = -6.36$.

13. The total differential of volume is $dV = 4\pi dr + \pi dh$. The coefficient of dr is greater than the coefficient of dh, so the volume is more sensitive to changes in the radius.

15. Using trigonometry, $\ell = x\tan\theta$, so $d\ell = \tan\theta dx + x\sec^2\theta d\theta$. With $\theta = 85°$ and $x = 30$, we have $d\ell = 11.43dx + 3949.38d\theta$. The measured length of the wall is much more sensitive to errors in θ than in x. While it can be difficult to compare sensitivities between measuring feet and measuring degrees (it is somewhat like "comparing apples to oranges"), here the coefficients are so different that the result is clear: a small error in degree has a much greater impact than a small error in distance.

17. $dw = 2xyz^3\,dx + x^2z^3\,dy + 3x^2yz^2\,dz$

19. $dx = 0.05$, $dy = -0.1$. $dz = 9(.05) + (-2)(-0.1) = 0.65$. So $f(3.05, 0.9) \approx 7 + 0.65 = 7.65$.

21. $dx = 0.5$, $dy = 0.1$, $dz = -0.2$.
$dw = 2(0.5) + (-3)(0.1) + 3.7(-0.2) = -0.04$, so $f(2.5, 4.1, 4.8) \approx -1 - 0.04 = -1.04$.

23. Everywhere except the origin.

Exercises 13.5

1. Because the parametric equations describe a level curve, z is constant for all t. Therefore $\frac{dz}{dt} = 0$.

3. $\frac{dx}{dt}$, and $\frac{\partial f}{\partial y}$

5. F

7.
   (a) $\frac{dz}{dt} = 3(2t) + 4(2) = 6t + 8$.
   (b) At $t = 1$, $\frac{dz}{dt} = 14$.

9.
   (a) $\frac{dz}{dt} = 5(-2\sin t) + 2(\cos t) = -10\sin t + 2\cos t$
   (b) At $t = \pi/4$, $\frac{dz}{dt} = -4\sqrt{2}$.

11.
   (a) $\frac{dz}{dt} = 2x(\cos t) + 4y(3\cos t)$.
   (b) At $t = \pi/4$, $x = \sqrt{2}/2$, $y = 3\sqrt{2}/2$, and $\frac{dz}{dt} = 19$.

13. $t = -4/3$; this corresponds to a minimum

15. $t = \tan^{-1}(1/5) + n\pi$, where n is an integer

17. We find that
$$\frac{dz}{dt} = 38\cos t\sin t.$$
Thus $\frac{dz}{dt} = 0$ when $t = \pi n$ or $\pi n + \pi/2$, where n is any integer.

19.
   (a) $\frac{\partial z}{\partial s} = 2xy(1) + x^2(2) = 2xy + 2x^2$;
   $\frac{\partial z}{\partial t} = 2xy(-1) + x^2(4) = -2xy + 4x^2$
   (b) With $s = 1$, $t = 0$, $x = 1$ and $y = 2$. Thus $\frac{\partial z}{\partial s} = 6$ and $\frac{\partial z}{\partial t} = 0$

21.
   (a) $\frac{\partial z}{\partial s} = 2x(\cos t) + 2y(\sin t) = 2x\cos t + 2y\sin t$;
   $\frac{\partial z}{\partial t} = 2x(-s\sin t) + 2y(s\cos t) = -2xs\sin t + 2ys\cos t$
   (b) With $s = 2$, $t = \pi/4$, $x = \sqrt{2}$ and $y = \sqrt{2}$. Thus $\frac{\partial z}{\partial s} = 4$ and $\frac{\partial z}{\partial t} = 0$

23. $f_x = 2x\tan y$, $f_y = x^2\sec^2 y$;
$$\frac{dy}{dx} = -\frac{2\tan y}{x\sec^2 y}$$

25. $f_x = \frac{(x+y^2)(2x) - (x^2+y)(1)}{(x+y^2)^2}$,
$f_y = \frac{(x+y^2)(1) - (x^2+y)(2y)}{(x+y^2)^2}$;
$\frac{dy}{dx} = -\frac{2x(x+y^2) - (x^2+y)}{x+y^2 - 2y(x^2+y)}$

27. $\frac{dz}{dt} = 2(4) + 1(-5) = 3$.

29. $\frac{\partial z}{\partial s} = -4(5) + 9(-2) = -38$,
$\frac{\partial z}{\partial t} = -4(7) + 9(6) = 26$.

Exercises 13.6

1. A partial derivative is essentially a special case of a directional derivative; it is the directional derivative in the direction of x or y, i.e., $\langle 1, 0\rangle$ or $\langle 0, 1\rangle$.

3. $\vec{u} = \langle 0, 1\rangle$

5. maximal, or greatest

7. $\nabla f = \langle -2xy + y^2 + y, -x^2 + 2xy + x\rangle$

9. $\nabla f = \left\langle \frac{-2x}{(x^2+y^2+1)^2}, \frac{-2y}{(x^2+y^2+1)^2}\right\rangle$

11. $\nabla f = \langle 2x - y - 7, 4y - x\rangle$

13. $\nabla f = \langle -2xy + y^2 + y, -x^2 + 2xy + x\rangle$; $\nabla f(2, 1) = \langle -2, 2\rangle$. Be sure to change all directions to unit vectors.
   (a) $2/5$ ($\vec{u} = \langle 3/5, 4/5\rangle$)
   (b) $-2/\sqrt{5}$ ($\vec{u} = \langle -1/\sqrt{5}, -2/\sqrt{5}\rangle$)

15. $\nabla f = \left\langle \frac{-2x}{(x^2+y^2+1)^2}, \frac{-2y}{(x^2+y^2+1)^2}\right\rangle$; $\nabla f(1, 1) = \langle -2/9, -2/9\rangle$. Be sure to change all directions to unit vectors.
   (a) 0 ($\vec{u} = \langle 1/\sqrt{2}, -1/\sqrt{2}\rangle$)
   (b) $2\sqrt{2}/9$ ($\vec{u} = \langle -1/\sqrt{2}, -1/\sqrt{2}\rangle$)

17. $\nabla f = \langle 2x - y - 7, 4y - x\rangle$; $\nabla f(4, 1) = \langle 0, 0\rangle$.
   (a) 0
   (b) 0

19. $\nabla f = \langle -2xy + y^2 + y, -x^2 + 2xy + x \rangle$

    (a) $\nabla f(2,1) = \langle -2, 2 \rangle$

    (b) $\|\nabla f(2,1)\| = \|\langle -2, 2 \rangle\| = \sqrt{8}$

    (c) $\langle 2, -2 \rangle$

    (d) $\langle 1/\sqrt{2}, 1/\sqrt{2} \rangle$

21. $\nabla f = \left\langle \frac{-2x}{(x^2+y^2+1)^2}, \frac{-2y}{(x^2+y^2+1)^2} \right\rangle$

    (a) $\nabla f(1,1) = \langle -2/9, -2/9 \rangle$.

    (b) $\|\nabla f(1,1)\| = \|\langle -2/9, -2/9 \rangle\| = 2\sqrt{2}/9$

    (c) $\langle 2/9, 2/9 \rangle$

    (d) $\langle 1/\sqrt{2}, -1/\sqrt{2} \rangle$

23. $\nabla f = \langle 2x - y - 7, 4y - x \rangle$

    (a) $\nabla f(4,1) = \langle 0, 0 \rangle$

    (b) 0

    (c) $\langle 0, 0 \rangle$

    (d) All directions give a directional derivative of 0.

25.

    (a) $\nabla F(x,y,z) = \langle 6xz^3 + 4y, 4x, 9x^2z^2 - 6z \rangle$

    (b) $113/\sqrt{3}$

27.

    (a) $\nabla F(x,y,z) = \langle 2xy^2, 2y(x^2 - z^2), -2y^2z \rangle$

    (b) 0

Exercises 13.7

1. Answers will vary. The displacement of the vector is one unit in the x-direction and 3 units in the z-direction, with no change in y. Thus along a line parallel to $\vec{v}$, the change in z is 3 times the change in x – i.e., a "slope" of 3. Specifically, the line in the x-z plane parallel to z has a slope of 3.

3. T

5.

    (a) $\ell_x(t) = \begin{cases} x = 2 + t \\ y = 3 \\ z = -48 - 12t \end{cases}$

    (b) $\ell_y(t) = \begin{cases} x = 2 \\ y = 3 + t \\ z = -48 - 40t \end{cases}$

    (c) $\ell_{\vec{u}}(t) = \begin{cases} x = 2 + t/\sqrt{10} \\ y = 3 + 3t/\sqrt{10} \\ z = -48 - 66\sqrt{2/5}t \end{cases}$

7.

    (a) $\ell_x(t) = \begin{cases} x = 4 + t \\ y = 2 \\ z = 2 + 3t \end{cases}$

    (b) $\ell_y(t) = \begin{cases} x = 4 \\ y = 2 + t \\ z = 2 - 5t \end{cases}$

    (c) $\ell_{\vec{u}}(t) = \begin{cases} x = 4 + t/\sqrt{2} \\ y = 2 + t/\sqrt{2} \\ z = 2 - \sqrt{2}t \end{cases}$

9. $\ell_{\vec{n}}(t) = \begin{cases} x = 2 - 12t \\ y = 3 - 40t \\ z = -48 - t \end{cases}$

11. $\ell_{\vec{n}}(t) = \begin{cases} x = 4 + 3t \\ y = 2 - 5t \\ z = 2 - t \end{cases}$

13. $(1.425, 1.085, -48.078)$, $(2.575, 4.915, -47.952)$

15. $(5.014, 0.31, 1.662)$ and $(2.986, 3.690, 2.338)$

17. $-12(x - 2) - 40(y - 3) - (z + 48) = 0$

19. $3(x - 4) - 5(y - 2) - (z - 2) = 0$ (Note that this tangent plane is the same as the original function, a plane.)

21. $\nabla F = \langle x/4, y/2, z/8 \rangle$; at P, $\nabla F = \langle 1/4, \sqrt{2}/2, \sqrt{6}/8 \rangle$

    (a) $\ell_{\vec{n}}(t) = \begin{cases} x = 1 + t/4 \\ y = \sqrt{2} + \sqrt{2}t/2 \\ z = \sqrt{6} + \sqrt{6}t/8 \end{cases}$

    (b) $\frac{1}{4}(x - 1) + \frac{\sqrt{2}}{2}(y - \sqrt{2}) + \frac{\sqrt{6}}{8}(z - \sqrt{6}) = 0$.

23. $\nabla F = \langle y^2 - z^2, 2xy, -2xz \rangle$; at P, $\nabla F = \langle 0, 4, 4 \rangle$

    (a) $\ell_{\vec{n}}(t) = \begin{cases} x = 2 \\ y = 1 + 4t \\ z = -1 + 4t \end{cases}$

    (b) $4(y - 1) + 4(z + 1) = 0$.

Exercises 13.8

1. F; it is the "other way around."

3. T

5. One critical point at $(-4, 2)$; $f_{xx} = 1$ and $D = 4$, so this point corresponds to a relative minimum.

7. One critical point at $(6, -3)$; $D = -4$, so this point corresponds to a saddle point.

9. Two critical points: at $(0, -1)$; $f_{xx} = 2$ and $D = -12$, so this point corresponds to a saddle point; at $(0, 1)$, $f_{xx} = 2$ and $D = 12$, so this corresponds to a relative minimum.

11. Critical points when x or y are 0. $D = -12x^2y^2$, so the test is inconclusive. (Some elementary thought shows that these are absolute minima.)

13. One critical point: $f_x = 0$ when $x = 3$; $f_y = 0$ when $y = 0$, so one critical point at $(3, 0)$, which is a relative maximum, where $f_{xx} = \frac{y^2 - 16}{(16 - (x-3)^2 - y^2)^{3/2}}$ and $D = \frac{16}{(16 - (x-3)^2 - y^2)^2}$. Both f_x and f_y are undefined along the circle $(x-3)^2 + y^2 = 16$; at any point along this curve, $f(x, y) = 0$, the absolute minimum of the function.

15. The triangle is bound by the lines $y = -1$, $y = 2x + 1$ and $y = -2x + 1$.
    Along $y = -1$, there is a critical point at $(0, -1)$.
    Along $y = 2x + 1$, there is a critical point at $(-3/5, -1/5)$.
    Along $y = -2x + 1$, there is a critical point at $(3/5, -1/5)$.
    The function f has one critical point, irrespective of the constraint, at $(0, -1/2)$.
    Checking the value of f at these four points, along with the three vertices of the triangle, we find the absolute maximum is at $(0, 1, 3)$ and the absolute minimum is at $(0, -1/2, 3/4)$.

17. The region has no "corners" or "vertices," just a smooth edge.
    To find critical points along the circle $x^2 + y^2 = 4$, we solve for y^2: $y^2 = 4 - x^2$. We can go further and state $y = \pm\sqrt{4 - x^2}$.
    We can rewrite f as $f(x) = x^2 + 2x + (4 - x^2) + \sqrt{4 - x^2} = 2x + 4 + \sqrt{4 - x^2}$. (We will return and use $-\sqrt{4 - x^2}$ later.)
    Solving $f'(x) = 0$, we get $x = \sqrt{2} \Rightarrow y = \sqrt{2}$. $f'(x)$ is also undefined at $x = \pm 2$, where $y = 0$.
    Using $y = -\sqrt{4 - x^2}$, we rewrite $f(x, y)$ as $f(x) = 2x + 4 - \sqrt{4 - x^2}$. Solving $f'(x) = 0$, we get $x = -\sqrt{2}$, $y = -\sqrt{2}$.
    The function f itself has a critical point at $(-1, -1)$.
    Checking the value of f at $(-1, -1)$, $(\sqrt{2}, \sqrt{2})$, $(-\sqrt{2}, -\sqrt{2})$, $(2, 0)$ and $(-2, 0)$, we find the absolute maximum is at $(\sqrt{2}, \sqrt{2}, 4 + 4\sqrt{2})$ and the absolute minimum is at $(-1, -1, -2)$.

Exercises 13.9

1. $\pm 2\sqrt{5}$ at $(\pm 4/\sqrt{5}, \pm 2/\sqrt{5})$

3. $(\pm 20/\sqrt{13}, \pm 30/\sqrt{13})$

5. $8abc/3\sqrt{3}$

7. Length $130/3$, height and width $65/3$.

9. $(0, \pm 1, 0)$

11. $(2, 1, 2)$

13. Max: 5 at $\pm(2, 2)$, min: $-9/2$ at $\pm(3/\sqrt{2}, -3/\sqrt{2})$.

15. $(2/\sqrt{3})^3 = 8/3\sqrt{3}$

Chapter 14

Exercises 14.1

1. $C(y)$, meaning that instead of being just a constant, like the number 5, it is a function of y, which acts like a constant when taking derivatives with respect to x.

3. curve to curve, then from point to point

5.
   (a) $18x^2 + 42x - 117$
   (b) -108

7.
   (a) $x^4/2 - x^2 + 2x - 3/2$
   (b) $23/15$

9.
   (a) $\sin^2 y$
   (b) $\pi/2$

11. $\int_1^4 \int_{-2}^1 dy\, dx$ and $\int_{-2}^1 \int_1^4 dx\, dy$.
    area of $R = 9$ units^2

13. $\int_2^4 \int_{x-1}^{7-x} dy\, dx$. The order $dx\, dy$ needs two iterated integrals as x is bounded above by two different functions. This gives:
    $$\int_1^3 \int_2^{y+1} dx\, dy + \int_3^5 \int_2^{7-y} dx\, dy.$$
    area of $R = 4$ units^2

15. $\int_0^1 \int_{x^4}^{\sqrt{x}} dy\, dx$ and $\int_0^1 \int_{y^2}^{\sqrt[4]{y}} dx\, dy$
    area of $R = 7/15$ units^2

17.

$y = 4 - x^2$
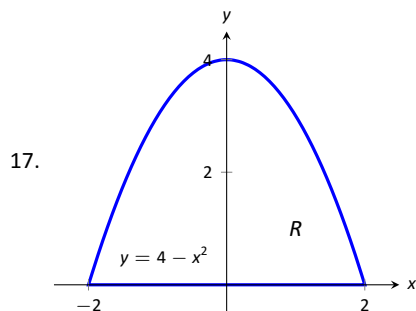
area of $R = \int_0^4 \int_{-\sqrt{4-y}}^{\sqrt{4-y}} dx\, dy$

19.

$x^2/16 + y^2/4 = 1$
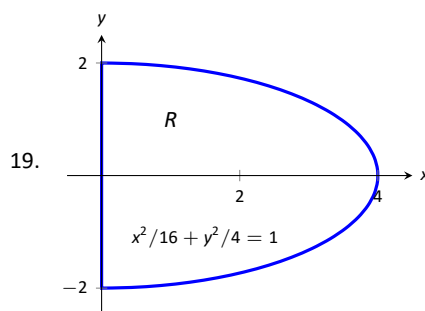
area of $R = \int_0^4 \int_{-\sqrt{4-x^2/4}}^{\sqrt{4-x^2/4}} dy\, dx$

21.

area of $R = \int_{-1}^{2} \int_{x^2}^{x+2} dy\, dx$
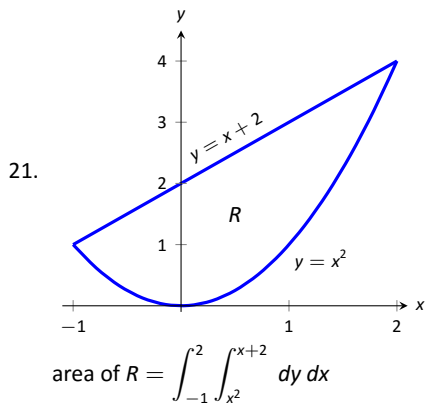
Exercises 14.2

1. volume

3. The double integral gives the signed volume under the surface. Since the surface is always positive, it is always above the x-y plane and hence produces only "positive" volume.

5. 6; $\int_{-1}^{1} \int_{1}^{2} \left(\dfrac{x}{y} + 3 \right) dy\, dx$

7. 112/3; $\int_{0}^{2} \int_{0}^{4-2y} \left(3x^2 - y + 2 \right) dx\, dy$

9. 16/5; $\int_{-1}^{1} \int_{0}^{1-x^2} (x + y + 2)\, dy\, dx$

11.

(a)
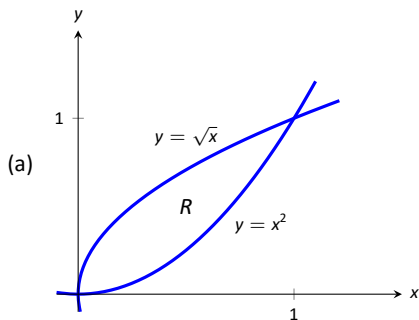
(b) $\int_{0}^{1} \int_{x^2}^{\sqrt{x}} x^2 y\, dy\, dx = \int_{0}^{1} \int_{y^2}^{\sqrt{y}} x^2 y\, dx\, dy.$

(c) $\frac{3}{56}$

13.

(a)
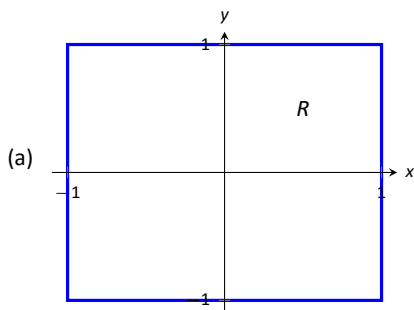
(b) $\int_{-1}^{1} \int_{-1}^{1} x^2 - y^2\, dy\, dx = \int_{-1}^{1} \int_{-1}^{1} x^2 - y^2\, dx\, dy.$

(c) 0

15.

(a)
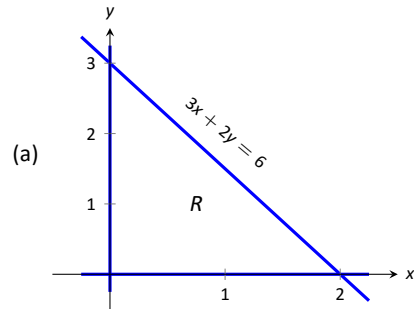
(b)

(c) $\int_{0}^{2} \int_{0}^{3-3/2x} \left(6 - 3x - 2y \right) dy\, dx = \int_{0}^{3} \int_{0}^{2-2/3y} \left(6 - 3x - 2y \right) dx\, dy.$

(d) 6

17.

(a)
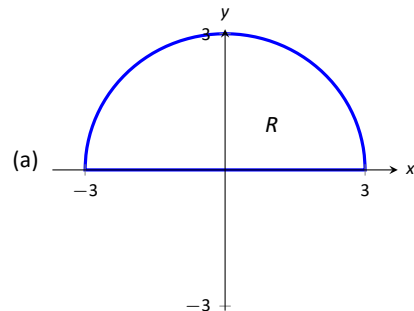
(b) $\int_{-3}^{3} \int_{0}^{\sqrt{9-x^2}} \left(x^3 y - x \right) dy\, dx = \int_{0}^{3} \int_{-\sqrt{9-y^2}}^{\sqrt{9-y^2}} \left(x^3 y - x \right) dx\, dy.$

(c) 0

19. Integrating e^{x^2} with respect to x is not possible in terms of elementary functions. $\int_{0}^{2} \int_{0}^{2x} e^{x^2}\, dy\, dx = e^4 - 1.$

21. Integrating $\int_{y}^{1} \dfrac{2y}{x^2 + y^2}\, dx$ gives $\tan^{-1}(1/y) - \pi/4$; integrating $\tan^{-1}(1/y)$ is hard. $\int_{0}^{1} \int_{0}^{x} \dfrac{2y}{x^2 + y^2}\, dy\, dx = \ln 2.$

23. average value of $f = 6/2 = 3$

25. average value of $f = \dfrac{112/3}{4} = 28/3$

Exercises 14.3

1. $f(r \cos\theta, r \sin\theta),\ r\, dr\, d\theta$

3. $\int_{0}^{2\pi} \int_{0}^{1} \left(3r \cos\theta - r \sin\theta + 4 \right) r\, dr\, d\theta = 4\pi$

5. $\int_{0}^{\pi} \int_{\cos\theta}^{3\cos\theta} \left(8 - r \sin\theta \right) r\, dr\, d\theta = 16\pi$

7. $\int_{0}^{2\pi} \int_{1}^{2} \left(\ln(r^2) \right) r\, dr\, d\theta = 2\pi \left(\ln 16 - 3/2 \right)$

9. $\displaystyle\int_{-\pi/2}^{\pi/2}\int_0^6 \left(r^2\cos^2\theta - r^2\sin^2\theta\right)r\,dr\,d\theta$

$\displaystyle =\int_{-\pi/2}^{\pi/2}\int_0^6 \left(r^2\cos(2\theta)\right)r\,dr\,d\theta = 0$

11. $\displaystyle\int_{-\pi/2}^{\pi/2}\int_0^5 \left(r^2\right)dr\,d\theta = 125\pi/3$

13. $\displaystyle\int_0^{\pi/4}\int_0^{\sqrt{8}} \left(r\cos\theta + r\sin\theta\right)r\,dr\,d\theta = 16\sqrt{2}/3$

15.

(a) This is impossible to integrate with rectangular coordinates as $e^{-(x^2+y^2)}$ does not have an antiderivative in terms of elementary functions.

(b) $\displaystyle\int_0^{2\pi}\int_0^a re^{r^2}\,dr\,d\theta = \pi(1-e^{-a^2})$.

(c) $\displaystyle\lim_{a\to\infty}\pi(1-e^{-a^2}) = \pi$. This implies that there is a finite volume under the surface $e^{-(x^2+y^2)}$ over the entire x-y plane.

Exercises 14.4

1. Because they are scalar multiples of each other.

3. "little masses"

5. M_x measures the moment about the x-axis, meaning we need to measure distance from the x-axis. Such measurements are measures in the y-direction.

7. $\bar{x} = 5.25$

9. $(\bar{x},\bar{y}) = (0,3)$

11. $M = 150$g;

13. $M = 2$lb

15. $M = 16\pi \approx 50.27$kg

17. $M = 54\pi \approx 169.65$lb

19. $M = 150$g; $M_y = 600$; $M_x = -75$; $(\bar{x},\bar{y}) = (4,-0.5)$

21. $M = 2$lb; $M_y = 0$; $M_x = 2/3$; $(\bar{x},\bar{y}) = (0,1/3)$

23. $M = 16\pi \approx 50.27$kg; $M_y = 4\pi$; $M_x = 4\pi$; $(\bar{x},\bar{y}) = (1/4,1/4)$

25. $M = 54\pi \approx 169.65$lb; $M_y = 0$; $M_x = 504$; $(\bar{x},\bar{y}) = (0,2.97)$

27. $I_x = 64/3$; $I_y = 64/3$; $I_O = 128/3$

29. $I_x = 16/3$; $I_y = 64/3$; $I_O = 80/3$

Exercises 14.5

1. arc length

3. surface areas

5. Intuitively, adding h to f only shifts f up (i.e., parallel to the z-axis) and does not change its shape. Therefore it will not change the surface area over R.

Analytically, $f_x = g_x$ and $f_y = g_y$; therefore, the surface area of each is computed with identical double integrals.

7. $\displaystyle S = \int_0^{2\pi}\int_0^{2\pi}\sqrt{1+\cos^2 x\cos^2 y + \sin^2 x\sin^2 y}\,dx\,dy$

9. $\displaystyle S = \int_{-1}^1\int_{-1}^1\sqrt{1+4x^2+4y^2}\,dx\,dy$

11. $\displaystyle S = \int_0^3\int_{-1}^1\sqrt{1+9+49}\,dx\,dy = 6\sqrt{59}\approx 46.09$

13. This is easier in polar:

$$S = \int_0^{2\pi}\int_0^4 r\sqrt{1+4r^2\cos^2 t + 4r^2\sin^2 t}\,dr\,d\theta$$
$$= \int_0^{2\pi}\int_0^4 r\sqrt{1+4r^2}\,dr\,d\theta$$
$$= \frac{\pi}{6}\left(65\sqrt{65}-1\right)\approx 273.87$$

15.

$$S = \int_0^2\int_0^{2x}\sqrt{1+1+4x^2}\,dy\,dx$$
$$= \int_0^2\left(2x\sqrt{2+4x^2}\right)dx$$
$$= \frac{26}{3}\sqrt{2}\approx 12.26$$

17. This is easier in polar:

$$S = \int_0^{2\pi}\int_0^5 r\sqrt{1+\frac{4r^2\cos^2 t + 4r^2\sin^2 t}{r^2\sin^2 t + r^2\cos^2 t}}\,dr\,d\theta$$
$$= \int_0^{2\pi}\int_0^5 r\sqrt{5}\,dr\,d\theta$$
$$= 25\pi\sqrt{5}\approx 175.62$$

19. Integrating in polar is easiest considering R:

$$S = \int_0^{2\pi}\int_0^1 r\sqrt{1+c^2+d^2}\,dr\,d\theta$$
$$= \int_0^{2\pi}\frac{1}{2}\left(\sqrt{1+c^2+d^2}\right)d\theta$$
$$= \pi\sqrt{1+c^2+d^2}.$$

The value of h does not matter as it only shifts the plane vertically (i.e., parallel to the z-axis). Different values of h do not create different ellipses in the plane.

Exercises 14.6

1. surface to surface, curve to curve and point to point

3. Answers can vary. From this section we used triple integration to find the volume of a solid region, the mass of a solid, and the center of mass of a solid.

5. $V = \int_{-1}^1\int_{-1}^1 \left(8 - x^2 - y^2 - (2x+y)\right)dx\,dy = 88/3$

7. $V = \int_0^\pi\int_0^x \left(\cos x\sin y + 2 - \sin x\cos y\right)dy\,dx = \pi^2 - \pi \approx 6.728$

9. *dz dy dx:* $\displaystyle\int_0^3\int_0^{1-x/3}\int_0^{2-2x/3-2y} dz\,dy\,dx$

dz dx dy: $\displaystyle\int_0^1\int_0^{3-3y}\int_0^{2-2x/3-2y} dz\,dx\,dy$

dy dz dx: $\displaystyle\int_0^3\int_0^{2-2x/3}\int_0^{1-x/3-z/2} dy\,dz\,dx$

dy dx dz: $\displaystyle\int_0^2\int_0^{3-3z/2}\int_0^{1-x/3-z/2} dy\,dx\,dz$

dx dz dy: $\displaystyle\int_0^1\int_0^{2-2y}\int_0^{3-3y-3z/2} dx\,dz\,dy$

dx dy dz: $\displaystyle\int_0^2\int_0^{1-z/2}\int_0^{3-3y-3z/2} dx\,dy\,dz$

$V = \displaystyle\int_0^3\int_0^{1-x/3}\int_0^{2-2x/3-2y} dz\,dy\,dx = 1.$

11. *dz dy dx:* $\displaystyle\int_0^2\int_{-2}^0\int_{y^2/2}^{-y} dz\,dy\,dx$

dz dx dy: $\displaystyle\int_{-2}^0\int_0^2\int_{y^2/2}^{-y} dz\,dx\,dy$

dy dz dx: $\displaystyle\int_0^2\int_0^2\int_{-\sqrt{2z}}^{-z} dy\,dz\,dx$

dy dx dz: $\displaystyle\int_0^2\int_0^2\int_{-\sqrt{2z}}^{-z} dy\,dx\,dz$

dx dz dy: $\displaystyle\int_{-2}^0\int_{y^2/2}^{-y}\int_0^2 dx\,dz\,dy$

dx dy dz: $\displaystyle\int_0^2\int_{-\sqrt{2z}}^{-z}\int_0^2 dx\,dy\,dz$

$V = \displaystyle\int_0^2\int_0^2\int_{-\sqrt{2z}}^{-z} dy\,dz\,dx = 4/3.$

13. *dz dy dx:* $\displaystyle\int_0^2\int_{1-x/2}^1\int_0^{2x+4y-4} dz\,dy\,dx$

dz dx dy: $\displaystyle\int_0^1\int_{2-2y}^2\int_0^{2x+4y-4} dz\,dx\,dy$

dy dz dx: $\displaystyle\int_0^2\int_0^{2x}\int_{z/4-x/2+1}^1 dy\,dz\,dx$

dy dx dz: $\displaystyle\int_0^4\int_{z/2}^2\int_{z/4-x/2+1}^1 dy\,dx\,dz$

dx dz dy: $\displaystyle\int_0^1\int_0^{4y}\int_{z/2-2y+2}^2 dx\,dz\,dy$

dx dy dz: $\displaystyle\int_0^4\int_{z/4}^1\int_{z/2-2y+2}^2 dx\,dy\,dz$

$V = \displaystyle\int_0^4\int_{z/4}^1\int_{z/2-2y+2}^2 dx\,dy\,dz = 4/3.$

15. *dz dy dx:* $\displaystyle\int_0^1\int_0^{1-x^2}\int_0^{\sqrt{1-y}} dz\,dy\,dx$

dz dx dy: $\displaystyle\int_0^1\int_0^{\sqrt{1-y}}\int_0^{\sqrt{1-y}} dz\,dx\,dy$

dy dz dx: $\displaystyle\int_0^1\int_0^x\int_0^{1-x^2} dy\,dz\,dx + \int_0^1\int_x^1\int_0^{1-z^2} dy\,dz\,dx$

dy dx dz: $\displaystyle\int_0^1\int_0^z\int_0^{1-z^2} dy\,dx\,dz + \int_0^1\int_z^1\int_0^{1-x^2} dy\,dx\,dz$

dx dz dy: $\displaystyle\int_0^1\int_0^{\sqrt{1-y}}\int_0^{\sqrt{1-y}} dx\,dz\,dy$

dx dy dz: $\displaystyle\int_0^1\int_0^{1-z^2}\int_0^{\sqrt{1-y}} dx\,dy\,dz$

Answers will vary. Neither order is particularly "hard." The order *dz dy dx* requires integrating a square root, so powers can be messy; the order *dy dz dx* requires two triple integrals, but each uses only polynomials.

17. 8

19. π

21. $M = 10$, $M_{yz} = 15/2$, $M_{xz} = 5/2$, $M_{xy} = 5$;
$(\bar{x}, \bar{y}, \bar{z}) = (3/4, 1/4, 1/2)$

23. $M = 16/5$, $M_{yz} = 16/3$, $M_{xz} = 104/45$, $M_{xy} = 32/9$;
$(\bar{x}, \bar{y}, \bar{z}) = (5/3, 13/18, 10/9) \approx (1.67, 0.72, 1.11)$

Exercises 14.7

1. 8π

3. $4\pi(8 - 3^{3/2})/3$

5.

7. $1 - \sin 2/2$

9. $2\pi ab$

11.

Chapter 15

Exercises 15.1

1. $1/2$

3. 23

5. 24π

7. -2π

9. 2π

11. 0

13.

15.

17.

Exercises 15.2

1. 0

3. No.

5. No.

7.

9. (b) No. *Hint:* Think of how *F* is defined.

Exercises 15.3

1. $16/15$

3. -5π

5. Yes. $F(x, y) = xy^2 + x^3$

7. Yes. $F(x, y) = 4x^2y + 2y^2 + 3x$

9.

11.

Exercises 15.4

1. 216π

3. $12\pi/5$

5.

7. $15/4$

9.

11.

Exercises 15.5

1. $2\sqrt{2}\pi^2$

3. $2/5$

5. $2\pi(\pi - 1)$

7. $67/15$

9. 6

11. Yes

13. No

15.

17.

19. Hint: Think of how a vector field $f(x, y) = P(x, y)i + Q(x, y)j$ in $\mathbb{R}^2$ can be extended in a natural way to be a vector field in $\mathbb{R}^3$.

Exercises 15.6

1. 0

3. $12\sqrt{x^2 + y^2 + z^2}$

5. $6(x + y + z)$

7. 12ρ

9. $-2zr^{-3}\vec{e}_r + r^{-2}\vec{e}_z$

11. $\text{div}\,\vec{f} = 2\rho^{-1} - \sin\theta \csc\phi + \cot\phi$;
$\text{curl}\,\vec{f} = \cot\phi\cos\theta\vec{e}_\rho + 2\vec{e}_\theta - 2\cos\theta\vec{e}_\phi$

13.

15.

17.

19.

21.

23.

25. Hint: Start by showing that $\vec{e}_r = \cos\theta\mathbf{i} + \sin\theta\mathbf{j}$, $\vec{e}_\theta = -\sin\theta\mathbf{i} + \cos\theta\mathbf{j}$, $\vec{e}_z = \mathbf{k}$.

27.

Index

Differentiation Rules

1. $\dfrac{d}{dx}(cx) = c$

2. $\dfrac{d}{dx}(u \pm v) = u' \pm v'$

3. $\dfrac{d}{dx}(u \cdot v) = uv' + u'v$

4. $\dfrac{d}{dx}\left(\dfrac{u}{v}\right) = \dfrac{vu' - uv'}{v^2}$

5. $\dfrac{d}{dx}(u(v)) = u'(v)v'$

6. $\dfrac{d}{dx}(c) = 0$

7. $\dfrac{d}{dx}(x) = 1$

8. $\dfrac{d}{dx}(x^n) = nx^{n-1}$

9. $\dfrac{d}{dx}(e^x) = e^x$

10. $\dfrac{d}{dx}(a^x) = \ln a \cdot a^x$

11. $\dfrac{d}{dx}(\ln x) = \dfrac{1}{x}$

12. $\dfrac{d}{dx}(\log_a x) = \dfrac{1}{\ln a} \cdot \dfrac{1}{x}$

13. $\dfrac{d}{dx}(\sin x) = \cos x$

14. $\dfrac{d}{dx}(\cos x) = -\sin x$

15. $\dfrac{d}{dx}(\csc x) = -\csc x \cot x$

16. $\dfrac{d}{dx}(\sec x) = \sec x \tan x$

17. $\dfrac{d}{dx}(\tan x) = \sec^2 x$

18. $\dfrac{d}{dx}(\cot x) = -\csc^2 x$

19. $\dfrac{d}{dx}(\sin^{-1} x) = \dfrac{1}{\sqrt{1-x^2}}$

20. $\dfrac{d}{dx}(\cos^{-1} x) = \dfrac{-1}{\sqrt{1-x^2}}$

21. $\dfrac{d}{dx}(\csc^{-1} x) = \dfrac{-1}{|x|\sqrt{x^2-1}}$

22. $\dfrac{d}{dx}(\sec^{-1} x) = \dfrac{1}{|x|\sqrt{x^2-1}}$

23. $\dfrac{d}{dx}(\tan^{-1} x) = \dfrac{1}{1+x^2}$

24. $\dfrac{d}{dx}(\cot^{-1} x) = \dfrac{-1}{1+x^2}$

25. $\dfrac{d}{dx}(\cosh x) = \sinh x$

26. $\dfrac{d}{dx}(\sinh x) = \cosh x$

27. $\dfrac{d}{dx}(\tanh x) = \operatorname{sech}^2 x$

28. $\dfrac{d}{dx}(\operatorname{sech} x) = -\operatorname{sech} x \tanh x$

29. $\dfrac{d}{dx}(\operatorname{csch} x) = -\operatorname{csch} x \coth x$

30. $\dfrac{d}{dx}(\coth x) = -\operatorname{csch}^2 x$

31. $\dfrac{d}{dx}(\cosh^{-1} x) = \dfrac{1}{\sqrt{x^2-1}}$

32. $\dfrac{d}{dx}(\sinh^{-1} x) = \dfrac{1}{\sqrt{x^2+1}}$

33. $\dfrac{d}{dx}(\operatorname{sech}^{-1} x) = \dfrac{-1}{x\sqrt{1-x^2}}$

34. $\dfrac{d}{dx}(\operatorname{csch}^{-1} x) = \dfrac{-1}{|x|\sqrt{1+x^2}}$

35. $\dfrac{d}{dx}(\tanh^{-1} x) = \dfrac{1}{1-x^2}$

36. $\dfrac{d}{dx}(\coth^{-1} x) = \dfrac{1}{1-x^2}$

Integration Rules

1. $\displaystyle\int c \cdot f(x)\, dx = c \int f(x)\, dx$

2. $\displaystyle\int f(x) \pm g(x)\, dx = \int f(x)\, dx \pm \int g(x)\, dx$

3. $\displaystyle\int 0\, dx = C$

4. $\displaystyle\int 1\, dx = x + C$

5. $\displaystyle\int x^n\, dx = \dfrac{1}{n+1}x^{n+1} + C, \quad n \neq -1$

6. $\displaystyle\int e^x\, dx = e^x + C$

7. $\displaystyle\int a^x\, dx = \dfrac{1}{\ln a} \cdot a^x + C$

8. $\displaystyle\int \dfrac{1}{x}\, dx = \ln|x| + C$

9. $\displaystyle\int \cos x\, dx = \sin x + C$

10. $\displaystyle\int \sin x\, dx = -\cos x + C$

11. $\displaystyle\int \tan x\, dx = -\ln|\cos x| + C$

12. $\displaystyle\int \sec x\, dx = \ln|\sec x + \tan x| + C$

13. $\displaystyle\int \csc x\, dx = -\ln|\csc x + \cot x| + C$

14. $\displaystyle\int \cot x\, dx = \ln|\sin x| + C$

15. $\displaystyle\int \sec^2 x\, dx = \tan x + C$

16. $\displaystyle\int \csc^2 x\, dx = -\cot x + C$

17. $\displaystyle\int \sec x \tan x\, dx = \sec x + C$

18. $\displaystyle\int \csc x \cot x\, dx = -\csc x + C$

19. $\displaystyle\int \cos^2 x\, dx = \dfrac{1}{2}x + \dfrac{1}{4}\sin(2x) + C$

20. $\displaystyle\int \sin^2 x\, dx = \dfrac{1}{2}x - \dfrac{1}{4}\sin(2x) + C$

21. $\displaystyle\int \dfrac{1}{x^2+a^2}\, dx = \dfrac{1}{a}\tan^{-1}\left(\dfrac{x}{a}\right) + C$

22. $\displaystyle\int \dfrac{1}{\sqrt{a^2-x^2}}\, dx = \sin^{-1}\left(\dfrac{x}{a}\right) + C$

23. $\displaystyle\int \dfrac{1}{x\sqrt{x^2-a^2}}\, dx = \dfrac{1}{a}\sec^{-1}\left(\dfrac{|x|}{a}\right) + C$

24. $\displaystyle\int \cosh x\, dx = \sinh x + C$

25. $\displaystyle\int \sinh x\, dx = \cosh x + C$

26. $\displaystyle\int \tanh x\, dx = \ln(\cosh x) + C$

27. $\displaystyle\int \coth x\, dx = \ln|\sinh x| + C$

28. $\displaystyle\int \dfrac{1}{\sqrt{x^2-a^2}}\, dx = \ln\left|x + \sqrt{x^2-a^2}\right| + C$

29. $\displaystyle\int \dfrac{1}{\sqrt{x^2+a^2}}\, dx = \ln\left|x + \sqrt{x^2+a^2}\right| + C$

30. $\displaystyle\int \dfrac{1}{a^2-x^2}\, dx = \dfrac{1}{2a}\ln\left|\dfrac{a+x}{a-x}\right| + C$

31. $\displaystyle\int \dfrac{1}{x\sqrt{a^2-x^2}}\, dx = \dfrac{1}{a}\ln\left(\dfrac{x}{a+\sqrt{a^2-x^2}}\right) + C$

32. $\displaystyle\int \dfrac{1}{x\sqrt{x^2+a^2}}\, dx = \dfrac{1}{a}\ln\left|\dfrac{x}{a+\sqrt{x^2+a^2}}\right| + C$
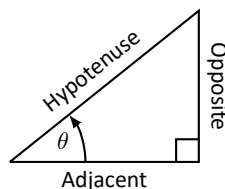
The Unit Circle
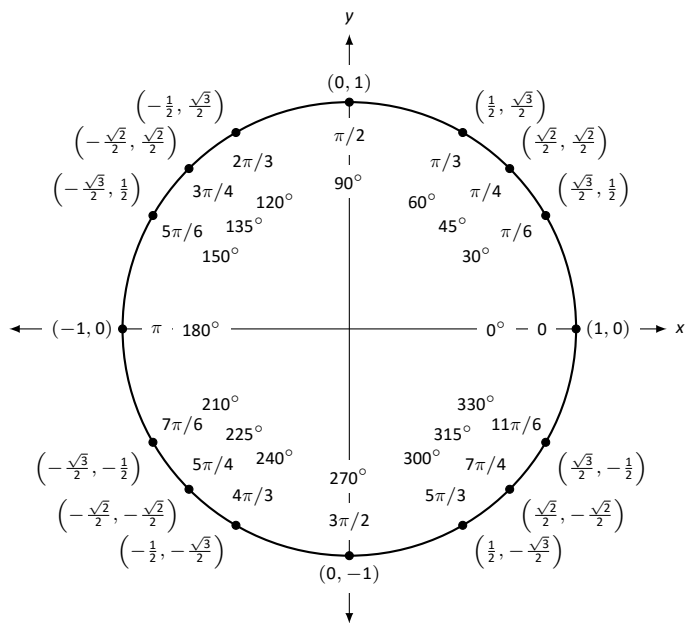
Definitions of the Trigonometric Functions

Unit Circle Definition

$$\sin \theta = y \qquad \cos \theta = x$$
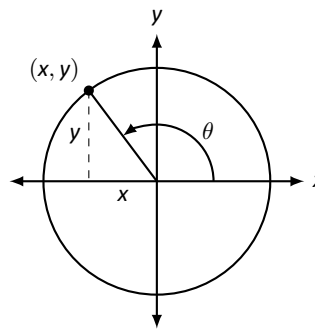
$$\csc \theta = \frac{1}{y} \qquad \sec \theta = \frac{1}{x}$$

$$\tan \theta = \frac{y}{x} \qquad \cot \theta = \frac{x}{y}$$

Right Triangle Definition

$$\sin \theta = \frac{O}{H} \qquad \csc \theta = \frac{H}{O}$$

$$\cos \theta = \frac{A}{H} \qquad \sec \theta = \frac{H}{A}$$

$$\tan \theta = \frac{O}{A} \qquad \cot \theta = \frac{A}{O}$$

Common Trigonometric Identities

Pythagorean Identities

$$\sin^2 x + \cos^2 x = 1$$
$$\tan^2 x + 1 = \sec^2 x$$
$$1 + \cot^2 x = \csc^2 x$$

Cofunction Identities

$$\sin\left(\frac{\pi}{2} - x\right) = \cos x \qquad \csc\left(\frac{\pi}{2} - x\right) = \sec x$$
$$\cos\left(\frac{\pi}{2} - x\right) = \sin x \qquad \sec\left(\frac{\pi}{2} - x\right) = \csc x$$
$$\tan\left(\frac{\pi}{2} - x\right) = \cot x \qquad \cot\left(\frac{\pi}{2} - x\right) = \tan x$$

Double Angle Formulas

$$\sin 2x = 2 \sin x \cos x$$
$$\cos 2x = \cos^2 x - \sin^2 x$$
$$= 2\cos^2 x - 1$$
$$= 1 - 2\sin^2 x$$
$$\tan 2x = \frac{2 \tan x}{1 - \tan^2 x}$$

Sum to Product Formulas

$$\sin x + \sin y = 2 \sin\left(\frac{x+y}{2}\right) \cos\left(\frac{x-y}{2}\right)$$
$$\sin x - \sin y = 2 \sin\left(\frac{x-y}{2}\right) \cos\left(\frac{x+y}{2}\right)$$
$$\cos x + \cos y = 2 \cos\left(\frac{x+y}{2}\right) \cos\left(\frac{x-y}{2}\right)$$
$$\cos x - \cos y = 2 \sin\left(\frac{x+y}{2}\right) \sin\left(\frac{y-x}{2}\right)$$

Power–Reducing Formulas

$$\sin^2 x = \frac{1 - \cos 2x}{2}$$
$$\cos^2 x = \frac{1 + \cos 2x}{2}$$
$$\tan^2 x = \frac{1 - \cos 2x}{1 + \cos 2x}$$

Even/Odd Identities

$$\sin(-x) = -\sin x$$
$$\cos(-x) = \cos x$$
$$\tan(-x) = -\tan x$$
$$\csc(-x) = -\csc x$$
$$\sec(-x) = \sec x$$
$$\cot(-x) = -\cot x$$

Product to Sum Formulas

$$\sin x \sin y = \frac{1}{2}\left(\cos(x-y) - \cos(x+y)\right)$$
$$\cos x \cos y = \frac{1}{2}\left(\cos(x-y) + \cos(x+y)\right)$$
$$\sin x \cos y = \frac{1}{2}\left(\sin(x+y) + \sin(x-y)\right)$$

Angle Sum/Difference Formulas

$$\sin(x \pm y) = \sin x \cos y \pm \cos x \sin y$$
$$\cos(x \pm y) = \cos x \cos y \mp \sin x \sin y$$
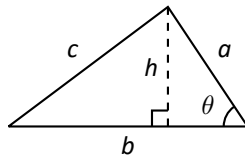$$\tan(x \pm y) = \frac{\tan x \pm \tan y}{1 \mp \tan x \tan y}$$

Areas and Volumes

Triangles

$h = a \sin \theta$
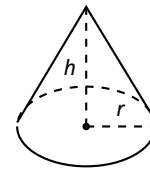
$\text{Area} = \dfrac{1}{2}bh$
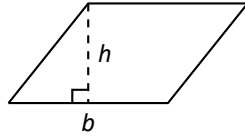
Law of Cosines:

$c^2 = a^2 + b^2 - 2ab \cos \theta$

Right Circular Cone

$\text{Volume} = \dfrac{1}{3}\pi r^2 h$

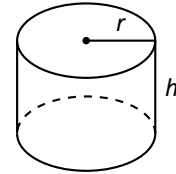$\text{Surface Area} =$

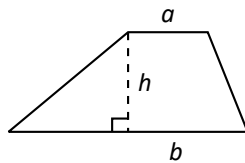$\pi r \sqrt{r^2 + h^2} + \pi r^2$

Parallelograms Area = bh

Right Circular Cylinder

$\text{Volume} = \pi r^2 h$
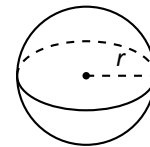
$\text{Surface Area} =$

$2\pi r h + 2\pi r^2$

Trapezoids Area = $\frac{1}{2}(a + b)h$

Sphere

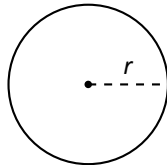$\text{Volume} = \dfrac{4}{3}\pi r^3$

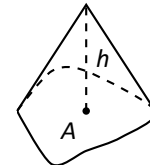$\text{Surface Area} = 4\pi r^2$

Circles

$\text{Area} = \pi r^2$

$\text{Circumference} = 2\pi r$

General Cone

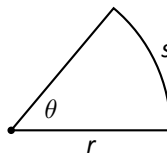$\text{Area of Base} = A$

$\text{Volume} = \dfrac{1}{3}Ah$

Sectors of Circles

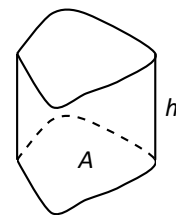θ in radians

$\text{Area} = \dfrac{1}{2}\theta r^2$

$s = r\theta$

General Right Cylinder

$\text{Area of Base} = A$

$\text{Volume} = Ah$

Algebra

Factors and Zeros of Polynomials

Let $p(x) = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0$ be a polynomial. If $p(a) = 0$, then a is a *zero* of the polynomial and a solution of the equation $p(x) = 0$. Furthermore, $(x - a)$ is a *factor* of the polynomial.

Fundamental Theorem of Algebra

An nth degree polynomial has n (not necessarily distinct) zeros. Although all of these zeros may be imaginary, a real polynomial of odd degree must have at least one real zero.

Quadratic Formula

If $p(x) = ax^2 + bx + c$, then the zeros of p are $x = \dfrac{-b \pm \sqrt{b^2 - 4ac}}{2a}$

Special Factoring

$$x^2 - a^2 = (x - a)(x + a) \qquad x^3 \pm a^3 = (x \pm a)(x^2 \mp ax + a^2) \qquad x^4 - a^4 = (x^2 - a^2)(x^2 + a^2)$$

Binomial Theorem

$$(x + y)^2 = x^2 + 2xy + y^2 \qquad\qquad (x + y)^3 = x^3 + 3x^2 y + 3xy^2 + y^3$$

$$(x + y)^4 = x^4 + 4x^3 y + 6x^2 y^2 + 4xy^3 + y^4 \qquad\qquad (x + y)^n = \sum_{i=0}^{n} \binom{n}{k} x^{n-k} y^k$$

Rational Zero Theorem

If $p(x) = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0$ has integer coefficients, then every *rational zero* of p is of the form $x = r/s$, where r is a factor of a_0 and s is a factor of a_n.

Factoring by Grouping

$$acx^3 + adx^2 + bcx + bd = ax^2(cs + d) + b(cx + d) = (ax^2 + b)(cx + d)$$

Arithmetic Operations

$$ab + ac = a(b + c) \qquad\qquad \frac{a}{b} + \frac{c}{d} = \frac{ad + bc}{bd} \qquad\qquad \frac{a + b}{c} = \frac{a}{c} + \frac{b}{c}$$

$$\frac{\left(\dfrac{a}{b}\right)}{\left(\dfrac{c}{d}\right)} = \left(\frac{a}{b}\right)\left(\frac{d}{c}\right) = \frac{ad}{bc} \qquad\qquad \frac{\left(\dfrac{a}{b}\right)}{c} = \frac{a}{bc} \qquad\qquad \frac{a}{\left(\dfrac{b}{c}\right)} = \frac{ac}{b}$$

$$a\left(\frac{b}{c}\right) = \frac{ab}{c} \qquad\qquad \frac{a - b}{c - d} = \frac{b - a}{d - c} \qquad\qquad \frac{ab + ac}{a} = b + c$$

Exponents and Radicals

$$a^0 = 1, \ \ a \neq 0 \qquad (ab)^x = a^x b^x \qquad a^x a^y = a^{x+y} \qquad \sqrt{a} = a^{1/2} \qquad \frac{a^x}{a^y} = a^{x-y} \qquad \sqrt[n]{a} = a^{1/n}$$

$$\left(\frac{a}{b}\right)^x = \frac{a^x}{b^x} \qquad \sqrt[n]{a^m} = a^{m/n} \qquad a^{-x} = \frac{1}{a^x} \qquad \sqrt[n]{ab} = \sqrt[n]{a}\sqrt[n]{b} \qquad (a^x)^y = a^{xy} \qquad \sqrt[n]{\frac{a}{b}} = \frac{\sqrt[n]{a}}{\sqrt[n]{b}}$$

Additional Formulas

Summation Formulas

$$\sum_{i=1}^{n} c = cn \qquad \sum_{i=1}^{n} i = \frac{n(n+1)}{2} \qquad \sum_{i=1}^{n} i^2 = \frac{n(n+1)(2n+1)}{6} \qquad \sum_{i=1}^{n} i^3 = \left(\frac{n(n+1)}{2}\right)^2$$

Trapezoidal Rule

$$\int_a^b f(x)\, dx \approx \frac{\Delta x}{2} \left[f(x_1) + 2f(x_2) + 2f(x_3) + \cdots + 2f(x_n) + f(x_{n+1})\right]$$

with Error $\leq \dfrac{(b-a)^3}{12n^2} \left[\max |f''(x)|\right]$

Simpson's Rule

$$\int_a^b f(x)\, dx \approx \frac{\Delta x}{3} \left[f(x_1) + 4f(x_2) + 2f(x_3) + 4f(x_4) + \cdots + 2f(x_{n-1}) + 4f(x_n) + f(x_{n+1})\right]$$

with Error $\leq \dfrac{(b-a)^5}{180n^4} \left[\max |f^{(4)}(x)|\right]$

Arc Length

$$L = \int_a^b \sqrt{1 + f'(x)^2}\, dx$$

Work Done by a Variable Force

$$W = \int_a^b F(x)\, dx$$

Force Exerted by a Fluid

$$F = \int_a^b w\, d(y)\, \ell(y)\, dy$$

Taylor Series Expansion for $f(x)$

$$p_n(x) = f(c) + f'(c)(x - c) + \frac{f''(c)}{2!}(x - c)^2 + \frac{f'''(c)}{3!}(x - c)^3 + \cdots + \frac{f^{(n)}(c)}{n!}(x - c)^n$$

Standard Form of Conic Sections

Parabola		Ellipse	Hyperbola	
Vertical axis	Horizontal axis		Foci and vertices on x-axis	Foci and vertices on y-axis
$y = \dfrac{x^2}{4p}$	$x = \dfrac{y^2}{4p}$	$\dfrac{x^2}{a^2} + \dfrac{y^2}{b^2} = 1$	$\dfrac{x^2}{a^2} - \dfrac{y^2}{b^2} = 1$	$\dfrac{y^2}{b^2} - \dfrac{x^2}{a^2} = 1$

Summary of Tests for Series

Test	Series	Condition(s) of Convergence	Condition(s) of Divergence	Comment
n^{th}-Term Test for Divergence	$\displaystyle\sum_{n=1}^{\infty} a_n$		$\displaystyle\lim_{n\to\infty} a_n \neq 0$	cannot show convergence.
Geometric Series	$\displaystyle\sum_{n=0}^{\infty} ar^n$	$\lvert r \rvert < 1$	$\lvert r \rvert \geq 1$	$\text{Sum} = \dfrac{a}{1-r}$
Telescoping Series	$\displaystyle\sum_{n=1}^{\infty} b_n - b_{n+m}$	$\displaystyle\lim_{n\to\infty} b_n = L$		$\text{Sum} = \left(\displaystyle\sum_{n=1}^{m} b_n\right) - L$
p-Series	$\displaystyle\sum_{n=1}^{\infty} \dfrac{1}{(an+b)^p}$	$p > 1$	$p \leq 1$	
p-Series For Logarithms	$\displaystyle\sum_{n=1}^{\infty} \dfrac{1}{(an+b)(\log n)^p}$	$p > 1$	$p \leq 1$	The base of the logarithm doesn't affect convergence.
Integral Test	$\displaystyle\sum_{n=1}^{\infty} a_n$	$\displaystyle\int_{1}^{\infty} a(n)\, dn$ converges	$\displaystyle\int_{1}^{\infty} a(n)\, dn$ diverges	$a_n = a(n)$ must be continuous and decreasing
Direct Comparison	$\displaystyle\sum_{n=1}^{\infty} a_n$	$\displaystyle\sum_{n=0}^{\infty} b_n$ converges and $0 \leq a_n \leq b_n$	$\displaystyle\sum_{n=0}^{\infty} b_n$ diverges and $0 \leq b_n \leq a_n$	
Limit Comparison	$\displaystyle\sum_{n=1}^{\infty} a_n$	$\displaystyle\sum_{n=0}^{\infty} b_n$ converges and $\displaystyle\lim_{n\to\infty} a_n/b_n \geq 0$	$\displaystyle\sum_{n=0}^{\infty} b_n$ diverges and $\displaystyle\lim_{n\to\infty} a_n/b_n > 0$ or $= \infty$	
Ratio Test	$\displaystyle\sum_{n=1}^{\infty} a_n$	$\displaystyle\lim_{n\to\infty} \dfrac{a_{n+1}}{a_n} < 1$	$\displaystyle\lim_{n\to\infty} \dfrac{a_{n+1}}{a_n} > 1$ or $= \infty$	$\{a_n\}$ must be positive
Root Test	$\displaystyle\sum_{n=1}^{\infty} a_n$	$\displaystyle\lim_{n\to\infty} (a_n)^{1/n} < 1$	$\displaystyle\lim_{n\to\infty} (a_n)^{1/n} > 1$ or $= \infty$	$\{a_n\}$ must be positive