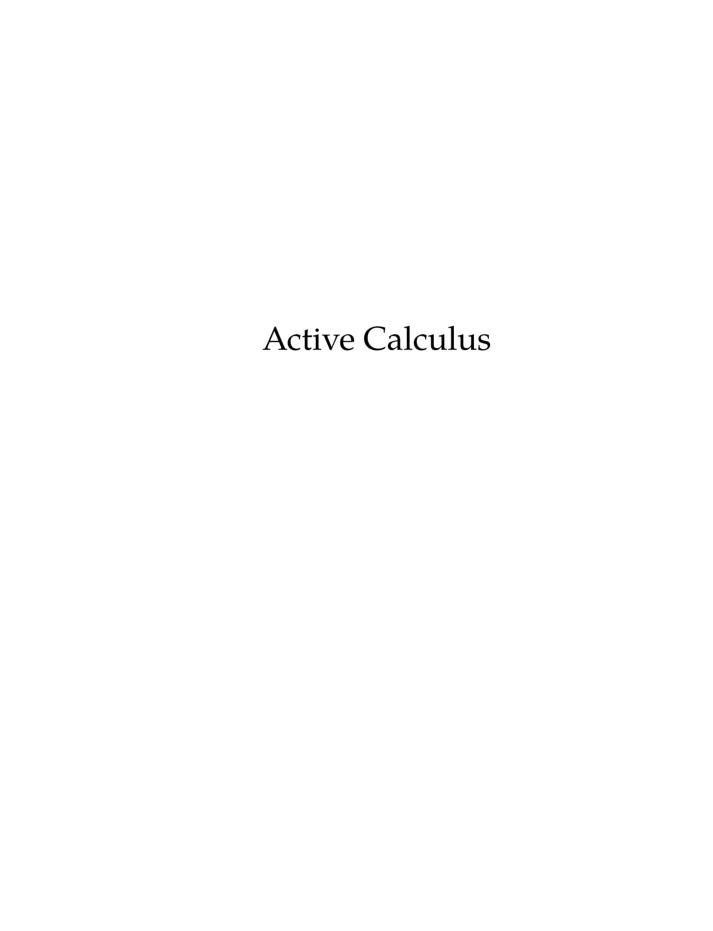
ACTIVE CALCULUS

2017 Edition



Matthew Boelkins
David Austin Steven Schlicker



Active Calculus

Matthew Boelkins Grand Valley State University

Contributing Authors

David Austin
Grand Valley State University
Steven Schlicker
Grand Valley State University

August 15, 2017

Cover Photo: James Haefner Photography

Edition: 2017

Website: faculty.gvsu.edu/boelkinm/Home/Active_Calculus.html

© 2012–2017 Matthew Boelkins

Permission is granted to copy, distribute and/or modify this document under the terms of the Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License. The work may be used for free by any party so long as attribution is given to the author(s), the work and its derivatives are used in the spirit of "share and share alike"; no party may sell this work or any of its derivatives for profit. All trademarks™ are the registered® marks of their respective owners. The graphic



that may appear in other locations in the text shows that the work is licensed with the Creative Commons, that the work may be used for free by any party so long as attribution is given to the author(s), that the work and its derivatives are used in the spirit of "share and share alike," and that no party may sell this work or any of its derivatives for profit, with the following exception: *it is entirely acceptable for university bookstores to sell bound photocopied copies of the activities workbook to students at their standard markup above the copying expense*. Full details may be found by visiting creativecommons.org/licenses/by-nc-sa/4.0/ or sending a letter to Creative Commons, 444 Castro Street, Suite 900, Mountain View, California, 94041, USA.

Acknowledgements

This text began as my sabbatical project in the winter semester of 2012, during which I wrote the preponderance of the materials for the first four chapters. For the sabbatical leave, I am indebted to Grand Valley State University for its support of the project and the time to write, as well as to my colleagues in the Department of Mathematics and the College of Liberal Arts and Sciences for their endorsement of the project as a valuable undertaking.

The beautiful full-color .eps graphics in the text are only possible because of David Austin of GVSU and Bill Casselman of the University of British Columbia. Building on their collective longstanding efforts to develop tools for high quality mathematical graphics, David wrote a library of Python routines that build on Bill's PiScript program, and David's routines are so easy to use that even I could generate graphics like the professionals that he and Bill are. I am deeply grateful to them both.

For the print-on-demand version of the 2016 text, I am thankful for the support of Lon Mitchell and Orthogonal Publishing L3C. For the first print edition of the text, Lon provided considerable guidance on LATEX and related typesetting issues, and volunteered his time throughout the production process. I met Lon at a special session devoted to open textbooks at the 2014 Joint Mathematics Meetings; I am grateful as well to the organizers of that session and subsequent ones who are part of a growing community of mathematicians committed to free and open texts. You can start to learn more about this work at the American Institute of Mathematics' open textbook site.

The current .html version of the text is possible only because of the amazing work of Rob Beezer and his development of the original Mathbook XML, soon to be known as PreTeXt. My ability to take advantage of Rob's work is largely due to the support of the American Institute of Mathematics, which funded me for a weeklong workshop in Mathbook XML in San Jose, CA, in April 2016. David Farmer's conversion script saved me hundreds of hours of work by taking my original IATEX source and converting it to MBX. Alex Jordan of Portland Community College has also been a tremendous help, and it is through Alex's fantastic work that live WeBWorK exercises are not only possible, but also included in the 2017 version. Alex has also contributed substantially to the .pdf version of the 2017 text.

Over my more than 15 years at GVSU, many of my colleagues have shared with me ideas and resources for teaching calculus. I am particularly indebted to David Austin, Will Dickinson, Paul Fishback, Jon Hodge, and Steve Schlicker for their contributions that improved my teaching of and thinking about calculus, including materials that I have modified and used over many different semesters with students. Parts of these ideas can be found throughout this text. In addition, Will Dickinson and Steve Schlicker provided me access to a large number of their electronic notes and activities from teaching of differential and integral calculus, and those ideas and materials have similarly impacted my work and writing in positive ways, with some of their problems and approaches finding parallel presentation here.



Shelly Smith of GVSU and Matt Delong of Taylor University both provided extensive comments on the first few chapters of early drafts, feedback that was immensely helpful in improving the text. As more and more people use the text, I am grateful to everyone who reads, edits, and uses this book, and hence contributes to its improvement through ongoing discussion.

Finally, I am grateful for all that my students have taught me over the years. Their responses and feedback have helped to shape me as a teacher, and I appreciate their willingness to wholeheartedly engage in the activities and approaches I've tried in class, to let me know how those affect their learning, and to help me learn and grow as an instructor. Early on, they provided useful editorial feedback on this text.

Any and all remaining errors or inconsistencies are mine. I will gladly take reader and user feedback to correct them, along with other suggestions to improve the text.

Contributors

A large and growing number of people have generously contributed to the development or improvement of the text. Contributing authors David Austin and Steven Schlicker have each written drafts of at least one full chapter of the text.

The following contributing editors have offered significant feedback that includes information about typographical errors or suggestions to improve the exposition.

DAVID AUSTIN HUGH McGUIRE

GVSU GVSU

Allan Bickle Ray Rosentrater

GVSU Westmont College

DAVID CLARK

GVS11 Luis Sanjuan

Conservatorio Profesional de Musica de Avila

WILL DICKINSON Spain

GVSU

STEVEN SCHLICKER

Charles Fortin GVSU

Champlain Regional College

St-Lambert, Quebec, Canada

Brian Stanley

Marcia Frobis Foothill Community College

GVSU

Amy Stone

Patti Hunter GVSU

Westmont College

MITCH KELLER ROBERT TALBERT

Washington and Lee University GVSU

Greg Thull

Dave Kung Greg TH

St. Mary's College of Maryland GVSU

GREG TH

GVSU

Paul Latiolais Sue Van Hattum

Portland State University Contra Costa College

Active Calculus: Our Goals

Several fundamental ideas in calculus are more than 2000 years old. As a formal subdiscipline of mathematics, calculus was first introduced and developed in the late 1600s, with key independent contributions from Sir Isaac Newton and Gottfried Wilhelm Leibniz. Mathematicians agree that the subject has been understood rigorously since the work of Augustin Louis Cauchy and Karl Weierstrass in the mid 1800s when the field of modern analysis was developed, in part to make sense of the infinitely small quantities on which calculus rests. As a body of knowledge, calculus has been completely understood for at least 150 years. The discipline is one of our great human intellectual achievements: among many spectacular ideas, calculus models how objects fall under the forces of gravity and wind resistance, explains how to compute areas and volumes of interesting shapes, enables us to work rigorously with infinitely small and infinitely large quantities, and connects the varying rates at which quantities change to the total change in the quantities themselves.

While each author of a calculus textbook certainly offers her own creative perspective on the subject, it is hardly the case that many of the ideas she presents are new. Indeed, the mathematics community broadly agrees on what the main ideas of calculus are, as well as their justification and their importance; the core parts of nearly all calculus textbooks are very similar. As such, it is our opinion that in the 21st century—an age where the internet permits seamless and immediate transmission of information—no one should be required to purchase a calculus text to read, to use for a class, or to find a coherent collection of problems to solve. Calculus belongs to humankind, not any individual author or publishing company. Thus, a main purpose of this work is to present a new calculus text that is *free*. In addition, instructors who are looking for a calculus text should have the opportunity to download the source files and make modifications that they see fit; thus this text is *open-source*. Since August 2013, *Active Calculus* has been endorsed by the American Institute of Mathematics and its Open Textbook Initiative.

In *Active Calculus*, we endeavor to actively engage students in learning the subject through an activity-driven approach in which the vast majority of the examples are completed by students. Where many texts present a general theory of calculus followed by substantial collections of worked examples, we instead pose problems or situations, consider possibilities, and then ask students to investigate and explore. Following key activities or examples, the presentation normally includes some overall perspective and a brief synopsis of general trends or properties, followed by formal statements of rules or theorems. While we often offer plausibility arguments for such results, rarely do we include formal proofs. It is not the intent of this text for the instructor or author to *demonstrate* to students that the ideas of calculus are coherent and true, but rather for students to *encounter* these ideas in a supportive, leading manner that enables them to begin to understand for themselves why calculus is both coherent and true. This approach is consistent with the growing body of scholarship that calls for students to be interactively engaged in class.



Moreover, this approach is consistent with the following goals:

- To have students engage in an active, inquiry-driven approach, where learners strive to construct solutions and approaches to ideas on their own, with appropriate support through questions posed, hints, and guidance from the instructor and text.
- To build in students intuition for why the main ideas in calculus are natural and true. Often, we do this through consideration of the instantaneous position and velocity of a moving object, a scenario that is common and familiar.
- To challenge students to acquire deep, personal understanding of calculus through reading the text and completing preview activities on their own, through working on activities in small groups in class, and through doing substantial exercises outside of class time.
- To strengthen students' written and oral communicating skills by having them write about and explain aloud the key ideas of calculus.

Features of the Text

Instructors and students alike will find several consistent features in the presentation, including:

- **Motivating Questions** At the start of each section, we list 2–3 *motivating questions* that provide motivation for why the following material is of interest to us. One goal of each section is to answer each of the motivating questions.
- **Preview Activities** Each section of the text begins with a short introduction, followed by a *preview activity*. This brief reading and the preview activity are designed to foreshadow the upcoming ideas in the remainder of the section; both the reading and preview activity are intended to be accessible to students *in advance* of class, and indeed to be completed by students before a day on which a particular section is to be considered.
- **Activities** A typical section in the text has three *activities*. These are designed to engage students in an inquiry-based style that encourages them to construct solutions to key examples on their own, working individually or in small groups.
- Exercises There are dozens of calculus texts with (collectively) tens of thousands of exercises. Rather than repeat standard and routine exercises in this text, we recommend the use of WeBWorK with its access to the Open Problem Library and around 20,000 calculus problems. In this text, each section includes a small number of anonymous WeBWorK exercises, as well as 3–4 challenging problems per section. The WeBWorK exercises are best completed in the .html version of the text. Almost every non-WeBWorK problem has multiple parts, requires the student to connect several key ideas, and expects that the student will do at least a modest amount of writing to answer the questions and explain their findings. For instructors interested in a more conventional source of exercises, consider the freely available *APEX Calculus* text by Greg Hartmann et al., available from www.apexcalculus.com.
- **Graphics** As much as possible, we strive to demonstrate key fundamental ideas visually, and to encourage students to do the same. Throughout the text, we use full-color¹ graphics to exemplify and magnify key ideas, and to use this graphical perspective alongside both numerical and algebraic representations of calculus.
- Links to interactive graphics Many of the ideas of calculus are best understood dynamically; java applets offer an often ideal format for investigations and demonstrations. Relying primarily on the work of David Austin of Grand Valley State University and Marc Renault of Shippensburg University, each of whom has developed a large library of applets for calculus, we frequently point the reader (through active links in the electronic versions of the text) to applets that are relevant for key ideas under con-

¹To keep cost low, the graphics in the print-on-demand version are in black and white. When the text itself refers to color in images, one needs to view the .html or .pdf electronically.



sideration.

Summary of Key Ideas Each section concludes with a summary of the key ideas encountered in the preceding section; this summary normally reflects responses to the motivating questions that began the section.

How to Use this Text

Because the text is free, any professor or student may use the electronic version of the text for no charge. For reading on laptops or mobile devices, the best electronic version to use is the .html one at gvsu.edu/s/0uo, but you may also download a full .pdf copy of the text from gvsu.edu/s/0vM, where there is also a link to a print-on-demand option for purchasing a bound, softcover version for under \$25. Other ancillary materials, such as WeBWorK .def files, an activities-only workbook, and sample computer laboratory activities are available upon direct request to the author via email at boelkinm@gvsu.edu. Furthermore, because the text is open-source, any instructor may acquire the full set of source files, which are available on GitHub.

This text may be used as a stand-alone textbook for a standard first semester college calculus course or as a supplement to a more traditional text. Chapters 1–4 address the typical topics for differential calculus, while Chapters 5–8 provide the standard topics of integral calculus, including a chapter on differential equations (Chapter 7) and on infinite series (Chapter 8).

Electronic Edition Because students and instructors alike have access to the book in electronic format, there are several advantages to the text over a traditional print text. One is that the text may be projected on a screen in the classroom (or even better, on a whiteboard) and the instructor may reference ideas in the text directly, add comments or notation or features to graphs, and indeed write right on the text itself. Students can do likewise, choosing to print only whatever portions of the text are needed for them. In addition, the electronic versions of the text includes live .html links to java applets, so student and instructor alike may follow those links to additional resources that lie outside the text itself. Finally, students can have access to a copy of the text anywhere they have a computer. The .html version is far superior to the .pdf version; this is especially true for viewing on a smartphone.

*Note.*In the .pdf version, there is not an obvious visual indicator of the live .html links, so some available information is suppressed. If you are using the text electronically in a setting with internet access, please know that it is assumed you are using the .html version.

Activities Workbook Each section of the text has a preview activity and at least three inclass activities embedded in the discussion. As it is the expectation that students will complete all of these activities, it is ideal for them to have room to work on them adjacent to the problem statements themselves. As a separate document, we have compiled a workbook of activities that includes only the individual activity prompts, along with space provided for students to write their responses. This workbook is the one printing expense that students will almost certainly have to undertake, and is available upon request.



Community of Users Because this text is free and open-source, we hope that as people use the text, they will contribute corrections, suggestions, and new material. The best way to communicate such feedback is by email to Matt Boelkins. I also have a blog at open-calculus.wordpress.com, at which we post new developments, other free resources, feedback, and other points of discussion.

Contents

A	knov	wledgements		V
Co	ontril	butors		vi
A	tive	Calculus: Our Goals		ix
Fe	ature	es of the Text		x i
Н	ow to	Use this Text		xiii
1	Und	derstanding the Derivative		1
	1.1	How do we measure velocity?		1
	1.2	The notion of limit		11
	1.3	The derivative of a function at a point		22
	1.4	The derivative function		34
	1.5	Interpreting, estimating, and using the derivative		45
	1.6	The second derivative		55
	1.7	Limits, Continuity, and Differentiability		70
	1.8	The Tangent Line Approximation		82
2	Con	nputing Derivatives		91
	2.1	Elementary derivative rules		91
	2.2	The sine and cosine functions		100
	2.3	The product and quotient rules		106
	2.4	Derivatives of other trigonometric functions		118
	2.5	The chain rule		
	2.6	Derivatives of Inverse Functions		134
	2.7	Derivatives of Functions Given Implicitly		145
	2.8	Using Derivatives to Evaluate Limits		153
3	Usi	ng Derivatives		165
	3.1	Using derivatives to identify extreme values		165
	3.2	Using derivatives to describe families of functions		
	3.3	Global Optimization		186
	3.4	Applied Optimization		
	3.5	Related Rates		
4	The	Definite Integral		209
	4.1	Determining distance traveled from velocity		
	4.2	Riemann Sums		



Contents

A SI	nort fable of filtegrals	329
A C1	hort Table of Integrals	529
8.6	Power Series	516
8.4		
8.3	Series of Real Numbers	
8.2		
8.1	1	
Seq		449
7.6	Population Growth and the Logistic Equation	436
7.3		
7.2	Qualitative behavior of solutions to DEs	
7.1	An Introduction to Differential Equations	
Diff	erential Equations	385
6.5	Improper Integrals	376
6.4	Physics Applications: Work, Force, and Pressure	
6.3	Density, Mass, and Center of Mass	354
6.2	Using Definite Integrals to Find Volume	
6.1		335
Usir	ng Definite Integrals	335
5.6	Numerical Integration	321
5.5	Other Options for Finding Algebraic Antiderivatives	312
5.4	Integration by Parts	
5.3		
5.2		
Eval	uating Integrals	269
4.4	The Fundamental Theorem of Calculus	254
4.3	The Definite Integral	
	4.4 Eval 5.1 5.2 5.3 5.4 5.5 5.6 Usin 6.1 6.2 6.3 6.4 6.5 Diff 7.1 7.2 7.3 7.4 7.5 7.6 Sequence 8.1 8.2 8.3 8.4 8.5 8.6	Evaluating Integrals 5.1 Constructing Accurate Graphs of Antiderivatives 5.2 The Second Fundamental Theorem of Calculus 5.3 Integration by Substitution 5.4 Integration by Parts 5.5 Other Options for Finding Algebraic Antiderivatives 5.6 Numerical Integration Using Definite Integrals 6.1 Using Definite Integrals to Find Area and Length 6.2 Using Definite Integrals to Find Volume 6.3 Density, Mass, and Center of Mass 6.4 Physics Applications: Work, Force, and Pressure 6.5 Improper Integrals Differential Equations 7.1 An Introduction to Differential Equations 7.2 Qualitative behavior of solutions to DEs 7.3 Euler's method 7.4 Separable differential equations 7.5 Modeling with differential equations 7.6 Population Growth and the Logistic Equation Sequences and Series 8.1 Sequences 8.2 Geometric Series 8.3 Series of Real Numbers 8.4 Alternating Series 8.5 Taylor Polynomials and Taylor Series 8.6 Power Series

2.8 Using Derivatives to Evaluate Limits

Motivating Questions

- How can derivatives be used to help us evaluate indeterminate limits of the form $\frac{0}{0}$?
- What does it mean to say that $\lim_{x\to\infty} f(x) = L$ and $\lim_{x\to a} f(x) = \infty$?
- How can derivatives assist us in evaluating indeterminate limits of the form $\frac{\infty}{\infty}$?

Because differential calculus is based on the definition of the derivative, and the definition of the derivative involves a limit, there is a sense in which all of calculus rests on limits. In addition, the limit involved in the limit definition of the derivative is one that always generates an indeterminate form of $\frac{0}{0}$. If f is a differentiable function for which f'(x) exists, then when we consider

$$f'(x) = \lim_{h \to 0} \frac{f(x+h) - f(x)}{h},$$

it follows that not only does $h \to 0$ in the denominator, but also $(f(x+h)-f(x)) \to 0$ in the numerator, since f is continuous. Thus, the fundamental form of the limit involved in the definition of f'(x) is $\frac{0}{0}$. Remember, saying a limit has an indeterminate form only means that we don't yet know its value and have more work to do: indeed, limits of the form $\frac{0}{0}$ can take on any value, as is evidenced by evaluating f'(x) for varying values of x for a function such as $f'(x) = x^2$.

Of course, we have learned many different techniques for evaluating the limits that result from the derivative definition, and including a large number of shortcut rules that enable us to evaluate these limits quickly and easily. In this section, we turn the situation upsidedown: rather than using limits to evaluate derivatives, we explore how to use derivatives to evaluate certain limits. This topic will combine several different ideas, including limits, derivative shortcuts, local linearity, and the tangent line approximation.

Preview Activity 2.8.1. Let *h* be the function given by $h(x) = \frac{x^5 + x - 2}{x^2 - 1}$.

- a. What is the domain of h?
- b. Explain why $\lim_{x\to 1} \frac{x^5 + x 2}{x^2 1}$ results in an indeterminate form.
- c. Next we will investigate the behavior of both the numerator and denominator of h near the point where x = 1. Let $f(x) = x^5 + x 2$ and $g(x) = x^2 1$. Find the local linearizations of f and g at g and g at g at g at g at g at g and g at g at g and g at g at g and g at g at g and g at g at g and g at g at g at g and g and g at g and g a
- d. Explain why $h(x) \approx \frac{L_f(x)}{L_g(x)}$ for x near a = 1.

e. Using your work from (c) and (d), evaluate

$$\lim_{x \to 1} \frac{L_f(x)}{L_g(x)}.$$

What do you think your result tells us about $\lim_{x\to 1} h(x)$?

f. Investigate the function h(x) graphically and numerically near x = 1. What do you think is the value of $\lim_{x\to 1} h(x)$?

2.8.1 Using derivatives to evaluate indeterminate limits of the form $\frac{0}{0}$.

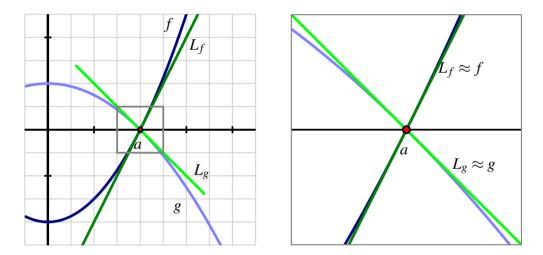


Figure 2.8.1: At left, the graphs of f and g near the value a, along with their tangent line approximations L_f and L_g at x = a. At right, zooming in on the point a and the four graphs.

The fundamental idea of Preview Activity 2.8.1 — that we can evaluate an indeterminate limit of the form $\frac{0}{0}$ by replacing each of the numerator and denominator with their local linearizations at the point of interest — can be generalized in a way that enables us to easily evaluate a wide range of limits. We begin by assuming that we have a function h(x) that can be written in the form $h(x) = \frac{f(x)}{g(x)}$ where f and g are both differentiable at x = a and for which f(a) = g(a) = 0. We are interested in finding a way to evaluate the indeterminate limit given by $\lim_{x\to a} h(x)$. In Figure 2.8.1, we see a visual representation of the situation involving such functions f and g. In particular, we see that both f and g have an g-intercept at the point where g and g are an interested in the limit of g and g at g at g and g at g at g and we can find their respective tangent line approximations g and g at g at g and g at g and g at g at g and g at g at g and g at g at g and g and g are both differentiable and g at g and g and g are both differentiable and g at g and g are both differentiable and g at g and g are both differentiable and g and g are both differentiable and g at g and g are both differentiable and g and

behaviors of f(x) and g(x) as $x \to a$ are key to understand. Here, we take advantage of the fact that each function and its tangent line approximation become indistinguishable as $x \to a$.

First, let's reall that $L_f(x) = f'(a)(x-a) + f(a)$ and $L_g(x) = g'(a)(x-a) + g(a)$. The critical observation we make is that when taking the limit, because x is getting arbitrarily close to a, we can replace f with L_f and replace g with L_g , and thus we observe that

$$\lim_{x \to a} \frac{f(x)}{g(x)} = \lim_{x \to a} \frac{L_f(x)}{L_g(x)}$$
$$= \lim_{x \to a} \frac{f'(a)(x-a) + f(a)}{g'(a)(x-a) + g(a)}.$$

Next, we remember a key fundamental assumption: that both f(a) = 0 and g(a) = 0, as this is precisely what makes the original limit indeterminate. Substituting these values for f(a) and g(a) in the limit above, we now have

$$\lim_{x \to a} \frac{f(x)}{g(x)} = \lim_{x \to a} \frac{f'(a)(x-a)}{g'(a)(x-a)}$$
$$= \lim_{x \to a} \frac{f'(a)}{g'(a)},$$

where the latter equality holds since x is approaching (but not equal to) a, so $\frac{x-a}{x-a} = 1$. Finally, we note that $\frac{f'(a)}{\sigma'(a)}$ is constant with respect to x, and thus

$$\lim_{x \to a} \frac{f(x)}{g(x)} = \frac{f'(a)}{g'(a)}.$$

We have, of course, implicitly made the assumption that $g'(a) \neq 0$, which is essential to the overall limit having the value $\frac{f'(a)}{g'(a)}$. We summarize our work above with the statement of L'Hôpital's Rule, which is the formal name of the result we have shown.

L'Hôpital's Rule

Let f and g be differentiable at x = a, and suppose that f(a) = g(a) = 0 and that $g'(a) \neq 0$. Then $\lim_{x \to a} \frac{f(x)}{g(x)} = \frac{f'(a)}{g'(a)}$.

In practice, we typically work with a slightly more general version of L'Hôpital's Rule, which states that (under the identical assumptions as the boxed rule above and the extra assumption that g' is continuous at x = a)

$$\lim_{x \to a} \frac{f(x)}{g(x)} = \lim_{x \to a} \frac{f'(x)}{g'(x)},$$

provided the righthand limit exists. This form reflects the fundamental benefit of L'Hôpital's Rule: if $\frac{f(x)}{g(x)}$ produces an indeterminate limit of form $\frac{0}{0}$ as $x \to a$, it is equivalent to consider

the limit of the quotient of the two functions' derivatives, $\frac{f'(x)}{g'(x)}$. For example, if we consider the limit from Preview Activity 2.8.1,

$$\lim_{x \to 1} \frac{x^5 + x - 2}{x^2 - 1},$$

by L'Hôpital's Rule we have that

$$\lim_{x \to 1} \frac{x^5 + x - 2}{x^2 - 1} = \lim_{x \to 1} \frac{5x^4 + 1}{2x} = \frac{6}{2} = 3.$$

By being able to replace the numerator and denominator with their respective derivatives, we often move from an indeterminate limit to one whose value we can easily determine.

Activity 2.8.2. Evaluate each of the following limits. If you use L'Hôpital's Rule, indicate where it was used, and be certain its hypotheses are met before you apply it.

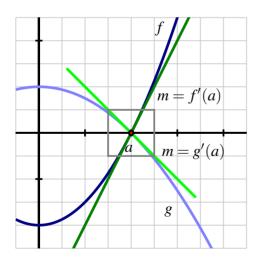
a.
$$\lim_{x\to 0} \frac{\ln(1+x)}{x}$$

c.
$$\lim_{x \to 1} \frac{2 \ln(x)}{1 - e^{x-1}}$$

b.
$$\lim_{x\to\pi} \frac{\cos(x)}{x}$$

d.
$$\lim_{x\to 0} \frac{\sin(x)-x}{\cos(2x)-1}$$

While L'Hôpital's Rule can be applied in an entirely algebraic way,



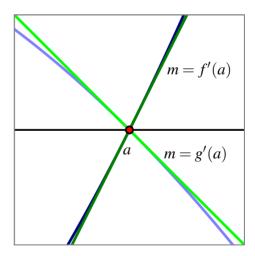


Figure 2.8.2: Two functions f and g that satisfy L'Hôpital's Rule.

it is important to remember that the genesis of the rule is graphical: the main idea is that the slopes of the tangent lines to f and g at x = a determine the value of the limit of $\frac{f(x)}{g(x)}$ as

 $x \to a$. We see this in Figure 2.8.2, which is a modified version of Figure 2.8.1, where we can see from the grid that f'(a) = 2 and g'(a) = -1, hence by L'Hôpital's Rule,

$$\lim_{x \to a} \frac{f(x)}{g(x)} = \frac{f'(a)}{g'(a)} = \frac{2}{-1} = -2.$$

Indeed, what we observe is that it's not the fact that f and g both approach zero that matters most, but rather the *rate* at which each approaches zero that determines the value of the limit. This is a good way to remember what L'Hôpital's Rule says: if f(a) = g(a) = 0, the the limit of $\frac{f(x)}{a(x)}$ as $x \to a$ is given by the ratio of the slopes of f and g at x = a.

Activity 2.8.3. In this activity, we reason graphically from the following figure to evaluate limits of ratios of functions about which some information is known.

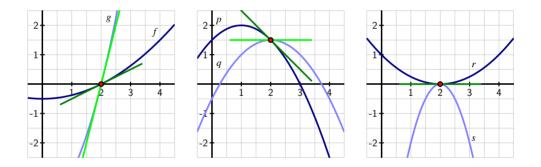


Figure 2.8.3: Three graphs referenced in the questions of Activity 2.8.3.

- a. Use the left-hand graph to determine the values of f(2), f'(2), g(2), and g'(2). Then, evaluate $\lim_{x\to 2}\frac{f(x)}{g(x)}$.
- b. Use the middle graph to find p(2), p'(2), q(2), and q'(2). Then, determine the value of $\lim_{x\to 2}\frac{p(x)}{q(x)}$.
- c. Use the right-hand graph to compute r(2), r'(2), s(2), s'(2). Explain why you cannot determine the exact value of $\lim_{x\to 2}\frac{r(x)}{s(x)}$ without further information being provided, but that you can determine the sign of $\lim_{x\to 2}\frac{r(x)}{s(x)}$. In addition, state what the sign of the limit will be, with justification.

2.8.2 Limits involving ∞

The concept of infinity, denoted ∞ , arises naturally in calculus, like it does in much of mathematics. It is important to note from the outset that ∞ is a concept, but not a number itself. Indeed, the notion of ∞ naturally invokes the idea of limits. Consider, for example, the function $f(x) = \frac{1}{x}$, whose graph is pictured in Figure 2.8.4.

We note that x = 0 is not in the domain of f, so we may naturally wonder what happens as $x \to 0$. As $x \to 0^+$, we observe that f(x) increases without bound. That is, we can make the value of f(x) as large as we like by taking x closer and closer (but not equal) to 0, while keeping x > 0. This is a good way to think about what infinity represents: a quantity is tending to infinity if there is no single number that the quantity is always less than.

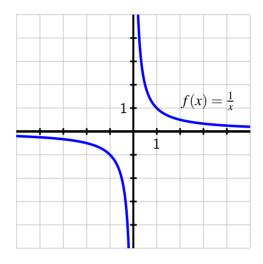


Figure 2.8.4: The graph of $f(x) = \frac{1}{x}$.

Recall that when we write $\lim_{x\to a} f(x) = L$, this means that can make f(x) as close to L as we'd like by taking x sufficiently close (but not equal) to a. We thus expand this notation and language to include the possibility that either L or a can be ∞ . For instance, for $f(x) = \frac{1}{x}$, we now write

$$\lim_{x\to 0^+}\frac{1}{x}=\infty,$$

by which we mean that we can make $\frac{1}{x}$ as large as we like by taking x sufficiently close (but not equal) to 0. In a similar way, we naturally write

$$\lim_{r\to\infty}\frac{1}{r}=0,$$

since we can make $\frac{1}{x}$ as close to 0 as we'd like by taking x sufficiently large (i.e., by letting x increase without bound).

In general, we understand the notation $\lim_{x\to a} f(x) = \infty$ to mean that we can make f(x) as large as we'd like by taking x sufficiently close (but not equal) to a, and the notation $\lim_{x\to\infty} f(x) = L$ to mean that we can make f(x) as close to L as we'd like by taking x sufficiently large. This notation applies to left- and right-hand limits, plus we can also use limits involving $-\infty$. For example, returning to Figure 2.8.4 and $f(x) = \frac{1}{x}$, we can say that

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

Finally, we write

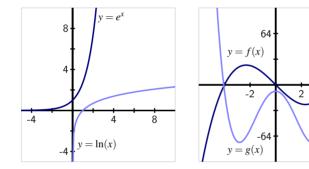
$$\lim_{x \to \infty} f(x) = \infty$$

when we can make the value of f(x) as large as we'd like by taking x sufficiently large. For example,

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

Note particularly that limits involving infinity identify *vertical* and *horizontal asymptotes* of a function. If $\lim_{x\to a} f(x) = \infty$, then x = a is a vertical asymptote of f, while if $\lim_{x\to\infty} f(x) = L$, then y = L is a horizontal asymptote of f. Similar statements can be made using $-\infty$, as well as with left- and right-hand limits as $x \to a^-$ or $x \to a^+$.

In precalculus classes, it is common to study the *end behavior* of certain families of functions, by which we mean the behavior of a function as $x \to \infty$ and as $x \to -\infty$. Here we briefly examine a library of some familiar functions and note the values of several limits involving ∞ .



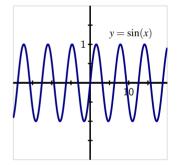


Figure 2.8.5: Graphs of some familiar functions whose end behavior as $x \to \pm \infty$ is known. In the middle graph, $f(x) = x^3 - 16x$ and $g(x) = x^4 - 16x^2 - 8$.

For the natural exponential function e^x , we note that $\lim_{x\to\infty} e^x = \infty$ and $\lim_{x\to-\infty} e^x = 0$, while for the related exponential decay function e^{-x} , observe that these limits are reversed, with $\lim_{x\to\infty} e^{-x} = 0$ and $\lim_{x\to-\infty} e^{-x} = \infty$. Turning to the natural logarithm function, we have $\lim_{x\to0^+} \ln(x) = -\infty$ and $\lim_{x\to\infty} \ln(x) = \infty$. While both e^x and $\ln(x)$ grow without bound as $x\to\infty$, the exponential function does so much more quickly than the logarithm function does. We'll soon use limits to quantify what we mean by "quickly."

For polynomial functions of the form $p(x) = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0$, the end behavior depends on the sign of a_n and whether the highest power n is even or odd. If n is even and a_n is positive, then $\lim_{x\to\infty} p(x) = \infty$ and $\lim_{x\to-\infty} p(x) = \infty$, as in the plot of g in Figure 2.8.5. If instead a_n is negative, then $\lim_{x\to\infty} p(x) = -\infty$ and $\lim_{x\to-\infty} p(x) = -\infty$. In the situation where n is odd, then either $\lim_{x\to\infty} p(x) = \infty$ and $\lim_{x\to-\infty} p(x) = -\infty$ (which occurs when a_n is positive, as in the graph of f in Figure 2.8.5), or $\lim_{x\to\infty} p(x) = -\infty$ and $\lim_{x\to-\infty} p(x) = \infty$ (when a_n is negative).

A function can fail to have a limit as $x \to \infty$. For example, consider the plot of the sine function at right in Figure 2.8.5. Because the function continues oscillating between -1 and 1 as $x \to \infty$, we say that $\lim_{x \to \infty} \sin(x)$ does not exist.

Finally, it is straightforward to analyze the behavior of any rational function as $x \to \infty$. Consider, for example, the function

$$q(x) = \frac{3x^2 - 4x + 5}{7x^2 + 9x - 10}.$$

Note that both $(3x^2 - 4x + 5) \to \infty$ as $x \to \infty$ and $(7x^2 + 9x - 10) \to \infty$ as $x \to \infty$. Here we say that $\lim_{x\to\infty} q(x)$ has indeterminate form $\frac{\infty}{\infty}$, much like we did when we encountered limits of the form $\frac{0}{0}$. We can determine the value of this limit through a standard algebraic approach. Multiplying the numerator and denominator each by $\frac{1}{x^2}$, we find that

$$\lim_{x \to \infty} q(x) = \lim_{x \to \infty} \frac{3x^2 - 4x + 5}{7x^2 + 9x - 10} \cdot \frac{\frac{1}{x^2}}{\frac{1}{x^2}}$$
$$= \lim_{x \to \infty} \frac{3 - 4\frac{1}{x} + 5\frac{1}{x^2}}{7 + 9\frac{1}{x} - 10\frac{1}{x^2}} = \frac{3}{7}$$

since $\frac{1}{x^2} \to 0$ and $\frac{1}{x} \to 0$ as $x \to \infty$. This shows that the rational function q has a horizontal asymptote at $y = \frac{3}{7}$. A similar approach can be used to determine the limit of any rational function as $x \to \infty$.

But how should we handle a limit such as

$$\lim_{x\to\infty}\frac{x^2}{e^x}?$$

Here, both $x^2 \to \infty$ and $e^x \to \infty$, but there is not an obvious algebraic approach that enables us to find the limit's value. Fortunately, it turns out that L'Hôpital's Rule extends to cases involving infinity.

L'Hôpital's Rule (∞)

If f and g are differentiable and both approach zero or both approach $\pm \infty$ as $x \to a$ (where a is allowed to be ∞), then

$$\lim_{x \to a} \frac{f(x)}{g(x)} = \lim_{x \to a} \frac{f'(x)}{g'(x)}.$$

(To be technically correct, we need to the additional hypothesis that $g'(x) \neq 0$ on an open interval that contains a or in every neighborhood of infinity if a is ∞ ; this is almost always met in practice.)

To evaluate $\lim_{x\to\infty} \frac{x^2}{e^x}$, we observe that we can apply L'Hôpital's Rule, since both $x^2\to\infty$ and $e^x\to\infty$. Doing so, it follows that

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

This updated limit is still indeterminate and of the form $\frac{\infty}{\infty}$, but it is simpler since 2x has replaced x^2 . Hence, we can apply L'Hôpital's Rule again, by which we find that

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

Now, since 2 is constant and $e^x \to \infty$ as $x \to \infty$, it follows that $\frac{2}{e^x} \to 0$ as $x \to \infty$, which shows that

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

Activity 2.8.4. Evaluate each of the following limits. If you use L'Hôpital's Rule, indicate where it was used, and be certain its hypotheses are met before you apply it.

a.
$$\lim_{x\to\infty} \frac{x}{\ln(x)}$$

d.
$$\lim_{x \to \frac{\pi}{2}^{-}} \frac{\tan(x)}{x - \frac{\pi}{2}}$$

e. $\lim_{x \to \infty} xe^{-x}$

a.
$$\lim_{x\to\infty} \frac{x}{\ln(x)}$$

b. $\lim_{x\to\infty} \frac{e^x + x}{2e^x + x^2}$

e.
$$\lim_{r\to\infty} xe^{-x}$$

c.
$$\lim_{x\to 0^+} \frac{\ln(x)}{\frac{1}{x}}$$

When we are considering the limit of a quotient of two functions $\frac{f(x)}{g(x)}$ that results in an indeterminate form of $\frac{\infty}{\infty}$, in essence we are asking which function is growing faster without bound. We say that the function *g* dominates the function *f* as $x \to \infty$ provided that

$$\lim_{x \to \infty} \frac{f(x)}{g(x)} = 0,$$

whereas f dominates g provided that $\lim_{x\to\infty}\frac{f(x)}{g(x)}=\infty$. Finally, if the value of $\lim_{x\to\infty}\frac{f(x)}{g(x)}$ is finite and nonzero, we say that f and g grow at the same rate. For example, from earlier work we know that $\lim_{x\to\infty}\frac{x^2}{e^x}=0$, so e^x dominates x^2 , while $\lim_{x\to\infty}\frac{3x^2-4x+5}{7x^2+9x-10}=\frac{3}{7}$, so $f(x)=3x^2-4x+5$ and $g(x)=7x^2+9x-10$ grow at the same rate.

Summary

• Derivatives be used to help us evaluate indeterminate limits of the form $\frac{0}{0}$ through L'Hôpital's Rule, which is developed by replacing the functions in the numerator and denominator with their tangent line approximations. In particular, if f(a) = g(a) = 0and f and q are differentiable at a, L'Hôpital's Rule tells us that

$$\lim_{x \to a} \frac{f(x)}{g(x)} = \lim_{x \to a} \frac{f'(x)}{g'(x)}.$$

• When we write $x \to \infty$, this means that x is increasing without bound. We thus use ∞ along with limit notation to write $\lim_{x\to\infty} f(x) = L$, which means we can make f(x) as close to L as we like by choosing x to be sufficiently large, and similarly $\lim_{x\to a} f(x) = \infty$, which means we can make f(x) as large as we like by choosing x sufficiently close to a.

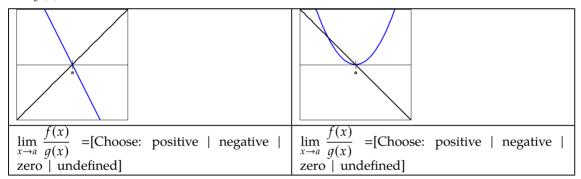
• A version of L'Hôpital's Rule also allows us to use derivatives to assist us in evaluating indeterminate limits of the form $\frac{\infty}{\infty}$. In particular, If f and g are differentiable and both approach zero or both approach $\pm \infty$ as $x \to a$ (where a is allowed to be ∞), then

$$\lim_{x \to a} \frac{f(x)}{g(x)} = \lim_{x \to a} \frac{f'(x)}{g'(x)}.$$

Exercises



1. For the figures below, determine the nature of $\lim_{x\to a} \frac{f(x)}{g(x)}$, if f(x) is shown as the blue curve and g(x) as the black curve.





2. Find the limit: $\lim_{x \to 4} \frac{\ln(x/4)}{x^2 - 16} =$ (*Enter* undefined *if the limit does not exist.*)



3. Compute the following limits using l'Hôpital's rule if appropriate. Use INF to denote ∞ and MINF to denote $-\infty$.

and MINF to denote
$$-\infty$$
.
$$\lim_{x \to 0} \frac{1 - \cos(7x)}{1 - \cos(6x)} = \lim_{x \to 1} \frac{7^x - 6^x - 1}{x^2 - 1} = \boxed{}$$



4. Evaluate the limit using L'Hopital's rule.

$$\lim_{x \to \infty} \frac{13x^2}{e^{8x}} = \frac{1}{1} \text{help (limits)}$$

5. Let f and g be differentiable functions about which the following information is known: f(3) = g(3) = 0, f'(3) = g'(3) = 0, f''(3) = -2, and g''(3) = 1. Let a new function h be given by the rule $h(x) = \frac{f(x)}{g(x)}$. On the same set of axes, sketch possible graphs of f and g near x = 3, and use the provided information to determine the value of

$$\lim_{x\to 3} h(x).$$

Provide explanation to support your conclusion.

6. Find all vertical and horizontal asymptotes of the function

$$R(x) = \frac{3(x-a)(x-b)}{5(x-a)(x-c)},$$

where a, b, and c are distinct, arbitrary constants. In addition, state all values of x for which R is not continuous. Sketch a possible graph of R, clearly labeling the values of a, b, and c.

- 7. Consider the function $g(x) = x^{2x}$, which is defined for all x > 0. Observe that $\lim_{x \to 0^+} g(x)$ is indeterminate due to its form of 0^0 . (Think about how we know that $0^k = 0$ for all k > 0, while $b^0 = 1$ for all $b \ne 0$, but that neither rule can apply to 0^0 .)
 - a. Let $h(x) = \ln(g(x))$. Explain why $h(x) = 2x \ln(x)$.
 - b. Next, explain why it is equivalent to write $h(x) = \frac{2\ln(x)}{\frac{1}{x}}$.
 - c. Use L'Hôpital's Rule and your work in (b) to compute $\lim_{x\to 0^+} h(x)$.
 - d. Based on the value of $\lim_{x\to 0^+} h(x)$, determine $\lim_{x\to 0^+} g(x)$.
- **8.** Recall we say that function g **dominates** function f provided that $\lim_{x\to\infty} f(x) = \infty$, $\lim_{x\to\infty} g(x) = \infty$, and $\lim_{x\to\infty} \frac{f(x)}{g(x)} = 0$.
 - a. Which function dominates the other: ln(x) or \sqrt{x} ?
 - b. Which function dominates the other: ln(x) or $\sqrt[n]{x}$? (n can be any positive integer)
 - c. Explain why e^x will dominate any polynomial function.
 - d. Explain why x^n will dominate ln(x) for any positive integer n.
 - e. Give any example of two nonlinear functions such that neither dominates the other.

Using Derivatives

3.1 Using derivatives to identify extreme values

Motivating Questions

- What are the critical numbers of a function *f* and how are they connected to identifying the most extreme values the function achieves?
- How does the first derivative of a function reveal important information about the behavior of the function, including the function's extreme values?
- How can the second derivative of a function be used to help identify extreme values of the function?

In many different settings, we are interested in knowing where a function achieves its least and greatest values. These can be important in applications — say to identify a point at which maximum profit or minimum cost occurs — or in theory to understand how to characterize the behavior of a function or a family of related functions. Consider the simple and familiar example of a parabolic function such as $s(t) = -16t^2 + 32t + 48$ (shown at left in Figure 3.1.1) that represents the height of an object tossed vertically: its maximum value occurs at the vertex of the parabola and represents the highest value that the object reaches. Moreover, this maximum value identifies an especially important point on the graph, the point at which the curve changes from increasing to decreasing.

More generally, for any function we consider, we can investigate where its lowest and highest points occur in comparison to points nearby or to all possible points on the graph. Given a function f, we say that f(c) is a global or absolute maximum provided that $f(c) \ge f(x)$ for all x in the domain of f, and similarly call f(c) a global or absolute minimum whenever $f(c) \le f(x)$ for all x in the domain of f. For instance, for the function g given at right in Figure 3.1.1, g has a global maximum of g(c), but g does not appear to have a global minimum, as the graph of g seems to decrease without bound. We note that the point (c,g(c)) marks a fundamental change in the behavior of g, where g changes from increasing to decreasing; similar things happen at both (a,g(a)) and (b,g(b)), although these points are not global mins or maxes.



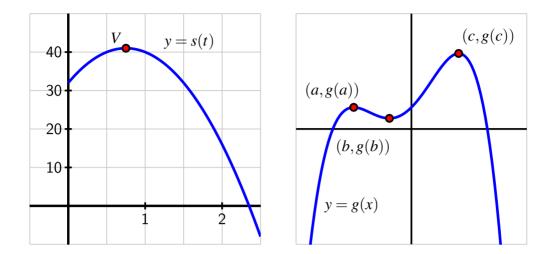


Figure 3.1.1: At left, $s(t) = -16t^2 + 24t + 32$ whose vertex is $(\frac{3}{4}, 41)$; at right, a function g that demonstrates several high and low points.

For any function f, we say that f(c) is a *local maximum* or *relative maximum* provided that $f(c) \ge f(x)$ for all x near c, while f(c) is called a *local* or *relative minimum* whenever $f(c) \le f(x)$ for all x near c. Any maximum or minimum may be called an *extreme value* of f. For example, in Figure 3.1.1, g has a relative minimum of g(b) at the point (b, g(b)) and a relative maximum of g(a) at (a, g(a)). We have already identified the global maximum of g(a) as g(c); this global maximum can also be considered a relative maximum.

We would like to use fundamental calculus ideas to help us identify and classify key function behavior, including the location of relative extremes. Of course, if we are given a graph of a function, it is often straightforward to locate these important behaviors visually. We investigate this situation in the following preview activity.

Preview Activity 3.1.1. Consider the function h given by the graph in Figure 3.1.2. Use the graph to answer each of the following questions.

- a. Identify all of the values of c for which h(c) is a local maximum of h.
- b. Identify all of the values of c for which h(c) is a local minimum of h.
- c. Does h have a global maximum on the interval [-3,3]? If so, what is the value of this global maximum?
- d. Does h have a global minimum on the interval [-3, 3]? If so, what is its value?
- e. Identify all values of c for which h'(c) = 0.

- f. Identify all values of c for which h'(c) does not exist.
- g. True or false: every relative maximum and minimum of h occurs at a point where h'(c) is either zero or does not exist.
- h. True or false: at every point where h'(c) is zero or does not exist, h has a relative maximum or minimum.

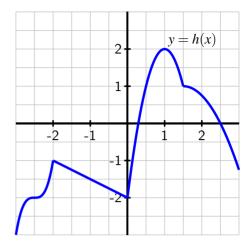


Figure 3.1.2: The graph of a function h on the interval [-3,3].

3.1.1 Critical numbers and the first derivative test

If a function has a relative extreme value at a point (c, f(c)), the function must change its behavior at c regarding whether it is increasing or decreasing before or after the point.

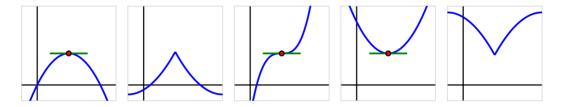


Figure 3.1.3: From left to right, a function with a relative maximum where its derivative is zero; a function with a relative maximum where its derivative is undefined; a function with neither a maximum nor a minimum at a point where its derivative is zero; a function with a relative minimum where its derivative is zero; and a function with a relative minimum where its derivative is undefined.

For example, if a continuous function has a relative maximum at c, such as those pictured in the two leftmost functions in Figure 3.1.3, then it is both necessary and sufficient that the function change from being increasing just before c to decreasing just after c. In the same way, a continuous function has a relative minimum at c if and only if the function changes from decreasing to increasing at c. See, for instance, the two functions pictured at right in Figure 3.1.3. There are only two possible ways for these changes in behavior to occur: either f'(c) = 0 or f'(c) is undefined.

Because these values of c are so important, we call them *critical numbers*. More specifically, we say that a function f has a *critical number* at x = c provided that c is in the domain of f, and f'(c) = 0 or f'(c) is undefined. Critical numbers provide us with the only possible locations where the function f may have relative extremes. Note that not every critical number produces a maximum or minimum; in the middle graph of Figure 3.1.3, the function pictured there has a horizontal tangent line at the noted point, but the function is increasing before and increasing after, so the critical number does not yield a location where the function is greater than every value nearby, nor less than every value nearby.

We also sometimes use the terminology that, when c is a critical number, that (c, f(c)) is a *critical point* of the function, or that f(c) is a *critical value*.

The *first derivative test* summarizes how sign changes in the first derivative indicate the presence of a local maximum or minimum for a given function.

First Derivative Test

If p is a critical number of a continuous function f that is differentiable near p (except possibly at x = p), then f has a relative maximum at p if and only if f' changes sign from positive to negative at p, and f has a relative minimum at p if and only if f' changes sign from negative to positive at p.

We consider an example to show one way the first derivative test can be used to identify the relative extreme values of a function.

Example 3.1.4. Let f be a function whose derivative is given by the formula $f'(x) = e^{-2x}(3-x)(x+1)^2$. Determine all critical numbers of f and decide whether a relative maximum, relative minimum, or neither occurs at each.

Solution. Since we already have f'(x) written in factored form, it is straightforward to find the critical numbers of f. Since f'(x) is defined for all values of x, we need only determine where f'(x) = 0. From the equation

$$e^{-2x}(3-x)(x+1)^2 = 0$$

and the zero product property, it follows that x = 3 and x = -1 are critical numbers of f. (Note particularly that there is no value of x that makes $e^{-2x} = 0$.)

Next, to apply the first derivative test, we'd like to know the sign of f'(x) at inputs near the critical numbers. Because the critical numbers are the only locations at which f' can change sign, it follows that the sign of the derivative is the same on each of the intervals created by the critical numbers: for instance, the sign of f' must be the same for every x < -1. We create a first derivative sign chart to summarize the sign of f' on the relevant intervals along with the corresponding behavior of f.

Figure 3.1.5: The first derivative sign chart for a function f whose derivative is given by the formula $f'(x) = e^{-2x}(3-x)(x+1)^2$.

The first derivative sign chart in Figure 3.1.5 comes from thinking about the sign of each of the terms in the factored form of f'(x) at one selected point in the interval under consideration. For instance, for x < -1, we could consider x = -2 and determine the sign of e^{-2x} , (3-x), and $(x+1)^2$ at the value x = -2. We note that both e^{-2x} and $(x+1)^2$ are positive regardless of the value of x, while (3-x) is also positive at x = -2. Hence, each of the three terms in f' is positive, which we indicate by writing "+ + +." Taking the product of three positive terms obviously results in a value that is positive, which we denote by the "+" in the interval to the left of x = -1 indicating the overall sign of f'. And, since f' is positive on that interval, we further know that f is increasing, which we summarize by writing "INC" to represent the corresponding behavior of f. In a similar way, we find that f' is positive and f is increasing on -1 < x < 3, and f' is negative and f is decreasing for x > 3.

Now, by the first derivative test, to find relative extremes of f we look for critical numbers at which f' changes sign. In this example, f' only changes sign at x=3, where f' changes from positive to negative, and thus f has a relative maximum at x=3. While f has a critical number at x=-1, since f is increasing both before and after x=-1, f has neither a minimum nor a maximum at f and f has neither a minimum nor a maximum at f has neither a minimum nor a maximum

Activity 3.1.2. Suppose that g(x) is a function continuous for every value of $x \ne 2$ whose first derivative is $g'(x) = \frac{(x+4)(x-1)^2}{x-2}$. Further, assume that it is known that g has a vertical asymptote at x = 2.

- a. Determine all critical numbers of q.
- b. By developing a carefully labeled first derivative sign chart, decide whether g has as a local maximum, local minimum, or neither at each critical number.
- c. Does *q* have a global maximum? global minimum? Justify your claims.
- d. What is the value of $\lim_{x\to\infty} g'(x)$? What does the value of this limit tell you about the long-term behavior of g?

e. Sketch a possible graph of y = g(x).

3.1.2 The second derivative test

Recall that the second derivative of a function tells us several important things about the behavior of the function itself. For instance, if f'' is positive on an interval, then we know that f' is increasing on that interval and, consequently, that f is concave up, which also tells us that throughout the interval the tangent line to y = f(x) lies below the curve at every point. In this situation where we know that f'(p) = 0, it turns out that the sign of the second derivative determines whether f has a local minimum or local maximum at the critical number p.

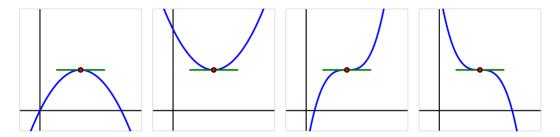


Figure 3.1.6: Four possible graphs of a function f with a horizontal tangent line at a critical point.

In Figure 3.1.6, we see the four possibilities for a function f that has a critical number p at which f'(p) = 0, provided f''(p) is not zero on an interval including p (except possibly at p). On either side of the critical number, f'' can be either positive or negative, and hence f can be either concave up or concave down. In the first two graphs, f does not change concavity at p, and in those situations, f has either a local minimum or local maximum. In particular, if f'(p) = 0 and f''(p) < 0, then we know f is concave down at p with a horizontal tangent line, and this guarantees f has a local maximum there. This fact, along with the corresponding statement for when f''(p) is positive, is stated in the second derivative test.

Second Derivative Test

If p is a critical number of a continuous function f such that f'(p) = 0 and $f''(p) \neq 0$, then f has a relative maximum at p if and only if f''(p) < 0, and f has a relative minimum at p if and only if f''(p) > 0.

In the event that f''(p) = 0, the second derivative test is inconclusive. That is, the test doesn't provide us any information. This is because if f''(p) = 0, it is possible that f has a local minimum, local maximum, or neither.¹

¹Consider the functions $f(x) = x^4$, $g(x) = -x^4$, and $h(x) = x^3$ at the critical point p = 0.

Just as a first derivative sign chart reveals all of the increasing and decreasing behavior of a function, we can construct a second derivative sign chart that demonstrates all of the important information involving concavity.

Example 3.1.7. Let f(x) be a function whose first derivative is $f'(x) = 3x^4 - 9x^2$. Construct both first and second derivative sign charts for f, fully discuss where f is increasing and decreasing and concave up and concave down, identify all relative extreme values, and sketch a possible graph of f.

Solution. Since we know $f'(x) = 3x^4 - 9x^2$, we can find the critical numbers of f by solving $3x^4 - 9x^2 = 0$. Factoring, we observe that

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

so that $x = 0, \pm \sqrt{3}$ are the three critical numbers of f. It then follows that the first derivative sign chart for f is given in Figure 3.1.8.

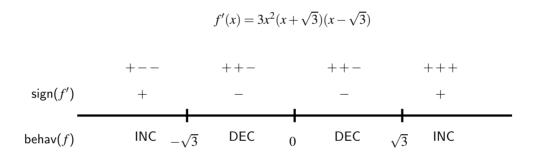


Figure 3.1.8: The first derivative sign chart for f when $f'(x) = 3x^4 - 9x^2 = 3x^2(x^2 - 3)$.

Thus, f is increasing on the intervals $(-\infty, -\sqrt{3})$ and $(\sqrt{3}, \infty)$, while f is decreasing on $(-\sqrt{3}, 0)$ and $(0, \sqrt{3})$. Note particularly that by the first derivative test, this information tells us that f has a local maximum at $x = -\sqrt{3}$ and a local minimum at $x = \sqrt{3}$. While f also has a critical number at x = 0, neither a maximum nor minimum occurs there since f' does not change sign at x = 0.

Next, we move on to investigate concavity. Differentiating $f'(x) = 3x^4 - 9x^2$, we see that $f''(x) = 12x^3 - 18x$. Since we are interested in knowing the intervals on which f'' is positive and negative, we first find where f''(x) = 0. Observe that

$$0 = 12x^3 - 18x = 12x\left(x^2 - \frac{3}{2}\right) = 12x\left(x + \sqrt{\frac{3}{2}}\right)\left(x - \sqrt{\frac{3}{2}}\right),$$

which implies that $x = 0, \pm \sqrt{\frac{3}{2}}$. Building a sign chart for f'' in the exact same way we do for f', we see the result shown in Figure 3.1.9.

Figure 3.1.9: The second derivative sign chart for f when $f''(x) = 12x^3 - 18x = 12x^2\left(x^2 - \sqrt{\frac{3}{2}}\right)$.

Therefore, f is concave down on the intervals $(-\infty, -\sqrt{\frac{3}{2}})$ and $(0, \sqrt{\frac{3}{2}})$, and concave up on $(-\sqrt{\frac{3}{2}}, 0)$ and $(\sqrt{\frac{3}{2}}, \infty)$.

Putting all of the above information together, we now see a complete and accurate possible graph of f in Figure 3.1.10.

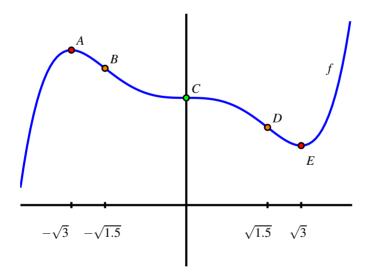


Figure 3.1.10: A possible graph of the function *f* in Example 3.1.7.

The point $A = (-\sqrt{3}, f(-\sqrt{3}))$ is a local maximum, as f is increasing prior to A and decreasing after; similarly, the point $E = (\sqrt{3}, f(\sqrt{3}))$ is a local minimum. Note, too, that f is concave down at A and concave up at B, which is consistent both with our second derivative sign chart and the second derivative test. At points B and D, concavity changes, as we saw in the

results of the second derivative sign chart in Figure 3.1.9. Finally, at point C, f has a critical point with a horizontal tangent line, but neither a maximum nor a minimum occurs there since f is decreasing both before and after C. It is also the case that concavity changes at C.

While we completely understand where f is increasing and decreasing, where f is concave up and concave down, and where f has relative extremes, we do not know any specific information about the y-coordinates of points on the curve. For instance, while we know that f has a local maximum at $x = -\sqrt{3}$, we don't know the value of that maximum because we do not know $f(-\sqrt{3})$. Any vertical translation of our sketch of f in Figure 3.1.10 would satisfy the given criteria for f.

Points B, C, and D in Figure 3.1.10 are locations at which the concavity of f changes. We give a special name to any such point: if p is a value in the domain of a continuous function f at which f changes concavity, then we say that (p, f(p)) is an *inflection point* of f. Just as we look for locations where f changes from increasing to decreasing at points where f'(p) = 0 or f'(p) is undefined, so too we find where f''(p) = 0 or f''(p) is undefined to see if there are points of inflection at these locations.

It is important at this point in our study to remind ourselves of the big picture that derivatives help to paint: the sign of the first derivative f' tells us *whether* the function f is increasing or decreasing, while the sign of the second derivative f'' tells us *how* the function f is increasing or decreasing.

Activity 3.1.3. Suppose that g is a function whose second derivative, g'', is given by the following graph.

- a. Find the *x*-coordinates of all points of inflection of *q*.
- b. Fully describe the concavity of *g* by making an appropriate sign chart.
- c. Suppose you are given that g'(-1.67857351) = 0. Is there is a local maximum, local minimum, or neither (for the function g) at this critical number of g, or is it impossible to say? Why?

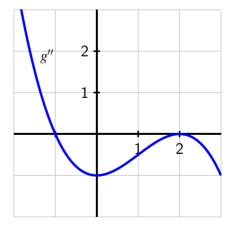


Figure 3.1.11: The graph of y = g''(x).

d. Assuming that g''(x) is a polynomial (and that all important behavior of g'' is seen

in the graph above), what degree polynomial do you think g(x) is? Why?

As we will see in more detail in the following section, derivatives also help us to understand families of functions that differ only by changing one or more parameters. For instance, we might be interested in understanding the behavior of all functions of the form $f(x) = a(x - h)^2 + k$ where a, h, and k are numbers that may vary. In the following activity, we investigate a particular example where the value of a single parameter has considerable impact on how the graph appears.

Activity 3.1.4. Consider the family of functions given by $h(x) = x^2 + \cos(kx)$, where k is an arbitrary positive real number.

- a. Use a graphing utility to sketch the graph of h for several different k-values, including k = 1, 3, 5, 10. Plot $h(x) = x^2 + \cos(3x)$ on the axes provided. What is the smallest value of k at which you think you can see (just by looking at the graph) at least one inflection point on the graph of h?
- b. Explain why the graph of h has no inflection points if $k \le \sqrt{2}$, but infinitely many inflection points if $k > \sqrt{2}$.
- c. Explain why, no matter the value of *k*, *h* can only have finitely many critical numbers.

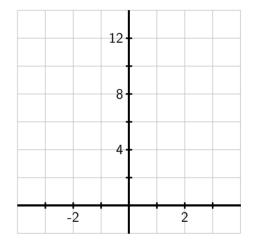


Figure 3.1.12: Axes for plotting y = h(x).

Summary

- The critical numbers of a continuous function f are the values of p for which f'(p) = 0 or f'(p) does not exist. These values are important because they identify horizontal tangent lines or corner points on the graph, which are the only possible locations at which a local maximum or local minimum can occur.
- Given a differentiable function f, whenever f' is positive, f is increasing; whenever f' is negative, f is decreasing. The first derivative test tells us that at any point where f changes from increasing to decreasing, f has a local maximum, while conversely at

any point where f changes from decreasing to increasing f has a local minimum.

• Given a twice differentiable function f, if we have a horizontal tangent line at x = pand f''(p) is nonzero, then the fact that f'' tells us the concavity of f will determine whether f has a maximum or minimum at x = p. In particular, if f'(p) = 0 and f''(p) < 0, then f is concave down at p and f has a local maximum there, while if f'(p) = 0 and f''(p) > 0, then f has a local minimum at p. If f'(p) = 0 and f''(p) = 0, then the second derivative does not tell us whether f has a local extreme at p or not.

Exercises

1. Use a graph below of $f(x) = 3e^{-9x^2}$ to estimate the x-values of any critical points and inflection points of f(x).



	•		
	\		
	.\		
		1.0	
-1			

critical points (enter as a comma-separated list): x =inflection points (enter as a comma-separated list): x =Next, use derivatives to find the x-values of any critical points and inflection points exactly. critical points (enter as a comma-separated list): x =inflection points (enter as a comma-separated list): x =



2. Find the inflection points of $f(x) = 2x^4 + 27x^3 - 21x^2 + 15$. (Give your answers as a comma separated list, e.g., 3,-2.)

inflection points =



3. The following shows graphs of three functions, A (in black), B (in blue), and C (in green). If these are the graphs of three functions f, f', and f'', identify which is which.





(For each enter A, B or C).

- 4. This problem concerns a function about which the following information is known:
 - *f* is a differentiable function defined at every real number *x*
 - f(0) = -1/2

• y = f'(x) has its graph given at center in Figure 3.1.13

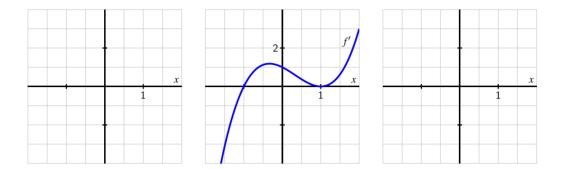


Figure 3.1.13: At center, a graph of y = f'(x); at left, axes for plotting y = f(x); at right, axes for plotting y = f''(x).

- a. Construct a first derivative sign chart for f. Clearly identify all critical numbers of f, where f is increasing and decreasing, and where f has local extrema.
- b. On the right-hand axes, sketch an approximate graph of y = f''(x).
- c. Construct a second derivative sign chart for f. Clearly identify where f is concave up and concave down, as well as all inflection points.
- d. On the left-hand axes, sketch a possible graph of y = f(x).
- **5.** Suppose that g is a differentiable function and g'(2) = 0. In addition, suppose that on 1 < x < 2 and 2 < x < 3 it is known that g'(x) is positive.
 - a. Does g have a local maximum, local minimum, or neither at x = 2? Why?
 - b. Suppose that g''(x) exists for every x such that 1 < x < 3. Reasoning graphically, describe the behavior of g''(x) for x-values near 2.
 - c. Besides being a critical number of g, what is special about the value x = 2 in terms of the behavior of the graph of g?
- **6.** Suppose that h is a differentiable function whose first derivative is given by the graph in Figure 3.1.14.

- a. How many real number solutions can the equation h(x) = 0 have? Why?
- b. If h(x) = 0 has two distinct real solutions, what can you say about the signs of the two solutions? Why?
- c. Assume that $\lim_{x\to\infty}h'(x) = 3$, as appears to be indicated in Figure 3.1.14. How will the graph of y = h(x) appear as $x\to\infty$? Why?
- d. Describe the concavity of y = h(x) as fully as you can from the provided information.

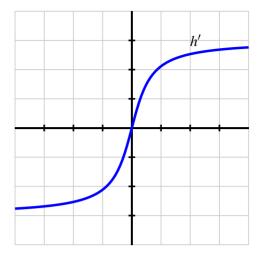


Figure 3.1.14: The graph of y = h'(x).

- 7. Let *p* be a function whose second derivative is $p''(x) = (x+1)(x-2)e^{-x}$.
 - a. Construct a second derivative sign chart for p and determine all inflection points of p.
 - b. Suppose you also know that $x = \frac{\sqrt{5}-1}{2}$ is a critical number of p. Does p have a local minimum, local maximum, or neither at $x = \frac{\sqrt{5}-1}{2}$? Why?
 - c. If the point $(2, \frac{12}{e^2})$ lies on the graph of y = p(x) and $p'(2) = -\frac{5}{e^2}$, find the equation of the tangent line to y = p(x) at the point where x = 2. Does the tangent line lie above the curve, below the curve, or neither at this value? Why?

3.3 Global Optimization

Motivating Questions

- What are the differences between finding relative extreme values and global extreme values of a function?
- How is the process of finding the global maximum or minimum of a function over the function's entire domain different from determining the global maximum or minimum on a restricted domain?
- For a function that is guaranteed to have both a global maximum and global minimum on a closed, bounded interval, what are the possible points at which these extreme values occur?

We have seen that we can use the first derivative of a function to determine where the function is increasing or decreasing, and the second derivative to know where the function is concave up or concave down. Each of these approaches provides us with key information that helps us determine the overall shape and behavior of the graph, as well as whether the function has a relative minimum or relative maximum at a given critical number. Remember that the difference between a relative maximum and a global maximum is that there is a relative maximum of f at x = p if $f(p) \ge f(x)$ for all x near p, while there is a global maximum at p if $f(p) \ge f(x)$ for all x in the domain of f.

For instance, in Figure 3.3.1, we see a function f that has a global maximum at x = c and a relative maximum at x = a, since f(c) is greater than f(x) for every value of x, while f(a) is only greater than the value of f(x) for x near a. Since the function appears to decrease without bound, f has no global minimum, though clearly f has a relative minimum at x = b.

Our emphasis in this section is on finding the global extreme values of a function (if they exist). In so doing, we will either be interested in the behavior of the function over its entire domain or on some restricted portion. The former situation is familiar and similar to work that we did in the two preceding sections of the text. We explore this through a particular example in the following preview activity.

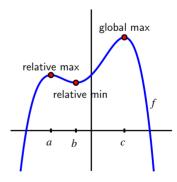


Figure 3.3.1: A function *f* with a global maximum, but no global minimum.

Preview Activity 3.3.1. Let $f(x) = 2 + \frac{3}{1 + (x+1)^2}$.

- a. Determine all of the critical numbers of f.
- b. Construct a first derivative sign chart for f and thus determine all intervals on which f is increasing or decreasing.
- c. Does *f* have a global maximum? If so, why, and what is its value and where is the maximum attained? If not, explain why.
- d. Determine $\lim_{x\to\infty} f(x)$ and $\lim_{x\to-\infty} f(x)$.
- e. Explain why f(x) > 2 for every value of x.
- f. Does *f* have a global minimum? If so, why, and what is its value and where is the minimum attained? If not, explain why.

3.3.1 Global Optimization

For the functions in Figure 3.3.1 and Preview Activity 3.3.1, we were interested in finding the global minimum and global maximum on the entire domain, which turned out to be $(-\infty, \infty)$ for each. At other times, our perspective on a function might be more focused due to some restriction on its domain. For example, rather than considering $f(x) = 2 + \frac{3}{1 + (x+1)^2}$ for every value of x, perhaps instead we are only interested in those x for which $0 \le x \le 4$, and we would like to know which values of x in the interval [0,4] produce the largest possible and smallest possible values of x. We are accustomed to critical numbers playing a key role in determining the location of extreme values of a function; now, by restricting the domain to an interval, it makes sense that the endpoints of the interval will also be important to consider, as we see in the following activity. When limiting ourselves to a particular interval, we will often refer to the *absolute* maximum or minimum value, rather than the *global* maximum or minimum.

Activity 3.3.2. Let
$$g(x) = \frac{1}{3}x^3 - 2x + 2$$
.

- a. Find all critical numbers of *q* that lie in the interval $-2 \le x \le 3$.
- b. Use a graphing utility to construct the graph of g on the interval $-2 \le x \le 3$.
- c. From the graph, determine the x-values at which the absolute minimum and absolute maximum of g occur on the interval [-2,3].
- d. How do your answers change if we instead consider the interval $-2 \le x \le 2$?
- e. What if we instead consider the interval $-2 \le x \le 1$?

In Activity 3.3.2, we saw how the absolute maximum and absolute minimum of a function on a closed, bounded interval [a, b], depend not only on the critical numbers of the function, but also on the selected values of a and b. These observations demonstrate several important facts that hold much more generally. First, we state an important result called the Extreme

Value Theorem.

The Extreme Value Theorem

If f is a continuous function on a closed interval [a, b], then f attains both an absolute minimum and absolute maximum on [a, b]. That is, for some value x_m such that $a \le a$ $x_m \le b$, it follows that $f(x_m) \le f(x)$ for all x in [a,b]. Similarly, there is a value x_M in [a, b] such that $f(x_M) \ge f(x)$ for all x in [a, b]. Letting $m = f(x_m)$ and $M = f(x_M)$, it follows that $m \leq f(x) \leq M$ for all x in [a, b].

The Extreme Value Theorem tells us that provided a function is continuous, on any closed interval [a, b] the function has to achieve both an absolute minimum and an absolute maximum. Note, however, that this result does not tell us where these extreme values occur, but rather only that they must exist. As seen in the examples of Activity 3.3.2, it is apparent that the only possible locations for relative extremes are either the endpoints of the interval or at a critical number (the latter being where a relative minimum or maximum could occur, which is a potential location for an absolute extreme).

Note 3.3.2. Thus, we have the following approach to finding the absolute maximum and minimum of a continuous function f on the interval [a, b]:

- find all critical numbers of *f* that lie in the interval;
- evaluate the function *f* at each critical number in the interval and at each endpoint of the interval;
- from among the noted function values, the smallest is the absolute minimum of f on the interval, while the largest is the absolute maximum.

Activity 3.3.3. Find the *exact* absolute maximum and minimum of each function on the stated interval.

a.
$$h(x) = xe^{-x}$$
, [0,3]

b.
$$p(t) = \sin(t) + \cos(t), [-\frac{\pi}{2}, \frac{\pi}{2}]$$

c.
$$q(x) = \frac{x^2}{x-2}$$
, [3,7]

c.
$$q(x) = \frac{x^2}{x-2}$$
, [3,7]
d. $f(x) = 4 - e^{-(x-2)^2}$, $(-\infty, \infty)$
e. $h(x) = xe^{-ax}$, $[0, \frac{2}{a}]$ $(a > 0)$

e.
$$h(x) = xe^{-ax}$$
, $[0, \frac{2}{a}]$ $(a > 0)$

e.
$$h(x) = xe^{-x}$$
, $[0, \frac{\pi}{a}]$ $(a > 0)$
f. $f(x) = b - e^{-(x-a)^2}$, $(-\infty, \infty)$, $a, b > 0$

One of the big lessons in finding absolute extreme values is the realization that the interval we choose has nearly the same impact on the problem as the function under consideration. Consider, for instance, the function pictured in Figure 3.3.3.

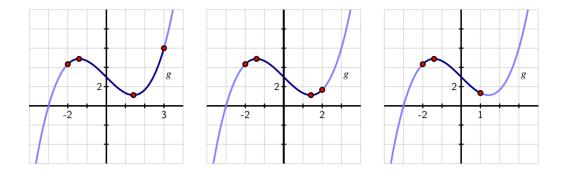


Figure 3.3.3: A function *g* considered on three different intervals.

In sequence, from left to right, as we see the interval under consideration change from [-2,3] to [-2,2] to [-2,1], we move from having two critical numbers in the interval with the absolute minimum at one critical number and the absolute maximum at the right endpoint, to still having both critical numbers in the interval but then with the absolute minimum and maximum at the two critical numbers, to finally having just one critical number in the interval with the absolute maximum at one critical number and the absolute minimum at one endpoint. It is particularly essential to always remember to only consider the critical numbers that lie within the interval.

3.3.2 Moving toward applications

In Section 3.4, we will focus almost exclusively on applied optimization problems: problems where we seek to find the absolute maximum or minimum value of a function that represents some physical situation. We conclude this current section with an example of one such problem because it highlights the role that a closed, bounded domain can play in finding absolute extrema. In addition, these problems often involve considerable preliminary work to develop the function which is to be optimized, and this example demonstrates that process.

Example 3.3.4. A 20 cm piece of wire is cut into two pieces. One piece is used to form a square and the other an equilateral triangle. How should the wire be cut to maximize the total area enclosed by the square and triangle? to minimize the area?

Solution. We begin by constructing a picture that exemplifies the given situation. The primary variable in the problem is where we decide to cut the wire. We thus label that point x, and note that the remaining portion of the wire then has length 20 - x

As shown in Figure 3.3.5, we see that the x cm of the wire that are used to form the equilateral triangle result in a triangle with three sides of length $\frac{x}{3}$. For the remaining 20-x cm of wire, the square that results will have each side of length $\frac{20-x}{4}$.

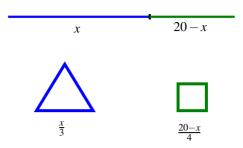


Figure 3.3.5: A 20 cm piece of wire cut into two pieces, one of which forms an equilateral triangle, the other which yields a square.

At this point, we note that there are obvious restrictions on x: in particular, $0 \le x \le 20$. In the extreme cases, all of the wire is being used to make just one figure. For instance, if x = 0, then all 20 cm of wire are used to make a square that is 5×5 .

Now, our overall goal is to find the absolute minimum and absolute maximum areas that can be enclosed. We note that the area of the triangle is $A_{\Delta} = \frac{1}{2}bh = \frac{1}{2} \cdot \frac{x}{3} \cdot \frac{x\sqrt{3}}{6}$, since the height of an equilateral triangle is $\sqrt{3}$ times half the length of the base. Further, the area of the square is $A_{\Box} = s^2 = \left(\frac{20-x}{4}\right)^2$. Therefore, the total area function is

$$A(x) = \frac{\sqrt{3}x^2}{36} + \left(\frac{20 - x}{4}\right)^2.$$

Again, note that we are only considering this function on the restricted domain [0, 20] and we seek its absolute minimum and absolute maximum.

Differentiating A(x), we have

$$A'(x) = \frac{\sqrt{3}x}{18} + 2\left(\frac{20-x}{4}\right)\left(-\frac{1}{4}\right) = \frac{\sqrt{3}}{18}x + \frac{1}{8}x - \frac{5}{2}.$$

Setting A'(x) = 0, it follows that $x = \frac{180}{4\sqrt{3}+9} \approx 11.3007$ is the only critical number of A, and we note that this lies within the interval [0, 20].

Evaluating A at the critical number and endpoints, we see that

•
$$A\left(\frac{180}{4\sqrt{3}+9}\right) = \frac{\sqrt{3}(\frac{180}{4\sqrt{3}+9})^2}{4} + \left(\frac{20 - \frac{180}{4\sqrt{3}+9}}{4}\right)^2 \approx 10.8741$$

- A(0) = 25
- $A(20) = \frac{\sqrt{3}}{36}(400) = \frac{100}{9}\sqrt{3} \approx 19.2450$

Thus, the absolute minimum occurs when $x \approx 11.3007$ and results in the minimum area of approximately 10.8741 square centimeters, while the absolute maximum occurs when we invest all of the wire in the square (and none in the triangle), resulting in 25 square centimeters of area. These results are confirmed by a plot of y = A(x) on the interval [0, 20], as shown in Figure 3.3.6.

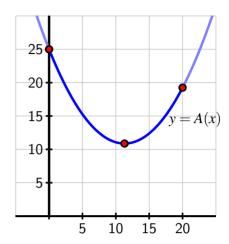


Figure 3.3.6: A plot of the area function from Example 3.3.4.

Activity 3.3.4. A piece of cardboard that is 10×15 (each measured in inches) is being made into a box without a top. To do so, squares are cut from each corner of the box and the remaining sides are folded up. If the box needs to be at least 1 inch deep and no more than 3 inches deep, what is the maximum possible volume of the box? what is the minimum volume? Justify your answers using calculus.

- a. Draw a labeled diagram that shows the given information. What variable should we introduce to represent the choice we make in creating the box? Label the diagram appropriately with the variable, and write a sentence to state what the variable represents.
- b. Determine a formula for the function V (that depends on the variable in (a)) that tells us the volume of the box.
- c. What is the domain of the function *V*? That is, what values of *x* make sense for input? Are there additional restrictions provided in the problem?
- d. Determine all critical numbers of the function V.
- e. Evaluate V at each of the endpoints of the domain and at any critical numbers that lie in the domain.
- f. What is the maximum possible volume of the box? the minimum?

The approaches shown in Example 3.3.4 and experienced in Activity 3.3.4 include standard steps that we undertake in almost every applied optimization problem: we draw a picture to demonstrate the situation, introduce one or more variables to represent quantities that are changing, work to find a function that models the quantity to be optimized, and then decide

an appropriate domain for that function. Once that work is done, we are in the familiar situation of finding the absolute minimum and maximum of a function over a particular domain, at which time we apply the calculus ideas that we have been studying to this point in Chapter 3.

Summary

- To find relative extreme values of a function, we normally use a first derivative sign chart and classify all of the function's critical numbers. If instead we are interested in absolute extreme values, we first decide whether we are considering the entire domain of the function or a particular interval.
- In the case of finding global extremes over the function's entire domain, we again use
 a first or second derivative sign chart in an effort to make overall conclusions about
 whether or not the function can have a absolute maximum or minimum. If we are
 working to find absolute extremes on a restricted interval, then we first identify all
 critical numbers of the function that lie in the interval.
- For a continuous function on a closed, bounded interval, the only possible points at which absolute extreme values occur are the critical numbers and the endpoints. Thus, to find said absolute extremes, we simply evaluate the function at each endpoint and each critical number in the interval, and then we compare the results to decide which is largest (the absolute maximum) and which is smallest (the absolute minimum).

Exercises

- 1. Based on the given information about each function, decide whether the function has global maximum, a global minimum, neither, both, or that it is not possible to say without more information. Assume that each function is twice differentiable and defined for all real numbers, unless noted otherwise. In each case, write one sentence to explain your conclusion.
 - a. f is a function such that f''(x) < 0 for every x.
 - b. g is a function with two critical numbers a and b (where a < b), and g'(x) < 0 for x < a, g'(x) < 0 for a < x < b, and g'(x) > 0 for x > b.
 - c. h is a function with two critical numbers a and b (where a < b), and h'(x) < 0 for x < a, h'(x) > 0 for a < x < b, and h'(x) < 0 for x > b. In addition, $\lim_{x \to \infty} h(x) = 0$ and $\lim_{x \to -\infty} h(x) = 0$.
 - d. p is a function differentiable everywhere except at x = a and p''(x) > 0 for x < a and p''(x) < 0 for x > a.
- **2.** For each family of functions that depends on one or more parameters, determine the function's absolute maximum and absolute minimum on the given interval.
 - a. $p(x) = x^3 a^2x$, [0, a] (a > 0)
 - b. $r(x) = axe^{-bx}$, $\left[\frac{1}{2b}, b\right](a, b > 0)$

c.
$$w(x) = a(1 - e^{-bx}), [b, 3b] (a, b > 0)$$

d.
$$s(x) = \sin(kx), [\frac{\pi}{3k}, \frac{5\pi}{6k}]$$

- **3.** For each of the functions described below (each continuous on [a, b]), state the location of the function's absolute maximum and absolute minimum on the interval [a, b], or say there is not enough information provided to make a conclusion. Assume that any critical numbers mentioned in the problem statement represent all of the critical numbers the function has in [a, b]. In each case, write one sentence to explain your answer.
 - a. $f'(x) \le 0$ for all x in [a, b]
 - b. g has a critical number at c such that a < c < b and g'(x) > 0 for x < c and g'(x) < 0 for x > c
 - c. h(a) = h(b) and h''(x) < 0 for all x in [a, b]
 - d. p(a) > 0, p(b) < 0, and for the critical number c such that a < c < b, p'(x) < 0 for x < c and p'(x) > 0 for x > c
- **4.** Let $s(t) = 3\sin(2(t \frac{\pi}{6})) + 5$. Find the exact absolute maximum and minimum of s on the provided intervals by testing the endpoints and finding and evaluating all relevant critical numbers of s.

a.
$$\left[\frac{\pi}{6}, \frac{7\pi}{6}\right]$$

c.
$$[0, 2\pi]$$

b.
$$[0, \frac{\pi}{2}]$$

d.
$$\left[\frac{\pi}{3}, \frac{5\pi}{6}\right]$$

3.4 Applied Optimization

Motivating Questions

 In a setting where a situation is described for which optimal parameters are sought, how do we develop a function that models the situation and use calculus to find the desired maximum or minimum?

Near the conclusion of Section 3.3, we considered two examples of optimization problems where determining the function to be optimized was part of a broader question. In Example 3.3.4, we sought to use a single piece of wire to build two geometric figures (an equilateral triangle and square) and to understand how various choices for how to cut the wire led to different values of the area enclosed. One of our conclusions was that in order to maximize the total combined area enclosed by the triangle and square, all of the wire must be used to make a square. In the subsequent Activity 3.3.4, we investigated how the volume of a box constructed from a piece of cardboard by removing squares from each corner and folding up the sides depends on the size of the squares removed.

Both of these problems exemplify situations where there is not a function explicitly provided to optimize. Rather, we first worked to understand the given information in the problem, drawing a figure and introducing variables, and then sought to develop a formula for a function that models the quantity (area or volume, in the two examples, respectively) to be optimized. Once the function was established, we then considered what domain was appropriate on which to pursue the desired absolute minimum or maximum (or both). At this point in the problem, we are finally ready to apply the ideas of calculus to determine and justify the absolute minimum or maximum. Thus, what is primarily different about problems of this type is that the problem-solver must do considerable work to introduce variables and develop the correct function and domain to represent the described situation.

Throughout what follows in the current section, the primary emphasis is on the reader solving problems. Initially, some substantial guidance is provided, with the problems progressing to require greater independence as we move along.

Preview Activity 3.4.1. According to U.S. postal regulations, the girth plus the length of a parcel sent by mail may not exceed 108 inches, where by "girth" we mean the perimeter of the smallest end. What is the largest possible volume of a rectangular parcel with a square end that can be sent by mail? What are the dimensions of the package of largest volume?

- a. Let *x* represent the length of one side of the square end and *y* the length of the longer side. Label these quantities appropriately on the image shown in Figure 3.4.1.
- b. What is the quantity to be optimized in this problem? Find a formula for this quantity in terms of *x* and *y*.

- c. The problem statement tells us that the parcel's girth plus length may not exceed 108 inches. In order to maximize volume, we assume that we will actually need the girth plus length to equal 108 inches. What equation does this produce involving *x* and *y*?
- d. Solve the equation you found in (c) for one of x or y (whichever is easier).
- e. Now use your work in (b) and (d) to determine a formula for the volume of the parcel so that this formula is a function of a single variable.
- f. Over what domain should we consider this function? Note that both x and y must be positive; how does the constraint that girth plus length is 108 inches produce intervals of possible values for x and y?
- g. Find the absolute maximum of the volume of the parcel on the domain you established in (f) and hence also determine the dimensions of the box of greatest volume. Justify that you've found the maximum using calculus.

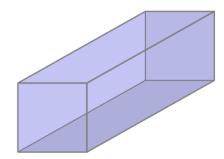


Figure 3.4.1: A rectangular parcel with a square end.

3.4.1 More applied optimization problems

Many of the steps in Preview Activity 3.4.1 are ones that we will execute in any applied optimization problem. We briefly summarize those here to provide an overview of our approach in subsequent questions.

Note 3.4.2.

- Draw a picture and introduce variables. It is essential to first understand what quantities are allowed to vary in the problem and then to represent those values with variables. Constructing a figure with the variables labeled is almost always an essential first step. Sometimes drawing several diagrams can be especially helpful to get a sense of the situation. A nice example of this can be seen at http://gvsu.edu/s/99, where the choice of where to bend a piece of wire into the shape of a rectangle determines both the rectangle's shape and area.
- Identify the quantity to be optimized as well as any key relationships among the variable quantities. Essentially this step involves writing equations that involve the vari-

ables that have been introduced: one to represent the quantity whose minimum or maximum is sought, and possibly others that show how multiple variables in the problem may be interrelated.

- Determine a function of a single variable that models the quantity to be optimized; this may involve using other relationships among variables to eliminate one or more variables in the function formula. For example, in Preview Activity 3.4.1, we initially found that $V = x^2y$, but then the additional relationship that 4x + y = 108 (girth plus length equals 108 inches) allows us to relate x and y and thus observe equivalently that y = 108 4x. Substituting for y in the volume equation yields $V(x) = x^2(108 4x)$, and thus we have written the volume as a function of the single variable x.
- Decide the domain on which to consider the function being optimized. Often the physical constraints of the problem will limit the possible values that the independent variable can take on. Thinking back to the diagram describing the overall situation and any relationships among variables in the problem often helps identify the smallest and largest values of the input variable.
- Use calculus to identify the absolute maximum and/or minimum of the quantity being optimized. This always involves finding the critical numbers of the function first. Then, depending on the domain, we either construct a first derivative sign chart (for an open or unbounded interval) or evaluate the function at the endpoints and critical numbers (for a closed, bounded interval), using ideas we've studied so far in Chapter 3.
- Finally, we make certain we have answered the question: does the question seek the
 absolute maximum of a quantity, or the values of the variables that produce the maximum? That is, finding the absolute maximum volume of a parcel is different from
 finding the dimensions of the parcel that produce the maximum.

Activity 3.4.2. A soup can in the shape of a right circular cylinder is to be made from two materials. The material for the side of the can costs \$0.015 per square inch and the material for the lids costs \$0.027 per square inch. Suppose that we desire to construct a can that has a volume of 16 cubic inches. What dimensions minimize the cost of the can?

- a. Draw a picture of the can and label its dimensions with appropriate variables.
- b. Use your variables to determine expressions for the volume, surface area, and cost of the can.
- c. Determine the total cost function as a function of a single variable. What is the domain on which you should consider this function?
- d. Find the absolute minimum cost and the dimensions that produce this value.

Familiarity with common geometric formulas is particularly helpful in problems like the one in Activity 3.4.2. Sometimes those involve perimeter, area, volume, or surface area. At other times, the constraints of a problem introduce right triangles (where the Pythagorean Theorem applies) or other functions whose formulas provide relationships among variables

present.

Activity 3.4.3. A hiker starting at a point P on a straight road walks east towards point Q, which is on the road and 3 kilometers from point P.

Two kilometers due north of point Q is a cabin. The hiker will walk down the road for a while, at a pace of 8 kilometers per hour. At some point Z between P and Q, the hiker leaves the road and makes a straight line towards the cabin through the woods, hiking at a pace of 3 kph, as pictured in Figure 3.4.3. In order to minimize the time to go from P to Z to the cabin, where should the hiker turn into the forest?

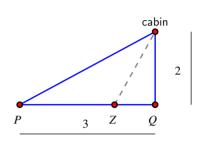


Figure 3.4.3: A hiker walks from *P* to *Z* to the cabin, as pictured.

In more geometric problems, we often use curves or functions to provide natural constraints. For instance, we could investigate which isosceles triangle that circumscribes a unit circle has the smallest area, which you can explore for yourself at http://gvsu.edu/s/9b. Or similarly, for a region bounded by a parabola, we might seek the rectangle of largest area that fits beneath the curve, as shown at http://gvsu.edu/s/9c. The next activity is similar to the latter problem.

Activity 3.4.4. Consider the region in the x-y plane that is bounded by the x-axis and the function $f(x) = 25 - x^2$. Construct a rectangle whose base lies on the x-axis and is centered at the origin, and whose sides extend vertically until they intersect the curve $y = 25 - x^2$. Which such rectangle has the maximum possible area? Which such rectangle has the greatest perimeter? Which has the greatest combined perimeter and area? (Challenge: answer the same questions in terms of positive parameters a and b for the function $f(x) = b - ax^2$.)

Activity 3.4.5. A trough is being constructed by bending a 4×24 (measured in feet) rectangular piece of sheet metal.

Two symmetric folds 2 feet apart will be made parallel to the longest side of the rectangle so that the trough has cross-sections in the shape of a trapezoid, as pictured in Figure 3.4.4. At what angle should the folds be made to produce the trough of maximum volume?



Figure 3.4.4: A cross-section of the trough formed by folding to an angle of θ .

Summary

• While there is no single algorithm that works in every situation where optimization is used, in most of the problems we consider, the following steps are helpful: draw a picture and introduce variables; identify the quantity to be optimized and find relationships among the variables; determine a function of a single variable that models the quantity to be optimized; decide the domain on which to consider the function being optimized; use calculus to identify the absolute maximum and/or minimum of the quantity being optimized.

Exercises

eBWork	1. An open box is to be made out of a 10-inch by 14-inch piece of cardboard by cutting out squares of equal size from the four corners and bending up the sides. Find the dimensions of the resulting box that has the largest volume.
	Dimensions of the bottom of the box: x
	Height of the box:
eBWork	2. A rectangular storage container with an open top is to have a volume of 22 cubic meters. The length of its base is twice the width. Material for the base costs 14 dollars per square meter. Material for the sides costs 8 dollars per square meter. Find the cost of materials for the cheapest such container.
	Total cost = (Round to the nearest penny and include monetary units. For example, if your answer is 1.095, enter \$1.10 including the
	dollar sign and second decimal place.)

with fencing parallel to one side of the rectangle (see the figure below). There are 520 feet of fencing available to complete the job. What is the largest possible total area of the four

pens?

Largest area =		(include units)
4. The top and bottom margins of area of printed material on the profit of the poster with the smallest a	poster is fixed at 386 square ce	
	printed material	
Width =		(include units)
Height =		(include units)
5. A rectangle is inscribed with $y = 12 - x^2$. What are the dimer Width = Height =		
6. A rectangular box with a squ The material for the side costs \$ \$3.00 per square foot. If you are it can contain? Justify your answ	\$1.50 per square foot and the willing to spend \$15 on the	e material for the bottom costs box, what is the largest volume
7. A farmer wants to start raisin angular pasture for the animals graze together. In order to minito enclose a large rectangular ar	to graze in. However, no two imize the amount of fencing	o different kinds of animals can

8. Two vertical poles of heights 60 ft and 80 ft stand on level ground, with their bases 100 ft apart. A cable that is stretched from the top of one pole to some point on the ground between the poles, and then to the top of the other pole. What is the minimum possible length of cable required? Justify your answer completely using calculus.

three segments of fence inside the large rectangle that are parallel to two existing sides. She has decided to purchase 7500 ft of fencing. What is the maximum possible area that each of

the four pens will enclose?

9. A company is designing propane tanks that are cylindrical with hemispherical ends. Assume that the company wants tanks that will hold 1000 cubic feet of gas, and that the ends are more expensive to make, costing \$5 per square foot, while the cylindrical barrel between the ends costs \$2 per square foot. Use calculus to determine the minimum cost to construct such a tank.

The Definite Integral

4.1 Determining distance traveled from velocity

Motivating Questions

- If we know the velocity of a moving body at every point in a given interval, can we determine the distance the object has traveled on the time interval?
- How is the problem of finding distance traveled related to finding the area under a certain curve?
- What does it mean to antidifferentiate a function and why is this process relevant to finding distance traveled?
- If velocity is negative, how does this impact the problem of finding distance traveled?

In the very first section of the text, we considered a situation where a moving object had a known position at time t. In particular, we stipulated that a tennis ball tossed into the air had its height s (in feet) at time t (in seconds) given by $s(t) = 64 - 16(t - 1)^2$. From this starting point, we investigated the average velocity of the ball on a given interval [a, b], computed by the difference quotient $\frac{s(b)-s(a)}{b-a}$, and eventually found that we could determine the exact instantaneous velocity of the ball at time t by taking the derivative of the position function,

$$s'(t) = \lim_{h \to 0} \frac{s(t+h) - s(t)}{h}.$$

Thus, given a differentiable position function, we are able to know the exact velocity of the moving object at any point in time.

Moreover, from this foundational problem involving position and velocity we have learned a great deal. Given a differentiable function f, we are now able to find its derivative and use this derivative to determine the function's instantaneous rate of change at any point in the domain, as well as to find where the function is increasing or decreasing, is concave up or concave down, and has relative extremes. The vast majority of the problems and applications we have considered have involved the situation where a particular function is known and we



seek information that relies on knowing the function's instantaneous rate of change. That is, we have typically proceeded from a function f to its derivative, f', and then used the meaning of the derivative to help us answer important questions.

In a much smaller number of situations so far, we have encountered the reverse situation where we instead know the derivative, f', and have tried to deduce information about f. It is this particular problem that will be the focus of our attention in most of Chapter 4: if we know the instantaneous rate of change of a function, are we able to determine the function itself? To begin, we start with a more focused question: if we know the instantaneous velocity of an object moving along a straight line path, can we determine its corresponding position function?

Preview Activity 4.1.1. Suppose that a person is taking a walk along a long straight path and walks at a constant rate of 3 miles per hour.

a. On the left-hand axes provided in Figure 4.1.1, sketch a labeled graph of the velocity function v(t) = 3.

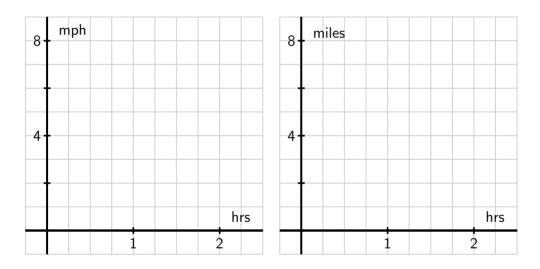


Figure 4.1.1: At left, axes for plotting y = v(t); at right, for plotting y = s(t).

Note that while the scale on the two sets of axes is the same, the units on the right-hand axes differ from those on the left. The right-hand axes will be used in question (d).

- b. How far did the person travel during the two hours? How is this distance related to the area of a certain region under the graph of y = v(t)?
- c. Find an algebraic formula, s(t), for the position of the person at time t, assuming that s(0) = 0. Explain your thinking.

- d. On the right-hand axes provided in Figure 4.1.1, sketch a labeled graph of the position function y = s(t).
- e. For what values of t is the position function s increasing? Explain why this is the case using relevant information about the velocity function v.

4.1.1 Area under the graph of the velocity function

In Preview Activity 4.1.1, we encountered a fundamental fact: when a moving object's velocity is constant (and positive), the area under the velocity curve over a given interval tells us the distance the object traveled.

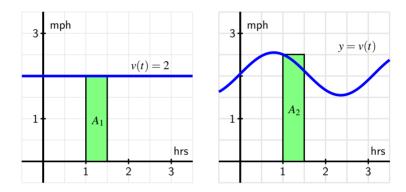


Figure 4.1.2: At left, a constant velocity function, at right, a non-constant velocity function.

As seen at left in Figure 4.1.2, if we consider an object moving at 2 miles per hour over the time interval [1, 1.5], then the area A_1 of the shaded region under y = v(t) on [1, 1.5] is

$$A_1 = 2 \frac{\text{miles}}{\text{hour}} \cdot \frac{1}{2} \text{ hours} = 1 \text{ mile.}$$

This principle holds in general simply due to the fact that distance equals rate times time, provided the rate is constant. Thus, if v(t) is constant on the interval [a, b], then the distance traveled on [a, b] is the area A that is given by

$$A = v(a)(b - a) = v(a)\Delta t$$

where Δt is the change in t over the interval. Note, too, that we could use any value of v(t) on the interval [a,b], since the velocity is constant; we simply chose v(a), the value at the interval's left endpoint. For several examples where the velocity function is piecewise constant, see http://gvsu.edu/s/9T.1

The situation is obviously more complicated when the velocity function is not constant. At the same time, on relatively small intervals on which v(t) does not vary much, the area

¹Marc Renault, calculus applets.

principle allows us to estimate the distance the moving object travels on that time interval. For instance, for the non-constant velocity function shown at right in Figure 4.1.2, we see that on the interval [1,1.5], velocity varies from v(1) = 2.5 down to $v(1.5) \approx 2.1$. Hence, one estimate for distance traveled is the area of the pictured rectangle,

$$A_2 = v(1)\Delta t = 2.5 \frac{\text{miles}}{\text{hour}} \cdot \frac{1}{2} \text{ hours} = 1.25 \text{ miles}.$$

Because v is decreasing on [1, 1.5] and the rectangle lies above the curve, clearly $A_2 = 1.25$ is an over-estimate of the actual distance traveled.

If we want to estimate the area under the non-constant velocity function on a wider interval, say [0,3], it becomes apparent that one rectangle probably will not give a good approximation. Instead, we could use the six rectangles pictured in Figure 4.1.3, find the area of each rectangle, and add up the total. Obviously there are choices to make and issues to understand: how many rectangles should we use? where should we evaluate the function to decide the rectangle's height? what happens if velocity is sometimes negative? can we attain the exact area under any non-constant curve?

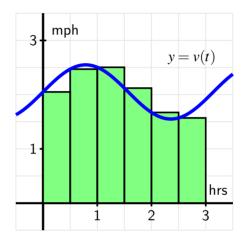


Figure 4.1.3: Using six rectangles to estimate the area under y = v(t) on [0,3].

These questions and more are ones we will study in what follows; for now it suffices to realize that the simple idea of the area of a rectangle gives us a powerful tool for estimating both distance traveled from a velocity function as well as the area under an arbitrary curve. To explore the setting of multiple rectangles to approximate area under a non-constant velocity function, see the applet found at http://gvsu.edu/s/9U. 2

Activity 4.1.2. Suppose that a person is walking in such a way that her velocity varies slightly according to the information given in Table 4.1.4 and graph given in Figure 4.1.5.

t	0.00	0.25	0.50	0.75	1.00	1.25	1.50	1.75	2.00
v(t)	1.500	1.789	1.938	1.992	2.000	2.008	2.063	2.211	2.500

Table 4.1.4: Velocity data for the person walking.

²Marc Renault, calculus applets.

- a. Using the grid, graph, and given data appropriately, estimate the distance traveled by the walker during the two hour interval from t=0 to t=2. You should use time intervals of width $\Delta t=0.5$, choosing a way to use the function consistently to determine the height of each rectangle in order to approximate distance traveled.
- How could you get a better approximation of the distance traveled on [0,2]? Explain, and then find this new estimate.
- c. Now suppose that you know that v is given by $v(t) = 0.5t^3 1.5t^2 + 1.5t + 1.5$. Remember that v is the derivative of the walker's position function, s. Find a formula for s so that s' = v.
- d. Based on your work in (c), what is the value of s(2) s(0)? What is the meaning of this quantity?

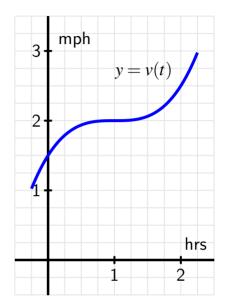


Figure 4.1.5: The graph of y = v(t).

4.1.2 Two approaches: area and antidifferentiation

When the velocity of a moving object is positive, the object's position is always increasing. While we will soon consider situations where velocity is negative and think about the ramifications of this condition on distance traveled, for now we continue to assume that we are working with a positive velocity function. In that setting, we have established that whenever v is actually constant, the exact distance traveled on an interval is the area under the velocity curve; furthermore, we have observed that when v is not constant, we can estimate the total distance traveled by finding the areas of rectangles that help to approximate the area under the velocity curve on the given interval. Hence, we see the importance of the problem of finding the area between a curve and the horizontal axis: besides being an interesting geometric question, in the setting of the curve being the (positive) velocity of a moving object, the area under the curve over an interval tells us the exact distance traveled on the interval. We can estimate this area any time we have a graph of the velocity function or a table of data that tells us some relevant values of the function.

In Activity 4.1.2, we also encountered an alternate approach to finding the distance traveled. In particular, if we know a formula for the instantaneous velocity, y = v(t), of the moving body at time t, then we realize that v must be the derivative of some corresponding position function s. If we can find a formula for s from the formula for v, it follows that we know the position of the object at time t. In addition, under the assumption that velocity is positive, the change in position over a given interval then tells us the distance traveled on that interval.

For a simple example, consider the situation from Preview Activity 4.1.1, where a person is walking along a straight line and has velocity function v(t) = 3 mph.

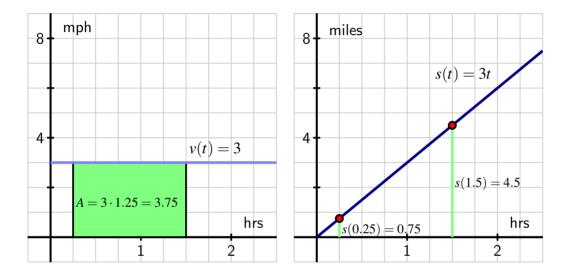


Figure 4.1.6: The velocity function v(t) = 3 and corresponding position function s(t) = 3t.

As pictured in Figure 4.1.6, we see the already noted relationship between area and distance traveled on the left-hand graph of the velocity function. In addition, because the velocity is constant at 3, we know that if s(t) = 3t, then s'(t) = 3, so s(t) = 3t is a function whose derivative is v(t). Furthermore, we now observe that s(1.5) = 4.5 and s(0.25) = 0.75, which are the respective locations of the person at times t = 0.25 and t = 1.5, and therefore

$$s(1.5) - s(0.25) = 4.5 - 0.75 = 3.75$$
 miles.

This is not only the change in position on [0.25, 1.5], but also precisely the distance traveled on [0.25, 1.5], which can also be computed by finding the area under the velocity curve over the same interval. There are profound ideas and connections present in this example that we will spend much of the remainder of Chapter 4 studying and exploring.

For now, it is most important to observe that if we are given a formula for a velocity function v, it can be very helpful to find a function s that satisfies s' = v. In this context, we say that

³Here we are making the implicit assumption that s(0) = 0; we will further discuss the different possibilities for values of s(0) in subsequent study.

s is an *antiderivative* of v. More generally, just as we say that f' is the derivative of f for a given function f, if we are given a function g and G is a function such that G' = g, we say that G is an *antiderivative* of g. For example, if $g(x) = 3x^2 + 2x$, an antiderivative of g is $G(x) = x^3 + x^2$, since G'(x) = g(x). Note that we say "an" antiderivative of g rather than "the" antiderivative of g because $H(x) = x^3 + x^2 + 5$ is also a function whose derivative is g, and thus H is another antiderivative of g.

Activity 4.1.3. A ball is tossed vertically in such a way that its velocity function is given by v(t) = 32 - 32t, where t is measured in seconds and v in feet per second. Assume that this function is valid for $0 \le t \le 2$.

- a. For what values of *t* is the velocity of the ball positive? What does this tell you about the motion of the ball on this interval of time values?
- b. Find an antiderivative, s, of v that satisfies s(0) = 0.
- c. Compute the value of $s(1) s(\frac{1}{2})$. What is the meaning of the value you find?
- d. Using the graph of y = v(t) provided in Figure 4.1.7, find the exact area of the region under the velocity curve between $t = \frac{1}{2}$ and t = 1. What is the meaning of the value you find?
- e. Answer the same questions as in (c) and (d) but instead using the interval [0, 1].
- f. What is the value of s(2) s(0)? What does this result tell you about the flight of the ball? How is this value connected to the provided graph of y = v(t)? Explain.

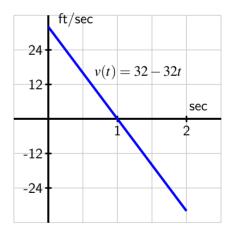


Figure 4.1.7: The graph of y = v(t).

4.1.3 When velocity is negative

Most of our work in this section has occurred under the assumption that velocity is positive. This hypothesis guarantees that the movement of the object under consideration is always in

a single direction, and hence ensures that the moving body's change in position is the same as the distance it travels on a given interval. As we saw in Activity 4.1.3, there are natural settings in which a moving object's velocity is negative; we would like to understand this scenario fully as well.

Consider a simple example where a person goes for a walk on a beach along a stretch of very straight shoreline that runs east-west. We can naturally assume that their initial position is s(0) = 0, and further stipulate that their position function increases as they move east from their starting location. For instance, a position of s = 1 mile represents being one mile east of the start location, while s = -1 tells us the person is one mile west of where they began walking on the beach. Now suppose the person walks in the following manner. From the outset at t = 0, the person walks due east at a constant rate of 3 mph for 1.5 hours. After 1.5 hours, the person stops abruptly and begins walking due west at the constant rate of 4 mph and does so for 0.5 hours. Then, after another abrupt stop and start, the person resumes walking at a constant rate of 3 mph to the east for one more hour. What is the total distance the person traveled on the time interval t = 0 to t = 3? What is the person's total change in position over that time?

On one hand, these are elementary questions to answer because the velocity involved is constant on each interval. From t = 0 to t = 1.5, the person traveled

$$D_{[0.1.5]} = 3$$
 miles per hour $\cdot 1.5$ hours = 4.5 miles.

Similarly, on t = 1.5 to t = 2, having a different rate, the distance traveled is

$$D_{[1.5.2]} = 4$$
 miles per hour \cdot 0.5 hours = 2 miles.

Finally, similar calculations reveal that in the final hour, the person walked

$$D_{[2,3]} = 3$$
 miles per hour · 1 hours = 3 miles,

so the total distance traveled is

$$D = D_{[0,1,5]} + D_{[1,5,2]} + D_{[2,3]} = 4.5 + 2 + 3 = 9.5$$
 miles.

Since the velocity on 1.5 < t < 2 is actually v = -4, being negative to indicate motion in the westward direction, this tells us that the person first walked 4.5 miles east, then 2 miles west, followed by 3 more miles east. Thus, the walker's total change in position is

change in position =
$$4.5 - 2 + 3 = 5.5$$
 miles.

While we have been able to answer these questions fairly easily, it is also important to think about this problem graphically in order that we can generalize our solution to the more complicated setting when velocity is not constant, as well as to note the particular impact that negative velocity has.

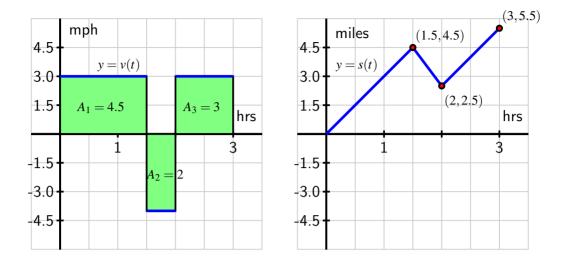


Figure 4.1.8: At left, the velocity function of the person walking; at right, the corresponding position function.

In Figure 4.1.8, we see how the distances we computed above can be viewed as areas: $A_1 = 4.5$ comes from taking rate times time (3·1.5), as do A_2 and A_3 for the second and third rectangles. The big new issue is that while A_2 is an area (and is therefore positive), because this area involves an interval on which the velocity function is negative, its area has a negative sign associated with it. This helps us to distinguish between distance traveled and change in position.

The distance traveled is the sum of the areas,

$$D = A_1 + A_2 + A_3 = 4.5 + 2 + 3 = 9.5$$
 miles.

But the change in position has to account for the sign associated with the area, where those above the t-axis are considered positive while those below the t-axis are viewed as negative, so that

$$s(3) - s(0) = (+4.5) + (-2) + (+3) = 5.5$$
 miles,

assigning the "-2" to the area in the interval [1.5, 2] because there velocity is negative and the person is walking in the "negative" direction. In other words, the person walks 4.5 miles in the positive direction, followed by two miles in the negative direction, and then 3 more miles in the positive direction. This affect of velocity being negative is also seen in the graph of the function y = s(t), which has a negative slope (specifically, its slope is -4) on the interval 1.5 < t < 2 since the velocity is -4 on that interval, which shows the person's position function is decreasing due to the fact that she is walking east, rather than west. On the intervals where she is walking west, the velocity function is positive and the slope of the position function s is therefore also positive.

To summarize, we see that if velocity is sometimes negative, this makes the moving object's change in position different from its distance traveled. By viewing the intervals on

which velocity is positive and negative separately, we may compute the distance traveled on each such interval, and then depending on whether we desire total distance traveled or total change in position, we may account for negative velocities that account for negative change in position, while still contributing positively to total distance traveled. We close this section with one additional activity that further explores the effects of negative velocity on the problem of finding change in position and total distance traveled.

Activity 4.1.4. Suppose that an object moving along a straight line path has its velocity v (in meters per second) at time t (in seconds) given by the piecewise linear function whose graph is pictured at left in Figure 4.1.9. We view movement to the right as being in the positive direction (with positive velocity), while movement to the left is in the negative direction.

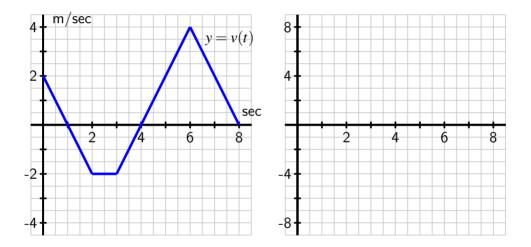


Figure 4.1.9: The velocity function of a moving object.

Suppose further that the object's initial position at time t = 0 is s(0) = 1.

- a. Determine the total distance traveled and the total change in position on the time interval $0 \le t \le 2$. What is the object's position at t = 2?
- b. On what time intervals is the moving object's position function increasing? Why? On what intervals is the object's position decreasing? Why?
- c. What is the object's position at t = 8? How many total meters has it traveled to get to this point (including distance in both directions)? Is this different from the object's total change in position on t = 0 to t = 8?
- d. Find the exact position of the object at t = 1, 2, 3, ..., 8 and use this data to sketch an accurate graph of y = s(t) on the axes provided at right in Figure 4.1.9. How

can you use the provided information about y = v(t) to determine the concavity of s on each relevant interval?

Summary

- If we know the velocity of a moving body at every point in a given interval and the velocity is positive throughout, we can estimate the object's distance traveled and in some circumstances determine this value exactly.
- In particular, when velocity is positive on an interval, we can find the total distance traveled by finding the area under the velocity curve and above the *t*-axis on the given time interval. We may only be able to estimate this area, depending on the shape of the velocity curve.
- An antiderivative of a function f is a new function F whose derivative is f. That is, F is an antiderivative of f provided that F' = f. In the context of velocity and position, if we know a velocity function v, an antiderivative of v is a position function s that satisfies s' = v. If v is positive on a given interval, say [a, b], then the change in position, s(b) s(a), measures the distance the moving object traveled on [a, b].
- In the setting where velocity is sometimes negative, this means that the object is sometimes traveling in the opposite direction (depending on whether velocity is positive or negative), and thus involves the object backtracking. To determine distance traveled, we have to think about the problem separately on intervals where velocity is positive and negative and account for the change in position on each such interval.

Exercises

1. A car comes to a stop six seconds after the driver applies the brakes. While the brakes are on, the following velocities are recorded:



Time since brakes applied (sec)	0	2	4	6
Velocity (ft/s)	88	45	16	0

Give lower and upper estimates (using all of the available data) for the distance the car traveled after the brakes were applied.

lower:		
upper:		
(for oacl	h include unite)	

On a sketch of velocity against time, show the lower and upper estimates you found above..

2. The velocity of a car is f(t) = 9t meters/second. Use a graph of f(t) to find the exact distance traveled by the car, in meters, from t = 0 to t = 10 seconds.





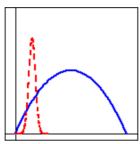


3. The velocity of a particle moving along the *x*-axis is given by f(t) = 12 - 4t cm/sec. Use a graph of f(t) to find the exact change in position of the particle from time t = 0 to t = 4 seconds.

change in position = (include units)



4. Two cars start at the same time and travel in the same direction along a straight road. The figure below gives the velocity, v (in km/hr), of each car as a function of time (in hr).



The velocity of car A is given by the solid, blue curve, and the velocity of car B by dashed, red curve.

(a)

Which car attains the larger maximum velocity?

(b)

Which stops first?

(c)

Which travels farther?



5. Suppose that an accelerating car goes from 0 mph to 66.8 mph in five seconds. Its velocity is given in the following table, converted from miles per hour to feet per second, so that all time measurements are in seconds. (Note: 1 mph is 22/15 feet per sec = 22/15 ft/s.) Find the average acceleration of the car over each of the first two seconds.

t	0	1	2	3	4	5
v(t)	0.00	33.41	57.91	75.73	89.09	98.00

 $average\ acceleration\ over\ the\ first\ second =$

(include units)

average aceleration over the second second = (include units)



6. The velocity function is $v(t) = t^2 - 6t + 8$ for a particle moving along a line. Find the displacement (net distance covered) of the particle during the time interval [-2,5].

displacement =

7. Along the eastern shore of Lake Michigan from Lake Macatawa (near Holland) to Grand Haven, there is a bike bath that runs almost directly north-south. For the purposes of this problem, assume the road is completely straight, and that the function s(t) tracks the position of the biker along this path in miles north of Pigeon Lake, which lies roughly halfway between the ends of the bike path.

Suppose that the biker's velocity function is given by the graph in Figure 4.1.10 on the time interval $0 \le t \le 4$ (where t is measured in hours), and that s(0) = 1.

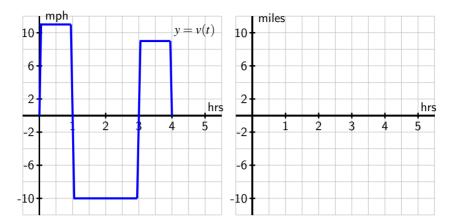


Figure 4.1.10: The graph of the biker's velocity, y = v(t), at left. At right, axes to plot an approximate sketch of y = s(t).

- a. Approximately how far north of Pigeon Lake was the cyclist when she was the greatest distance away from Pigeon Lake? At what time did this occur?
- b. What is the cyclist's total change in position on the time interval $0 \le t \le 2$? At t = 2, was she north or south of Pigeon Lake?
- c. What is the total distance the biker traveled on $0 \le t \le 4$? At the end of the ride, how close was she to the point at which she started?
- d. Sketch an approximate graph of y = s(t), the position function of the cyclist, on the interval $0 \le t \le 4$. Label at least four important points on the graph of s.
- **8.** A toy rocket is launched vertically from the ground on a day with no wind. The rocket's vertical velocity at time t (in seconds) is given by v(t) = 500 32t feet/sec.
 - a. At what time after the rocket is launched does the rocket's velocity equal zero? Call this time value a. What happens to the rocket at t = a?
 - b. Find the value of the total area enclosed by y = v(t) and the t-axis on the interval $0 \le t \le a$. What does this area represent in terms of the physical setting of the problem?
 - c. Find an antiderivative s of the function v. That is, find a function s such that s'(t) = v(t).
 - d. Compute the value of s(a) s(0). What does this number represent in terms of the physical setting of the problem?
 - e. Compute s(5) s(1). What does this number tell you about the rocket's flight?
- **9.** An object moving along a horizontal axis has its instantaneous velocity at time t in seconds given by the function v pictured in Figure 4.1.11, where v is measured in feet/sec.

Assume that the curves that make up the parts of the graph of y = v(t) are either portions of straight lines or portions of circles.

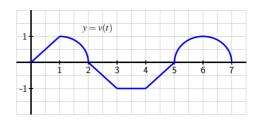


Figure 4.1.11: The graph of y = v(t), the velocity function of a moving object.

- a. Determine the exact total distance the object traveled on $0 \le t \le 2$.
- b. What is the value and meaning of s(5) s(2), where y = s(t) is the position function of the moving object?
- c. On which time interval did the object travel the greatest distance: [0, 2], [2, 4], or [5, 7]?
- d. On which time interval(s) is the position function *s* increasing? At which point(s) does *s* achieve a relative maximum?

10. Filters at a water treatment plant become dirtier over time and thus become less effective; they are replaced every 30 days. During one 30-day period, the rate at which pollution passes through the filters into a nearby lake (in units of particulate matter per day) is measured every 6 days and is given in the following table. The time t is measured in days since the filters were replaced.

Day, t	0	6	12	18	24	30
Rate of pollution in units per day, $p(t)$	7	8	10	13	18	35

Table 4.1.12: Pollution data for the water filters.

- a. Plot the given data on a set of axes with time on the horizontal axis and the rate of pollution on the vertical axis.
- b. Explain why the amount of pollution that entered the lake during this 30-day period would be given exactly by the area bounded by y = p(t) and the t-axis on the time interval [0,30].
- c. Estimate the total amount of pollution entering the lake during this 30-day period. Carefully explain how you determined your estimate.

4.2 Riemann Sums

Motivating Questions

- How can we use a Riemann sum to estimate the area between a given curve and the horizontal axis over a particular interval?
- What are the differences among left, right, middle, and random Riemann sums?
- How can we write Riemann sums in an abbreviated form?

In Section 4.1, we learned that if we have a moving object with velocity function v, whenever v(t) is positive, the area between y=v(t) and the t-axis over a given time interval tells us the distance traveled by the object over that time period; in addition, if v(t) is sometimes negative and we view the area of any region below the t-axis as having an associated negative sign, then the sum of these signed areas over a given interval tells us the moving object's change in position over the time interval.

For instance, for the velocity function given in Figure 4.2.1, if the areas of shaded regions are A_1 , A_2 , and A_3 as labeled, then the total distance D traveled by the moving object on [a, b] is

$$D = A_1 + A_2 + A_3$$

while the total change in the object's position on [a, b] is

$$s(b) - s(a) = A_1 - A_2 + A_3$$
.

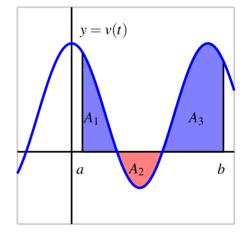


Figure 4.2.1: A velocity function that is sometimes negative.

Because the motion is in the negative direction on the interval where v(t) < 0, we subtract A_2 when determining the object's total change in position.

Of course, finding D and s(b) - s(a) for the situation given in Figure 4.2.1 presumes that we can actually find the areas represented by A_1 , A_2 , and A_3 . In most of our work in Section 4.1, such as in Activities 4.1.3 and Activity 4.1.4, we worked with velocity functions that were either constant or linear, so that by finding the areas of rectangles and triangles, we could find the area bounded by the velocity function and the horizontal axis exactly. But when the

curve that bounds a region is not one for which we have a known formula for area, we are unable to find this area exactly. Indeed, this is one of our biggest goals in Chapter 4: to learn how to find the exact area bounded between a curve and the horizontal axis for as many different types of functions as possible.

To begin, we expand on the ideas in Activity 4.1.2, where we encountered a nonlinear velocity function and approximated the area under the curve using four and eight rectangles, respectively. In the following preview activity, we focus on three different options for deciding how to find the heights of the rectangles we will use.

Preview Activity 4.2.1. A person walking along a straight path has her velocity in miles per hour at time t given by the function $v(t) = 0.25t^3 - 1.5t^2 + 3t + 0.25$, for times in the interval $0 \le t \le 2$. The graph of this function is also given in each of the three diagrams in Figure 4.2.2.

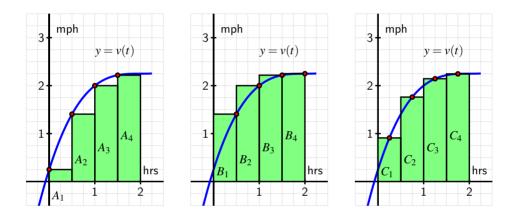


Figure 4.2.2: Three approaches to estimating the area under y = v(t) on the interval [0,2].

Note that in each diagram, we use four rectangles to estimate the area under y = v(t) on the interval [0,2], but the method by which the four rectangles' respective heights are decided varies among the three individual graphs.

a. How are the heights of rectangles in the left-most diagram being chosen? Explain, and hence determine the value of

$$S = A_1 + A_2 + A_3 + A_4$$

by evaluating the function y = v(t) at appropriately chosen values and observing the width of each rectangle. Note, for example, that

$$A_3 = v(1) \cdot \frac{1}{2} = 2 \cdot \frac{1}{2} = 1.$$

b. Explain how the heights of rectangles are being chosen in the middle diagram and find the value of

$$T = B_1 + B_2 + B_3 + B_4.$$

c. Likewise, determine the pattern of how heights of rectangles are chosen in the right-most diagram and determine

$$U = C_1 + C_2 + C_3 + C_4.$$

d. Of the estimates S, T, and U, which do you think is the best approximation of D, the total distance the person traveled on [0,2]? Why?

4.2.1 Sigma Notation

It is apparent from several different problems we have considered that sums of areas of rectangles is one of the main ways to approximate the area under a curve over a given interval. Intuitively, we expect that using a larger number of thinner rectangles will provide a way to improve the estimates we are computing. As such, we anticipate dealing with sums with a large number of terms. To do so, we introduce the use of so-called *sigma notation*, named for the Greek letter Σ , which is the capital letter S in the Greek alphabet.

For example, say we are interested in the sum

$$1 + 2 + 3 + \cdots + 100$$
.

which is the sum of the first 100 natural numbers. Sigma notation provides a shorthand notation that recognizes the general pattern in the terms of the sum. It is equivalent to write

$$\sum_{k=1}^{100} k = 1 + 2 + 3 + \dots + 100.$$

We read the symbol $\sum_{k=1}^{100} k$ as "the sum from k equals 1 to 100 of k." The variable k is usually called the index of summation, and the letter that is used for this variable is immaterial. Each sum in sigma notation involves a function of the index; for example,

$$\sum_{k=1}^{10} (k^2 + 2k) = (1^2 + 2 \cdot 1) + (2^2 + 2 \cdot 2) + (3^2 + 2 \cdot 3) + \dots + (10^2 + 2 \cdot 10),$$

and more generally,

$$\sum_{k=1}^{n} f(k) = f(1) + f(2) + \dots + f(n).$$

Sigma notation allows us the flexibility to easily vary the function being used to track the pattern in the sum, as well as to adjust the number of terms in the sum simply by changing the value of n. We test our understanding of this new notation in the following activity.

Activity 4.2.2. For each sum written in sigma notation, write the sum long-hand and evaluate the sum to find its value. For each sum written in expanded form, write the sum in sigma notation.

a.
$$\sum_{k=1}^{5} (k^2 + 2)$$

b. $\sum_{i=3}^{6} (2i - 1)$
c. $3 + 7 + 11 + 15 + \dots + 27$
d. $4 + 8 + 16 + 32 + \dots + 256$
e. $\sum_{i=1}^{6} \frac{1}{2^i}$

4.2.2 Riemann Sums

When a moving body has a positive velocity function y = v(t) on a given interval [a, b], we know that the area under the curve over the interval is the total distance the body travels on [a, b]. While this is the fundamental motivating force behind our interest in the area bounded by a function, we are also interested more generally in being able to find the exact area bounded by y = f(x) on an interval [a, b], regardless of the meaning or context of the function f. For now, we continue to focus on determining an accurate estimate of this area through the use of a sum of the areas of rectangles, doing so in the setting where $f(x) \ge 0$ on [a, b]. Throughout, unless otherwise indicated, we also assume that f is continuous on [a, b].

The first choice we make in any such approximation is the number of rectangles.

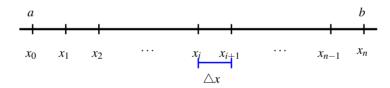


Figure 4.2.3: Subdividing the interval [a, b] into n subintervals of equal length Δx .

If we say that the total number of rectangles is n, and we desire n rectangles of equal width to subdivide the interval [a, b], then each rectangle must have width $\Delta x = \frac{b-a}{n}$. We observe further that $x_1 = x_0 + \Delta x$, $x_2 = x_0 + 2\Delta x$, and thus in general $x_i = a + i\Delta x$, as pictured in Figure 4.2.3.

We use each subinterval $[x_i, x_{i+1}]$ as the base of a rectangle, and next must choose how to decide the height of the rectangle that will be used to approximate the area under y = f(x) on the subinterval. There are three standard choices: use the left endpoint of each subinterval, the right endpoint of each subinterval, or the midpoint of each. These are precisely the options encountered in Preview Activity 4.2.1 and seen in Figure 4.2.2. We next explore how these choices can be reflected in sigma notation.

If we now consider an arbitrary positive function f on [a, b] with the interval subdivided as shown in Figure 4.2.3, and choose to use left endpoints, then on each interval of the form

 $[x_i, x_{i+1}]$, the area of the rectangle formed is given by

$$A_{i+1} = f(x_i) \cdot \Delta x$$
,

as seen in Figure 4.2.4.

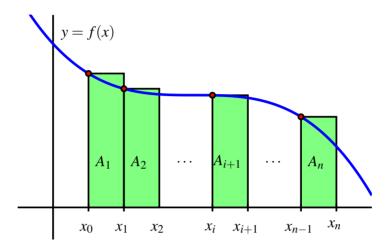


Figure 4.2.4: Subdividing the interval [a, b] into n subintervals of equal length Δx and approximating the area under y = f(x) over [a, b] using left rectangles.

If we let L_n denote the sum of the areas of rectangles whose heights are given by the function value at each respective left endpoint, then we see that

$$L_n = A_1 + A_2 + \dots + A_{i+1} + \dots + A_n$$

= $f(x_0) \cdot \Delta x + f(x_1) \cdot \Delta x + \dots + f(x_i) \cdot \Delta x + \dots + f(x_{n-1}) \cdot \Delta x$.

In the more compact sigma notation, we have

$$L_n = \sum_{i=0}^{n-1} f(x_i) \Delta x.$$

Note particularly that since the index of summation begins at 0 and ends at n-1, there are indeed n terms in this sum. We call L_n the *left Riemann sum* for the function f on the interval [a, b].

There are now two fundamental issues to explore: the number of rectangles we choose to use and the selection of the pattern by which we identify the height of each rectangle. It is best to explore these choices dynamically, and the applet found at http://gvsu.edu/s/a9 is a particularly useful one. There we see the image shown in Figure 4.2.5, but with the opportunity to adjust the slider bars for the left endpoint and the number of subintervals.

¹Marc Renault, Geogebra Calculus Applets.

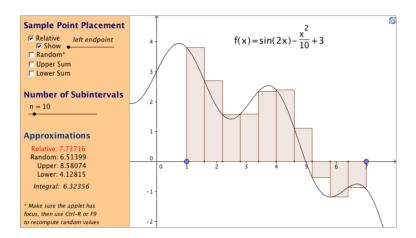


Figure 4.2.5: A snapshot of the applet found at http://gvsu.edu/s/a9.

By moving the sliders, we can see how the heights of the rectangles change as we consider left endpoints, midpoints, and right endpoints, as well as the impact that a larger number of narrower rectangles has on the approximation of the exact area bounded by the function and the horizontal axis.

To see how the Riemann sums for right endpoints and midpoints are constructed, we consider Figure 4.2.6.

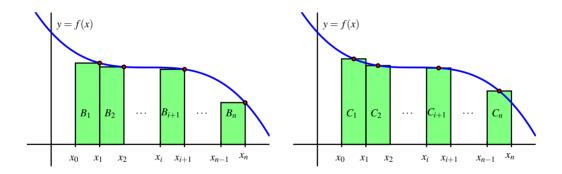


Figure 4.2.6: Riemann sums using right endpoints and midpoints.

For the sum with right endpoints, we see that the area of the rectangle on an arbitrary interval $[x_i, x_{i+1}]$ is given by $B_{i+1} = f(x_{i+1}) \cdot \Delta x$, so that the sum of all such areas of rectangles is given by

$$R_n = B_1 + B_2 + \dots + B_{i+1} + \dots + B_n$$

$$= f(x_1) \cdot \Delta x + f(x_2) \cdot \Delta x + \dots + f(x_{i+1}) \cdot \Delta x + \dots + f(x_n) \cdot \Delta x$$

$$= \sum_{i=1}^n f(x_i) \Delta x.$$

We call R_n the *right Riemann sum* for the function f on the interval [a,b]. For the sum that uses midpoints, we introduce the notation

$$\overline{x}_{i+1} = \frac{x_i + x_{i+1}}{2}$$

so that \overline{x}_{i+1} is the midpoint of the interval $[x_i, x_{i+1}]$. For instance, for the rectangle with area C_1 in Figure 4.2.6, we now have

$$C_1 = f(\overline{x}_1) \cdot \Delta x$$
.

Hence, the sum of all the areas of rectangles that use midpoints is

$$M_n = C_1 + C_2 + \dots + C_{i+1} + \dots + C_n$$

= $f(\overline{x_1}) \cdot \Delta x + f(\overline{x_2}) \cdot \Delta x + \dots + f(\overline{x_{i+1}}) \cdot \Delta x + \dots + f(\overline{x_n}) \cdot \Delta x$
= $\sum_{i=1}^n f(\overline{x_i}) \Delta x$,

and we say that M_n is the *middle Riemann sum* for f on [a,b].

When $f(x) \ge 0$ on [a, b], each of the Riemann sums L_n , R_n , and M_n provides an estimate of the area under the curve y = f(x) over the interval [a, b]; momentarily, we will discuss the meaning of Riemann sums in the setting when f is sometimes negative. We also recall that in the context of a nonnegative velocity function y = v(t), the corresponding Riemann sums are approximating the distance traveled on [a, b] by the moving object with velocity function v.

There is a more general way to think of Riemann sums, and that is to not restrict the choice of where the function is evaluated to determine the respective rectangle heights. That is, rather than saying we'll always choose left endpoints, or always choose midpoints, we simply say that a point x_{i+1}^* will be selected at random in the interval $[x_i, x_{i+1}]$ (so that $x_i \le x_{i+1}^* \le x_{i+1}$), which makes the Riemann sum given by

$$f(x_1^*) \cdot \Delta x + f(x_2^*) \cdot \Delta x + \dots + f(x_{i+1}^*) \cdot \Delta x + \dots + f(x_n^*) \cdot \Delta x = \sum_{i=1}^n f(x_i^*) \Delta x.$$

At http://gvsu.edu/s/a9, the applet noted earlier and referenced in Figure 4.2.5, by unchecking the "relative" box at the top left, and instead checking "random," we can easily explore the effect of using random point locations in subintervals on a given Riemann sum. In computational practice, we most often use L_n , R_n , or M_n , while the random Riemann sum is useful in theoretical discussions. In the following activity, we investigate several different Riemann sums for a particular velocity function.

Activity 4.2.3. Suppose that an object moving along a straight line path has its velocity in feet per second at time t in seconds given by $v(t) = \frac{2}{9}(t-3)^2 + 2$.

a. Carefully sketch the region whose exact area will tell you the value of the distance the object traveled on the time interval $2 \le t \le 5$.

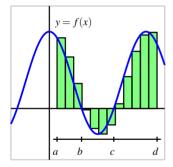
- b. Estimate the distance traveled on [2,5] by computing L_4 , R_4 , and M_4 .
- c. Does averaging L_4 and R_4 result in the same value as M_4 ? If not, what do you think the average of L_4 and R_4 measures?
- d. For this question, think about an arbitrary function f, rather than the particular function v given above. If f is positive and increasing on [a,b], will L_n overestimate or under-estimate the exact area under f on [a,b]? Will R_n over- or underestimate the exact area under f on [a,b]? Explain.

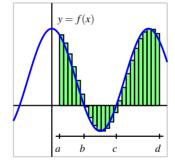
4.2.3 When the function is sometimes negative

For a Riemann sum such as

$$L_n = \sum_{i=0}^{n-1} f(x_i) \Delta x,$$

we can of course compute the sum even when f takes on negative values. We know that when f is positive on [a, b], the corresponding left Riemann sum L_n estimates the area bounded by f and the horizontal axis over the interval.





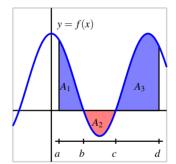


Figure 4.2.7: At left and center, two left Riemann sums for a function f that is sometimes negative; at right, the areas bounded by f on the interval [a, d].

For a function such as the one pictured in Figure 4.2.7, where in the first figure a left Riemann sum is being taken with 12 subintervals over [a,d], we observe that the function is negative on the interval $b \le x \le c$, and so for the four left endpoints that fall in [b,c], the terms $f(x_i)\Delta x$ have negative function values. This means that those four terms in the Riemann sum produce an estimate of the *opposite* of the area bounded by y = f(x) and the x-axis on [b,c].

In Figure 4.2.7, we also see evidence that by increasing the number of rectangles used in a Riemann sum, it appears that the approximation of the area (or the opposite of the area) bounded by a curve appears to improve. For instance, in the middle graph, we use 24 left rectangles, and from the shaded areas, it appears that we have decreased the error from the

approximation that uses 12. When we proceed to Section 4.3, we will discuss the natural idea of letting the number of rectangles in the sum increase without bound.

For now, it is most important for us to observe that, in general, any Riemann sum of a continuous function f on an interval [a, b] approximates the difference between the area that lies above the horizontal axis on [a, b] and under f and the area that lies below the horizontal axis on [a, b] and above f. In the notation of Figure 4.2.7, we may say that

$$L_{24} \approx A_1 - A_2 + A_3$$

where L_{24} is the left Riemann sum using 24 subintervals shown in the middle graph, and A_1 and A_3 are the areas of the regions where f is positive on the interval of interest, while A_2 is the area of the region where f is negative. We will also call the quantity $A_1 - A_2 + A_3$ the *net signed area* bounded by f over the interval [a,d], where by the phrase "signed area" we indicate that we are attaching a minus sign to the areas of regions that fall below the horizontal axis.

Finally, we recall from the introduction to this present section that in the context where the function f represents the velocity of a moving object, the total sum of the areas bounded by the curve tells us the total distance traveled over the relevant time interval, while the total net signed area bounded by the curve computes the object's change in position on the interval.

Activity 4.2.4. Suppose that an object moving along a straight line path has its velocity v (in feet per second) at time t (in seconds) given by

$$v(t) = \frac{1}{2}t^2 - 3t + \frac{7}{2}.$$

- a. Compute M_5 , the middle Riemann sum, for v on the time interval [1,5]. Be sure to clearly identify the value of Δt as well as the locations of t_0, t_1, \cdots, t_5 . In addition, provide a careful sketch of the function and the corresponding rectangles that are being used in the sum.
- b. Building on your work in (a), estimate the total change in position of the object on the interval [1, 5].
- c. Building on your work in (a) and (b), estimate the total distance traveled by the object on [1,5].
- d. Use appropriate computing technology^a to compute M_{10} and M_{20} . What exact value do you think the middle sum eventually approaches as n increases without bound? What does that number represent in the physical context of the overall problem?

^aFor instance, consider the applet at http://gvsu.edu/s/a9 and change the function and adjust the locations of the blue points that represent the interval endpoints a and b.

Summary

- A Riemann sum is simply a sum of products of the form $f(x_i^*)\Delta x$ that estimates the area between a positive function and the horizontal axis over a given interval. If the function is sometimes negative on the interval, the Riemann sum estimates the difference between the areas that lie above the horizontal axis and those that lie below the axis.
- The three most common types of Riemann sums are left, right, and middle sums, plus we can also work with a more general, random Riemann sum. The only difference among these sums is the location of the point at which the function is evaluated to determine the height of the rectangle whose area is being computed in the sum. For a left Riemann sum, we evaluate the function at the left endpoint of each subinterval, while for right and middle sums, we use right endpoints and midpoints, respectively.
- The left, right, and middle Riemann sums are denoted L_n , R_n , and M_n , with formulas

$$L_n = f(x_0)\Delta x + f(x_1)\Delta x + \dots + f(x_{n-1})\Delta x = \sum_{i=0}^{n-1} f(x_i)\Delta x,$$

$$R_n = f(x_1)\Delta x + f(x_2)\Delta x + \dots + f(x_n)\Delta x = \sum_{i=1}^n f(x_i)\Delta x,$$

$$M_n = f(\overline{x}_1)\Delta x + f(\overline{x}_2)\Delta x + \dots + f(\overline{x}_n)\Delta x = \sum_{i=1}^n f(\overline{x}_i)\Delta x,$$

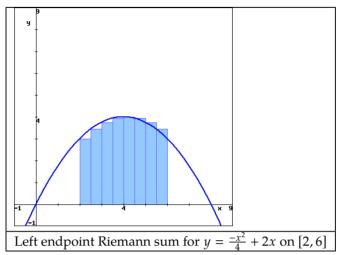
where $x_0 = a$, $x_i = a + i\Delta x$, and $x_n = b$, using $\Delta x = \frac{b-a}{n}$. For the midpoint sum, $\overline{x}_i = (x_{i-1} + x_i)/2$.

Exercises



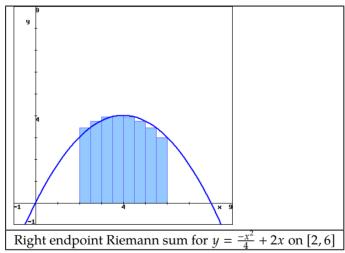
1. The rectangles in the graph below illustrate a left endpoint Riemann sum for $f(x) = \frac{-x^2}{4} + 2x$ on the interval [2, 6].

The value of this left endpoint Riemann sum is _______, and this Riemann sum is [Choose: [select an answer] | an overestimate of | equal to | an underestimate of | there is ambiguity] the area of the region enclosed by y = f(x), the x-axis, and the vertical lines x = 2 and x = 6.



The rectangles in the graph below illustrate a right endpoint Riemann sum for $f(x) = \frac{-x^2}{4} + 2x$ on the interval [2, 6].

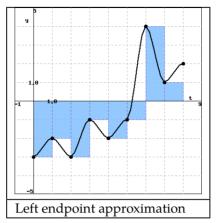
The value of this right endpoint Riemann sum is _______, and this Riemann sum is [Choose: [select an answer] | an overestimate of | equal to | an underestimate of | there is ambiguity] the area of the region enclosed by y = f(x), the x-axis, and the vertical lines x = 2 and x = 6.



2. Your task is to estimate how far an object traveled during the time interval $0 \le t \le 8$, but you only have the following data about the velocity of the object.

time (sec)	0	1	2	3	4	5	6	7	8
velocity (feet/sec)	-3	-2	-3	-1	-2	-1	4	1	2

To get an idea of what the velocity function might look like, you pick up a black pen, plot the data points, and connect them by curves. Your sketch looks something like the black curve in the graph below.

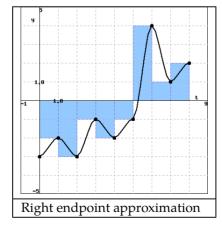


You decide to use a left endpoint Riemann sum to estimate the total displacement. So, you pick up a blue pen and draw rectangles whose height is determined by the velocity measurement at the left endpoint of each one-second interval. By using the left endpoint Riemann sum as an approximation, you are assuming that the actual velocity is approximately constant on each one-second interval (or, equivalently, that the actual acceleration is approximately zero on each one-second interval), and that the velocity and acceleration have discontinuous jumps every second. This assumption is probably incorrect because it is likely that the velocity and acceleration change continuously over time. However, you decide to use this approximation anyway since it seems like a reasonable approximation to the actual velocity given the limited amount of data.

(A) Using the left endpoint Riemann sum, find approximately how far the object traveled. Your answers must include the correct units.

Total displacement =		
Total distance traveled	I =	

Using the same data, you also decide to estimate how far the object traveled using a right endpoint Riemann sum. So, you sketch the curve again with a black pen, and draw rectangles whose height is determined by the velocity measurement at the right endpoint of each one-second interval.



(B) Using the right endpoint Riemann sum, find approximately how far the object traveled.

our answers must include the correct units.
otal displacement =
otal distance traveled =
• On a sketch of $y = e^x$, represent the left Riemann sum with $n = 2$ approximating $\int_2^3 e^x dx$.
Write out the terms of the sum, but do not evaluate it:
um = +
On another sketch, represent the right Riemann sum with $n=2$ approximating $\int_2^3 e^x dx$.
Vrite out the terms of the sum, but do not evaluate it:
um = +
Which sum is an overestimate?
Vhich sum is an underestimate?

- **4.** Consider the function f(x) = 3x + 4.
 - a. Compute M_4 for y = f(x) on the interval [2,5]. Be sure to clearly identify the value of Δx , as well as the locations of x_0, x_1, \ldots, x_4 . Include a careful sketch of the function and the corresponding rectangles being used in the sum.
 - b. Use a familiar geometric formula to determine the exact value of the area of the region bounded by y = f(x) and the x-axis on [2, 5].
 - c. Explain why the values you computed in (a) and (b) turn out to be the same. Will this be true if we use a number different than n = 4 and compute M_n ? Will L_4 or R_4 have the same value as the exact area of the region found in (b)?
 - d. Describe the collection of functions g for which it will always be the case that M_n , regardless of the value of n, gives the exact net signed area bounded between the function g and the x-axis on the interval [a,b].
- **5.** Let *S* be the sum given by

$$S = ((1.4)^2 + 1) \cdot 0.4 + ((1.8)^2 + 1) \cdot 0.4 + ((2.2)^2 + 1) \cdot 0.4 + ((2.6)^2 + 1) \cdot 0.4 + ((3.0)^2 + 1) \cdot 0.4.$$

- a. Assume that S is a right Riemann sum. For what function f and what interval [a, b] is S an approximation of the area under f and above the x-axis on [a, b]? Why?
- b. How does your answer to (a) change if *S* is a left Riemann sum? a middle Riemann sum?
- c. Suppose that *S* really is a right Riemann sum. What is geometric quantity does *S* approximate?
- d. Use sigma notation to write a new sum *R* that is the right Riemann sum for the same function, but that uses twice as many subintervals as *S*.
- **6.** A car traveling along a straight road is braking and its velocity is measured at several different points in time, as given in the following table.

seconds, t	0	0.3	0.6	0.9	1.2	1.5	1.8
Velocity in ft/sec, $v(t)$	100	88	74	59	40	19	0

Table 4.2.8: Data for the braking car.

- a. Plot the given data on a set of axes with time on the horizontal axis and the velocity on the vertical axis.
- b. Estimate the total distance traveled during the car the time brakes using a middle Riemann sum with 3 subintervals.
- c. Estimate the total distance traveled on [0, 1.8] by computing L_6 , R_6 , and $\frac{1}{2}(L_6 + R_6)$.
- d. Assuming that v(t) is always decreasing on [0, 1.8], what is the maximum possible distance the car traveled before it stopped? Why?
- 7. The rate at which pollution escapes a scrubbing process at a manufacturing plant increases over time as filters and other technologies become less effective. For this particular example, assume that the rate of pollution (in tons per week) is given by the function r that is pictured in Figure 4.2.9.
 - a. Use the graph to estimate the value of M_4 on the interval [0,4].
 - b. What is the meaning of M_4 in terms of the pollution discharged by the plant?
 - c. Suppose that $r(t) = 0.5e^{0.5t}$. Use this formula for r to compute L_5 on [0,4].
 - d. Determine an upper bound on the total amount of pollution that can escape the plant during the pictured four week time period that is accurate within an error of at most one ton of pollution.

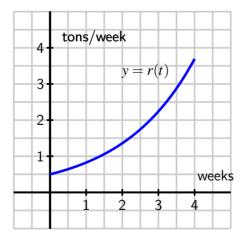


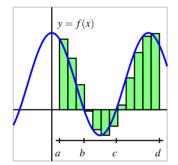
Figure 4.2.9: The rate, r(t), of pollution in tons per week.

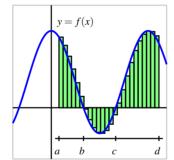
4.3 The Definite Integral

Motivating Questions

- How does increasing the number of subintervals affect the accuracy of the approximation generated by a Riemann sum?
- What is the definition of the definite integral of a function *f* over the interval [*a*, *b*]?
- What does the definite integral measure exactly, and what are some of the key properties of the definite integral?

In Figure 4.2.7, which is repeated below as Figure 4.3.1, we see visual evidence that increasing the number of rectangles in a Riemann sum improves the accuracy of the approximation of the net signed area that is bounded by the given function on the interval under consideration.





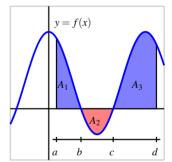


Figure 4.3.1: At left and center, two left Riemann sums for a function f that is sometimes negative; at right, the exact areas bounded by f on the interval [a, d].

We thus explore the natural idea of allowing the number of rectangles to increase without bound in an effort to compute the exact net signed area bounded by a function on an interval. In addition, it is important to think about the differences among left, right, and middle Riemann sums and the different results they generate as the value of n increases. As we have done throughout our investigations with area, we begin with functions that are exclusively positive on the interval under consideration.

Preview Activity 4.3.1. Consider the applet found at http://gvsu.edu/s/a 9^a . There, you will initially see the situation shown in Figure 4.3.2.

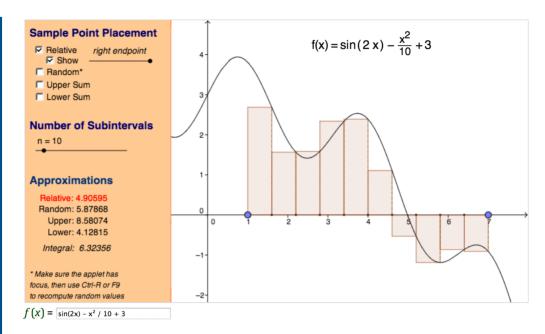


Figure 4.3.2: A right Riemann sum with 10 subintervals for the function $f(x) = \sin(2x) - \frac{x^2}{10} + 3$ on the interval [1, 7]. The value of the sum is $R_{10} = 4.90595$.

Note that the value of the chosen Riemann sum is displayed next to the word "relative," and that you can change the type of Riemann sum being computed by dragging the point on the slider bar below the phrase "sample point placement."

Explore to see how you can change the window in which the function is viewed, as well as the function itself. You can set the minimum and maximum values of x by clicking and dragging on the blue points that set the endpoints; you can change the function by typing a new formula in the "f(x)" window at the bottom; and you can adjust the overall window by "panning and zooming" by using the Shift key and the scrolling feature of your mouse. More information on how to pan and zoom can be found at http://gvsu.edu/s/Fl.

Work accordingly to adjust the applet so that it uses a left Riemann sum with n = 5 subintervals for the function is f(x) = 2x + 1. You should see the updated figure shown in Figure 4.3.3. Then, answer the following questions.

- a. Update the applet (and view window, as needed) so that the function being considered is f(x) = 2x + 1 on [1,4], as directed above. For this function on this interval, compute L_n , M_n , R_n for n = 5, n = 25, and n = 100. What appears to be the exact area bounded by f(x) = 2x + 1 and the x-axis on [1,4]?
- b. Use basic geometry to determine the exact area bounded by f(x) = 2x + 1 and the

- x-axis on [1, 4].
- c. Based on your work in (a) and (b), what do you observe occurs when we increase the number of subintervals used in the Riemann sum?
- d. Update the applet to consider the function $f(x) = x^2 + 1$ on the interval [1,4] (note that you need to enter "x ^ 2 + 1" for the function formula). Use the applet to compute L_n , M_n , R_n for n = 5, n = 25, and n = 100. What do you conjecture is the exact area bounded by $f(x) = x^2 + 1$ and the x-axis on [1,4]?
- e. Why can we not compute the exact value of the area bounded by $f(x) = x^2 + 1$ and the *x*-axis on [1, 4] using a formula like we did in (b)?

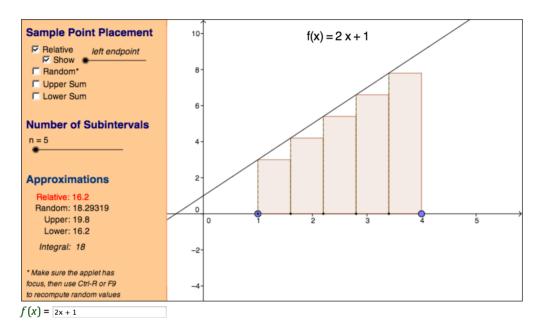


Figure 4.3.3: A left Riemann sum with 5 subintervals for the function f(x) = 2x + 1 on the interval [1, 4]. The value of the sum is $L_5 = 16.2$.

4.3.1 The definition of the definite integral

In both examples in Preview Activity 4.3.1, we saw that as the number of rectangles got larger and larger, the values of L_n , M_n , and R_n all grew closer and closer to the same value. It turns out that this occurs for any continuous function on an interval [a, b], and even more generally for a Riemann sum using any point x_{i+1}^* in the interval $[x_i, x_{i+1}]$. Said differently, as we let $n \to \infty$, it doesn't really matter where we choose to evaluate the function within a

^aMarc Renault, Shippensburg University, Geogebra Applets for Calclulus, http://gvsu.edu/s/5p.

given subinterval, because

$$\lim_{n\to\infty} L_n = \lim_{n\to\infty} R_n = \lim_{n\to\infty} M_n = \lim_{n\to\infty} \sum_{i=1}^n f(x_i^*) \Delta x.$$

That these limits always exist (and share the same value) for a continuous¹ function f allows us to make the following definition.

Definition 4.3.4. The **definite integral** of a continuous function f on the interval [a, b], denoted $\int_a^b f(x) dx$, is the real number given by

$$\int_a^b f(x) dx = \lim_{n \to \infty} \sum_{i=1}^n f(x_i^*) \Delta x,$$

where $\Delta x = \frac{b-a}{n}$, $x_i = a + i\Delta x$ (for i = 0, ..., n), and x_i^* satisfies $x_{i-1} \le x_i^* \le x_i$ (for i = 1, ..., n).

We call the symbol \int the *integral sign*, the values a and b the *limits of integration*, and the function f the *integrand*. The process of determining the real number $\int_a^b f(x) dx$ is called *evaluating the definite integral*. While we will come to understand that there are several different interpretations of the value of the definite integral, for now the most important is that $\int_a^b f(x) dx$ measures the net signed area bounded by y = f(x) and the x-axis on the interval [a, b].

For example, in the notation of the definite integral, if f is the function pictured in Figure 4.3.5 and A_1 , A_2 , and A_3 are the exact areas bounded by f and the x-axis on the respective intervals [a, b], [b, c], and [c, d], then

$$\int_{a}^{b} f(x) dx = A_{1}, \int_{b}^{c} f(x) dx = -A_{2},$$

$$\int_{c}^{d} f(x) dx = A_{3},$$
and
$$\int_{a}^{d} f(x) dx = A_{1} - A_{2} + A_{3}.$$

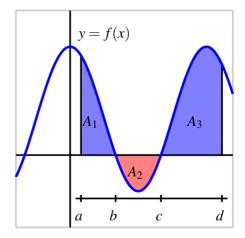


Figure 4.3.5: A continuous function f on the interval [a, d].

¹It turns out that a function need not be continuous in order to have a definite integral. For our purposes, we assume that the functions we consider are continuous on the interval(s) of interest. It is straightforward to see that any function that is piecewise continuous on an interval of interest will also have a well-defined definite integral.

We can also use definite integrals to express the change in position and distance traveled by a moving object. In the setting of a velocity function v on an interval [a,b], it follows from our work above and in preceding sections that the change in position, s(b) - s(a), is given by

$$s(b) - s(a) = \int_a^b v(t) dt.$$

If the velocity function is nonnegative on [a,b], then $\int_a^b v(t) \, dt$ tells us the distance the object traveled. When velocity is sometimes negative on [a,b], the areas bounded by the function on intervals where v does not change sign can be found using integrals, and the sum of these values will tell us the distance the object traveled.

If we wish to compute the value of a definite integral using the definition, we have to take the limit of a sum. While this is possible to do in select circumstances, it is also tedious and time-consuming; moreover, computing these limits does not offer much additional insight into the meaning or interpretation of the definite integral. Instead, in Section 4.4, we will learn the Fundamental Theorem of Calculus, a result that provides a shortcut for evaluating a large class of definite integrals. This will enable us to determine the exact net signed area bounded by a continuous function and the x-axis in many circumstances, including examples such as $\int_1^4 (x^2 + 1) \, dx$, which we approximated by Riemann sums in Preview Activity 4.3.1.

For now, our goal is to understand the meaning and properties of the definite integral, rather than how to actually compute its value using ideas in calculus. Thus, we temporarily rely on the net signed area interpretation of the definite integral and observe that if a given curve produces regions whose areas we can compute exactly through known area formulas, we can thus compute the exact value of the integral.

For instance, if we wish to evaluate the definite integral $\int_1^4 (2x+1) dx$, we can observe that the region bounded by this function and the *x*-axis is the trapezoid shown in Figure 4.3.6, and by the known formula for the area of a trapezoid, its area is $A = \frac{1}{2}(3+9) \cdot 3 = 18$, so

$$\int_{1}^{4} (2x+1) \, dx = 18.$$

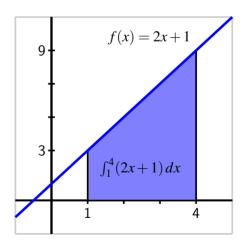


Figure 4.3.6: The area bounded by f(x) = 2x + 1 and the *x*-axis on the interval [1, 4].

Activity 4.3.2. Use known geometric formulas and the net signed area interpretation of the definite integral to evaluate each of the definite integrals below.



b.
$$\int_{-1}^{4} (2-2x) dx$$

a.
$$\int_0^1 3x \, dx$$

b. $\int_{-1}^4 (2 - 2x) \, dx$
c. $\int_{-1}^1 \sqrt{1 - x^2} \, dx$

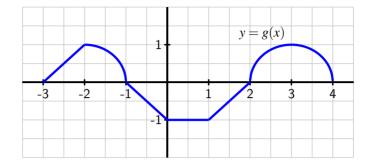


Figure 4.3.7: A function *q* that is piecewise defined; each piece of the function is part of a circle or part of a line.

d. $\int_{-3}^{4} g(x) dx$, where g is the function pictured in Figure 4.3.7. Assume that each portion of q is either part of a line or part of a circle.

4.3.2 Some properties of the definite integral

With the perspective that the definite integral of a function f over an interval [a, b] measures the net signed area bounded by f and the x-axis over the interval, we naturally arrive at several different standard properties of the definite integral. In addition, it is helpful to remember that the definite integral is defined in terms of Riemann sums that fundamentally consist of the areas of rectangles.

If we consider the definite integral $\int_a^a f(x) dx$ for any real number a, it is evident that no area is being bounded because the interval begins and ends with the same point. Hence,

If *f* is a continuous function and *a* is a real number, then $\int_a^a f(x) dx = 0$.

Next, we consider the results of subdividing a given interval. In Figure 4.3.8, we see that

$$\int_{a}^{b} f(x) dx = A_{1}, \int_{b}^{c} f(x) dx = A_{2},$$
and
$$\int_{a}^{c} f(x) dx = A_{1} + A_{2},$$

which is indicative of the following general rule.

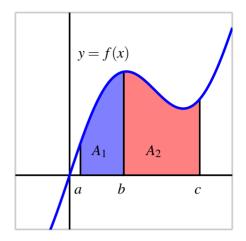


Figure 4.3.8: The area bounded by y = f(x) on the interval [a, c].

If f is a continuous function and a, b, and c are real numbers, then

$$\int_a^c f(x) dx = \int_a^b f(x) dx + \int_b^c f(x) dx.$$

While this rule is most apparent in the situation where a < b < c, it in fact holds in general for any values of a, b, and c. This result is connected to another property of the definite integral, which states that if we reverse the order of the limits of integration, we change the sign of the integral's value.

If f is a continuous function and a and b are real numbers, then

$$\int_{b}^{a} f(x) dx = -\int_{a}^{b} f(x) dx.$$

This result makes sense because if we integrate from a to b, then in the defining Riemann sum $\Delta x = \frac{b-a}{n}$, while if we integrate from b to a, $\Delta x = \frac{a-b}{n} = -\frac{b-a}{n}$, and this is the only change in the sum used to define the integral.

There are two additional properties of the definite integral that we need to understand. Recall that when we worked with derivative rules in Chapter 2, we found that both the Constant Multiple Rule and the Sum Rule held. The Constant Multiple Rule tells us that if f is a differentiable function and k is a constant, then

$$\frac{d}{dx}[kf(x)] = kf'(x),$$

and the Sum Rule states that if f and g are differentiable functions, then

$$\frac{d}{dx}[f(x) + g(x)] = f'(x) + g'(x).$$

These rules are useful because they enable us to deal individually with the simplest parts of certain functions and take advantage of the elementary operations of addition and multiplying by a constant. They also tell us that the process of taking the derivative respects addition and multiplying by constants in the simplest possible way.

It turns out that similar rules hold for the definite integral. First, let's consider the situation pictured in Figure 4.3.9,

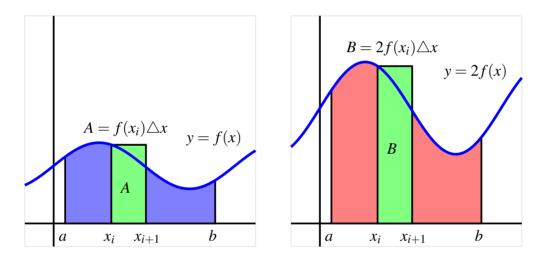


Figure 4.3.9: The areas bounded by y = f(x) and y = 2f(x) on [a, b].

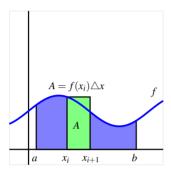
where we examine the effect of multiplying a function by a factor of 2 on the area it bounds with the x-axis. Because multiplying the function by 2 doubles its height at every x-value, we see that if we consider a typical rectangle from a Riemann sum, the difference in area comes from the changed height of the rectangle: $f(x_i)$ for the original function, versus $2f(x_i)$ in the doubled function, in the case of left sum. Hence, in Figure 4.3.9, we see that for the pictured rectangles with areas A and B, it follows B = 2A. As this will happen in every such rectangle, regardless of the value of n and the type of sum we use, we see that in the limit, the area of the red region bounded by y = 2f(x) will be twice that of the area of the blue region bounded by y = f(x). As there is nothing special about the value 2 compared to an arbitrary constant k, it turns out that the following general principle holds.

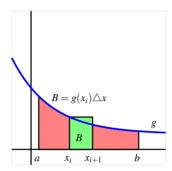
Constant Multiple Rule

If f is a continuous function and k is any real number then

$$\int_a^b k \cdot f(x) \, dx = k \int_a^b f(x) \, dx.$$

Finally, we see a similar situation geometrically with the sum of two functions f and g.





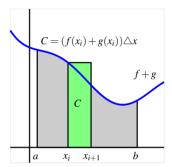


Figure 4.3.10: The areas bounded by y = f(x) and y = g(x) on [a, b], as well as the area bounded by y = f(x) + g(x).

In particular, as shown in Figure 4.3.10, if we take the sum of two functions f and g, at every point in the interval, the height of the function f+g is given by $(f+g)(x_i)=f(x_i)+g(x_i)$, which is the sum of the individual function values of f and g (taken at left endpoints). Hence, for the pictured rectangles with areas A, B, and C, it follows that C=A+B, and because this will occur for every such rectangle, in the limit the area of the gray region will be the sum of the areas of the blue and red regions. Stated in terms of definite integrals, we have the following general rule.

Sum Rule

If *f* and *g* are continuous functions, then

$$\int_{a}^{b} [f(x) + g(x)] dx = \int_{a}^{b} f(x) dx + \int_{a}^{b} g(x) dx.$$

More generally, the Constant Multiple and Sum Rules can be combined to make the observation that for any continuous functions f and g and any constants c and k,

$$\int_a^b \left[c f(x) \pm k g(x) \right] dx = c \int_a^b f(x) \, dx \pm k \int_a^b g(x) \, dx.$$

Activity 4.3.3. Suppose that the following information is known about the functions f,

- $\int_0^2 f(x) dx = -3$; $\int_2^5 f(x) dx = 2$ $\int_0^2 g(x) dx = 4$; $\int_2^5 g(x) dx = -1$ $\int_0^2 x^2 dx = \frac{8}{3}$; $\int_2^5 x^2 dx = \frac{117}{3}$

 - $\int_0^2 x^3 dx = 4$; $\int_2^5 x^3 dx = \frac{609}{4}$

Use the provided information and the rules discussed in the preceding section to evaluate each of the following definite integrals.

a.
$$\int_5^2 f(x) \, dx$$

d.
$$\int_2^5 (3x^2 - 4x^3) dx$$

b.
$$\int_0^5 g(x) \, dx$$

e.
$$\int_{5}^{0} (2x^3 - 7g(x)) dx$$

a.
$$\int_{5}^{2} f(x) dx$$

b. $\int_{0}^{5} g(x) dx$
c. $\int_{0}^{5} (f(x) + g(x)) dx$

4.3.3 How the definite integral is connected to a function's average value

One of the most valuable applications of the definite integral is that it provides a way to meaningfully discuss the average value of a function, even for a function that takes on infinitely many values. Recall that if we wish to take the average of n numbers y_1, y_2, \ldots, y_n , we do so by computing

$$AVG = \frac{y_1 + y_2 + \dots + y_n}{n}.$$

Since integrals arise from Riemann sums in which we add *n* values of a function, it should not be surprising that evaluating an integral is something like averaging the output values of a function. Consider, for instance, the right Riemann sum R_n of a function f, which is given by

$$R_n = f(x_1)\Delta x + f(x_2)\Delta x + \dots + f(x_n)\Delta x = (f(x_1) + f(x_2) + \dots + f(x_n))\Delta x.$$

Since $\Delta x = \frac{b-a}{n}$, we can thus write

$$R_n = (f(x_1) + f(x_2) + \dots + f(x_n)) \cdot \frac{b - a}{n}$$

$$= (b - a) \frac{f(x_1) + f(x_2) + \dots + f(x_n)}{n}.$$
(4.3.1)

Here, we see that the right Riemann sum with *n* subintervals is the length of the interval (b-a) times the average of the n function values found at the right endpoints. And just as with our efforts to compute area, we see that the larger the value of n we use, the more accurate our average of the values of f will be. Indeed, we will define the average value of f on [a, b] to be

$$f_{\text{AVG}[a,b]} = \lim_{n \to \infty} \frac{f(x_1) + f(x_2) + \dots + f(x_n)}{n}.$$

But we also know that for any continuous function f on [a,b], taking the limit of a Riemann sum leads precisely to the definite integral. That is, $\lim_{n\to\infty} R_n = \int_a^b f(x) \, dx$, and thus taking the limit as $n\to\infty$ in Equation (4.3.1), we have that

$$\int_{a}^{b} f(x) dx = (b - a) \cdot f_{\text{AVG}[a,b]}.$$
 (4.3.2)

Solving Equation (4.3.2) for $f_{AVG[a,b]}$, we have the following general principle.

Average value of a function

If f is a continuous function on [a, b], then its average value on [a, b] is given by the formula

$$f_{\text{AVG}[a,b]} = \frac{1}{b-a} \cdot \int_a^b f(x) \, dx.$$

Observe that Equation (4.3.2) tells us another way to interpret the definite integral: the definite integral of a function f from a to b is the length of the interval (b-a) times the average value of the function on the interval. In addition, Equation (4.3.2) has a natural visual interpretation when the function f is nonnegative on [a, b].

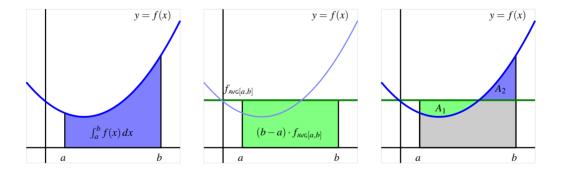


Figure 4.3.11: A function y = f(x), the area it bounds, and its average value on [a, b].

Consider Figure 4.3.11, where we see at left the shaded region whose area is $\int_a^b f(x) dx$, at center the shaded rectangle whose dimensions are (b-a) by $f_{\text{AVG}[a,b]}$, and at right these two figures superimposed. Specifically, note that in dark green we show the horizontal line $y = f_{\text{AVG}[a,b]}$. Thus, the area of the green rectangle is given by $(b-a) \cdot f_{\text{AVG}[a,b]}$, which is precisely the value of $\int_a^b f(x) dx$. Said differently, the area of the blue region in the left figure is the same as that of the green rectangle in the center figure; this can also be seen by observing that

the areas A_1 and A_2 in the rightmost figure appear to be equal. Ultimately, the average value of a function enables us to construct a rectangle whose area is the same as the value of the definite integral of the function on the interval. The java applet² at http://gvsu.edu/s/az provides an opportunity to explore how the average value of the function changes as the interval changes, through an image similar to that found in Figure 4.3.11.

Activity 4.3.4. Suppose that $v(t) = \sqrt{4 - (t - 2)^2}$ tells us the instantaneous velocity of a moving object on the interval $0 \le t \le 4$, where t is measured in minutes and v is measured in meters per minute.

- a. Sketch an accurate graph of y = v(t). What kind of curve is $y = \sqrt{4 (t-2)^2}$?
- b. Evaluate $\int_0^4 v(t) dt$ exactly.
- c. In terms of the physical problem of the moving object with velocity v(t), what is the meaning of $\int_0^4 v(t) dt$? Include units on your answer.
- d. Determine the exact average value of v(t) on [0,4]. Include units on your answer.
- e. Sketch a rectangle whose base is the line segment from t=0 to t=4 on the t-axis such that the rectangle's area is equal to the value of $\int_0^4 v(t) \, dt$. What is the rectangle's exact height?
- f. How can you use the average value you found in (d) to compute the total distance traveled by the moving object over [0, 4]?

Summary

- Any Riemann sum of a continuous function f on an interval [a, b] provides an estimate of the net signed area bounded by the function and the horizontal axis on the interval. Increasing the number of subintervals in the Riemann sum improves the accuracy of this estimate, and letting the number of subintervals increase without bound results in the values of the corresponding Riemann sums approaching the exact value of the enclosed net signed area.
- When we take the just described limit of Riemann sums, we arrive at what we call the definite integral of f over the interval [a,b]. In particular, the symbol $\int_a^b f(x) dx$ denotes the definite integral of f over [a,b], and this quantity is defined by the equation

$$\int_a^b f(x) dx = \lim_{n \to \infty} \sum_{i=1}^n f(x_i^*) \Delta x,$$

where $\Delta x = \frac{b-a}{n}$, $x_i = a + i\Delta x$ (for i = 0, ..., n), and x_i^* satisfies $x_{i-1} \le x_i^* \le x_i$ (for i = 1, ..., n).

²David Austin, http://gvsu.edu/s/5r.

- The definite integral $\int_a^b f(x) dx$ measures the exact net signed area bounded by f and the horizontal axis on [a,b]; in addition, the value of the definite integral is related to what we call the average value of the function on [a,b]: $f_{AVG[a,b]} = \frac{1}{b-a} \cdot \int_a^b f(x) dx$. In the setting where we consider the integral of a velocity function v, $\int_a^b v(t) dt$ measures the exact change in position of the moving object on [a,b]; when v is nonnegative, $\int_a^b v(t) dt$ is the object's distance traveled on [a,b].
- The definite integral is a sophisticated sum, and thus has some of the same natural properties that finite sums have. Perhaps most important of these is how the definite integral respects sums and constant multiples of functions, which can be summarized by the rule

$$\int_a^b \left[c f(x) \pm k g(x) \right] dx = c \int_a^b f(x) dx \pm k \int_a^b g(x) dx$$

where f and g are continuous functions on [a, b] and c and k are arbitrary constants.

Exercises

1. Use the following figure, which shows a graph of f(x) to find each of the indicated integrals.





Note that the first area (with vertical, red shading) is 55 and the second (with oblique, black shading) is 5.

$$A. \int_{a}^{b} f(x)dx =$$

$$B. \int_{b}^{c} f(x)dx =$$

$$C. \int_{a}^{c} f(x) dx =$$

$$D. \int_{a}^{c} |f(x)| dx = \boxed{$$

2. Use the graph of f(x) shown below to find the following integrals.



A.
$$\int_{-4}^{0} f(x)dx =$$

B. If the vertical red shaded area in the graph has area A, estimate: $\int_{-4}^{6} f(x)dx =$ (Your estimate may be written in terms of A.)

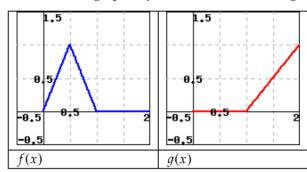


3. Find the average value of f(x) = 6x + 5 over [2, 6]

average value =

West Hotel

4. The figure below to the left is a graph of f(x), and below to the right is g(x).



(a)

What is the average value of f(x) on $0 \le x \le 2$?

avg value =

(b)

What is the average value of q(x) on $0 \le x \le 2$?

avg value =

(c)

What is the average value of $f(x) \cdot g(x)$ on $0 \le x \le 2$?

avg value =

(d)

Is the following statement true?

$$Average(f) \cdot Average(g) = Average(f \cdot g)$$



5. Use the figure below, which shows the graph of y = f(x), to answer the following questions.



A. Estimate the integral: $\int_{-3}^{3} f(x) dx \approx$

(You will certainly want to use an enlarged version of the graph to obtain your estimate.)

B. Which of the following average values of f is larger?



6. Suppose $\int_{10}^{14.5} f(x)dx = 5$, $\int_{10}^{11.5} f(x)dx = 6$, $\int_{13}^{14.5} f(x)dx = 8$.

 $\int_{11.5}^{13} f(x) dx = \boxed{}$

$$\int_{13}^{11.5} (5f(x) - 6) dx = \boxed{}$$

- 7. The velocity of an object moving along an axis is given by the piecewise linear function v that is pictured in Figure 4.3.12. Assume that the object is moving to the right when its velocity is positive, and moving to the left when its velocity is negative. Assume that the given velocity function is valid for t = 0 to t = 4.
 - a. Write an expression involving definite integrals whose value is the total change in position of the object on the interval [0, 4].
 - b. Use the provided graph of *v* to determine the value of the total change in position on [0,4].
 - c. Write an expression involving definite integrals whose value is the total distance traveled by the object on [0,4]. What is the exact value of the total distance traveled on [0,4]?
 - d. What is the object's exact average velocity on [0, 4]?
 - e. Find an algebraic formula for the object's position function on [0, 1.5] that satisfies s(0) = 0.

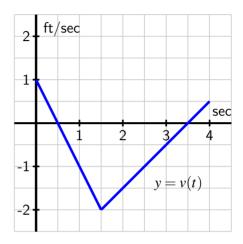


Figure 4.3.12: The velocity function of a moving object.

- 8. Suppose that the velocity of a moving object is given by v(t) = t(t-1)(t-3), measured in feet per second, and that this function is valid for $0 \le t \le 4$.
 - a. Write an expression involving definite integrals whose value is the total change in position of the object on the interval [0, 4].
 - b. Use appropriate technology (such as http://gvsu.edu/s/a9³) to compute Riemann sums to estimate the object's total change in position on [0, 4]. Work to ensure that your estimate is accurate to two decimal places, and explain how you know this to be the case.
 - c. Write an expression involving definite integrals whose value is the total distance traveled by the object on [0,4].
 - d. Use appropriate technology to compute Riemann sums to estimate the object's total distance travelled on [0, 4]. Work to ensure that your estimate is accurate to two decimal places, and explain how you know this to be the case.

³Marc Renault, Shippensburg University.

Chapter 4 The Definite Integral

- e. What is the object's average velocity on [0, 4], accurate to two decimal places?
- **9.** Consider the graphs of two functions f and g that are provided in Figure 4.3.13. Each piece of f and g is either part of a straight line or part of a circle.

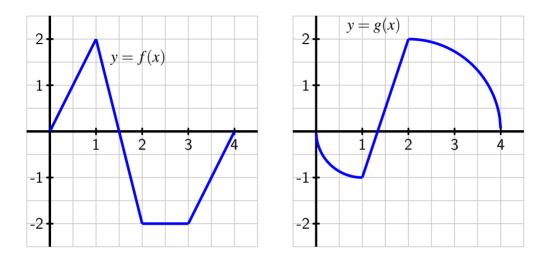


Figure 4.3.13: Two functions f and g.

- a. Determine the exact value of $\int_0^1 [f(x) + g(x)] dx$.
- b. Determine the exact value of $\int_1^4 [2f(x) 3g(x)] dx$.
- c. Find the exact average value of h(x) = g(x) f(x) on [0, 4].
- d. For what constant c does the following equation hold?

$$\int_0^4 c \, dx = \int_0^4 [f(x) + g(x)] \, dx$$

- **10.** Let $f(x) = 3 x^2$ and $g(x) = 2x^2$.
 - a. On the interval [-1,1], sketch a labeled graph of y = f(x) and write a definite integral whose value is the exact area bounded by y = f(x) on [-1,1].
 - b. On the interval [-1,1], sketch a labeled graph of y = g(x) and write a definite integral whose value is the exact area bounded by y = g(x) on [-1,1].
 - c. Write an expression involving a difference of definite integrals whose value is the exact area that lies between y = f(x) and y = g(x) on [-1, 1].
 - d. Explain why your expression in (c) has the same value as the single integral $\int_{-1}^{1} [f(x) g(x)] dx$.

e. Explain why, in general, if $p(x) \ge q(x)$ for all x in [a,b], the exact area between y = p(x) and y = q(x) is given by

$$\int_{a}^{b} [p(x) - q(x)] dx.$$

4.4 The Fundamental Theorem of Calculus

Motivating Questions

- How can we find the exact value of a definite integral without taking the limit of a Riemann sum?
- What is the statement of the Fundamental Theorem of Calculus, and how do antiderivatives of functions play a key role in applying the theorem?
- What is the meaning of the definite integral of a rate of change in contexts other than when the rate of change represents velocity?

Much of our work in Chapter 4 has been motivated by the velocity-distance problem: if we know the instantaneous velocity function, v(t), for a moving object on a given time interval [a,b], can we determine its exact distance traveled on [a,b]? In the vast majority of our discussion in Sections 4.1- Section 4.3, we have focused on the fact that this distance traveled is connected to the area bounded by y=v(t) and the t-axis on [a,b]. In particular, for any nonnegative velocity function y=v(t) on [a,b], we know that the exact area bounded by the velocity curve and the t-axis on the interval tells us the total distance traveled, which is also the value of the definite integral $\int_a^b v(t) \, dt$. In the situation where velocity is sometimes negative, the total area bounded by the velocity function still tells us distance traveled, while the net signed area that the function bounds tells us the object's change in position.

Recall, for instance, the introduction to Section 4.2, where we observed that for the velocity function in Figure 4.4.1, the total distance D traveled by the moving object on [a, b] is

$$D = A_1 + A_2 + A_3,$$

while the total change in the object's position on [a, b] is

$$s(b) - s(a) = A_1 - A_2 + A_3$$
.

The areas A_1 , A_2 , and A_3 , which are each given by definite integrals, may be computed through limits of Riemann sums (and in select special circumstances through familiar geometric formulas).

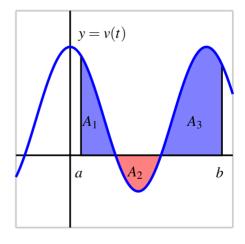


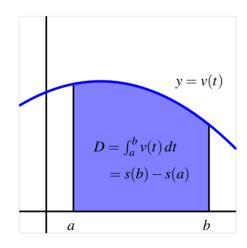
Figure 4.4.1: A velocity function that is sometimes negative.

In the present section we turn our attention to an alternate approach, similar to the one we encountered in Activity 4.1.3. To explore these ideas further, we consider the following preview activity.

Preview Activity 4.4.1. A student with a third floor dormitory window 32 feet off the ground tosses a water balloon straight up in the air with an initial velocity of 16 feet per second. It turns out that the instantaneous velocity of the water balloon is given by v(t) = -32t + 16, where v is measured in feet per second and t is measured in seconds.

- a. Let s(t) represent the height of the water balloon above ground at time t, and note that s is an antiderivative of v. That is, v is the derivative of s: s'(t) = v(t). Find a formula for s(t) that satisfies the initial condition that the balloon is tossed from 32 feet above ground. In other words, make your formula for s satisfy s(0) = 32.
- b. When does the water balloon reach its maximum height? When does it land?
- c. Compute $s(\frac{1}{2}) s(0)$, $s(2) s(\frac{1}{2})$, and s(2) s(0). What do these represent?
- d. What is the total vertical distance traveled by the water balloon from the time it is tossed until the time it lands?
- e. Sketch a graph of the velocity function y = v(t) on the time interval [0,2]. What is the total net signed area bounded by y = v(t) and the t-axis on [0,2]? Answer this question in two ways: first by using your work above, and then by using a familiar geometric formula to compute areas of certain relevant regions.

4.4.1 The Fundamental Theorem of Calculus



Consider the setting where we know the position function s(t) of an object moving along an axis, as well as its corresponding velocity function v(t), and for the moment let us assume that v(t) is positive on [a, b]. Then, as shown in Figure 4.4.2, we know two different perspectives on the distance, D, the object travels: one is that D = s(b) - s(a), which is the object's change in position. The other is that the distance traveled is the area under the velocity curve, which is given by the definite integral, so $D = \int_a^b v(t) dt$.

Figure 4.4.2: Finding distance traveled when we know a velocity function *v*.

Of course, since both of these expressions tell us the distance traveled, it follows that they are equal, so

$$s(b) - s(a) = \int_{a}^{b} v(t) dt.$$
 (4.4.1)

Furthermore, we know that Equation (4.4.1) holds even when velocity is sometimes negative, since s(b) - s(a) is the object's change in position over [a, b], which is simultaneously measured by the total net signed area on [a, b] given by $\int_a^b v(t) dt$.

Perhaps the most powerful part of Equation (4.4.1) lies in the fact that we can compute the integral's value if we can find a formula for s. Remember, s and v are related by the fact that v is the derivative of s, or equivalently that s is an antiderivative of v. For example, if we have an object whose velocity is $v(t) = 3t^2 + 40$ feet per second (which is always nonnegative), and wish to know the distance traveled on the interval [1, 5], we have that

$$D = \int_{1}^{5} v(t) dt = \int_{1}^{5} (3t^{2} + 40) dt = s(5) - s(1),$$

where s is an antiderivative of v. We know that the derivative of t^3 is $3t^2$ and that the derivative of 40t is 40, so it follows that if $s(t) = t^3 + 40t$, then s is a function whose derivative is $v(t) = s'(t) = 3t^2 + 40$, and thus we have found an antiderivative of v. Therefore,

$$D = \int_{1}^{5} 3t^{2} + 40 dt = s(5) - s(1)$$
$$= (5^{3} + 40 \cdot 5) - (1^{3} + 40 \cdot 1) = 284 \text{ feet.}$$

Note the key lesson of this example: to find the distance traveled, we needed to compute the area under a curve, which is given by the definite integral. But to evaluate the integral, we found an antiderivative, s, of the velocity function, and then computed the total change in s on the interval. In particular, observe that we have found the exact area of the region shown in Figure 4.4.3, and done so without a familiar formula (such as those for the area of a triangle or circle) and without directly computing the limit of a Riemann sum.

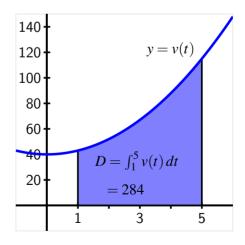


Figure 4.4.3: The exact area of the region enclosed by $v(t) = 3t^2 + 40$ on [1, 5].

As we proceed to thinking about contexts other than just velocity and position, it is advantageous to have a shorthand symbol for a function's antiderivative. In the general setting of a continuous function f, we will often denote an antiderivative of f by F, so that the relationship between F and f is that F'(x) = f(x) for all relevant x. Using the notation V in place of s (so that V is an antiderivative of s) in Equation (4.4.1), we find it is equivalent to write that

$$V(b) - V(a) = \int_{a}^{b} v(t) dt.$$
 (4.4.2)

Now, in the general setting of wanting to evaluate the definite integral $\int_a^b f(x) \, dx$ for an arbitrary continuous function f, we could certainly think of f as representing the velocity of some moving object, and x as the variable that represents time. And again, Equations (4.4.1) and (4.4.2) hold for any continuous velocity function, even when v is sometimes negative. This leads us to see that Equation (4.4.2) tells us something even more important than the change in position of a moving object: it offers a shortcut route to evaluating any definite integral, provided that we can find an antiderivative of the integrand. The Fundamental Theorem of Calculus (FTC) summarizes these observations.

Fundamental Theorem of Calculus

If f is a continuous function on [a, b], and F is any antiderivative of f, then $\int_a^b f(x) dx = F(b) - F(a)$.

A common alternate notation for F(b) - F(a) is

$$F(b) - F(a) = F(x)|_a^b,$$

where we read the righthand side as "the function F evaluated from a to b." In this notation, the FTC says that

$$\int_a^b f(x) \, dx = F(x)|_a^b.$$

The FTC opens the door to evaluating exactly a wide range of integrals. In particular, if we are interested in a definite integral for which we can find an antiderivative F for the integrand f, then we can evaluate the integral exactly. For instance since $\frac{d}{dx}[\frac{1}{3}x^3] = x^2$, the FTC tells us that

$$\int_0^1 x^2 dx = \frac{1}{3} x^3 \Big|_0^1$$
$$= \frac{1}{3} (1)^3 - \frac{1}{3} (0)^3$$
$$= \frac{1}{3}.$$

But finding an antiderivative can be far from simple; in fact, often finding a formula for an antiderivative is very hard or even impossible. While we can differentiate just about any function, even some relatively simple ones don't have an elementary antiderivative. A significant portion of integral calculus (which is the main focus of second semester college calculus) is devoted to understanding the problem of finding antiderivatives.

Activity 4.4.2. Use the Fundamental Theorem of Calculus to evaluate each of the following integrals exactly. For each, sketch a graph of the integrand on the relevant interval and write one sentence that explains the meaning of the value of the integral in terms of the (net signed) area bounded by the curve.

a.
$$\int_{-1}^{4} (2-2x) dx$$

c.
$$\int_0^1 e^x dx$$

a.
$$\int_{-1}^{4} (2 - 2x) dx$$

b. $\int_{0}^{\frac{\pi}{2}} \sin(x) dx$

d.
$$\int_{-1}^{1} x^5 dx$$

e.
$$\int_0^2 (3x^3 - 2x^2 - e^x) dx$$

4.4.2 Basic antiderivatives

The general problem of finding an antiderivative is difficult. In part, this is due to the fact that we are trying to undo the process of differentiating, and the undoing is much more difficult than the doing. For example, while it is evident that an antiderivative of f(x) $\sin(x)$ is $F(x) = -\cos(x)$ and that an antiderivative of $g(x) = x^2$ is $G(x) = \frac{1}{3}x^3$, combinations of f and g can be far more complicated. Consider such functions as

$$5\sin(x) - 4x^2$$
, $x^2\sin(x)$, $\frac{\sin(x)}{x^2}$, and $\sin(x^2)$.

What is involved in trying to find an antiderivative for each? From our experience with derivative rules, we know that while derivatives of sums and constant multiples of basic functions are simple to execute, derivatives involving products, quotients, and composites of familiar functions are much more complicated. Thus, it stands to reason that antidifferentiating products, quotients, and composites of basic functions may be even more challenging. We defer our study of all but the most elementary antiderivatives to later in the text.

We do note that each time we have a function for which we know its derivative, we have a function-derivative pair, which also leads us to knowing the antiderivative of a function. For instance, since we know that

$$\frac{d}{dx}[-\cos(x)] = \sin(x),$$

it follows that $F(x) = -\cos(x)$ is an antiderivative of $f(x) = \sin(x)$. It is equivalent to say that $f(x) = \sin(x)$ is the derivative of $F(x) = -\cos(x)$, and thus F and f together form the function-derivative pair. Clearly, every basic derivative rule leads us to such a pair, and thus to a known antiderivative. In Activity 4.4.3, we will construct a list of most of the basic antiderivatives we know at this time. Furthermore, those rules will enable us to antidifferentiate sums and constant multiples of basic functions. For example, if $f(x) = 5\sin(x) - 4x^2$, note that since $-\cos(x)$ is an antiderivative of $\sin(x)$ and $\frac{1}{3}x^3$ is an antiderivative of x^2 , it follows that

$$F(x) = -5\cos(x) - \frac{4}{3}x^3$$

is an antiderivative of f, by the sum and constant multiple rules for differentiation.

Finally, before proceeding to build a list of common functions whose antiderivatives we know, we revisit the fact that each function has more than one antiderivative. Because the derivative of any constant is zero, any time we seek an arbitrary antiderivative, we may add a constant of our choice. For instance, if we want to determine an antiderivative of $g(x) = x^2$, we know that $G(x) = \frac{1}{3}x^3$ is one such function. But we could alternately have chosen $G(x) = \frac{1}{3}x^3 + 7$, since in this case as well, $G'(x) = x^2$. In some contexts later on in calculus, it is important to discuss the most general antiderivative of a function. If $g(x) = x^2$, we say that the *general antiderivative* of g is

$$G(x) = \frac{1}{3}x^3 + C,$$

where C represents an arbitrary real number constant. Regardless of the formula for g, including +C in the formula for its antiderivative G results in the most general possible antiderivative.

Our primary current interest in antiderivatives is for use in evaluating definite integrals by the Fundamental Theorem of Calculus. In that situation, the arbitrary constant *C* is irrelevant, and thus we usually omit it. To see why, consider the definite integral

$$\int_0^1 x^2 dx.$$

For the integrand $g(x) = x^2$, suppose we find and use the general antiderivative $G(x) = \frac{1}{3}x^3 + C$. Then, by the FTC,

$$\int_0^1 x^2 dx = \frac{1}{3}x^3 + C \Big|_0^1$$

$$= \left(\frac{1}{3}(1)^3 + C\right) - \left(\frac{1}{3}(0)^3 + C\right)$$

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

$$= \frac{1}{3}.$$

Specifically, we observe that the C-values appear as opposites in the evaluation of the integral and thus do not affect the definite integral's value. In the same way, the potential inclusion of +C with the antiderivative has no bearing on any definite integral, and thus we generally choose to omit this possible constant whenever we evaluate an integral using the Fundamental Theorem of Calculus.

In the following activity, we work to build a list of basic functions whose antiderivatives we already know.

Activity 4.4.3. Use your knowledge of derivatives of basic functions to complete the above table of antiderivatives. For each entry, your task is to find a function F whose derivative is the given function f. When finished, use the FTC and the results in the table to evaluate the three given definite integrals.

given function, $f(x)$	antiderivative, $F(x)$
k, (k is constant)	
x^n , $n \neq -1$	
$\frac{1}{x}$, $x > 0$	
sin(x)	
$\cos(x)$	
sec(x) tan(x)	
$\csc(x)\cot(x)$	
$sec^2(x)$	
$csc^2(x)$	
e^x	
$a^x (a > 1)$	
$\frac{1}{1+x^2}$	
$\frac{1}{\sqrt{1-x^2}}$	

Table 4.4.4: Familiar basic functions and their antiderivatives.

a.
$$\int_0^1 (x^3 - x - e^x + 2) dx$$

b.
$$\int_0^{\pi/3} (2\sin(t) - 4\cos(t) + \sec^2(t) - \pi) dt$$

c.
$$\int_0^1 (\sqrt{x} - x^2) dx$$

4.4.3 The total change theorem

As we use the Fundamental Theorem of Calculus to evaluate definite integrals, it is essential that we remember and understand the meaning of the numbers we find. We briefly summarize three key interpretations to date.

- For a moving object with instantaneous velocity v(t), the object's change in position on the time interval [a,b] is given by $\int_a^b v(t) \, dt$, and whenever $v(t) \ge 0$ on [a,b], $\int_a^b v(t) \, dt$ tells us the total distance traveled by the object on [a,b].
- For any continuous function f, its definite integral $\int_a^b f(x) dx$ represents the total net

signed area bounded by y = f(x) and the x-axis on [a, b], where regions that lie below the x-axis have a minus sign associated with their area.

• The value of a definite integral is linked to the average value of a function: for a continuous function f on [a,b], its average value $f_{AVG[a,b]}$ is given by

$$f_{\text{AVG}[a,b]} = \frac{1}{b-a} \int_a^b f(x) \, dx.$$

The Fundamental Theorem of Calculus now enables us to evaluate exactly (without taking a limit of Riemann sums) any definite integral for which we are able to find an antiderivative of the integrand.

A slight change in notational perspective allows us to gain even more insight into the meaning of the definite integral. To begin, recall Equation (4.4.2), where we wrote the Fundamental Theorem of Calculus for a velocity function v with antiderivative V as

$$V(b) - V(a) = \int_a^b v(t) dt.$$

If we instead replace V with s (which represents position) and replace v with s' (since velocity is the derivative of position), Equation (4.4.2) equivalently reads

$$s(b) - s(a) = \int_{a}^{b} s'(t) dt.$$
 (4.4.3)

In words, this version of the FTC tells us that the total change in the object's position function on a particular interval is given by the definite integral of the position function's derivative over that interval.

Of course, this result is not limited to only the setting of position and velocity. Writing the result in terms of a more general function f, we have the Total Change Theorem.

Total Change Theorem

If f is a continuously differentiable function on [a, b] with derivative f', then $f(b) - f(a) = \int_a^b f'(x) dx$. That is, the definite integral of the derivative of a function on [a, b] is the total change of the function itself on [a, b].

The Total Change Theorem tells us more about the relationship between the graph of a function and that of its derivative. Recall Figure 1.4.1, which provided one of the first times we saw that heights on the graph of the derivative function come from slopes on the graph of the function itself. That observation occurred in the context where we knew f and were seeking f'; if now instead we think about knowing f' and seeking information about f, we can instead say the following:

differences in heights on f correspond to net signed areas bounded by f'.

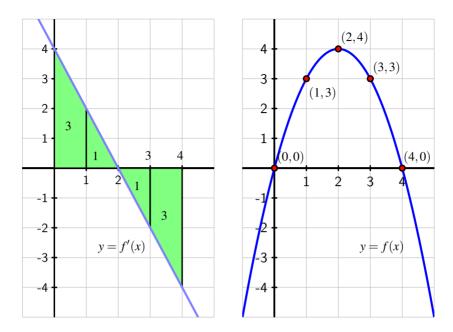


Figure 4.4.5: The graphs of f'(x) = 4 - 2x (at left) and an antiderivative $f(x) = 4x - x^2$ at right. Differences in heights on f correspond to net signed areas bounded by f'.

To see why this is so, say we consider the difference f(1) - f(0). Note that this value is 3, in part because f(1) = 3 and f(0) = 0, but also because the net signed area bounded by y = f'(x) on [0,1] is 3. That is, $f(1) - f(0) = \int_0^1 f'(x) dx$. A similar pattern holds throughout, including the fact that since the total net signed area bounded by f' on [0,4] is 0, $\int_0^4 f'(x) dx = 0$, so it must be that f(4) - f(0) = 0, so f(4) = f(0).

Beyond this general observation about area, the Total Change Theorem enables us to consider interesting and important problems where we know the rate of change, and answer key questions about the function whose rate of change we know.

Example 4.4.6. Suppose that pollutants are leaking out of an underground storage tank at a rate of r(t) gallons/day, where t is measured in days. It is conjectured that r(t) is given by the formula $r(t) = 0.0069t^3 - 0.125t^2 + 11.079$ over a certain 12-day period. The graph of y = r(t) is given in Figure 4.4.7. What is the meaning of $\int_4^{10} r(t) \, dt$ and what is its value? What is the average rate at which pollutants are leaving the tank on the time interval $4 \le t \le 10$?

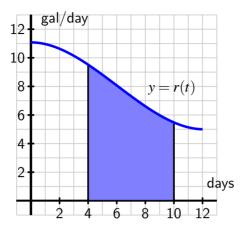


Figure 4.4.7: The rate r(t) of pollution leaking from a tank, measured in gallons per day.

Solution. We know that since $r(t) \ge 0$, the value of $\int_4^{10} r(t) dt$ is the area under the curve on the interval [4, 10]. If we think about this area from the perspective of a Riemann sum, the rectangles will have heights measured in gallons per day and widths measured in days, thus the area of each rectangle will have units of

$$\frac{\text{gallons}}{\text{day}} \cdot \text{days} = \text{gallons}.$$

Thus, the definite integral tells us the total number of gallons of pollutant that leak from the tank from day 4 to day 10. The Total Change Theorem tells us the same thing: if we let R(t) denote the function that measures the total number of gallons of pollutant that have leaked from the tank up to day t, then R'(t) = r(t), and

$$\int_{4}^{10} r(t) dt = R(10) - R(4),$$

which is the total change in the function that measures total gallons leaked over time, thus the number of gallons that have leaked from day 4 to day 10.

To compute the exact value, we use the Fundamental Theorem of Calculus. Antidifferentiating $r(t) = 0.0069t^3 - 0.125t^2 + 11.079$, we find that

$$\int_{4}^{10} 0.0069t^{3} - 0.125t^{2} + 11.079 dt = 0.0069 \cdot \frac{1}{4} t^{4} - 0.125 \cdot \frac{1}{3} t^{3} + 11.079t \Big|_{4}^{10}$$

$$\approx 44.282.$$

Thus, approximately 44.282 gallons of pollutant leaked over the six day time period.

To find the average rate at which pollutant leaked from the tank over $4 \le t \le 10$, we want to compute the average value of r on [4, 10]. Thus,

$$r_{\text{AVG}[4,10]} = \frac{1}{10 - 4} \int_{4}^{10} r(t) \, dt \approx \frac{44.282}{6} = 7.380,$$

which has its units measured in gallons per day.

Activity 4.4.4. During a 40-minute workout, a person riding an exercise machine burns calories at a rate of c calories per minute, where the function y = c(t) is given in Figure 4.4.8. On the interval $0 \le t \le 10$, the formula for c is $c(t) = -0.05t^2 + t + 10$, while on $30 \le t \le 40$, its formula is $c(t) = -0.05t^2 + 3t - 30$.

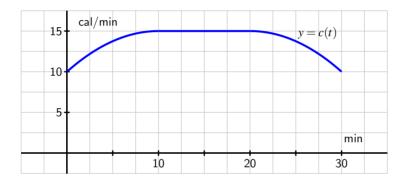


Figure 4.4.8: The rate c(t) at which a person exercising burns calories, measured in calories per minute.

- a. What is the exact total number of calories the person burns during the first 10 minutes of her workout?
- b. Let C(t) be an antiderivative of c(t). What is the meaning of C(40) C(0) in the context of the person exercising? Include units on your answer.
- c. Determine the exact average rate at which the person burned calories during the 40-minute workout.
- d. At what time(s), if any, is the instantaneous rate at which the person is burning calories equal to the average rate at which she burns calories, on the time interval $0 \le t \le 40$?

Summary

- We can find the exact value of a definite integral without taking the limit of a Riemann sum or using a familiar area formula by finding the antiderivative of the integrand, and hence applying the Fundamental Theorem of Calculus.
- The Fundamental Theorem of Calculus says that if f is a continuous function on [a,b]and F is an antiderivative of f, then

$$\int_a^b f(x) \, dx = F(b) - F(a).$$

Hence, if we can find an antiderivative for the integrand f, evaluating the definite integral comes from simply computing the change in F on [a, b].

• A slightly different perspective on the FTC allows us to restate it as the Total Change Theorem, which says that

$$\int_a^b f'(x) \, dx = f(b) - f(a),$$

for any continuously differentiable function f. This means that the definite integral of the instantaneous rate of change of a function f on an interval [a, b] is equal to the total change in the function f on [a, b].

Exercises

1. Use the following figure, which shows a graph of f(x) to find each of the indicated integrals.





Note that the first area (with vertical, red shading) is 9 and the second (with oblique, black shading) is 3.

$$A. \int_{a}^{b} f(x)dx =$$

$$B. \int_{b}^{c} f(x)dx =$$

B.
$$\int_b f(x)dx =$$
C. $\int_a^c f(x)dx =$

$$D. \int_{a}^{c} |f(x)| dx =$$

2. Use the graph of f(x) shown below to find the following integrals.



Chapter 4 The Definite Integral



$$A. \int_{-3}^{0} f(x) dx =$$

B. If the vertical red shaded area in the graph has area A, estimate: $\int_{-3}^{5} f(x)dx =$ (Your estimate may be written in terms of A.)

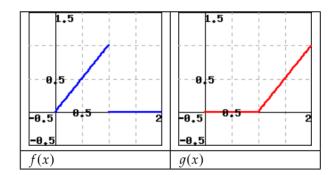
極
WeBWorl

3. Find the average value of f(x) = 4x + 4 over [4, 9]

average value =

極
HieRillork

4. The figure below to the left is a graph of f(x), and below to the right is g(x).



(a)

What is the average value of f(x) on $0 \le x \le 2$?

avg value =

(b)

What is the average value of g(x) on $0 \le x \le 2$?

avg value =

(c)

What is the average value of $f(x) \cdot g(x)$ on $0 \le x \le 2$?

avg value =

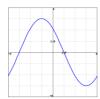
(d)

Is the following statement true?

 $Average(f) \cdot Average(g) = Average(f \cdot q)$



5. Use the figure below, which shows the graph of y = f(x), to answer the following questions.



A. Estimate the integral: $\int_{-3}^{3} f(x) dx \approx$

(You will certainly want to use an enlarged version of the graph to obtain your estimate.)

B. Which of the following average values of f is larger?

6. Suppose
$$\int_{3}^{7.5} f(x)dx = 4$$
, $\int_{3}^{4.5} f(x)dx = 1$, $\int_{6}^{7.5} f(x)dx = 1$.
$$\int_{4.5}^{6} f(x)dx = \boxed{ }$$

$$\int_{6}^{4.5} (4f(x) - 1)dx = \boxed{ }$$



7. The instantaneous velocity (in meters per minute) of a moving object is given by the function v as pictured in Figure 4.4.9. Assume that on the interval $0 \le t \le 4$, v(t) is given by $v(t) = -\frac{1}{4}t^3 + \frac{3}{2}t^2 + 1$, and that on every other interval v is piecewise linear, as shown.

- a. Determine the exact distance traveled by the object on the time interval $0 \le t \le 4$.
- b. What is the object's average velocity on [12, 24]?
- c. At what time is the object's acceleration greatest?
- d. Suppose that the velocity of the object is increased by a constant value *c* for all values of *t*. What value of *c* will make the object's total distance traveled on [12, 24] be 210 meters?

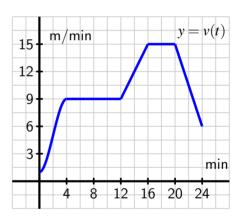


Figure 4.4.9: The velocity function of a moving body.

8. A function f is given piecewise by the formula

$$f(x) = \begin{cases} -x^2 + 2x + 1, & \text{if } 0 \le x < 2\\ -x + 3, & \text{if } 2 \le x < 3\\ x^2 - 8x + 15, & \text{if } 3 \le x \le 5 \end{cases}$$

a. Determine the exact value of the net signed area enclosed by f and the x-axis on the interval [2,5].

- b. Compute the exact average value of f on [0,5].
- c. Find a formula for a function g on $5 \le x \le 7$ so that if we extend the above definition of f so that f(x) = g(x) if $5 \le x \le 7$, it follows that $\int_0^7 f(x) \, dx = 0$.
- 9. When an aircraft attempts to climb as rapidly as possible, its climb rate (in feet per minute) decreases as altitude increases, because the air is less dense at higher altitudes. Given below is a table showing performance data for a certain single engine aircraft, giving its climb rate at various altitudes, where c(h) denotes the climb rate of the airplane at an altitude h.

h (feet)	0	1000	2000	3000	4000	5000	6000	7000	8000	9000	10,000
c (ft/min)	925	875	830	780	730	685	635	585	535	490	440

Let a new function called m(h) measure the number of minutes required for a plane at altitude h to climb the next foot of altitude.

- a. Determine a similar table of values for m(h) and explain how it is related to the table above. Be sure to explain the units.
- b. Give a careful interpretation of a function whose derivative is m(h). Describe what the input is and what the output is. Also, explain in plain English what the function tells us.
- c. Determine a definite integral whose value tells us exactly the number of minutes required for the airplane to ascend to 10,000 feet of altitude. Clearly explain why the value of this integral has the required meaning.
- d. Use the Riemann sum M_5 to estimate the value of the integral you found in (c). Include units on your result.
- **10.** In Chapter 1, we showed that for an object moving along a straight line with position function s(t), the object's "average velocity on the interval [a, b]" is given by

$$AV_{[a,b]} = \frac{s(b) - s(a)}{b - a}.$$

More recently in Chapter 4, we found that for an object moving along a straight line with velocity function v(t), the object's "average value of its velocity function on [a, b]" is

$$v_{\text{AVG}[a,b]} = \frac{1}{b-a} \int_{a}^{b} v(t) dt.$$

Are the "average velocity on the interval [a, b]" and the "average value of the velocity function on [a, b]" the same thing? Why or why not? Explain.

Evaluating Integrals

5.1 Constructing Accurate Graphs of Antiderivatives

Motivating Questions

- Given the graph of a function's derivative, how can we construct a completely accurate graph of the original function?
- How many antiderivatives does a given function have? What do those antiderivatives all have in common?
- Given a function f, how does the rule $A(x) = \int_0^x f(t) dt$ define a new function A?

A recurring theme in our discussion of differential calculus has been the question "Given information about the derivative of an unknown function f, how much information can we obtain about f itself?" For instance, in Activity 1.8.3, we explored the situation where the graph of g = f'(x) was known (along with the value of f at a single point) and endeavored to sketch a possible graph of f near the known point. In Example 3.1.4 — and indeed throughout Section 3.1 — we investigated how the first derivative test enables us to use information regarding f' to determine where the original function f is increasing and decreasing, as well as where f has relative extreme values. Further, if we know a formula or graph of f', by computing f'' we can find where the original function f is concave up and concave down. Thus, the combination of knowing f' and f'' enables us to fully understand the shape of the graph of f.

We returned to this question in even more detail in Section 4.1; there, we considered the situation where we knew the instantaneous velocity of a moving object and worked from that information to determine as much information as possible about the object's position function. We found key connections between the net-signed area under the velocity function and the corresponding change in position of the function; in Section 4.4, the Total Change Theorem further illuminated these connections between f' and f in a more general setting, such as the one found in Figure 4.4.5, showing that the total change in the value of f over an interval [a,b] is determined by the exact net-signed area bounded by f' and the x-axis on the same interval.



Chapter 5 Evaluating Integrals

In what follows, we explore these issues still further, with a particular emphasis on the situation where we possess an accurate graph of the derivative function along with a single value of the function f. From that information, we desire to completely determine an accurate graph of f that not only represents correctly where f is increasing, decreasing, concave up, and concave down, but also allows us to find an accurate function value at any point of interest to us.

Preview Activity 5.1.1. Suppose that the following information is known about a function f: the graph of its derivative, y = f'(x), is given in Figure 5.1.1. Further, assume that f' is piecewise linear (as pictured) and that for $x \le 0$ and $x \ge 6$, f'(x) = 0. Finally, it is given that f(0) = 1.

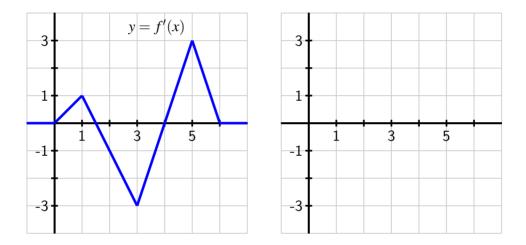


Figure 5.1.1: At left, the graph of y = f'(x); at right, axes for plotting y = f(x).

- a. On what interval(s) is f an increasing function? On what intervals is f decreasing?
- b. On what interval(s) is *f* concave up? concave down?
- c. At what point(s) does *f* have a relative minimum? a relative maximum?
- d. Recall that the Total Change Theorem tells us that

$$f(1) - f(0) = \int_0^1 f'(x) dx.$$

What is the exact value of f(1)?

- e. Use the given information and similar reasoning to that in (d) to determine the exact value of f(2), f(3), f(4), f(5), and f(6).
- f. Based on your responses to all of the preceding questions, sketch a complete and accurate graph of y = f(x) on the axes provided, being sure to indicate the behavior of f for x < 0 and x > 6.

5.1.1 Constructing the graph of an antiderivative

Preview Activity 5.1.1 demonstrates that when we can find the exact area under a given graph on any given interval, it is possible to construct an accurate graph of the given function's antiderivative: that is, we can find a representation of a function whose derivative is the given one. While we have considered this question at different points throughout our study, it is important to note here that we now can determine not only the overall shape of the antiderivative, but also the actual *height* of the antiderivative at any point of interest.

Indeed, this is one key consequence of the Fundamental Theorem of Calculus: if we know a function f and wish to know information about its antiderivative, F, provided that we have some starting point a for which we know the value of F(a), we can determine the value of F(b) via the definite integral. In particular, since $F(b) - F(a) = \int_a^b f(x) \, dx$, it follows that

$$F(b) = F(a) + \int_{a}^{b} f(x) dx.$$
 (5.1.1)

Moreover, in the discussion surrounding Figure 4.4.5, we made the observation that differences in heights of a function correspond to net-signed areas bounded by its derivative. Rephrasing this in terms of a given function f and its antiderivative F, we observe that on an interval [a,b],

differences in heights on the antiderivative (such as F(b) - F(a)) correspond to the netsigned area bounded by the original function on the interval [a,b] ($\int_a^b f(x) dx$).

For example, say that $f(x) = x^2$ and that we are interested in an antiderivative of f that satisfies F(1) = 2. Thinking of a = 1 and b = 2 in Equation (5.1.1), it follows from the Fundamental Theorem of Calculus that

$$F(2) = F(1) + \int_{1}^{2} x^{2} dx$$
$$= 2 + \frac{1}{3}x^{3}\Big|_{1}^{2}$$
$$= 2 + \left(\frac{8}{3} - \frac{1}{3}\right)$$
$$= \frac{13}{3}.$$

In this way, we see that if we are given a function f for which we can find the exact net-signed area bounded by f on a given interval, along with one value of a corresponding antiderivative F, we can find any other value of F that we seek, and in this way construct a completely

accurate graph of F. We have two main options for finding the exact net-signed area: using the Fundamental Theorem of Calculus (which requires us to find an algebraic formula for an antiderivative of the given function f), or, in the case where f has nice geometric properties, finding net-signed areas through the use of known area formulas.

Activity 5.1.2. Suppose that the function y = f(x) is given by the graph shown in Figure 5.1.2, and that the pieces of f are either portions of lines or portions of circles. In addition, let F be an antiderivative of f and say that F(0) = -1. Finally, assume that for $x \le 0$ and $x \ge 7$, f(x) = 0.

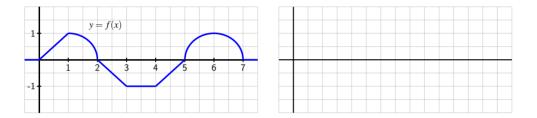


Figure 5.1.2: At left, the graph of y = f(x).

- a. On what interval(s) is *F* an increasing function? On what intervals is *F* decreasing?
- b. On what interval(s) is *F* concave up? concave down? neither?
- c. At what point(s) does *F* have a relative minimum? a relative maximum?
- d. Use the given information to determine the exact value of F(x) for x = 1, 2, ..., 7. In addition, what are the values of F(-1) and F(8)?
- e. Based on your responses to all of the preceding questions, sketch a complete and accurate graph of y = F(x) on the axes provided, being sure to indicate the behavior of F for x < 0 and x > 7. Clearly indicate the scale on the vertical and horizontal axes of your graph.
- f. What happens if we change one key piece of information: in particular, say that G is an antiderivative of f and G(0) = 0. How (if at all) would your answers to the preceding questions change? Sketch a graph of G on the same axes as the graph of F you constructed in (e).

5.1.2 Multiple antiderivatives of a single function

In the final question of Activity 5.1.2, we encountered a very important idea: a given function f has more than one antiderivative. In addition, any antiderivative of f is determined uniquely by identifying the value of the desired antiderivative at a single point. For example,

suppose that f is the function given at left in Figure 5.1.3,

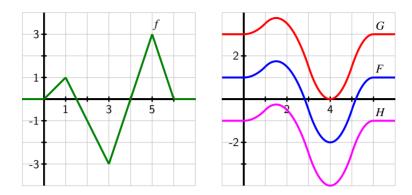


Figure 5.1.3: At left, the graph of y = f(x). At right, three different antiderivatives of f.

and we say that *F* is an antiderivative of *f* that satisfies F(0) = 1.

Then, using Equation (5.1.1), we can compute F(1) = 1.5, F(2) = 1.5, F(3) = -0.5, F(4) = -2, F(5) = -0.5, and F(6) = 1, plus we can use the fact that F' = f to ascertain where F is increasing and decreasing, concave up and concave down, and has relative extremes and inflection points. Through work similar to what we encountered in Preview Activity 5.1.1 and Activity 5.1.2, we ultimately find that the graph of F is the one given in blue in Figure 5.1.3.

If we instead chose to consider a function G that is an antiderivative of f but has the property that G(0) = 3, then G will have the exact same shape as F (since both share the derivative f), but G will be shifted vertically away from the graph of F, as pictured in red in Figure 5.1.3. Note that $G(1) - G(0) = \int_0^1 f(x) \, dx = 0.5$, just as F(1) - F(0) = 0.5,, but since G(0) = 3, G(1) = G(0) + 0.5 = 3.5, whereas F(1) = F(0) + 0.5 = 1.5, since F(0) = 1. In the same way, if we assigned a different initial value to the antiderivative, say H(0) = -1, we would get still another antiderivative, as shown in magenta in Figure 5.1.3.

This example demonstrates an important fact that holds more generally:

If G and H are both antiderivatives of a function f, then the function G-H must be constant.

To see why this result holds, observe that if G and H are both antiderivatives of f, then G' = f and H' = f. Hence,

$$\frac{d}{dx}[G(x) - H(x)] = G'(x) - H'(x) = f(x) - f(x) = 0.$$

Since the only way a function can have derivative zero is by being a constant function, it follows that the function G - H must be constant.

Further, we now see that if a function has a single antiderivative, it must have infinitely many: we can add any constant of our choice to the antiderivative and get another antiderivative.

For this reason, we sometimes refer to the *general antiderivative* of a function f. For example, if $f(x) = x^2$, its general antiderivative is $F(x) = \frac{1}{3}x^3 + C$, where we include the "+C" to indicate that F includes all of the possible antiderivatives of f. To identify a particular antiderivative of f, we must be provided a single value of the antiderivative F (this value is often called an *initial condition*). In the present example, suppose that condition is F(2) = 3; substituting the value of 2 for x in $F(x) = \frac{1}{3}x^3 + C$, we find that

$$3 = \frac{1}{3}(2)^3 + C,$$

and thus $C=3-\frac{8}{3}=\frac{1}{3}$. Therefore, the particular antiderivative in this case is $F(x)=\frac{1}{3}x^3+\frac{1}{3}$.

Activity 5.1.3. For each of the following functions, sketch an accurate graph of the antiderivative that satisfies the given initial condition. In addition, sketch the graph of two additional antiderivatives of the given function, and state the corresponding initial conditions that each of them satisfy. If possible, find an algebraic formula for the antiderivative that satisfies the initial condition.

- a. original function: g(x) = |x| 1; initial condition: G(-1) = 0; interval for sketch: [-2,2]
- b. original function: $h(x) = \sin(x)$; initial condition: H(0) = 1; interval for sketch: $[0, 4\pi]$
- c. original function: $p(x) = \begin{cases} x^2, & \text{if } 0 < x \le 1 \\ -(x-2)^2, & \text{if } 1 < x < 2; \text{ initial condition: } P(0) = 1; \\ 0 & \text{otherwise} \end{cases}$ interval for sketch: [-1,3]

5.1.3 Functions defined by integrals

In Equation (5.1.1), we found an important rule that enables us to compute the value of the antiderivative F at a point b, provided that we know F(a) and can evaluate the definite integral from a to b of f. Again, that rule is

$$F(b) = F(a) + \int_a^b f(x) \, dx.$$

In several examples, we have used this formula to compute several different values of F(b) and then plotted the points (b, F(b)) to assist us in generating an accurate graph of F. That suggests that we may want to think of b, the upper limit of integration, as a variable itself. To that end, we introduce the idea of an *integral function*, a function whose formula involves a definite integral.

Given a continuous function f, we define the corresponding integral function A according to the rule

$$A(x) = \int_{a}^{x} f(t) dt.$$
 (5.1.2)

Note particularly that because we are using the variable x as the independent variable in the function A, and x determines the other endpoint of the interval over which we integrate (starting from a), we need to use a variable other than x as the variable of integration. A standard choice is t, but any variable other than x is acceptable.

One way to think of the function A is as the "net-signed area from a up to x" function, where we consider the region bounded by y = f(t) on the relevant interval. For example, in Figure 5.1.4, we see a given function f pictured at left, and its corresponding area function (choosing a = 0), $A(x) = \int_0^x f(t) dt$ shown at right.

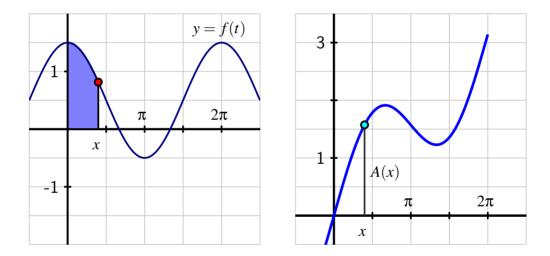


Figure 5.1.4: At left, the graph of the given function f. At right, the area function $A(x) = \int_0^x f(t) dt$.

Note particularly that the function A measures the net-signed area from t=0 to t=x bounded by the curve y=f(t); this value is then reported as the corresponding height on the graph of y=A(x). It is even more natural to think of this relationship between f and A dynamically. At http://gvsu.edu/s/cz, we find a java applet that brings the static picture in Figure 5.1.4 to life. There, the user can move the red point on the function f and see how the corresponding height changes at the light blue point on the graph of A.

The choice of a is somewhat arbitrary. In the activity that follows, we explore how the value of a affects the graph of the integral function, as well as some additional related is-

¹David Austin, Grand Valley State University

sues.

Activity 5.1.4. Suppose that g is given by the graph at left in Figure 5.1.5 and that A is the corresponding integral function defined by $A(x) = \int_1^x g(t) dt$.

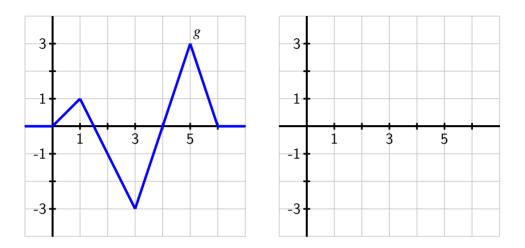


Figure 5.1.5: At left, the graph of y = g(t); at right, axes for plotting y = A(x), where A is defined by the formula $A(x) = \int_1^x g(t) dt$.

- a. On what interval(s) is *A* an increasing function? On what intervals is *A* decreasing? Why?
- b. On what interval(s) do you think *A* is concave up? concave down? Why?
- c. At what point(s) does *A* have a relative minimum? a relative maximum?
- d. Use the given information to determine the exact values of A(0), A(1), A(2), A(3), A(4), A(5), and A(6).
- e. Based on your responses to all of the preceding questions, sketch a complete and accurate graph of y = A(x) on the axes provided, being sure to indicate the behavior of A for x < 0 and x > 6.
- f. How does the graph of *B* compare to *A* if *B* is instead defined by $B(x) = \int_0^x g(t) dt$?

Summary

• Given the graph of a function f, we can construct the graph of its antiderivative F provided that (a) we know a starting value of F, say F(a), and (b) we can evaluate the

integral $\int_a^b f(x) dx$ exactly for relevant choices of a and b. For instance, if we wish to know F(3), we can compute $F(3) = F(a) + \int_a^3 f(x) dx$. When we combine this information about the function values of F together with our understanding of how the behavior of F' = f affects the overall shape of F, we can develop a completely accurate graph of the antiderivative F.

- Because the derivative of a constant is zero, if F is an antiderivative of f, it follows that G(x) = F(x) + C will also be an antiderivative of f. Moreover, any two antiderivatives of a function f differ precisely by a constant. Thus, any function with at least one antiderivative in fact has infinitely many, and the graphs of any two antiderivatives will differ only by a vertical translation.
- Given a function f, the rule $A(x) = \int_a^x f(t) dt$ defines a new function A that measures the net-signed area bounded by f on the interval [a, x]. We call the function A the integral function corresponding to f.

Exercises

1. Use the graph of f(x) shown below to find the following integrals.



$$A. \int_{-\pi}^{0} f(x)dx = \boxed{}$$

B. If the vertical red shaded area in the graph has area A, estimate: $\int_{-5}^{7} f(x)dx =$ (Your estimate may be written in terms of A.)

2. Consider the graph of the function f(x) shown below.





(Click on the graph for a larger version)

A. Estimate the integral

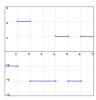
$$\int_0^7 f(x)dx \approx \boxed{}$$

B. If F is an antiderivative of the same function f and F(0) = 40, estimate F(7):

3. Assume f' is given by the graph below. Suppose f is continuous and that f(5) = 0.



Chapter 5 Evaluating Integrals



Sketch, on a sheet of work paper, an accurate graph of f, and use it to find each of

f(0) = 0	
and	
f(7) =	

Then find the value of the integral:

$$\int_0^7 f'(x) dx =$$
(Note that you can do this in two different ways!)



4. The figure below shows f.



If F' = f and F(0) = 0, find F(b) for b = 1, 2, 3, 4, 5, 6, and fill these values in the following table.

b	1	2	3	4	5	6
F(b)						

5. A moving particle has its velocity given by the quadratic function v pictured in Figure 5.1.6. In addition, it is given that $A_1 = \frac{7}{6}$ and $A_2 = \frac{8}{3}$, as well as that for the corresponding position function s, s(0) = 0.5.

- a. Use the given information to determine s(1), s(3), s(5), and s(6).
- b. On what interval(s) is *s* increasing? On what interval(s) is *s* decreasing?
- c. On what interval(s) is *s* concave up? On what interval(s) is *s* concave down?
- d. Sketch an accurate, labeled graph of *s* on the axes at right in Figure 5.1.6.
- e. Note that $v(t) = -2 + \frac{1}{2}(t-3)^2$. Find a formula for s.

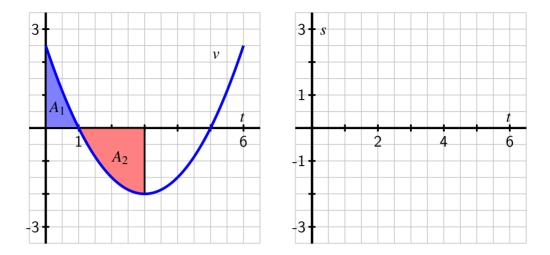


Figure 5.1.6: At left, the given graph of *v*. At right, axes for plotting *s*.

6. A person exercising on a treadmill experiences different levels of resistance and thus burns calories at different rates, depending on the treadmill's setting. In a particular workout, the rate at which a person is burning calories is given by the piecewise constant function c pictured in Figure 5.1.7. Note that the units on c are "calories per minute."

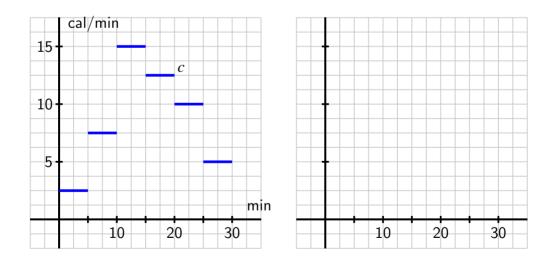


Figure 5.1.7: At left, the given graph of *c*. At right, axes for plotting *C*.

a. Let C be an antiderivative of c. What does the function C measure? What are its units?

- b. Assume that C(0) = 0. Determine the exact value of C(t) at the values t = 5, 10, 15, 20, 25, 30.
- c. Sketch an accurate graph of *C* on the axes provided at right in Figure 5.1.7. Be certain to label the scale on the vertical axis.
- d. Determine a formula for C that does not involve an integral and is valid for $5 \le t \le 10$.
- 7. Consider the piecewise linear function f given in Figure 5.1.8. Let the functions A, B, and C be defined by the rules $A(x) = \int_{-1}^{x} f(t) dt$, $B(x) = \int_{0}^{x} f(t) dt$, and $C(x) = \int_{1}^{x} f(t) dt$.
 - a. For the values x = -1, 0, 1, ..., 6, make a table that lists corresponding values of A(x), B(x), and C(x).
 - b. On the axes provided in Figure 5.1.8, sketch the graphs of *A*, *B*, and *C*.
 - c. How are the graphs of *A*, *B*, and *C* related?
 - d. How would you best describe the relationship between the function *A* and the function *f*?

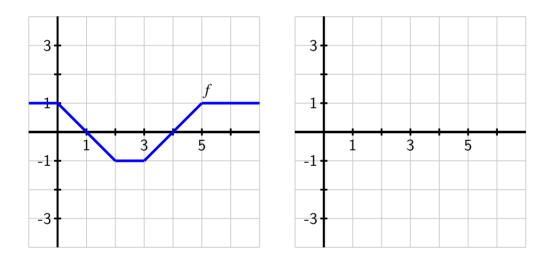


Figure 5.1.8: At left, the given graph of *f*. At right, axes for plotting *A*, *B*, and *C*.

5.2 The Second Fundamental Theorem of Calculus

Motivating Questions

- How does the integral function $A(x) = \int_1^x f(t) dt$ define an antiderivative of f?
- What is the statement of the Second Fundamental Theorem of Calculus?
- How do the First and Second Fundamental Theorems of Calculus enable us to formally see how differentiation and integration are almost inverse processes?

In Section 4.4, we learned the Fundamental Theorem of Calculus (FTC), which from here forward will be referred to as the *First* Fundamental Theorem of Calculus, as in this section we develop a corresponding result that follows it. In particular, recall that the First FTC tells us that if f is a continuous function on [a, b] and F is any antiderivative of f (that is, F' = f), then

$$\int_a^b f(x) \, dx = F(b) - F(a).$$

We have typically used this result in two settings: (1) where f is a function whose graph we know and for which we can compute the exact area bounded by f on a certain interval [a,b], we can compute the change in an antiderivative F over the interval; and (2) where f is a function for which it is easy to determine an algebraic formula for an antiderivative, we may evaluate the integral exactly and hence determine the net-signed area bounded by the function on the interval. For the former, see Preview Activity 5.1.1 or Activity 5.1.2. For the latter, we can easily evaluate exactly integrals such as

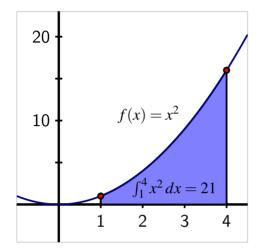
$$\int_1^4 x^2 dx,$$

since we know that the function $F(x) = \frac{1}{3}x^3$ is an antiderivative of $f(x) = x^2$. Thus,

$$\int_{1}^{4} x^{2} dx = \frac{1}{3} x^{3} \Big|_{1}^{4}$$
$$= \frac{1}{3} (4)^{3} - \frac{1}{3} (1)^{3}$$
$$= 21.$$

Here we see that the First FTC can be viewed from at least two perspectives: first, as a tool to find the difference F(b)-F(a) for an antiderivative F of the integrand f. In this situation, we need to be able to determine the value of the integral $\int_a^b f(x) dx$ exactly, perhaps through known geometric formulas for area. It is possible that we may not have a formula for F itself. From a second perspective, the First FTC provides a way to find the exact value of a definite integral, and hence a certain net-signed area exactly, by finding an antiderivative of

the integrand and evaluating its total change over the interval. In this latter case, we need to know a formula for the antiderivative F, as this enables us to compute net-signed areas exactly through definite integrals, as demonstrated in Figure 5.2.1.



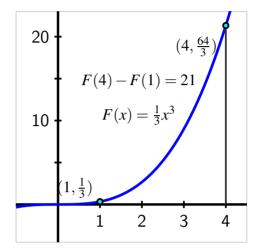


Figure 5.2.1: At left, the graph of $f(x) = x^2$ on the interval [1, 4] and the area it bounds. At right, the antiderivative function $F(x) = \frac{1}{3}x^3$, whose total change on [1, 4] is the value of the definite integral at left.

We recall further that the value of a definite integral may have additional meaning depending on context: change in position when the integrand is a velocity function, total pollutant leaked from a tank when the integrand is the rate at which pollution is leaking, or other total changes that correspond to a given rate function that is the integrand. In addition, the value of the definite integral is always connected to the average value of a continuous function on a given interval: $f_{\text{AVG}[a,b]} = \frac{1}{b-a} \int_a^b f(x) \, dx$.

Next, remember that in the last part of Section 5.1, we studied integral functions of the form $A(x) = \int_c^x f(t) dt$. Figure 5.1.4 is a particularly important image to keep in mind as we work with integral functions, and the corresponding java applet at gvsu.edu/s/cz is likewise foundational to our understanding of the function A. In what follows, we use the First FTC to gain additional understanding of the function $A(x) = \int_c^x f(t) dt$, where the integrand f is given (either through a graph or a formula), and c is a constant. In particular, we investigate further the special nature of the relationship between the functions A and f.

Preview Activity 5.2.1. Consider the function *A* defined by the rule

$$A(x) = \int_1^x f(t) \, dt,$$

where f(t) = 4 - 2t.

- a. Compute A(1) and A(2) exactly.
- b. Use the First Fundamental Theorem of Calculus to find a formula for A(x) that does not involve integrals. That is, use the first FTC to evaluate $\int_{1}^{x} (4-2t) dt$.
- c. Observe that *f* is a linear function; what kind of function is *A*?
- d. Using the formula you found in (b) that does not involve integrals, compute A'(x).
- e. While we have defined f by the rule f(t) = 4 2t, it is equivalent to say that f is given by the rule f(x) = 4 2x. What do you observe about the relationship between A and f?

5.2.1 The Second Fundamental Theorem of Calculus

The result of Preview Activity 5.2.1 is not particular to the function f(t) = 4 - 2t, nor to the choice of "1" as the lower bound in the integral that defines the function A. For instance, if we let $f(t) = \cos(t) - t$ and set $A(x) = \int_2^x f(t) \, dt$, then we can determine a formula for A without integrals by the First FTC. Specifically,

$$A(x) = \int_{2}^{x} (\cos(t) - t) dt$$

$$= \sin(t) - \frac{1}{2}t^{2}\Big|_{2}^{x}$$

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

Differentiating A(x), since $(\sin(2) - 2)$ is constant, it follows that

$$A'(x) = \cos(x) - x$$
.

and thus we see that A'(x) = f(x). This tells us that for this particular choice of f, A is an antiderivative of f. More specifically, since $A(2) = \int_2^2 f(t) \, dt = 0$, A is the only antiderivative of f for which A(2) = 0.

In general, if *f* is any continuous function, and we define the function *A* by the rule

$$A(x) = \int_{c}^{x} f(t) dt,$$

where c is an arbitrary constant, then we can show that A is an antiderivative of f. To see why, let's demonstrate that A'(x) = f(x) by using the limit definition of the derivative. Doing so, we observe that

$$A'(x) = \lim_{h \to 0} \frac{A(x+h) - A(x)}{h}$$
$$= \lim_{h \to 0} \frac{\int_{c}^{x+h} f(t) dt - \int_{c}^{x} f(t) dt}{h}$$

$$= \lim_{h \to 0} \frac{\int_{x}^{x+h} f(t) dt}{h},$$
 (5.2.1)

where Equation (5.2.1) in the preceding chain follows from the fact that $\int_c^x f(t) dt + \int_x^{x+h} f(t) dt = \int_c^{x+h} f(t) dt$. Now, observe that for small values of h,

$$\int_{x}^{x+h} f(t) dt \approx f(x) \cdot h,$$

by a simple left-hand approximation of the integral. Thus, as we take the limit in Equation (5.2.1), it follows that

$$A'(x) = \lim_{h \to 0} \frac{\int_{x}^{x+h} f(t) dt}{h} = \lim_{h \to 0} \frac{f(x) \cdot h}{h} = f(x).$$

Hence, A is indeed an antiderivative of f. In addition, $A(c) = \int_{c}^{c} f(t) dt = 0$. The preceding argument demonstrates the truth of the Second Fundamental Theorem of Calculus, which we state as follows.

The Second Fundamental Theorem of Calculus

If f is a continuous function and c is any constant, then f has a unique antiderivative A that satisfies A(c) = 0, and that antiderivative is given by the rule $A(x) = \int_{c}^{x} f(t) dt$.

Activity 5.2.2. Suppose that f is the function given in Figure 5.2.2 and that f is a piecewise function whose parts are either portions of lines or portions of circles, as pictured.

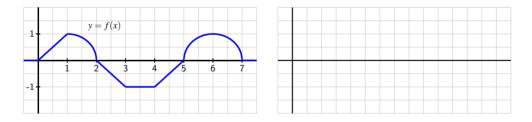


Figure 5.2.2: At left, the graph of y = f(x). At right, axes for sketching y = A(x).

In addition, let *A* be the function defined by the rule $A(x) = \int_2^x f(t) dt$.

- a. What does the Second FTC tell us about the relationship between A and f?
- b. Compute A(1) and A(3) exactly.
- c. Sketch a precise graph of y = A(x) on the axes at right that accurately reflects where A is increasing and decreasing, where A is concave up and concave down, and the exact values of A at x = 0, 1, ..., 7.

- d. How is *A* similar to, but different from, the function *F* that you found in Activity 5.1.2?
- e. With as little additional work as possible, sketch precise graphs of the functions $B(x) = \int_3^x f(t) dt$ and $C(x) = \int_1^x f(t) dt$. Justify your results with at least one sentence of explanation.

5.2.2 Understanding Integral Functions

The Second FTC provides us with a means to construct an antiderivative of any continuous function. In particular, if we are given a continuous function g and wish to find an antiderivative of G, we can now say that

$$G(x) = \int_{c}^{x} g(t) dt$$

provides the rule for such an antiderivative, and moreover that G(c) = 0. Note especially that we know that G'(x) = g(x). We sometimes want to write this relationship between G and g from a different notational perspective. In particular, observe that

$$\frac{d}{dx} \left[\int_{c}^{x} g(t) dt \right] = g(x). \tag{5.2.2}$$

This result can be particularly useful when we're given an integral function such as G and wish to understand properties of its graph by recognizing that G'(x) = g(x), while not necessarily being able to exactly evaluate the definite integral $\int_c^x g(t) dt$. To see how this is the case, we consider the following example.

Example 5.2.3. Investigate the behavior of the integral function

$$E(x) = \int_0^x e^{-t^2} dt.$$

Solution. E is closely related to the well known $error function^1$, a function that is particularly important in probability and statistics. It turns out that the function e^{-t^2} does not have an elementary antiderivative that we can express without integrals. That is, whereas a function such as f(t) = 4 - 2t has elementary antiderivative $F(t) = 4t - t^2$, we are unable to find a simple formula for an antiderivative of e^{-t^2} that does not involve a definite integral. We will learn more about finding (complicated) algebraic formulas for antiderivatives without definite integrals in the chapter on infinite series.

Returning our attention to the function E, while we cannot evaluate E exactly for any value other than x = 0, we still can gain a tremendous amount of information about the function

¹The error function is defined by the rule $\operatorname{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} dt$ and has the key property that $0 \le \operatorname{erf}(x) < 1$ for all $x \ge 0$ and moreover that $\lim_{x \to \infty} \operatorname{erf}(x) = 1$.

E. To begin, applying the rule in Equation (5.2.2) to E, it follows that

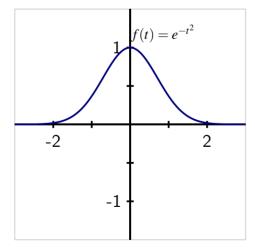
$$E'(x) = \frac{d}{dx} \left[\int_0^x e^{-t^2} dt \right] = e^{-x^2},$$

so we know a formula for the derivative of E. Moreover, we know that E(0) = 0. This information is precisely the type we were given in problems such as the one in Activity 3.1.2 and others in Section 3.1, where we were given information about the derivative of a function, but lacked a formula for the function itself.

Here, using the first and second derivatives of E, along with the fact that E(0) = 0, we can determine more information about the behavior of E. First, with $E'(x) = e^{-x^2}$, we note that for all real numbers x, $e^{-x^2} > 0$, and thus E'(x) > 0 for all x. Thus E is an always increasing function. Further, we note that as $x \to \infty$, $E'(x) = e^{-x^2} \to 0$, hence the slope of the function E tends to zero as $x \to \infty$ (and similarly as $x \to -\infty$). Indeed, it turns out (due to some more sophisticated analysis) that E has horizontal asymptotes as x increases or decreases without bound.

In addition, we can observe that $E''(x) = -2xe^{-x^2}$, and that E''(0) = 0, while E''(x) < 0 for x > 0 and E''(x) > 0 for x < 0. This information tells us that E is concave up for x < 0 and concave down for x > 0 with a point of inflection at x = 0.

The only thing we lack at this point is a sense of how big E can get as x increases. If we use a midpoint Riemann sum with 10 subintervals to estimate E(2), we see that $E(2) \approx 0.8822$; a similar calculation to estimate E(3) shows little change ($E(3) \approx 0.8862$), so it appears that as x increases without bound, E approaches a value just larger than 0.886, which aligns with the fact that E has horizontal asymptotes. Putting all of this information together (and using the symmetry of $E(1) = e^{-t^2}$), we see the results shown in Figure 5.2.4.



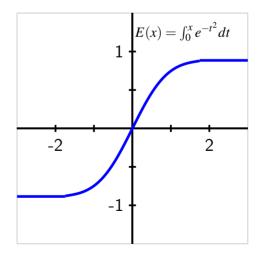


Figure 5.2.4: At left, the graph of $f(t) = e^{-t^2}$. At right, the integral function $E(x) = \int_0^x e^{-t^2} dt$, which is the unique antiderivative of f that satisfies E(0) = 0.

Again, E is the antiderivative of $f(t) = e^{-t^2}$ that satisfies E(0) = 0. Moreover, the values on the graph of y = E(x) represent the net-signed area of the region bounded by $f(t) = e^{-t^2}$ from 0 up to x. We see that the value of E increases rapidly near zero but then levels off as x increases since there is less and less additional accumulated area bounded by $f(t) = e^{-t^2}$ as x increases.

Activity 5.2.3. Suppose that $f(t) = \frac{t}{1+t^2}$ and $F(x) = \int_0^x f(t) dt$.

- a. On the axes at left in Figure 5.2.5, plot a graph of $f(t) = \frac{t}{1+t^2}$ on the interval $-10 \le t \le 10$. Clearly label the vertical axes with appropriate scale.
- b. What is the key relationship between F and f, according to the Second FTC?
- c. Use the first derivative test to determine the intervals on which *F* is increasing and decreasing.
- d. Use the second derivative test to determine the intervals on which F is concave up and concave down. Note that f'(t) can be simplified to be written in the form $f'(t) = \frac{1-t^2}{(1+t^2)^2}$.
- e. Using technology appropriately, estimate the values of F(5) and F(10) through appropriate Riemann sums.
- f. Sketch an accurate graph of y = F(x) on the righthand axes provided, and clearly label the vertical axes with appropriate scale.

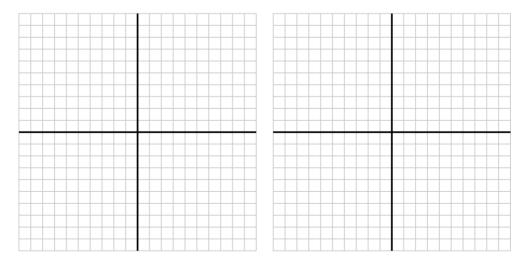


Figure 5.2.5: Axes for plotting f and F.

5.2.3 Differentiating an Integral Function

We have seen that the Second FTC enables us to construct an antiderivative F of any continuous function f by defining F by the corresponding integral function $F(x) = \int_c^x f(t) \, dt$. Said differently, if we have a function of the form $F(x) = \int_c^x f(t) \, dt$, then we know that $F'(x) = \frac{d}{dx} \left[\int_c^x f(t) \, dt \right] = f(x)$. This shows that integral functions, while perhaps having the most complicated formulas of any functions we have encountered, are nonetheless particularly simple to differentiate. For instance, if

$$F(x) = \int_{\pi}^{x} \sin(t^2) \, dt,$$

then by the Second FTC, we know immediately that

$$F'(x) = \sin(x^2).$$

Stating this result more generally for an arbitrary function f, we know by the Second FTC that

$$\frac{d}{dx}\left[\int_{a}^{x}f(t)\,dt\right]=f(x).$$

In words, the last equation essentially says that "the derivative of the integral function whose integrand is f, is f." In this sense, we see that if we first integrate the function f from t = a to t = x, and then differentiate with respect to x, these two processes "undo" one another.

Taking a different approach, say we begin with a function f(t) and differentiate with respect to t. What happens if we follow this by integrating the result from t = a to t = x? That is, what can we say about the quantity

$$\int_{a}^{x} \frac{d}{dt} [f(t)] dt?$$

Here, we use the First FTC and note that f(t) is an antiderivative of $\frac{d}{dt}[f(t)]$. Applying this result and evaluating the antiderivative function, we see that

$$\int_{a}^{x} \frac{d}{dt} [f(t)] dt = f(t) \Big|_{a}^{x}$$
$$= f(x) - f(a).$$

Thus, we see that if we apply the processes of first differentiating f and then integrating the result from a to x, we return to the function f, minus the constant value f(a). So in this situation, the two processes almost undo one another, up to the constant f(a).

The observations made in the preceding two paragraphs demonstrate that differentiating and integrating (where we integrate from a constant up to a variable) are almost inverse processes. In one sense, this should not be surprising: integrating involves antidifferentiating, which reverses the process of differentiating. On the other hand, we see that there is

some subtlety involved, as integrating the derivative of a function does not quite produce the function itself. This is connected to a key fact we observed in Section 5.1, which is that any function has an entire family of antiderivatives, and any two of those antiderivatives differ only by a constant.

Activity 5.2.4. Evaluate each of the following derivatives and definite integrals. Clearly cite whether you use the First or Second FTC in so doing.

a.
$$\frac{d}{dx} \left[\int_4^x e^{t^2} dt \right]$$

b.
$$\int_{-2}^{x} \frac{d}{dt} \left[\frac{t^4}{1+t^4} \right] dt$$

c.
$$\frac{d}{dx} \left[\int_x^1 \cos(t^3) dt \right]$$

d.
$$\int_{3}^{x} \frac{d}{dt} [\ln(1+t^2)] dt$$

a.
$$\frac{d}{dx} \left[\int_{4}^{x} e^{t^{2}} dt \right]$$
b.
$$\int_{-2}^{x} \frac{d}{dt} \left[\frac{t^{4}}{1+t^{4}} \right] dt$$
c.
$$\frac{d}{dx} \left[\int_{x}^{1} \cos(t^{3}) dt \right]$$
d.
$$\int_{3}^{x} \frac{d}{dt} \left[\ln(1+t^{2}) \right] dt$$
e.
$$\frac{d}{dx} \left[\int_{4}^{x^{3}} \sin(t^{2}) dt \right].$$

Summary

- For a continuous function f, the integral function $A(x) = \int_1^x f(t) dt$ defines an antiderivative of f.
- The Second Fundamental Theorem of Calculus is the formal, more general statement of the preceding fact: if f is a continuous function and c is any constant, then A(x) = a $\int_{c}^{x} f(t) dt$ is the unique antiderivative of f that satisfies A(c) = 0.
- Together, the First and Second FTC enable us to formally see how differentiation and integration are almost inverse processes through the observations that

$$\int_{c}^{x} \frac{d}{dt} [f(t)] dt = f(x) - f(c)$$

and

$$\frac{d}{dx}\left[\int_{c}^{x}f(t)\,dt\right]=f(x).$$

Exercises

1. Let $g(x) = \int_2^x f(t) dt$, where f(t) is given in the figure below.





Find each of the following:

$$A. g(2) =$$
 $B. g'(3) =$

C. The interval (with endpoints given to the nearest 0.25) where g is concave up:

(Give your answer as an interval or a list of intervals, e.g., (-infinity,8] or (1,5),(7,10), or enter nonefor no intervals.)

D. The value of *x* where *g* takes its maximum on the interval $0 \le x \le 8$.

$$x =$$



2. Find the derivative: $\frac{d}{dx} \int_{x}^{a} \cos(\tan(t)) dt =$



3. Find a good numerical approximation to F(9) for the function with the properties that $F'(x) = e^{-x^2/5}$ and F(0) = 2.

4. Let g be the function pictured at left in Figure 5.2.6, and let F be defined by $F(x) = \int_2^x g(t) dt$. Assume that the shaded areas have values $A_1 = 4.29$, $A_2 = 12.75$, $A_3 = 0.36$, and $A_4 = 1.79$. Assume further that the portion of A_2 that lies between x = 0.5 and x = 2 is 6.06.

Sketch a carefully labeled graph of *F* on the axes provided, and include a written analysis of how you know where *F* is zero, increasing, decreasing, CCU, and CCD.

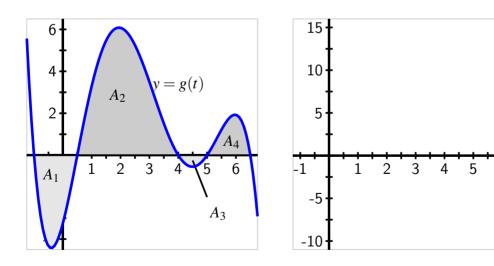


Figure 5.2.6: At left, the graph of *q*. At right, axes for plotting *F*.

5. The tide removes sand from the beach at a small ocean park at a rate modeled by the function

$$R(t) = 2 + 5\sin\left(\frac{4\pi t}{25}\right)$$

A pumping station adds sand to the beach at rate modeled by the function

$$S(t) = \frac{15t}{1+3t}$$

Both R(t) and S(t) are measured in cubic yards of sand per hour, t is measured in hours, and the valid times are $0 \le t \le 6$. At time t = 0, the beach holds 2500 cubic yards of sand.

- a. What definite integral measures how much sand the tide will remove during the time period $0 \le t \le 6$? Why?
- b. Write an expression for Y(x), the total number of cubic yards of sand on the beach at time x. Carefully explain your thinking and reasoning.
- c. At what instantaneous rate is the total number of cubic yards of sand on the beach at time t = 4 changing?
- d. Over the time interval $0 \le t \le 6$, at what time t is the amount of sand on the beach least? What is this minimum value? Explain and justify your answers fully.
- **6.** When an aircraft attempts to climb as rapidly as possible, its climb rate (in feet per minute) decreases as altitude increases, because the air is less dense at higher altitudes. Given below is a table showing performance data for a certain single engine aircraft, giving its climb rate at various altitudes, where c(h) denotes the climb rate of the airplane at an altitude h.

h (feet)	0	1000	2000	3000	4000	5000	6000	7000	8000	9000	10,000
c (ft/min)	925	875	830	780	730	685	635	585	535	490	440

Table 5.2.7: Data for the climbing aircraft.

Let a new function m, that also depends on h, (say y = m(h)) measure the number of minutes required for a plane at altitude h to climb the next foot of altitude.

- a. Determine a similar table of values for m(h) and explain how it is related to the table above. Be sure to discuss the units on m.
- b. Give a careful interpretation of a function whose derivative is m(h). Describe what the input is and what the output is. Also, explain in plain English what the function tells us.
- c. Determine a definite integral whose value tells us exactly the number of minutes required for the airplane to ascend to 10,000 feet of altitude. Clearly explain why the value of this integral has the required meaning.
- d. Determine a formula for a function M(h) whose value tells us the exact number of minutes required for the airplane to ascend to h feet of altitude.
- e. Estimate the values of M(6000) and M(10000) as accurately as you can. Include units on your results.

5.3 Integration by Substitution

Motivating Questions

- How can we begin to find algebraic formulas for antiderivatives of more complicated algebraic functions?
- What is an indefinite integral and how is its notation used in discussing antiderivatives?
- How does the technique of *u*-substitution work to help us evaluate certain indefinite integrals, and how does this process rely on identifying function-derivative pairs?

In Section 4.4, we learned the key role that antiderivatives play in the process of evaluating definite integrals exactly. In particular, the Fundamental Theorem of Calculus tells us that if F is any antiderivative of f, then

$$\int_a^b f(x) \, dx = F(b) - F(a).$$

Furthermore, we realized that each elementary derivative rule developed in Chapter 2 leads to a corresponding elementary antiderivative, as summarized in Table 4.4.4. Thus, if we wish to evaluate an integral such as

$$\int_0^1 \left(x^3 - \sqrt{x} + 5^x \right) dx,$$

it is straightforward to do so, since we can easily antidifferentiate $f(x) = x^3 - \sqrt{x} + 5^x$. In particular, since a function F whose derivative is f is given by $F(x) = \frac{1}{4}x^4 - \frac{2}{3}x^{3/2} + \frac{1}{\ln(5)}5^x$, the Fundamental Theorem of Calculus tells us that

$$\int_0^1 \left(x^3 - \sqrt{x} + 5^x \right) dx = \left. \frac{1}{4} x^4 - \frac{2}{3} x^{3/2} + \frac{1}{\ln(5)} 5^x \right|_0^1$$

$$= \left(\frac{1}{4} (1)^4 - \frac{2}{3} (1)^{3/2} + \frac{1}{\ln(5)} 5^1 \right) - \left(0 - 0 + \frac{1}{\ln(5)} 5^0 \right)$$

$$= -\frac{5}{12} + \frac{4}{\ln(5)}.$$

Because an algebraic formula for an antiderivative of f enables us to evaluate the definite integral $\int_a^b f(x) \, dx$ exactly, we see that we have a natural interest in being able to find such algebraic antiderivatives. Note that we emphasize *algebraic* antiderivatives, as opposed to any antiderivative, since we know by the Second Fundamental Theorem of Calculus that $G(x) = \int_a^x f(t) \, dt$ is indeed an antiderivative of the given function f, but one that still involves a definite integral. One of our main goals in this section and the one following is to develop understanding, in select circumstances, of how to "undo" the process of differentiation in order to find an algebraic antiderivative for a given function.

Preview Activity 5.3.1. In Section 2.5, we learned the Chain Rule and how it can be applied to find the derivative of a composite function. In particular, if u is a differentiable function of x, and f is a differentiable function of u(x), then

$$\frac{d}{dx}\left[f(u(x))\right] = f'(u(x)) \cdot u'(x).$$

In words, we say that the derivative of a composite function c(x) = f(u(x)), where f is considered the "outer" function and u the "inner" function, is "the derivative of the outer function, evaluated at the inner function, times the derivative of the inner function."

a. For each of the following functions, use the Chain Rule to find the function's derivative. Be sure to label each derivative by name (e.g., the derivative of g(x) should be labeled g'(x)).

i.
$$g(x) = e^{3x}$$
 iii. $p(x) = \arctan(2x)$ ii. $h(x) = \sin(5x + 1)$ iv. $q(x) = (2 - 7x)^4$ v. $r(x) = 3^{4-11x}$

b. For each of the following functions, use your work in (a) to help you determine the general antiderivative^a of the function. Label each antiderivative by name (e.g., the antiderivative of m should be called M). In addition, check your work by computing the derivative of each proposed antiderivative.

i.
$$m(x) = e^{3x}$$
 iv. $v(x) = (2 - 7x)^3$ ii. $n(x) = \cos(5x + 1)$ v. $w(x) = 3^{4-11x}$ iii. $s(x) = \frac{1}{1+4x^2}$

c. Based on your experience in parts (a) and (b), conjecture an antiderivative for each of the following functions. Test your conjectures by computing the derivative of each proposed antiderivative.

i.
$$a(x) = \cos(\pi x)$$

ii. $c(x) = xe^{x^2}$
ii. $b(x) = (4x + 7)^{11}$

[&]quot;Recall that the general antiderivative of a function includes "+C" to reflect the entire family of functions that share the same derivative.

5.3.1 Reversing the Chain Rule: First Steps

In Preview Activity 5.3.1, we saw that it is usually straightforward to antidifferentiate a function of the form

$$h(x) = f(u(x)),$$

whenever f is a familiar function whose antiderivative is known and u(x) is a linear function. For example, if we consider

$$h(x) = (5x - 3)^6$$

in this context the outer function f is $f(u) = u^6$, while the inner function is u(x) = 5x - 3. Since the antiderivative of f is $F(u) = \frac{1}{7}u^7 + C$, we see that the antiderivative of h is

$$H(x) = \frac{1}{7}(5x - 3)^7 \cdot \frac{1}{5} + C = \frac{1}{35}(5x - 3)^7 + C.$$

The inclusion of the constant $\frac{1}{5}$ is essential precisely because the derivative of the inner function is u'(x) = 5. Indeed, if we now compute H'(x), we find by the Chain Rule (and Constant Multiple Rule) that

$$H'(x) = \frac{1}{35} \cdot 7(5x - 3)^6 \cdot 5 = (5x - 3)^6 = h(x),$$

and thus H is indeed the general antiderivative of h.

Hence, in the special case where the outer function is familiar and the inner function is linear, we can antidifferentiate composite functions according to the following rule.

If h(x) = f(ax + b) and F is a known algebraic antiderivative of f, then the general antiderivative of h is given by

$$H(x) = \frac{1}{a}F(ax+b) + C.$$

When discussing antiderivatives, it is often useful to have shorthand notation that indicates the instruction to find an antiderivative. Thus, in a similar way to how the notation

$$\frac{d}{dx}\big[f(x)\big]$$

represents the derivative of f(x) with respect to x, we use the notation of the *indefinite integral*,

$$\int f(x) \, dx$$

to represent the general antiderivative of f with respect to x. For instance, returning to the earlier example with $h(x) = (5x-3)^6$ above, we can rephrase the relationship between h and its antiderivative H through the notation

$$\int (5x-3)^6 dx = \frac{1}{35}(5x-6)^7 + C.$$

When we find an antiderivative, we will often say that we evaluate an indefinite integral; said differently, the instruction to evaluate an indefinite integral means to find the general antiderivative. Just as the notation $\frac{d}{dx}[\Box]$ means "find the derivative with respect to x of \Box ," the notation $\int \Box dx$ means "find a function of x whose derivative is \Box ."

Activity 5.3.2. Evaluate each of the following indefinite integrals. Check each antiderivative that you find by differentiating.

a.
$$\int \sin(8-3x) dx$$

a.
$$\int \sin(8 - 3x) dx$$

b. $\int \sec^2(4x) dx$
c. $\int \frac{1}{11x - 9} dx$

c.
$$\int \frac{1}{11x-9} dx$$

d.
$$\int \csc(2x+1)\cot(2x+1)\,dx$$

e.
$$\int \frac{1}{\sqrt{1-16x^2}} dx$$

f.
$$\int 5^{-x} dx$$

5.3.2 Reversing the Chain Rule: *u*-substitution

Of course, a natural question arises from our recent work: what happens when the inner function is not a linear function? For example, can we find antiderivatives of such functions as

$$g(x) = xe^{x^2}$$
 and $h(x) = e^{x^2}$?

It is important to explicitly remember that differentiation and antidifferentiation are essentially inverse processes; that they are not quite inverse processes is due to the +C that arises when antidifferentiating. This close relationship enables us to take any known derivative rule and translate it to a corresponding rule for an indefinite integral. For example, since

$$\frac{d}{dx}\left[x^5\right] = 5x^4,$$

we can equivalently write

$$\int 5x^4 dx = x^5 + C.$$

Recall that the Chain Rule states that

$$\frac{d}{dx}\left[f(g(x))\right] = f'(g(x)) \cdot g'(x).$$

Restating this relationship in terms of an indefinite integral,

$$\int f'(g(x))g'(x) \, dx = f(g(x)) + C. \tag{5.3.1}$$

Hence, Equation (5.3.1) tells us that if we can take a given function and view its algebraic structure as f'(q(x))q'(x) for some appropriate choices of f and q, then we can antidifferentiate the function by reversing the Chain Rule. It is especially notable that both g(x) and g'(x) appear in the form of f'(g(x))g'(x); we will sometimes say that we seek to *identify a function-derivative pair* when trying to apply the rule in Equation (5.3.1).

In the situation where we can identify a function-derivative pair, we will introduce a new variable u to represent the function g(x). Observing that with u = g(x), it follows in Leibniz notation that $\frac{du}{dx} = g'(x)$, so that in terms of differentials¹, du = g'(x) dx. Now converting the indefinite integral of interest to a new one in terms of u, we have

$$\int f'(g(x))g'(x) dx = \int f'(u) du.$$

Provided that f' is an elementary function whose antiderivative is known, we can now easily evaluate the indefinite integral in u, and then go on to determine the desired overall antiderivative of f'(g(x))g'(x). We call this process u-substitution. To see u-substitution at work, we consider the following example.

Example 5.3.1. Evaluate the indefinite integral

$$\int x^3 \cdot \sin(7x^4 + 3) \, dx$$

and check the result by differentiating.

Solution. We can make two key algebraic observations regarding the integrand, $x^3 \cdot \sin(7x^4 + 3)$. First, $\sin(7x^4 + 3)$ is a composite function; as such, we know we'll need a more sophisticated approach to antidifferentiating. Second, x^3 is almost the derivative of $(7x^4 + 3)$; the only issue is a missing constant. Thus, x^3 and $(7x^4 + 3)$ are nearly a function-derivative pair. Furthermore, we know the antiderivative of $f(u) = \sin(u)$. The combination of these observations suggests that we can evaluate the given indefinite integral by reversing the chain rule through u-substitution.

Letting u represent the inner function of the composite function $\sin(7x^4 + 3)$, we have $u = 7x^4 + 3$, and thus $\frac{du}{dx} = 28x^3$. In differential notation, it follows that $du = 28x^3 dx$, and thus $x^3 dx = \frac{1}{28} du$. We make this last observation because the original indefinite integral may now be written

$$\int \sin(7x^4+3) \cdot x^3 \, dx,$$

and so by substituting the expressions in u for x (specifically u for $7x^4 + 3$ and $\frac{1}{28} du$ for $x^3 dx$), it follows that

$$\int \sin(7x^4 + 3) \cdot x^3 dx = \int \sin(u) \cdot \frac{1}{28} du.$$

Now we may evaluate the original integral by first evaluating the easier integral in u, followed by replacing u by the expression $7x^4 + 3$. Doing so, we find

$$\int \sin(7x^4 + 3) \cdot x^3 dx = \int \sin(u) \cdot \frac{1}{28} du$$

¹If we recall from the definition of the derivative that $\frac{du}{dx} \approx \frac{\Delta u}{\Delta x}$ and use the fact that $\frac{du}{dx} = g'(x)$, then we see that $g'(x) \approx \frac{\Delta u}{\Delta x}$. Solving for Δu , $\Delta u \approx g'(x)\Delta x$. It is this last relationship that, when expressed in "differential" notation enables us to write du = g'(x) dx in the change of variable formula.

$$= \frac{1}{28} \int \sin(u) du$$

$$= \frac{1}{28} (-\cos(u)) + C$$

$$= -\frac{1}{28} \cos(7x^4 + 3) + C.$$

To check our work, we observe by the Chain Rule that

$$\frac{d}{dx} \left[-\frac{1}{28} \cos(7x^4 + 3) \right] = -\frac{1}{28} \cdot (-1) \sin(7x^4 + 3) \cdot 28x^3 = \sin(7x^4 + 3) \cdot x^3,$$

which is indeed the original integrand.

An essential observation about our work in Example 5.3.1 is that the u-substitution only worked because the function multiplying $\sin(7x^4 + 3)$ was x^3 . If instead that function was x^2 or x^4 , the substitution process may not (and likely would not) have worked. This is one of the primary challenges of antidifferentiation: slight changes in the integrand make tremendous differences. For instance, we can use u-substitution with $u = x^2$ and du = 2x dx to find that

$$\int xe^{x^2} dx = \int e^u \cdot \frac{1}{2} du$$
$$= \frac{1}{2} \int e^u du$$
$$= \frac{1}{2} e^u + C$$
$$= \frac{1}{2} e^{x^2} + C.$$

If, however, we consider the similar indefinite integral

$$\int e^{x^2} dx,$$

the missing x to multiply e^{x^2} makes the u-substitution $u = x^2$ no longer possible. Hence, part of the lesson of u-substitution is just how specialized the process is: it only applies to situations where, up to a missing constant, the integrand that is present is the result of applying the Chain Rule to a different, related function.

Activity 5.3.3. Evaluate each of the following indefinite integrals by using these steps:

- Find two functions within the integrand that form (up to a possible missing constant) a function-derivative pair;
- Make a substitution and convert the integral to one involving u and du;
- Evaluate the new integral in *u*;
- Convert the resulting function of *u* back to a function of *x* by using your earlier substitution;

• Check your work by differentiating the function of *x*. You should come up with the integrand originally given.

a.
$$\int \frac{x^2}{5x^3+1} \, dx$$

c.
$$\int \frac{\cos(\sqrt{x})}{\sqrt{x}} dx$$

a.
$$\int \frac{x^2}{5x^3+1} dx$$

b.
$$\int e^x \sin(e^x) dx$$

5.3.3 Evaluating Definite Integrals via *u*-substitution

We have just introduced *u*-substitution as a means to evaluate indefinite integrals of functions that can be written, up to a constant multiple, in the form f(q(x))q'(x). This same technique can be used to evaluate definite integrals involving such functions, though we need to be careful with the corresponding limits of integration. Consider, for instance, the definite integral

$$\int_2^5 x e^{x^2} dx.$$

Whenever we write a definite integral, it is implicit that the limits of integration correspond to the variable of integration. To be more explicit, observe that

$$\int_2^5 x e^{x^2} \, dx = \int_{x=2}^{x=5} x e^{x^2} \, dx.$$

When we execute a *u*-substitution, we change the *variable* of integration; it is essential to note that this also changes the *limits* of integration. For instance, with the substitution $u = x^2$ and du = 2x dx, it also follows that when x = 2, $u = 2^2 = 4$, and when x = 5, $u = 5^2 = 25$. Thus, under the change of variables of *u*-substitution, we now have

$$\int_{x=2}^{x=5} xe^{x^2} dx = \int_{u=4}^{u=25} e^u \cdot \frac{1}{2} du$$
$$= \frac{1}{2} e^u \Big|_{u=4}^{u=25}$$
$$= \frac{1}{2} e^{25} - \frac{1}{2} e^4.$$

Alternatively, we could consider the related indefinite integral $\int xe^{x^2} dx$, find the antiderivative $\frac{1}{2}e^{x^2}$ through *u*-substitution, and then evaluate the original definite integral. From that perspective, we'd have

$$\int_{2}^{5} x e^{x^{2}} dx = \frac{1}{2} e^{x^{2}} \Big|_{2}^{5}$$
$$= \frac{1}{2} e^{25} - \frac{1}{2} e^{4},$$

which is, of course, the same result.

Activity 5.3.4. Evaluate each of the following definite integrals exactly through an appropriate u-substitution.

a.
$$\int_1^2 \frac{x}{1+4x^2} \, dx$$

c.
$$\int_{2/\pi}^{4/\pi} \frac{\cos(\frac{1}{x})}{x^2} dx$$

a.
$$\int_{1}^{2} \frac{x}{1+4x^{2}} dx$$

b. $\int_{0}^{1} e^{-x} (2e^{-x} + 3)^{9} dx$

Summary

- To begin to find algebraic formulas for antiderivatives of more complicated algebraic functions, we need to think carefully about how we can reverse known differentiation rules. To that end, it is essential that we understand and recall known derivatives of basic functions, as well as the standard derivative rules.
- The indefinite integral provides notation for antiderivatives. When we write " $\int f(x) dx$," we mean "the general antiderivative of f." In particular, if we have functions f and Fsuch that F' = f, the following two statements say the exact thing:

$$\frac{d}{dx}[F(x)] = f(x)$$
 and $\int f(x) dx = F(x) + C$.

That is, *f* is the derivative of *F*, and *F* is an antiderivative of *f*.

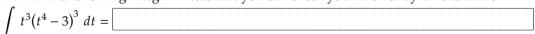
• The technique of *u*-substitution helps us to evaluate indefinite integrals of the form $\int f(q(x))g'(x) dx$ through the substitutions u = q(x) and du = g'(x) dx, so that

$$\int f(g(x))g'(x) dx = \int f(u) du.$$

A key part of choosing the expression in x to be represented by u is the identification of a function-derivative pair. To do so, we often look for an "inner" function q(x) that is part of a composite function, while investigating whether q'(x) (or a constant multiple of q'(x) is present as a multiplying factor of the integrand.

Exercises

1. Find the following integral. Note that you can check your answer by differentiation.





2. Find the general antiderivative F(x) of the function f(x) given below. Note that you can check your answer by differentiation.



$$f(x) = 2x^3 \sin(x^4)$$

antiderivative F(x) =



3. Find the following integral. Note that you can check your answer by differentiation.

$$\int \frac{\ln^7(z)}{z} dz = \boxed{$$



4. Find the following integral. Note that you can check your answer by differentiation.

$$\int \frac{e^{5x}}{5 + e^{5x}} dx = \boxed{$$



5. Find the following integral. Note that you can check your answer by differentiation.

$$\int \frac{7e^{2\sqrt{y}}}{\sqrt{y}} \, dy = \boxed{}$$



6. Use the Fundamental Theorem of Calculus to find

$$\int_{5\pi/2}^{3\pi} e^{\sin(q)} \cdot \cos(q) dq = \boxed{$$

- **7.** This problem centers on finding antiderivatives for the basic trigonometric functions other than sin(x) and cos(x).
 - a. Consider the indefinite integral $\int \tan(x) dx$. By rewriting the integrand as $\tan(x) = \frac{\sin(x)}{\cos(x)}$ and identifying an appropriate function-derivative pair, make a u-substitution and hence evaluate $\int \tan(x) dx$.
 - b. In a similar way, evaluate $\int \cot(x) dx$.
 - c. Consider the indefinite integral

$$\int \frac{\sec^2(x) + \sec(x)\tan(x)}{\sec(x) + \tan(x)} dx.$$

Evaluate this integral using the substitution $u = \sec(x) + \tan(x)$.

- d. Simplify the integrand in (c) by factoring the numerator. What is a far simpler way to write the integrand?
- e. Combine your work in (c) and (d) to determine $\int \sec(x) dx$.
- f. Using (c)-(e) as a guide, evaluate $\int \csc(x) dx$.
- **8.** Consider the indefinite integral $\int x\sqrt{x-1} dx$.
 - a. At first glance, this integrand may not seem suited to substitution due to the presence of x in separate locations in the integrand. Nonetheless, using the composite function $\sqrt{x-1}$ as a guide, let u=x-1. Determine expressions for both x and dx in terms of u.
 - b. Convert the given integral in x to a new integral in u.
 - c. Evaluate the integral in (b) by noting that $\sqrt{u} = u^{1/2}$ and observing that it is now possible to rewrite the integrand in u by expanding through multiplication.
 - d. Evaluate each of the integrals $\int x^2 \sqrt{x-1} \, dx$ and $\int x \sqrt{x^2-1} \, dx$. Write a paragraph to discuss the similarities among the three indefinite integrals in this problem and the role of substitution and algebraic rearrangement in each.

- **9.** Consider the indefinite integral $\int \sin^3(x) dx$.
 - a. Explain why the substitution $u = \sin(x)$ will not work to help evaluate the given integral.
 - b. Recall the Fundamental Trigonometric Identity, which states that $\sin^2(x) + \cos^2(x) = 1$. By observing that $\sin^3(x) = \sin(x) \cdot \sin^2(x)$, use the Fundamental Trigonometric Identity to rewrite the integrand as the product of $\sin(x)$ with another function.
 - c. Explain why the substitution u = cos(x) now provides a possible way to evaluate the integral in (b).
 - d. Use your work in (a)-(c) to evaluate the indefinite integral $\int \sin^3(x) dx$.
 - e. Use a similar approach to evaluate $\int \cos^3(x) dx$.
- **10.** For the town of Mathland, MI, residential power consumption has shown certain trends over recent years. Based on data reflecting average usage, engineers at the power company have modeled the town's rate of energy consumption by the function

$$r(t) = 4 + \sin(0.263t + 4.7) + \cos(0.526t + 9.4).$$

Here, t measures time in hours after midnight on a typical weekday, and r is the rate of consumption in megawatts² at time t. Units are critical throughout this problem.

- a. Sketch a carefully labeled graph of r(t) on the interval [0,24] and explain its meaning. Why is this a reasonable model of power consumption?
- b. Without calculating its value, explain the meaning of $\int_0^{24} r(t) dt$. Include appropriate units on your answer.
- c. Determine the exact amount of power Mathland consumes in a typical day.
- d. What is Mathland's average rate of energy consumption in a given 24-hour period? What are the units on this quantity?

²The unit *megawatt* is itself a rate, which measures energy consumption per unit time. A *megawatt-hour* is the total amount of energy that is equivalent to a constant stream of 1 megawatt of power being sustained for 1 hour.

5.4 Integration by Parts

Motivating Questions

- How do we evaluate indefinite integrals that involve products of basic functions such as $\int x \sin(x) dx$ and $\int xe^x dx$?
- What is the method of integration by parts and how can we consistently apply it to integrate products of basic functions?
- How does the algebraic structure of functions guide us in identifying u and dv in using integration by parts?

In Section 5.3, we learned the technique of u-substitution for evaluating indefinite integrals that involve certain composite functions. For example, the indefinite integral $\int x^3 \sin(x^4) dx$ is perfectly suited to u-substitution, since not only is there a composite function present, but also the inner function's derivative (up to a constant) is multiplying the composite function. Through u-substitution, we learned a general situation where recognizing the algebraic structure of a function can enable us to find its antiderivative.

It is natural to ask similar questions to those we considered in Section 5.3 about functions with a different elementary algebraic structure: those that are the product of basic functions. For instance, suppose we are interested in evaluating the indefinite integral

$$\int x \sin(x) \, dx.$$

Here, there is not a composite function present, but rather a product of the basic functions f(x) = x and $g(x) = \sin(x)$. From our work in Section 2.3 with the Product Rule, we know that it is relatively complicated to compute the derivative of the product of two functions, so we should expect that antidifferentiating a product should be similarly involved. In addition, intuitively we expect that evaluating $\int x \sin(x) dx$ will involve somehow reversing the Product Rule.

To that end, in Preview Activity 5.4.1 we refresh our understanding of the Product Rule and then investigate some indefinite integrals that involve products of basic functions.

Preview Activity 5.4.1. In Section 2.3, we developed the Product Rule and studied how it is employed to differentiate a product of two functions. In particular, recall that if f and g are differentiable functions of x, then

$$\frac{d}{dx}\left[f(x)\cdot g(x)\right] = f(x)\cdot g'(x) + g(x)\cdot f'(x).$$

a. For each of the following functions, use the Product Rule to find the function's derivative. Be sure to label each derivative by name (e.g., the derivative of g(x) should be labeled g'(x)).

i.
$$g(x) = x \sin(x)$$

iv.
$$q(x) = x^2 \cos(x)$$

ii.
$$h(x) = xe^x$$

v.
$$r(x) = e^x \sin(x)$$

iii.
$$p(x) = x \ln(x)$$

b. Use your work in (a) to help you evaluate the following indefinite integrals. Use differentiation to check your work.

i.
$$\int xe^x + e^x \, dx$$

iv.
$$\int x \cos(x) + \sin(x) dx$$

ii.
$$\int e^x (\sin(x) + \cos(x)) dx$$

v.
$$\int 1 + \ln(x) dx$$

iii.
$$\int 2x \cos(x) - x^2 \sin(x) dx$$

c. Observe that the examples in (b) work nicely because of the derivatives you were asked to calculate in (a). Each integrand in (b) is precisely the result of differentiating one of the products of basic functions found in (a). To see what happens when an integrand is still a product but not necessarily the result of differentiating an elementary product, we consider how to evaluate

$$\int x \cos(x) \, dx.$$

i. First, observe that

$$\frac{d}{dx}\left[x\sin(x)\right] = x\cos(x) + \sin(x).$$

Integrating both sides indefinitely and using the fact that the integral of a sum is the sum of the integrals, we find that

$$\int \left(\frac{d}{dx} \left[x \sin(x)\right]\right) dx = \int x \cos(x) dx + \int \sin(x) dx.$$

In this last equation, evaluate the indefinite integral on the left side as well as the rightmost indefinite integral on the right.

- ii. In the most recent equation from (i.), solve the equation for the expression $\int x \cos(x) dx$.
- iii. For which product of basic functions have you now found the antiderivative?

5.4.1 Reversing the Product Rule: Integration by Parts

Problem (c) in Preview Activity 5.4.1 provides a clue for how we develop the general technique known as Integration by Parts, which comes from reversing the Product Rule. Recall that the Product Rule states that

$$\frac{d}{dx}[f(x)g(x)] = f(x)g'(x) + g(x)f'(x).$$

Integrating both sides of this equation indefinitely with respect to x, it follows that

$$\int \frac{d}{dx} \left[f(x)g(x) \right] dx = \int f(x)g'(x) dx + \int g(x)f'(x) dx. \tag{5.4.1}$$

On the left in Equation (5.4.1), we recognize that we have the indefinite integral of the derivative of a function which, up to an additional constant, is the original function itself. Temporarily omitting the constant that may arise, we equivalently have

$$f(x)g(x) = \int f(x)g'(x) \, dx + \int g(x)f'(x) \, dx. \tag{5.4.2}$$

The most important thing to observe about Equation (5.4.2) is that it provides us with a choice of two integrals to evaluate. That is, in a situation where we can identify two functions f and g, if we can integrate f(x)g'(x), then we know the indefinite integral of g(x)f'(x), and vice versa. To that end, we choose the first indefinite integral on the left in Equation (5.4.2) and solve for it to generate the rule

$$\int f(x)g'(x) \, dx = f(x)g(x) - \int g(x)f'(x) \, dx. \tag{5.4.3}$$

Often we express Equation (5.4.3) in terms of the variables u and v, where u = f(x) and v = g(x). Note that in differential notation, du = f'(x) dx and dv = g'(x) dx, and thus we can state the rule for Integration by Parts in its most common form as follows.

$$\int u\,dv = uv - \int v\,du.$$

To apply Integration by Parts, we look for a product of basic functions that we can identify as u and dv. If we can antidifferentiate dv to find v, and evaluating $\int v \, du$ is not more difficult than evaluating $\int u \, dv$, then this substitution usually proves to be fruitful. To demonstrate, we consider the following example.

Example 5.4.1. Evaluate the indefinite integral

$$\int x \cos(x) \, dx$$

using Integration by Parts.

Solution. Whenever we are trying to integrate a product of basic functions through Integration by Parts, we are presented with a choice for u and dv. In the current problem, we can either let u = x and $dv = \cos(x) dx$, or let $u = \cos(x)$ and dv = x dx. While there is not a universal rule for how to choose u and dv, a good guideline is this: do so in a way that $\int v du$ is at least as simple as the original problem $\int u dv$.

In this setting, this leads us to choose u = x and $dv = \cos(x) dx$, from which it follows that du = 1 dx and $v = \sin(x)$. With this substitution, the rule for Integration by Parts tells us that

$$\int x \cos(x) dx = x \sin(x) - \int \sin(x) \cdot 1 dx.$$

At this point, all that remains to do is evaluate the (simpler) integral $\int \sin(x) \cdot 1 dx$. Doing so, we find

$$\int x \cos(x) \, dx = x \sin(x) - (-\cos(x)) + C = x \sin(x) + \cos(x) + C.$$

There are at least two additional important observations to make from Example 5.4.1. First, the general technique of Integration by Parts involves trading the problem of integrating the product of two functions for the problem of integrating the product of two related functions. In particular, we convert the problem of evaluating $\int u \, dv$ for that of evaluating $\int v \, du$. This perspective clearly shapes our choice of u and v. In Example 5.4.1, the original integral to evaluate was $\int x \cos(x) \, dx$, and through the substitution provided by Integration by Parts, we were instead able to evaluate $\int \sin(x) \cdot 1 \, dx$. Note that the original function x was replaced by its derivative, while $\cos(x)$ was replaced by its antiderivative. Second, observe that when we get to the final stage of evaluating the last remaining antiderivative, it is at this step that we include the integration constant, +C.

Activity 5.4.2. Evaluate each of the following indefinite integrals. Check each antiderivative that you find by differentiating.

a.
$$\int te^{-t} dt$$

c.
$$\int z \sec^2(z) dz$$

b.
$$\int 4x \sin(3x) dx$$

d.
$$\int x \ln(x) dx$$

5.4.2 Some Subtleties with Integration by Parts

There are situations where Integration by Parts is not an obvious choice, but the technique is appropriate nonetheless. One guide to understanding why is the observation that integration by parts allows us to replace one function in a product with its derivative while

¹Observe that if we considered the alternate choice, and let $u = \cos(x)$ and dv = x dx, then $du = -\sin(x) dx$ and $v = \frac{1}{2}x^2$, from which we would write $\int x \cos(x) dx = \frac{1}{2}x^2 \cos(x) - \int \frac{1}{2}x^2 (-\sin(x)) dx$. Thus we have replaced the problem of integrating $x \cos(x)$ with that of integrating $\frac{1}{2}x^2 \sin(x)$; the latter is clearly more complicated, which shows that this alternate choice is not as helpful as the first choice.

replacing the other with its antiderivative. For instance, consider the problem of evaluating

$$\int \arctan(x) \, dx.$$

Initially, this problem seems ill-suited to Integration by Parts, since there does not appear to be a product of functions present. But if we note that $\arctan(x) = \arctan(x) \cdot 1$, and realize that we know the derivative of $\arctan(x)$ as well as the antiderivative of 1, we see the possibility for the substitution $u = \arctan(x)$ and dv = 1 dx. We explore this substitution further in Activity 5.4.3.

In a related problem, if we consider $\int t^3 \sin(t^2) \, dt$, two key observations can be made about the algebraic structure of the integrand: there is a composite function present in $\sin(t^2)$, and there is not an obvious function-derivative pair, as we have t^3 present (rather than simply t) multiplying $\sin(t^2)$. This problem exemplifies the situation where we sometimes use both u-substitution and Integration by Parts in a single problem. If we write $t^3 = t \cdot t^2$ and consider the indefinite integral

$$\int t \cdot t^2 \cdot \sin(t^2) \, dt,$$

we can use a mix of the two techniques we have recently learned. First, let $z=t^2$ so that $dz=2t\ dt$, and thus $t\ dt=\frac{1}{2}\ dz$. (We are using the variable z to perform a "z-substitution" since u will be used subsequently in executing Integration by Parts.) Under this z-substitution, we now have

$$\int t \cdot t^2 \cdot \sin(t^2) \, dt = \int z \cdot \sin(z) \cdot \frac{1}{2} \, dz.$$

The remaining integral is a standard one that can be evaluated by parts. This, too, is explored further in Activity 5.4.3.

The problems briefly introduced here exemplify that we sometimes must think creatively in choosing the variables for substitution in Integration by Parts, as well as that it is entirely possible that we will need to use the technique of substitution for an additional change of variables within the process of integrating by parts.

Activity 5.4.3. Evaluate each of the following indefinite integrals, using the provided hints.

- a. Evaluate $\int \arctan(x) dx$ by using Integration by Parts with the substitution $u = \arctan(x)$ and dv = 1 dx.
- b. Evaluate $\int \ln(z) dz$. Consider a similar substitution to the one in (a).
- c. Use the substitution $z = t^2$ to transform the integral $\int t^3 \sin(t^2) dt$ to a new integral in the variable z, and evaluate that new integral by parts.
- d. Evaluate $\int s^5 e^{s^3} ds$ using an approach similar to that described in (c).
- e. Evaluate $\int e^{2t} \cos(e^t) dt$. You will find it helpful to note that $e^{2t} = e^t \cdot e^t$.

5.4.3 Using Integration by Parts Multiple Times

We have seen that the technique of Integration by Parts is well suited to integrating the product of basic functions, and that it allows us to essentially trade a given integrand for a new one where one function in the product is replaced by its derivative, while the other is replaced by its antiderivative. The main goal in this trade of $\int u \, dv$ for $\int v \, du$ is to have the new integral not be more challenging to evaluate than the original one. At times, it turns out that it can be necessary to apply Integration by Parts more than once in order to ultimately evaluate a given indefinite integral.

For example, if we consider $\int t^2 e^t dt$ and let $u = t^2$ and $dv = e^t dt$, then it follows that du = 2t dt and $v = e^t$, thus

$$\int t^2 e^t dt = t^2 e^t - \int 2t e^t dt.$$

The integral on the righthand side is simpler to evaluate than the one on the left, but it still requires Integration by Parts. Now letting u = 2t and $dv = e^t dt$, we have du = 2 dt and $v = e^t$, so that

$$\int t^2 e^t dt = t^2 e^t - \left(2te^t - \int 2e^t dt\right).$$

Note the key role of the parentheses, as it is essential to distribute the minus sign to the entire value of the integral $\int 2te^t dt$. The final integral on the right in the most recent equation is a basic one; evaluating that integral and distributing the minus sign, we find

$$\int t^2 e^t dt = t^2 e^t - 2t e^t + 2e^t + C.$$

Of course, situations are possible where even more than two applications of Integration by Parts may be necessary. For instance, in the preceding example, it is apparent that if the integrand was t^3e^t instead, we would have to use Integration by Parts three times.

Next, we consider the slightly different scenario presented by the definite integral $\int e^t \cos(t) dt$. Here, we can choose to let u be either e^t or $\cos(t)$; we pick $u = \cos(t)$, and thus $dv = e^t dt$. With $du = -\sin(t) dt$ and $v = e^t$, Integration by Parts tells us that

$$\int e^t \cos(t) dt = e^t \cos(t) - \int e^t (-\sin(t)) dt,$$

or equivalently that

$$\int e^t \cos(t) dt = e^t \cos(t) + \int e^t \sin(t) dt$$
 (5.4.4)

Observe that the integral on the right in Equation (5.4.4), $\int e^t \sin(t) dt$, while not being more complicated than the original integral we want to evaluate, it is essentially identical to $\int e^t \cos(t) dt$. While the overall situation isn't necessarily better than what we started with,

the problem hasn't gotten worse. Thus, we proceed by integrating by parts again. This time we let $u = \sin(t)$ and $dv = e^t dt$, so that $du = \cos(t) dt$ and $v = e^t$, which implies

$$\int e^t \cos(t) dt = e^t \cos(t) + \left(e^t \sin(t) - \int e^t \cos(t) dt \right)$$
 (5.4.5)

We seem to be back where we started, as two applications of Integration by Parts has led us back to the original problem, $\int e^t \cos(t) dt$. But if we look closely at Equation (5.4.5), we see that we can use algebra to solve for the value of the desired integral. In particular, adding $\int e^t \cos(t) dt$ to both sides of the equation, we have

$$2\int e^t \cos(t) dt = e^t \cos(t) + e^t \sin(t),$$

and therefore

$$\int e^t \cos(t) dt = \frac{1}{2} \left(e^t \cos(t) + e^t \sin(t) \right) + C.$$

Note that since we never actually encountered an integral we could evaluate directly, we didn't have the opportunity to add the integration constant C until the final step, at which point we include it as part of the most general antiderivative that we sought from the outset in evaluating an indefinite integral.

Activity 5.4.4. Evaluate each of the following indefinite integrals.

- a. $\int x^2 \sin(x) dx$ b. $\int t^3 \ln(t) dt$ c. $\int e^z \sin(z) dz$ d. $\int s^2 e^{3s} ds$ e. $\int t \arctan(t) dt$ (*Hint*: At a certain point in this problem, it is very helpful to note that $\frac{t^2}{1+t^2} = 1 \frac{1}{1+t^2}$.)

5.4.4 Evaluating Definite Integrals Using Integration by Parts

Just as we saw with *u*-substitution in Section 5.3, we can use the technique of Integration by Parts to evaluate a definite integral. Say, for example, we wish to find the exact value of

$$\int_0^{\pi/2} t \sin(t) dt.$$

One option is to evaluate the related indefinite integral to find that $\int t \sin(t) dt = -t \cos(t) +$ $\sin(t) + C$, and then use the resulting antiderivative along with the Fundamental Theorem of Calculus to find that

$$\int_0^{\pi/2} t \sin(t) dt = \left(-t \cos(t) + \sin(t) \right) \Big|_0^{\pi/2}$$

$$= \left(-\frac{\pi}{2} \cos(\frac{\pi}{2}) + \sin(\frac{\pi}{2}) \right) - \left(-0 \cos(0) + \sin(0) \right)$$

$$= 1.$$

Alternatively, we can apply Integration by Parts and work with definite integrals throughout. In this perspective, it is essential to remember to evaluate the product uv over the given limits of integration. To that end, using the substitution u = t and $dv = \sin(t) dt$, so that du = dt and $v = -\cos(t)$, we write

$$\int_0^{\pi/2} t \sin(t) dt = -t \cos(t) \Big|_0^{\pi/2} - \int_0^{\pi/2} (-\cos(t)) dt$$

$$= -t \cos(t) \Big|_0^{\pi/2} + \sin(t) \Big|_0^{\pi/2}$$

$$= \left(-\frac{\pi}{2} \cos(\frac{\pi}{2}) + \sin(\frac{\pi}{2}) \right) - (-0 \cos(0) + \sin(0))$$

$$= 1.$$

As with any substitution technique, it is important to remember the overall goal of the problem, to use notation carefully and completely, and to think about our end result to ensure that it makes sense in the context of the question being answered.

5.4.5 When *u*-substitution and Integration by Parts Fail to Help

As we close this section, it is important to note that both integration techniques we have discussed apply in relatively limited circumstances. In particular, it is not hard to find examples of functions for which neither technique produces an antiderivative; indeed, there are many, many functions that appear elementary but that do not have an elementary algebraic antiderivative. For instance, if we consider the indefinite integrals

$$\int e^{x^2} dx$$
 and $\int x \tan(x) dx$,

neither u-substitution nor Integration by Parts proves fruitful. While there are other integration techniques, some of which we will consider briefly, none of them enables us to find an algebraic antiderivative for e^{x^2} or $x \tan(x)$. There are at least two key observations to make: one, we do know from the Second Fundamental Theorem of Calculus that we can construct an integral antiderivative for each function; and two, antidifferentiation is much, much harder in general than differentiation. In particular, we observe that $F(x) = \int_0^x e^{t^2} dt$ is an antiderivative of $f(x) = e^{x^2}$, and $G(x) = \int_0^x t \tan(t) dt$ is an antiderivative of $g(x) = x \tan(x)$. But finding an elementary algebraic formula that doesn't involve integrals for either F or G turns out not only to be impossible through u-substitution or Integration by Parts, but indeed impossible altogether.

Summary

- Through the method of Integration by Parts, we can evaluate indefinite integrals that involve products of basic functions such as $\int x \sin(x) dx$ and $\int x \ln(x) dx$ through a substitution that enables us to effectively trade one of the functions in the product for its derivative, and the other for its antiderivative, in an effort to find a different product of functions that is easier to integrate.
- If we are given an integral whose algebraic structure we can identify as a product of basic functions in the form $\int f(x)g'(x) dx$, we can use the substitution u = f(x) and dv = q'(x) dx and apply the rule

$$\int u\,dv = uv - \int v\,du$$

in an effort to evaluate the original integral $\int f(x)g'(x) dx$ by instead evaluating $\int v du =$ $\int f'(x)g(x) dx$.

• When deciding to integrate by parts, we normally have a product of functions present in the integrand and we have to select both u and dv. That selection is guided by the overall principal that we desire the new integral $\int v \, du$ to not be any more difficult or complicated than the original integral $\int u \, dv$. In addition, it is often helpful to recognize if one of the functions present is much easier to differentiate than antidifferentiate (such as ln(x)), in which case that function often is best assigned the variable u. For sure, when choosing dv, the corresponding function must be one that we can antidifferentiate.

Exercises



1. For each of the following integrals, indicate whether integration by substitution or integration by parts is more appropriate, or if neither method is appropriate. Do not evaluate the integrals.

1.
$$\int x \sin x \, dx$$

2.
$$\int \frac{x^2}{1-x^2} dx$$

2.
$$\int \frac{x^2}{1+x^3} dx$$

3. $\int x^2 e^{x^3} dx$

4.
$$\int_{0}^{1} x^{2} \cos(x^{3}) dx$$

5. $\int_{0}^{1} \frac{1}{\sqrt{3x+1}} dx$

5.
$$\int \frac{1}{\sqrt{3x+1}} dx$$

(Note that because this is multiple choice, you will not be able to see which parts of the problem you got correct.)



2. Use integration by parts to evaluate the integral.

$$\int 3x \cos(2x) \, dx = \boxed{ + c}$$



3. Find the integral

$$\int (z+1)e^{4z}dz =$$

4. Evaluate the definite integral.

$$\int_{0}^{4} te^{-t} dt = \boxed{$$



- **5.** Let $f(t) = te^{-2t}$ and $F(x) = \int_0^x f(t) dt$.
 - a. Determine F'(x).
 - b. Use the First FTC to find a formula for *F* that does not involve an integral.
 - c. Is F an increasing or decreasing function for x > 0? Why?
- **6.** Consider the indefinite integral given by $\int e^{2x} \cos(e^x) dx$.
 - a. Noting that $e^{2x} = e^x \cdot e^x$, use the substitution $z = e^x$ to determine a new, equivalent integral in the variable z.
 - b. Evaluate the integral you found in (a) using an appropriate technique.
 - c. How is the problem of evaluating $\int e^{2x} \cos(e^{2x}) dx$ different from evaluating the integral in (a)? Do so.
 - d. Evaluate each of the following integrals as well, keeping in mind the approach(es) used earlier in this problem:
 - $\int e^{2x} \sin(e^x) dx$
 - $\int e^{3x} \sin(e^{3x}) dx$
 - $\int xe^{x^2}\cos(e^{x^2})\sin(e^{x^2})\,dx$
- 7. For each of the following indefinite integrals, determine whether you would use u-substitution, integration by parts, neither*, or both to evaluate the integral. In each case, write one sentence to explain your reasoning, and include a statement of any substitutions used. (That is, if you decide in a problem to let $u = e^{3x}$, you should state that, as well as that $du = 3e^{3x} dx$.) Finally, use your chosen approach to evaluate each integral. (* one of the following problems does not have an elementary antiderivative and you are not expected to actually evaluate this integral; this will correspond with a choice of "neither" among those given.)
 - a. $\int x^2 \cos(x^3) dx$
 - b. $\int x^5 \cos(x^3) dx$ (Hint: $x^5 = x^2 \cdot x^3$)
 - c. $\int x \ln(x^2) dx$
 - d. $\int \sin(x^4) dx$
 - e. $\int x^3 \sin(x^4) dx$
 - f. $\int x^7 \sin(x^4) dx$

5.5 Other Options for Finding Algebraic Antiderivatives

Motivating Questions

- How does the method of partial fractions enable any rational function to be antidifferentiated?
- What role have integral tables historically played in the study of calculus and how can a table be used to evaluate integrals such as $\int \sqrt{a^2 + u^2} du$?
- What role can a computer algebra system play in the process of finding antiderivatives?

In the preceding sections, we have learned two very specific antidifferentiation techniques: u-substitution and integration by parts. The former is used to reverse the chain rule, while the latter to reverse the product rule. But we have seen that each only works in very specialized circumstances. For example, while $\int xe^{x^2} dx$ may be evaluated by u-substitution and $\int xe^x dx$ by integration by parts, neither method provides a route to evaluate $\int e^{x^2} dx$. That fact is not a particular shortcoming of these two antidifferentiation techniques, as it turns out there does not exist an elementary algebraic antiderivative for e^{x^2} . Said differently, no matter what antidifferentiation methods we could develop and learn to execute, none of them will be able to provide us with a simple formula that does not involve integrals for a function F(x) that satisfies $F'(x) = e^{x^2}$.

In this section of the text, our main goals are to better understand some classes of functions that can always be antidifferentiated, as well as to learn some options for so doing. At the same time, we want to recognize that there are many functions for which an algebraic formula for an antiderivative does not exist, and also appreciate the role that computing technology can play in helping us find antiderivatives of other complicated functions. Throughout, it is helpful to remember what we have learned so far: how to reverse the chain rule through *u*-substitution, how to reverse the product rule through integration by parts, and that overall, there are subtle and challenging issues to address when trying to find antiderivatives.

Preview Activity 5.5.1. For each of the indefinite integrals below, the main question is to decide whether the integral can be evaluated using u-substitution, integration by parts, a combination of the two, or neither. For integrals for which your answer is affirmative, state the substitution(s) you would use. It is not necessary to actually evaluate any of the integrals completely, unless the integral can be evaluated immediately using a familiar basic antiderivative.

a.
$$\int x^2 \sin(x^3) dx$$
, $\int x^2 \sin(x) dx$, $\int \sin(x^3) dx$, $\int x^5 \sin(x^3) dx$

b.
$$\int \frac{1}{1+x^2} dx$$
, $\int \frac{x}{1+x^2} dx$, $\int \frac{2x+3}{1+x^2} dx$, $\int \frac{e^x}{1+(e^x)^2} dx$,

c.
$$\int x \ln(x) dx$$
, $\int \frac{\ln(x)}{x} dx$, $\int \ln(1+x^2) dx$, $\int x \ln(1+x^2) dx$,

c.
$$\int x \ln(x) dx$$
, $\int \frac{\ln(x)}{x} dx$, $\int \ln(1+x^2) dx$, $\int x \ln(1+x^2) dx$, d. $\int x\sqrt{1-x^2} dx$, $\int \frac{1}{\sqrt{1-x^2}} dx$, $\int \frac{x}{\sqrt{1-x^2}} dx$, $\int \frac{1}{x\sqrt{1-x^2}} dx$,

5.5.1 The Method of Partial Fractions

The method of partial fractions is used to integrate rational functions, and essentially involves reversing the process of finding a common denominator. For example, suppose we have the function $R(x) = \frac{5x}{x^2 - x - 2}$ and want to evaluate

$$\int \frac{5x}{x^2 - x - 2} \, dx.$$

Thinking algebraically, if we factor the denominator, we can see how R might come from the sum of two fractions of the form $\frac{A}{r-2} + \frac{B}{r+1}$. In particular, suppose that

$$\frac{5x}{(x-2)(x+1)} = \frac{A}{x-2} + \frac{B}{x+1}.$$

Multiplying both sides of this last equation by (x-2)(x+1), we find that

$$5x = A(x + 1) + B(x - 2).$$

Since we want this equation to hold for every value of x, we can use insightful choices of specific x-values to help us find A and B. Taking x = -1, we have

$$5(-1) = A(0) + B(-3),$$

and thus $B = \frac{5}{3}$. Choosing x = 2, it follows

$$5(2) = A(3) + B(0),$$

so $A = \frac{10}{3}$. Therefore, we now know that

$$\int \frac{5x}{x^2 - x - 2} \, dx = \int \frac{10/3}{x - 2} + \frac{5/3}{x + 1} \, dx.$$

This equivalent integral expression is straightforward to evaluate, and hence we find that

$$\int \frac{5x}{x^2 - x - 2} \, dx = \frac{10}{3} \ln|x - 2| + \frac{5}{3} \ln|x + 1| + C.$$

It turns out that for any rational function $R(x) = \frac{P(x)}{O(x)}$ where the degree of the polynomial P is less than the degree of the polynomial Q, the method of partial fractions can be used to rewrite the function as a sum of simpler rational functions of one of the following forms:

$$\frac{A}{x-c}$$
, $\frac{A}{(x-c)^n}$, $\frac{Ax+B}{x^2+k}$, or $\frac{Ax+B}{(x^2+k)^n}$

If the degree of P is greater than or equal to the degree of Q, long division may be used to write R as the sum of a polynomial plus a rational function where the numerator's degree is less than the denominator's.

where A, B, and c are real numbers, and k is a positive real number. Because each of these basic forms is one we can antidifferentiate, partial fractions enables us to antidifferentiate any rational function. A computer algebra system such as Maple, Mathematica, or WolframAlpha can be used to find the partial fraction decomposition of any rational function. In Wolfram Alpha, entering

partial fraction $5x/(x^2-x-2)$

results in the output

$$\frac{5x}{x^2 - x - 2} = \frac{10}{3(x - 2)} + \frac{5}{3(x + 1)}.$$

We will primarily use technology to generate partial fraction decompositions of rational functions, and then work from there to evaluate the integrals of interest using established methods.

Activity 5.5.2. For each of the following problems, evaluate the integral by using the partial fraction decomposition provided.

a.
$$\int \frac{1}{x^2 - 2x - 3} dx$$
, given that $\frac{1}{x^2 - 2x - 3} = \frac{1/4}{x - 3} - \frac{1/4}{x + 1}$

a.
$$\int \frac{1}{x^2 - 2x - 3} dx$$
, given that $\frac{1}{x^2 - 2x - 3} = \frac{1/4}{x - 3} - \frac{1/4}{x + 1}$
b. $\int \frac{x^2 + 1}{x^3 - x^2} dx$, given that $\frac{x^2 + 1}{x^3 - x^2} = -\frac{1}{x} - \frac{1}{x^2} + \frac{2}{x - 1}$
c. $\int \frac{x - 2}{x^4 + x^2} dx$, given that $\frac{x - 2}{x^4 + x^2} = \frac{1}{x} - \frac{2}{x^2} + \frac{-x + 2}{1 + x^2}$

c.
$$\int \frac{x-2}{x^4+x^2} dx$$
, given that $\frac{x-2}{x^4+x^2} = \frac{1}{x} - \frac{2}{x^2} + \frac{-x+2}{1+x^2}$

5.5.2 Using an Integral Table

Calculus has a long history, with key ideas going back as far as Greek mathematicians in 400-300 BC. Its main foundations were first investigated and understood independently by Isaac Newton and Gottfried Wilhelm Leibniz in the late 1600s, making the modern ideas of calculus well over 300 years old. It is instructive to realize that until the late 1980s, the personal computer essentially did not exist, so calculus (and other mathematics) had to be done by hand for roughly 300 years. During the last 30 years, however, computers have revolutionized many aspects of the world we live in, including mathematics. In this section we take a short historical tour to precede the following discussion of the role computer algebra systems can play in evaluating indefinite integrals. In particular, we consider a class of integrals involving certain radical expressions that, until the advent of computer algebra systems, were often evaluated using an integral table.

As seen in the short table of integrals found in Appendix A, there are also many forms of integrals that involve $\sqrt{a^2 \pm w^2}$ and $\sqrt{w^2 - a^2}$. These integral rules can be developed using a technique known as trigonometric substitution that we choose to omit; instead, we will simply accept the results presented in the table. To see how these rules are needed and used, consider the differences among

$$\int \frac{1}{\sqrt{1-x^2}} dx, \quad \int \frac{x}{\sqrt{1-x^2}} dx, \quad \text{and} \quad \int \sqrt{1-x^2} dx.$$

The first integral is a familiar basic one, and results in $\arcsin(x) + C$. The second integral can be evaluated using a standard *u*-substitution with $u = 1 - x^2$. The third, however, is not familiar and does not lend itself to *u*-substitution.

In Appendix A, we find the rule

(h)
$$\int \sqrt{a^2 - u^2} \, du = \frac{u}{2} \sqrt{a^2 - u^2} + \frac{a^2}{2} \arcsin \frac{u}{a} + C.$$

Using the substitutions a = 1 and u = x (so that du = dx), it follows that

$$\int \sqrt{1 - x^2} \, dx = \frac{x}{2} \sqrt{1 - x^2} - \frac{1}{2} \arcsin x + C.$$

One important point to note is that whenever we are applying a rule in the table, we are doing a u-substitution. This is especially key when the situation is more complicated than allowing u = x as in the last example. For instance, say we wish to evaluate the integral

$$\int \sqrt{9 + 64x^2} \, dx.$$

Here, we want to use Rule (c) from the table, and do so with a = 3 and u = 8x; we also choose the "+" option in the rule. With this substitution, it follows that du = 8dx, so $dx = \frac{1}{8}du$. Applying this substitution,

$$\int \sqrt{9 + 64x^2} \, dx = \int \sqrt{9 + u^2} \cdot \frac{1}{8} \, du = \frac{1}{8} \int \sqrt{9 + u^2} \, du.$$

By Rule (c), we now find that

$$\int \sqrt{9 + 64x^2} \, dx = \frac{1}{8} \left(\frac{u}{2} \sqrt{u^2 + 9} + \frac{9}{2} \ln|u + \sqrt{u^2 + 9}| + C \right)$$
$$= \frac{1}{8} \left(\frac{8x}{2} \sqrt{64x^2 + 9} + \frac{9}{2} \ln|8x + \sqrt{64x^2 + 9}| + C \right).$$

In problems such as this one, it is essential that we not forget to account for the factor of $\frac{1}{8}$ that must be present in the evaluation.

Activity 5.5.3. For each of the following integrals, evaluate the integral using u-substitution and/or an entry from the table found in Appendix A.

a.
$$\int \sqrt{x^2 + 4} \, dx$$

b.
$$\int \frac{x}{\sqrt{x^2 + 4}} \, dx$$

c.
$$\int \frac{2}{\sqrt{16+25x^2}} dx$$

b.
$$\int \frac{x}{\sqrt{x^2+4}} dx$$

d.
$$\int \frac{1}{x^2 \sqrt{49-36x^2}} dx$$

5.5.3 Using Computer Algebra Systems

A computer algebra system (CAS) is a computer program that is capable of executing symbolic mathematics. For a simple example, if we ask a CAS to solve the equation $ax^2 + bx + c = 0$ for the variable x, where a, b, and c are arbitrary constants, the program will return $x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$. While research to develop the first CAS dates to the 1960s, these programs became more common and publicly available in the early 1990s. Two prominent early examples are the programs Maple and Mathematica, which were among the first computer algebra systems to offer a graphical user interface. Today, Maple and Mathematica are exceptionally powerful professional software packages that are capable of executing an amazing array of sophisticated mathematical computations. They are also very expensive, as each is a proprietary program. The CAS SAGE is an open-source, free alternative to Maple and Mathematica.

For the purposes of this text, when we need to use a CAS, we are going to turn instead to a similar, but somewhat different computational tool, the web-based "computational knowledge engine" called *WolframAlpha*. There are two features of *WolframAlpha* that make it stand out from the CAS options mentioned above: (1) unlike *Maple* and *Mathematica*, *WolframAlpha* is free (provided we are willing to suffer through some pop-up advertising); and (2) unlike any of the three, the syntax in *WolframAlpha* is flexible. Think of *WolframAlpha* as being a little bit like doing a Google search: the program will interpret what is input, and then provide a summary of options.

If we want to have WolframAlpha evaluate an integral for us, we can provide it syntax such as

integrate x^2 dx

to which the program responds with

$$\int x^2 dx = \frac{x^3}{3} + \text{constant.}$$

While there is much to be enthusiastic about regarding CAS programs such as *WolframAlpha*, there are several things we should be cautious about: (1) a CAS only responds to exactly what is input; (2) a CAS can answer using powerful functions from highly advanced mathematics; and (3) there are problems that even a CAS cannot do without additional human insight.

Although (1) likely goes without saying, we have to be careful with our input: if we enter syntax that defines a function other than the problem of interest, the CAS will work with precisely the function we define. For example, if we are interested in evaluating the integral

$$\int \frac{1}{16 - 5x^2} \, dx,$$

and we mistakenly enter

integrate 1/16 - 5x^2 dx

a CAS will (correctly) reply with

$$\frac{1}{16}x - \frac{5}{3}x^3$$
.

It is essential that we are sufficiently well-versed in antidifferentiation to recognize that this function cannot be the one that we seek: integrating a rational function such as $\frac{1}{16-5x^2}$, we expect the logarithm function to be present in the result.

Regarding (2), even for a relatively simple integral such as $\int \frac{1}{16-5x^2} dx$, some CASs will invoke advanced functions rather than simple ones. For instance, if we use *Maple* to execute the command

$$int(1/(16-5*x^2), x);$$

the program responds with

$$\int \frac{1}{16 - 5x^2} dx = \frac{\sqrt{5}}{20} \operatorname{arctanh}(\frac{\sqrt{5}}{4}x).$$

While this is correct (save for the missing arbitrary constant, which *Maple* never reports), the inverse hyperbolic tangent function is not a common nor familiar one; a simpler way to express this function can be found by using the partial fractions method, and happens to be the result reported by *WolframAlpha*:

$$\int \frac{1}{16 - 5x^2} dx = \frac{1}{8\sqrt{5}} \left(\log(4\sqrt{5} + 5\sqrt{x}) - \log(4\sqrt{5} - 5\sqrt{x}) \right) + \text{constant.}$$

Using sophisticated functions from more advanced mathematics is sometimes the way a CAS says to the user "I don't know how to do this problem." For example, if we want to evaluate

$$\int e^{-x^2} dx,$$

and we ask WolframAlpha to do so, the input

integrate $exp(-x^2)dx$

results in the output

$$\int e^{-x^2} dx = \frac{\sqrt{\pi}}{2} \operatorname{erf}(x) + \operatorname{constant.}$$

The function " $\operatorname{erf}(x)$ " is the *error function*, which is actually defined by an integral:

$$\operatorname{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} dt.$$

So, in producing output involving an integral, the CAS has basically reported back to us the very question we asked.

Finally, as remarked at (3) above, there are times that a CAS will actually fail without some additional human insight. If we consider the integral

$$\int (1+x)e^x \sqrt{1+x^2e^{2x}} \, dx$$

and ask WolframAlpha to evaluate

int
$$(1+x)* \exp(x)* \operatorname{sqrt}(1+x^2* \exp(2x)) dx$$
,

the program thinks for a moment and then reports

(no result found in terms of standard mathematical functions)

But in fact this integral is not that difficult to evaluate. If we let $u = xe^x$, then $du = (1 + x)e^x dx$, which means that the preceding integral has form

$$\int (1+x)e^x \sqrt{1+x^2e^{2x}} \, dx = \int \sqrt{1+u^2} \, du,$$

which is a straightforward one for any CAS to evaluate.

So, the above observations regarding computer algebra systems lead us to proceed with some caution: while any CAS is capable of evaluating a wide range of integrals (both definite and indefinite), there are times when the result can mislead us. We must think carefully about the meaning of the output, whether it is consistent with what we expect, and whether or not it makes sense to proceed.

Summary

- The method of partial fractions enables any rational function to be antidifferentiated, because any polynomial function can be factored into a product of linear and irreducible quadratic terms. This allows any rational function to be written as the sum of a polynomial plus rational terms of the form $\frac{A}{(x-c)^n}$ (where n is a natural number) and $\frac{Bx+C}{x^2+k}$ (where k is a positive real number).
- Until the development of computing algebra systems, integral tables enabled students of calculus to more easily evaluate integrals such as $\int \sqrt{a^2 + u^2} \, du$, where a is a positive real number. A short table of integrals may be found in Appendix A.
- Computer algebra systems can play an important role in finding antiderivatives, though
 we must be cautious to use correct input, to watch for unusual or unfamiliar advanced
 functions that the CAS may cite in its result, and to consider the possibility that a CAS
 may need further assistance or insight from us in order to answer a particular question.

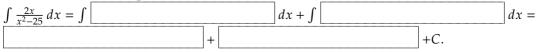
Exercises



1. Calculate the integral below by partial fractions and by using the indicated substitution. Be sure that you can show how the results you obtain are the same.

$$\int \frac{2x}{x^2 - 25} \, dx$$

First, rewrite this with partial fractions:



(Note that you should not include the +C in your entered answer, as it has been provided at the end of the expression.)

Next, use the substitution $w = x^2 - 25$ to find the integral:

$$\int \frac{2x}{x^2 - 25} \, dx = \int \boxed{ +C.}$$

(For the second answer blank, give your antiderivative in terms of the variable w. Again, note that you should not include the +C in your answer.)

2. Calculate the integral:

$$\int \frac{1}{(x+6)(x+8)} dx = \boxed{}$$

3. Calculate the integral

$$\int \frac{7x+3}{x^2 - 3x + 2} dx = \boxed{}$$

4. Consider the following indefinite integral.

$$\int \frac{6x^3 + 8x^2 + 2x + 6}{x^4 + 1x^2} \, dx$$

The integrand has partial fractions decomposition:

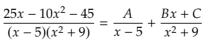
$$\frac{a}{x^2} + \frac{b}{x} + \frac{cx+d}{x^2+1}$$

where h =c =d =

Now integrate term by term to evaluate the integral.

Answer: +C

5. The form of the partial fraction decomposition of a rational function is given below.



$$A =$$
 $B =$ $C =$ Now evaluate the indefinite integral.

$$\int \frac{25x - 10x^2 - 45}{(x - 5)(x^2 + 9)} dx = \boxed{ }$$

6. For each of the following integrals involving rational functions, (1) use a CAS to find the partial fraction decomposition of the integrand; (2) evaluate the integral of the resulting

Chapter 5 Evaluating Integrals

function without the assistance of technology; (3) use a CAS to evaluate the original integral to test and compare your result in (2).

a.
$$\int \frac{x^3 + x + 1}{x^4 - 1} dx$$

b.
$$\int \frac{x^5 + x^2 + 3}{x^3 - 6x^2 + 11x - 6} dx$$

c.
$$\int \frac{x^2 - x - 1}{(x - 3)^3} dx$$

7. For each of the following integrals involving radical functions, (1) use an appropriate u-substitution along with Appendix A to evaluate the integral without the assistance of technology, and (2) use a CAS to evaluate the original integral to test and compare your result in (1).

a.
$$\int \frac{1}{x\sqrt{9x^2+25}} dx$$

b.
$$\int x\sqrt{1+x^4}\,dx$$

c.
$$\int e^x \sqrt{4 + e^{2x}} dx$$

d.
$$\int \frac{\tan(x)}{\sqrt{9-\cos^2(x)}} dx$$

8. Consider the indefinite integral given by

$$\int \frac{\sqrt{x + \sqrt{1 + x^2}}}{x} \, dx.$$

- a. Explain why u-substitution does not offer a way to simplify this integral by discussing at least two different options you might try for u.
- b. Explain why integration by parts does not seem to be a reasonable way to proceed, either, by considering one option for u and dv.
- c. Is there any line in the integral table in Appendix Athat is helpful for this integral?
- d. Evaluate the given integral using Wolfram Alpha. What do you observe?

5.6 Numerical Integration

Motivating Questions

- How do we accurately evaluate a definite integral such as $\int_0^1 e^{-x^2} dx$ when we cannot use the First Fundamental Theorem of Calculus because the integrand lacks an elementary algebraic antiderivative? Are there ways to generate accurate estimates without using extremely large values of n in Riemann sums?
- What is the Trapezoid Rule, and how is it related to left, right, and middle Riemann sums?
- How are the errors in the Trapezoid Rule and Midpoint Rule related, and how can they be used to develop an even more accurate rule?

When we were first exploring the problem of finding the net-signed area bounded by a curve, we developed the concept of a Riemann sum as a helpful estimation tool and a key step in the definition of the definite integral. In particular, as we found in Section 4.2, recall that the left, right, and middle Riemann sums of a function f on an interval [a, b] are denoted L_n , R_n , and M_n , with formulas

$$L_n = f(x_0)\Delta x + f(x_1)\Delta x + \dots + f(x_{n-1})\Delta x = \sum_{i=0}^{n-1} f(x_i)\Delta x,$$
 (5.6.1)

$$R_n = f(x_1)\Delta x + f(x_2)\Delta x + \dots + f(x_n)\Delta x = \sum_{i=1}^n f(x_i)\Delta x,$$
 (5.6.2)

$$M_n = f(\overline{x}_1)\Delta x + f(\overline{x}_2)\Delta x + \dots + f(\overline{x}_n)\Delta x = \sum_{i=1}^n f(\overline{x}_i)\Delta x,$$
 (5.6.3)

where $x_0 = a$, $x_i = a + i\Delta x$, $x_n = b$, and $\Delta x = \frac{b-a}{n}$. For the middle sum, note that $\overline{x}_i = (x_{i-1} + x_i)/2$.

Further, recall that a Riemann sum is essentially a sum of (possibly signed) areas of rectangles, and that the value of n determines the number of rectangles, while our choice of left endpoints, right endpoints, or midpoints determines how we use the given function to find the heights of the respective rectangles we choose to use. Visually, we can see the similarities and differences among these three options in Figure 5.6.1, where we consider the function $f(x) = \frac{1}{20}(x-4)^3 + 7$ on the interval [1, 8], and use 5 rectangles for each of the Riemann sums.

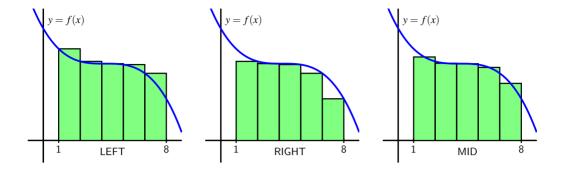


Figure 5.6.1: Left, right, and middle Riemann sums for y = f(x) on [1, 8] with 5 subintervals.

While it is a good exercise to compute a few Riemann sums by hand, just to ensure that we understand how they work and how varying the function, the number of subintervals, and the choice of endpoints or midpoints affects the result, it is of course the case that using computing technology is the best way to determine L_n , R_n , and M_n going forward. Any computer algebra system will offer this capability; as we saw in Preview Activity 4.3.1, a straightforward option that happens to also be freely available online is the applet 1 at http://gvsu.edu/s/a9.

Note that we can adjust the formula for f(x), the window of x- and y-values of interest, the number of subintervals, and the method. See Preview Activity 4.3.1 for any needed reminders on how the applet works.

In what follows in this section we explore several different alternatives, including left, right, and middle Riemann sums, for estimating definite integrals. One of our main goals in the upcoming section is to develop formulas that enable us to estimate definite integrals accurately without having to use exceptionally large numbers of rectangles.

Preview Activity 5.6.1. As we begin to investigate ways to approximate definite integrals, it will be insightful to compare results to integrals whose exact values we know. To that end, the following sequence of questions centers on $\int_0^3 x^2 dx$.

- a. Use the applet at http://gvsu.edu/s/a9 with the function $f(x) = x^2$ on the window of x values from 0 to 3 to compute L_3 , the left Riemann sum with three subintervals.
- b. Likewise, use the applet to compute R_3 and M_3 , the right and middle Riemann sums with three subintervals, respectively.
- c. Use the Fundamental Theorem of Calculus to compute the exact value of $I = \int_0^3 x^2 dx$.

¹Marc Renault, Shippensburg University

- d. We define the *error* in an approximation of a definite integral to be the difference between the integral's exact value and the approximation's value. What is the error that results from using L_3 ? From R_3 ? From R_3 ?
- e. In what follows in this section, we will learn a new approach to estimating the value of a definite integral known as the Trapezoid Rule. The basic idea is to use trapezoids, rather than rectangles, to estimate the area under a curve. What is the formula for the area of a trapezoid with bases of length b_1 and b_2 and height h?
- f. Working by hand, estimate the area under $f(x) = x^2$ on [0,3] using three subintervals and three corresponding trapezoids. What is the error in this approximation? How does it compare to the errors you calculated in (d)?

5.6.1 The Trapezoid Rule

Throughout our work to date with developing and estimating definite integrals, we have used the simplest possible quadrilaterals (that is, rectangles) to subdivide regions with complicated shapes. It is natural, however, to wonder if other familiar shapes might serve us even better. In particular, our goal is to be able to accurately estimate $\int_a^b f(x) dx$ without having to use extremely large values of n in Riemann sums.

To this end, we consider an alternative to L_n , R_n , and M_n , know as the *Trapezoid Rule*. The fundamental idea is simple: rather than using a rectangle to estimate the (signed) area bounded by y = f(x) on a small interval, we use a trapezoid. For example, in Figure 5.6.2, we estimate the area under the pictured curve using three subintervals and the trapezoids that result from connecting the corresponding points on the curve with straight lines.

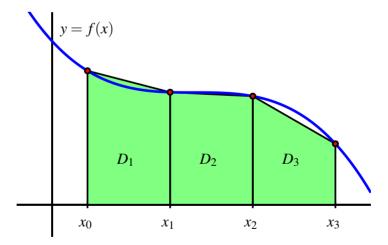


Figure 5.6.2: Estimating $\int_a^b f(x) dx$ using three subintervals and trapezoids, rather than rectangles, where $a = x_0$ and $b = x_3$.

The biggest difference between the Trapezoid Rule and a left, right, or middle Riemann sum is that on each subinterval, the Trapezoid Rule uses two function values, rather than one, to estimate the (signed) area bounded by the curve. For instance, to compute D_1 , the area of the trapezoid generated by the curve y = f(x) in Figure 5.6.2 on $[x_0, x_1]$, we observe that the left base of this trapezoid has length $f(x_0)$, while the right base has length $f(x_1)$. In addition, the height of this trapezoid is $x_1 - x_0 = \Delta x = \frac{b-a}{3}$. Since the area of a trapezoid is the average of the bases times the height, we have

$$D_1 = \frac{1}{2}(f(x_0) + f(x_1)) \cdot \Delta x.$$

Using similar computations for D_2 and D_3 , we find that T_3 , the trapezoidal approximation to $\int_a^b f(x) dx$ is given by

$$T_3 = D_1 + D_2 + D_3$$

= $\frac{1}{2} (f(x_0) + f(x_1)) \cdot \Delta x + \frac{1}{2} (f(x_1) + f(x_2)) \cdot \Delta x + \frac{1}{2} (f(x_2) + f(x_3)) \cdot \Delta x$.

Because both left and right endpoints are being used, we recognize within the trapezoidal approximation the use of both left and right Riemann sums. In particular, rearranging the expression for T_3 by removing factors of $\frac{1}{2}$ and Δx , grouping the left endpoint evaluations of f, and grouping the right endpoint evaluations of f, we see that

$$T_3 = \frac{1}{2} \left[\left(f(x_0) + f(x_1) + f(x_2) \right) \right] \Delta x + \left(f(x_1) + \frac{1}{2} \left[f(x_2) + f(x_3) \right) \right] \Delta x. \tag{5.6.4}$$

At this point, we observe that two familiar sums have arisen. Since the left Riemann sum L_3 is $L_3 = f(x_0)\Delta x + f(x_1)\Delta x + f(x_2)\Delta x$, and the right Riemann sum is $R_3 = f(x_1)\Delta x + f(x_2)\Delta x + f(x_3)\Delta x$, substituting L_3 and R_3 for the corresponding expressions in Equation (5.6.4), it follows that $T_3 = \frac{1}{2} [L_3 + R_3]$. We have thus seen the main ideas behind a very important result: using trapezoids to estimate the (signed) area bounded by a curve is the same as averaging the estimates generated by using left and right endpoints.

The Trapezoid Rule

The trapezoidal approximation, T_n , of the definite integral $\int_a^b f(x) dx$ using n subintervals is given by the rule

$$T_n = \left[\frac{1}{2} (f(x_0) + f(x_1)) + \frac{1}{2} (f(x_1) + f(x_2)) + \dots + \frac{1}{2} (f(x_{n-1}) + f(x_n)) \right] \Delta x.$$

$$= \sum_{i=0}^{n-1} \frac{1}{2} (f(x_i) + f(x_{i+1})) \Delta x.$$

Moreover, $T_n = \frac{1}{2} [L_n + R_n]$.

Activity 5.6.2. In this activity, we explore the relationships among the errors generated by left, right, midpoint, and trapezoid approximations to the definite integral $\int_1^2 \frac{1}{x^2} dx$

- a. Use the First FTC to evaluate $\int_1^2 \frac{1}{x^2} dx$ exactly.
- b. Use appropriate computing technology to compute the following approximations for $\int_1^2 \frac{1}{x^2} dx$: T_4 , M_4 , T_8 , and M_8 .
- c. Let the *error* of an approximation be the difference between the exact value of the definite integral and the resulting approximation. For instance, if we let $E_{T,4}$ represent the error that results from using the trapezoid rule with 4 subintervals to estimate the integral, we have

$$E_{T,4} = \int_1^2 \frac{1}{x^2} dx - T_4.$$

Similarly, we compute the error of the midpoint rule approximation with 8 subintervals by the formula

$$E_{M,8} = \int_{1}^{2} \frac{1}{x^2} dx - M_8.$$

Based on your work in (a) and (b) above, compute $E_{T,4}$, $E_{T,8}$, $E_{M,4}$, $E_{M,8}$.

- d. Which rule consistently over-estimates the exact value of the definite integral? Which rule consistently under-estimates the definite integral?
- e. What behavior(s) of the function $f(x) = \frac{1}{x^2}$ lead to your observations in (d)?

5.6.2 Comparing the Midpoint and Trapezoid Rules

We know from the definition of the definite integral of a continuous function f, that if we let n be large enough, we can make the value of any of the approximations L_n , R_n , and M_n as close as we'd like (in theory) to the exact value of $\int_a^b f(x) dx$. Thus, it may be natural to wonder why we ever use any rule other than L_n or R_n (with a sufficiently large n value) to estimate a definite integral. One of the primary reasons is that as $n \to \infty$, $\Delta x = \frac{b-a}{n} \to 0$, and thus in a Riemann sum calculation with a large n value, we end up multiplying by a number that is very close to zero. Doing so often generates roundoff error, as representing numbers close to zero accurately is a persistent challenge for computers.

Hence, we are exploring ways by which we can estimate definite integrals to high levels of precision, but without having to use extremely large values of n. Paying close attention to patterns in errors, such as those observed in Activity 5.6.2, is one way to begin to see some alternate approaches.

To begin, we make a comparison of the errors in the Midpoint and Trapezoid rules from two different perspectives. First, consider a function of consistent concavity on a given interval, and picture approximating the area bounded on that interval by both the Midpoint and

Trapezoid rules using a single subinterval.

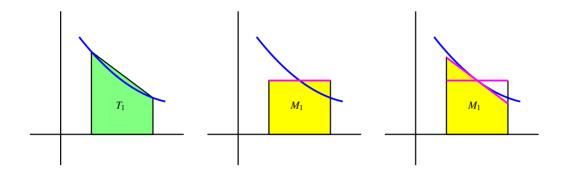


Figure 5.6.3: Estimating $\int_a^b f(x) dx$ using a single subinterval: at left, the trapezoid rule; in the middle, the midpoint rule; at right, a modified way to think about the midpoint rule.

As seen in Figure 5.6.3, it is evident that whenever the function is concave up on an interval, the Trapezoid Rule with one subinterval, T_1 , will overestimate the exact value of the definite integral on that interval. Moreover, from a careful analysis of the line that bounds the top of the rectangle for the Midpoint Rule (shown in magenta), we see that if we rotate this line segment until it is tangent to the curve at the point on the curve used in the Midpoint Rule (as shown at right in Figure 5.6.3), the resulting trapezoid has the same area as M_1 , and this value is less than the exact value of the definite integral. Hence, when the function is concave up on the interval, M_1 underestimates the integral's true value.

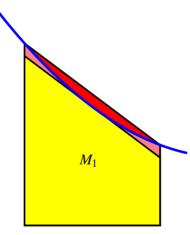


Figure 5.6.4: Comparing the error in estimating $\int_a^b f(x) dx$ using a single subinterval: in red, the error from the Trapezoid rule; in light red, the error from the Midpoint rule.

These observations extend easily to the situation where the function's concavity remains consistent but we use higher values of n in the Midpoint and Trapezoid Rules. Hence, whenever f is concave up on [a,b], T_n will overestimate the value of $\int_a^b f(x) dx$, while M_n will underestimate $\int_a^b f(x) dx$. The reverse observations are true in the situation where f is concave down.

Next, we compare the size of the errors between M_n and T_n . Again, we focus on M_1 and T_1 on an interval where the concavity of f is consistent. In Figure 5.6.4, where the error of the Trapezoid Rule is shaded in red, while the error of the Midpoint Rule is shaded lighter red, it is visually apparent that the error in the Trapezoid Rule is more significant. To see how much more significant, let's consider two examples and some particular computations.

If we let $f(x) = 1 - x^2$ and consider $\int_0^1 f(x) dx$, we know by the First FTC that the exact value of the integral is

$$\int_0^1 (1 - x^2) \, dx = x - \frac{x^3}{3} \Big|_0^1 = \frac{2}{3}.$$

Using appropriate technology to compute M_4 , M_8 , T_4 , and T_8 , as well as the corresponding errors $E_{M,4}$, $E_{M,8}$, $E_{T,4}$, and $E_{T,8}$, as we did in Activity 5.6.2, we find the results summarized in Table 5.6.5. Note that in the table, we also include the approximations and their errors for the example $\int_1^2 \frac{1}{v^2} dx$ from Activity 5.6.2.

Rule	$\int_0^1 (1 - x^2) dx = 0.\overline{6}$	error	$\int_{1}^{2} \frac{1}{x^{2}} dx = 0.5$	error
T_4	0.65625	-0.0104166667	0.5089937642	0.0089937642
M_4	0.671875	0.0052083333	0.4955479365	-0.0044520635
T_8	0.6640625	-0.0026041667	0.5022708502	0.0022708502
M_8	0.66796875	0.0013020833	0.4988674899	-0.0011325101

Table 5.6.5: Calculations of T_4 , M_4 , T_8 , and M_8 , along with corresponding errors, for the definite integrals $\int_0^1 (1-x^2) dx$ and $\int_1^2 \frac{1}{x^2} dx$.

Recall that for a given function f and interval [a,b], $E_{T,4}=\int_a^b f(x)\,dx-T_4$ calculates the difference between the exact value of the definite integral and the approximation generated by the Trapezoid Rule with n=4. If we look at not only $E_{T,4}$, but also the other errors generated by using T_n and M_n with n=4 and n=8 in the two examples noted in Table 5.6.5, we see an evident pattern. Not only is the sign of the error (which measures whether the rule generates an over- or under-estimate) tied to the rule used and the function's concavity, but the magnitude of the errors generated by T_n and M_n seems closely connected. In particular, the errors generated by the Midpoint Rule seem to be about half the size of those generated by the Trapezoid Rule.

That is, we can observe in both examples that $E_{M,4} \approx -\frac{1}{2}E_{T,4}$ and $E_{M,8} \approx -\frac{1}{2}E_{T,8}$, which demonstrates a property of the Midpoint and Trapezoid Rules that turns out to hold in general: for a function of consistent concavity, the error in the Midpoint Rule has the opposite sign and approximately half the magnitude of the error of the Trapezoid Rule. Said symbol-

ically,

$$E_{M,n}\approx -\frac{1}{2}E_{T,n}.$$

This important relationship suggests a way to combine the Midpoint and Trapezoid Rules to create an even more accurate approximation to a definite integral.

5.6.3 Simpson's Rule

When we first developed the Trapezoid Rule, we observed that it can equivalently be viewed as resulting from the average of the Left and Right Riemann sums:

$$T_n = \frac{1}{2}(L_n + R_n).$$

Whenever a function is always increasing or always decreasing on the interval [a, b], one of L_n and R_n will over-estimate the true value of $\int_a^b f(x) dx$, while the other will under-estimate the integral. Said differently, the errors found in L_n and R_n will have opposite signs; thus, averaging L_n and R_n eliminates a considerable amount of the error present in the respective approximations. In a similar way, it makes sense to think about averaging M_n and T_n in order to generate a still more accurate approximation.

At the same time, we've just observed that M_n is typically about twice as accurate as T_n . Thus, we instead choose to use the weighted average

$$S_{2n} = \frac{2M_n + T_n}{3}. ag{5.6.5}$$

The rule for S_{2n} giving by Equation (5.6.5) is usually known as Simpson's Rule.² Note that we use " S_{2n} " rather that " S_n " since the n points the Midpoint Rule uses are different from the n points the Trapezoid Rule uses, and thus Simpson's Rule is using 2n points at which to evaluate the function. We build upon the results in Table 5.6.5 to see the approximations generated by Simpson's Rule. In particular, in Table 5.6.6, we include all of the results in Table 5.6.5, but include additional results for $S_8 = \frac{2M_4 + T_4}{3}$ and $S_{16} = \frac{2M_8 + T_8}{3}$.

²Thomas Simpson was an 18th century mathematician; his idea was to extend the Trapezoid rule, but rather than using straight lines to build trapezoids, to use quadratic functions to build regions whose area was bounded by parabolas (whose areas he could find exactly). Simpson's Rule is often developed from the more sophisticated perspective of using interpolation by quadratic functions.

Rule	$\int_0^1 (1 - x^2) dx = 0.\overline{6}$	error	$\int_{1}^{2} \frac{1}{x^2} dx = 0.5$	error
T_4	0.65625	-0.0104166667	0.5089937642	0.0089937642
M_4	0.671875	0.0052083333	0.4955479365	-0.0044520635
S_8	0.6666666667	0	0.5000298792	0.0000298792
T_8	0.6640625	-0.0026041667	0.5022708502	0.0022708502
M_8	0.66796875	0.0013020833	0.4988674899	-0.0011325101
S_{16}	0.6666666667	0	0.5000019434	0.0000019434

Table 5.6.6: Table 5.6.5 updated to include S_8 , S_{16} , and the corresponding errors.

The results seen in Table 5.6.6 are striking. If we consider the S_{16} approximation of $\int_1^2 \frac{1}{x^2} dx$, the error is only $E_{S,16} = 0.0000019434$. By contrast, $L_8 = 0.5491458502$, so the error of that estimate is $E_{L,8} = -0.0491458502$. Moreover, we observe that generating the approximations for Simpson's Rule is almost no additional work: once we have L_n , R_n , and M_n for a given value of n, it is a simple exercise to generate T_n , and from there to calculate S_{2n} . Finally, note that the error in the Simpson's Rule approximations of $\int_0^1 (1-x^2) dx$ is zero!³

These rules are not only useful for approximating definite integrals such as $\int_0^1 e^{-x^2} dx$, for which we cannot find an elementary antiderivative of e^{-x^2} , but also for approximating definite integrals in the setting where we are given a function through a table of data.

Activity 5.6.3. A car traveling along a straight road is braking and its velocity is measured at several different points in time, as given in the following table. Assume that v is continuous, always decreasing, and always decreasing at a decreasing rate, as is suggested by the data.

seconds, t	0	0.3	0.6	0.9	1.2	1.5	1.8
Velocity in ft/sec, $v(t)$	100	99	96	90	80	50	0

Table 5.6.7: Data for the braking car.

- a. Plot the given data on the set of axes provided in Figure 5.6.8 with time on the horizontal axis and the velocity on the vertical axis.
- b. What definite integral will give you the exact distance the car traveled on [0, 1.8]?
- c. Estimate the total distance traveled on [0, 1.8] by computing L_3 , R_3 , and T_3 . Which of these under-estimates the true distance traveled?

³Similar to how the Midpoint and Trapezoid approximations are exact for linear functions, Simpson's Rule approximations are exact for quadratic and cubic functions. See additional discussion on this issue later in the section and in the exercises.

- d. Estimate the total distance traveled on [0, 1.8] by computing M_3 . Is this an over- or under-estimate? Why?
- e. Using your results from (c) and (d), improve your estimate further by using Simpson's Rule.
- f. What is your best estimate of the average velocity of the car on [0,1.8]? Why? What are the units on this quantity?

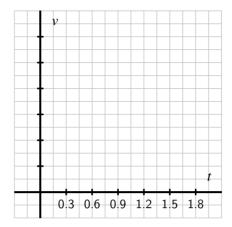


Figure 5.6.8: Axes for plotting the data in Activity 5.6.3.

5.6.4 Overall observations regarding L_n , R_n , T_n , M_n , and S_{2n} .

As we conclude our discussion of numerical approximation of definite integrals, it is important to summarize general trends in how the various rules over- or under-estimate the true value of a definite integral, and by how much. To revisit some past observations and see some new ones, we consider the following activity.

Activity 5.6.4. Consider the functions $f(x) = 2 - x^2$, $g(x) = 2 - x^3$, and $h(x) = 2 - x^4$, all on the interval [0, 1]. For each of the questions that require a numerical answer in what follows, write your answer exactly in fraction form.

- a. On the three sets of axes provided in Figure 5.6.9, sketch a graph of each function on the interval [0, 1], and compute L_1 and R_1 for each. What do you observe?
- b. Compute M_1 for each function to approximate $\int_0^1 f(x) dx$, $\int_0^1 g(x) dx$, and $\int_0^1 h(x) dx$, respectively.
- c. Compute T_1 for each of the three functions, and hence compute S_2 for each of the three functions.
- d. Evaluate each of the integrals $\int_0^1 f(x) dx$, $\int_0^1 g(x) dx$, and $\int_0^1 h(x) dx$ exactly using the First FTC.
- e. For each of the three functions f, g, and h, compare the results of L_1 , R_1 , M_1 , T_1 , and S_2 to the true value of the corresponding definite integral. What patterns do you observe?

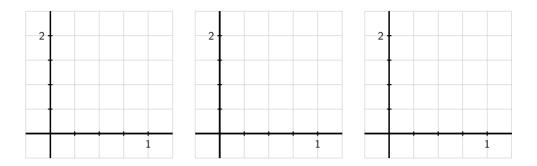


Figure 5.6.9: Axes for plotting the functions in Activity 5.6.4.

The results seen in the examples in Activity 5.6.4 generalize nicely. For instance, for any function f that is decreasing on [a,b], L_n will over-estimate the exact value of $\int_a^b f(x) dx$, and for any function f that is concave down on [a,b], M_n will over-estimate the exact value of the integral. An excellent exercise is to write a collection of scenarios of possible function behavior, and then categorize whether each of L_n , R_n , T_n , and M_n is an over- or underestimate.

Finally, we make two important notes about Simpson's Rule. When T. Simpson first developed this rule, his idea was to replace the function f on a given interval with a quadratic function that shared three values with the function f. In so doing, he guaranteed that this new approximation rule would be exact for the definite integral of any quadratic polynomial. In one of the pleasant surprises of numerical analysis, it turns out that even though it was designed to be exact for quadratic polynomials, Simpson's Rule is exact for any cubic polynomial: that is, if we are interested in an integral such as $\int_2^5 (5x^3 - 2x^2 + 7x - 4) \, dx$, S_{2n} will always be exact, regardless of the value of n. This is just one more piece of evidence that shows how effective Simpson's Rule is as an approximation tool for estimating definite integrals.⁴

Summary

- For a definite integral such as $\int_0^1 e^{-x^2} dx$ when we cannot use the First Fundamental Theorem of Calculus because the integrand lacks an elementary algebraic antiderivative, we can estimate the integral's value by using a sequence of Riemann sum approximations. Typically, we start by computing L_n , R_n , and M_n for one or more chosen values of n.
- The Trapezoid Rule, which estimates $\int_a^b f(x) dx$ by using trapezoids, rather than rectangles, can also be viewed as the average of Left and Right Riemann sums. That is, $T_n = \frac{1}{2}(L_n + R_n)$.

⁴One reason that Simpson's Rule is so effective is that S_{2n} benefits from using 2n + 1 points of data. Because it combines M_n , which uses n midpoints, and T_n , which uses the n + 1 endpoints of the chosen subintervals, S_{2n} takes advantage of the maximum amount of information we have when we know function values at the endpoints and midpoints of n subintervals.

• The Midpoint Rule is typically twice as accurate as the Trapezoid Rule, and the signs of the respective errors of these rules are opposites. Hence, by taking the weighted average $S_n = \frac{2M_n + T_n}{3}$, we can build a much more accurate approximation to $\int_a^b f(x) \, dx$ by using approximations we have already computed. The rule for S_n is known as Simpson's Rule, which can also be developed by approximating a given continuous function with pieces of quadratic polynomials.

Exercises



1. Note: for this problem, because later answers depend on earlier ones, you must enter answers for all answer blanks for the problem to be correctly graded. If you would like to get feedback before you completed all computations, enter a "1" for each answer you did not yet compute and then submit the problem. (But note that this will, obviously, result in a problem submission.)

(a) What is the exact value of $\int_0^3 e^x dx$?

$$\int_0^3 e^x dx = \boxed{}$$

Find LEFT(2), RIGHT(2), TRAP(2), MID(2), and SIMP(2); compute the error for each.

	LEFT(2)	RIGHT(2)	TRAP(2)	MID(2)	SIMP(2)
value					
error					

(c) Repeat part (b) with n = 4 (instead of n = 2).

	LEFT(4)	RIGHT(4)	TRAP(4)	MID(4)	SIMP(4)
value					
error					

(*d*)

For each rule in part (b), as n goes from n = 2 to n = 4, does the error go down approximately as you would expect? Explain by calculating the ratios of the errors:

Error LEFT(2)/Error LEFT(4) =

Error RIGHT(2)/Error RIGHT(4) =

Error TRAP(2)/Error TRAP(4) =

Error MID(2)/Error MID(4) =

Error SIMP(2) / Error SIMP(4) = $\[$ (Be sure that you can explain in words why these do (or don't) make sense.)



2. Using the figure showing f(x) below, order the following approximations to the integral $\int_0^3 f(x) dx$ and its exact value from smallest to largest.



Enter each of "LEFT(n)", "RIGHT(n)", "TRAP(n)", "MID(n)" and "Exact" in one of the following answer blanks to indicate the correct ordering:



3. Using a fixed number of subdivisions, we approximate the integrals of f and g on the interval shown in the figure below.

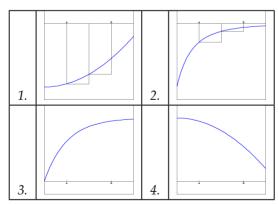




(The function f(x) is shown in blue, and g(x) in black.) For which function, f or g is LEFT more accurate? For which function, *f* or *g* is RIGHT more accurate? For which function, *f* or *q* is MID more accurate? For which function, *f* or *q* is TRAP more accurate?

4. Consider the four functions shown below. On the first two, an approximation for $\int_a^b f(x) dx$ is shown.





- 1. For graph number 1, Which integration method is shown?
- Is this method an over- or underestimate?
- 2. For graph number 2, Which integration method is shown?
- Is this method an over- or underestimate?
- 3. On a copy of graph number 3, sketch an estimate with n = 2 subdivisions using the midpoint rule.

Is this method an over- or underestimate?

4. On a copy of graph number 4, sketch an estimate with n=2 subdivisions using the trapezoid rule.

Is this method an over- or underestimate?

- **5.** Consider the definite integral $\int_0^1 x \tan(x) dx$.
 - a. Explain why this integral cannot be evaluated exactly by using either *u*-substitution or by integrating by parts.
 - b. Using 4 subintervals, compute L_4 , R_4 , M_4 , T_4 , and S_4 .
 - c. Which of the approximations in (b) is an over-estimate to the true value of $\int_0^1 x \tan(x) dx$? Which is an under-estimate? How do you know?
- **6.** For an unknown function f(x), the following information is known.
 - *f* is continuous on [3, 6];
 - *f* is either always increasing or always decreasing on [3, 6];
 - *f* has the same concavity throughout the interval [3, 6];
 - As approximations to $\int_3^6 f(x) dx$, $L_4 = 7.23$, $R_4 = 6.75$, and $M_4 = 7.05$.
 - a. Is *f* increasing or decreasing on [3, 6]? What data tells you?
 - b. Is *f* concave up or concave down on [3, 6]? Why?
 - c. Determine the best possible estimate you can for $\int_3^6 f(x) dx$, based on the given information.
- 7. The rate at which water flows through Table Rock Dam on the White River in Branson, MO, is measured in thousands of cubic feet per second (TCFS). As engineers open the floodgates, flow rates are recorded according to the following chart.

seconds, t	0	10	20	30	40	50	60
flow in TCFS, $r(t)$	2000	2100	2400	3000	3900	5100	6500

Table 5.6.10: Water flow data.

- a. What definite integral measures the total volume of water to flow through the dam in the 60 second time period provided by the table above?
- b. Use the given data to calculate M_n for the largest possible value of n to approximate the integral you stated in (a). Do you think M_n over- or under-estimates the exact value of the integral? Why?
- c. Approximate the integral stated in (a) by calculating S_n for the largest possible value of n, based on the given data.
- d. Compute $\frac{1}{60}S_n$ and $\frac{2000+2100+2400+3000+3900+5100+6500}{7}$. What quantity do both of these values estimate? Which is a more accurate approximation?

Using Definite Integrals

6.1 Using Definite Integrals to Find Area and Length

Motivating Questions

- How can we use definite integrals to measure the area between two curves?
- How do we decide whether to integrate with respect to *x* or with respect to *y* when we try to find the area of a region?
- How can a definite integral be used to measure the length of a curve?

Early on in our work with the definite integral, we learned that if we have a nonnegative velocity function, v, for an object moving along an axis, the area under the velocity function between a and b tells us the distance the object traveled on that time interval. Moreover, based on the definition of the definite integral, that area is given precisely by $\int_a^b v(t) \, dt$. Indeed, for any nonnegative function f on an interval [a,b], we know that $\int_a^b f(x) \, dx$ measures the area bounded by the curve and the x-axis between x = a and x = b.

Through our upcoming work in the present section and chapter, we will explore how definite integrals can be used to represent a variety of different physically important properties. In Preview Activity 6.1.1, we begin this investigation by seeing how a single definite integral may be used to represent the area between two curves.

Preview Activity 6.1.1. Consider the functions given by $f(x) = 5 - (x - 1)^2$ and g(x) = 4 - x.

- a. Use algebra to find the points where the graphs of f and g intersect.
- b. Sketch an accurate graph of f and g on the axes provided, labeling the curves by name and the intersection points with ordered pairs.
- c. Find and evaluate exactly an integral expression that represents the area between y = f(x) and the x-axis on the interval between the intersection points of f and g.



- d. Find and evaluate exactly an integral expression that represents the area between y = g(x) and the x-axis on the interval between the intersection points of f and g.
- e. What is the exact area between f and g between their intersection points? Why?

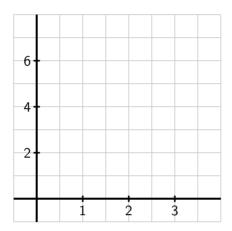


Figure 6.1.1: Axes for plotting f and g in Preview Activity 6.1.1

6.1.1 The Area Between Two Curves

Through Preview Activity 6.1.1, we encounter a natural way to think about the area between two curves: the area between the curves is the area beneath the upper curve minus the area below the lower curve. For the functions $f(x) = (x-1)^2 + 1$ and g(x) = x + 2, shown in Figure 6.1.2,

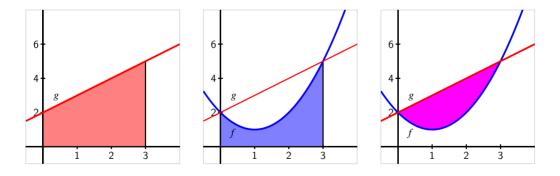


Figure 6.1.2: The areas bounded by the functions $f(x) = (x-1)^2 + 1$ and g(x) = x + 2 on the interval [0, 3].

we see that the upper curve is g(x) = x + 2, and that the graphs intersect at (0, 2) and (3, 5).

Note that we can find these intersection points by solving the system of equations given by $y = (x - 1)^2 + 1$ and y = x + 2 through substitution: substituting x + 2 for y in the first equation yields $x + 2 = (x - 1)^2 + 1$, so $x + 2 = x^2 - 2x + 1 + 1$, and thus

$$x^2 - 3x = x(x - 3) = 0$$

from which it follows that x = 0 or x = 3. Using y = x + 2, we find the corresponding y-values of the intersection points.

On the interval [0,3], the area beneath g is

$$\int_0^3 (x+2) \, dx = \frac{21}{2},$$

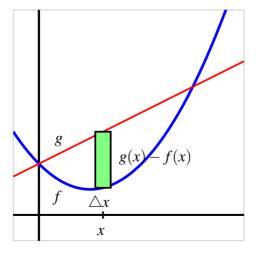
while the area under f on the same interval is

$$\int_0^3 [(x-1)^2 + 1] \, dx = 6.$$

Thus, the area between the curves is

$$A = \int_0^3 (x+2) \, dx - \int_0^3 \left[(x-1)^2 + 1 \right] dx = \frac{21}{2} - 6 = \frac{9}{2}. \tag{6.1.1}$$

A slightly different perspective is also helpful here: if we take the region between two curves and slice it up into thin vertical rectangles (in the same spirit as we originally sliced the region between a single curve and the x-axis in Section 4.2), then we see that the height of a typical rectangle is given by the difference between the two functions. For example, for the rectangle shown at left in Figure 6.1.3,



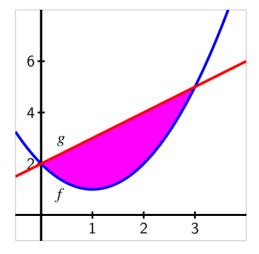


Figure 6.1.3: The area bounded by the functions $f(x) = (x-1)^2 + 1$ and g(x) = x + 2 on the interval [0,3].

we see that the rectangle's height is g(x) - f(x), while its width can be viewed as Δx , and thus the area of the rectangle is

$$A_{\text{rect}} = (g(x) - f(x))\Delta x.$$

The area between the two curves on [0,3] is thus approximated by the Riemann sum

$$A \approx \sum_{i=1}^{n} (g(x_i) - f(x_i)) \Delta x,$$

and then as we let $n \to \infty$, it follows that the area is given by the single definite integral

$$A = \int_0^3 (g(x) - f(x)) dx.$$
 (6.1.2)

In many applications of the definite integral, we will find it helpful to think of a "representative slice" and how the definite integral may be used to add these slices to find the exact value of a desired quantity. Here, the integral essentially sums the areas of thin rectangles.

Finally, whether we think of the area between two curves as the difference between the area bounded by the individual curves (as in (6.1.1)) or as the limit of a Riemann sum that adds the areas of thin rectangles between the curves (as in (6.1.2)), these two results are the same, since the difference of two integrals is the integral of the difference:

$$\int_0^3 g(x) \, dx - \int_0^3 f(x) \, dx = \int_0^3 (g(x) - f(x)) \, dx.$$

Moreover, our work so far in this section exemplifies the following general principle.

If two curves y = g(x) and y = f(x) intersect at (a, g(a)) and (b, g(b)), and for all x such that $a \le x \le b$, $g(x) \ge f(x)$, then the area between the curves is $A = \int_a^b (g(x) - f(x)) \, dx$.

Activity 6.1.2. In each of the following problems, our goal is to determine the area of the region described. For each region, (i) determine the intersection points of the curves, (ii) sketch the region whose area is being found, (iii) draw and label a representative slice, and (iv) state the area of the representative slice. Then, state a definite integral whose value is the exact area of the region, and evaluate the integral to find the numeric value of the region's area.

- a. The finite region bounded by $y = \sqrt{x}$ and $y = \frac{1}{4}x$.
- b. The finite region bounded by $y = 12 2x^2$ and $y = x^2 8$.
- c. The area bounded by the *y*-axis, $f(x) = \cos(x)$, and $g(x) = \sin(x)$, where we consider the region formed by the first positive value of x for which f and g intersect.

d. The finite regions between the curves $y = x^3 - x$ and $y = x^2$.

6.1.2 Finding Area with Horizontal Slices

At times, the shape of a geometric region may dictate that we need to use horizontal rectangular slices, rather than vertical ones. For instance, consider the region bounded by the parabola $x = y^2 - 1$ and the line y = x - 1, pictured in Figure 6.1.4. First, we observe that by solving the second equation for x and writing x = y + 1, we can eliminate a variable through substitution and find that $y + 1 = y^2 - 1$, and hence the curves intersect where $y^2 - y - 2 = 0$. Thus, we find y = -1 or y = 2, so the intersection points of the two curves are (0, -1) and (3, 2).

We see that if we attempt to use vertical rectangles to slice up the area, at certain values of x (specifically from x = -1 to x = 0, as seen in the center graph of Figure 6.1.4), the curves that govern the top and bottom of the rectangle are one and the same. This suggests, as shown in the rightmost graph in the figure, that we try using horizontal rectangles as a way to think about the area of the region.

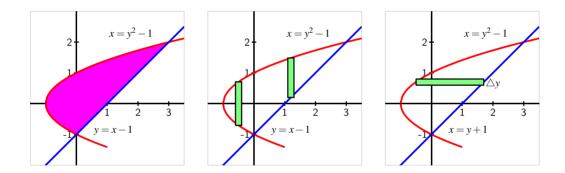


Figure 6.1.4: The area bounded by the functions $x = y^2 - 1$ and y = x - 1 (at left), with the region sliced vertically (center) and horizontally (at right).

For such a horizontal rectangle, note that its width depends on y, the height at which the rectangle is constructed. In particular, at a height y between y = -1 and y = 2, the right end of a representative rectangle is determined by the line, x = y + 1, while the left end of the rectangle is determined by the parabola, $x = y^2 - 1$, and the thickness of the rectangle is Δy .

Therefore, the area of the rectangle is

$$A_{\text{rect}} = [(y+1) - (y^2 - 1)]\Delta y,$$

from which it follows that the area between the two curves on the y-interval [-1,2] is approximated by the Riemann sum

$$A \approx \sum_{i=1}^{n} [(y_i + 1) - (y_i^2 - 1)] \Delta y.$$

Taking the limit of the Riemann sum, it follows that the area of the region is

$$A = \int_{y=-1}^{y=2} [(y+1) - (y^2 - 1)] \, dy. \tag{6.1.3}$$

We emphasize that we are integrating with respect to y; this is dictated by the fact that we chose to use horizontal rectangles whose widths depend on y and whose thickness is denoted Δy . It is a straightforward exercise to evaluate the integral in Equation (6.1.3) and find that $A = \frac{9}{2}$.

Just as with the use of vertical rectangles of thickness Δx , we have a general principle for finding the area between two curves, which we state as follows.

If two curves x = g(y) and x = f(y) intersect at (g(c), c) and (g(d), d), and for all y such that $c \le y \le d$, $g(y) \ge f(y)$, then the area between the curves is

$$A = \int_{y=c}^{y=d} (g(y) - f(y)) \, dy.$$

Activity 6.1.3. In each of the following problems, our goal is to determine the area of the region described. For each region, (i) determine the intersection points of the curves, (ii) sketch the region whose area is being found, (iii) draw and label a representative slice, and (iv) state the area of the representative slice. Then, state a definite integral whose value is the exact area of the region, and evaluate the integral to find the numeric value of the region's area. *Note well:* At the step where you draw a representative slice, you need to make a choice about whether to slice vertically or horizontally.

- a. The finite region bounded by $x = y^2$ and $x = 6 2y^2$.
- b. The finite region bounded by $x = 1 y^2$ and $x = 2 2y^2$.
- c. The area bounded by the *x*-axis, $y = x^2$, and y = 2 x.
- d. The finite regions between the curves $x = y^2 2y$ and y = x.

6.1.3 Finding the length of a curve

In addition to being able to use definite integrals to find the areas of certain geometric regions, we can also use the definite integral to find the length of a portion of a curve. We use the same fundamental principle: we take a curve whose length we cannot easily find, and slice it up into small pieces whose lengths we can easily approximate. In particular, we take a given curve and subdivide it into small approximating line segments, as shown at left in Figure 6.1.5.

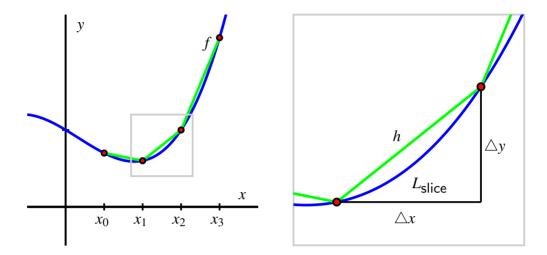


Figure 6.1.5: At left, a continuous function y = f(x) whose length we seek on the interval $a = x_0$ to $b = x_3$. At right, a close up view of a portion of the curve.

To see how we find such a definite integral that measures arc length on the curve y = f(x) from x = a to x = b, we think about the portion of length, $L_{\rm slice}$, that lies along the curve on a small interval of length Δx , and estimate the value of $L_{\rm slice}$ using a well-chosen triangle. In particular, if we consider the right triangle with legs parallel to the coordinate axes and hypotenuse connecting two points on the curve, as seen at right in Figure 6.1.5, we see that the length, h, of the hypotenuse approximates the length, $L_{\rm slice}$, of the curve between the two selected points. Thus,

$$L_{\rm slice} \approx h = \sqrt{(\Delta x)^2 + (\Delta y)^2}.$$

By algebraically rearranging the expression for the length of the hypotenuse, we see how a definite integral can be used to compute the length of a curve. In particular, observe that by removing a factor of $(\Delta x)^2$, we find that

$$\begin{split} L_{\text{slice}} &\approx \sqrt{(\Delta x)^2 + (\Delta y)^2} \\ &= \sqrt{(\Delta x)^2 \left(1 + \frac{(\Delta y)^2}{(\Delta x)^2}\right)} \\ &= \sqrt{1 + \frac{(\Delta y)^2}{(\Delta x)^2} \cdot \Delta x}. \end{split}$$

Furthermore, as $n \to \infty$ and $\Delta x \to 0$, it follows that $\frac{\Delta y}{\Delta x} \to \frac{dy}{dx} = f'(x)$. Thus, we can say that

$$L_{\rm slice} \approx \sqrt{1 + f'(x)^2} \Delta x.$$

Taking a Riemann sum of all of these slices and letting $n \to \infty$, we arrive at the following fact.

Given a differentiable function f on an interval [a, b], the total arc length, L, along the curve y = f(x) from x = a to x = b is given by

$$L = \int_a^b \sqrt{1 + f'(x)^2} \, dx.$$

Activity 6.1.4. Each of the following questions somehow involves the arc length along a curve.

- a. Use the definition and appropriate computational technology to determine the arc length along $y = x^2$ from x = -1 to x = 1.
- b. Find the arc length of $y = \sqrt{4 x^2}$ on the interval $-2 \le x \le 2$. Find this value in two different ways: (a) by using a definite integral, and (b) by using a familiar property of the curve.
- c. Determine the arc length of $y = xe^{3x}$ on the interval [0, 1].
- d. Will the integrals that arise calculating arc length typically be ones that we can evaluate exactly using the First FTC, or ones that we need to approximate? Why?
- e. A moving particle is traveling along the curve given by $y = f(x) = 0.1x^2 + 1$, and does so at a constant rate of 7 cm/sec, where both x and y are measured in cm (that is, the curve y = f(x) is the path along which the object actually travels; the curve is not a "position function"). Find the position of the particle when t = 4 sec, assuming that when t = 0, the particle's location is (0, f(0)).

Summary

• To find the area between two curves, we think about slicing the region into thin rectangles. If, for instance, the area of a typical rectangle on the interval x = a to x = b is given by $A_{\text{rect}} = (g(x) - f(x))\Delta x$, then the exact area of the region is given by the definite integral

$$A = \int_a^b (g(x) - f(x)) dx.$$

• The shape of the region usually dictates whether we should use vertical rectangles of thickness Δx or horizontal rectangles of thickness Δy . We desire to have the height of the rectangle governed by the difference between two curves: if those curves are best thought of as functions of y, we use horizontal rectangles, whereas if those curves are best viewed as functions of x, we use vertical rectangles.

• The arc length, *L*, along the curve y = f(x) from x = a to x = b is given by

$$L = \int_a^b \sqrt{1 + f'(x)^2} \, dx.$$

Exercises

The area is

1. Find the area of the region between $y = x^{1/2}$ and $y = x^{1/4}$ for $0 \le x \le 1$. area =



2. Find the area between $y = 7 \sin x$ and $y = 10 \cos x$ over the interval $[0, \pi]$. Sketch the curves if necessary.



A =

3. Sketch the region enclosed by $x + y^2 = 42$ and x + y = 0. Decide whether to integrate with respect to x or y, and then find the area of the region.



4. Find the arc length of the graph of the function $f(x) = 9\sqrt{x^3}$ from x = 5 to x = 8. arc length =



- 5. Find the exact area of each described region.
 - a. The finite region between the curves x = y(y-2) and x = -(y-1)(y-3).
 - b. The region between the sine and cosine functions on the interval $\left[\frac{\pi}{4}, \frac{3\pi}{4}\right]$.
 - c. The finite region between $x = y^2 y 2$ and y = 2x 1.
 - d. The finite region between y = mx and $y = x^2 1$, where m is a positive constant.
- **6.** Let $f(x) = 1 x^2$ and $g(x) = ax^2 a$, where a is an unknown positive real number. For what value(s) of a is the area between the curves f and g equal to 2?
- 7. Let $f(x) = 2 x^2$. Recall that the average value of any continuous function f on an interval [a, b] is given by $\frac{1}{b-a} \int_a^b f(x) dx$.
 - a. Find the average value of $f(x) = 2 x^2$ on the interval $[0, \sqrt{2}]$. Call this value r.
 - b. Sketch a graph of y = f(x) and y = r. Find their intersection point(s).
 - c. Show that on the interval $[0, \sqrt{2}]$, the amount of area that lies below y = f(x) and above y = r is equal to the amount of area that lies below y = r and above y = f(x).
 - d. Will the result of (c) be true for any continuous function and its average value on any interval? Why?

6.2 Using Definite Integrals to Find Volume

Motivating Questions

- How can we use a definite integral to find the volume of a three-dimensional solid
 of revolution that results from revolving a two-dimensional region about a particular
 axis?
- In what circumstances do we integrate with respect to *y* instead of integrating with respect to *x*?
- What adjustments do we need to make if we revolve about a line other than the *x* or *y*-axis?

Just as we can use definite integrals to add the areas of rectangular slices to find the exact area that lies between two curves, we can also employ integrals to determine the volume of certain regions that have cross-sections of a particular consistent shape.

As a very elementary example, consider a cylinder of radius 2 and height 3, as pictured in Figure 6.2.1. While we know that we can compute the area of any circular cylinder by the formula $V = \pi r^2 h$, if we think about slicing the cylinder into thin pieces, we see that each is a cylinder of radius r = 2 and height (thickness) Δx . Hence, the volume of a representative slice is

$$V_{\text{slice}} = \pi \cdot 2^2 \cdot \Delta x$$
.

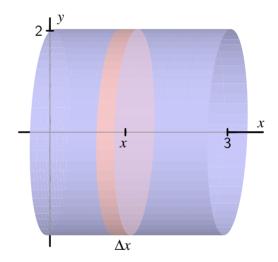


Figure 6.2.1: A right circular cylinder.

Letting $\Delta x \to 0$ and using a definite integral to add the volumes of the slices, we find that

$$V = \int_0^3 \pi \cdot 2^2 \, dx.$$

Moreover, since $\int_0^3 4\pi \, dx = 12\pi$, we have found that the volume of the cylinder is 12π . The principal problem of interest in our upcoming work will be to find the volume of certain solids whose cross-sections are all thin cylinders (or washers) and to do so by using a definite

integral. To that end, we first consider another familiar shape in Preview Activity 6.2.1: a circular cone.

Preview Activity 6.2.1. Consider a circular cone of radius 3 and height 5, which we view horizontally as pictured in Figure 6.2.2. Our goal in this activity is to use a definite integral to determine the volume of the cone.

- a. Find a formula for the linear function y = f(x) that is pictured in Figure 6.2.2.
- b. For the representative slice of thickness Δx that is located horizontally at a location x (somewhere between x=0 and x=5), what is the radius of the representative slice? Note that the radius depends on the value of x.
- c. What is the volume of the representative slice you found in (b)?
- d. What definite integral will sum the volumes of the thin slices across the full horizontal span of the cone? What is the exact value of this definite integral?
- e. Compare the result of your work in (d) to the volume of the cone that comes from using the formula $V_{\text{cone}} = \frac{1}{3}\pi r^2 h$.

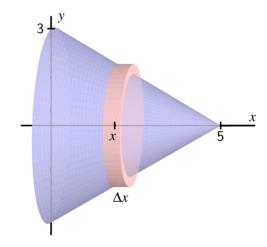


Figure 6.2.2: The circular cone described in Preview Activity 6.2.1

6.2.1 The Volume of a Solid of Revolution

A solid of revolution is a three dimensional solid that can be generated by revolving one or more curves around a fixed axis. For example, we can think of a circular cylinder as a solid of revolution: in Figure 6.2.1, this could be accomplished by revolving the line segment from (0,2) to (3,2) about the x-axis. Likewise, the circular cone in Figure 6.2.2 is the solid of revolution generated by revolving the portion of the line $y = 3 - \frac{3}{5}x$ from x = 0 to x = 5 about the x-axis. It is particularly important to notice in any solid of revolution that if we

slice the solid perpendicular to the axis of revolution, the resulting cross-section is circular.

We consider two examples to highlight some of the natural issues that arise in determining the volume of a solid of revolution.

Example 6.2.3. Find the volume of the solid of revolution generated when the region *R* bounded by $y = 4 - x^2$ and the *x*-axis is revolved about the *x*-axis.

Solution. First, we observe that $y = 4 - x^2$ intersects the *x*-axis at the points (-2,0) and (2,0). When we take the region *R* that lies between the curve and the *x*-axis on this interval and revolve it about the *x*-axis, we get the three-dimensional solid pictured in Figure 6.2.4.

Taking a representative slice of the solid located at a value x that lies between x = -2 and x = 2, we see that the thickness of such a slice is Δx (which is also the height of the cylinder-shaped slice), and that the radius of the slice is determined by the curve $y = 4 - x^2$. Hence, we find that

$$V_{\text{slice}} = \pi (4 - x^2)^2 \Delta x,$$

since the volume of a cylinder of radius r and height h is $V = \pi r^2 h$.

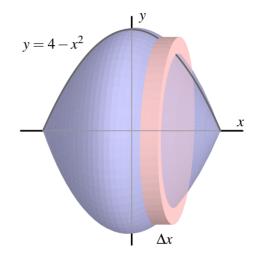


Figure 6.2.4: The solid of revolution in Example 6.2.3.

Using a definite integral to sum the volumes of the representative slices, it follows that

$$V = \int_{-2}^{2} \pi (4 - x^2)^2 \, dx.$$

It is straightforward to evaluate the integral and find that the volume is $V = \frac{512}{15}\pi$.

For a solid such as the one in Example 6.2.3, where each cross-section is a cylindrical disk, we first find the volume of a typical cross-section (noting particularly how this volume depends on x), and then we integrate over the range of x-values through which we slice the solid in order to find the exact total volume. Often, we will be content with simply finding the integral that represents the sought volume; if we desire a numeric value for the integral, we typically use a calculator or computer algebra system to find that value.

The general principle we are using to find the volume of a solid of revolution generated by a single curve is often called the *disk method*.

The Disk Method

If y = r(x) is a nonnegative continuous function on [a, b], then the volume of the solid of revolution generated by revolving the curve about the x-axis over this interval is given by

$$V = \int_a^b \pi r(x)^2 \, dx.$$

A different type of solid can emerge when two curves are involved, as we see in the following example.

Example 6.2.5. Find the volume of the solid of revolution generated when the finite region R that lies between $y = 4 - x^2$ and y = x + 2 is revolved about the x-axis.

Solution. First, we must determine where the curves $y = 4 - x^2$ and y = x + 2 intersect. Substituting the expression for y from the second equation into the first equation, we find that $x + 2 = 4 - x^2$. Rearranging, it follows that

$$x^2 + x - 2 = 0,$$

and the solutions to this equation are x = -2 and x = 1. The curves therefore cross at (-2, 0) and (1, 1).

When we take the region R that lies between the curves and revolve it about the x-axis, we get the three-dimensional solid pictured at left in Figure 6.2.6.

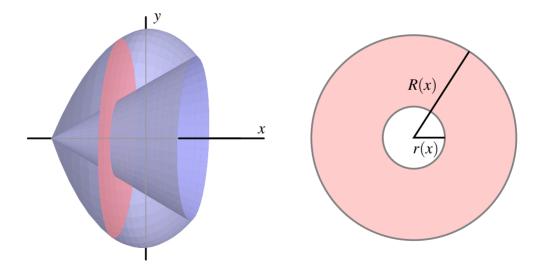


Figure 6.2.6: At left, the solid of revolution in Example 6.2.5. At right, a typical slice with inner radius r(x) and outer radius R(x).

Immediately we see a major difference between the solid in this example and the one in

Example 6.2.3: here, the three-dimensional solid of revolution isn't "solid" in the sense that it has open space in its center. If we slice the solid perpendicular to the axis of revolution, we observe that in this setting the resulting representative slice is not a solid disk, but rather a *washer*, as pictured at right in Figure 6.2.6. Moreover, at a given location x between x = -2 and x = 1, the small radius r(x) of the inner circle is determined by the curve y = x + 2, so r(x) = x + 2. Similarly, the big radius R(x) comes from the function $y = 4 - x^2$, and thus $R(x) = 4 - x^2$.

Thus, to find the volume of a representative slice, we compute the volume of the outer disk and subtract the volume of the inner disk. Since

$$\pi R(x)^2 \Delta x - \pi r(x)^2 \Delta x = \pi [R(x)^2 - r(x)^2] \Delta x,$$

it follows that the volume of a typical slice is

$$V_{\text{slice}} = \pi [(4 - x^2)^2 - (x + 2)^2] \Delta x.$$

Hence, using a definite integral to sum the volumes of the respective slices across the integral, we find that

$$V = \int_{-2}^{1} \pi [(4 - x^2)^2 - (x + 2)^2] dx.$$

Evaluating the integral, the volume of the solid of revolution is $V = \frac{108}{5}\pi$.

The general principle we are using to find the volume of a solid of revolution generated by a single curve is often called the *washer method*.

The Washer Method

If y = R(x) and y = r(x) are nonnegative continuous functions on [a, b] that satisfy $R(x) \ge r(x)$ for all x in [a, b], then the volume of the solid of revolution generated by revolving the region between them about the x-axis over this interval is given by

$$V = \int_{a}^{b} \pi [R(x)^{2} - r(x)^{2}] dx.$$

Activity 6.2.2. In each of the following questions, draw a careful, labeled sketch of the region described, as well as the resulting solid that results from revolving the region about the stated axis. In addition, draw a representative slice and state the volume of that slice, along with a definite integral whose value is the volume of the entire solid. It is not necessary to evaluate the integrals you find.

- a. The region *S* bounded by the *x*-axis, the curve $y = \sqrt{x}$, and the line x = 4; revolve *S* about the *x*-axis.
- b. The region *S* bounded by the *y*-axis, the curve $y = \sqrt{x}$, and the line y = 2; revolve *S* about the *x*-axis.

- c. The finite region *S* bounded by the curves $y = \sqrt{x}$ and $y = x^3$; revolve *S* about the *x*-axis.
- d. The finite region *S* bounded by the curves $y = 2x^2 + 1$ and $y = x^2 + 4$; revolve *S* about the *x*-axis.
- e. The region *S* bounded by the *y*-axis, the curve $y = \sqrt{x}$, and the line y = 2; revolve *S* about the *y*-axis. How is this problem different from the one posed in part (b)?

6.2.2 Revolving about the *y*-axis

As seen in Activity 6.2.2, problem (e), the problem changes considerably when we revolve a given region about the y-axis. Foremost, this is due to the fact that representative slices now have thickness Δy , which means that it becomes necessary to integrate with respect to y. Let's consider a particular example to demonstrate some of the key issues.

Example 6.2.7. Find the volume of the solid of revolution generated when the finite region R that lies between $y = \sqrt{x}$ and $y = x^4$ is revolved about the y-axis.

Solution. We observe that these two curves intersect when x = 1, hence at the point (1, 1). When we take the region R that lies between the curves and revolve it about the y-axis, we get the three-dimensional solid pictured at left in Figure 6.2.8.

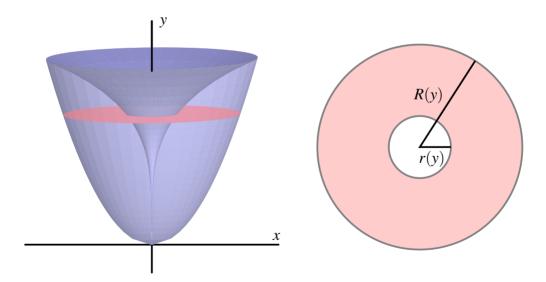


Figure 6.2.8: At left, the solid of revolution in Example 6.2.7. At right, a typical slice with inner radius r(y) and outer radius R(y).

Now, it is particularly important to note that the thickness of a representative slice is Δy , and that the slices are only cylindrical washers in nature when taken perpendicular to the

y-axis. Hence, we envision slicing the solid horizontally, starting at y = 0 and proceeding up to y = 1. Because the inner radius is governed by the curve $y = \sqrt{x}$, but from the perspective that x is a function of y, we solve for x and get $x = y^2 = r(y)$. In the same way, we need to view the curve $y = x^4$ (which governs the outer radius) in the form where x is a function of y, and hence $x = \sqrt[4]{y}$. Therefore, we see that the volume of a typical slice is

$$V_{\text{slice}} = \pi [R(y)^2 - r(y)^2] = \pi [\sqrt[4]{y^2} - (y^2)^2] \Delta y.$$

Using a definite integral to sum the volume of all the representative slices from y = 0 to y = 1, the total volume is

$$V = \int_{y=0}^{y=1} \pi \left[\sqrt[4]{y^2} - (y^2)^2 \right] dy.$$

It is straightforward to evaluate the integral and find that $V = \frac{7}{15}\pi$.

Activity 6.2.3. In each of the following questions, draw a careful, labeled sketch of the region described, as well as the resulting solid that results from revolving the region about the stated axis. In addition, draw a representative slice and state the volume of that slice, along with a definite integral whose value is the volume of the entire solid. It is not necessary to evaluate the integrals you find.

- a. The region *S* bounded by the *y*-axis, the curve $y = \sqrt{x}$, and the line y = 2; revolve *S* about the *y*-axis.
- b. The region *S* bounded by the *x*-axis, the curve $y = \sqrt{x}$, and the line x = 4; revolve *S* about the *y*-axis.
- c. The finite region *S* in the first quadrant bounded by the curves y = 2x and $y = x^3$; revolve *S* about the *x*-axis.
- d. The finite region *S* in the first quadrant bounded by the curves y = 2x and $y = x^3$; revolve *S* about the *y*-axis.
- e. The finite region *S* bounded by the curves $x = (y 1)^2$ and y = x 1; revolve *S* about the *y*-axis

6.2.3 Revolving about horizontal and vertical lines other than the coordinate axes

Just as we can revolve about one of the coordinate axes (y = 0 or x = 0), it is also possible to revolve around any horizontal or vertical line. Doing so essentially adjusts the radii of cylinders or washers involved by a constant value. A careful, well-labeled plot of the solid of revolution will usually reveal how the different axis of revolution affects the definite integral we set up. Again, an example is instructive.

Example 6.2.9. Find the volume of the solid of revolution generated when the finite region S that lies between $y = x^2$ and y = x is revolved about the line y = -1.

Solution.

Graphing the region between the two curves in the first quadrant between their points of intersection ((0,0) and (1,1)) and then revolving the region about the line y = -1, we see the solid shown in Figure 6.2.10. Each slice of the solid perpendicular to the axis of revolution is a washer, and the radii of each washer are governed by the curves $y = x^2$ and y = x. But we also see that there is one added change: the axis of revolution adds a fixed length to each radius. In particular, the inner radius of a typical slice, r(x), is given by $r(x) = x^2 + 1$, while the outer radius is R(x) = x + 1.

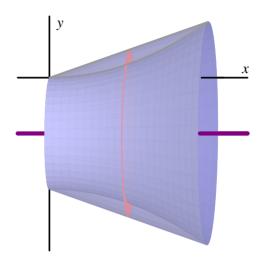


Figure 6.2.10: The solid of revolution described in Example 6.2.9.

Therefore, the volume of a typical slice is

$$V_{\rm slice} = \pi [R(x)^2 - r(x)^2] \Delta x = \pi \left[(x+1)^2 - (x^2+1)^2 \right] \Delta x.$$

Finally, we integrate to find the total volume, and

$$V = \int_0^1 \pi \left[(x+1)^2 - (x^2+1)^2 \right] dx = \frac{7}{15} \pi.$$

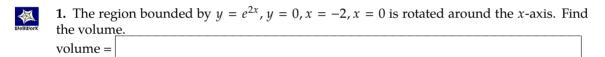
Activity 6.2.4. In each of the following questions, draw a careful, labeled sketch of the region described, as well as the resulting solid that results from revolving the region about the stated axis. In addition, draw a representative slice and state the volume of that slice, along with a definite integral whose value is the volume of the entire solid. It is not necessary to evaluate the integrals you find. For each prompt, use the finite region S in the first quadrant bounded by the curves y = 2x and $y = x^3$.

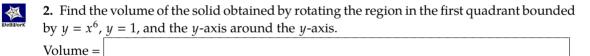
- a. Revolve *S* about the line y = -2.
- c. Revolve *S* about the line x = -1.
- b. Revolve *S* about the line y = 4.
- d. Revolve *S* about the line x = 5.

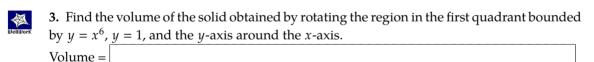
Summary

- We can use a definite integral to find the volume of a three-dimensional solid of revolution that results from revolving a two-dimensional region about a particular axis by taking slices perpendicular to the axis of revolution which will then be circular disks or washers.
- If we revolve about a vertical line and slice perpendicular to that line, then our slices are horizontal and of thickness Δy . This leads us to integrate with respect to y, as opposed to with respect to x when we slice a solid vertically.
- If we revolve about a line other than the *x* or *y*-axis, we need to carefully account for the shift that occurs in the radius of a typical slice. Normally, this shift involves taking a sum or difference of the function along with the constant connected to the equation for the horizontal or vertical line; a well-labeled diagram is usually the best way to decide the new expression for the radius.

Exercises







4. Find the volume of the solid obtained by rotating the region in the first quadrant bounded by $y = x^6$, y = 1, and the *y*-axis about the line y = -5.

 $y = x^2, \quad y = 1$ about the line y = 4 .

5. Find the volume of the solid obtained by rotating the region bounded by the curves

Answer: 6. Find the volume of the solid obtained by rotating the region bounded by the given curves

$$y = x^2$$
, $x = y^2$

about the line x = -5

Answer:	

- 7. Consider the curve $f(x) = 3\cos(\frac{x^3}{4})$ and the portion of its graph that lies in the first quadrant between the *y*-axis and the first positive value of *x* for which f(x) = 0. Let *R* denote the region bounded by this portion of *f*, the *x*-axis, and the *y*-axis.
 - a. Set up a definite integral whose value is the exact arc length of f that lies along the upper boundary of R. Use technology appropriately to evaluate the integral you find.
 - b. Set up a definite integral whose value is the exact area of *R*. Use technology appropriately to evaluate the integral you find.
 - c. Suppose that the region *R* is revolved around the *x*-axis. Set up a definite integral whose value is the exact volume of the solid of revolution that is generated. Use technology appropriately to evaluate the integral you find.
 - d. Suppose instead that *R* is revolved around the *y*-axis. If possible, set up an integral expression whose value is the exact volume of the solid of revolution and evaluate the integral using appropriate technology. If not possible, explain why.
- **8.** Consider the curves given by $y = \sin(x)$ and $y = \cos(x)$. For each of the following problems, you should include a sketch of the region/solid being considered, as well as a labeled representative slice.
 - a. Sketch the region R bounded by the y-axis and the curves $y = \sin(x)$ and $y = \cos(x)$ up to the first positive value of x at which they intersect. What is the exact intersection point of the curves?
 - b. Set up a definite integral whose value is the exact area of *R*.
 - c. Set up a definite integral whose value is the exact volume of the solid of revolution generated by revolving *R* about the *x*-axis.
 - d. Set up a definite integral whose value is the exact volume of the solid of revolution generated by revolving R about the y-axis.
 - e. Set up a definite integral whose value is the exact volume of the solid of revolution generated by revolving R about the line y = 2.
 - f. Set up a definite integral whose value is the exact volume of the solid of revolution generated by revolving R about the x = -1.
- **9.** Consider the finite region *R* that is bounded by the curves $y = 1 + \frac{1}{2}(x 2)^2$, $y = \frac{1}{2}x^2$, and x = 0.
 - a. Determine a definite integral whose value is the area of the region enclosed by the two curves.
 - b. Find an expression involving one or more definite integrals whose value is the volume of the solid of revolution generated by revolving the region R about the line y = -1.
 - c. Determine an expression involving one or more definite integrals whose value is the volume of the solid of revolution generated by revolving the region *R* about the *y*-axis.
 - d. Find an expression involving one or more definite integrals whose value is the perimeter of the region R.

6.3 Density, Mass, and Center of Mass

Motivating Questions

- How are mass, density, and volume related?
- How is the mass of an object with varying density computed?
- What is is the center of mass of an object, and how are definite integrals used to compute it?

We have seen in several different circumstances how studying the units on the integrand and variable of integration enables us to better understand the meaning of a definite integral. For instance, if v(t) is the velocity of an object moving along an axis, measured in feet per second, while t measures time in seconds, then both the definite integral and its Riemann sum approximation,

$$\int_a^b v(t) dt \approx \sum_{i=1}^n v(t_i) \Delta t,$$

have their overall units given by the product of the units of v(t) and t:

$$(feet/sec) \cdot (sec) = feet.$$

Thus, $\int_a^b v(t) dt$ measures the total change in position (in feet) of the moving object.

This type of unit analysis will be particularly helpful to us in what follows. To begin, in the following preview activity we consider two different definite integrals where the integrand is a function that measures how a particular quantity is distributed over a region and think about how the units on the integrand and the variable of integration indicate the meaning of the integral.

Preview Activity 6.3.1. In each of the following scenarios, we consider the distribution of a quantity along an axis.

- a. Suppose that the function $c(x) = 200 + 100e^{-0.1x}$ models the density of traffic on a straight road, measured in cars per mile, where x is number of miles east of a major interchange, and consider the definite integral $\int_0^2 (200 + 100e^{-0.1x}) dx$.
 - i. What are the units on the product $c(x) \cdot \Delta x$?
 - ii. What are the units on the definite integral and its Riemann sum approximation given by

$$\int_0^2 c(x) dx \approx \sum_{i=1}^n c(x_i) \Delta x?$$

- iii. Evaluate the definite integral $\int_0^2 c(x) dx = \int_0^2 (200 + 100e^{-0.1x}) dx$ and write one sentence to explain the meaning of the value you find.
- b. On a 6 foot long shelf filled with books, the function *B* models the distribution of the weight of the books, in pounds per inch, where *x* is the number of inches from the left end of the bookshelf. Let B(x) be given by the rule $B(x) = 0.5 + \frac{1}{(x+1)^2}$.
 - i. What are the units on the product $B(x) \cdot \Delta x$?
 - ii. What are the units on the definite integral and its Riemann sum approximation given by

$$\int_{12}^{36} B(x) dx \approx \sum_{i=1}^{n} B(x_i) \Delta x?$$

iii. Evaluate the definite integral $\int_0^{72} B(x) dx = \int_0^{72} \left(0.5 + \frac{1}{(x+1)^2}\right) dx$ and write one sentence to explain the meaning of the value you find.

6.3.1 Density

The *mass* of a quantity, typically measured in metric units such as grams or kilograms, is a measure of the amount of a quantity. In a corresponding way, the *density* of an object measures the distribution of mass per unit volume. For instance, if a brick has mass 3 kg and volume 0.002 m^3 , then the density of the brick is

$$\frac{3kg}{0.002m^3} = 1500 \frac{kg}{m^3}.$$

As another example, the mass density of water is 1000 kg/m³. Each of these relationships demonstrate the following general principle.

For an object of constant density d, with mass m and volume V,

$$d = \frac{m}{V}$$
, or $m = d \cdot V$.

But what happens when the density is not constant?

If we consider the formula $m = d \cdot V$, it is reminiscent of two other equations that we have used frequently in recent work: for a body moving in a fixed direction, distance = rate \cdot time, and, for a rectangle, its area is given by $A = l \cdot w$. These formulas hold when the principal quantities involved, such as the rate the body moves and the height of the rectangle, are *constant*. When these quantities are not constant, we have turned to the definite integral for assistance. The main idea in each situation is that by working with small slices of the quantity that is varying, we can use a definite integral to add up the values of small pieces on which the quantity of interest (such as the velocity of a moving object) are approximately constant.

For example, in the setting where we have a nonnegative velocity function that is not constant, over a short time interval Δt we know that the distance traveled is approximately $v(t)\Delta t$, since v(t) is almost constant on a small interval, and for a constant rate, distance = rate \cdot time. Similarly, if we are thinking about the area under a nonnegative function f whose value is changing, on a short interval Δx the area under the curve is approximately the area of the rectangle whose height is f(x) and whose width is Δx : $f(x)\Delta x$. Both of these principles are represented visually in Figure 6.3.1.

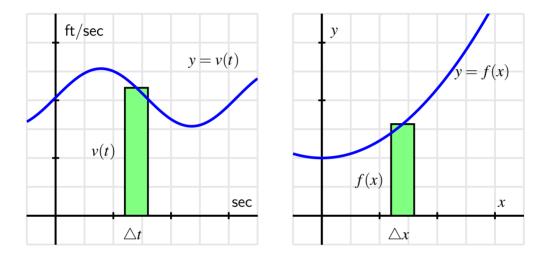


Figure 6.3.1: At left, estimating a small amount of distance traveled, $v(t)\Delta t$, and at right, a small amount of area under the curve, $f(x)\Delta x$.

In a similar way, if we consider the setting where the density of some quantity is not constant, the definite integral enables us to still compute the overall mass of the quantity. Throughout, we will focus on problems where the density varies in only one dimension, say along a single axis, and think about how mass is distributed relative to location along the axis.

Let's consider a thin bar of length b that is situated so its left end is at the origin, where x=0, and assume that the bar has constant cross-sectional area of 1 cm². We let the function $\rho(x)$ represent the mass density function of the bar, measured in grams per cubic centimeter. That is, given a location x, $\rho(x)$ tells us approximately how much mass will be found in a one-centimeter wide slice of the bar at x.

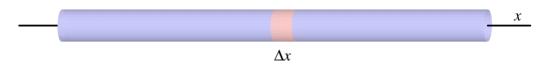


Figure 6.3.2: A thin bar of constant cross-sectional area 1 cm² with density function $\rho(x)$ g/cm³.

If we now consider a thin slice of the bar of width Δx , as pictured in Figure 6.3.2, the volume of such a slice is the cross-sectional area times Δx . Since the cross-sections each have constant area 1 cm², it follows that the volume of the slice is $1\Delta x$ cm³. Moreover, since mass is the product of density and volume (when density is constant), we see that the mass of this given slice is approximately

$$\text{mass}_{\text{slice}} \approx \rho(x) \frac{\text{g}}{\text{cm}^3} \cdot 1\Delta x \text{ cm}^3 = \rho(x) \cdot \Delta x \text{ g}.$$

Hence, for the corresponding Riemann sum (and thus for the integral that it approximates),

$$\sum_{i=1}^{n} \rho(x_i) \Delta x \approx \int_{0}^{b} \rho(x) \, dx,$$

we see that these quantities measure the mass of the bar between 0 and b. (The Riemann sum is an approximation, while the integral will be the exact mass.)

At this point, we note that we will be focused primarily on situations where mass is distributed relative to horizontal location, x, for objects whose cross-sectional area is constant. In that setting, it makes sense to think of the density function $\rho(x)$ with units "mass per unit length," such as g/cm. Thus, when we compute $\rho(x) \cdot \Delta x$ on a small slice Δx , the resulting units are $g/cm \cdot cm = g$, which thus measures the mass of the slice. The general principle follows.

For an object of constant cross-sectional area whose mass is distributed along a single axis according to the function $\rho(x)$ (whose units are units of mass per unit of length), the total mass, M of the object between x = a and x = b is given by

$$M = \int_a^b \rho(x) \, dx.$$

Activity 6.3.2. Consider the following situations in which mass is distributed in a non-constant manner.

- a. Suppose that a thin rod with constant cross-sectional area of 1 cm² has its mass distributed according to the density function $\rho(x) = 2e^{-0.2x}$, where x is the distance in cm from the left end of the rod, and the units on $\rho(x)$ are g/cm. If the rod is 10 cm long, determine the exact mass of the rod.
- b. Consider the cone that has a base of radius 4 m and a height of 5 m. Picture the cone lying horizontally with the center of its base at the origin and think of the cone as a solid of revolution.
 - i. Write and evaluate a definite integral whose value is the volume of the cone.
 - ii. Next, suppose that the cone has uniform density of 800 kg/m³. What is the mass of the solid cone?
 - iii. Now suppose that the cone's density is not uniform, but rather that the cone

is most dense at its base. In particular, assume that the density of the cone is uniform across cross sections parallel to its base, but that in each such cross section that is a distance x units from the origin, the density of the cross section is given by the function $\rho(x) = 400 + \frac{200}{1+x^2}$, measured in kg/m³. Determine and evaluate a definite integral whose value is the mass of this cone of non-uniform density. Do so by first thinking about the mass of a given slice of the cone x units away from the base; remember that in such a slice, the density will be *essentially constant*.

c. Let a thin rod of constant cross-sectional area 1 cm² and length 12 cm have its mass be distributed according to the density function $\rho(x) = \frac{1}{25}(x-15)^2$, measured in g/cm. Find the exact location z at which to cut the bar so that the two pieces will each have identical mass.

6.3.2 Weighted Averages

The concept of an average is a natural one, and one that we have used repeatedly as part of our understanding of the meaning of the definite integral. If we have n values $a_1, a_2, ..., a_n$, we know that their average is given by

$$\frac{a_1+a_2+\cdots+a_n}{n},$$

and for a quantity being measured by a function f on an interval [a, b], the average value of the quantity on [a, b] is

$$\frac{1}{b-a}\int_a^b f(x)\,dx.$$

As we continue to think about problems involving the distribution of mass, it is natural to consider the idea of a *weighted* average, where certain quantities involved are counted more in the average.

A common use of weighted averages is in the computation of a student's GPA, where grades are weighted according to credit hours. Let's consider the scenario in Table 6.3.3.

class	grade	grade points	credits
chemistry	B+	3.3	5
calculus	A-	3.7	4
history	B-	2.7	3
psychology	B-	2.7	3

Table 6.3.3: A college student's semester grades.

If all of the classes were of the same weight (i.e., the same number of credits), the student's GPA would simply be calculated by taking the average

$$\frac{3.3 + 3.7 + 2.7 + 2.7}{4} = 3.1.$$

But since the chemistry and calculus courses have higher weights (of 5 and 4 credits respectively), we actually compute the GPA according to the weighted average

$$\frac{3.3 \cdot 5 + 3.7 \cdot 4 + 2.7 \cdot 3 + 2.7 \cdot 3}{5 + 4 + 3 + 3} = 3.1\overline{6}.$$

The weighted average reflects the fact that chemistry and calculus, as courses with higher credits, have a greater impact on the students' grade point average. Note particularly that in the weighted average, each grade gets multiplied by its weight, and we divide by the sum of the weights.

In the following activity, we explore further how weighted averages can be used to find the balancing point of a physical system.

Activity 6.3.3. For quantities of equal weight, such as two children on a teeter-totter, the balancing point is found by taking the average of their locations. When the weights of the quantities differ, we use a weighted average of their respective locations to find the balancing point.

- a. Suppose that a shelf is 6 feet long, with its left end situated at x = 0. If one book of weight 1 lb is placed at $x_1 = 0$, and another book of weight 1 lb is placed at $x_2 = 6$, what is the location of \overline{x} , the point at which the shelf would (theoretically) balance on a fulcrum?
- b. Now, say that we place four books on the shelf, each weighing 1 lb: at $x_1 = 0$, at $x_2 = 2$, at $x_3 = 4$, and at $x_4 = 6$. Find \overline{x} , the balancing point of the shelf.
- c. How does \overline{x} change if we change the location of the third book? Say the locations of the 1-lb books are $x_1 = 0$, $x_2 = 2$, $x_3 = 3$, and $x_4 = 6$.
- d. Next, suppose that we place four books on the shelf, but of varying weights: at $x_1 = 0$ a 2-lb book, at $x_2 = 2$ a 3-lb book, at $x_3 = 4$ a 1-lb book, and at $x_4 = 6$ a 1-lb book. Use a weighted average of the locations to find \overline{x} , the balancing point of the shelf. How does the balancing point in this scenario compare to that found in (b)?
- e. What happens if we change the location of one of the books? Say that we keep everything the same in (d), except that $x_3 = 5$. How does \overline{x} change?
- f. What happens if we change the weight of one of the books? Say that we keep everything the same in (d), except that the book at $x_3 = 4$ now weighs 2 lbs. How does \overline{x} change?
- g. Experiment with a couple of different scenarios of your choosing where you move one of the books to the left, or you decrease the weight of one of the books.
- h. Write a couple of sentences to explain how adjusting the location of one of the books or the weight of one of the books affects the location of the balancing point of the shelf. Think carefully here about how your changes should be considered relative to the location of the balancing point \overline{x} of the current scenario.

6.3.3 Center of Mass

In Activity 6.3.3, we saw that the balancing point of a system of point-masses¹ (such as books on a shelf) is found by taking a weighted average of their respective locations. In the activity, we were computing the *center of mass* of a system of masses distributed along an axis, which is the balancing point of the axis on which the masses rest.

Center of Mass (point-masses)

For a collection of n masses $m_1, ..., m_n$ that are distributed along a single axis at the locations $x_1, ..., x_n$, the *center of mass* is given by

$$\overline{x} = \frac{x_1 m_1 + x_2 m_2 + \dots + x_n m_n}{m_1 + m_2 + \dots + m_n}.$$

What if we instead consider a thin bar over which density is distributed continuously? If the density is constant, it is obvious that the balancing point of the bar is its midpoint. But if density is not constant, we must compute a weighted average. Let's say that the function $\rho(x)$ tells us the density distribution along the bar, measured in g/cm. If we slice the bar into small sections, this enables us to think of the bar as holding a collection of adjacent point-masses. For a slice of thickness Δx at location x_i , note that the mass of the slice, m_i , satisfies $m_i \approx \rho(x_i)\Delta x$.

Taking n slices of the bar, we can approximate its center of mass by

$$\overline{x} \approx \frac{x_1 \cdot \rho(x_1) \Delta x + x_2 \cdot \rho(x_2) \Delta x + \dots + x_n \cdot \rho(x_n) \Delta x}{\rho(x_1) \Delta x + \rho(x_2) \Delta x + \dots + \rho(x_n) \Delta x}.$$

Rewriting the sums in sigma notation, it follows that

$$\overline{x} \approx \frac{\sum_{i=1}^{n} x_i \cdot \rho(x_i) \Delta x}{\sum_{i=1}^{n} \rho(x_i) \Delta x}.$$
(6.3.1)

Moreover, it is apparent that the greater the number of slices, the more accurate our estimate of the balancing point will be, and that the sums in Equation (6.3.1) can be viewed as Riemann sums. Hence, in the limit as $n \to \infty$, we find that the center of mass is given by the quotient of two integrals.

Center of Mass (continuous mass distribution)

For a thin rod of density $\rho(x)$ distributed along an axis from x = a to x = b, the center of mass of the rod is given by

$$\overline{x} = \frac{\int_a^b x \rho(x) \, dx}{\int_a^b \rho(x) \, dx}.$$

¹In the activity, we actually used *weight* rather than *mass*. Since weight is proportional to mass, the computations for the balancing point result in the same location regardless of whether we use weight or mass. The gravitational constant is present in both the numerator and denominator of the weighted average.

Note particularly that the denominator of \overline{x} is the mass of the bar, and that this quotient of integrals is simply the continuous version of the weighted average of locations, x, along the bar.

Activity 6.3.4. Consider a thin bar of length 20 cm whose density is distributed according to the function $\rho(x) = 4 + 0.1x$, where x = 0 represents the left end of the bar. Assume that ρ is measured in g/cm and x is measured in cm.

- a. Find the total mass, *M*, of the bar.
- b. Without doing any calculations, do you expect the center of mass of the bar to be equal to 10, less than 10, or greater than 10? Why?
- c. Compute \overline{x} , the exact center of mass of the bar.
- d. What is the average density of the bar?
- e. Now consider a different density function, given by $p(x) = 4e^{0.020732x}$, also for a bar of length 20 cm whose left end is at x = 0. Plot both $\rho(x)$ and $\rho(x)$ on the same axes. Without doing any calculations, which bar do you expect to have the greater center of mass? Why?
- f. Compute the exact center of mass of the bar described in (e) whose density function is $p(x) = 4e^{0.020732x}$. Check the result against the prediction you made in (e).

Summary

- For an object of constant density D, with volume V and mass m, we know that $m = D \cdot V$.
- If an object with constant cross-sectional area (such as a thin bar) has its density distributed along an axis according to the function $\rho(x)$, then we can find the mass of the object between x = a and x = b by

$$m = \int_a^b \rho(x) \, dx.$$

• For a system of point-masses distributed along an axis, say m_1, \ldots, m_n at locations x_1, \ldots, x_n , the center of mass, \overline{x} , is given by the weighted average

$$\overline{x} = \frac{\sum_{i=1}^{n} x_i m_i}{\sum_{i=1}^{n} m_i}.$$

If instead we have mass continuously distributed along an axis, such as by a density function $\rho(x)$ for a thin bar of constant cross-sectional area, the center of mass of the portion of the bar between x = a and x = b is given by

$$\overline{x} = \frac{\int_a^b x \rho(x) \, dx}{\int_a^b \rho(x) \, dx}.$$

In each situation, \bar{x} represents the balancing point of the system of masses or of the portion of the bar.

Exercises

	極
ı	WeBWork

1. A rod has length 3 meters. At a distance <i>x</i> meters from its left end, the density of the rod
is given by $\delta(x) = 4 + 5x$ g/m.
(a) Complete the Riemann sum for the total mass of the rod (use Dx in place of Δx):
$mass = \Sigma$
(b) Convert the Riemann sum to an integral and find the exact mass.
mass = (in all do unito)
(include units)

極
Helliork

2. A ro	d with de	nsity $\delta(x) = 7 + \sin(x)$ lies on the x-axis between $x = 0$ and $x = \pi$. Find the	he
mass ar	nd center	of mass of the rod.	
mass =			
center c	of mass =		



3. Suppose that the density of cars (in cars per mile) down a 20-mile stretch of the Pennsylvania Turnpike is approximated by $\delta(x) = 350 \left(2 + \sin\left(4\sqrt{x} + 0.175\right)\right)$, at a distance x miles from the Breezewood toll plaza. Sketch a graph of this function for $0 \le x \le 20$.

(a) Complete the Riemann sum that approximates the total number of cars on this 20-mile stretch (use Dx instead of Δx):

Number = Σ		
(b) Find the	total number of cars on the 20-mile stretch.	
Number =		



4. A point mass of 1 grams located 6 centimeters to the left of the origin and a point mass of 3 grams located 7 centimeters to the right of the origin are connected by a thin, light rod. Find the center of mass of the system.

```
Center of Mass = [Choose: to the left of the origin | to the right of the origin | (at the origin)] (include units in your center of mass)
```

- **5.** Let a thin rod of length a have density distribution function $\rho(x) = 10e^{-0.1x}$, where x is measured in cm and ρ in grams per centimeter.
 - a. If the mass of the rod is 30 g, what is the value of *a*?
 - b. For the 30g rod, will the center of mass lie at its midpoint, to the left of the midpoint, or to the right of the midpoint? Why?
 - c. For the 30g rod, find the center of mass, and compare your prediction in (b).
 - d. At what value of *x* should the 30g rod be cut in order to form two pieces of equal mass?

- **6.** Consider two thin bars of constant cross-sectional area, each of length 10 cm, with respective mass density functions $\rho(x) = \frac{1}{1+x^2}$ and $p(x) = e^{-0.1x}$.
 - a. Find the mass of each bar.
 - b. Find the center of mass of each bar.
 - c. Now consider a new 10 cm bar whose mass density function is $f(x) = \rho(x) + p(x)$.
 - a. Explain how you can easily find the mass of this new bar with little to no additional work.
 - b. Similarly, compute $\int_0^{10} x f(x) dx$ as simply as possible, in light of earlier computations.
 - c. True or false: the center of mass of this new bar is the average of the centers of mass of the two earlier bars. Write at least one sentence to say why your conclusion makes sense.
- 7. Consider the curve given by $y = f(x) = 2xe^{-1.25x} + (30 x)e^{-0.25(30 x)}$.
 - a. Plot this curve in the window x = 0...30, y = 0...3 (with constrained scaling so the units on the x and y axis are equal), and use it to generate a solid of revolution about the x-axis. Explain why this curve could generate a reasonable model of a baseball bat.
 - b. Let *x* and *y* be measured in inches. Find the total volume of the baseball bat generated by revolving the given curve about the *x*-axis. Include units on your answer
 - c. Suppose that the baseball bat has constant weight density, and that the weight density is 0.6 ounces per cubic inch. Find the total weight of the bat whose volume you found in (b).
 - d. Because the baseball bat does not have constant cross-sectional area, we see that the amount of weight concentrated at a location x along the bat is determined by the volume of a slice at location x. Explain why we can think about the function $\rho(x) = 0.6\pi f(x)^2$ (where f is the function given at the start of the problem) as being the weight density function for how the weight of the baseball bat is distributed from x = 0 to x = 30.
 - e. Compute the center of mass of the baseball bat.

6.4 Physics Applications: Work, Force, and Pressure

Motivating Questions

- How do we measure the work accomplished by a varying force that moves an object a certain distance?
- What is the total force exerted by water against a dam?
- How are both of the above concepts and their corresponding use of definite integrals similar to problems we have encountered in the past involving formulas such as "distance equals rate times time" and "mass equals density times volume"?

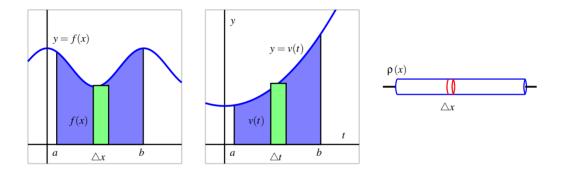


Figure 6.4.1: Three settings where we compute the accumulation of a varying quantity: the area under y = f(x), the distance traveled by an object with velocity y = v(t), and the mass of a bar with density function $y = \rho(x)$.

In our work to date with the definite integral, we have seen several different circumstances where the integral enables us to measure the accumulation of a quantity that varies, provided the quantity is approximately constant over small intervals. For instance, based on the fact that the area of a rectangle is $A = l \cdot w$, if we wish to find the area bounded by a nonnegative curve y = f(x) and the x-axis on an interval [a,b], a representative slice of width Δx has area $A_{\rm slice} = f(x)\Delta x$, and thus as we let the width of the representative slice tend to zero, we find that the exact area of the region is

$$A = \int_a^b f(x) \, dx.$$

In a similar way, if we know that the velocity of a moving object is given by the function y = v(t), and we wish to know the distance the object travels on an interval [a, b] where v(t)

is nonnegative, we can use a definite integral to generalize the fact that $d=r\cdot t$ when the rate, r, is constant. More specifically, on a short time interval Δt , v(t) is roughly constant, and hence for a small slice of time, $d_{\rm slice}=v(t)\Delta t$, and so as the width of the time interval Δt tends to zero, the exact distance traveled is given by the definite integral

$$d = \int_a^b v(t) \, dt.$$

Finally, when we recently learned about the mass of an object of non-constant density, we saw that since $M=D\cdot V$ (mass equals density times volume, provided that density is constant), if we can consider a small slice of an object on which the density is approximately constant, a definite integral may be used to determine the exact mass of the object. For instance, if we have a thin rod whose cross sections have constant density, but whose density is distributed along the x axis according to the function $y=\rho(x)$, it follows that for a small slice of the rod that is Δx thick, $M_{\rm slice}=\rho(x)\Delta x$. In the limit as $\Delta x\to 0$, we then find that the total mass is given by

$$M = \int_a^b \rho(x) \, dx.$$

Note that all three of these situations are similar in that we have a basic rule $(A = l \cdot w, d = r \cdot t, M = D \cdot V)$ where one of the two quantities being multiplied is no longer constant; in each, we consider a small interval for the other variable in the formula, calculate the approximate value of the desired quantity (area, distance, or mass) over the small interval, and then use a definite integral to sum the results as the length of the small intervals is allowed to approach zero. It should be apparent that this approach will work effectively for other situations where we have a quantity of interest that varies.

We next turn to the notion of *work*: from physics, a basic principal is that work is the product of force and distance. For example, if a person exerts a force of 20 pounds to lift a 20-pound weight 4 feet off the ground, the total work accomplished is

$$W = F \cdot d = 20 \cdot 4 = 80$$
 foot-pounds.

If force and distance are measured in English units (pounds and feet), then the units on work are *foot-pounds*. If instead we work in metric units, where forces are measured in Newtons and distances in meters, the units on work are *Newton-meters*.

Of course, the formula $W = F \cdot d$ only applies when the force is constant while it is exerted over the distance d. In Preview Activity 6.4.1, we explore one way that we can use a definite integral to compute the total work accomplished when the force exerted varies.

Preview Activity 6.4.1. A bucket is being lifted from the bottom of a 50-foot deep well; its weight (including the water), B, in pounds at a height h feet above the water is given by the function B(h). When the bucket leaves the water, the bucket and water together weigh B(0) = 20 pounds, and when the bucket reaches the top of the well, B(50) = 12 pounds. Assume that the bucket loses water at a constant rate (as a function of height,

- *h*) throughout its journey from the bottom to the top of the well.
 - a. Find a formula for B(h).
 - b. Compute the value of the product $B(5)\Delta h$, where $\Delta h=2$ feet. Include units on your answer. Explain why this product represents the approximate work it took to move the bucket of water from h=5 to h=7.
 - c. Is the value in (b) an over- or under-estimate of the actual amount of work it took to move the bucket from h = 5 to h = 7? Why?
 - d. Compute the value of the product $B(22)\Delta h$, where $\Delta h = 0.25$ feet. Include units on your answer. What is the meaning of the value you found?
 - e. More generally, what does the quantity $W_{\text{slice}} = B(h)\Delta h$ measure for a given value of h and a small positive value of h?
 - f. Evaluate the definite integral $\int_0^{50} B(h) dh$. What is the meaning of the value you find? Why?

6.4.1 Work

Because work is calculated by the rule $W = F \cdot d$, whenever the force F is constant, it follows that we can use a definite integral to compute the work accomplished by a varying force. For example, suppose that in a setting similar to the problem posed in Preview Activity 6.4.1, we have a bucket being lifted in a 50-foot well whose weight at height h is given by $B(h) = 12 + 8e^{-0.1h}$.

In contrast to the problem in the preview activity, this bucket is not leaking at a constant rate; but because the weight of the bucket and water is not constant, we have to use a definite integral to determine the total work that results from lifting the bucket. Observe that at a height h above the water, the approximate work to move the bucket a small distance Δh is

$$W_{\text{slice}} = B(h)\Delta h = (12 + 8e^{-0.1h})\Delta h.$$

Hence, if we let Δh tend to 0 and take the sum of all of the slices of work accomplished on these small intervals, it follows that the total work is given by

$$W = \int_0^{50} B(h) \, dh = \int_0^{50} (12 + 8e^{-0.1h}) \, dh.$$

While is a straightforward exercise to evaluate this integral exactly using the First Fundamental Theorem of Calculus, in applied settings such as this one we will typically use computing technology to find accurate approximations of integrals that are of interest to us. Here, it turns out that $W = \int_0^{50} (12 + 8e^{-0.1h}) \, dh \approx 679.461$ foot-pounds.

Our work in Preview Activity 6.4.1 and in the most recent example above employs the following important general principle.

For an object being moved in the positive direction along an axis, x, by a force F(x), the total work to move the object from a to b is given by

$$W = \int_a^b F(x) \, dx.$$

Activity 6.4.2. Consider the following situations in which a varying force accomplishes work.

- a. Suppose that a heavy rope hangs over the side of a cliff. The rope is 200 feet long and weighs 0.3 pounds per foot; initially the rope is fully extended. How much work is required to haul in the entire length of the rope? (Hint: set up a function F(h) whose value is the weight of the rope remaining over the cliff after h feet have been hauled in.)
- b. A leaky bucket is being hauled up from a 100 foot deep well. When lifted from the water, the bucket and water together weigh 40 pounds. As the bucket is being hauled upward at a constant rate, the bucket leaks water at a constant rate so that it is losing weight at a rate of 0.1 pounds per foot. What function B(h) tells the weight of the bucket after the bucket has been lifted h feet? What is the total amount of work accomplished in lifting the bucket to the top of the well?
- c. Now suppose that the bucket in (b) does not leak at a constant rate, but rather that its weight at a height h feet above the water is given by $B(h) = 25 + 15e^{-0.05h}$. What is the total work required to lift the bucket 100 feet? What is the average force exerted on the bucket on the interval h = 0 to h = 100?
- d. From physics, $Hooke's\ Law$ for springs states that the amount of force required to hold a spring that is compressed (or extended) to a particular length is proportionate to the distance the spring is compressed (or extended) from its natural length. That is, the force to compress (or extend) a spring x units from its natural length is F(x) = kx for some constant k (which is called the *spring constant*.) For springs, we choose to measure the force in pounds and the distance the spring is compressed in feet. Suppose that a force of 5 pounds extends a particular spring 4 inches (1/3 foot) beyond its natural length.
 - i. Use the given fact that F(1/3) = 5 to find the spring constant k.
 - ii. Find the work done to extend the spring from its natural length to 1 foot beyond its natural length.
 - iii. Find the work required to extend the spring from 1 foot beyond its natural length to 1.5 feet beyond its natural length.

6.4.2 Work: Pumping Liquid from a Tank

In certain geographic locations where the water table is high, residential homes with basements have a peculiar feature: in the basement, one finds a large hole in the floor, and in the hole, there is water. For example, in Figure 6.4.2 we see a *sump crock*¹. Essentially, a sump crock provides an outlet for water that may build up beneath the basement floor; of course, as that water rises, it is imperative that the water not flood the basement.

Hence, in the crock we see the presence of a floating pump that sits on the surface of the water: this pump is activated by elevation, so when the water level reaches a particular height, the pump turns on and pumps a certain portion of the water out of the crock, hence relieving the water buildup beneath the foundation. One of the questions we'd like to answer is: how much work does a sump pump accomplish?



Figure 6.4.2: A sump crock.

To that end, let's suppose that we have a sump crock that has the shape of a frustum of a cone, as pictured in Figure 6.4.3. Assume that the crock has a diameter of 3 feet at its surface, a diameter of 1.5 feet at its base, and a depth of 4 feet. In addition, suppose that the sump pump is set up so that it pumps the water vertically up a pipe to a drain that is located at ground level just outside a basement window. To accomplish this, the pump must send the water to a location 9 feet above the surface of the sump crock.

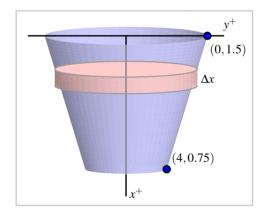


Figure 6.4.3: A sump crock with approximately cylindrical cross-sections that is 4 feet deep, 1.5 feet in diameter at its base, and 3 feet in diameter at its top.

It turns out to be advantageous to think of the depth below the surface of the crock as being the independent variable, so, in problems such as this one we typically let the positive x-axis

¹Image credit to www.warreninspect.com/basement-moisture.

point down, and the positive *y*-axis to the right, as pictured in the figure. As we think about the work that the pump does, we first realize that the pump sits on the surface of the water, so it makes sense to think about the pump moving the water one "slice" at a time, where it takes a thin slice from the surface, pumps it out of the tank, and then proceeds to pump the next slice below.

For the sump crock described in this example, each slice of water is cylindrical in shape. We see that the radius of each approximately cylindrical slice varies according to the linear function y = f(x) that passes through the points (0, 1.5) and (4, 0.75), where x is the depth of the particular slice in the tank; it is a straightforward exercise to find that f(x) = 1.5 - 0.1875x. Now we are prepared to think about the overall problem in several steps: (a) determining the volume of a typical slice; (b) finding the weight² of a typical slice (and thus the force that must be exerted on it); (c) deciding the distance that a typical slice moves; and (d) computing the work to move a representative slice. Once we know the work it takes to move one slice, we use a definite integral over an appropriate interval to find the total work.

Consider a representative cylindrical slice that sits on the surface of the water at a depth of *x* feet below the top of the crock. It follows that the approximate volume of that slice is given by

$$V_{\text{slice}} = \pi f(x)^2 \Delta x = \pi (1.5 - 0.1875x)^2 \Delta x.$$

Since water weighs 62.4 lb/ft³, it follows that the approximate weight of a representative slice, which is also the approximate force the pump must exert to move the slice, is

$$F_{\text{slice}} = 62.4 \cdot V_{\text{slice}} = 62.4\pi (1.5 - 0.1875x)^2 \Delta x.$$

Because the slice is located at a depth of x feet below the top of the crock, the slice being moved by the pump must move x feet to get to the level of the basement floor, and then, as stated in the problem description, be moved another 9 feet to reach the drain at ground level outside a basement window. Hence, the total distance a representative slice travels is

$$d_{\text{slice}} = x + 9.$$

Finally, we note that the work to move a representative slice is given by

$$W_{\text{slice}} = F_{\text{slice}} \cdot d_{\text{slice}} = 62.4\pi (1.5 - 0.1875x)^2 \Delta x \cdot (x + 9),$$

since the force to move a particular slice is constant.

We sum the work required to move slices throughout the tank (from x = 0 to x = 4), let $\Delta x \rightarrow 0$, and hence

$$W = \int_0^4 62.4\pi (1.5 - 0.1875x)^2 (x+9) \, dx,$$

which, when evaluated using appropriate technology, shows that the total work is $W = 10970.5\pi$ foot-pounds.

The preceding example demonstrates the standard approach to finding the work required to empty a tank filled with liquid. The main task in each such problem is to determine

²We assume that the weight density of water is 62.4 pounds per cubic foot.

the volume of a representative slice, followed by the force exerted on the slice, as well as the distance such a slice moves. In the case where the units are metric, there is one key difference: in the metric setting, rather than weight, we normally first find the mass of a slice. For instance, if distance is measured in meters, the mass density of water is 1000 kg/m^3 . In that setting, we can find the mass of a typical slice (in kg). To determine the force required to move it, we use F = ma, where m is the object's mass and a is the gravitational constant 9.81 N/kg^3 . That is, in metric units, the weight density of water is 9810 N/m^3 .

Activity 6.4.3. In each of the following problems, determine the total work required to accomplish the described task. In parts (b) and (c), a key step is to find a formula for a function that describes the curve that forms the side boundary of the tank.

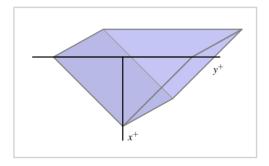


Figure 6.4.4: A trough with triangular ends, as described in Activity 6.4.3, part (c).

- a. Consider a vertical cylindrical tank of radius 2 meters and depth 6 meters. Suppose the tank is filled with 4 meters of water of mass density 1000 kg/m^3 , and the top 1 meter of water is pumped over the top of the tank.
- b. Consider a hemispherical tank with a radius of 10 feet. Suppose that the tank is full to a depth of 7 feet with water of weight density 62.4 pounds/ft^3 , and the top 5 feet of water are pumped out of the tank to a tanker truck whose height is 5 feet above the top of the tank.
- c. Consider a trough with triangular ends, as pictured in Figure 6.4.4, where the tank is 10 feet long, the top is 5 feet wide, and the tank is 4 feet deep. Say that the trough is full to within 1 foot of the top with water of weight density 62.4 pounds/ft 3 , and a pump is used to empty the tank until the water remaining in the tank is 1 foot deep.

6.4.3 Force due to Hydrostatic Pressure

When a dam is built, it is imperative to for engineers to understand how much force water will exert against the face of the dam. The first thing we realize is the force exerted by the fluid is related to the natural concept of pressure. The pressure a force exerts on a region is measured in units of force per unit of area: for example, the air pressure in a tire is often measured in pounds per square inch (PSI). Hence, we see that the general relationship is given by

$$P = \frac{F}{A}$$
, or $F = P \cdot A$,

where P represents pressure, F represents force, and A the area of the region being considered. Of course, in the equation F = PA, we assume that the pressure is constant over the entire region A.

Most people know from experience that the deeper one dives underwater while swimming, the greater the pressure that is exerted by the water. This is due to the fact that the deeper one dives, the more water there is right on top of the swimmer: it is the force that "column" of water exerts that determines the pressure the swimmer experiences. To get water pressure measured in its standard units (pounds per square foot), we say that the total water pressure is found by computing the total weight of the column of water that lies above a region of area 1 square foot at a fixed depth. Such a rectangular column with a 1×1 base and a depth of d feet has volume $V = 1 \cdot 1 \cdot d$ ft³, and thus the corresponding weight of the water overhead is 62.4d. Since this is also the amount of force being exerted on a 1 square foot region at a depth d feet underwater, we see that P = 62.4d (lbs/ft²) is the pressure exerted by water at depth d.

The understanding that P = 62.4d will tell us the pressure exerted by water at a depth of d, along with the fact that F = PA, will now enable us to compute the total force that water exerts on a dam, as we see in the following example.

Example 6.4.5. Consider a trapezoid-shaped dam that is 60 feet wide at its base and 90 feet wide at its top, and assume the dam is 25 feet tall with water that rises to within 5 feet of the top of its face. Water weighs 62.4 pounds per cubic foot. How much force does the water exert against the dam?

Solution. First, we sketch a picture of the dam, as shown in Figure 6.4.6. Note that, as in problems involving the work to pump out a tank, we let the positive x-axis point down.

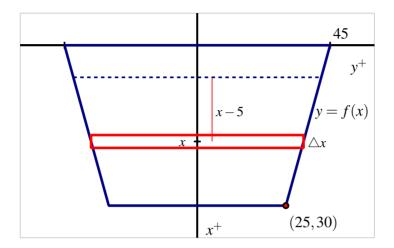


Figure 6.4.6: A trapezoidal dam that is 25 feet tall, 60 feet wide at its base, 90 feet wide at its top, with the water line 5 feet down from the top of its face.

It is essential to use the fact that pressure is constant at a fixed depth. Hence, we consider a slice of water at constant depth on the face, such as the one shown in the figure. First, the approximate area of this slice is the area of the pictured rectangle. Since the width of that rectangle depends on the variable x (which represents the how far the slice lies from the top of the dam), we find a formula for the function y = f(x) that determines one side of the face of the dam. Since f is linear, it is straightforward to find that $y = f(x) = 45 - \frac{3}{5}x$. Hence, the approximate area of a representative slice is

$$A_{\text{slice}} = 2f(x)\Delta x = 2(45 - \frac{3}{5}x)\Delta x.$$

At any point on this slice, the depth is approximately constant, and thus the pressure can be considered constant. In particular, we note that since x measures the distance to the top of the dam, and because the water rises to within 5 feet of the top of the dam, the depth of any point on the representative slice is approximately (x - 5). Now, since pressure is given by P = 62.4d, we have that at any point on the representative slice

$$P_{\text{slice}} = 62.4(x - 5).$$

Knowing both the pressure and area, we can find the force the water exerts on the slice. Using F = PA, it follows that

$$F_{\text{slice}} = P_{\text{slice}} \cdot A_{\text{slice}} = 62.4(x - 5) \cdot 2(45 - \frac{3}{5}x)\Delta x.$$

Finally, we use a definite integral to sum the forces over the appropriate range of x-values. Since the water rises to within 5 feet of the top of the dam, we start at x = 5 and slice all the

way to the bottom of the dam, where x = 30. Hence,

$$F = \int_{x=5}^{x=30} 62.4(x-5) \cdot 2(45 - \frac{3}{5}x) \, dx.$$

Using technology to evaluate the integral, we find $F \approx 1.248 \times 10^6$ pounds.

Activity 6.4.4. In each of the following problems, determine the total force exerted by water against the surface that is described.

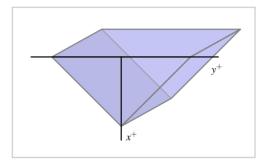


Figure 6.4.7: A trough with triangular ends, as described in Activity 6.4.4, part (c).

- a. Consider a rectangular dam that is 100 feet wide and 50 feet tall, and suppose that water presses against the dam all the way to the top.
- b. Consider a semicircular dam with a radius of 30 feet. Suppose that the water rises to within 10 feet of the top of the dam.
- c. Consider a trough with triangular ends, as pictured in Figure 6.4.7, where the tank is 10 feet long, the top is 5 feet wide, and the tank is 4 feet deep. Say that the trough is full to within 1 foot of the top with water of weight density 62.4 pounds/ft³. How much force does the water exert against one of the triangular ends?

While there are many different formulas that we use in solving problems involving work, force, and pressure, it is important to understand that the fundamental ideas behind these problems are similar to several others that we've encountered in applications of the definite integral. In particular, the basic idea is to take a difficult problem and somehow slice it into more manageable pieces that we understand, and then use a definite integral to add up these simpler pieces.

Summary

- To measure the work accomplished by a varying force that moves an object, we subdivide the problem into pieces on which we can use the formula $W = F \cdot d$, and then use a definite integral to sum the work accomplished on each piece.
- To find the total force exerted by water against a dam, we use the formula $F = P \cdot A$ to measure the force exerted on a slice that lies at a fixed depth, and then use a definite integral to sum the forces across the appropriate range of depths.
- Because work is computed as the product of force and distance (provided force is constant), and the force water exerts on a dam can be computed as the product of pressure and area (provided pressure is constant), problems involving these concepts are similar to earlier problems we did using definite integrals to find distance (via "distance equals rate times time") and mass ("mass equals density times volume").

Exercises

JEBWORK	1. A tank in the shape of an inverted right circular cone has height 4 meters and radius 2 meters. It is filled with 2 meters of hot chocolate. Find the work required to empty the tank by pumping the hot chocolate over the top of the tank. The density of hot chocolate is $\delta = 1080 \text{ kg/m}^3$. Your answer must include the correct units.
	Work =
JeBWork	2. A fuel oil tank is an upright cylinder, buried so that its circular top is 10 feet beneath ground level. The tank has a radius of 7 feet and is 21 feet high, although the current oil level is only 17 feet deep. Calculate the work required to pump all of the oil to the surface. Oil weighs 50lb/ft^3 .
	Work =
	(include units)
JeBWork	3. A rectangular swimming pool 40 ft long, 15 ft wide, and 16 ft deep is filled with water to a depth of 15 ft. Use an integral to find the work required to pump all the water out over the top. (Take as the density of water $\delta = 62.4 \text{lb/ft}^3$.)
	Work =
	(include units)
JeBWork	4. Water in a cylinder of height 10 ft and radius 3 ft is to be pumped out. The density of water is 62.4 lb/ft ³ . Find the work required if (a) The tank is full of water and the water is to be pumped over the top of the tank.
	Work =
	(include units)
	(b) The tank is full of water and the water must be pumped to a height 4 ft above the top of
	the tank.
	Work =

(include ur	nits)	
(c) The dep	oth of water in the tank is 8 ft and the water must be pumped over the top of th	ıe
tank		
Work =		
(include ur	nits)	

5. A lobster tank in a restaurant is 1.25 m long by 1 m wide by 80 cm deep. Taking the density of water to be 1000kg/m^3 , find the water forces

density of water to be 1000kg/m ³ , find the water	er forces
on the bottom of the tank: Force =	
on each of the larger sides of the tank: Force =	
on each of the smaller sides of the tank: Force =	

6. Consider the curve $f(x) = 3\cos(\frac{x^3}{4})$ and the portion of its graph that lies in the first quadrant between the *y*-axis and the first positive value of *x* for which f(x) = 0. Let *R* denote the region bounded by this portion of *f*, the *x*-axis, and the *y*-axis. Assume that *x* and *y* are each measured in feet.

(include units for each, and use $q = 9.8 \text{ m/s}^2$)

- a. Picture the coordinate axes rotated 90 degrees clockwise so that the positive x-axis points straight down, and the positive y-axis points to the right. Suppose that R is rotated about the x axis to form a solid of revolution, and we consider this solid as a storage tank. Suppose that the resulting tank is filled to a depth of 1.5 feet with water weighing 62.4 pounds per cubic foot. Find the amount of work required to lower the water in the tank until it is 0.5 feet deep, by pumping the water to the top of the tank.
- b. Again picture the coordinate axes rotated 90 degrees clockwise so that the positive *x*-axis points straight down, and the positive *y*-axis points to the right. Suppose that *R*, together with its reflection across the *x*-axis, forms one end of a storage tank that is 10 feet long. Suppose that the resulting tank is filled completely with water weighing 62.4 pounds per cubic foot. Find a formula for a function that tells the amount of work required to lower the water by *h* feet.
- c. Suppose that the tank described in (b) is completely filled with water. Find the total force due to hydrostatic pressure exerted by the water on one end of the tank.
- 7. A cylindrical tank, buried on its side, has radius 3 feet and length 10 feet. It is filled completely with water whose weight density is 62.4 lbs/ft³, and the top of the tank is two feet underground.
 - a. Set up, but do not evaluate, an integral expression that represents the amount of work required to empty the top half of the water in the tank to a truck whose tank lies 4.5 feet above ground.
 - b. With the tank now only half-full, set up, but do not evaluate an integral expression that represents the total force due to hydrostatic pressure against one end of the tank.



6.5 Improper Integrals

Motivating Questions

- What are improper integrals and why are they important?
- What does it mean to say that an improper integral converges or diverges?
- What are some typical improper integrals that we can classify as convergent or divergent?

Another important application of the definite integral regards how the likelihood of certain events can be measured. For example, consider a company that manufactures incandescent light bulbs, and suppose that based on a large volume of test results, they have determined that the fraction of light bulbs that fail between times t = a and t = b of use (where t is measured in months) is given by

$$\int_{a}^{b} 0.3e^{-0.3t} dt$$
.

For example, the fraction of light bulbs that fail during their third month of use is given by

$$\int_{2}^{3} 0.3e^{-0.3t} dt = -e^{-0.3t} \Big|_{2}^{3}$$
$$= -e^{-0.9} + e^{-0.6}$$
$$\approx 0.1422.$$

Thus about 14.22% of all lightbulbs fail between t = 2 and t = 3. Clearly we could adjust the limits of integration to measure the fraction of light bulbs that fail during any time period of interest.

Preview Activity 6.5.1. A company with a large customer base has a call center that receives thousands of calls a day. After studying the data that represents how long callers wait for assistance, they find that the function $p(t) = 0.25e^{-0.25t}$ models the time customers wait in the following way: the fraction of customers who wait between t = a and t = b minutes is given by

$$\int_a^b p(t)\,dt.$$

Use this information to answer the following questions.

- a. Determine the fraction of callers who wait between 5 and 10 minutes.
- b. Determine the fraction of callers who wait between 10 and 20 minutes.
- c. Next, let's study the fraction who wait up to a certain number of minutes:

- i. What is the fraction of callers who wait between 0 and 5 minutes?
- ii. What is the fraction of callers who wait between 0 and 10 minutes?
- iii. Between 0 and 15 minutes? Between 0 and 20?
- d. Let F(b) represent the fraction of callers who wait between 0 and b minutes. Find a formula for F(b) that involves a definite integral, and then use the First FTC to find a formula for F(b) that does not involve a definite integral.
- e. What is the value of the limit $\lim_{b\to\infty} F(b)$? What is its meaning in the context of the problem?

6.5.1 Improper Integrals Involving Unbounded Intervals

In light of our example with light bulbs that fail, as well as with the problem involving customer wait time in Preview Activity 6.5.1, we see that it is natural to consider questions where we desire to integrate over an interval whose upper limit grows without bound. For example, if we are interested in the fraction of light bulbs that fail within the first b months of use, we know that the expression

$$\int_0^b 0.3e^{-0.3t} \, dt$$

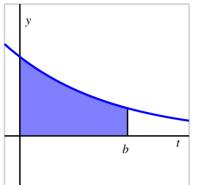
measures this value. To think about the fraction of light bulbs that fail *eventually*, we understand that we wish to find

$$\lim_{b \to \infty} \int_0^b 0.3e^{-0.3t} \, dt,$$

for which we will also use the notation

$$\int_0^\infty 0.3e^{-0.3t} dt. \tag{6.5.1}$$

Note particularly that we are studying the area of an unbounded region, as pictured in Figure 6.5.1.



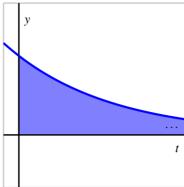


Figure 6.5.1: At left, the area bounded by $p(t) = 0.3e^{-0.3t}$ on the finite interval [0, b]; at right, the result of letting $b \to \infty$. By "···" in the righthand figure, we mean that the region extends to the right without bound.

Anytime we are interested in an integral for which the interval of integration is unbounded (that is, one for which at least one of the limits of integration involves ∞), we say that the integral is *improper*. For instance, the integrals

$$\int_{1}^{\infty} \frac{1}{x^2} dx$$
, $\int_{-\infty}^{0} \frac{1}{1+x^2} dx$, and $\int_{-\infty}^{\infty} e^{-x^2} dx$

are all improper due to having limits of integration that involve ∞ . We investigate the value of any such integral by replacing the improper integral with a limit of proper integrals; for an improper integral such as $\int_0^\infty f(x) \, dx$, we write

$$\int_0^\infty f(x) \, dx = \lim_{b \to \infty} \int_0^b f(x) \, dx.$$

We can then attempt to evaluate $\int_0^b f(x) dx$ using the First FTC, after which we can evaluate the limit. An immediate and important question arises: is it even possible for the area of such an unbounded region to be finite? The following activity explores this issue and others in more detail.

Activity 6.5.2. In this activity we explore the improper integrals $\int_1^\infty \frac{1}{x} dx$ and $\int_1^\infty \frac{1}{x^{3/2}} dx$.

- a. First we investigate $\int_1^\infty \frac{1}{x} dx$.
 - i. Use the First FTC to determine the exact values of $\int_1^{10} \frac{1}{x} dx$, $\int_1^{1000} \frac{1}{x} dx$, and $\int_1^{100000} \frac{1}{x} dx$. Then, use your calculator to compute a decimal approximation of each result

- ii. Use the First FTC to evaluate the definite integral $\int_1^b \frac{1}{x} dx$ (which results in an expression that depends on b).
- iii. Now, use your work from (ii.) to evaluate the limit given by

$$\lim_{b\to\infty}\int_1^b\frac{1}{x}\,dx.$$

- b. Next, we investigate $\int_1^\infty \frac{1}{x^{3/2}} dx$.
 - i. Use the First FTC to determine the exact values of $\int_1^{10} \frac{1}{x^{3/2}} dx$, $\int_1^{1000} \frac{1}{x^{3/2}} dx$, and $\int_1^{100000} \frac{1}{x^{3/2}} dx$. Then, use your calculator to compute a decimal approximation of each result.
 - ii. Use the First FTC to evaluate the definite integral $\int_1^b \frac{1}{x^{3/2}} dx$ (which results in an expression that depends on b).
 - iii. Now, use your work from (ii.) to evaluate the limit given by

$$\lim_{b\to\infty}\int_1^b \frac{1}{x^{3/2}} \, dx.$$

- c. Plot the functions $y = \frac{1}{x}$ and $y = \frac{1}{x^{3/2}}$ on the same coordinate axes for the values x = 0...10. How would you compare their behavior as x increases without bound? What is similar? What is different?
- d. How would you characterize the value of $\int_1^\infty \frac{1}{x} dx$? of $\int_1^\infty \frac{1}{x^{3/2}} dx$? What does this tell us about the respective areas bounded by these two curves for $x \ge 1$?

6.5.2 Convergence and Divergence

Our work so far has suggested that when we consider a nonnegative function f on an interval $[1, \infty]$, such as $f(x) = \frac{1}{x}$ or $f(x) = \frac{1}{x^{3/2}}$, there are at least two possibilities for the value of $\lim_{b\to\infty} \int_1^b f(x) \, dx$: the limit is finite or infinite. With these possibilities in mind, we introduce the following terminology.

If f(x) is nonnegative for $x \ge a$, then we say that the improper integral $\int_a^\infty f(x) dx$ converges provided that

$$\lim_{b\to\infty}\int_a^b f(x)\,dx$$

exists and is finite. Otherwise, we say that $\int_a^\infty f(x) dx$ diverges.

We normally restrict our interest to improper integrals for which the integrand is nonnegative. Further, we note that our primary interest is in functions f for which $\lim_{x\to\infty} f(x) = 0$,

for if the function f does not approach 0 as $x \to \infty$, then it is impossible for $\int_a^\infty f(x) dx$ to converge.

Activity 6.5.3. Determine whether each of the following improper integrals converges or diverges. For each integral that converges, find its exact value.

a.
$$\int_{1}^{\infty} \frac{1}{r^2} dx$$

b.
$$\int_0^\infty e^{-x/4} dx$$

a.
$$\int_{1}^{\infty} \frac{1}{x^{2}} dx$$

b. $\int_{0}^{\infty} e^{-x/4} dx$
c. $\int_{2}^{\infty} \frac{9}{(x+5)^{2/3}} dx$

d.
$$\int_4^\infty \frac{3}{(x+2)^{5/4}} dx$$

e.
$$\int_0^\infty x e^{-x/4} dx$$

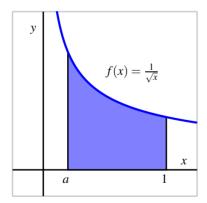
f. $\int_1^\infty \frac{1}{x!} dx$, where *p* is a positive real

6.5.3 Improper Integrals Involving Unbounded Integrands

It is also possible for an integral to be improper due to the integrand being unbounded on the interval of integration. For example, if we consider

$$\int_0^1 \frac{1}{\sqrt{x}} \, dx,$$

we see that because $f(x) = \frac{1}{\sqrt{x}}$ has a vertical asymptote at x = 0, f is not continuous on [0, 1], and the integral is attempting to represent the area of the unbounded region shown at right in Figure 6.5.2.



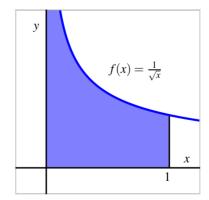


Figure 6.5.2: At left, the area bounded by $f(x) = \frac{1}{\sqrt{x}}$ on the finite interval [a, 1]; at right, the result of letting $a \to 0^+$, where we see that the shaded region will extend vertically without bound.

Just as we did with improper integrals involving infinite limits, we address the problem

of the integrand being unbounded by replacing such an improper integral with a limit of proper integrals. For example, to evaluate $\int_0^1 \frac{1}{\sqrt{x}} dx$, we replace 0 with a and let a approach 0 from the right. Thus,

$$\int_0^1 \frac{1}{\sqrt{x}} \, dx = \lim_{a \to 0^+} \int_a^1 \frac{1}{\sqrt{x}} \, dx,$$

and then we evaluate the proper integral $\int_a^1 \frac{1}{\sqrt{x}} dx$, followed by taking the limit. In the same way as with improper integrals involving unbounded regions, we will say that the improper integral converges provided that this limit exists, and diverges otherwise. In the present example, we observe that

$$\int_{0}^{1} \frac{1}{\sqrt{x}} dx = \lim_{a \to 0^{+}} \int_{a}^{1} \frac{1}{\sqrt{x}} dx$$
$$= \lim_{a \to 0^{+}} 2\sqrt{x} \Big|_{a}^{1}$$
$$= \lim_{a \to 0^{+}} 2\sqrt{1} - 2\sqrt{a}$$
$$= 2,$$

and therefore the improper integral $\int_0^1 \frac{1}{\sqrt{x}} dx$ converges (to the value 2).

We have to be particularly careful with unbounded integrands, for they may arise in ways that may not initially be obvious. Consider, for instance, the integral

$$\int_{1}^{3} \frac{1}{(x-2)^2} \, dx.$$

At first glance we might think that we can simply apply the Fundamental Theorem of Calculus by antidifferentiating $\frac{1}{(x-2)^2}$ to get $-\frac{1}{x-2}$ and then evaluate from 1 to 3. Were we to do so, we would be erroneously applying the FTC because $f(x) = \frac{1}{(x-2)^2}$ fails to be continuous throughout the interval, as seen in Figure 6.5.3.

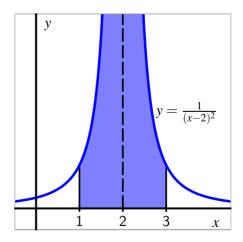


Figure 6.5.3: The function $f(x) = \frac{1}{(x-2)^2}$ on an interval including x = 2.

Such an incorrect application of the FTC leads to an impossible result (-2), which would itself suggest that something we did must be wrong. Indeed, we must address the vertical asymptote in $f(x) = \frac{1}{(x-2)^2}$ at x = 2 by writing

$$\int_{1}^{3} \frac{1}{(x-2)^{2}} dx = \lim_{a \to 2^{-}} \int_{1}^{a} \frac{1}{(x-2)^{2}} dx + \lim_{b \to 2^{+}} \int_{b}^{3} \frac{1}{(x-2)^{2}} dx$$

and then evaluate two separate limits of proper integrals. For instance, doing so for the integral with a approaching 2 from the left, we find

$$\int_{1}^{2} \frac{1}{(x-2)^{2}} dx = \lim_{a \to 2^{-}} \int_{1}^{a} \frac{1}{(x-2)^{2}} dx$$

$$= \lim_{a \to 2^{-}} -\frac{1}{(x-2)} \Big|_{1}^{a}$$

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

$$= \infty,$$

since $\frac{1}{a-2} \to -\infty$ as a approaches 2 from the left. Thus, the improper integral $\int_1^2 \frac{1}{(x-2)^2} dx$ diverges; similar work shows that $\int_2^3 \frac{1}{(x-2)^2} dx$ also diverges. From either of these two results, we can conclude that that the original integral, $\int_1^3 \frac{1}{(x-2)^2} dx$ diverges, too.

Activity 6.5.4. For each of the following definite integrals, decide whether the integral is improper or not. If the integral is proper, evaluate it using the First FTC. If the integral is improper, determine whether or not the integral converges or diverges; if the integral converges, find its exact value.

a.
$$\int_0^1 \frac{1}{x^{1/3}} dx$$

d.
$$\int_{-2}^{2} \frac{1}{x^2} dx$$

b.
$$\int_0^2 e^{-x} dx$$

e.
$$\int_0^{\pi/2} \tan(x) \, dx$$

a.
$$\int_0^1 \frac{1}{x^{1/3}} dx$$

b. $\int_0^2 e^{-x} dx$
c. $\int_1^4 \frac{1}{\sqrt{4-x}} dx$

f.
$$\int_0^1 \frac{1}{\sqrt{1-x^2}} \, dx$$

Summary

- An integral $\int_a^b f(x) dx$ can be improper if at least one of a or b is $\pm \infty$, making the interval unbounded, or if f has a vertical asymptote at x = c for some value of c that satisfies $a \le c \le b$. One reason that improper integrals are important is that certain probabilities can be represented by integrals that involve infinite limits.
- When we encounter an improper integral, we work to understand it by replacing the

improper integral with a limit of proper integrals. For instance, we write

$$\int_{a}^{\infty} f(x) dx = \lim_{b \to \infty} \int_{a}^{b} f(x) dx,$$

and then work to determine whether the limit exists and is finite. For any improper integral, if the resulting limit of proper integrals exists and is finite, we say the improper integral converges. Otherwise, the improper integral diverges.

• An important class of improper integrals is given by

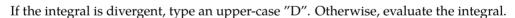
$$\int_{1}^{\infty} \frac{1}{x^{p}} \, dx$$

where p is a positive real number. We can show that this improper integral converges whenever p > 1, and diverges whenever 0 . A related class of improperintegrals is $\int_0^1 \frac{1}{x^p} dx$, which converges for $0 , and diverges for <math>p \ge 1$.

Exercises

1. Consider the integral

$$\int_0^3 \frac{-8}{x\sqrt{x}} \, dx$$



2. Calculate the integral below, if it converges. If it does not converge, enter *diverges* for your

answer.
$$\int_{2}^{\infty} 1x^2 e^{-x^3} dx = \boxed{}$$

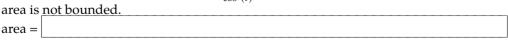
3. Calculate the integral, if it converges. If it diverges, enter *diverges* for your answer.

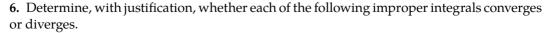
$$\int_{-\infty}^{1} \frac{e^{5x}}{1 + e^{5x}} \, dx = \boxed{$$

4. Calculate the integral, if it converges. If it diverges, enter *diverges* for your answer.

$$\int_{-1}^{1} \frac{1}{v} \, dv = \boxed{$$

5. Find the area under the curve $y = \frac{1}{\cos^2(t)}$ between t = 0 and $t = \pi/2$. Enter *diverges* if the





a.
$$\int_{e}^{\infty} \frac{\ln(x)}{x} dx$$

b.
$$\int_{e}^{\infty} \frac{1}{x \ln(x)} dx$$



極

Chapter 6 Using Definite Integrals

c.
$$\int_{e}^{\infty} \frac{1}{x(\ln(x))^2} dx$$

d. $\int_{e}^{\infty} \frac{1}{x(\ln(x))^{p}} dx$, where p is a positive real number

e.
$$\int_0^1 \frac{\ln(x)}{x} dx$$

f.
$$\int_0^1 \ln(x) \, dx$$

7. Sometimes we may encounter an improper integral for which we cannot easily evaluate the limit of the corresponding proper integrals. For instance, consider $\int_1^\infty \frac{1}{1+x^3} dx$. While it is hard (or perhaps impossible) to find an antiderivative for $\frac{1}{1+x^3}$, we can still determine whether or not the improper integral converges or diverges by comparison to a simpler one. Observe that for all x > 0, $1 + x^3 > x^3$, and therefore

$$\frac{1}{1+x^3}<\frac{1}{x^3}.$$

It therefore follows that

$$\int_{1}^{b} \frac{1}{1+x^3} \, dx < \int_{1}^{b} \frac{1}{x^3} \, dx$$

for every b>1. If we let $b\to\infty$ so as to consider the two improper integrals $\int_1^\infty \frac{1}{1+x^3}\,dx$ and $\int_1^\infty \frac{1}{x^3}\,dx$, we know that the larger of the two improper integrals converges. And thus, since the smaller one lies below a convergent integral, it follows that the smaller one must converge, too. In particular, $\int_1^\infty \frac{1}{1+x^3}\,dx$ must converge, even though we never explicitly evaluated the corresponding limit of proper integrals. We use this idea and similar ones in the exercises that follow.

- a. Explain why $x^2 + x + 1 > x^2$ for all $x \ge 1$, and hence show that $\int_1^\infty \frac{1}{x^2 + x + 1} dx$ converges by comparison to $\int_1^\infty \frac{1}{x^2} dx$.
- b. Observe that for each x > 1, ln(x) < x. Explain why

$$\int_{2}^{b} \frac{1}{x} \, dx < \int_{2}^{b} \frac{1}{\ln(x)} \, dx$$

for each b > 2. Why must it be true that $\int_2^b \frac{1}{\ln(x)} dx$ diverges?

c. Explain why $\sqrt{\frac{x^4+1}{x^4}} > 1$ for all x > 1. Then, determine whether or not the improper integral

$$\int_{1}^{\infty} \frac{1}{x} \cdot \sqrt{\frac{x^4 + 1}{x^4}} \, dx$$

converges or diverges.

APPENDIX A

A Short Table of Integrals

a.
$$\int \frac{du}{a^2 + u^2} = \frac{1}{a} \arctan \frac{u}{a} + C$$

b.
$$\int \frac{du}{\sqrt{u^2+a^2}} = \ln|u + \sqrt{u^2 \pm a^2}| + C$$

c.
$$\int \sqrt{u^2 \pm a^2} du = \frac{u}{2} \sqrt{u^2 \pm a^2} \pm \frac{a^2}{2} \ln|u + \sqrt{u^2 \pm a^2}| + C$$

d.
$$\int \frac{u^2 du}{\sqrt{u^2 + a^2}} = \frac{u}{2} \sqrt{u^2 \pm a^2} \mp \frac{a^2}{2} \ln|u + \sqrt{u^2 \pm a^2}| + C$$

e.
$$\int \frac{du}{u\sqrt{u^2+a^2}} = -\frac{1}{a} \ln \left| \frac{a+\sqrt{u^2+a^2}}{u} \right| + C$$

f.
$$\int \frac{du}{u\sqrt{u^2-a^2}} = \frac{1}{a} \sec^{-1} \frac{u}{a} + C$$

g.
$$\int \frac{du}{\sqrt{a^2 - u^2}} = \arcsin \frac{u}{a} + C$$

h.
$$\int \sqrt{a^2 - u^2} \, du = \frac{u}{2} \sqrt{a^2 - u^2} + \frac{a^2}{2} \arcsin \frac{u}{a} + C$$

i.
$$\int \frac{u^2}{\sqrt{a^2 - u^2}} du = -\frac{u}{2} \sqrt{a^2 - u^2} + \frac{a^2}{2} \arcsin \frac{u}{a} + C$$

j.
$$\int \frac{du}{u\sqrt{a^2-u^2}} = -\frac{1}{a} \ln \left| \frac{a+\sqrt{a^2-u^2}}{u} \right| + C$$

k.
$$\int \frac{du}{u^2 \sqrt{a^2 - u^2}} = -\frac{\sqrt{a^2 - u^2}}{a^2 u} + C$$

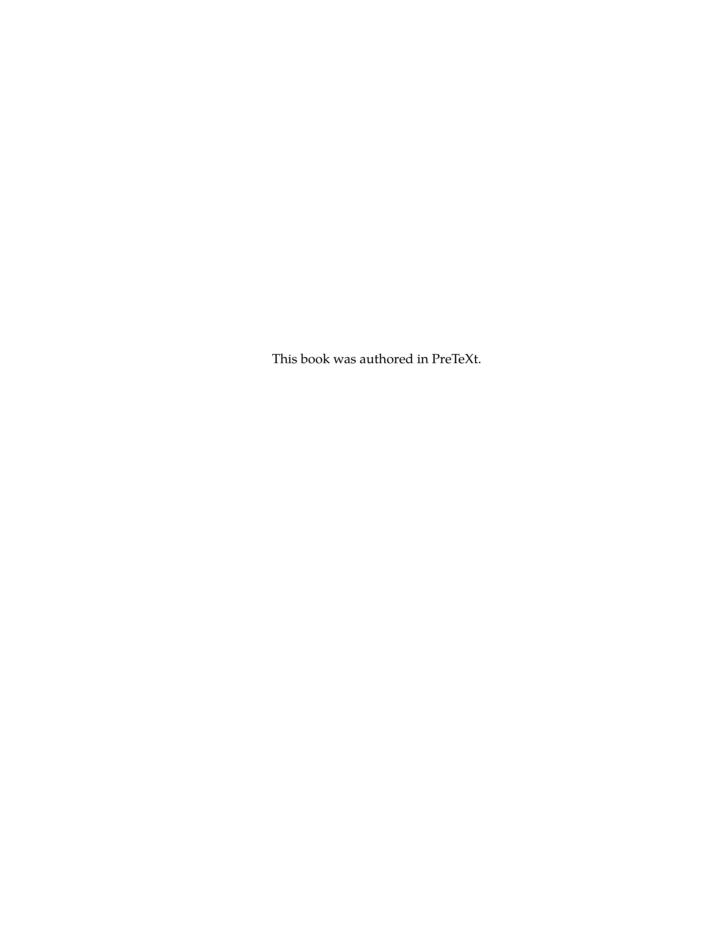
Index

u-substitution, 295	converge
	sequence, 452
absolute convergence, 490	convergence
acceleration, 62	absolute, 490
alternating series, 486	conditional, 490
alternating series estimation theorem,	convergent sequence, 452
489	cosecant, 119
alternating series test, 488	cotangent, 119
antiderivative, 215	critical number, 168
general, 259, 274	critical point, 168
graph, 271	critical value, 168
antidifferentiation, 213	cusp, 77
arc length, 340, 342	
arcsine, 138	decreasing, 55
area, 336	definite integral
under velocity function, 211	constant multiple rule, 245
asymptote, 159	definition, 239
horizontal, 159	sum rule, 245
vertical, 159	density, 355
autonomous, 392	derivative
average rate of change, 22	arcsine, 140
average value, 247	constant function, 93
average value of a function, 246	cosine, 103
average velocity, 2	cotangent, 120
	definition, 24, 36
backward difference, 48	exponential function, 93
	inverse, 142
carrying capacity, 441	logarithm, 137
central difference, 48	power function, 93
chain rule, 127	sine, 103
codomain, 135	tangent, 120
composition, 124	difference quotient, 47
concave down, 61	differentiable, 24, 75
concave up, 61	differential equation, 386
concavity, 60	autonomous, 392
conditional convergence, 490	first order, 392
constant multiple rule, 94	solution, 389
continuous, 74	disk method, 346
continuous at $x = a$, 74	distance traveled, 213

INDEX

diverge	indeterminate form, 14
sequence, 452	infinite series, 469
Divergence Test, 473	infinity, 158
domain, 135	inflection point, 173
dominates, 163	initial condition, 274
·	instantaneous rate of change, 23, 46
equilibrium solution, 402	instantaneous velocity, 3, 16
stable, 403	integral function, 274
unstable, 403	integral sign, 240
error, 325	integral test, 474, 477
error functin, 285	integrand, 240
error function, 317	integration by parts, 304
Euler's Method, 410	interval of convergence, 507
error, 415	
extreme value, 166	L'Hôpital's rule, 155
extreme value theorem, 188	L'Hôpital's rule (at infinity), 160
,	Lagrange error bound, 511
Fibonacci sequence, 449	left limit, 71
first derivative test, 168	lemniscate, 146
foot-pound, 365	limit
forward difference, 48	definition, 13
FTC, 257	one-sided, 71
function, 135	limit comparison test, 477
function-derivative pair, 296	limits of integration, 240
Fundamental Theorem of Calculus	local linearization, 84
First, 281	locally linear, 77
Second, 283	logistic, 440
fundamental theorem of calculus, 255,	logistic equation, 440
257	solution, 442
	,
geometric series, 459	Maclaurin series, 506
common ratio, 459	mass, 355
geometric sum, 458	maximum
	absolute, 165
harmonic series, 474	global, 165
Hooke's Law, 367	local, 166
	relative, 166
implicit function, 146	midpoint rule
improper integral, 377	error, 325
converges, 379	minimum
diverges, 379	absolute, 165
unbounded integrand, 381	global, 165
unbounded region of integration,	local, 166
378	relative, 166
increasing, 55	·
indefinite integral, 294	net signed area, 231
evaluate, 295	Newton's Law of Cooling, 393

Newton-meter, 365	series, 469
	converges, 471
one-to-one, 135	diverges, 471
onto, 135	geometric, 459
1 6 212	sigma notation, 225
partial fractions, 313	Simpson's rule, 328
partial sum, 460, 471	slope field, 397, 399
per capita growth rate, 438	solid of revolution, 345
position, 2	stable, 403
power series, 517	sum rule, 95
power series differentiation and	
integration theorem, 522	tangent, 119
product rule, 108	tangent line, 25
austiont mula 110	equation, 83
quotient rule, 110	Taylor polynomial
ratio test, 479	error, 509
related rates, 201	Taylor polynomials, 500
Riemann sum, 227	Taylor series, 504, 505
left, 227	interval of convergence, 507
middle, 229	radius of convergence, 509
right, 229	total change theorem, 260, 261
right limit, 71	trapezoid rule, 323
	error, 325
secant, 119	triangular numbers, 449
secant line, 25	trigonometry, 118
second derivative, 58	fundamental trigonometric identity,
second derivative test, 170	118
second fundamental theorem of calculus,	
284	unstable, 403
separable, 420, 421	
sequence, 451	washer method, 348
term, 451	weighted average, 328, 358
sequence of partial sums, 471	work, 365, 366



MTH 252 Lab Supplement

Contents

Antiderivatives	1
Trigonometric Substitution	3
Approximate Integrals Technology Lab (Optional)	4
Error Bound Formulas	6
The Comparison Test for Improper Integrals	7
Cross-Sectional Volumes	8
Method of Cylindrical Shells	11
Mean Value Theorem for Integrals	

Antiderivatives

1. Find the general antiderivative of each function. Check your answer by taking the derivative of the second column.

Function	Antiderivative
$f(x) = 4x^3$	
$f(x) = x^5$	
$f(x) = x^n$	
$f(x) = \sqrt[3]{x^2} - \frac{2}{x^4}$	
$f(x) = \frac{x^4 + 2\sqrt{x}}{x}$	
$f(x) = \frac{1}{x}$	
$f(x) = e^x$	

$f(x) = \cos(x)$	
$f(x) = \sin(x)$	
$f(x) = \sec^2(x)$	
$f(x) = \frac{1}{\sqrt{1 - x^2}}$	
$f(x) = -\frac{1}{\sqrt{1 - x^2}}$	
$f(x) = \frac{1}{1+x^2}$	

- 2. Find equations for f' and f, given f''(x) = x, f'(0) = 1, f(0) = 0.
- 3. Find equations for g' and g, given $g''(t) = 2e^t + 3\cos(t)$, g'(0) = 1, g(0) = 3.
- 4. A stone is dropped off a cliff. It hits the ground 6 seconds later. How high is the cliff? (Hint: Acceleration due to gravity is a constant -32 ft/sec².)
- 5. A car traveling 84 ft/s begins to decelerate at a constant rate of 14 ft/s². After how many seconds does the car come to a stop and how far will the car have traveled before stopping?

6. You are the question designer! Create a limit problem where you must correctly apply L'Hospital's Rule twice in order to evaluate the limit. The ground work is laid out for you below. Determine the functions f'(x), g'(x), f(x), and g(x) to result in the following limit. Don't forget about the criteria needed in order to apply L'Hospital's Rule.

$$\lim_{x \to 0} \frac{f(x)}{g(x)} \stackrel{\text{L'H}}{=} \lim_{x \to 0} \frac{f'(x)}{g'(x)} \leftarrow \text{After } 1^{\text{st}} \text{ application of L'Hospital's Rule}$$

$$\stackrel{\text{L'H}}{=} \lim_{x \to 0} \frac{f''(x)}{g''(x)} \leftarrow \text{After } 2^{\text{nd}} \text{ application of L'Hospital's Rule}$$

$$= \lim_{x \to 0} \frac{e^x + \cos(x)}{12x - 6} \leftarrow \text{The resulting limit}$$

$$= \frac{e^0 + \cos(0)}{12(0) - 6} \leftarrow \text{Evaluating the limit}$$

$$= \frac{1 + 1}{-6}$$

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

Trigonometric Substitution

1. Evaluate the integral using the recommended trigonometric substitution:

b.
$$\int_{0}^{5} \frac{1}{(25+x^2)^2} dx$$
, $x = 5 \tan \theta$

c.
$$\int \frac{x^2}{\sqrt{4-x^2}} dx, \quad x = 2\sin\theta$$

- 2. Evaluate the integral using the necessary trigonometric substitution:
 - a. $\int_{-6}^{6} \sqrt{36 x^2} dx$

← Check your answer using geometry!

- b. $\int \frac{1}{x\sqrt{x^2 16}} dx$
- c. $\int \frac{1}{x^2 \sqrt{1-x^2}} dx$
- d. $\int_{-2}^{2} \frac{1}{(4+x^2)^{3/2}} dx$

Approximate Integrals Technology Lab (Optional)

Use the link below to be taken to the pre-created GeoGebra applet:

https://www.geogebra.org/m/SKAvZYtn

1. Consider the integral $\int_{0}^{1} xe^{2x^2} dx$. Using the GeoGebra applet, fill in the table with the given approximations using the sliders, as needed. Round to 8 decimal places.

	Approximation	Over or under?
M ₆		
T ₆		
M ₁₆		
T ₁₆		

Based on the shape of the curve, explain how you determined whether $M_{\rm N}$ and $T_{\rm N}$ were over- or under-approximations.

2. Confirm GeoGebra's result for T_6 by hand using the general formula for T_N (that is, the formula with 1, 2, 2, ..., 2, 2, 1). Be sure to state whether your result for T_6 matches the one that GeoGebra gave. Round your final answer to 8 decimal places.

- 3. Determine the *exact* value of $\int_0^1 xe^{2x^2} dx$ using an integration technique we've learned in class. State the technique being used. Show all work with proper notation, and give both an exact value and an approximate value, rounding to 8 decimal places. Is this pretty close to the approximation from #2? What's the error?
- 4. Change the function and sliders in GeoGebra so that the applet approximates the integral $\int_{-1}^{1} x \sin(\pi x) dx$. (To change the function, double-click in the

	Approximation		
M ₄			
T ₄			
S_8			
M ₁₆			
T ₁₆			
S ₃₂			

Algebra window and edit. Use pi for π .)

Fill in the table.

- 5. Confirm GeoGebra's result for S_8 by hand using the general formula for S_N (the one with $1, 4, 2, 4, \ldots, 4, 2, 4, 1$). Be sure to state whether your result for S_8 matches the one that GeoGebra gave. Do not use decimal approximations in your work; keep values exact until the very end, and then round your final answer to 8 decimal places.\
- 6. Determine the *exact* value of $\int_{-1}^{1} x \sin(\pi x) dx$ using an integration technique we've learned in class. State the technique being used. Show all work with proper notation, and give both an exact value and an approximate value, rounding to 8 decimal places. Is this pretty close to the approximation from #5? What's the error?

Error Bound Formulas

$$Error(T_N) \le \frac{K_2(b-a)^3}{12N^2}$$
 $Error(M_N) \le \frac{K_2(b-a)^3}{24N^2}$ $Error(S_N) \le \frac{K_4(b-a)^5}{180N^4}$

- 1. Find the maximum possible error associated in using T_{10} to approximate $\int_{0}^{1} e^{-4x} dx$.
- 2. Find the value of N for which S_N approximates $\int_0^{\pi/2} \sin(2x) dx$ with an error of at most 0.001.

- 3. Consider $\int x \ln(x) dx$. Answer the following, rounding all approximations to 6 decimal places.
 - a. Find the maximum possible error associated in using S₄ to approximate $\int_{\mathcal{X}} \ln(x) dx.$
 - b. Use S₄ to approximate $\int_{1}^{3} x \ln(x) dx$.
 - c. Evaluate $\int_{1}^{3} x \ln(x) dx$. Give an exact and approximate value.
 - d. Determine the actual error when using S₄ to approximate $\int_{1}^{3} x \ln(x) dx$.

The Comparison Test for Improper Integrals

1. Use the Comparison Test to determine whether the following integrals converge or diverge. Your answer should be a sentence of how the Comparison Test was used and what your final conclusion is.

$$\int_{1}^{\infty} \frac{1}{\sqrt{x^4 + 1}} dx$$

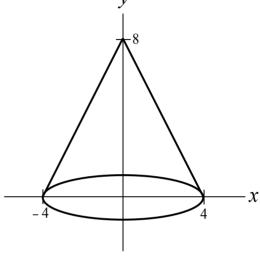
b.
$$\int_{1}^{\infty} \frac{1}{x^{1/3} + x^{3}} dx$$
c.
$$\int_{2}^{\infty} \frac{1}{\sqrt[3]{x^{2} - 1}} dx$$

$$\int_{2}^{\infty} \frac{1}{\sqrt[3]{x^2 - 1}} dx$$

2. Determine whether the improper integral $\int_{0}^{\infty} \frac{x}{(x^2+1)^3} dx$ converges or diverges by using the Comparison Test. If it converges, evaluate it.

Cross-Sectional Volumes

- 1. Consider a cone with base of radius 4 and height 8 (see figure).
 - a. What shape are the cross-sections perpendicular to the *y*-axis? Are the widths of the cross-sections Δx or Δy ?

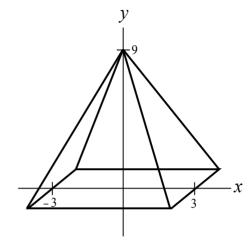


b. Draw the cross-section at the following *y*-values and determine the radius and area.

y-value	Radius	Area
y = 0		
<i>y</i> = 4		
y = 8		

- c. Notice in part b that the radii are given by the *x*-coordinate (verify this in the table above). Find an equation that relates *x* and *y* so that we can get the radii in terms of *y* instead of *x*. Use this to get the area of the cross-sections in terms of *y*. Why do we need to do this?
- d. Integrate the area formula found in part c to find the volume of the cone. What should the bounds of integration be? Check your answer using the formula $V = \frac{1}{3}\pi r^2 h$.

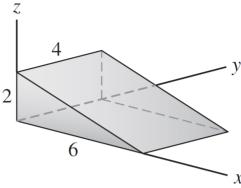
- 2. Consider a pyramid with square base of dimension 6 by 6 and height of 9 (see figure).
 - a. What shape are the cross-sections perpendicular to the *y*-axis? Are the widths of the cross-sections Δx or Δy ?



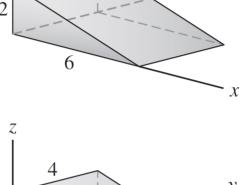
b. Draw an arbitrary cross-section and determine the formula for its area. (Find an equation that relates *x* and *y* and use as necessary in creating your formula.) Should this area formula be in terms of *x* or *y*?

c. Integrate the area formula found in part c with respect to the appropriate variable (so either dx or dy) to find the volume of the pyramid. What should the bounds of integration be? Check your answer using the formula $V = \frac{1}{3}b^2h.$

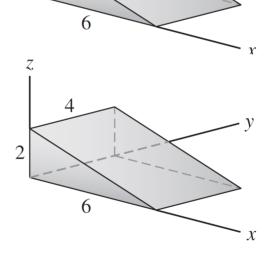
- 3. Calculate the volume of the ramp shown by integrating the area of the crosssections perpendicular to each axis. Draw an arbitrary cross-section for each and state what the shape is. You should get the same answer.
 - a. x-axis



b. y-axis



c. z-axis



- d. Which approach from part (a)—(c) was easiest? Reflect on why.
- e. Check your answer using the formula $V = \frac{1}{2}lwh$.

Method of Cylindrical Shells

Students can use the Method of Cylindrical Shells in many of the Active Calculus Section 6.2 questions. They can do some of the exercises twice: once using disks/washers and again using cylindrical shells.

Problems that lend themselves well to the Method of Cylindrical Shells:

Activity 6.2.2 e

Activity 6.2.3 a, b, d, e

Activity 6.2.4 c, d

Exercise 6.2.5 #2, 6, 7d, 8d, 8f, 9c

Mean Value Theorem for Integrals

The Mean Value Theorem for Integrals: If f is continuous on [a, b], then there exists a number c in [a, b] such that

$$f(c) = f_{ave} = \frac{1}{b - a} \int_{a}^{b} f(x) dx$$

That is,

$$\int_{a}^{b} f(x)dx = f(c)(b-a)$$

- 1. For each of the following:
 - Find the average value $f_{ave} = \frac{1}{b-a} \int_{a}^{b} f(x) dx$ on the given interval.
 - Find c such that $f(c) = f_{ave}$.
 - Sketch the graph of f and a rectangle whose area is the same as the area under the graph of f on the given interval.

a.
$$g(x) = 3x^2 + 3$$
 on $[-1, 3]$

b.
$$g(x) = \frac{1}{x} \text{ on } [1, e]$$

c.
$$f(x) = \cos(2x)$$
 on $\left[0, \frac{\pi}{4}\right]$

d.
$$g(t) = \frac{1}{1+t^2}$$
 on [-1, 1]

- 2. A ball is thrown vertically upwards from ground level with an initial velocity of 96 ft/sec. It's height, h(t) in feet, as a function of time, t in seconds, is given by $h(t) = 96t 16t^2$.
 - a. Find the average height of the ball during the time period that it's in the air.
 - b. Determine when the ball's height is equal to its average height.
 - c. Find the average velocity of the ball during the time period that it's in the air.
 - d. Determine when the ball's velocity is equal to its average velocity.
 - e. Find the average acceleration of the ball during the time period that it's in the air.
 - f. Determine when the ball's acceleration is equal to its average acceleration.