



CALCULUS

Dynamic Mathematics

Single and Multivariable

Fourth Edition

Chuck Garner, Ph.D.

Rockdale Magnet School for Science and Technology



Raleigh London Toronto Bangalore

RMMT PUBLISHING

Conyers

Calculus: Dynamic Mathematics

Single and Multivariable

Fourth Edition

Printed by LULU, published by RMMT PUBLISHING.

This book was produced directly from the author's \LaTeX files.

Figures were drawn by the author using the \TeX draw package and PSTricks.

Set in *STIX2*, designed by Ross Mills and John Hudson in 2016, and *Optima*, designed by Hermann Zapf in 1958.

Cover and titlepage photograph:

“Roller Coaster in Amusement Park”

by Paul Brennan. Photo courtesy of www.publicdomainpictures.net.

Printed and bound in the United States of America.

Contents

Preface	v
Acknowledgements	xi
1 RATE EQUATIONS	1
1.1 Water Flowing Into a Tank,	2
1.2 Population Explosion,	4
1.3 Constrained Growth,	7
1.4 Falling Through the Sea,	11
1.5 A Measles Epidemic,	13
1.6 Euler's Method,	16
1.7 Slope Fields,	20
1.8 Preparation and Extension,	22
2 RATES AND DERIVATIVES	29
2.1 Local Linearity,	29
2.2 Have You Reached the Limit?,	32
2.3 Pathological Cases,	36
2.4 The Derivative,	40
2.5 Trigonometric Derivatives,	44
2.6 Powers, Products, and Quotients,	47
2.7 Compositions and the Chain Rule,	52
2.8 The Value of Theorems,	56
2.9 Preparation and Extension,	62
3 USING THE DERIVATIVE	67
3.1 Differentials and Linearizations,	67
3.2 Related Rate Problems,	71
3.3 Derivatives That Are Not Explicit,	75
3.4 Inverse Trigonometric Derivatives,	81
3.5 More Value of Theorems,	84
3.6 l'Hôpital's Rule,	88
3.7 Infinite Limits,	92

3.8 Preparation and Extension, 96

4 AREAS AND ANTIDERIVATIVES 101

- 4.1 Finite Series, 101
- 4.2 Approximating Areas by Rectangles, 107
- 4.3 Trapezoids, Tables, and Graphs, 112
- 4.4 Definite Integrals, 118
- 4.5 Antiderivatives, 126
- 4.6 Properties of Areas, 128
- 4.7 Preparation and Extension, 135

5 THE INTEGRAL AS A FUNCTION 141

- 5.1 The Fundamental Theorem of Calculus, 141
- 5.2 The Natural Logarithm, 147
- 5.3 The Exponential Function, 151
- 5.4 Derivatives of Logarithms and Exponentials, 155
- 5.5 Hyperbolic Functions, 157
- 5.6 The Substitution Method, 161
- 5.7 Integrals Involving Infinity, 167
- 5.8 Preparation and Extension, 172

6 MODELING PHYSICAL PHENOMENA 179

- 6.1 Maxima and Minima of Functions, 179
- 6.2 Optimization Problems, 185
- 6.3 Inflections and Concavity, 192
- 6.4 Position, Velocity, Acceleration, 196
- 6.5 Graphical and Tabular Applications, 202
- 6.6 Cross-Sections, or Volume Part 1, 207
- 6.7 Disks and Washers, or Volume Part 2, 212
- 6.8 Cylindrical Shells, or Volume Part 3, 217
- 6.9 Arc Length and Surface Area, 222
- 6.10 Preparation and Extension, 226

7 DIFFERENTIAL EQUATIONS 235

- 7.1 Rate Equations: The Sequel, 235
- 7.2 Separable Differential Equations, 239
- 7.3 Growth and Decay, 244
- 7.4 Integration by Parts, 248
- 7.5 Integration by Partial Fractions, 253
- 7.6 Logistic Differential Equations, 259
- 7.7 Useful Substitutions, 263
- 7.8 Preparation and Extension, 268

8 INFINITE SERIES 279

- 8.1 Approximating Polynomials, 279
- 8.2 Taylor Polynomials, 283

- 8.3 Infinite Sequences, 289
- 8.4 Summing the Infinite, 293
- 8.5 Testing for Convergence, Part One, 300
- 8.6 Making Comparisons, 307
- 8.7 Alternating Series, 312
- 8.8 Absolute and Conditional Convergence, 315
- 8.9 Testing for Convergence, Part Two, 319
- 8.10 Power Series, 324
- 8.11 Creating New Series From Old, 329
- 8.12 Taylor Series, 335
- 8.13 Preparation Problems, 346

9 NEW TYPES OF FUNCTIONS 351

- 9.1 Vectors, 351
- 9.2 Modeling Motion with Vectors, 355
- 9.3 Parametric Equations, 361
- 9.4 More with Motion, 367
- 9.5 Polar Functions, 370
- 9.6 The Calculus of Polar Functions, 374
- 9.7 Preparation and Extension, 380

10 LINEAR ALGEBRA 385

- 10.1 The Geometry of Vectors, 385
- 10.2 Vectors in Three Dimensions, 388
- 10.3 Determinants, 393
- 10.4 Linear Independence, 396
- 10.5 Systems of Linear Equations, 400
- 10.6 Matrices, 407
- 10.7 Invertibility, 413
- 10.8 Eigenvalues and Eigenvectors, 419
- 10.9 Orthogonal Matrices, 424
- 10.10 Preparation and Extension, 429

11 PARTIAL DERIVATIVES 435

- 11.1 Multivariable Functions, 435
- 11.2 The Partial Derivative, 443
- 11.3 The Total Differential, 450
- 11.4 The Jacobian, 455
- 11.5 Differentials of Composite Functions, 460
- 11.6 The General Chain Rule, 462
- 11.7 Implicit Functions, 465
- 11.8 Gradients and Directional Derivatives, 469
- 11.9 Tangents and Normals, 474
- 11.10 Extrema of Surfaces, 480
- 11.11 The Lagrange Multiplier, 486
- 11.12 Two Differential Operators, 491
- 11.13 Preparation and Extension, 496

12	MULTIPLE INTEGRALS	503
12.1	Double Integrals,	503
12.2	The Theory of Double Integrals,	507
12.3	Substitution Revisited,	512
12.4	Integration in the Physical World,	518
12.5	Triple Integrals,	522
12.6	Line Integrals,	525
12.7	Green's Theorem,	533
12.8	The Vector Form of Line Integrals,	536
12.9	Independence of Path,	540
12.10	Independence of Path Using Vectors,	545
12.11	Line Integrals in Space,	549
12.12	Surface Integrals,	554
12.13	Gauss' Divergence Theorem,	559
12.14	Stokes' Theorem,	564
12.15	Independence of Path Redux,	567
12.16	Preparation and Extension,	572
13	MORE DIFFERENTIAL EQUATIONS	575
13.1	Linear First-Order Differential Equations,	575
13.2	Homogeneous Differential Equations,	584
13.3	Linear Second-Order Differential Equations,	590
13.4	The Existence Theorem,	594
13.5	Clairaut Equations and Orthogonal Families,	599
13.6	Series Solutions,	604
13.7	Predator and Prey,	609
13.8	Preparation and Extension,	615
A	CHALLENGE PROBLEMS	619
B	FORMULAS	627
C	SUCCESS IN MATHEMATICS	631
	List of Named Theorems	637
	References	641
	Answers	645
	Index	663

Preface

I would like to throttle the man who wrote this book.

— Hermann Weyl

To throw yet another calculus book into the marketplace is a laughable exercise indeed. Why bother? Aren't there plenty of good books from which one may learn calculus? So why another? The only reason one has for writing a new book on an old subject is that it must present some new viewpoint or some new reason for existence. Those reasons are embodied in one simple statement: This book presents calculus the way I think it should be taught to high school students.

I originally wrote this book because the existing calculus textbooks a) are too big to lug around in your hands or in your backpack, b) have too many exercises, c) try to do far too much to please hundreds of different needs, and d) those books that are directed towards high school are slaves to the Advanced Placement curriculum.

"Advanced Placement" and "AP" are registered trademarks of the College Board, which is not associated or affiliated with the production of this book.

Nearly every word of this book has appeared in separate books over three editions in the last 11 years. So nothing here has not been classroom tested and revised numerous times. However, this is the first time everything has been collected in one volume. My goals for the book from 11 years ago remain unchanged, and they still resonate today. This book is not the typical 1200-page calculus doorstop. That is because this book was written with *teaching* and *learning* in mind. Each section was written as if I was teaching in front of a class – nearly everything written is what I would say to my students, write on the board, or have them discuss. As such, there are no “supplemental” or “optional” sections: everything is taught and all is (hopefully!) learned.

Many calculus books are huge, not only for including unused content, but also for including vast numbers of unattempted problems. There are simply too many exercises for a student to complete in one year in most calculus textbooks. In this book, *every problem is attempted by every student*. Problems have been chosen carefully to be important, relevant, and worthwhile. As a result, there are not many problems at the end of a section. Should students need additional practice with mastering the skills of, say, the chain rule, such material is easily accessed elsewhere—in other books or on the internet—and as such I felt it was not worthwhile to include buckets of drill exercises.

Proofs are an important aspect of this book. The study of calculus begins the transition between computation-driven mathematics to

proof-driven mathematics. Yet proofs are often neglected in the teaching of high school calculus. The book is some attempt at rectifying that situation. I have tried to present rigorous proofs without delving into the higher analysis of the real numbers. Where rigor is not possible, or where such a proof would deliver more confusion than insight, I present an intuitive argument based on assumptions of certain properties of the real numbers. As the student progresses through the book, the rigor increases slightly. But in every case, the proofs are written in an accessible way in order to give high school students practice at thinking about proofs and reading proofs, as well as understanding the logic behind a statement of a theorem.

Most calculus books that are geared towards high school students are written with the goal of the student passing the AP exam. This is not an indictment of College Board's goals for calculus. Indeed, there is much that is good about the focus on understanding, justifications, and application. However, there is also much lacking: hyperbolic functions, real uses of Euler's method, other indeterminate cases for l'Hôpital's rule besides the $0/0$ case, and trigonometric substitution, among others. Students learning nothing but the topics listed in the Course Description will do well on the Exam, but is the AP Exam the main reason a high school student should learn calculus? I don't think so. I believe that the student should learn calculus well in order *to learn calculus*, one of the most important intellectual endeavors of the human race, and *to apply calculus*, one of the most important tools in scientific application and prediction. The AP exam is extra; students who know calculus well will, of course, do well on the AP exam. This book does include every topic in the AP Calculus BC Course Description, for those teachers or students who wish to learn calculus for the AP exam.

Finally, I wish to reiterate that this book was written for one need and one need only: to present calculus to gifted high school students. There is no other motivation for this book: it is not intended for "Advanced Placement," for "liberal arts majors," for "future math teachers," or for "engineering majors." This book does not try to be all things to all instructors. It is simply a concise textbook for gifted high school students—students who may choose to become liberal arts majors, math teachers, or engineers.

Organization. Every topic in a traditional single-variable or multivariable calculus course is included to create a comprehensive high school calculus course. The following is a chapter-by-chapter summary of the book. The following is a chapter-by-chapter summary of the book.

CHAPTER 1: RATE EQUATIONS. This is the most crucial chapter of the book. Here, the students do number crunching to predict physical phenomena based on the rate at which the phenomena occur. They use Euler's method extensively before it is made formulaic. This sets up the students to understand how to work with and use rates to make predictions. Students also must make decisions about the validity of the data, determine constants used in exponential and logistic growth, and judge the validity of their predictions. Euler's method is used throughout the first volume as an investigative tool. In addition, problems in Section 1.4 set up the idea of Riemann sums and foreshadow the Fundamental Theorem of Calculus.

CHAPTER 2: RATES AND DERIVATIVES. After dwelling on the numerical and graphical, we introduce the algebraic as a way to make our predictive abilities more precise. Following this is a traditional approach to derivative rules, but with more of a focus on presenting proofs of major theorems, and less of a focus on the details of limits. Applications of differentiation are deferred to Chapters 3 and 6 so that we may use integration in applications as well.

CHAPTER 3: USING THE DERIVATIVE. This chapter starts with the notion of the differential and then applies it to approximations, related rates, and implicit differentiation. The Mean Value Theorem continues the focus on proofs, and then we use derivatives to return to limits with l'Hôpital's Rule and limits involving infinity.

CHAPTER 4: AREAS AND ANTIDERIVATIVES. Like the arrangement in the first three chapters, we introduce the numerical and graphical first (Sections 1 through 3) before the algebraic (Sections 4 and 5). The sixth section of this short chapter wraps up the question of area computations.

CHAPTER 5: THE INTEGRAL AS A FUNCTION. The Fundamental Theorem of Calculus is stated and proved before students learn any antiderivative rules (other than basic reversal of differentiation) so that students don't get the impression that the Fundamental Theorem is just about antiderivatives. It is this chapter where the logarithmic (in terms of an integral) and the exponential (in terms of the inverse of the logarithm) functions are defined.

CHAPTER 6: MODELING PHYSICAL PHENOMENA. Here we find the traditional applications of single-variable calculus. Most books separate these applications so that they occur closer to the respective differentiation or integration chapters; I feel that the applications are more meaningful and relevant when presented after differentiation and integration have been introduced. Certainly, problems in position, velocity, and acceleration are more realistic when one has both derivatives and integrals at one's disposal, as are other problems from physics (Sections 6.4 and 6.5). The treatment of volume of solids of revolution is deliberately slow-paced to give students time to become familiar with these new methods.

CHAPTER 7: DIFFERENTIAL EQUATIONS. Here, we take another look at the situations in Chapter 1, but this time from the perspective of having built a huge toolbox of more precise techniques.

CHAPTER 8: INFINITE SERIES. The chapter begins in a similar vein as Chapter 4 did, with a development of extending the idea of a linear approximation to polynomial approximations and then Taylor polynomials representing transcendental functions. Then we raise the question of the convergence of an "infinite-degree" polynomial. Over the course of the chapter, we answer that question through convergence tests, power series, Maclaurin series, and finally Taylor series.

CHAPTER 9: NEW TYPES OF FUNCTIONS. This introduces vectors, parametric equations, and polar coordinates as new types of functions. We use these to model curves whose " $y = f(x)$ " form is not regarded as a function. We also use vectors and parametric equations to solve motion problems.

CHAPTER 10: LINEAR ALGEBRA. This is a short introduction to the concepts of linear algebra. The theory of three-dimensional vectors

is fully developed, as are matrices and determinants. All of this is needed for applications of partial derivatives, but many calculus books give this a quick and incomplete treatment. Linear algebra is useful and important in its own right and not just a tool for calculus in three dimensions. Continuing that theme, in the remainder of the chapter are topics traditional in linear algebra courses, but not usually included in calculus books. It is a shame, as eigenvectors and eigenvalues fit nicely with differential equations that model predator-prey problems (which are included in Chapter 4).

CHAPTER 11: PARTIAL DERIVATIVES. A more traditional approach to partial derivatives is found here, but with a few twists. Mirroring the development of single-variable derivatives in Volume 1, the partial derivative is introduced numerically before the algebraic aspect takes over. Applications appear throughout this chapter: biophysics, business, and chemistry, among others. The Jacobian and its purpose is introduced as a differential matrix and used extensively in the next chapter. Differentials of composite functions and the general chain rule lead into the Implicit Function Theorem. We also discuss extrema, including use of the Lagrange multiplier. Eigenvalues and eigenvectors also make a return when discussing extrema on quadratic surfaces.

CHAPTER 12: MULTIPLE INTEGRALS. The multiple integrals only run the first five sections, then line and surface integrals, and how to express and evaluate them as multiple integrals, take over for the remainder of the chapter. Many applications are included throughout. All theorems in this chapter are stated in terms of rectangular equations as well as in terms of vectors. Transforming regions and integrals from rectangular to polar is used throughout; however, spherical and cylindrical coordinates are not used. The reason for the omission of these coordinate systems is that I believe both systems can be easily learned by the student if the need arises, given a firm background in polar coordinates. The three major theorems of multivariable integrals are developed and explained (Green, Gauss, and Stokes), as well as the concept of line integrals which are independent of path.

CHAPTER 13: MORE DIFFERENTIAL EQUATIONS. This short chapter includes simple techniques that rely on partial derivatives and infinite series. In a sense this chapter is an “application” of multivariable ideas to single variable differential equations.

APPENDICES. The appendices are split into three sections. The first is a list of 40 *Challenge Problems* offered to the students as extra credit in the class. They are difficult, and no answers are provided. The second section is a list of formulas. Also included is a blank unit circle diagram that students are encouraged to fill in. The third section is an essay concerning being successful in a mathematics class. It includes suggestions for studying, homework, asking questions in class, getting extra help, and preparing for a test.

At the end of each chapter is a page titled “At the End of the Chapter, You Should Be Able To. . .”. This page acts as an outline of the topics and skills students should have by the close of each chapter. It is one thing to have a goal at the beginning of each section, or a summary of each section at the end of the section. However, students need to develop the idea of a cumulative assessment, and they sometimes have trouble preparing for such an assessment. This new feature of the book

should help students study for a test and to help focus their review on any deficits they may have.

Chapters 1 through 9 and the appendices constitute a complete AP Calculus BC for those instructors who teach BC.

Problems and Exercises. Most sections have at least one “Exercise” in the text. These exercises are designed to be completed by the student at that point in the lesson. They extend recent concepts, foreshadow upcoming ideas, or are for practicing new skills. I use these as classwork assignments, completed individually or in small groups.

One section common to each chapter is the section entitled “Preparation and Extension.” The Preparation problems are designed to prepare students for a chapter test. The problems here are similar to test questions that I write and are a good measure of how well prepared the student is for the test. At the beginning of each Preparation section is a smattering of review problems as well. They are offered to help students retain previously learned concepts. The Extension problems are meant to challenge the students by extending the concepts and skills learned in that chapter. Groups of three or four students work these problems together and they have a minimum of three weeks to complete them. Very often we forget that mathematics is a actually a group activity, and students should have opportunities to model real mathematical problem-solving as practiced by mathematicians.

History. In 2002, inspired by a *Book of Exercises* written by calculus teacher Sergio Stadler at the Marist School in Atlanta, I wrote my own supplement to the calculus textbook my classes used that focused on AP-style questions and problems, as well as some extra topics not currently included in the AP curriculum. What I thought would be a small project blossomed into a 250-page *AP Calculus Problem Book*. About two-thirds of the problems were original, but I must confess that the best problems I stole from various sources. The *Problem Book* was self-published every school year (that is, I made the copies, collated them, and punched holes in the papers and ran the comb bindings through) which allowed me to make corrections and additions (or subtractions) easily. As the years went by, three unforeseen consequences arose from using the book.

First, I noticed that my students preferred the *Problem Book* to the regular textbook. They seemed more willing to do the problems and carry it around as opposed to the big textbook. This led to eventually dropping the published textbook altogether, and using the *Problem Book* as the sole “text.”

Second, I realized that I could put everything my students needed into one place. In future editions, I included the syllabus, labs, old tests, challenge problems, study tips, group projects, practice exams, and other items. This allowed the students to simply keep up with one book where everything was located instead of keeping up with multiple handouts. And, by including everything in one book I didn’t visit the copy machine throughout the school year to make all the handouts and worksheets.

The third consequence was a positive influence on how I teach. I began to include problems specifically designed to be worked in pairs or groups, so that more collaborative learning was taking place. It also allowed me the flexibility to assign homework problems on a class-by-

class basis. I included group problem sets for those occasions when a substitute was necessary, so no day was ever wasted. These features exist in the present work—the “Exercises” are designed to be worked in class by pairs of students, and the “Extension Problems” are used as problem sets for groups of three or four students to tackle over a period of weeks.

But, students missed the examples of a textbook, and I grew complacent with teaching calculus the same way each year. So I started to wonder how best to shake things up. With the announcement that Georgia was moving to an integrated mathematics curriculum, I was charged with re-aligning the “calculus-and-beyond” classes at my school. That meant taking another look at the post-calculus courses: Multivariable Calculus, Linear Algebra, History of Mathematics, and Discrete Mathematics. The latter two changed little, and I still teach them both. Material from linear algebra and multivariable calculus were always approached together in a single year, and I wanted to include differential equations to a greater extent than in the past. So I took this opportunity to rethink my approach to calculus as a whole, and devised my own two-year calculus curriculum that would encompass everything required for students to do well on standardized tests, other single-variable calculus topics, as well as some differential equations, linear algebra, and multivariable calculus. Thankfully, my administration encouraged my ideas and was completely supportive.

The content of Discrete Mathematics did not change but the name did. It was relaunched as “Advanced Finite Mathematics”. We also added a course in introductory operations research called “Mathematics of Industry and Government”.

Technicalities. I began work on the book in June 2007. Besides creating something for my classes, I also approached this as an exercise in \LaTeX . I started using \LaTeX in late 2003, so I was familiar with the program (there was a learning curve: the code for *Problem Book* is quite rough) and knew the possibilities. I started this book on a Windows machine using the fantastic MikTeX distribution with the T_EXnicCenter editor, but I soon migrated to a MacBook where I now use the wonderful MacT_EX and the T_EXShop editor.

When I first learned how to draw figures in \LaTeX , I used the first package I found, which was T_EXdraw. T_EXdraw generates PostScript drawings and was suitable for the limited pictures I needed at the time. However, as this book progressed, I needed more than T_EXdraw was capable of doing. Not wanting to redo every single figure, I needed another PostScript-output package. Hence, I started using the PSTricks bundle of packages. The resulting figures in this book are all PSTricks code.

CHUCK GARNER
Conyers, Georgia
June 2021

Chapter 1

RATE EQUATIONS

HAVE YOU EVER HEARD SOMEONE—maybe you—say any of the following?

“The population is growing more slowly.”

“The plane is landing smoothly.”

“The economy is picking up.”

“The tax rate is constant.”

“The unemployment rate is decreasing.”

“Stock prices have peaked.”

What do these statements have in common? They are all describing **change**. What do these statements have to do with calculus? *Calculus is the mathematics of change.*

We begin our journey by discussing statements of this type. Each of the statements above describes a **rate of change**, a fundamental concept in mathematics and science. In this chapter we will investigate these rates through numerical and graphical analysis. Rates will usually be described by **rate equations**.

To be a bit more precise about it, rate equations are used to describe, predict, and model real-world phenomena, such as the population of Georgia, the growth of yeast bacteria, and the disposal of nuclear waste.

We will model these situations, introduce some useful mathematical tools for investigating these models, and test our predictions against real-world data. In later chapters, we will develop algebraic and analytical methods for investigating rate equations.

1.1 Water Flowing Into a Tank

In the mathematics I can report no deficiency, except that it be that men do not sufficiently understand the excellent use of the pure mathematics, in that they do remedy and cure many defects in the wit and faculties intellectual. For if the wit be too dull, they sharpen it; if too wandering, they fix it; if too inherent in the sense, they abstract it. So that as tennis is a game of no use in itself, but of great use in respect that it maketh a quick eye and a body ready to put itself into all postures; so in the mathematics, that use which is collateral and intervenient is no less worthy than that which is principal and intended.

— Roger Bacon

Before we begin with *rate equations*, we should begin simply with *rates*. Consider the three situations below, each with a graph plotting the time it takes to fill a tank with water, versus the volume of water in the tank.

Situation I. Water is flowing into a tank that is initially empty at a constant rate of 40 gallons per minute for 6 minutes. Since the rate is constant, we know that there must be $6 \times 40 = 240$ gallons in the tank. The graph is therefore a line with endpoints $(0, 0)$ and $(6, 240)$. The slope of the line is 40, the same as the rate. This gives us our first interpretation of a rate: **slope**.

Situation II. The tap, closed at the beginning, is gradually opened over 6 minutes, so that at 6 minutes the tank contains 200 gallons.

Here, the graph is concave up, meaning that the rate of water flow increases. The important question here is this: *What is the average rate of water flow over the 6-minute interval?* The answer is the slope of the line that connects the endpoints $(0, 0)$ and $(6, 200)$. Thus, the **average rate of change** is $\frac{200}{6} \approx 33.33$ gallons per minute.

Situation III. The tap, opened so that water flows at a rate of 40 gallons per minute, is gradually closed, so that at 6 minutes the tank contains 160 gallons. The graph is concave down, meaning that the rate of water flow decreases. The average rate of change is $\frac{160}{6} \approx 26.67$ gallons per minute.

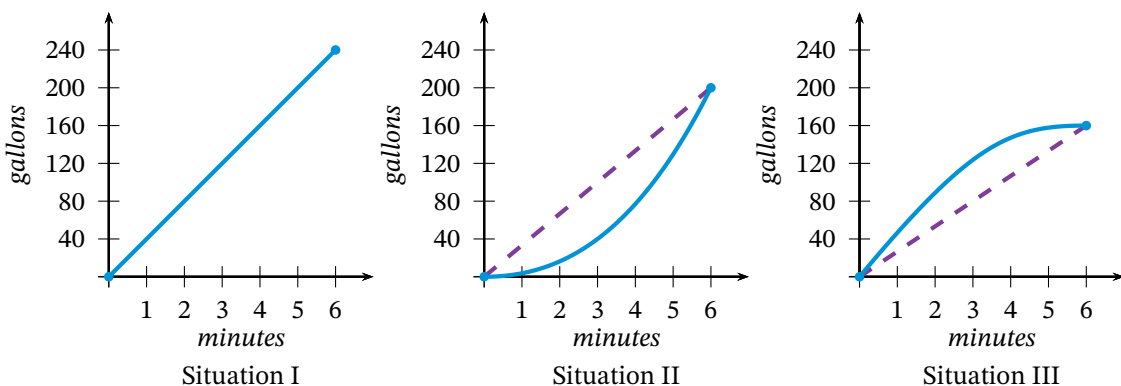


Figure 1.1 – Three different situations for water flowing into a tank

The dashed lines in the graphs above connect the endpoints over which we compute the average rate of change. Such a line is called a **secant line**.

Exercise 1.1.1 At time $t = 0$, water begins to flow from a tap into an empty tank at the rate of 40 gallons per minute. This flow rate is held constant for 2 minutes, then the tap is gradually closed until at $t = 4$ the flow rate is 5 gallons per minute. This new rate is held constant for the final 2 minutes, so that at time $t = 6$, the tank contains 120 gallons.

- Draw a graph of the volume of water in the tank over the first 6 minutes.
- What is the average rate of flow over the interval from 2 to 4 minutes? Draw a secant line on your graph that conveys this information.
- If we wanted to put the entire 120 gallons into the tank at a constant rate, what would that constant rate be?
- Now suppose that a pump is started at $t = 2$, and, for the next 4 minutes, pumps water out of the tank at a constant rate of 15 gallons per minute. The graph you drew in part (a) now represents the total amount of water flowing into the tank at time t . On the same axes as your volume graph, draw the graph that represents the total amount of water pumped out at time t . How do you now interpret the *total* amount of water at time t ?
- What is the total amount of water in the tank at time $t = 6$?
- On your graph from part (e), show how to find the point at which the water level in the tank is a maximum.
- What is the flow rate into the tank at the point found in part (f)?

Let us introduce some convenient mathematical shorthand for the rate of change. Take *Situation I* from the previous page. If we let V represent the volume of water in the tank, then the rate at which water flows into the tank will be denoted by V' . (The symbol $'$ is a *prime mark* so that V' is read “ V prime.”) For instance, in *Situation I*, we have $V' = 40$.

Notice that in *Situation II* and *Situation III*, the rate is not constant, and so V' cannot be expressed as easily as in *Situation I*. One of the questions that calculus helps us answer is that of finding and interpreting expressions for such nonconstant rates. In the next section, we will develop one such technique.

The material in this section and the exercise is adapted from Taylor (1992).

Exercise 1.1.2 Referring to the tank in Exercise 1.1.1, suppose the pump does not operate, so that water is only flowing *into* the tank. Find:

- the exact time $t = k$ for which the average flow rate into the tank over the first k minutes is 25;
- the exact time $t = m$ for which the average flow rate into the tank over the interval from m minutes to 6 minutes is 15;
- a time interval from p to $p + 4$ minutes over which the average flow rate is the same as the average flow rate over the entire 6 minutes.

Problems for Section 1.1

- Illustrate the results in Exercise 1.1.2 using secant lines on the graph.
- Again referring to the tank in the exercise, now suppose that the pump does operate. If the pump continues past 6 minutes operating at 15 gallons per minute, and water continues past 6 minutes flowing in at 5 gallons per minute, then at what time will the tank be empty?
- A tank is being filled at a variable rate. The only thing known is that between 1 and 2 minutes, the

- average inflow rate is 18 gallons per minute, and the amount of water in the tank at 2 minutes is 42 gallons. What is your best estimate for the amount of water in the tank at 2.5 minutes?
- 4 A tank is being filled at a variable rate. The only thing known is that between 1 and 2 minutes, the average inflow rate is 18 gallons per minute, and the amount of water in the tank at 2 minutes is 36 gallons. What is your best estimate for the amount of water in the tank at 2.5 minutes?
- 5 A tank has 20 gallons of water in it, then a tap is opened and water flows into the tank. From 0 to 3 seconds, the average flow rate is 6 gallons per second. From 3 to 5 seconds, the average flow rate is 15 gallons per second. How many gallons of water are there in the tank at 5 seconds?
- 6 Suppose the equation of volume of water in a tank at time t for the first 10 seconds is given by $V(t) = 0.3t^2$. Find the average inflow rate over (a) the first 5 seconds; (b) seconds 2 to 5; (c) seconds 4 to 5; (d) seconds 4.5 to 5; (e) seconds 4.9 to 5. As the time interval gets shorter, to what value does the average inflow rate approach?
- 7 Suppose the volume of water that is flowing out of a full tank is given by $V(t) = 1024 - 0.25t^3$. The time at which the tank is empty is when $t = c$. Find the value of c , then use it to find the average outflow rate over the interval from $c - 1$ to c seconds.

1.2 Population Explosion

In the company of friends, writers can discuss their books, economists the state of the economy, lawyers their latest cases, and businessmen their latest acquisitions, but mathematicians cannot discuss their mathematics at all. And the more profound their work, the less understandable it is.

— Alfred Adler

Thomas Malthus (1766-1834) was an English economist. His essay was titled "An Essay on the Principle of Population."

Thomas Malthus wrote in an essay in 1798 that the growth of the human population is very different from that of the growth of the food supply. He concluded that the human population would grow so fast that famine would be a global epidemic. Malthus, in fact, referred to a rate equation in his essay that simply says that the rate of population growth is proportional to the population. (In simpler terms, this means that the greater the population, the faster the rate of growth.)

Let us consider his statement again with a bit of mathematical symbolism. We let P represent the population; that is, the number of people. Since the growth of that population is related to P , let us denote the rate of growth of that population by P' . Hence Malthus' rate equation becomes the proportion

$$P' = kP$$

where the real number k is called the **constant of proportionality**.

We will now use **Malthus' equation** to predict the population of the state of Georgia in the year 2030. Table 1.1 lists the population of the state of Georgia according to U. S. Census information for the years 1900 to 2010 in 10-year intervals.

The first obstacle we face is how to find the constant of proportionality. We cannot find it exactly, but we can approximate it using the data in the table. Notice that if we solve Malthus' equation for k we get

$$k = \frac{P'}{P}.$$

So we need only to divide values of P' by P to find k . But how do we find values of P' ? We use the average rate of change in the population over every 10-year interval. This means we treat the years as *single data*

Year	Population (millions)
1900	2.216
1910	2.609
1920	2.896
1930	2.909
1940	3.124
1950	3.445
1960	3.943
1970	4.590
1980	5.463
1990	6.478
2000	8.186
2010	9.688

Table 1.1 – Population of the state of Georgia, rounded to the nearest thousand. Data from the U.S. Census Bureau, 2013. (quickfacts.census.gov/qfd/states/13000.html).

points, not as representing 10-year intervals. Thus, in the following computations, the years have been divided by 10 so they act as single data points. For instance, for the years 1900 to 1910, we have

$$P' \approx \frac{2.609 - 2.216}{191 - 190} = \frac{0.393}{1} = 0.393.$$

Therefore,

$$k = \frac{P'}{P} \approx \frac{0.393}{2.609} = 0.151.$$

Notice that we have rounded to the nearest thousandth which is appropriate for this estimation; however, more decimal places may be necessary in other computations. Also, we have not used the population figures in millions, but have treated the figures as if they were decimals. This will cause no problem, as long as we remember that *our predicted answers are in millions*. Most interestingly, we somewhat arbitrarily assumed P was the population in 1910 in the final calculation above. There is no mathematically valid reason to choose the 1910 population rather than the 1900 population. But we can offset this randomness by: finding the slope between the years 1910 and 1920; dividing again by the 1910 value; and taking the average of this with the previous result. This gives an approximate measure of the rate of change during the year 1910, rather than between 1900 and 1910, or between 1910 and 1920.

The slope between 1910 and 1920 is

$$P' \approx \frac{2.896 - 2.609}{192 - 191} = 0.287;$$

the k -value is

$$k = \frac{P'}{P} \approx \frac{0.287}{2.609} = 0.110.$$

Thus, the average of our two k -values at the year 1910 is

$$\frac{0.151 + 0.11}{2} = 0.1305.$$

We have used four decimal places rather than the three we used before in order to better compare this k -value with later results.

Mathematically, we have computed what is called the **symmetric difference** around the point (1910, 2.609). The symmetric difference is slope “around” a point, using x -values (the years in this case) that are spaced the same distance apart on either side of the point. Since the years are always 10 years apart, our “averaging” the two k -values is equivalent to finding the slope between the points (1900, 2.216) and (1920, 2.896), then dividing by the population at 1910 (again treating the years as single data points):

$$\frac{2.896 - 2.216}{192 - 190} = \frac{0.680}{2} = 0.340,$$

then

$$k \approx \frac{0.34}{2.609} = 0.1303.$$

Exercise 1.2.1 What are the units of measure for P' ?

Exercise 1.2.2 Create a table for the years 1910 to 2000 in 10-year intervals, and compute the approximate value of k for each year. Use the k -value of 0.1303 for the year 1910. Why should there be no entry for the year 2010?

Exercise 1.2.3 Malthus' equation only admits one value of k . What is the best value we should adopt for our model? Justify your methods and reasons for deciding upon your value of k . Would the value $k = 0.1316$ be appropriate? Can you determine what computations led to that value?

Let us assume $k = 0.1316$. Then Malthus' equation for the population of Georgia is $P' = 0.1316P$. This equation describes a rate of change; that is, this equation tells us how much to add (P') to the present value (P in 1900) to get the next value (P in 1910).

We have that the change in population is 0.1316 times the population. Hence, if the population in 1900 is 2.216 million, we can approximate the population in 1910 by computing

$$0.1316 \times 2.216 = 0.292$$

and then adding that to the population in 1900:

$$2.216 + 0.292 = 2.508 \text{ million.}$$

To approximate the population in 1920, we repeat the process, but this time using the population in 1910.

Note that the calculated population for 1910 does not agree with the given population in 1910. In fact, Table 1.2 gives the actual population and the population as determined by the model.

None of the estimated and actual numbers agree! But that's to be expected. After all, Malthus' equation is only a model of a real-world situation. This data is still useful, however, because one can readily approximate the answer to the original question: What will be the population of Georgia in the year 2030?

Notice that we cannot use the population in 2020 since we do not have that data available. Hence, we must approximate the population in 2020, and then use that approximation to estimate the population in 2030.

We compute:

$$(9.688 \times 0.1316) + 9.688 = 10.963 \text{ million in 2020, and} \quad (1.1)$$

$$(10.963 \times 0.1316) + 10.963 = 12.406 \text{ million in 2030.} \quad (1.2)$$

We pause briefly to clarify some rate concepts. Earlier, we used the symmetric difference to compute a rate of change *during* the year 1910. In effect, we have estimated the **instantaneous rate of change** of the population in 1910. The instantaneous rate of change is the concept described by a given rate equation.

For instance, if we are given the rate equation $P' = 0.2P$, where P is a population, then given any value of P we can compute the instantaneous rate of change *at that value of P* . Notice that this is not an average rate of change, but the change *exactly* at the given value of P .

Hence, we can safely talk about the the change *at* a point, instead of the change *around* a point.

Year	Actual (millions)	Estimated (millions)
1910	2.609	2.508
1920	2.896	2.838
1930	2.909	3.211
1940	3.124	3.634
1950	3.445	4.112
1960	3.943	4.653
1970	4.590	5.265
1980	5.463	5.958
1990	6.478	6.742
2000	8.186	7.630
2010	9.688	8.634

Table 1.2 – Actual and modeled population of Georgia.

Problems for Section 1.2

- 1 Notice that the population of Georgia rises dramatically in the second half of the 20th century. Let us compute a new value of k by averaging the five intermediate k -values in your table from 1960 to 2000. Use this new k to make a new prediction of the population of Georgia in 2030. Is this a better prediction than 12.239 million? How do you know?
- 2 If the more recent data is more relevant, then the newest data should be the best! Compute the slope between the data points for the years 2000 and 2010, and let this be the value of P' so that the value of k can be found. Make another prediction of the population in 2030. Is this a better prediction? Justify your answer.
- 3 The Government of the State of Georgia projects a population in Georgia of 12.189 million in 2020 and 14.688 million in 2030. Experiment with your calculator to determine a value of k that could arrive at these predictions.
- 4 Is it possible to predict the population of Georgia in the year 2025? Either explain why it cannot be done, or give a method for doing so.
- 5 Below is the population data for Gwinnett County, Georgia. Find an appropriate value of k for Malthus' equation that models the growth of Gwinnett County. Is the value of k for Gwinnett County larger or smaller than the value of k for the state of Georgia? What does this imply about the growth of Gwinnett County? Predict the population of Gwinnett County in the year 2030.
- 6 Either by hand or with a calculator, plot the population data for Gwinnett County shown below. (You may want to convert the years to a more reasonable scale; for instance, 1940 to 0, 1950 to 1, etc.) What curve does the data resemble? Malthus' equation is more commonly referred to as the **exponential growth equation**—does this make sense based on the data?

POPULATION OF GWINNETT COUNTY

Year	Population (thousands)
1940	29.1
1950	32.3
1960	43.5
1970	72.3
1980	166.9
1990	352.9
2000	588.4
2010	805.3

1.3 Constrained Growth

Our population and our use of the finite resources of planet Earth are growing exponentially, along with our technical ability to change the environment for good or ill.

— Stephen Hawking

The equation in Section 1.2 is not the best model to use for population. As Malthus himself pointed out, the food supply will not support the growth of the human population. So what happens when the food supply runs out? The resources will not be able to support unlimited population growth, and the growth rate will slow to a crawl, or “level off.” A model to account for this type of **constrained growth** was first introduced by Pierre-Francois Verhulst (1804-1849), but it was not until Swedish biologist Tor Carlson used real data to develop a mathematical model in 1913 that constrained growth was taken seriously.

How can Malthus' equation be modified to produce a new constrained growth rate equation? Since the population P must level off, this implies that the rate of population growth P' must decrease towards zero. There must also be a maximum sustainable population; let us call this maximum sustainable population M . (M is also called the **carrying capacity** of the population.) The closer P is to M , the closer

the Chain Rules.

$$\begin{aligned} y &= x^2 \arctan(3x - 1) \\ \frac{dy}{dx} &= 2x \cdot \arctan(3x - 1) + x^2 \cdot \frac{3}{9x^2 - 6x + 2} \\ &= 2x \arctan(3x - 1) + \frac{3x^2}{9x^2 - 6x + 2}, \end{aligned}$$

where we have the 3 produced by the Chain Rule. \blacklozenge

Example 3.4.9

We find the derivative of $y = 7 \operatorname{arcsec}(x^2)$. We have

$$\begin{aligned} y &= 7 \operatorname{arcsec}(x^2) \\ \frac{dy}{dx} &= 7 \cdot \frac{1}{|x^2| \sqrt{(x^2)^2 - 1}} \cdot (2x) = \frac{14x}{|x^2| \sqrt{x^4 - 1}} = \frac{14}{|x| \sqrt{x^4 - 1}} \end{aligned}$$

as the derivative. \blacklozenge

Problems for Section 3.4

Evaluate each of the following expressions. Do not use a calculator. **16** Which of the following are undefined?

- 1** $\sin^{-1}\left(-\frac{\sqrt{3}}{2}\right)$ **4** $\tan(\sec^{-1}(2))$
2 $\tan^{-1}(-\sqrt{3})$ **5** $\sin(\arcsin(0.3))$
3 $\sin(\arctan(1))$ **6** $\arcsin(\sin(\pi))$
7 $\arccos(\cos(-\frac{\pi}{4}))$

- $\arccos(1.5)$, $\operatorname{arcsec}(1.5)$, $\arctan(1.5)$,
 $\operatorname{arcsec}(0.3)$, $\arcsin(2.4)$.

17 This problem is an application of the Extreme Value Theorem.

- (a) Since $y = \arcsin(x)$ is defined only on $[-1, 1]$, we can find the extrema by examining the graph. What are the extreme points?
 (b) What are the extreme points of $y = \operatorname{arcsec}(x)$? Are they local or global?
 (c) What are the extreme points of $y = \arctan(x)$?

Find the derivatives of the following functions.

- 8** $y = \sec^{-1}(5x)$ **12** $y = \tan^{-1}\left(\frac{3}{x}\right)$
9 $y = \cos^{-1}(2x - 3)$ **13** $y = \arccos\left(\frac{1}{x}\right)$
10 $y = \arctan(2x - 3)$ **14** $y = 2 \sin^{-1} \sqrt{1 - 2x^2}$
11 $y = \operatorname{arcsec}(3x^2)$ **15** $y = \arcsin(1 - x)$

18 Find the derivative of $y = 4 \arctan(x)$. What is the equation of the tangent line when $x = 1$?

3.5 More Value of Theorems

Don't just read it; fight it! Ask your own questions, look for your own examples, discover your own proofs. Is the hypothesis necessary? Is the converse true? What happens in the classical special case? What about the degenerate cases? Where does the proof use the hypothesis?

— Jacques Hadamard

We know that a function such as $f(x) = 7$ has a derivative of zero. We begin this section with a Lemma concerning this fact.

LEMMA 3.A Let $f(x)$ be differentiable on (a, b) . The function $f(x)$ is constant on $[a, b]$ if and only if $f'(x) = 0$.

Proof. Assume $f(x)$ is constant on $[a, b]$. Then $f(x) = r$ for all x in $[a, b]$, where r is some real number. Hence,

$$f'(x) = \lim_{\Delta x \rightarrow 0} \frac{f(x + \Delta x) - f(x)}{\Delta x} = \lim_{\Delta x \rightarrow 0} \frac{r - r}{\Delta x} = 0.$$

Next, assume $f'(x) = 0$ on (a, b) . Then f is continuous on $[a, b]$, and we have

$$\begin{aligned} 0 &= \lim_{\Delta x \rightarrow 0} \frac{f(x + \Delta x) - f(x)}{\Delta x} \\ 0 &= \frac{f(x + \Delta x) - f(x)}{\Delta x} \\ 0 &= f(x + \Delta x) - f(x) \\ f(x) &= f(x + \Delta x). \end{aligned}$$

Since x is an arbitrary point in (a, b) , we have that all y -values are the same; hence, f must be constant on $[a, b]$. ■

Exercise 3.5.1 Draw coordinate axes and pick any two points on your graph that have the same y -values and different x -values. Draw any differentiable function between your 2 points (make it as “curvy” as you wish, but make sure it meets the qualifications). Does your function have a horizontal tangent somewhere between your two points?

You may have difficulty drawing a differentiable function that doesn't have a horizontal tangent in the above exercise—simply because it is not possible to do so! The proof of the theorem below explains why.

The proof is adapted from Apostol (1967).

THEOREM 3.B (Rolle's Theorem) Let $f(x)$ be differentiable on (a, b) and continuous on $[a, b]$ where $f(a) = f(b)$. Then there is some point c in (a, b) such that $f'(c) = 0$.

In a proof by contradiction, we assume the opposite of the conclusion and show that this leads to a contradiction of known facts. The result is that the opposite of the conclusion must be false, and so the conclusion must be true.

Proof. We will use the method of proof by contradiction. Assume $f'(x) \neq 0$ for every x in (a, b) .

By the Extreme Value Theorem, f has a global maximum M and a global minimum m . Fermat's Test indicates that neither extreme value can be given by an x -value within (a, b) or else f' is zero there; hence, both extrema occur at the endpoints of the interval. But since $f(a) = f(b)$, we have $M = m$ and thus f is constant on (a, b) . By Lemma 3.A, this implies $f'(x) = 0$ everywhere on (a, b) ; thus, we have a contradiction. Hence, there is at least one c in (a, b) such that $f'(c) = 0$. ■

Example 3.5.2

Consider the function $f(x) = x^4 - 2x^2$ on the interval $[-2, 2]$. Since $f(-2) = f(2) = 8$, and f is differentiable and continuous, we may apply Rolle's Theorem. The theorem tells us that there is a point in the interval $[-2, 2]$ where the derivative of f is zero. We now find that point.

Since $f'(x) = 4x^3 - 4x = 4x(x^2 - 1)$ and we know that there is some point

where $f' = 0$, we set the derivative equal to zero and solve: $4x(x^2 - 1) = 0$ gives solutions $x = -1$, $x = 0$, and $x = 1$. All of these are in the interval $[-2, 2]$ and so all of these points satisfy the conclusion of Rolle's Theorem.

(Note that Rolle's Theorem simply says that there is some point in the interval where the derivative is zero; there could be multiple points where the derivative is zero, as this example illustrates.) ♦

Over any interval, must it be true that the slope between the endpoints of the interval is equal to the derivative of a differentiable function within the interval? Surprisingly, this answer is yes, and it is due to the Mean Value Theorem.

The Mean Value Theorem is another one of those Big Important Theorems!

THEOREM 3.C (The Mean Value Theorem) *Let f be a function continuous on $[a, b]$ and differentiable on (a, b) . Then, for some c in (a, b) ,*

$$\frac{f(b) - f(a)}{b - a} = f'(c)$$

provided $b \neq a$.

Proof. We note that the quantity

$$k = \frac{f(b) - f(a)}{b - a}$$

is constant since it depends only on the real constants a and b . We introduce the function $P(x) = f(x) - kx$ on $[a, b]$, where k is defined above. Then we have

$$P(x) = f(x) - kx = f(x) - \frac{f(b) - f(a)}{b - a}x = \frac{f(x)[b - a] - x[f(b) - f(a)]}{b - a}$$

so that

$$P(a) = f(a) - ka = \frac{f(a)[b - a] - a[f(b) - f(a)]}{b - a} = \frac{bf(a) - af(b)}{b - a}$$

and

$$P(b) = f(b) - kb = \frac{f(b)[b - a] - b[f(b) - f(a)]}{b - a} = \frac{-af(b) + bf(a)}{b - a}$$

implying that $P(a) = P(b)$. Hence, we may apply Rolle's Theorem. Therefore, there exists a point c in (a, b) such that $P'(c) = 0$. We have

$$\begin{aligned} P'(c) &= f'(c) - k = 0 \\ f'(c) &= k \\ f'(c) &= \frac{f(b) - f(a)}{b - a} \end{aligned}$$

which proves the theorem. ■

Some consider Rolle's Theorem as a special case of the MVT. But it isn't really: did you notice we needed Rolle to prove the MVT?

This is called the Mean Value Theorem for a reason. Geometrically, we recognize the left side as the slope of the secant line between the points $(a, f(a))$ and $(b, f(b))$. The theorem then says that the slope of the secant must be equal to the slope of the tangent to the curve at a point somewhere between $x = a$ and $x = b$. See Figure 3.6.

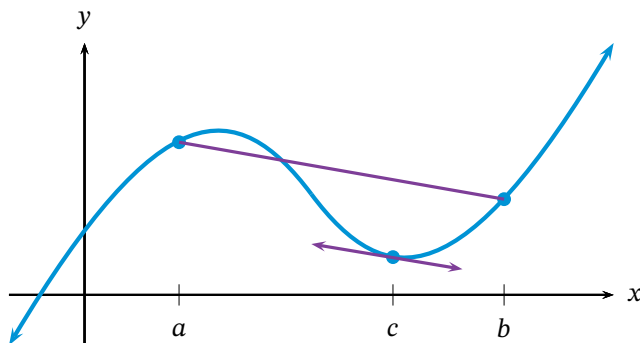


Figure 3.6 – Demonstrating the Mean Value Theorem. The secant line over $[a, b]$ has the same slope as the tangent line at $x = c$

This should make some intuitive sense: If you averaged 50 MPH on a trip in your car, then there would be *at least* one instant during your trip when your speed was *exactly* 50 MPH. In other words, the average rate of change is equal to the instantaneous rate of change at some point.

Notice that the Mean Value Theorem is another “existence theorem” – we know a point c exists, but we are given no method to find it. Fortunately, finding the value of c is not that difficult!

Well, it depends on how complicated the derivative is. . .

Example 3.5.3

Let us find the appropriate value of c given by the Mean Value Theorem for the function $f(x) = 4x^2 - x - 6$ on $[1, 3]$.

First, we compute the slope of the secant:

$$\frac{f(3) - f(1)}{3 - 1} = \frac{27 + 3}{2} = 15.$$

Next, we find that $f'(x) = 8x - 1$. Thus we take the secant slope (15), set it equal to the tangent slope ($8x - 1$) and solve to get $8x - 1 = 15$, or $x = 2$. Thus, 2 is the value of c guaranteed to exist according to the Mean Value Theorem. ♦

Exercise 3.5.4 Find the value of c guaranteed to exist by the Mean Value Theorem for the function $f(x) = x + \cos(x)$ over the interval $[0, \frac{\pi}{2}]$.

Exercise 3.5.5 Let a be a positive real number. Explain why the Mean Value Theorem cannot be used on the function $g(x) = \sqrt[3]{x}$ over the interval $[-a, a]$.

The Mean Value Theorem may be expressed in a different form. Consider multiplying both sides of

$$\frac{f(b) - f(a)}{b - a} = f'(c)$$

by $b - a$. Then we may write

$$f(b) - f(a) = f'(c)(b - a).$$

Now add $f(a)$ to both sides to get

$$f(b) = f(a) + f'(c)(b - a).$$

Since $b - a$ is the change in x -values, and $f'(c)$ is the slope, we have yet again an expression for a tangent line.

Problems for Section 3.5

1 Explain why the Mean Value Theorem does not apply to the function $y = |x|$ on any interval that includes zero.

In the following, verify the three conditions required by Rolle's Theorem and then find a suitable number c guaranteed to exist by Rolle's Theorem.

2 $f(x) = 2x^2 - 11x + 15$ on $[\frac{5}{2}, 3]$

3 $g(x) = x^3 + 5x^2 - x - 5$ on $[-5, -1]$

4 $p(x) = 4x^{4/3} - 6x^{1/3}$ on $[0, 6]$

5 $k(x) = \frac{x^2 - 4}{x^2 + 4}$ on $[-2, 2]$

In the following problems, verify the two conditions required by the Mean Value Theorem and then find a suitable number c guaranteed to exist by the Mean Value Theorem.

6 $f(x) = 4x^2 - x - 6$ on $[0, 2]$

7 $g(x) = \frac{x - 1}{x + 2}$ on $[0, 2]$

8 $p(x) = 3x^{2/3} - 2x$ on $[0, 1]$

9 $k(x) = x^4 - 3x$ on $[1, 3]$

10 The function

$$f(x) = \begin{cases} x & 0 \leq x < 1 \\ 0 & x = 1 \end{cases}$$

is zero at $x = 0$ and $x = 1$, and differentiable on $(0, 1)$, but its derivative on $(0, 1)$ is never zero. Doesn't this contradict Rolle's Theorem?

(Finney et al., 2001, p. 261)

11 Define a function f as follows:

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

(a) Sketch the graph of $f(x)$ in the interval $[0, 2]$.

(b) Show that f satisfies the conditions required by the Mean Value Theorem over the interval $[0, 2]$, and determine the values of c guaranteed to exist by the Mean Value Theorem.

(Apostol, 1967, p. 186)

12 A trucker was cited for speeding on a toll road with speed limit 70 MPH. He argued to the judge that he wasn't speeding. He handed in his ticket from a toll booth showing that in 3 hours he had covered 219 miles on the toll road. Determine whether the trucker was actually speeding.

13 Let $f(x) = 1 - x^{2/3}$. Show that $f(1) = f(-1) = 0$, but that $f'(x) \neq 0$ in the interval $[-1, 1]$. Explain how this is possible, in view of Rolle's Theorem.

(Apostol, 1967, p. 186)

14 To save up for a car, you take a job working 10 hours a week at the library. For the first six weeks, the library pays you \$8 an hour. After that, you earn \$11.50 an hour. You put all the money you earn each week in a savings account. On the day you start work your savings account already holds \$200. Let $S(t)$ be the function that describes the amount in your savings account t weeks after your library job begins.

(a) Find the values $S(3)$, $S(6)$, $S(8)$, $S'(3)$, $S'(6)$, and $S'(8)$, if possible, and describe their meaning in practical terms. If it is not possible to find one or more of these values, explain why.

(b) Write an equation for the function $S(t)$. Be sure that your equation correctly produces the values you calculated in part (a).

(c) Sketch a labeled graph of $S(t)$. By looking at the graph, determine whether $S(t)$ is continuous and whether it is differentiable. Explain the practical significance of your answers.

(d) Show algebraically that $S(t)$ is continuous but not differentiable.

(Taalman and Kohn, 2014, p. 186)

3.6 l'Hôpital's Rule

A man is like a fraction whose numerator is what he is and whose denominator is what he thinks of himself. The larger the denominator, the smaller the fraction.

— Leo Tolstoy

In the previous chapter, we discovered the values of the following

contained 25,000 gallons of oil, approximately how many more hours will elapse in the worst case before all the oil has spilled? In the best case? (Finney et al., 2001, p. 362)

- 5 Assume the following function f is a decreasing continuous function on the interval $0 \leq x \leq 4$ and that the following is a table showing some function values.

x	0	1	1.5	3	4
$f(x)$	4	3	2	1.5	1

- (a) Use the trapezoid rule to approximate the area under f . (Be careful! The subintervals are not

all the same width!)

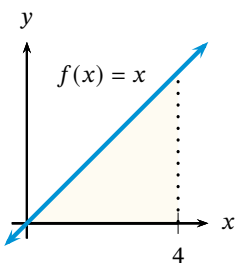
- (b) Is your answer to part (a) larger or smaller than the actual area under f ? How do you know?
- 6 Consider the area under $y = 5$ over the interval $-10 \leq x \leq 10$.
- (a) Sketch the region and find the area using a basic geometry area formula.
- (b) Use the trapezoid rule on 5 equal subintervals and compare this with the answer from part (a). Does this result make sense? Why or why not?

4.4 Definite Integrals

Success and failure have much in common that is good. Both mean you're trying.

— Frank Tyger

As we have seen, Riemann sums provide an approximation to the area under a curve, and the smaller the subinterval width, the more rectangles are used, and the better our approximation becomes. In this section, we will explore this idea of using more and rectangles (on smaller and smaller subintervals) to determine exact areas under certain curves. To facilitate this, we need new symbolism to represent the exact area under f .



The exact area under a bounded function $f(x)$ on the interval $[a, b]$ is denoted by

$$\int_a^b f(x) dx.$$

The symbol \int_a^b is called the **definite integral** from a to b .

For example, using basic geometry, it is easy to determine the exact area under the function $f(x) = x$ from 0 to 4. Since the area is that of a right triangle of base 4 and height 4, we see that the exact area must be $\frac{1}{2}(4)(4) = 8$. Hence, we say that the definite integral is equal to 8 or, in symbols,

$$\int_0^4 f(x) dx = \int_0^4 x dx = 8.$$

Previously, in Section 4.2, we spent some time discussing the fact that areas can be found under bounded functions. We clarify this discussion of exact areas (definite integrals) with an additional fact concerning such functions.

THEOREM 4.F Let f be an increasing bounded function on the interval $[a, b]$. Let A be any number that satisfies

$$\Delta x \sum_{k=0}^{n-1} f(x_k) < A < \Delta x \sum_{k=1}^n f(x_k)$$

for every integer $n > 1$. Then $A = \int_a^b f dx$ as $n \rightarrow \infty$.

Note that there is a similar statement for decreasing bounded functions. As the argument is very similar, we will omit the proof for decreasing bounded functions.

Proof. The series above are nothing more than the definitions of LRS and RRS, respectively. We already know that these are approximations, so there must be some number A such that $\text{LRS} < A < \text{RRS}$. By the definition of a definite integral, we know that $\int_a^b f dx$ satisfies the same inequalities as A . So we must prove that A is the definite integral. We begin by finding the difference between the left and right Riemann sums.

$$\begin{aligned} \text{RRS} - \text{LRS} &= \Delta x \sum_{k=1}^n f(x_k) - \Delta x \sum_{k=0}^{n-1} f(x_k) \\ &= \Delta x \left(\sum_{k=1}^n f(x_k) - \sum_{k=0}^{n-1} f(x_k) \right) \\ &= \Delta x \left(\sum_{k=1}^{n-1} f(x_k) + f(x_n) - \left[f(x_0) + \sum_{k=1}^{n-1} f(x_k) \right] \right) \\ &= \Delta x \left(f(x_n) - f(x_0) + \sum_{k=1}^{n-1} f(x_k) - \sum_{k=1}^{n-1} f(x_k) \right) \\ &= \Delta x (f(x_n) - f(x_0)) \\ &= \frac{b-a}{n} (f(b) - f(a)) \end{aligned}$$

which we can write as

$$\text{RRS} - \text{LRS} = \frac{k}{n}$$

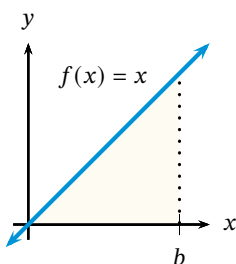
where $k = (b-a)(f(b) - f(a))$ is constant since it only depends on constants a and b . Clearly, the difference between the left and right Riemann sums must be greater than the difference between A and the definite integral, since each of these is between the Riemann sums. Then

$$0 \leq \left| A - \int_a^b f dx \right| \leq \frac{k}{n}$$

for every $n > 1$. Letting $n \rightarrow \infty$, we see that $\frac{k}{n} \rightarrow 0$, so that this inequality approaches the equality $A - \int_a^b f dx = 0$, or $A = \int_a^b f dx$. ■

The preceding proof gives us an indication as to how to compute the definite integral. Let us re-do the previous example. This time, we want the area from 0 to any number b .

Again, using basic geometry, we may determine the exact area under the function $f(x) = x$ from 0 to b . The area is that of a right



triangle of base b and height b , so the exact area must be $\frac{1}{2}(b)(b) = \frac{1}{2}b^2$. Hence,

$$\int_0^b f(x) dx = \int_0^b x dx = \frac{b^2}{2}.$$

But how can we determine the exact area under the standard parabola $f(x) = x^2$ over the interval $[0, b]$? For this, we rely on the previous theorem and the formulas for sum of powers of integers from Section 4.1. We will use a Riemann sum and find its limit as $n \rightarrow \infty$.

We wish to use n equal subintervals. Since the interval $[0, b]$ has length $b - 0 = b$, each subinterval has length

$$\Delta x = \frac{b - 0}{n} = \frac{b}{n}.$$

Since the theorem guarantees that the answer is the same whether we use a left-hand or right-hand Riemann sum, we can choose whichever one we want; let us choose a right-hand sum. Hence, the x -coordinates we use are the right-hand endpoints of the subintervals $[0, \Delta x]$, $[\Delta x, 2\Delta x]$, $[2\Delta x, 3\Delta x]$, and so on, up to $[(n-1)\Delta x, n\Delta x]$. We can represent each right-hand endpoint as $x_k = k\Delta x = kb/n$ for $k = 1, 2, 3, \dots, n$. Thus, we see that

$$\int_0^b x^2 dx = \lim_{n \rightarrow \infty} \sum_{k=1}^n x_k^2 \Delta x = \lim_{n \rightarrow \infty} \sum_{k=1}^n \left(\frac{kb}{n}\right)^2 \left(\frac{b}{n}\right).$$

Now we expand this sum, and take the limit.

$$\begin{aligned} \int_0^b x^2 dx &= \lim_{n \rightarrow \infty} \sum_{k=1}^n \left(\frac{kb}{n}\right)^2 \left(\frac{b}{n}\right) = \lim_{n \rightarrow \infty} \sum_{k=1}^n \left(\frac{k^2 b^2}{n^2} \cdot \frac{b}{n}\right) \\ &= \lim_{n \rightarrow \infty} \sum_{k=1}^n \left(\frac{k^2 b^3}{n^3}\right) = \lim_{n \rightarrow \infty} \frac{b^3}{n^3} \sum_{k=1}^n k^2 \\ &= \lim_{n \rightarrow \infty} \frac{b^3}{n^3} \left(\frac{n(n+1)(2n+1)}{6}\right) = \lim_{n \rightarrow \infty} \frac{b^3 (2n^3 + 3n^2 + n)}{6n^3} \\ &= \lim_{n \rightarrow \infty} \frac{b^3 \cdot 2n^3}{6n^3} = \frac{2b^3}{6} = \frac{b^3}{3}. \end{aligned}$$

Hence, we find that

$$\int_0^b x^2 dx = \frac{b^3}{3}.$$

Exercise 4.4.1 Compute $\int_0^4 x^2 dx$.

We may use a similar procedure to find the area under $f(x) = x^3$

over the interval $[0, b]$. We have

$$\begin{aligned} \int_0^b x^3 dx &= \lim_{n \rightarrow \infty} \sum_{k=1}^n \left(\frac{kb}{n} \right)^3 \left(\frac{b}{n} \right) = \lim_{n \rightarrow \infty} \sum_{k=1}^n \left(\frac{k^3 b^3}{n^3} \cdot \frac{b}{n} \right) \\ &= \lim_{n \rightarrow \infty} \sum_{k=1}^n \left(\frac{k^3 b^4}{n^4} \right) = \lim_{n \rightarrow \infty} \frac{b^4}{n^4} \sum_{k=1}^n k^3 \\ &= \lim_{n \rightarrow \infty} \frac{b^4}{n^4} \left(\frac{n(n+1)}{2} \right)^2 = \lim_{n \rightarrow \infty} \frac{b^4 (n^4 + 2n^3 + n^2)}{4n^4} \\ &= \lim_{n \rightarrow \infty} \frac{b^4 \cdot n^4}{4n^4} = \frac{b^4}{4}. \end{aligned}$$

If you are thinking that the area under x^4 over $[0, b]$ is $b^5/5$, you are correct!

By now, you should see a relationship between the function under which we want the area, and the area. As this applies to power functions, we will prove the the following theorem, which establishes the relationship.

THEOREM 4.G (Definite Integral of Power Functions) *If p is a positive integer and $b > 0$, then*

$$\int_0^b x^p dx = \frac{b^{p+1}}{p+1}.$$

Proof. We begin with the restatement of an algebraic fact for integer $p \geq 1$:

$$a^p - b^p = (a - b) (a^{p-1} + a^{p-2}b + a^{p-3}b^2 + \cdots + ab^{p-2} + b^{p-1}),$$

where there are exactly p terms in the rightmost parentheses. Let $a = n+1$ and $b = n$ in the identity above. Then, since $a - b = n+1 - n = 1$, we have, for any positive integer $n \geq 1$,

$$\begin{aligned} (n+1)^{p+1} - n^{p+1} &= (n+1)^p + (n+1)^{p-1}n + \cdots + (n+1)n^{p-1} + n^p \\ &< (n+1)^p + (n+1)^{p-1}(n+1) + \cdots \\ &\quad + (n+1)(n+1)^{p-1} + (n+1)^p \\ &= (n+1)^p + (n+1)^p + \cdots + (n+1)^p + (n+1)^p \\ &= (p+1)(n+1)^p. \end{aligned}$$

Similarly, we find that $(n+1)^{p+1} - n^{p+1} > (p+1)n^p$. Hence,

$$n^p < \frac{(n+1)^{p+1} - n^{p+1}}{p+1} < (n+1)^p.$$

Now let n run from all integer values from 0 to some positive integer k , and add up each term on all sides of the inequality. We get

$$\begin{aligned} 0^p + 1^p + \cdots + k^p &< \frac{(k+1)^{p+1} - k^{p+1} + k^{p+1} - (k-1)^{p+1} + \cdots + 1^{p+1} - 0^{p+1}}{p+1} \\ &< 1^p + 2^p + \cdots + (k+1)^p \end{aligned}$$

which simplifies to

$$\sum_{n=1}^k n^p < \frac{(k+1)^{p+1}}{p+1} < \sum_{n=1}^{k+1} n^p.$$

Changing indicies slightly, we get the more useful form

$$\sum_{n=1}^{k-1} n^p < \frac{k^{p+1}}{p+1} < \sum_{n=1}^k n^p$$

which holds for every positive integer $n \geq 1$ and every positive integer $p \geq 1$. Multiplication of these inequalities by b^{p+1}/k^{p+1} gives us

$$\frac{b}{k} \sum_{n=1}^{k-1} \left(\frac{nb}{k}\right)^p < \frac{b^{p+1}}{p+1} < \frac{b}{k} \sum_{n=1}^k \left(\frac{nb}{k}\right)^p.$$

Letting $f(x) = x^p$, and $x_n = nb/k$ for $n = 0, 1, 2, \dots, k$, then these inequalities become

$$\frac{b}{k} \sum_{n=1}^{k-1} f(x_n) < \frac{b^{p+1}}{p+1} < \frac{b}{k} \sum_{n=1}^k f(x_n)$$

which are exactly the expressions for the left and right Riemann sums of x^p on the interval $[0, b]$. Since $f(x) = x^p$ is increasing on this interval, we may apply Theorem 4.F. Thus,

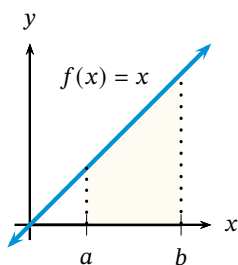
$$\int_0^b x^p dx = \frac{b^{p+1}}{p+1}. \quad \blacksquare$$

Example 4.4.2

With this Theorem, we may calculate the following:

$$\int_0^2 x^6 dx = \frac{2^7}{7} = \frac{128}{7}.$$

In this manner, we have easily found the exact area under the curve $y = x^6$ on the interval $[0, 2]$ to be $\frac{128}{7}$. \blacklozenge



Exercise 4.4.3 Compute $\int_0^4 x^3 dx$.

Finding the area over $[0, b]$ is a step in the right direction, but we want to make this definite integral as general as possible. So next we find the area under $f(x) = x$ over the interval $[a, b]$.

Again, using basic geometry, we may determine the exact area under the function $f(x) = x$ from a to b . The area is that of a right triangle of base b and height b , with another right triangle of base a and height a cut off from it. The exact area must then be $\frac{1}{2}b^2 - \frac{1}{2}a^2$. Hence,

$$\int_a^b x dx = \frac{b^2}{2} - \frac{a^2}{2}.$$

Exercise 4.4.4 Compute $\int_3^5 x dx$.

Exercise 4.4.5 Conjecture the formula for $\int_a^b x^p dx$.

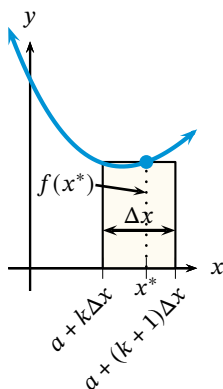


Figure 4.8 – The k th rectangle in a Riemann sum. The height of the rectangle is determined by $f(x^*)$. The width of the rectangle is $\Delta x = (b - a)/n$, where n is the number of subintervals. The left-hand side of the rectangle is a distance of $k\Delta x$ from the right endpoint a and the right-hand side is $k\Delta x + \Delta x = (k + 1)\Delta x$ from a .

You may have been wondering about the significance of the differential dx in the expression for the definite integral. This can be thought of as arising from the definition of Riemann sums for the area between f and the x -axis.

Consider for instance the area under a function $f(x)$ over the interval $[a, b]$. If we use rectangles to approximate the area, we must split the interval up into n subintervals. Thus the width of each one is $(b - a)/n = \Delta x$, and the height of each one is given by $f(x^*)$, where x^* may be the left endpoint, right endpoint, or midpoint of $[a, b]$ (actually, x^* can be any point within the subinterval). Hence, the area of any rectangle is *height* \times *width* $= f(x^*)\Delta x$. By using more and more rectangles — that is, letting $n \rightarrow \infty$, which forces $\Delta x \rightarrow 0$ — we arrive at infinitely many rectangles with height described by $f(x)$ and with infinitely small widths now described by dx rather than Δx . Symbolically,

$$\lim_{n \rightarrow \infty} \sum_{k=1}^n f(x_k^*)\Delta x = \int_a^b f(x) dx.$$

Thus, as we pass to the infinite, the measurable quantity Δx becomes the infinitely small quantity dx .

Example 4.4.6

Let us make the relationship between Riemann sums and definite integrals more explicit. In this example, we will compute a definite integral using Riemann sums as the number of subintervals goes to infinity.

Suppose we wish to compute $\int_1^5 3x^2 dx$. By definition,

$$\int_1^5 3x^2 dx = \lim_{n \rightarrow \infty} \sum_{k=1}^n 3(x_k^*)^2 \Delta x.$$

Since we need n subintervals, we set $\Delta x = (5 - 1)/n = 4/n$. Choosing the right hand endpoint of each subinterval gives us

$$x_k^* = 1 + k\Delta x = 1 + \frac{4k}{n}$$

so that

$$3(x_k^*)^2 = 3 \left(1 + \frac{4k}{n}\right)^2 = 3 \left(1 + \frac{8k}{n} + \frac{16k^2}{n^2}\right) = 3 + \frac{24k}{n} + \frac{48k^2}{n^2}.$$

Now we are ready to evaluate the limit.

$$\begin{aligned} \int_1^5 3x^2 dx &= \lim_{n \rightarrow \infty} \sum_{k=1}^n 3(x_k^*)^2 \Delta x = \lim_{n \rightarrow \infty} \sum_{k=1}^n \left(3 + \frac{24k}{n} + \frac{48k^2}{n^2}\right) \left(\frac{4}{n}\right) \\ &= \lim_{n \rightarrow \infty} \left(\sum_{k=1}^n 3 + \frac{24}{n} \sum_{k=1}^n k + \frac{48}{n^2} \sum_{k=1}^n k^2 \right) \left(\frac{4}{n}\right) \\ &= \lim_{n \rightarrow \infty} \left(3n + \frac{24}{n} \cdot \frac{n(n+1)}{2} + \frac{48}{n^2} \cdot \frac{n(n+1)(2n+1)}{6} \right) \left(\frac{4}{n}\right) \\ &= \lim_{n \rightarrow \infty} \left(12 + \frac{48(n+1)}{n} + \frac{32(n+1)(2n+1)}{n^2} \right) \\ &= 12 + 48 + 64 = 124. \end{aligned}$$

This agrees with the answer obtained from using the Power Rule. \blacklozenge

It is also important to know how to reverse the process. That is, given a Riemann sum, what is the definite integral the sum represents?

Example 4.4.7

We will find the definite integral represented by the Riemann sum

$$\lim_{n \rightarrow \infty} \sum_{i=1}^n \left(2 + \frac{3i}{n}\right)^3 \frac{3}{n}.$$

Note that the fraction on the end, $3/n$, must be the width of each subinterval: $\Delta x = 3/n$. The interval over which we want the sum must then have a total width of 3. The expression $2 + 3i/n$ must be the x -coordinate; hence, $x_i^* = 2 + 3i/n$. In terms of Δx , this is $x_i^* = 2 + i\Delta x$. This allows us to see that 2 must be the starting point of the definite integral. Since the width is 3, the definite integral must be from 2 to $2 + 3 = 5$. The summation can then be written in terms of x -coordinates, and this gives us our integral.

$$\lim_{n \rightarrow \infty} \sum_{i=1}^n \left(2 + \frac{3i}{n}\right)^3 \frac{3}{n} = \lim_{n \rightarrow \infty} \sum_{i=1}^n (x_i^*)^3 \Delta x = \int_2^5 x^3 dx. \quad \blacklozenge$$

We close this section by stating a useful property of definite integrals, the proof of which is left as a problem.

THEOREM 4.H (The Linearity Property) *Let f and g be bounded continuous functions on the interval $[a, b]$. Let k_1 and k_2 be real constants. Then*

$$\int_a^b [k_1 f(x) + k_2 g(x)] dx = k_1 \int_a^b f(x) dx + k_2 \int_a^b g(x) dx.$$

This is the property that allows us to find the area under any polynomial. Consider a polynomial such as $f(x) = 3x^2 + 7x + 1$. This linearity property says that we can find the area under $f(x) = 3x^2 + 7x + 1$ by finding the area under each of x^2 , x , and 1. Then multiply the area under x^2 by 3 and the area under x by 7. Finally, add the areas under $3x^2$, $7x$, and 1 together.

The following example illustrates this idea. We will compute the area under a polynomial.

Example 4.4.8

Since areas may be added or subtracted from other areas, we can compute

$$\begin{aligned} \int_0^2 (x^3 + 5x^2 - x) dx &= \int_0^2 x^3 dx + 5 \int_0^2 x^2 dx - \int_0^2 x dx \\ &= \frac{2^4}{4} + 5 \cdot \frac{2^3}{3} - \frac{2^2}{2} \\ &= 4 + 5 \cdot \frac{8}{3} - 2 = \frac{52}{3}. \end{aligned}$$

Hence, we have a method to find the area under any polynomial! \blacklozenge

Problems for Section 4.4

- 1 What is the area under the horizontal line $y = 5$ over the interval $[-2, 2]$? What is $\int_{-2}^2 5 \, dx$? Make a conjecture concerning the value of $\int_a^b C \, dx$ where C is any real constant.
- 2 Very often we use the *evaluation notation* to aid in computing definite integrals. For instance, writing " $x^2|_2^3$ " indicates that we should evaluate x^2 at $x = 3$ and at $x = 2$, then compute the difference: $x^2|_2^3 = 3^2 - 2^2 = 5$. Thus, we see that Theorem 4.G could be restated as

$$\int_0^b x^p \, dx = \left. \frac{x^{p+1}}{p+1} \right|_0^b = \frac{b^{p+1}}{p+1}.$$

In practice, we would then write

$$\int_1^2 x^5 \, dx = \left. \frac{x^6}{6} \right|_1^2 = \frac{2^6}{6} - \frac{1}{6} = \frac{21}{2}.$$

Use this notation to compute

$$\int_2^4 (3x^2 - x + 2) \, dx.$$

- 3 Evaluate the following.

(a) $\int_{-1}^5 (x+2) \, dx$ (c) $\int_2^8 (x-1) \, dx$

(b) $\int_1^7 x \, dx$ (d) $\int_6^{12} (x-5) \, dx$

(e) $\int_{1+k}^{7+k} (x-k) \, dx$ where k is constant

- (f) What do you notice about the answers to the preceding problems? Why do you think this is? Explain your answer in terms of area.

Evaluate the following definite integrals. Remember to use the properties of definite integrals when possible.

4 $\int_{-1}^2 (x^2 + 4x - 1) \, dx$

5 $\int_2^4 (x-1)(x+2) \, dx$

6 $\int_{-3}^{-1} (3x^2 + 2x + 1) \, dx$

7 $\int_5^2 x \, dx$

8 $\int_{1/2}^{3/2} (4x - 6) \, dx$

9 $\int_0^1 (x^{10} - 8x^7 + 20) \, dx$

10 $\int_{-1}^1 (2x-1)(x+1) \, dx$

- 11 Use the limit definition (as shown in Example 4.4.6) to compute $\int_3^7 2x \, dx$.

Write each Riemann sum as a definite integral.

12 $\lim_{n \rightarrow \infty} \sum_{i=1}^n 6 \left(\frac{3i}{n} \right) \frac{3}{n}$

13 $\lim_{n \rightarrow \infty} \sum_{i=1}^n \left(\frac{3i}{n} \right)^3 \frac{3}{n}$

14 $\lim_{n \rightarrow \infty} \sum_{i=1}^n \left(\frac{3i}{n} \right) \left[\left(\frac{3i}{n} \right) - 5 \right] \frac{3}{n}$

15 $\lim_{n \rightarrow \infty} \sum_{i=1}^n \left(2 + \frac{3i}{n} \right)^4 \frac{3}{n}$

16 $\lim_{n \rightarrow \infty} \sum_{i=1}^n \left(\frac{3i}{n} \right)^3 \frac{3}{n}$

17 $\lim_{n \rightarrow \infty} \sum_{i=1}^n \frac{i^4}{n^5}$

18 $\lim_{n \rightarrow \infty} \sum_{i=1}^n \frac{n}{n^2 + i^2}$

19 $\lim_{n \rightarrow \infty} \sum_{i=1}^n \frac{i^3}{n^4}$

20 $\lim_{n \rightarrow \infty} \sum_{i=1}^n \frac{1}{n} \sqrt{\frac{i}{n}}$

21 $\lim_{n \rightarrow \infty} \sum_{i=1}^n \frac{80i}{n^2}$

- 6 Let R be the region bounded by $y = e^x$, $y = 2$, and $x = 0$. Find the volume of the solid whose base is bounded by the region R and the cross sections perpendicular to the x -axis are semicircles.
- 7 Let R be the region bounded by $y = x^2$ and $y = x$. Find the volume of the solid whose base is bounded by the region R and the cross sections perpendicular to the x -axis are semicircles.
- 8 Let R be the region bounded by $y = \frac{1}{16}x^2$ and $y = 2$. Find the volume of the solid whose base is bounded by the region R and the cross sections perpendicular to the x -axis are rectangles whose height is twice that of the side in the plane of the base.
- 9 Find the volume of the solid whose base is bounded by the curve $y = 2\sqrt{\sin(x)}$, the lines $x = 0$, $x = \pi$, and $y = 0$, and the cross sections perpendicular to the x -axis are squares.

6.7 Disks and Washers, or Volume Part 2

When we ask advice, we are usually looking for an accomplice.
— Joseph-Louis Lagrange

In this section we consider solids formed by revolving a plane about an axis. We describe this procedure by establishing the known volume formula for a cylinder, $V = \pi r^2 h$.

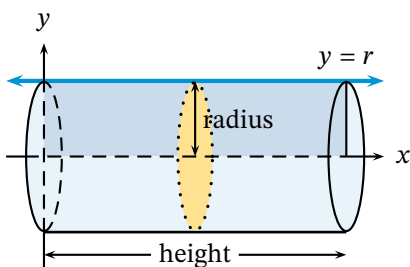


Figure 6.13 – A cylinder formed by revolving a rectangle about the x -axis.

Consider the region in the coordinate plane bounded by the line $y = r$, for constant r , and the x -axis over the interval $[0, h]$, for constant h . This region is a rectangle of dimensions r and h . But rather than treat this rectangle as a base for which we build cross sections, we will *revolve* the rectangle about the x -axis to form a solid. This creates circular cross sections of radius r with centers on the x -axis, as shown in Figure 6.13.

Since each cross section is a circle, we sum all the areas of all circles to find the volume over the entire length of the solid (the interval $[0, h]$). Hence, using the same reasoning as in the previous section,

$$V = \int_0^h \pi r^2 dx = \pi r^2 x \Big|_0^h = \pi r^2 h.$$

Consider what we have done. We took a curve $f(x)$ (which in the case of the cylinder was simply $f(x) = r$) which bounds a region with the x -axis over an interval $[a, b]$ (which in the case of the cylinder was $[0, h]$) and revolved the region around the x -axis, thereby creating a solid.

By revolving, we create circular cross sections whose radius is $f(x)$ since this is the distance from the center to the edge of each cross section. So the area of each cross section is

$$A(x) = \pi [f(x)]^2.$$

Integrating (or summing) all the circular cross sections from $x = a$ to $x = b$, we have

$$V = \int_a^b \pi [f(x)]^2 dx = \pi \int_a^b [f(x)]^2 dx.$$

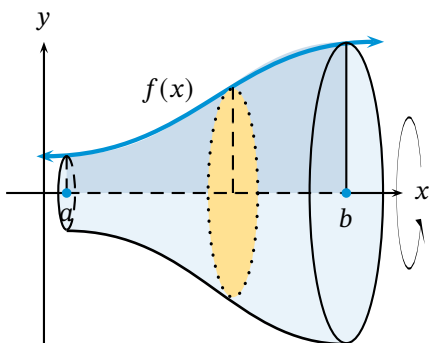


Figure 6.14 – The disk method in general.

Figure 6.14 shows a solid formed by revolving a region bounded by a general curve $f(x)$ over the interval $[a, b]$. The figure also shows a typical circular cross section; note the radius of the cross section is determined by the function $f(x)$.

Exercise 6.7.1 Establish the formula for the volume of a cone of radius r and height h using an appropriately situated region in the plane that is then revolved about the x -axis.

We clarify the method with the following definition.

Given a region R in the coordinate plane bounded by $f(x)$ and the x -axis over the interval $[a, b]$, then the volume of the solid generated by revolving R about the x -axis is given by

$$V = \pi \int_a^b [f(x)]^2 dx.$$

This method is called the **disk method** of finding volume.

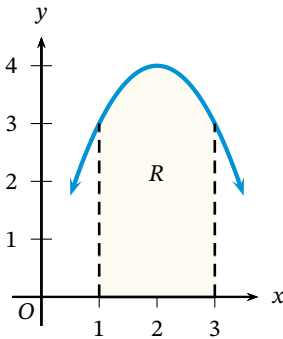


Figure 6.15 – The region in Example 6.7.2.

Example 6.7.2

Consider the region R in the plane bounded by the curve $f(x) = 4x - x^2$ and the x -axis over the interval $[1, 3]$. This region is shown in Figure 6.15. Revolving this region about the x -axis generates a solid with circular cross sections (or, disks). Hence, the volume of the solid is

$$\begin{aligned} V &= \pi \int_1^3 [f(x)]^2 dx = \pi \int_1^3 (4x - x^2)^2 dx \\ &= \pi \int_1^3 (16x^2 - 8x^3 + x^4) dx = \pi \left(\frac{16}{3}x^3 - 2x^4 + \frac{1}{5}x^5 \right) \Big|_1^3 \\ &= \pi \left(144 - 162 + \frac{243}{5} \right) - \pi \left(\frac{16}{3} - 2 + \frac{1}{5} \right) = \frac{406\pi}{15}. \end{aligned}$$

The calculator gives the approximate volume as 85.302. ♦

Example 6.7.3

Suppose the region R is bounded by the curve $y = \exp(x) + 1$ and the x -axis over the interval $[\ln(2), \ln(8)]$. Revolving this region about the x -axis again produces a solid with circular cross sections. Thus, the volume of this solid is

$$\begin{aligned} V &= \pi \int_{\ln(2)}^{\ln(8)} [f(x)]^2 dx = \pi \int_{\ln(2)}^{\ln(8)} (e^x + 1)^2 dx \\ &= \pi \int_{\ln(2)}^{\ln(8)} (e^{2x} + 2e^x + 1) dx = \pi \left(\frac{1}{2}e^{2x} + 2e^x + x \right) \Big|_{\ln(2)}^{\ln(8)} \\ &= \pi \left(\frac{1}{2}e^{\ln(64)} + 2e^{\ln(8)} + \ln(8) - 2 - 4 - \ln(2) \right) = \pi(42 + \ln(4)) \end{aligned}$$

By the calculator, we have 136.302. ♦

Suppose we change the axis of revolution to something other than the x -axis. This is what we do in the next example.

Example 6.7.4

The region R is bounded by $f(x) = \exp(x)$ and the line $y = -1$ over $[\ln(2), \ln(8)]$.

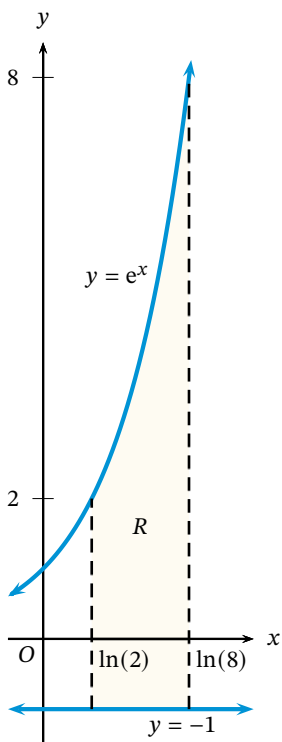


Figure 6.16 – The region in Example 6.7.4.

It is revolved around the line $y = -1$. How do we find the volume?
 To calculate this volume, we recall that the lengths of the radii are needed. By revolving around the line $y = -1$, we are making the radii 1 unit longer than if we had revolved around the line $y = 0$. In Figure 6.17, we see the situation. Note that the radii are no longer simply defined by the curve $y = e^x$. They are now one unit longer: the radii are actually $e^x + 1$.

It is this “new” function we must integrate. The volume of this solid is given by

$$V = \pi \int_{\ln(2)}^{\ln(8)} (e^x + 1)^2 dx,$$

which is the same integral as in the previous example, so the volume is $\pi(42 + \ln(4))$. ♦

Exercise 6.7.5 The region R is bounded by $y = 10 - x^3$ and the line $y = 2$ over the interval $[1, 2]$. Find the volume of the solid generated by revolving R about the line $y = 2$.

Notice that we have thus far always bounded our regions with a curve and a line, and the line has been the axis of revolution. What happens if the axis of revolution is not the line that bounds the region?

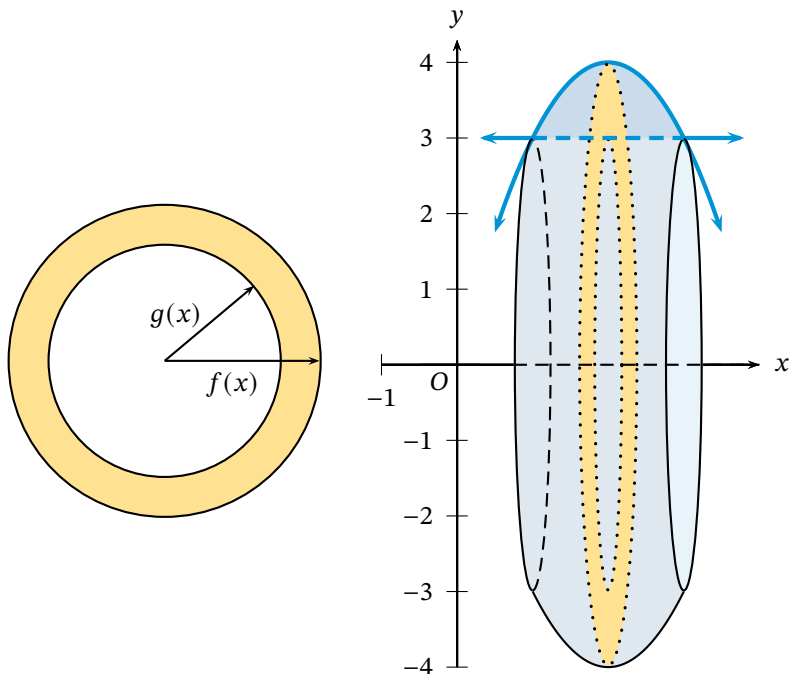


Figure 6.17 – The cross sections (left) and the resulting solid from Example 6.7.6 (right).

Example 6.7.6

The region R is bounded by $f(x) = 4x - x^2$ and the line $g(x) = 3$. It is revolved about the x -axis. What is the volume of the resulting solid?

The solid created has a hole through the middle of it; in other words, it is a doughnut-like shape. The shape of the outer edge is determined by the parabola, while the shape of the “hole” is determined by the line. The solid is shown in Figure 6.17.

This causes us to consider the cross sections once more, because they are not disks this time. Due to the hole, each cross section, while still circular, has a circular hole in it. The cross sections are like rings (or like the washers one uses with bolts and screws) and can be seen in Figure 6.17. The area of the cross section is then the area of the larger circle minus the area of the inner circle (the hole). Note that the radii of the circles are given by the distance from the center to the outer and inner edges; these distances are given by the functions f and g . Hence, the area formula is

$$A(x) = \pi[f(x)]^2 - \pi[g(x)]^2 = \pi \left([f(x)]^2 - [g(x)]^2 \right).$$

By integrating $A(x)$ over the interval, we obtain the volume.

Back to the problem: Notice that we were not explicitly given an interval over which to integrate. However, we can find it. The interval is given by the x -coordinates of the intersection points of f and g . Setting $f(x) = g(x)$ gives $4x - x^2 = 3$ whose solutions are 1 and 3, from which we get an interval of $[1, 3]$.

Thus, using the washer method, we get a volume of

$$\begin{aligned} V &= \pi \int_1^3 \left([f(x)]^2 - [g(x)]^2 \right) dx = \pi \int_1^3 \left([4x - x^2]^2 - 3^2 \right) dx \\ &= \pi \int_1^3 \left(16x^2 - 8x^3 + x^4 - 9 \right) dx = \pi \left(\frac{16}{3}x^3 - 2x^4 + \frac{1}{5}x^5 - 9x \right) \Big|_1^3 = \frac{136\pi}{15}, \end{aligned}$$

or 28.484. ♦

The method used in the previous example has a name.

Given a region R in the coordinate plane bounded above by $f(x)$ and below by $g(x)$ over the interval $[a, b]$, then the volume of the solid generated by revolving R about the x -axis is given by

$$V = \pi \int_a^b \left([f(x)]^2 - [g(x)]^2 \right) dx.$$

This method is called the **washer method** of finding volume.

Next, we use this washer method again to find the volume of a solid of revolution.

Example 6.7.7

The region R is bounded by $f(x) = 8x - x^2$ and $g(x) = x^2$. Find the volume of the solid generated as R is revolved about the x -axis.

As in the previous example, we must find the intersection points so we know what interval we are to integrate over. Setting $8x - x^2 = x^2$ gives $x = 0$ and $x = 4$; hence the interval is $[0, 4]$.

Note that the cross sections perpendicular to the axis of revolution (the x -axis in this case) are washers. This can be seen in Figure 6.18. Hence the

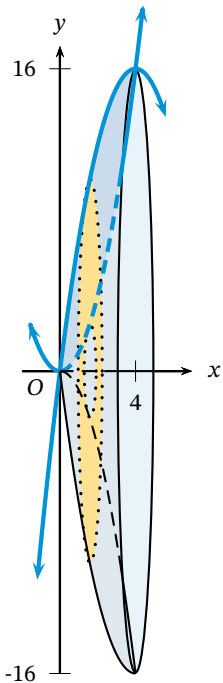


Figure 6.18 – A solid formed by revolving a region bounded by two parabolas.

volume is given by

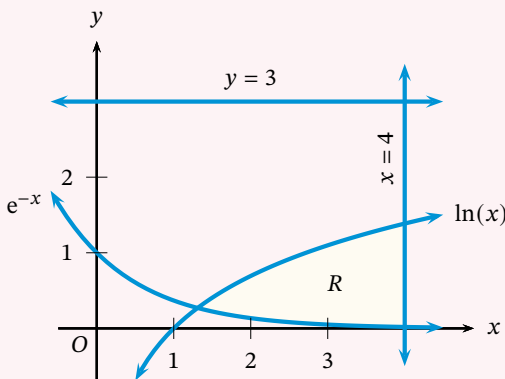
$$\begin{aligned}
 V &= \pi \int_0^4 \left([8x - x^2]^2 - [x^2]^2 \right) dx \\
 &= \pi \int_0^4 (64x^2 - 16x^3 + x^4 - x^4) dx \\
 &= \pi \int_0^4 (64x^2 - 16x^3) dx \\
 &= \pi \left(\frac{64}{3}x^3 - 4x^4 \right) \Big|_0^4 = \frac{1024\pi}{3}
 \end{aligned}$$

or 1072.33. ♦

Finally, we look at an example where we use the graphing calculator to help us.

Example 6.7.8

Suppose the region R is bounded by $f(x) = \ln(x)$, $g(x) = e^{-x}$, and the line $x = 4$. Find the volume of the solid generated by revolving R about the line $y = 3$. (See the figure below.)



To calculate the volume, we must first determine the length of the radii. Since the region being revolved is not bounded by the axis of revolution, there will be a hole created in the solid, and therefore we must determine the length of the radii from the center to the inner and outer edges of the region.

Hence, the length of the radius to the outer edge of the region is equal to the distance between the axis of revolution and the function. In this case, that is $G(x) = 3 - e^{-x}$. The distance from the axis of revolution to the inner function is $F(x) = 3 - \ln(x)$.

Next, we must find the interval. Using the calculator, we find the intersection point of $Y1 = 3 - e^{-X}$ and $Y2 = 3 - \ln(X)$ to be $X = 1.309799586$. Storing this as A so we may use it later (press $X \text{ STO } A$), we set up the volume integral.

This time, because the axis of revolution is *above* the region, G is the outer radius and F is the inner radius. Hence, we need to integrate $G^2 - F^2$. The volume integral is then

$$V = \pi \int_{1.309799586}^4 \left([3 - e^{-x}]^2 - [3 - \ln(x)]^2 \right) dx.$$

Using the calculator, we enter

$$\pi * \text{fnInt}((Y1)^2 - (Y2)^2, X, A, 4)$$

where we have the functions entered in Y1 and Y2, and A is the stored x -coordinate of the intersection point. The calculator gives the approximation 34.391. ♦

Problems for Section 6.7

The region R is bounded by the given curves and/or lines. Find the volume of the solid generated by revolving R around the given axis. You may find graphing the curves on your calculator useful.

1 $y = \cos(x)$, $y = \sin(x)$, $x = 0$, $x = \pi/4$; axis: x -axis

2 $y = 1/\sqrt{x}$, $x = e$, $x = e^3$, $y = 0$; axis: x -axis

3 $y = 3 - x^2$, $y = -1$; axis: $y = -1$

4 $y = 16x - 4x^2$, $y = 0$; axis: $y = -20$

5 $y = (x + 3)^3$, $y = 0$, $x = 2$; axis: $y = -1$

6 For the following problems, set up the required volume integral, then evaluate that integral on your calculator.

- (a) The region R is bounded by the curve $y = \sin x \cos x$ and the x -axis from $x = 0$ to $x = \frac{\pi}{2}$. Find the volume of the solid generated by revolving R about the x -axis.

(b) The region R is bounded by the curve $y = e^x$ and the lines $y = 2$ and $x = -1$. Find the volume of the solid generated by revolving R about the line $y = e$.

(c) The region R is bounded by the curve $16y^2 + 9x^2 = 144$ and the line $4y = 3x + 12$ in Quadrant II. Find the volume of the solid generated by revolving R about the x -axis.

7 (Calculator) Let R be the region bounded by the graph of $f(x) = e^{2x-x^2}$ and the graph of $g(x) = \sec(x)$.

- (a) Find the area of the region R .
- (b) The region R is revolved about the x -axis. Find the volume of the resulting solid.
- (c) The region R is revolved about the line $y = 3$. Find the volume of the resulting solid.
- (d) The region R is the base of a solid whose cross sections perpendicular to the x -axis are squares. Find the volume of this solid.

6.8 Cylindrical Shells, or Volume Part 3

Anyone who cannot cope with mathematics is not fully human. At best he is a tolerable subhuman who has learned to wear shoes, bathe, and not make messes in the house.

— Robert A. Heinlein

You may have noticed when we discussed disks and washers that we never revolved our region about any vertical axis. It is to this problem we now turn. Revolving a region around the y -axis can be done more efficiently with a method other than disks or washers. (That is not to say that it is impossible to use the disk and washer method. See problem 9 for instance.) To describe this new method we compute the volume of the solid generated by revolving around the y -axis the region bounded by $f(x) = (x - 2)^2$ in the first quadrant.

Instead of creating circular cross sections as in the disk method, we consider this solid built of many hollow cylinders, each inside another. Figure 6.19 shows the region, the solid, and a few such “cylindrical shells.” Although we are not using a cross section method, we still compute volume by a sum; in this case, we sum all the cylindrical shells.

The thickness of each shell is infinitesimally small (as with the disks). As such, it is as if the shells consist only of the *surface* of a cylinder. The areas of the surfaces is what we sum. The area of each surface is simply that of a rectangle—a rectangle that has been

The shell is just the outer portion of a cylinder without a top or bottom. You can think of a cylindrical shell as the label on a can of soup.

Chapter 8

INFINITE SERIES

HUNDREDS OF YEARS AGO, mathematicians considered infinite series simply as polynomials with infinitely many terms. It was therefore assumed that such “infinite polynomials” had the same properties as finite polynomials, but they do not. The question of convergence must be considered. In this chapter, we attempt to answer that question by introducing techniques for determining convergence of certain kinds of infinite series. However, we will begin by extending the tangent line concept to show how terms of a polynomial may be used as an approximator to a function.

8.1 Approximating Polynomials

To admit error and cut losses is rare among individuals, unknown among states. States function only in terms of what those in control perceive as power or personal ambition, and both of those wear blinkers.

— Barbara Tuchman, *A Distant Mirror: The Calamitous 14th Century*

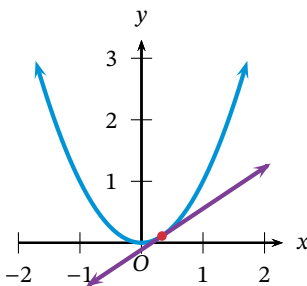


Figure 8.1 – The graph of $y = x^2$ and a tangent line.

Many applied problems cannot be solved exactly. This is why we have approximation techniques like Riemann sums that allow us to estimate answers. In this section, we use knowledge of the derivatives of a function to arrive at an approximation of that function.

We will begin with investigating one-degree polynomial approximations to a function $f(x)$ at a point $x = a$; in other words, we begin with finding a line that approximates values near $f(a)$. Such a line we have already discovered: it is the tangent line to $f(x)$ at $x = a$. Notice that the tangent line and f are equal when $x = a$, and for values of x “close” to a , the tangent line has values that are “close” to the function. In Figure 8.1 we have the graph of $f(x) = x^2$ and the tangent line to f at $x = \frac{1}{3}$. By established methods, we can easily find the tangent line to be

$$g(x) = \frac{1}{9} + \frac{2}{3}\left(x - \frac{1}{3}\right).$$

Note that when $x = \frac{1}{3}$, both the function f and the tangent line g have

the same value. For values close to $x = \frac{1}{3}$, the tangent line provides an approximation to f . Indeed, when $x = \frac{2}{3}$, we see that

$$f\left(\frac{2}{3}\right) = \frac{4}{9} \quad \text{and} \quad g\left(\frac{2}{3}\right) = \frac{2}{9}.$$

Exercise 8.1.1 In your calculator on the Y= menu, enter $Y1 = X^2$ and $Y2 = 1/9 + 2(X-1/3)/3$. In the Table Setup, set $TblStart=0$ and $\Delta Tbl=0.05$. When does the difference between the function and its tangent line become greater than 0.5?

We may redefine the tangent line to reflect its application as an approximation tool.

The **linearization**, or **linear approximation** of a function f at $x = a$ is

$$P_1(x) = f(a) + f'(a)(x - a)$$

where a is called the **center** of the approximation. We use the notation $P_1(x)$ to signify that this is a one-degree polynomial approximation.

Example 8.1.2

Let us find the linearization of $f(x) = \sin(x)$ centered at $x = 0$, and use it to approximate $\sin(0.12)$.

The tangent line is easily found to be

$$P_1(x) = \sin(0) + \cos(0)(x - 0) = x.$$

At $x = 0.12$, we have $P_1(0.12) = 0.12$; hence, $\sin(0.12) \approx 0.12$. (Note that a calculator gives $\sin(0.12) \approx 0.1197$.)

We remark that using this linear approximation to estimate $\sin(4.5)$ would be useless: the tangent line gives 4.5, but we know $-1 \leq \sin(x) \leq 1$. Hence, we see what happens when we move too far from the center. If we wanted to approximate $\sin(4.5)$, we should find a new tangent line centered somewhere closer to 4.5. ♦

Example 8.1.3

We can use linear approximations to approximate numerical values such as $\sqrt{77}$. We do this by finding the linearization of $f(x) = \sqrt{x}$ centered at $x = 81$. Why $x = 81$? Because the value we are interested in approximating is when $x = 77$, it would be foolish indeed to center this at, say, $x = 1$ since that center would be too far away from the value we wish to approximate. We choose 81 since it is the closest perfect square to 77.

The derivative of f is $f'(x) = \frac{1}{2\sqrt{x}}$. The linearization is then

$$P_1(x) = f(81) + f'(81)(x - 81) = 9 + \frac{1}{18}(x - 81).$$

So $P_1(77) = 8.778 \approx \sqrt{77}$. (Again, the calculator gives $\sqrt{77} \approx 8.775$.) ♦

Exercise 8.1.4 Find the linearization of $\ln(x)$ centered at $x = 1$ and use it to approximate $\ln(0.9)$.

Exercise 8.1.5 Let k be any nonzero constant. Find the linearization of

$f(x) = (1+x)^k$ centered at $x = 0$. Use this linearization to approximate $\sqrt{1.06}$; $\sqrt[3]{1.06}$; $\frac{1}{1.06}$; and $\frac{1}{(1.06)^2}$.

So now the question is how do we get a better approximation than the linear one? Adding a quadratic term to our approximation would be beneficial since it could match the curvature of the function. So we should use something like the following, to approximate $f(x)$ centered at $x = a$:

$$P_2(x) = A + B(x - a) + C(x - a)^2.$$

How do we know what the coefficients A , B , and C are? Well, certainly we want the values of f and P_1 to be the same at $x = a$. Therefore, we need $f(a) = P_2(a) = A$. Also, we want P_2 and f to be tangent; thus, we need $f'(a) = P_2'(a)$. Computing $P_2'(x)$ we get

$$P_2'(x) = B + 2C(x - a)$$

so that $P_2'(a) = B = f'(a)$. This means that so far we have

$$P_2(x) = f(a) + f'(a)(x - a) + C(x - a)^2$$

and all that remains is to determine C .

We added the quadratic term so the approximating polynomial can curve as the function does. This means that we need the *concavity* of $P_2(x)$ to match that of $f(x)$ at $x = a$. This implies that we need $f''(a) = P_2''(a)$. Computing,

$$P_2''(x) = 2C$$

so then $P_2''(a) = 2C = f''(a)$; this implies $C = \frac{1}{2}f''(a)$. Now we make the following definition.

The **quadratic approximation** of a function f at $x = a$ is

$$P_2(x) = f(a) + f'(a)(x - a) + \frac{1}{2}f''(a)(x - a)^2$$

where a is the center.

Example 8.1.6

In Exercise 8.1.4, you found that $P_1(x) = x - 1$ was the linear approximation of $f(x) = \ln(x)$ at $x = 1$. Now we find the quadratic approximation at $x = 1$.

Since $f''(x) = -1/x^2$, we see that $f''(1) = -1$. Then the quadratic approximation is

$$P_2(x) = x - 1 - \frac{1}{2}(x - 1)^2.$$

Now we have a better approximation for $\ln(0.9)$:

$$P_2(0.9) = 0.9 - 1 - \frac{1}{2}(0.9 - 1)^2 = -0.1 - 0.005 = -0.105.$$

The calculator gives $\ln(0.9) \approx -0.1053$. ♦

Exercise 8.1.7 Find the quadratic approximation for $f(x) = \sin(2x) + \cos(x)$ centered at $x = 0$. Graph both $f(x)$ and $P_2(x)$ on your calculator on the same axes. What is an approximate interval over which the difference between f

and P_2 is less than 0.5?

It is natural to ask about a cubic approximation. In fact, we can determine such an approximation using previous methods. Here, we want to determine coefficients A , B , C , and D such that

$$P_3(x) = A + B(x - a) + C(x - a)^2 + D(x - a)^3$$

where, again, $x = a$ is the center. We will end up going through the same procedure as with the quadratic approximation to obtain A , B , and C . Thus, we know $A = f(a)$, $B = f'(a)$, and $C = \frac{1}{2}f''(a)$. It only remains to determine D . Following the pattern established, it makes sense that we need the third derivatives of f and P_3 to be equal at $x = a$. Computing, we have

$$P_3'(x) = B + 2C(x - a) + 3D(x - a)^2$$

$$P_3''(x) = 2C + 6D(x - a)$$

$$P_3'''(x) = 6D$$

so then $P_3'''(a) = f'''(a) = 6D$, or $D = \frac{1}{6}f'''(a)$. And, once more, we state this as a definition.

The **cubic approximation** of a function f at $x = a$ is

$$P_3(x) = f(a) + f'(a)(x - a) + \frac{1}{2}f''(a)(x - a)^2 + \frac{1}{6}f'''(a)(x - a)^3$$

where a is the center.

Example 8.1.8

We now find the cubic approximation to $f(x) = \sin(x)$ centered at $x = 0$. We have

$$f'(x) = \cos(x), \quad f''(x) = -\sin(x), \quad \text{and} \quad f'''(x) = -\cos(x).$$

Thus, $f'(0) = 1$, $f''(0) = 0$, and $f'''(0) = -1$. The cubic approximation is therefore

$$P_3(x) = 0 + 1 \cdot (x - 0) + \frac{1}{2} \cdot 0 \cdot (x - 0)^2 + \frac{1}{6}(-1)(x - 0)^3 = x - \frac{1}{6}x^3.$$

Now we approximate

$$\sin(0.12) \approx P_3(0.12) = 0.12 - \frac{1}{6}(0.12)^3 = 0.119712.$$

This gives a better approximation to $\sin(0.12)$ than before. We can see how well the cubic approximation “fits” the graph of $f(x) = \sin(x)$ in Figure 8.2. ♦

Exercise 8.1.9 Find the cubic approximation for $f(x) = \frac{1}{x} + x^2 - 1$ centered at $x = 1$ and use it to approximate $f(1.1)$.

Example 8.1.10

Consider this question: Approximately how much does the area of a square increase if its side length increases from 5 cm to 5.3 cm?

This is one situation in which polynomial approximations are useful. We will find the linearization of the area function, $A(s) = s^2$, where s is the side

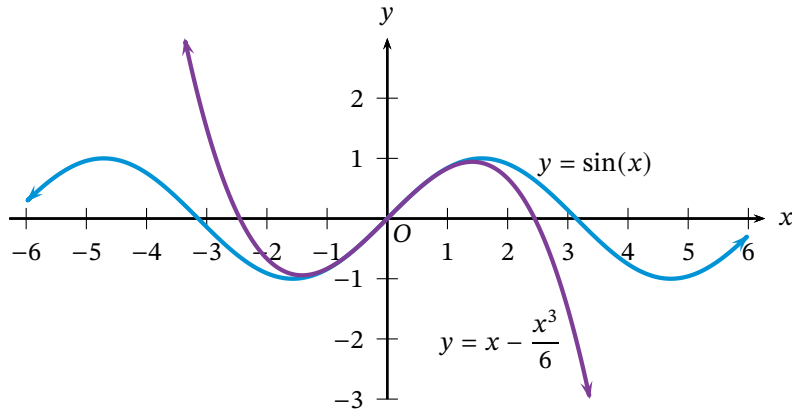


Figure 8.2 – The graph of $y = \sin(x)$ and the graph of its cubic approximating polynomial, $y = x - \frac{1}{6}x^3$.

length, centered at $s = 5$. Since $A'(s) = 2s$, we have

$$P_1(s) = A(5) + A'(5)(s - 5) = 5^2 + 2(5)(s - 5) = 25 + 10(s - 5).$$

Therefore, we see that $P_1(s) = 10s - 25$. This, however, is the approximate area of the square, not the difference between the areas when $s = 5$ and $s = 5.3$. The difference would be, since $A(5) = 25$,

$$P_1(5.3) - A(5) \approx 10(5.3) - 25 - 25 = 53 - 50 = 3$$

and the square's area increases approximately 3 square centimeters. ♦

Problems for Section 8.1

Find the linear, quadratic, and cubic approximations to each of the following, all centered at $x = 0$.

1 $f(x) = \cos(x)$

3 $f(x) = e^x$

2 $f(x) = \tan(x)$

4 $f(x) = \ln(1 + x)$

5 The best quadratic approximation for an unknown function g centered at $x = -2$ is the polynomial

$$\pi - \frac{1}{2}(x + 2) - 7(x + 2)^2.$$

Find $g(-2)$, $g'(-2)$, and $g''(-2)$.

(Cohen and Henle, 2005, p. 514)

- 6 A spherical balloon is inflated so that its radius increases from 10 cm to 10.15 cm. By approximately how much was its volume increased?
- 7 Prove that if F is a quadratic polynomial, then the quadratic approximation to F centered at any point is F itself.
- 8 The function $f(x) = x^{2/3}$ does not have a polynomial approximation of any degree centered at $x = 0$. Why?

8.2 Taylor Polynomials

Errors using inadequate data are much less than those using no data at all.

— Charles Babbage

Thanks to the work in previous section, it should be obvious that the approximating polynomial idea could be extended past the cubic. Indeed, that is true! Of course, there are some things to be careful

Chapter 9

NEW TYPES OF FUNCTIONS

THE PRESENT CHAPTER focuses on the different ways we can model motion. Those ways are with *vectors*, *parametric equations*, and *polar coordinates*. Vectors provide an easy to talk of a particle's direction and distance. Parametric equations make explicit the relationship between motion and time. Polar coordinates are introduced as a more intuitive way of expressing circular motion. All of these topics enable us to describe any curve as a kind of function.

9.1 Vectors

Biographical history, as taught in our public schools, is still largely a history of boneheads: ridiculous kings and queens, paranoid political leaders, compulsive voyagers, ignorant generals – the flotsam and jetsam of historical currents. The men who radically altered history, the great scientists and mathematicians, are seldom mentioned, if at all.

— Martin Gardner

Suppose you wanted to travel from your home to Taco Bell. There are many different ways you could get there: the direct route, on the highway; by numerous back-roads; maybe you could stop by your friend's home and take them with you. But no matter which path you take from your home to Taco Bell, two things about your journey are identical. After arriving at Taco Bell, the distance you are from home has not changed and the direction from your home has not changed. (Notice we are not concerned with the time it takes to travel the various ways from your home to Taco Bell, simply that you started at home and arrived at Taco Bell.)

In any journey, the aspects of distance (from starting point to ending point) and direction (in the sense of the “bird's eye view” a map provides) are most important. We encapsulate this notion mathematically with the idea of a **vector**. A vector is a mathematical quantity that carries with it two pieces of information: direction and distance.

Let us use a coordinate plane to refine these ideas. Suppose we start at point $A(2, 3)$ and we wish to arrive at point $B(7, 8)$. We denote this

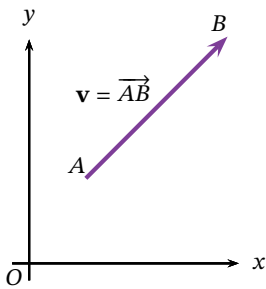


Figure 9.1 – A vector.

graphically with an arrow whose tip is at the destination, point B in this case (see Figure 9.1). We may then say that the vector \mathbf{v} represents the **displacement** from A to B ; i.e., $\mathbf{v} = \overrightarrow{AB}$.

Note that the only things known about \mathbf{v} are the direction and distance. The distance from A to B is found by using the standard distance formula:

$$\sqrt{(8 - 3)^2 + (7 - 2)^2} = \sqrt{5^2 + 5^2} = 2\sqrt{5},$$

and the direction can be considered as the angle θ between the vector \overrightarrow{AB} and a line from A parallel to the x -axis, where θ satisfies

$$\tan(\theta) = \frac{8 - 3}{7 - 2} = \frac{5}{5} = 1.$$

Thus the vector $\mathbf{v} = \overrightarrow{AB}$ has a distance of $2\sqrt{5}$ and a direction of $\arctan(1) = 45^\circ$.

The important aspect of a vector is that a vector is considered a **free vector**. That is, the precise location of the vector is irrelevant as long as direction and distance are preserved. Consider, for example, shifting the vector in Figure 9.1 so that point A coincides with the origin. Then point B is also shifted to another point; call this point B' with coordinates $(8 - 3, 7 - 2) = (5, 5)$. So we have a vector $\overrightarrow{OB'}$ with the same direction and distance, as in Figure 9.2.

Since $\overrightarrow{OB'}$ has the same direction and distance as $\mathbf{v} = \overrightarrow{AB}$, we can say that $\mathbf{v} = \overrightarrow{OB'} = \overrightarrow{AB}$. So it makes no real difference where the vector starts; only what its distance and direction are. Thus, we can henceforth consider every vector as starting (for convenience) at the origin.

Of course, if we consider every vector as starting at the origin, then the only other numerical values that define the vector are the coordinates of its endpoint. Hence, we can properly talk of the vector in Figure 9.2 as the vector $\mathbf{v} = \langle 5, 5 \rangle$, where we use angular brackets to distinguish this from a coordinate point. Now we are ready to formalize the notion of distance and direction of a vector.

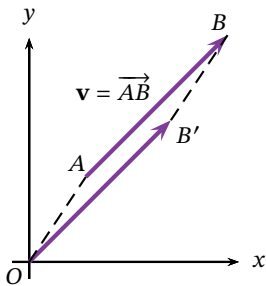


Figure 9.2 – Vectors are free.

The vector $\mathbf{v} = \langle v_1, v_2 \rangle$ has a distance, or **magnitude**, defined to be the **norm** of the vector:

$$\|\mathbf{v}\| = \|\langle v_1, v_2 \rangle\| = \sqrt{v_1^2 + v_2^2}.$$

The vector also has a direction, defined to be the angle θ , where

$$\theta = \arctan\left(\frac{v_2}{v_1}\right).$$

Given two vectors $\mathbf{u} = \langle u_1, u_2 \rangle$ and $\mathbf{v} = \langle v_1, v_2 \rangle$, we define **vector addition** as

$$\mathbf{u} + \mathbf{v} = \langle u_1 + v_1, u_2 + v_2 \rangle.$$

Clearly, vector subtraction is similarly defined.

It is “easy” to perform vector addition, but why define it this way? Because we want to combine two displacements to obtain one overall

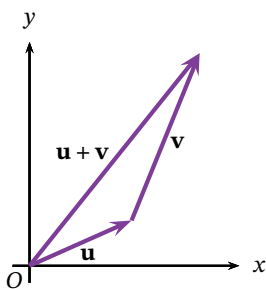


Figure 9.3 – Vector addition.

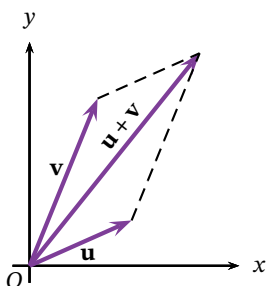


Figure 9.4 – The parallelogram property of vectors.

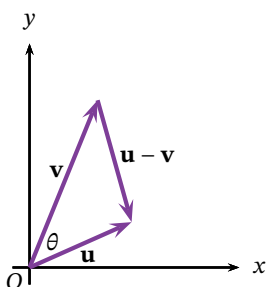


Figure 9.5 – The angle between two vectors.

displacement. Suppose we add together the vectors $\mathbf{u} = \langle 3, 1 \rangle$ and $\mathbf{v} = \langle 2, 5 \rangle$. Supposing vector \mathbf{v} begins at the end of \mathbf{u} , we may describe the overall displacement as the vector $\mathbf{u} + \mathbf{v} = \langle 5, 6 \rangle$. Figure 9.3 shows us the placement of the vectors.

The interesting property exhibited in the diagram of Figure 9.3 is made explicit in Figure 9.4. Since vectors are free, we may move the vector \mathbf{v} to the origin. Then the vectors \mathbf{u} and \mathbf{v} form two sides of a parallelogram, with the sum of the two vectors as the diagonal. (The difference of the two vectors is the other diagonal of the parallelogram.) We call this property the **parallelogram property** of vectors.

Given a real number, or **scalar**, c and a vector $\mathbf{v} = \langle v_1, v_2 \rangle$, the **scalar product** of c and \mathbf{v} is defined to be

$$c\mathbf{v} = c \langle v_1, v_2 \rangle = \langle cv_1, cv_2 \rangle.$$

Again, suppose $\mathbf{v} = \langle 2, 5 \rangle$. Then the vector $7\mathbf{v} = \langle 14, 35 \rangle$ represents a seven-fold increase in the magnitude of the vector, leaving the vector's direction unchanged. In other words, the factor of 7 *scales* the vector. Multiplication by a scalar t , where $0 < t < 1$ will shrink the vector. Multiplication by a negative scalar scales the vector as well, but in the *opposite* direction.

Exercise 9.1.1 Let $\mathbf{p} = \langle 3, 5 \rangle$ and $\mathbf{q} = \langle -1, 4 \rangle$. Sketch the following vectors on the same graph: \mathbf{q} , $2\mathbf{p}$, $-2\mathbf{q}$, $\mathbf{p} + \mathbf{q}$ and $\mathbf{p} - \mathbf{q}$.

Exercise 9.1.2 Suppose $\mathbf{v} = \langle v_1, v_2 \rangle$. Show that $\|\mathbf{v}\|^2 = v_1^2 + v_2^2$.

Suppose we have the vectors $\mathbf{u} = \langle u_1, u_2 \rangle$ and $\mathbf{v} = \langle v_1, v_2 \rangle$. We may find the angle between them by appealing to the Law of Cosines. The Law of Cosines applies to any triangle. The triangle we will be using is formed by the vectors \mathbf{u} , \mathbf{v} , and $\mathbf{u} - \mathbf{v}$ (see Figure 9.5), and we will be using it to find the angle θ . Since the Law of Cosines applies to the lengths of the sides of triangle, we use the magnitudes of the three vectors, and *not* the vectors themselves. To proceed, we will rewrite the Law of Cosines for $\cos(\theta)$, then substitute our vector magnitudes. Recalling that side c is opposite the angle, we have

$$c^2 = a^2 + b^2 - 2ab \cos(\theta)$$

$$c^2 - a^2 - b^2 = -2ab \cos(\theta)$$

$$\begin{aligned} \cos(\theta) &= \frac{a^2 + b^2 - c^2}{2ab} = \frac{\|\mathbf{u}\|^2 + \|\mathbf{v}\|^2 - \|\mathbf{u} - \mathbf{v}\|^2}{2\|\mathbf{u}\|\|\mathbf{v}\|} \\ &= \frac{u_1^2 + u_2^2 + v_1^2 + v_2^2 - (u_1 - v_1)^2 - (u_2 - v_2)^2}{2\|\mathbf{u}\|\|\mathbf{v}\|} \\ &= \frac{u_1^2 + u_2^2 + v_1^2 + v_2^2 - u_1^2 + 2u_1v_1 - v_1^2 - u_2^2 + 2u_2v_2 - v_2^2}{2\|\mathbf{u}\|\|\mathbf{v}\|} \\ &= \frac{2u_1v_1 + 2u_2v_2}{2\|\mathbf{u}\|\|\mathbf{v}\|} = \frac{u_1v_1 + u_2v_2}{\|\mathbf{u}\|\|\mathbf{v}\|}. \end{aligned}$$

Thus, the cosine of the angle between the two vectors is found by the above fraction. Clearly, to find the measure of the angle, one finds the inverse cosine of the fraction.

We may write the above more compactly. The denominator is simply

the product of the magnitudes of the two vectors. The quantity in the numerator is actually quite useful; we make the following definition.

The **dot product** of two vectors $\mathbf{u} = \langle u_1, u_2 \rangle$ and $\mathbf{v} = \langle v_1, v_2 \rangle$ is defined to be

$$\mathbf{u} \cdot \mathbf{v} = u_1 v_1 + u_2 v_2.$$

The cosine of the angle θ between the two vectors is given by

$$\cos(\theta) = \frac{\mathbf{u} \cdot \mathbf{v}}{\|\mathbf{u}\| \|\mathbf{v}\|}.$$

Two vectors are **orthogonal** if and only if the dot product of two vectors is zero.

The expression for the cosine of the angle between two vectors is surprisingly useful, as you will see!

Note that the dot product of two vectors is *not* another vector, but is a real-numbered scalar.

Exercise 9.1.3 *Orthogonal* simply means *perpendicular* in this context. So why are vectors perpendicular if the dot product is zero? What happens if the angle between them is 90° ?

Next, we turn to another definition.

A **unit vector** is a vector whose magnitude is 1. Any vector \mathbf{v} may be scaled to a unit vector by dividing by its magnitude; the unit vector in the same direction as $\mathbf{v} = \langle v_1, v_2 \rangle$ but with a magnitude of 1 is given by

$$\frac{\mathbf{v}}{\|\mathbf{v}\|} = \left\langle \frac{v_1}{\|\mathbf{v}\|}, \frac{v_2}{\|\mathbf{v}\|} \right\rangle.$$

The two **standard unit vectors** in the plane are

$$\mathbf{i} = \langle 1, 0 \rangle \quad \text{and} \quad \mathbf{j} = \langle 0, 1 \rangle.$$

Suppose we are given the vector $\mathbf{u} = \langle 3, 1 \rangle$. Its norm is $\|\langle 3, 1 \rangle\| = \sqrt{10}$. Then the unit vector that points in the same direction as \mathbf{u} would be

$$\frac{\mathbf{u}}{\|\mathbf{u}\|} = \left\langle \frac{3}{\sqrt{10}}, \frac{1}{\sqrt{10}} \right\rangle.$$

As verification, let us compute the magnitude of the new vector $\left\langle \frac{3}{\sqrt{10}}, \frac{1}{\sqrt{10}} \right\rangle$. We have

$$\left\| \left\langle \frac{3}{\sqrt{10}}, \frac{1}{\sqrt{10}} \right\rangle \right\| = \sqrt{\frac{9}{10} + \frac{1}{10}} = 1.$$

The standard unit vectors \mathbf{i} and \mathbf{j} are far more convenient than we may think at first glance. Consider the vector $\mathbf{v} = \langle 2, 5 \rangle$. We may write this vector as the *sum of standard unit vectors*. We have

$$2\mathbf{i} + 5\mathbf{j} = 2 \langle 1, 0 \rangle + 5 \langle 0, 1 \rangle = \langle 2, 0 \rangle + \langle 0, 5 \rangle = \langle 2, 5 \rangle = \mathbf{v}.$$

Hence, we may say $\mathbf{v} = 2\mathbf{i} + 5\mathbf{j}$. (See Figure 9.6.) This gives us an alternate way to notate vectors.

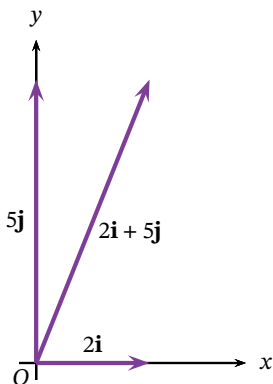


Figure 9.6 – Using standard unit vectors.

Exercise 9.1.4 Let $\mathbf{p} = 3\mathbf{i} + 5\mathbf{j}$ and $\mathbf{q} = -\mathbf{i} + 4\mathbf{j}$. Compute $\mathbf{p} \cdot \mathbf{q}$. Compute $2\mathbf{p} + 3\mathbf{q}$ and write your answer in terms of the standard unit vectors.

Problems for Section 9.1

- 1 The points $A(-3, 2)$, $B(1, 5)$, and $C(0, -4)$ in the plane determine a triangle.
- Determine the vectors \overrightarrow{AB} , \overrightarrow{BC} , and \overrightarrow{CA} .
 - Find the lengths of the sides of the triangle.
 - Graph the three vectors in part (a), making sure to put the arrowheads in the correct position. Why were you asked to find \overrightarrow{CA} and not \overrightarrow{AC} ?
- 2 Let $\mathbf{u} = \langle 2, -5 \rangle$, $\mathbf{v} = \langle 1, 3 \rangle$, and $\mathbf{w} = \langle 6, -2 \rangle$.
- Compute $\|\mathbf{u}\|$, $\|\mathbf{v}\|$, and $\|\mathbf{w}\|$.
 - Find the unit vector that has the same direction as \mathbf{w} .
 - Compute $\mathbf{u} \cdot \mathbf{v}$, $\mathbf{u} \cdot \mathbf{w}$, and $\mathbf{v} \cdot \mathbf{w}$.
 - Compute $(\mathbf{u} + \mathbf{v}) \cdot \mathbf{w}$ and compare this with the value of $(\mathbf{u} \cdot \mathbf{w}) + (\mathbf{v} \cdot \mathbf{w})$. What can you conclude about using addition and the dot product?
 - Compute $\|\mathbf{u}\|^2$ and compare this with the value of $\mathbf{u} \cdot \mathbf{u}$. What can you conclude about the dot product of a vector with itself?
 - Which pair of vectors given above are orthogonal? Find the angles between the other two pairs of non-orthogonal vectors.
- 3 In the text, we defined the direction of a vector as the tangent of the angle the vector makes with the x -axis. Although convenient, it is more precise (for later applications) to define a vector's direction with reference to *both* axes. Let θ be the angle $\mathbf{v} = \langle v_1, v_2 \rangle$ makes with the x -axis, and let φ be the angle \mathbf{v} makes with the y -axis. Then these so-called **direction angles** are given by
- $$\cos(\theta) = \frac{v_1}{\|\mathbf{v}\|} \quad \text{and} \quad \cos(\varphi) = \frac{v_2}{\|\mathbf{v}\|}.$$
- Compute the cosines of the direction angles for the three vectors given in Problem 2.
- 4 Consider the vector $\mathbf{0} = \langle 0, 0 \rangle$. This is called the **zero vector**.
- What is the graph of the zero vector?
 - Given any vector $\mathbf{v} = v_1\mathbf{i} + v_2\mathbf{j}$, what is $\mathbf{0} + \mathbf{v}$? What is $\mathbf{0} \cdot \mathbf{v}$?
 - Suppose we have two vectors \mathbf{p} and \mathbf{q} such that $\mathbf{p} + \mathbf{q} = \mathbf{0}$. What must be true of the directions and magnitudes of \mathbf{p} and \mathbf{q} ?
 - Is it true to say that the zero vector is orthogonal to *every* vector? Explain.
- 5 Will the magnitude of a vector ever be negative? Will it ever be zero? Justify your answers.

9.2 Modeling Motion with Vectors

As the sun eclipses the stars by its brilliance, so the man of knowledge will eclipse the fame of others in assemblies of the people if he proposes algebraic problems, and still more if he solves them.

— Brahmagupta

One advantage of using the standard unit vectors is they distinctly separate the horizontal and vertical components of a vector. This is useful when we use functions that model motion. Using vectors in a function that models motion allows us to see the horizontal and vertical motion as distinct, with the resultant motion being the sum of the horizontal and vertical motions. We make another definition.

A function $\mathbf{F}(t) = x(t)\mathbf{i} + y(t)\mathbf{j}$ is a **vector-valued function** where $x(t)$ and $y(t)$ are the usual real-valued functions of the variable t . The variable t is called the **parameter** of the function.

Chapter 10

LINEAR ALGEBRA

MANY STUDENTS WHO learn linear algebra consider it to be the mathematics of matrices. While matrices are a central component of linear algebra, matrices are not the fundamental idea, only the representation for that idea. The idea is a *linear transformation*. By a linear transformation, we mean any set of linear equations that describes how to transform one set of points to another. The theory and concept behind linear transformations gives the study its name: *linear algebra*.

Because we are interested in linear transformations in general (that is, we will investigate more than simply transformation in two-dimensions), we will begin with a treatment of three-dimensional vectors. This will result in a natural justification for the matrix representation of linear transformations. We will also develop some of the theory and applications of linear algebra.

10.1 The Geometry of Vectors

*Mathematical reality lies outside us,
... our function is to discover or observe it.*

— G. H. Hardy

Vectors are well-suited to describe geometrical concepts. We use this section to re-interpret some familiar ideas in two dimensions with vectors, and then to extend these ideas to three dimensions. We begin with lines in the plane.

Suppose we have a point $P_0(x_0, y_0)$ on a line and a vector $\mathbf{n} = \langle a, b \rangle$ that is perpendicular to the line. Then another point $P(x, y)$ is on the line precisely when the vectors \mathbf{n} and $\overrightarrow{P_0P}$ are perpendicular. (See Figure 10.1.) We have the following definition.

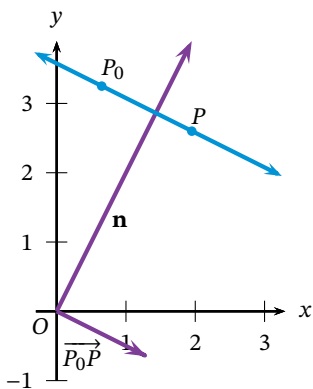


Figure 10.1 – The equation of the line $\overline{P_0P}$ is derived by taking the dot product of the vector from P_0 to P and a vector normal to the line

Suppose $P_0(x_0, y_0)$ is a point on a line and $\mathbf{n} = \langle a, b \rangle$ is a vector perpendicular to the line. Then the equation of the line is

$$\mathbf{n} \cdot \overrightarrow{P_0P} = 0$$

where P is any point on the line.

Notice that this dot product results in the familiar equation of a line. We have

$$\begin{aligned} \mathbf{n} \cdot \overrightarrow{P_0P} &= \langle a, b \rangle \cdot \langle x - x_0, y - y_0 \rangle \\ &= a(x - x_0) + b(y - y_0) = 0 \end{aligned}$$

which simplifies to $ax + by = d$, where $d = ax_0 + by_0$.

Example 10.1.1

Suppose the point $P_0(4, 5)$ is on a line and the vector $\mathbf{n} = \langle -3, 1 \rangle$ is perpendicular to the line. Then, for any point $P(x, y)$ on the line, we have

$$\langle -3, 1 \rangle \cdot \langle x - 4, y - 5 \rangle = -3(x - 4) + (y - 5) = -3x + 12 + y - 5 = 0,$$

or $3x - y = 7$. ♦

This new way to write a line is very useful, as we will now show. Given an equation of a line $ax + by = d$, we see that the vector $\langle a, b \rangle$ is perpendicular to the line. This is simply what we learned in algebra: the slopes of perpendicular lines are negative reciprocals. Note that the slope of $ax + by = d$ is $-\frac{b}{a}$ and the slope of the vector $\langle a, b \rangle$ is $\frac{b}{a}$.

The representation of a line by an equation $\mathbf{n} \cdot \overrightarrow{P_0P} = 0$ is not unique. Any point on the line can replace P_0 , and a nonzero multiple of \mathbf{n} can replace \mathbf{n} without changing the set of points P that satisfy the equation. We can use some of this freedom by requiring the vector \mathbf{n} to be a unit vector. The result is called the **normalized equation**, which then becomes

$$\mathbf{u} \cdot \overrightarrow{P_0P} = 0 \quad \text{where} \quad \mathbf{u} = \frac{\mathbf{n}}{\|\mathbf{n}\|}.$$

Example 10.1.2

Suppose the point $P_0(-1, -1)$ lies on a line and $\mathbf{n} = \langle 3, 4 \rangle$ is orthogonal to the line. Then the normalized equation of the line is

$$\frac{\langle 3, 4 \rangle}{\|\langle 3, 4 \rangle\|} \cdot \langle x + 1, y + 1 \rangle = \frac{3}{5}(x + 1) + \frac{4}{5}(y + 1) = 0,$$

or $\frac{3}{5}x + \frac{4}{5}y = -\frac{7}{5}$. ♦

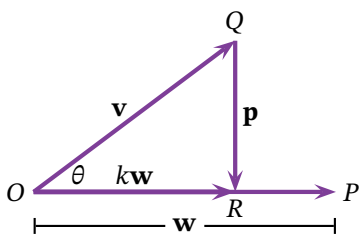


Figure 10.2 – The projection of \mathbf{v} on \mathbf{w} is $k\mathbf{w}$.

At this point, it is prudent to give another interpretation of the dot product of two vectors. Consider a vector $\mathbf{w} = \overrightarrow{OP}$ and a vector $\mathbf{v} = \overrightarrow{OQ}$. If we drop a perpendicular from Q to \overrightarrow{OP} , as in Figure 10.2, we create a vector $\mathbf{p} = \overrightarrow{QR}$ where R is the foot of the perpendicular. The vector \overrightarrow{OR} is called the **projection** of \mathbf{v} on \mathbf{w} . Note that the vector \overrightarrow{OR} , which lies on \mathbf{w} , is a scalar multiple of \mathbf{w} ; i.e., $\overrightarrow{OR} = k\mathbf{w}$.

Since $\cos \theta = \pm \|k\mathbf{w}\| / \|\mathbf{v}\|$, we have that

$$\mathbf{v} \cdot \mathbf{w} = \|\mathbf{v}\| \|\mathbf{w}\| \cos \theta = \pm \|\mathbf{v}\| \|\mathbf{w}\| \frac{\|k\mathbf{w}\|}{\|\mathbf{v}\|} = \pm \|\mathbf{w}\| \|k\mathbf{w}\|$$

where the sign is chosen as positive if θ is an acute angle or negative if θ is an obtuse angle. Hence, we have the following alternate definition of the dot product.

The dot product of \mathbf{v} and \mathbf{w} is

$$\mathbf{v} \cdot \mathbf{w} = \pm \|\mathbf{w}\| \|\text{projection of } \mathbf{v} \text{ on } \mathbf{w}\| = \pm \|\mathbf{w}\| \|k\mathbf{w}\| = \pm k \|\mathbf{w}\|^2$$

This interpretation of the dot product helps to prove the following.

LEMMA 10.A (Distance from a Point to a Line) Let $\mathbf{u} \cdot \overrightarrow{P_0P} = 0$ be the normalized equation of a line and let P_1 be any point in the plane. Then the distance from P_1 to the line is given by $D = |\mathbf{u} \cdot \overrightarrow{P_0P_1}|$.

Proof. From point P_1 we drop a perpendicular to $\overrightarrow{P_0P}$, meeting the vector at point Q . Then the distance we seek is given by $D = \|\overrightarrow{P_1Q}\|$. Recall that \mathbf{u} is a unit vector perpendicular to the line. Hence, the projection of $\overrightarrow{P_0P_1}$ on \mathbf{u} is exactly $\overrightarrow{P_1Q}$. Thus,

$$\mathbf{u} \cdot \overrightarrow{P_0P_1} = \pm \|\mathbf{u}\| \|\overrightarrow{P_1Q}\| = \pm \|\mathbf{u}\| D.$$

Therefore, since \mathbf{u} is a unit vector and distance is always positive, we have

$$|\mathbf{u} \cdot \overrightarrow{P_0P_1}| = D. \quad \blacksquare$$

Example 10.1.3

To find the distance from the point $(2, -1)$ to the line $3x + 4y = 9$, we first write the normalized equation of the line. The vector perpendicular to the line is $\mathbf{n} = \langle 3, 4 \rangle$, so the unit vector is $\langle \frac{3}{5}, \frac{4}{5} \rangle$. Then the normalized equation is $\frac{3}{5}x + \frac{4}{5}y - \frac{9}{5} = 0$. Now we substitute the point $(2, -1)$ into the normalized expression to get

$$\frac{3}{5}(2) + \frac{4}{5}(-1) - \frac{9}{5} = \frac{6}{5} - \frac{4}{5} - \frac{9}{5} = -\frac{7}{5},$$

which indicates that $|\frac{7}{5}| = \frac{7}{5}$ is the distance from the point to the line. \blacklozenge

Examining distances with vectors naturally leads to measuring other geometric properties. One such property is area.

THEOREM 10.B (Area of a Parallelogram) Let vectors \mathbf{v} and \mathbf{w} lie in the plane such that they form adjacent sides of a parallelogram. Then the area of the parallelogram is given by

$$A = \|\mathbf{v}\| \|\mathbf{w}\| \sin \theta$$

where θ is the angle between the two vectors.

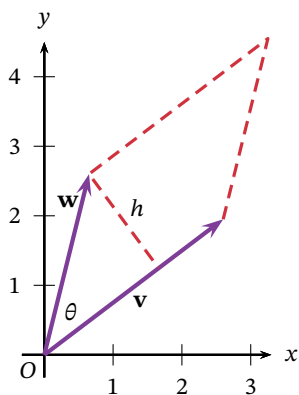


Figure 10.3 – The area of a parallelogram.

Proof. Suppose $\mathbf{w} = \langle w_1, w_2 \rangle$ and $\mathbf{v} = \langle v_1, v_2 \rangle$ both emanate from the origin with an angle of θ between them. Without loss of generality, assume \mathbf{v} is long enough so that when a perpendicular is dropped from the end of \mathbf{w} , the perpendicular intersects \mathbf{v} , as in Figure 10.3. Call the length of the perpendicular h . Then, since $\sin \theta = h/\|\mathbf{w}\|$, we have $h = \|\mathbf{w}\| \sin(\theta)$. Thus, because the base is $\|\mathbf{v}\|$, the area of the parallelogram is

$$A = bh = \|\mathbf{v}\| \|\mathbf{w}\| \sin(\theta),$$

and the theorem is proved. ■

Exercise 10.1.4 Find the distance from the point $(4, -7)$ to the line through the point $(-3, 1)$ and perpendicular to the vector $\langle 2, 3 \rangle$.

Exercise 10.1.5 Find the area of the parallelogram formed by $\langle 3, \sqrt{3} \rangle$ and $\langle 0, 9 \rangle$.

Problems for Section 10.1

- 1 Find the equation of the line through $(3, -1)$ and perpendicular to $\langle -1, 2 \rangle$.
- 2 Find the unit vector that is perpendicular to the line $2x - y + 4 = 0$.
- 3 Find the distance from the point $(-1, 7)$ to the line $4x - y - 11 = 0$.
- 4 Find the area of the parallelogram formed by the vectors $\langle -2, 4 \rangle$ and $\langle 3, 3 \rangle$. What is the area of the triangle formed by these vectors?
- 5 Given that $\mathbf{v} = \langle x, y \rangle$, explain why $\|\mathbf{v}\| = r$ is the equation of a circle of radius r . What is the circle's center?
- 6 Given three points $R(2, 5)$, $S(-4, 2)$, and $T(1, -1)$ in the plane, find:
 - (a) the equation of the line through S if \overrightarrow{TR} is the normal vector;
 - (b) the distance from R to the line found in part (a);
 - (c) the unit vector perpendicular to the line through R and S ;
 - (d) the area of the parallelogram formed by \overrightarrow{SR} and \overrightarrow{ST} ;
 - (e) the equations of the line through R that is perpendicular to the line through S and T ;
 - (f) the angles of the triangle formed by vertices R , S , and T .

10.2 Vectors in Three Dimensions

There exists, if I am not mistaken, an entire world which is the totality of mathematical truths, to which we have access only with our mind, just as a world of physical reality exists, the one like the other independent of ourselves, both of divine creation.

— Charles Hermite

Vectors in three-dimensional space are analogous to vectors in the plane but with a third component to represent the third dimension. A vector in space is $\mathbf{v} = \langle v_1, v_2, v_3 \rangle$ where the third component can be considered “height” (if the first two are “length” and “width”). This leads to a third standard unit vector, and modifications to the familiar \mathbf{i} and \mathbf{j} :

$$\mathbf{i} = \langle 1, 0, 0 \rangle, \quad \mathbf{j} = \langle 0, 1, 0 \rangle, \quad \mathbf{k} = \langle 0, 0, 1 \rangle.$$

Problems for Section 10.8

Find the eigenvalues and eigenvectors of the following matrices.

$$1 \begin{bmatrix} -3 & 2 \\ 8 & 3 \end{bmatrix}$$

$$4 \begin{bmatrix} \pi & 2 \\ 0 & 2\pi \end{bmatrix}$$

$$2 \begin{bmatrix} 0 & 1 & -2 \\ 2 & 1 & 0 \\ 4 & -2 & 5 \end{bmatrix}$$

$$5 \begin{bmatrix} 5 & -2 & 8 \\ -4 & 0 & -5 \\ -4 & 2 & -7 \end{bmatrix}$$

$$3 \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$$

$$6 \begin{bmatrix} 0 & 1 & -2 \\ -6 & 5 & -4 \\ 0 & 0 & 3 \end{bmatrix}$$

7 Prove that the product of the eigenvalues of a square matrix is equal to its determinant.

8 Prove that every square matrix is similar to itself.

9 Prove that if A is similar to B and B is similar to C , then A is similar to C .

10.9 Orthogonal Matrices

In many cases, mathematics is an escape from reality. The mathematician finds his own monastic niche and happiness in pursuits that are disconnected from external affairs. Some practice it as if using a drug. Chess sometimes plays a similar role. In their unhappiness over the events of this world, some immerse themselves in a kind of self-sufficiency in mathematics. (Some have engaged in it for this reason alone.)

— Stanislaw Ulam

We begin this section with a new matrix operation in which we interchange the rows and columns of a matrix.

The **transpose** of a matrix A , denoted A^T , is obtained by swapping rows of A with the columns of A . That is, if $A = [a_{ij}]$, then $A^T = [a_{ji}]$.

Example 10.9.1

We have that if $A = \begin{bmatrix} 1 & 0 \\ -4 & 5 \\ 4 & 3 \end{bmatrix}$, then $A^T = \begin{bmatrix} 1 & -4 & 4 \\ 0 & 5 & 3 \end{bmatrix}$.

Clearly, if the matrix is square, so is its transpose. \blacklozenge

The transpose obeys several rules, as the theorem below indicates.

THEOREM 10.P (Properties of the Transpose) Suppose A and B are matrices and k is a constant. Then

i. $(A^T)^T = A$.

ii. $(kA)^T = kA^T$.

iii. $(A + B)^T = A^T + B^T$.

iv. $(AB)^T = B^T A^T$.

v. If A is nonsingular, then $(A^T)^{-1} = (A^{-1})^T$.

Proof. Proofs of statements (i) and (ii) are left as Problem 15.

To prove statement (iii), we set $C = A + B$ and $D = A^T + B^T$. Then $c_{ij} = a_{ij} + b_{ij}$ and $d_{ij} = a_{ji} + b_{ji}$ for all i and j . Then $C^T = [c_{ji}] = [a_{ji} + b_{ji}] = [d_{ij}] = D$, which proves the statement.

To prove statement (iv), we let A be an $m \times p$ matrix and B a $p \times n$ matrix so that the multiplication of A and B is defined. Now set $E = AB = [e_{ij}]$ and $F = B^T A^T = [f_{ij}]$, where

$$f_{ij} = b_{1i}a_{j1} + \cdots + b_{pi}a_{jp} = a_{j1}b_{1i} + \cdots + a_{jp}b_{pi} = e_{ji}$$

for $i = 1, 2, \dots, n$ and $j = 1, 2, \dots, m$. Thus, $F = E^T$, and the statement follows.

Finally, to prove statement (v), write $AA^{-1} = I$. Then by statement (iv), $(AA^{-1})^T = (A^{-1})^T A^T = I^T = I$, so that A^T has an inverse $(A^T)^{-1}$, which, since inverses are unique, must be equal to $(A^{-1})^T$. ■

With the transpose, we may succinctly write a column vector as the transpose of the corresponding row vector:

$$\begin{bmatrix} x_1 & x_2 & x_3 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} x_1 & x_2 & x_3 \end{bmatrix} \begin{bmatrix} y_1 & y_2 & y_3 \end{bmatrix}^T$$

Note that this gives us a way to express the dot product of \mathbf{x} and \mathbf{y} as the product of the “matrix” \mathbf{x} and the “matrix” \mathbf{y}^T . Hence, $\mathbf{x} \cdot \mathbf{y} = \mathbf{xy}^T$.

A square matrix M for which $M = M^T$ is called a **symmetric matrix**.

Example 10.9.2

Both the matrices

$$\begin{bmatrix} 1 & 3 \\ 3 & 4 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} 2 & 1 & 0 \\ 1 & -1 & -3 \\ 0 & -3 & 5 \end{bmatrix}$$

are symmetric. ◆

Concerning symmetric matrices, we have the following interesting result linking symmetry with eigenvalues.

THEOREM 10.Q *Suppose A is an $n \times n$ symmetric matrix. Then any two eigenvectors that are associated with distinct eigenvalues are orthogonal.*

Proof. For symmetric matrix A , suppose λ_1 has eigenvector \mathbf{x}_1 and λ_2 has eigenvector \mathbf{x}_2 . Then

$$\begin{aligned} \lambda_2(\mathbf{x}_1 \cdot \mathbf{x}_2) &= (\lambda_2 \mathbf{x}_1) \mathbf{x}_2^T \\ &= \mathbf{x}_1 (\lambda_2 \mathbf{x}_2^T) \\ &= \mathbf{x}_1 A \mathbf{x}_2^T, \end{aligned}$$

and since A is symmetric, $A = A^T$, so that we have

$$\begin{aligned}
 &= \mathbf{x}_1 A^T \mathbf{x}_2^T \\
 &= (\mathbf{x}_1 A^T) \mathbf{x}_2^T \\
 &= (A \mathbf{x}_1^T)^T \mathbf{x}_2^T \\
 &= (\lambda_1 \mathbf{x}_1^T)^T \mathbf{x}_2^T \\
 &= \lambda_1 (\mathbf{x}_1^T)^T \mathbf{x}_2^T \\
 &= \lambda_1 (\mathbf{x}_1 \mathbf{x}_2^T) \\
 &= \lambda_1 (\mathbf{x}_1 \cdot \mathbf{x}_2).
 \end{aligned}$$

Thus, $\lambda_1 (\mathbf{x}_1 \cdot \mathbf{x}_2) = \lambda_2 (\mathbf{x}_1 \cdot \mathbf{x}_2)$, or $(\lambda_1 - \lambda_2) (\mathbf{x}_1 \cdot \mathbf{x}_2) = 0$. Since $\lambda_1 \neq \lambda_2$, we must have that $\mathbf{x}_1 \cdot \mathbf{x}_2 = 0$. Therefore, since the dot product is zero, the eigenvectors are orthogonal. ■

We introduce some notation. The symbol $\text{diag}(a_{11}, a_{22}, \dots, a_{nn})$ denotes the matrix whose every entry is zero except for the entries along the main diagonal; such a matrix is called a **diagonal matrix**. For example,

$$\text{diag}(2, 7, -1) = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 7 & 0 \\ 0 & 0 & -1 \end{bmatrix}.$$

Note that every diagonal matrix $\text{diag}(a_{11}, a_{22}, \dots, a_{nn})$ is also symmetric.

The next theorem relates the concept of similarity to the eigenvalues of a matrix.

COROLLARY 10.R *Let the $n \times n$ matrix A have n distinct eigenvalues λ_i for $i = 1, 2, \dots, n$. Then A is similar to $\text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n)$.*

Proof. Let A have n distinct eigenvalues λ_i . Then A also has n distinct eigenvectors \mathbf{x}_i so that $A\mathbf{x}_i = \lambda_i\mathbf{x}_i$ for every i . Let C be the matrix whose column vectors are the eigenvectors of A ; in other words,

$$C = [\mathbf{x}_1 \quad \mathbf{x}_2 \quad \cdots \quad \mathbf{x}_n].$$

Then we have

$$\begin{aligned}
 AC &= A [\mathbf{x}_1 \quad \mathbf{x}_2 \quad \cdots \quad \mathbf{x}_n] \\
 &= A \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1n} \\ x_{21} & x_{22} & \cdots & x_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{nn} \end{bmatrix} \\
 &= \begin{bmatrix} \lambda_1 x_{11} & \lambda_2 x_{12} & \cdots & \lambda_n x_{1n} \\ \lambda_1 x_{21} & \lambda_2 x_{22} & \cdots & \lambda_n x_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \lambda_1 x_{n1} & \lambda_2 x_{n2} & \cdots & \lambda_n x_{nn} \end{bmatrix} \\
 &= \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1n} \\ x_{21} & x_{22} & \cdots & x_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{nn} \end{bmatrix} \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix} \\
 &= C \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n).
 \end{aligned}$$

Since C is the matrix of eigenvectors arising from n distinct eigenvalues, by Theorem 10.Q, we know that C consists of mutually orthogonal columns; hence, $\det C$ cannot be zero, which means that C is nonsingular. Therefore, we may multiply both sides of $AC = C\text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n)$ by C^{-1} to get $\text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n) = C^{-1}AC$. We conclude that $A \sim \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n)$. ■

Since a matrix is always similar to its diagonal matrix of its eigenvalues, we have a computationally easier way to study the effect of large matrices. For example, let $D = \text{diag}(\lambda_1, \dots, \lambda_n)$ and suppose we wish to calculate A^2 . Since we have $A = CDC^{-1}$, then

$$A^2 = (CDC^{-1})(CDC^{-1}) = CD(C^{-1}C)DC^{-1} = CD^2C^{-1}.$$

Similarly,

$$A^3 = (CD^2C^{-1})(CDC^{-1}) = CD^2(CC^{-1})DC^{-1} = CD^3C^{-1}.$$

In this manner, we see that

$$A^k = CD^kC^{-1}.$$

Powers of diagonal matrices are very easy to compute:

$$D^k = [\text{diag}(\lambda_1, \dots, \lambda_n)]^k = \text{diag}(\lambda_1^k, \dots, \lambda_n^k).$$

Example 10.9.3

From Example 10.8.1, we know the eigenvalues and eigenvectors of

$$A = \begin{bmatrix} 1 & 2 \\ 3 & 2 \end{bmatrix}.$$

Thus, instead of performing three complicated multiplications, we compute

$$\begin{aligned} A^4 &= \begin{bmatrix} -1 & 2 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} -1 & 0 \\ 0 & 4 \end{bmatrix}^4 \begin{bmatrix} -1 & 2 \\ 1 & 3 \end{bmatrix}^{-1} \\ &= \begin{bmatrix} -1 & 2 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} (-1)^4 & 0 \\ 0 & 4^4 \end{bmatrix} \begin{bmatrix} \frac{3}{5} & -\frac{2}{5} \\ -\frac{1}{5} & -\frac{1}{5} \end{bmatrix} \\ &= \begin{bmatrix} 103 & 102 \\ 153 & 154 \end{bmatrix}. \end{aligned}$$

Computing A to any other large power is much simpler than multiplying A to itself over and over again. ♦

Questions of orthogonality bring us to the following definition.

A square matrix M is called an **orthogonal matrix** if $MM^T = I$. In other words, M is orthogonal if and only if $M^T = M^{-1}$.

There are two consequences of this definition. The first is that every orthogonal matrix must also be nonsingular. The second is that, due to the commutativity of a matrix and its inverse, $MM^T = M^T M = I$.

Let us give a reason for this definition. We may call a matrix M such that $MM^T = I$ orthogonal because each row vector is orthogonal to every other row vector, and each column vector is orthogonal to every other column vector. In particular, we have the following theorem.

THEOREM 10.5 A square matrix M is orthogonal if and only if each column and row vector is a unit vector and each column and row vector is orthogonal to every other column and row vector.

Proof. Let us denote the columns of M by \mathbf{m}_i for $i = 1, 2, \dots, n$. Then

$$M = [\mathbf{m}_1 \quad \mathbf{m}_2 \quad \cdots \quad \mathbf{m}_n] \quad \text{and} \quad M^T = \begin{bmatrix} \mathbf{m}_1^T \\ \mathbf{m}_2^T \\ \vdots \\ \mathbf{m}_n^T \end{bmatrix}$$

so that

$$M^T M = \begin{bmatrix} \mathbf{m}_1^T \mathbf{m}_1 & \mathbf{m}_1^T \mathbf{m}_2 & \cdots & \mathbf{m}_1^T \mathbf{m}_n \\ \mathbf{m}_2^T \mathbf{m}_1 & \mathbf{m}_2^T \mathbf{m}_2 & \cdots & \mathbf{m}_2^T \mathbf{m}_n \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{m}_n^T \mathbf{m}_1 & \mathbf{m}_n^T \mathbf{m}_2 & \cdots & \mathbf{m}_n^T \mathbf{m}_n \end{bmatrix} = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix} = I.$$

Thus, since $\mathbf{m}_i^T \mathbf{m}_i = 1$ for each i , each column vector is a unit vector; since $\mathbf{m}_i^T \mathbf{m}_j = 0$ when $i \neq j$, the column vectors are orthogonal.

The proof for the row vectors is similar and is omitted. ■

Exercise 10.9.4 Is the matrix $B = \frac{1}{7} \begin{bmatrix} 2 & 3 & 6 \\ 3 & -6 & 2 \\ -6 & -2 & 3 \end{bmatrix}$ orthogonal? Explain.

Exercise 10.9.5 What is the determinant of B ?

Example 10.9.6

The matrix

$$R(\theta) = \begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix}$$

is orthogonal for every θ . Note that

$$\begin{aligned} R^T R &= \begin{bmatrix} \cos(\theta) & \sin(\theta) \\ -\sin(\theta) & \cos(\theta) \end{bmatrix} \begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix} \\ &= \begin{bmatrix} \cos^2(\theta) + \sin^2(\theta) & 0 \\ 0 & \cos^2(\theta) + \sin^2(\theta) \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I. \end{aligned}$$

Also the magnitude of each column is 1, and the columns are orthogonal for every θ . ◆

The matrix $R(\theta)$ in the above example is called a **rotational matrix**. This matrix represents a counterclockwise rotation of θ about the origin.

Example 10.9.7

The point $(2, 0)$ is rotated to the point $(0, 2)$ when the matrix $R(\frac{\pi}{2})$ is applied. To wit:

$$\begin{bmatrix} \cos(\frac{\pi}{2}) & -\sin(\frac{\pi}{2}) \\ \sin(\frac{\pi}{2}) & \cos(\frac{\pi}{2}) \end{bmatrix} \begin{bmatrix} 2 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 2 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 2 \end{bmatrix}.$$

In fact we may rotate any point (x, y) .

$$\begin{bmatrix} \cos(\frac{\pi}{2}) & -\sin(\frac{\pi}{2}) \\ \sin(\frac{\pi}{2}) & \cos(\frac{\pi}{2}) \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} -y \\ x \end{bmatrix}.$$

Hence, the rotation sends the point (x, y) to the point $(-y, x)$.

If a locus of points are described by some equation, say $y = f(x)$, then this allows us to rotate all such points through an angle of $\frac{\pi}{2}$. If $y = x^2$, then the point (x, x^2) becomes $(-x^2, x)$. In other words, the upward positive parabola $y = x^2$ rotates to become $x = -y^2$, which describes a parabola that opens to the left. ♦

Problems for Section 10.9

1 What is $\det R(\theta)$?

Find the transpose of each matrix.

2 $\begin{bmatrix} 1 & 2 & 3 \\ 7 & 0 & 5 \end{bmatrix}$

4 $\begin{bmatrix} 2 \\ 1 \\ -6 \end{bmatrix}$

3 $\begin{bmatrix} 2 & 8 & -4 & 1 \end{bmatrix}$

Find the values of a and b so that each matrix is symmetric.

5 $\begin{bmatrix} -3 & 3a - 1 \\ 2a & 7 \end{bmatrix}$

6 $\begin{bmatrix} 1 & a & 8 \\ b - a & 9 & 4 + a \\ 8 & b & 3 \end{bmatrix}$

Are the following matrices orthogonal? Justify your answers.

7 $\begin{bmatrix} \frac{\sqrt{3}}{2} & -\frac{1}{2} \\ \frac{1}{2} & \frac{\sqrt{3}}{2} \end{bmatrix}$

10 $\begin{bmatrix} \cos(\theta) & -\sin(\theta) & 0 \\ \sin(\theta) & \cos(\theta) & 0 \\ 0 & 0 & 1 \end{bmatrix}$

8 $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$

11 $\begin{bmatrix} 2 & 1 & 2 \\ 1 & 2 & -2 \\ -2 & 2 & 1 \end{bmatrix}$

9 $\begin{bmatrix} \frac{1}{\sqrt{3}} & \frac{2}{\sqrt{6}} & 0 \\ \frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{6}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{2}} \end{bmatrix}$

12 Suppose $A = \begin{bmatrix} 3 & 1 \\ 4 & 3 \end{bmatrix}$.

(a) Find a matrix B and a nonsingular matrix C so that $A = CBC^{-1}$.

(b) Compute A^5 .

Each matrix is a rotational matrix. Determine the angle of rotation, then determine the new, rotated equations of the lines $x = 0$ and $y = 0$.

13 $\begin{bmatrix} \frac{1}{2} & -\frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & \frac{1}{2} \end{bmatrix}$

14 $\begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}$

15 Prove statements (i) and (ii) of Theorem 10.P.

16 Suppose A is an orthogonal matrix.

(a) Prove that $\det A = \pm 1$.

(b) Prove that A^T and A^{-1} are also orthogonal.

10.10 Preparation and Extension

One of the penalties for refusing to participate in politics is that you end up being governed by your inferiors.

— Plato

Preparation Problems for Chapter 10

Find the velocity and acceleration vectors, given the position vector $\mathbf{p}(t)$.

1 $\mathbf{p}(t) = \langle 6te^{-2t}, 8e^{-2t}, -16t^2 \rangle$

2 $\mathbf{p}(t) = \langle 6e^{-3t}, \sin(3t), t^3 - 6t \rangle$

Find the position vector, given either the velocity vector $\mathbf{v}(t)$ or the acceleration vector $\mathbf{a}(t)$.

3 $\mathbf{v}(t) = \langle 10, 3e^t, -32t + 8 \rangle$, $\mathbf{p}(0) = \langle 0, 6, 20 \rangle$

4 $\mathbf{v}(t) = \langle t + 1, t^2, e^{-t/3} \rangle$, $\mathbf{p}(0) = \langle 4, 0, -3 \rangle$

Pricing information from future.aae.wisc.edu/tab/prices.html, retrieved January 2011.

(Note that at the critical point, the price of butter becomes 279 cents per pound and cheese becomes 470 cents per pound. According to the webpage *Understanding Dairy Markets*, by Brian W. Gould of the University of Wisconsin, in 2009 the average national cost of butter was \$2.80 per pound and that of cheddar cheese was \$4.70 per pound. ♦

Problems for Section 11.10

For each of the following surfaces, find critical points and test those points for extrema.

1 $z = 2x^3 - xy^2 + 5x^2 - y^2$

2 $z = x^3 - 12xy + 8y^3$

3 $z = 2x^2 - xy - 3y^2 - 3x + 7y$

4 $z = x \cosh(y)$

5 $z = x^2 - 2x(\sin(y) + \cos(y)) + 1$

6 $z = x^2 + y^2 + 2x^{-1}y^{-2}$

7 $z = x^3 + y^2 - 3x^2 - 3y - 9x$

8 (Calculator) Products A and B are jointly produced. The price of product A is $P_A(x) = 300 - x$ and that of product B is $P_B(y) = 150 - 4y$. The joint cost is

$$C(x, y) = 2x^2 + \frac{1}{2}y^2 + xy + 30.$$

What is the maximum profit of this production model, and what prices should be established for A and B to achieve the maximum profit?

11.11 The Lagrange Multiplier

The reader will find no figures in this work. The methods which I set forth do not require either constructions or geometrical or mechanical reasonings: but only algebraic operations, subject to a regular and uniform rule of procedure.

— Joseph-Louis Lagrange

Often we are required to maximize or minimize a function $f(x, y, z)$ subject to a condition $g(x, y, z)$. For instance we may want to maximize profit, but under the constraint of the availability of supplies. So in this section we develop methods to find extrema of functions subject to additional **constraints**, or **side conditions**.

The viability of such methods is a direct result of the existence of the following type of extrema. Given a function and a constraint, the effect of the constraint on the function is to restrict the function's domain to a closed subset E of the domain D .

Let $z = f(x, y)$ be a continuous function defined on a closed region E formed from a domain D of the xy -plane. If P_0 is a point of E such that $f(P_0) \geq f(x, y)$ for all points (x, y) in E , then P_0 is an **absolute maximum**. If P_0 is a point of E such that $f(P_0) \leq f(x, y)$ for all points (x, y) in E , then P_0 is an **absolute minimum**.

This, combined with the following theorem, is exploited to find extrema of functions with constraints.

THEOREM 11.M Suppose D is a bounded domain of the xy -plane. If $z = f(x, y)$ is defined and continuous in the closed region E formed of

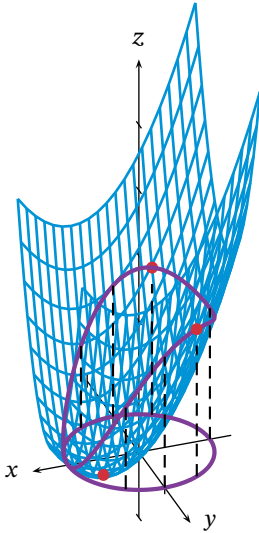


Figure 11.10 – Figure for Example 11.11.1. The points of absolute extrema of the surface under the constraint are marked.

D and its boundary, then f has an absolute maximum and an absolute minimum in E .

This theorem is the three-dimensional analogue of the Extreme Value Theorem. Since every continuous function has absolute extrema on a closed region, we are guaranteed to find extrema. However, in the course of our search we may find (as in the last section) saddle points or other local extrema as well. One method to find absolute extrema, which is illustrated in the example below, relies on being able to write the constraint in terms of one variable. This method for finding absolute extrema is known as the *direct method*.

Example 11.11.1

Let us find the absolute maximum and absolute minimum of $z = x^2 + 2y^2 - x$ on the unit disk $x^2 + y^2 \leq 1$.

Since $\frac{\partial z}{\partial x} = 2x - 1$ and $\frac{\partial z}{\partial y} = 4y$, we have a critical point at $(\frac{1}{2}, 0)$. Note that this critical point is within the unit disk, so it is a candidate for absolute extrema.

Now we investigate the function values on the boundary of the unit disk. Figure 11.10 shows the surface z and the projection of the constraint onto the surface. It is along this projected curve that we find extrema. To do this, we substitute the boundary equation $x^2 + y^2 = 1$ into z to get $z = 2 - x - x^2 = (2 + x)(1 - x)$, which is valid for $-1 \leq x \leq 1$. This function (which is now a single-variable function) has the critical points $x = -1$, $x = -\frac{1}{2}$, and $x = 1$ on the boundary.

Finally, to determine absolute extrema, we evaluate z at these points and see which is the smallest and which is the largest. At the point $(\frac{1}{2}, 0)$, we compute $z = -\frac{1}{4}$. On the boundary, we may compute values of z using the equation $z = (2 + x)(1 - x)$. Hence, $x = -1$ gives $z = 2$; $x = -\frac{1}{2}$ gives $z = \frac{9}{4}$; and $x = 1$ gives $z = 0$. Thus we conclude that the absolute maximum is $\frac{9}{4}$ which occurs at the two points $(-\frac{1}{2}, \pm\frac{\sqrt{3}}{2})$ and the absolute minimum is $-\frac{1}{4}$ which occurs at $(\frac{1}{2}, 0)$. (Since the minimum occurs inside the unit disk, it is also a relative minimum.) ♦

This direct method is very useful in the simple cases, but is implausible for more complicated functions and constraints: it may be impossible to solve the constraint equation for one of the variables. So although the direct method is fine, we need a more general method. This method is called the **Lagrange multiplier** method and its basis is the following theorem.

THEOREM 11.N (Lagrange's Theorem) *Let functions $f(\mathbf{x})$ and $g(\mathbf{x})$ have continuous partial derivatives in a neighborhood of \mathbf{x}_0 . If \mathbf{x}_0 gives an extreme of f subject to the constraint $g(\mathbf{x}) = 0$ where $\nabla g(\mathbf{x}_0) \neq \mathbf{0}$, then $\nabla f(\mathbf{x}_0) = \lambda \nabla g(\mathbf{x}_0)$ for some constant λ .*

Proof. Suppose the point \mathbf{x}_0 gives an extreme of $f(\mathbf{x})$ where $g(\mathbf{x}) = 0$ is given as a constraint. Let $C(t)$ be any differentiable curve lying on the surface f that passes through the point $f(\mathbf{x}_0)$. Then there is a value t_0 such that $C(t_0) = \mathbf{x}_0$. Then the function $f(C(t))$ has an extreme for $t = t_0$. The derivative of $f(C(t))$ at the point t_0 is therefore zero. In other words, by the Chain Rule,

$$\left. \frac{d}{dt} f(C(t)) \right|_{t_0} = \frac{\partial f}{\partial x} \frac{dx}{dt} + \frac{\partial f}{\partial y} \frac{dy}{dt} + \frac{\partial f}{\partial z} \frac{dz}{dt} \Big|_{t_0} = \nabla f(\mathbf{x}_0) \cdot C'(t_0) = 0.$$

Joseph-Louis Lagrange, 18th century Italian-born French mathematician, advanced all fields of analysis, number theory, and celestial mechanics. He first wrote about the multiplier 1764, and fully realized the method in his groundbreaking book *Mechanical Analysis* of 1788.

This implies that $\nabla f(\mathbf{x}_0)$ is normal to the curve C at $f(\mathbf{x}_0)$. Since the curve C was an arbitrary curve passing through $f(\mathbf{x}_0)$, we have that $\nabla f(\mathbf{x}_0)$ is normal to every curve passing through $f(\mathbf{x}_0)$.

Since \mathbf{x}_0 maximizes f subject to $g = 0$, it must be that g passes through the point $f(\mathbf{x}_0)$ also. By definition, $\nabla g(\mathbf{x}_0)$ is normal to the surface $g = 0$; but so is $\nabla f(\mathbf{x}_0)$. Thus, $\nabla g(\mathbf{x}_0)$ is parallel to $\nabla f(\mathbf{x}_0)$. Therefore there is a constant λ such that $\nabla f(\mathbf{x}_0) = \lambda \nabla g(\mathbf{x}_0)$. ■

So we are using λ for the Lagrange multiplier and we used λ for eigenvalues... is there some connection?

The constant λ in Theorem 11.N is called the **Lagrange multiplier**. This method consists of solving the system of equations

$$\begin{cases} \nabla f(\mathbf{x}_0) = \lambda \nabla g(\mathbf{x}_0) \\ g(\mathbf{x}_0) = 0 \end{cases}$$

where λ is treated as an extra variable to ensure the number of equations and variables is identical. Like with other methods of finding extrema, not all solutions to the system will give absolute extrema. Just as with critical points, we must determine which of all the solutions will give extrema. Fortunately, this is simply done by evaluating f at our possible candidates. The following example illustrates the method.

Example 11.11.2

To find the extreme values of $z = x + 2y$ on the circle $x^2 + y^2 = 1$, we compute the gradients

$$\nabla f = \langle 1, 2 \rangle \quad \text{and} \quad \nabla g = \langle 2x, 2y \rangle$$

and then form the system

$$\begin{aligned} 1 &= 2\lambda x \\ 2 &= 2\lambda y \\ x^2 + y^2 &= 1. \end{aligned}$$

We begin with the first equation to get that $\lambda = \frac{1}{2x}$. Substituting this into the second, we obtain $2x = y$. Then we use this relation in the last equation to write $x^2 + (2x)^2 = 1$. Solving this equation gives $x = \pm 1/\sqrt{5}$; then $y = \pm 2/\sqrt{5}$. The maximum is then $z = 1/\sqrt{5} + 4/\sqrt{5} = \sqrt{5}$ and the minimum is $z = -1/\sqrt{5} - 4/\sqrt{5} = -\sqrt{5}$. ♦

Exercise 11.11.3 Find the extrema of $z = x + y$ given the constraint $x^2 + y^2 \leq 1$.

Example 11.11.4

To find the point on the surface $x^2 + xy - z^2 + 4 = 0$ that is closest to the origin requires us to minimize the distance function $D = \sqrt{x^2 + y^2 + z^2}$. However, to minimize this function is the same as minimizing $D^2 = x^2 + y^2 + z^2$, which is simpler to do!

For this problem, our function to minimize is distance, which makes the surface the constraint. So we have the system

$$\begin{aligned} 2x &= \lambda(2x + y) \\ 2y &= \lambda x \\ 2z &= -2\lambda z \\ x^2 + xy - z^2 + 4 &= 0 \end{aligned}$$

The third equation simplifies to $\lambda = -1$, and we use this value in the first equation to get that $y = -4x$. Using this relation and the value of λ in the second equation, we find that $x = 0$. Thus, $y = 0$ and $z = \pm 2$. Hence, there are two points on the surface closest to the origin: $(0, 0, 2)$ and $(0, 0, -2)$. \blacklozenge

Some Applications. The method of Lagrange multipliers allows us to find extrema of a wide variety of problems. The next few examples illustrate business and physics applications.

Example 11.11.5

A rocket is launched with a constant acceleration of a feet per second per second. The rocket's height after t seconds is given by $f(t, a) = \frac{1}{2}(a - 32)t^2$ feet. Fuel usage for t seconds is proportional to a^2t and the limited fuel capacity of the rocket satisfies $a^2t = 8000$. We are to find the value of a that maximizes the height that the rocket attains when the fuel runs out.

In order to find this value, we maximize $f(t, a)$ given the constraint $a^2t - 8000 = 0$. Hence, we have the system

$$\begin{aligned}(a - 32)t &= \lambda a^2 \\ \frac{1}{2}t^2 &= 2\lambda at \\ a^2t - 8000 &= 0.\end{aligned}$$

The first equation gives us $\lambda = (a - 32)t/a^2$, which, when substituted into the second equation, gives us

$$\begin{aligned}\frac{1}{2}t^2 &= \frac{2a(a - 32)t^2}{a^2} \\ a^2t^2 &= 4a(a - 32)t^2 \\ 3a^2t^2 &= 128at^2 \\ a &= \frac{128}{3}.\end{aligned}$$

(Note that the solutions $a = 0$ and $t = 0$ are rejected because they do not satisfy the constraint.) Hence, with an acceleration of $\frac{128}{3}$, we have

$$t = \frac{8000}{a^2} = \frac{8000}{(128/3)^2} = \frac{1125}{256} \approx 4.395 \text{ seconds.}$$

The height the rocket attains when the fuel runs out is $f(4.395, \frac{128}{3}) = \frac{1}{2}(\frac{128}{3} - 32)(\frac{1125}{256})^2 \approx 102.997$ feet. \blacklozenge

Next we demonstrate an application to business and economics. In 1928 Charles Cobb and Paul Douglas published a study in which they modeled the growth of the American economy during the period 1899–1922. They considered a simplified view of the economy in which production output is determined by the amount of labor involved and the amount of capital invested. Their model is the function

$$P(K, L) = aK^mL^n$$

where K is the capitol investment (value of all facilities and equipment in a year), L is the labor investment (person-hours per year), P is the total production (value of goods produced per year), and a , m , and n are constants reflecting the types of capitol, labor, and available technology. This is known as the **Cobb-Douglas production function**.

This production function is now known as being too simplified for complex economies like the United States, but can still describe some economies in emerging markets.

Example 11.11.6

Suppose the production of a company follows the Cobb-Douglas production model $P = 300K^{1/3}L^{2/3}$. However, cost constraints on the business force $5K + 2L \leq 180$. How do we maximize production?

The answer is easy: Lagrange multipliers! Although the constraint is an inequality, we may approach this problem in the same manner as the others. The only thing we must take into consideration is that the Lagrange multiplier method only gives us extrema *on the boundary* of the constraint. So the first partials of the function must be searched for critical points also. We begin with that. The partials of P are

$$\frac{\partial P}{\partial K} = 100K^{-2/3}L^{2/3} \quad \text{and} \quad \frac{\partial P}{\partial L} = 200K^{1/3}L^{-1/3}$$

and we must reject the only solutions of $K = 0$ and $L = 0$. So we continue with the system

$$\begin{aligned} 100K^{-2/3}L^{2/3} &= 5\lambda \\ 200K^{1/3}L^{-1/3} &= 2\lambda \\ 5K + 2L - 180 &= 0 \end{aligned}$$

The first equation tells us $\lambda = 20L^{2/3}/K^{2/3}$. We substitute this into the second equation to get $200K^{1/3}/L^{1/3} = 40L^{2/3}/K^{2/3}$, or $5K = L$. Using this relation in the third equation, we get $L = 60$; whence $K = 12$. Therefore the production is maximized at $(12, 60)$ giving $P = 300(12^{1/3})(60^{2/3}) \approx \$10,526.46$ per year. ♦

Quadratic Forms. An important theoretical application of Lagrange multipliers is in extrema of quadratic forms on the unit circle. A quadratic form is a surface of the form

$$f(x, y) = ax^2 + 2bxy + cy^2$$

for constants a, b , and c . (Ellipsoids, paraboloids, and hyperboloids have such equations.) To maximize such a function, we proceed as before, with the constraint $g(x, y) = x^2 + y^2 - 1$. We have the system

$$\begin{aligned} 2ax + 2by &= 2\lambda x \\ 2bx + 2cy &= 2\lambda y \\ x^2 + y^2 - 1 &= 0. \end{aligned}$$

But notice that the first two equations can be represented as the matrix equation $A\mathbf{x} = \lambda\mathbf{x}$ where

$$A = \begin{bmatrix} a & b \\ b & c \end{bmatrix} \quad \text{and} \quad \mathbf{x} = \begin{bmatrix} x \\ y \end{bmatrix}.$$

Moreover, \mathbf{x} must meet the constraint $x^2 + y^2 = 1$; hence $\|\mathbf{x}\| = 1$. Therefore the Lagrange multiplier is actually an eigenvalue of A , and \mathbf{x} is a unit eigenvector of A .

Hence, the eigenvalues are the critical points of f on the unit circle: particularly, the absolute maximum of f on the unit circle is the largest eigenvalue; the absolute minimum of f is the smallest eigenvalue. The corresponding unit eigenvectors are the values that give the critical points.

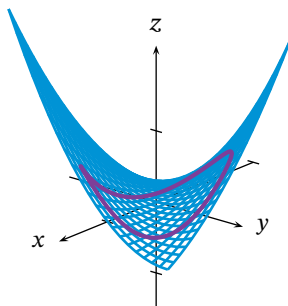


Figure 11.11 – A quadratic form (a hyperboloid of one sheet) with the unit circle projected on its surface.

So here's the connection!

Example 11.11.7

To find the extrema of $x^2 + 4xy - 2y^2$ on the unit circle, we compute the eigenvalues and eigenvectors of the associated matrix.

The matrix is $\begin{bmatrix} 1 & 2 \\ 2 & -2 \end{bmatrix}$ with eigenvalues -3 and 2 and unit eigenvectors $\frac{1}{\sqrt{5}} \langle 1, -2 \rangle$ and $\frac{1}{\sqrt{5}} \langle 2, 1 \rangle$. Hence, the maximum of f is 2 , given by $(\frac{2}{\sqrt{5}}, \frac{1}{\sqrt{5}})$; and the minimum is -3 , given by $(\frac{1}{\sqrt{5}}, -\frac{2}{\sqrt{5}})$. \blacklozenge

Problems for Section 11.11

Find the absolute extrema of the function f subject to the given constraint.

- 1 $f(x, y) = 3x^2 + 3y^2 + 5$, subject to $x - y = 1$
- 2 $f(x, y) = x^3y$, subject to $2x + y = 5$
- 3 $f(x, y) = x^3y$, subject to $\sqrt{x} + \sqrt{y} = 1$
- 4 $f(x, y) = x^3 - y^3$, subject to $x - y = 2$
- 5 $f(x, y) = e^{x+y}$, subject to $x^2 + y^2 = 2$
- 6 $f(x, y, z) = x^2 + 2y^2 + 4z^2$, subject to $x^2 + y^2 + z^2 = 1$
- 7 Find the point on the surface $z = (y+1)^2 - (x-2)^2 + 1$ nearest to the point $(2, -1, -2)$.
- 8 A closed rectangular box is to be made so that its volume is 60 cubic feet. The costs of the material for the top and bottom are 10 cents per square foot and 20 cents per square foot, respectively. The cost of the sides is 2 cents per square foot.
- (a) Determine the cost function $C(x, y)$, where x and y are the length and width of the box.
- (b) Find the dimensions of the box that will give the minimum cost and then compute the minimum cost. (Zill, 1985, p. 794)
- 9 A clothing company makes two types of overcoats, and the cost of manufacturing these overcoats is $C(x, y) = 2x^2 + 6y^2 + 4xy + 10$. If a total of 20 overcoats can be made daily, how many of each type should be made to minimize the cost? What is the minimum cost?
- 10 (Calculator) Maximize the production given by $P(K, L) = 10K^{0.6}L^{0.4}$ with constraint $2K + 3L \leq 60$.
- 11 Find the extrema of the quadratic form $13x^2 - 8xy - 2y^2$ on the unit circle.

11.12 Two Differential Operators

A modern branch of mathematics, having achieved the art of dealing with the infinitely small, can now yield solutions in other more complex problems of motion, which used to appear insoluble. This modern branch of mathematics, unknown to the ancients, when dealing with problems of motion, admits the conception of the infinitely small, and so conforms to the chief condition of motion (absolute continuity) and thereby corrects the inevitable error which the human mind cannot avoid when dealing with separate elements of motion instead of examining continuous motion. In seeking the laws of historical movement just the same thing happens. The movement of humanity, arising as it does from innumerable human wills, is continuous. To understand the laws of this continuous movement is the aim of history. Only by taking an infinitesimally small unit for observation (the differential of history, that is, the individual tendencies of man) and attaining to the art of integrating them (that is, finding the sum of these infinitesimals) can we hope to arrive at the laws of history.

— Leo Tolstoy

In this final section of the present chapter, we introduce two differential operators. We have already encountered one differential operator, the gradient. Recall that the gradient of a function $f(x, y, z)$ is the vector

$$\nabla f = \left\langle \frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \frac{\partial f}{\partial z} \right\rangle.$$

The gradient ∇ by itself is meaningless; it is an operator, and as such, we take the gradient of something. The two new differential operators

12.4 Integration in the Physical World

We speak of invention: it would be more correct to speak of discovery. The distinction between these two words is well known: discovery concerns a phenomenon, a law, a being which already existed, but had not been perceived. Columbus discovered America: it existed before him; on the contrary, Franklin invented the lightning rod: before him there had never been any lightning rod.

Such a distinction has proved less evident than it appears at first glance. Torricelli has observed that when one inverts a closed tube on a mercury trough, the mercury ascends to a certain determinate height: this is a discovery; but in doing this, he has invented the barometer; and there are plenty of examples of scientific results which are just as much discoveries as inventions.

— Jacques Hadamard

In this section, we present six applications of double integrals. Some of these applications we have already encountered, and some are three-dimensional analogues to familiar concepts. All are important to the physical sciences. We begin with two straightforward ideas.

I. Volume. If $f(x, y)$ is the equation of a surface, then $V = \iint_R f \, dx \, dy$ is the volume between the surface and the xy -plane.

II. Area. For $f(x, y) = 1$, we have $A = \iint_R dx \, dy$ as the area of R .

These definitions of volume and area arise from the definition of double integrals. Note, however, that we may use single-variable integrals to compute the same things.

For instance, consider the region R bounded by the x -axis, the line $x = 3$, and the curve $y = x^2$. Using a single-variable integral, we may compute

$$\int_0^3 x^2 \, dx = \left. \frac{x^3}{3} \right|_0^3 = 9.$$

On the other hand, we may also compute the same area using a double integral:

$$\int_0^3 \int_0^{x^2} dy \, dx = \int_0^3 y \Big|_0^{x^2} dx = \int_0^3 x^2 \, dx = 9.$$

The same goes for volume. In Volume 1, we determined the volume of a solid with defined cross-sections. The following is one such problem.

Example 12.4.1

Suppose the region R in the xy -plane, bounded by the x -axis, the line $x = 3$, and the curve $y = x^2$, is the base of a solid whose cross-sections are isosceles right triangles with a leg of each triangle in R . We may find the volume by first finding an expression for the area of a cross section. Since the cross sections are triangles of height x^2 and base x^2 , we have $A(x) = \frac{1}{2}x^4$. Then we compute

$$\int_0^3 \frac{x^4}{2} \, dx = \left. \frac{x^5}{10} \right|_0^3 = 24.3.$$

On the other hand, we could use a double integral. The region R for our double integral is the same, but in this interpretation, we want the volume of

the solid above R and below the plane $y = z$. Hence, we compute

$$\int_0^3 \int_0^{x^2} y \, dy \, dx = \int_0^3 \frac{y^2}{2} \Big|_0^{x^2} \, dx = \int_0^3 \frac{x^4}{2} \, dx = 24.3.$$

Of course, using a double integral in place of a single integral may complicate things if one cannot easily determine the surface under which we wish to find the volume. ♦

The next application concerns the mass of a **lamina**, or thin plate, sheet, or layer of a larger composite structure. A lamina will usually be defined by a region R .

III. Mass. If $\rho(x, y)$ is the equation of the density of a lamina (in mass per unit area), then $M = \iint_R \rho \, dx \, dy$ is the mass of R .

IV. Center of Mass. If ρ is density, then the center of mass (\bar{x}, \bar{y}) of the lamina represented by R is given by

$$\bar{x} = \frac{1}{M} \iint_R x\rho \, dx \, dy, \quad \bar{y} = \frac{1}{M} \iint_R y\rho \, dx \, dy$$

where M is the mass of R .

The use of the Greek letter rho (ρ) is traditional notation for density.

Example 12.4.2

We compute the center of mass of the lamina covering the triangular region with vertices $(0, 0)$, $(1, 0)$, and $(1, 2)$, given that the plate's density is $\rho(x, y) = 12x + 12y + 6$.

First, we find M , the mass of the plate:

$$\begin{aligned} \int_0^1 \int_0^{2x} (12x + 12y + 6) \, dy \, dx &= \int_0^1 (12xy + 6y^2 + 6y) \Big|_0^{2x} \, dx \\ &= \int_0^1 (48x^2 + 12x) \, dx = 16x^3 + 6x^2 \Big|_0^1 = 22. \end{aligned}$$

Now we have

$$\begin{aligned} \bar{x} &= \frac{1}{22} \int_0^1 \int_0^{2x} (12x^2 + 12xy + 6x) \, dy \, dx \\ &= \frac{1}{22} \int_0^1 (12x^2y + 6xy^2 + 6xy) \Big|_0^{2x} \, dx = \frac{1}{22} \int_0^1 (48x^3 + 12x^2) \, dx \\ &= \frac{1}{22} (12x^4 + 4x^3) \Big|_0^1 = \frac{1}{22} (16) = \frac{8}{11}. \end{aligned}$$

Similarly, we find $\bar{y} = \frac{9}{11}$; hence, the center of mass is at the point $(\frac{8}{11}, \frac{9}{11})$. ♦

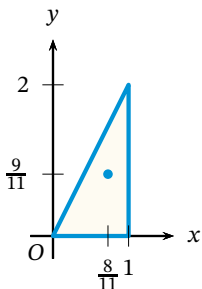


Figure 12.8 – The center of mass of the lamina from Example 12.4.2.

V. Moment of Inertia. If ρ is density, and R represents a lamina, the moments of inertia about the x -axis and y -axis are

$$I_x = \iint_R y^2 \rho \, dx \, dy, \quad I_y = \iint_R x^2 \rho \, dx \, dy,$$

and the moment of inertia about the origin is $I_O = I_x + I_y$.

The moment of inertia quantifies the resistance of a physical object to angular acceleration. The moment of inertia is to rotational motion as mass is to linear motion. An object's moment of inertia depends on its shape and the distribution of mass within that shape: the greater the concentration of material away from the object's center, the larger the moment of inertia.

This is calculated in the above manner because the definition of the moment of inertia is the product of the mass times the square of the distance from an axis. The mass is $\iint_R \rho \, dx \, dy$ and the distance from the x -axis to a point of mass is y ; hence the moment around the x -axis is $I_x = \iint_R y^2 \rho \, dx \, dy$. Likewise for the moment around the y -axis.

Example 12.4.3

We compute the moments of inertia of the lamina covering the triangular region with vertices $(0, 0)$, $(1, 0)$, and $(1, 2)$, given that the plate's density is $\rho(x, y) = 12x + 12y + 6$.

The moment of inertia about the x -axis is given by

$$\begin{aligned} I_x &= \int_0^1 \int_0^{2x} (12xy^2 + 12y^3 + 6y^2) \, dy \, dx \\ &= \int_0^1 (4xy^3 + 3y^4 + 2y^3) \Big|_0^{2x} = \int_0^1 (80x^4 + 16x^3) \, dx \\ &= 16x^5 + 4x^4 \Big|_0^1 = 20. \end{aligned}$$

Similarly, the moment of inertia about the y -axis is $\frac{63}{5}$; hence the moment of inertia about the origin is $\frac{63}{5} + 20 = \frac{163}{5}$. ♦

VI. Surface Area. The surface area of a surface in space over a region R in the xy -plane is proved as the following theorem.

THEOREM 12.C (Area of a Surface) If $z = f(x, y)$ is defined and has continuous partial derivatives in $R \subseteq D$, then the surface area of a surface in space is

$$S = \iint_R \sqrt{1 + \left(\frac{\partial z}{\partial x}\right)^2 + \left(\frac{\partial z}{\partial y}\right)^2} \, dx \, dy.$$

Proof. Let (x_i, y_i) be a point of the i th rectangle of the subdivision of $R \subseteq D$. (See Figure 12.9.) Then the tangent plane at (x_i, y_i) is

$$z - z_i = \frac{\partial z}{\partial x}(x - x_i) + \frac{\partial z}{\partial y}(y - y_i)$$

where the partials are evaluated at (x_i, y_i) . Let S_i be the area of the part of tangent plane above the i th rectangle. Then S_i is the area of a parallelogram whose projection is a rectangle on R with area A_i . Let \mathbf{n}_i be the normal to z at (x_i, y_i) ; hence,

$$\mathbf{n}_i = -\frac{\partial z}{\partial x} \mathbf{i} - \frac{\partial z}{\partial y} \mathbf{j} + \mathbf{k}.$$

Then $S_i = A_i \sec(\gamma_i)$, where γ_i is the angle between \mathbf{n}_i and \mathbf{k} . Note that

$$\cos(\gamma_i) = \frac{\mathbf{n}_i \cdot \mathbf{k}}{\|\mathbf{n}_i\| \|\mathbf{k}\|} = \frac{1}{\sqrt{1 + \left(\frac{\partial z}{\partial x}\right)^2 + \left(\frac{\partial z}{\partial y}\right)^2}}$$

so that

$$\sec(\gamma_i) = \sqrt{1 + \left(\frac{\partial z}{\partial x}\right)^2 + \left(\frac{\partial z}{\partial y}\right)^2}.$$

Hence, if we let d_i denote the diagonal of the i th rectangle,

$$\begin{aligned} \lim_{n \rightarrow \infty} \sum_{i=1}^n S_i &= \lim_{n \rightarrow \infty} \sum_{i=1}^n A_i \sec(\gamma_i) \\ &= \lim_{n \rightarrow \infty} \sum_{i=1}^n A_i \sqrt{1 + \left(\frac{\partial z}{\partial x}\right)^2 + \left(\frac{\partial z}{\partial y}\right)^2} \\ &= \iint_R \sqrt{1 + \left(\frac{\partial z}{\partial x}\right)^2 + \left(\frac{\partial z}{\partial y}\right)^2} \, dx \, dy \end{aligned}$$

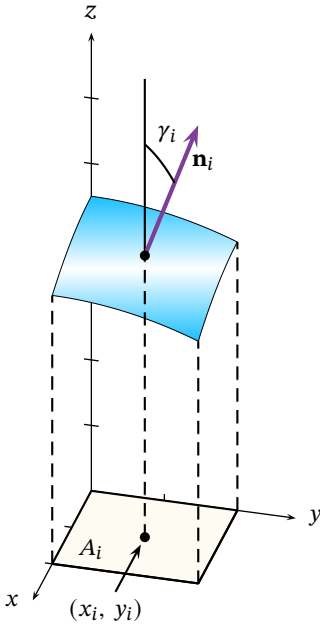


Figure 12.9 – The derivation of the surface area.

Example 12.4.4

We use the previous theorem to compute the surface area of the paraboloid $z = x^2 + y^2$ bounded by $x^2 + y^2 = 4$.

We have that $-\sqrt{4 - x^2} \leq y \leq \sqrt{4 - x^2}$ and $-2 \leq x \leq 2$. Thus,

$$A = \iint_R \sqrt{1 + (2x)^2 + (2y)^2} \, dx \, dy = \iint_R \sqrt{1 + 4x^2 + 4y^2} \, dx \, dy.$$

This integral is tedious; so we change coordinates to polar. Then the region is the complete circular disk of radius 2. Hence, $0 \leq r \leq 2$ and $0 \leq \theta \leq 2\pi$, and, since the Jacobian is r , we have

$$\begin{aligned} A &= \int_0^{2\pi} \int_0^2 r \sqrt{4r^2 + 1} \, dr \, d\theta = \int_0^{2\pi} \frac{1}{12} (4r^2 + 1)^{3/2} \Big|_0^2 \, d\theta \\ &= \int_0^{2\pi} \frac{1}{12} (17^{3/2} - 1) \, d\theta = \frac{\pi}{6} (17^{3/2} - 1). \end{aligned}$$

This is approximately 36.177. ♦

Problems for Section 12.4

- 1 Show the computations that verify the values of \bar{y} from Example 12.4.2 and I_y from Example 12.4.3.
- 2 Compute the moment of inertia about the origin for the triangular lamina with vertices $(0, 0)$, $(1, 0)$, and $(0, 3)$ and density $\rho(x, y) = 6$.
- 3 Compute the center of mass of the unit square in the first quadrant with density $\rho(x, y) = y \arctan(x)$.
- 4 Using a double integral, compute the volume of the solid whose base is the region R in the xy -plane bounded by the x -axis, the line $x = \frac{\pi}{2}$, the line $x = 0$, and the curve $y = \cos(x)$, and whose cross-sections are rectangles with width $\cos(x)$ and height 5.
- 5 Compute the surface area of the hemisphere $z = \sqrt{1 - x^2 - y^2}$ using a double integral.
- 6 Compute the surface area of the cone $2x^2 + 2y^2 = 5z^2$, $z \geq 0$, bounded by $x^2 + y^2 = 4$.