APEX Calculus

for University of Lethbridge

APEX Calculus

for University of Lethbridge

Gregory Hartman, Ph.D.
Virginia Military Institute

Sean Fitzpatrick, Ph.D., Editor
University of Lethbridge

Alex Jordan, Ph.D., Editor
Portland Community College

Carly Vollet, M.S., Editor
Portland Community College

August 17, 2023

Contributors to the 4th Edition: Jennifer Bowen, Troy Siemers, Brian Heinold, Dimplekumar Chalishajar

Edition: 5

Website: www.apexcalculus.com

©2021 Gregory Hartman

Licensed to the public under Creative Commons Attribution-Noncommercial 4.0 International Public License

Thanks

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

Preface

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

This text comprises a three—volume series on Calculus. The first part covers material taught in many "Calc 1" courses: limits, derivatives, and the basics of integration, found in Chapters 1 through 6.1. The second text covers material often taught in "Calc 2:" integration and its applications, including an introduction to differential equations, along with an introduction to sequences, series and Taylor Polynomials, found in Chapters 5 through 8. The third text covers topics common in "Calc 3" or "multivariable calc:" parametric equations, polar coordinates, vector-valued functions, and functions of more than one variable, found in Chapters 10 through 15. All three are available separately for free at apexcalculus.com[1], and HTML versions of the book can be found at opentext.uleth.ca[2].

These three texts are intended to work together and make one cohesive text, *APEX Calculus*, which can also be downloaded from the website.

Printing the entire text as one volume makes for a large, heavy, cumbersome book. One can certainly only print the pages they currently need, but some prefer to have a nice, bound copy of the text. Therefore this text has been split into these three manageable parts, each of which can be purchased for about $15 at Amazon.com[3].

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

Each chapter and section starts with an introduction of the coming material, hopefully setting the stage for "why you should care," and ends with a look ahead to see how the just—learned material helps address future problems.

- *Please read the text.*

  It is written to explain the concepts of Calculus. There are numerous examples to demonstrate the meaning of definitions, the truth of theorems, and the application of mathematical techniques. When you encounter a sentence you don't understand, read it again. If it still doesn't make sense, read on anyway, as sometimes confusing sentences are explained by later sentences.

[1] apexcalculus.com
[2] opentext.uleth.ca/calculus.html
[3] amazon.com

- *You don't have to read every equation.*

  The examples generally show "all" the steps needed to solve a problem. Sometimes reading through each step is helpful; sometimes it is confusing. When the steps are illustrating a new technique, one probably should follow each step closely to learn the new technique. When the steps are showing the mathematics needed to find a number to be used later, one can usually skip ahead and see how that number is being used, instead of getting bogged down in reading how the number was found.

- *Most proofs have been omitted.*

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

Interactive, 3D Graphics. Versions 3.0 and 4.0 of the textbook include interactive, 3D graphics in the pdf version. Nearly all graphs of objects in space can be rotated, shifted, and zoomed in/out so the reader can better understand the object illustrated. However, the only pdf viewers that support these 3D graphics are Adobe Reader Acrobat (and only the versions for PC/Mac/Unix/Linux computers, not tablets or smartphones).

The latest version of the book, which is authored in PreTeXt, is available in html. In html, the 3D graphics are rendered using WebGL, and should work in any modern web browser.

Interactive graphics are no longer supported within the pdf, but clicking on any 3D graphic within the pdf will take you directly to the interactive version on the web.

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

Each text is available as a free .pdf, protected by a Creative Commons Attribution - Noncommercial 4.0 copyright. That means you can give the .pdf to anyone you like, print it in any form you like, and even edit the original content and redistribute it. If you do the latter, you must clearly reference this work and you cannot sell your edited work for money.

We encourage others to adapt this work to fit their own needs. One might add sections that are "missing" or remove sections that your students won't

need. The source files can be found at github.com/APEXCalculus[4].

You can learn more at www.vmi.edu/APEX[5].

First PreTeXt Edition (Version 5.0). Key changes from Version 4.0 to 5.0:

- The underlying source code has been completely rewritten, to use the PreTeXt[6] language, instead of the original LaTeX.
- Using PreTeXt allows us to produce the books in multiple formats, including html, which is both more accessible and more interactive than the original pdf. html versions of the book can be found at opentext.uleth.ca[7].
- The appendix on differential equations from the "Calculus for Quarters" version of the book has been included as Chapter 8, just after applications of integration. Chapters 8 — 14 are now numbered 9 — 15 as a result.
- In the html version of the book, many of the exercises are now interactive, and powered by WeBWorK.

Key changes from Version 3.0 to 4.0:

- Numerous typographical and "small" mathematical corrections (again, thanks to all my close readers!).
- "Large" mathematical corrections and adjustments. There were a number of places in Version 3.0 where a definition/theorem was not correct as stated. See www.apexcalculus.com[8] for more information.
- More useful numbering of Examples, Theorems, etc. . "Definition 11.4.2" refers to the second definition of Chapter 11, Section 4.
- The addition of Section 13.7: Triple Integration with Cylindrical and Spherical Coordinates
- The addition of Chapter 14: Vector Analysis.

[4] github.com/APEXCalculus
[5] www.vmi.edu/APEX
[6] pretextbook.org
[7] opentext.uleth.ca/calculus.html
[8] apexcalculus.com

x

A Brief History of Calculus

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

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

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

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

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

Contents

Thanks v

Preface vii

A Brief History of Calculus xi

I Math 1565: Accelerated Calculus I

1 Limits 3

1.1 An Introduction To Limits 3
1.2 Epsilon-Delta Definition of a Limit 11
1.3 Finding Limits Analytically 19
1.4 One-Sided Limits 31
1.5 Continuity 40
1.6 Limits Involving Infinity 52

2 Derivatives 65

2.1 Instantaneous Rates of Change: The Derivative 65
2.2 Interpretations of the Derivative 82
2.3 Basic Differentiation Rules 89
2.4 The Product and Quotient Rules 97
2.5 The Chain Rule 107
2.6 Implicit Differentiation. 117
2.7 Derivatives of Inverse Functions 128

3 The Graphical Behavior of Functions 135

3.1 Extreme Values 135
3.2 The Mean Value Theorem 143
3.3 Increasing and Decreasing Functions 149
3.4 Concavity and the Second Derivative 157
3.5 Curve Sketching 166

4 Applications of the Derivative **175**

4.1 Newton's Method 175
4.2 Related Rates. 181
4.3 Optimization 190
4.4 Differentials 197
4.5 Taylor Polynomials 205
4.6 L'Hospital's Rule 217

5 Integration **227**

5.1 Antiderivatives and Indefinite Integration 227
5.2 The Definite Integral 236
5.3 Riemann Sums 248
5.4 The Fundamental Theorem of Calculus 263
5.5 Numerical Integration 277
5.6 Substitution 293
5.7 Hyperbolic Functions 310

II Math 2565: Accelerated Calculus II

6 Techniques of Antidifferentiation **325**

6.1 Integration by Parts. 325
6.2 Trigonometric Integrals 335
6.3 Trigonometric Substitution 346
6.4 Partial Fraction Decomposition 356
6.5 Improper Integration 364

7 Applications of Integration **375**

7.1 Area Between Curves 376
7.2 Volume by Cross-Sectional Area; Disk and Washer Methods . . 385
7.3 The Shell Method 395
7.4 Arc Length and Surface Area 404
7.5 Work 414
7.6 Fluid Forces 424

8 Differential Equations **433**

8.1 Graphical and Numerical Solutions to Differential Equations . . 433
8.2 Separable Differential Equations 446
8.3 First Order Linear Differential Equations 452
8.4 Modeling with Differential Equations 460

9 Sequences and Series **471**

9.1 Sequences. 471
9.2 Infinite Series. 485
9.3 Integral and Comparison Tests 501
9.4 Ratio and Root Tests 512
9.5 Alternating Series and Absolute Convergence 519
9.6 Power Series 529
9.7 Taylor Series 540

10 Curves in the Plane **553**

10.1 Conic Sections 553
10.2 Parametric Equations 567
10.3 Calculus and Parametric Equations. 577
10.4 Introduction to Polar Coordinates 587
10.5 Calculus and Polar Functions 601

11 Introduction to Functions of Several Variables **613**

11.1 Introduction to Multivariable Functions 613
11.2 Limits and Continuity of Multivariable Functions 621
11.3 Partial Derivatives 630

III Math 2575: Accelerated Calculus III

12 Vectors **645**

12.1 Introduction to Cartesian Coordinates in Space 645
12.2 An Introduction to Vectors 661
12.3 The Dot Product 674
12.4 The Cross Product 687
12.5 Lines 697
12.6 Planes 706

13 Vector Valued Functions **715**

13.1 Vector-Valued Functions 715
13.2 Calculus and Vector-Valued Functions 721
13.3 The Calculus of Motion 733
13.4 Unit Tangent and Normal Vectors 746
13.5 The Arc Length Parameter and Curvature 755

14 Functions of Several Variables, Continued **767**

14.1 Differentiability and the Total Differential 767
14.2 The Multivariable Chain Rule 777
14.3 Directional Derivatives. 786
14.4 Tangent Lines, Normal Lines, and Tangent Planes 796
14.5 Extreme Values 806
14.6 The Derivative as a Linear Transformation 815
14.7 Constrained Optimization and Lagrange Multipliers 824
14.8 Hessians and the General Second Derivative Test 830

15 Multiple Integration **837**

15.1 Iterated Integrals and Area 837
15.2 Double Integration and Volume. 846
15.3 Double Integration with Polar Coordinates 857
15.4 Center of Mass 865
15.5 Surface Area 877
15.6 Volume Between Surfaces and Triple Integration 884
15.7 Triple Integration with Cylindrical and Spherical Coordinates . . 907
15.8 Change of Variables in Multiple Integrals. 918

16 Vector Analysis **941**

16.1 Introduction to Line Integrals 942
16.2 Vector Fields 952
16.3 Line Integrals over Vector Fields 961
16.4 Flow, Flux, Green's Theorem and the Divergence Theorem . . . 971
16.5 Parametrized Surfaces and Surface Area 981
16.6 Surface Integrals. 993
16.7 The Divergence Theorem and Stokes' Theorem 1000

Appendices

A Answers to Selected Exercises **1017**

B Quick Reference **1103**

B.1 Differentiation Formulas 1103
B.2 Integration Formulas 1104
B.3 Trigonometry Reference 1106
B.4 Areas and Volumes 1108
B.5 Algebra. 1109
B.6 Additional Formulas 1111
B.7 Summary of Tests for Series 1112

Back Matter

Index **1113**

Part III

Math 2575: Accelerated Calculus III

Chapter 12

Vectors

This chapter begins with moving our mathematics out of the plane and into "space." That is, we begin to think mathematically not only in two dimensions, but in three. With this foundation, we can explore vectors both in the plane and in space.

12.1 Introduction to Cartesian Coordinates in Space

Up to this point in this text we have considered mathematics in a 2-dimensional world. We have plotted graphs on the xy-plane using rectangular and polar coordinates and found the area of regions in the plane. We have considered properties of *solid* objects, such as volume and surface area, but only by first defining a curve in the plane and then rotating it out of the plane.

While there is wonderful mathematics to explore in "2D," we live in a "3D" world and eventually we will want to apply mathematics involving this third dimension. In this section we introduce Cartesian coordinates in space and explore basic surfaces. This will lay a foundation for much of what we do in the remainder of the text.

Each point P in space can be represented with an ordered triple, $P = (a, b, c)$, where a, b and c represent the relative position of P along the x-, y- and z-axes, respectively. Each axis is perpendicular to the other two.

Visualizing points in space on paper can be problematic, as we are trying to represent a 3-dimensional concept on a 2-dimensional medium. We cannot draw three lines representing the three axes in which each line is perpendicular to the other two. Despite this issue, standard conventions exist for plotting shapes in space that we will discuss that are more than adequate.

One convention is that the axes must conform to the *right hand rule*. This rule states that when the index finger of the right hand is extended in the direction of the positive x-axis, and the middle finger (bent "inward" so it is perpendicular to the palm) points along the positive y-axis, then the extended thumb will point in the direction of the positive z-axis. (It may take some thought to verify this, but this system is inherently different from the one created by using the "left hand rule.")

As long as the coordinate axes are positioned so that they follow this rule, it does not matter how the axes are drawn on paper. There are two popular methods that we briefly discuss.

In Figure 12.1.2 we see the point $P = (2, 1, 3)$ plotted on a set of axes. The basic convention here is that the xy-plane is drawn in its standard way, with

youtu.be/watch?v=zqaTPul--ZU

Figure 12.1.1 Video introduction to Section 12.1

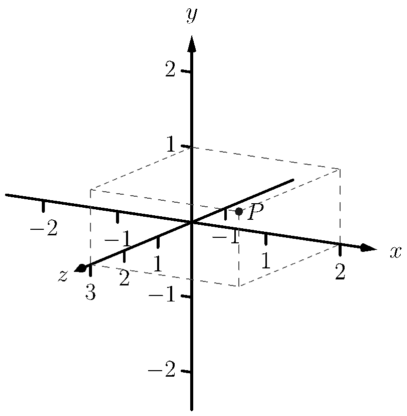


Figure 12.1.2 Plotting the point $P = (2, 1, 3)$ in space

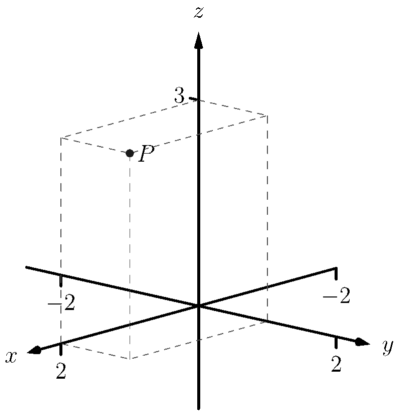


Figure 12.1.3 Plotting the point $P = (2, 1, 3)$ in space with a perspective used in this text

the z-axis down to the left. The perspective is that the paper represents the xy-plane and the positive z axis is coming up, off the page. This method is preferred by many engineers. Because it can be hard to tell where a single point lies in relation to all the axes, dashed lines have been added to let one see how far along each axis the point lies.

One can also consider the xy-plane as being a horizontal plane in, say, a room, where the positive z-axis is pointing up. When one steps back and looks at this room, one might draw the axes as shown in Figure 12.1.3. The same point P is drawn, again with dashed lines. This point of view is preferred by most mathematicians, and is the convention adopted by this text.

Just as the x- and y-axes divide the plane into four *quadrants*, the x-, y-, and z-coordinate planes divide space into eight *octants*. The octant in which x, y, and z are positive is called the *first octant*. We do not name the other seven octants in this text.

youtu.be/watch?v=zNUbi5ahC0Y

Figure 12.1.5 Video presentation of Definition 12.1.4

12.1.1 Measuring Distances

It is of critical importance to know how to measure distances between points in space. The formula for doing so is based on measuring distance in the plane, and is known (in both contexts) as the Euclidean measure of distance.

Definition 12.1.4 Distance In Space.

Let $P = (x_1, y_1, z_1)$ and $Q = (x_2, y_2, z_2)$ be points in space. The distance D between P and Q is

$$D = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2 + (z_2 - z_1)^2}.$$

We refer to the line segment that connects points P and Q in space as $\overline{PQ}$, and refer to the length of this segment as $\left\|\overline{PQ}\right\|$. The above distance formula allows us to compute the length of this segment.

Video Solution

youtu.be/watch?v=EF-aTKoAQsU

Example 12.1.6 Length of a line segment.

Let $P = (1, 4, -1)$ and let $Q = (2, 1, 1)$. Draw the line segment $\overline{PQ}$ and find its length.

Solution. The points P and Q are plotted in Figure 12.1.7; no special consideration need be made to draw the line segment connecting these two points; simply connect them with a straight line. One *cannot* actually measure this line on the page and deduce anything meaningful; its true length must be measured analytically. Applying Definition 12.1.4, we have

$$\left\|\overline{PQ}\right\| = \sqrt{(2-1)^2 + (1-4)^2 + (1-(-1))^2} = \sqrt{14} \approx 3.74.$$

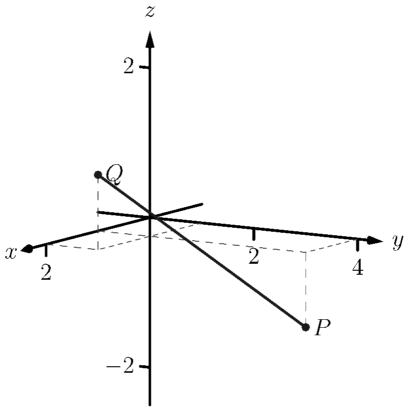


Figure 12.1.7 Plotting points P and Q in Example 12.1.6

12.1.2 Spheres

Just as a circle is the set of all points in the *plane* equidistant from a given point (its center), a sphere is the set of all points in *space* that are equidistant from a given point. Definition 12.1.4 allows us to write an equation of the sphere.

We start with a point $C = (a, b, c)$ which is to be the center of a sphere with radius r. If a point $P = (x, y, z)$ lies on the sphere, then P is r units from C;

that is,

$$\left\|\overrightarrow{PC}\right\| = \sqrt{(x-a)^2 + (y-b)^2 + (z-c)^2} = r.$$

Squaring both sides, we get the standard equation of a sphere in space with center at $C = (a, b, c)$ with radius r, as given in the following Key Idea.

Key Idea 12.1.8 Standard Equation of a Sphere in Space.

The standard equation of the sphere with radius r, centered at $C = (a, b, c)$, is

$$(x-a)^2 + (y-b)^2 + (z-c)^2 = r^2.$$

youtu.be/watch?v=Xzs2ita0nI0

Figure 12.1.9 Video presentation of Key Idea 12.1.8

Example 12.1.10 Equation of a sphere.

Find the center and radius of the sphere defined by $x^2 + 2x + y^2 - 4y + z^2 - 6z = 2$.

Solution. To determine the center and radius, we must put the equation in standard form. This requires us to complete the square (three times).

$$\begin{aligned} x^2 + 2x + y^2 - 4y + z^2 - 6z &= 2 \\ (x^2 + 2x + 1) + (y^2 - 4y + 4) + (z^2 - 6z + 9) - 14 &= 2 \\ (x+1)^2 + (y-2)^2 + (z-3)^2 &= 16 \end{aligned}$$

The sphere is centered at $(-1, 2, 3)$ and has a radius of 4.

Video Solution

youtu.be/watch?v=O0VtyYNXeig

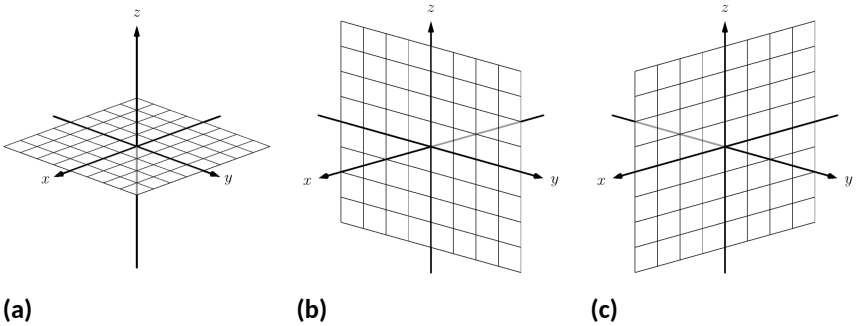
The equation of a sphere is an example of an implicit function defining a surface in space. In the case of a sphere, the variables x, y and z are all used. We now consider situations where surfaces are defined where one or two of these variables are absent.

12.1.3 Introduction to Planes in Space

The coordinate axes naturally define three planes (shown in Figure 12.1.12), the *coordinate planes*: the xy-plane, the yz-plane and the xz-plane. The xy-plane is characterized as the set of all points in space where the z-value is 0. This, in fact, gives us an equation that describes this plane: $z = 0$. Likewise, the xz-plane is all points where the y-value is 0, characterized by $y = 0$.

youtu.be/watch?v=8yImf_yW-Ys

Figure 12.1.11 Video introduction to Subsection 12.1.3



(a) **(b)** **(c)**

Figure 12.1.12 The xy-plane in (a), the yz-plane in (b) and the xz-plane in (c)

The equation $x = 2$ describes all points in space where the x-value is 2. This is a plane, parallel to the yz-coordinate plane, shown in Figure 12.1.13.

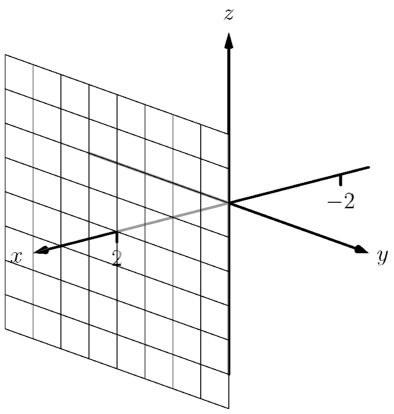


Figure 12.1.13 The plane $x = 2$

Video Solution

youtu.be/watch?v=-9luj-GlMqA

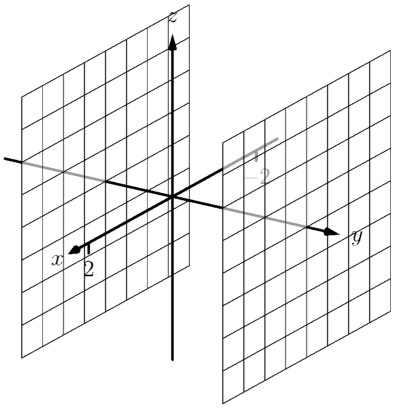


Figure 12.1.15 Sketching the boundaries of a region in Example 12.1.14

youtu.be/watch?v=9wusVjjItyY

Figure 12.1.16 Video introduction to Subsection 12.1.4

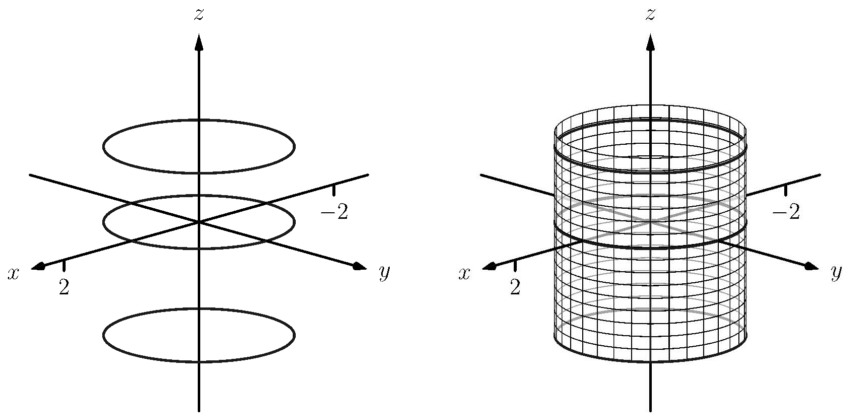
Example 12.1.14 Regions defined by planes.

Sketch the region defined by the inequalities $-1 \leq y \leq 2$.

Solution. The region is all points between the planes $y = -1$ and $y = 2$. These planes are sketched in Figure 12.1.15, which are parallel to the xz-plane. Thus the region extends infinitely in the x and z directions, and is bounded by planes in the y direction.

12.1.4 Cylinders

The equation $x = 1$ obviously lacks the y and z variables, meaning it defines points where the y and z coordinates can take on any value. Now consider the equation $x^2 + y^2 = 1$ *in space.* In *the plane*, this equation describes a circle of radius 1, centered at the origin. In space, the z coordinate is not specified, meaning it can take on any value. In Figure 12.1.17(a), we show part of the graph of the equation $x^2 + y^2 = 1$ by sketching 3 circles: the bottom one has a constant z-value of -1.5, the middle one has a z-value of 0 and the top circle has a z-value of 1. By plotting *all* possible z-values, we get the surface shown in Figure 12.1.17(b). This surface looks like a "tube," or a "cylinder"; mathematicians call this surface a *cylinder* for an entirely different reason.



(a) **(b)**

Figure 12.1.17 Sketching $x^2 + y^2 = 1$

Definition 12.1.18 Cylinder.

Let C be a curve in a plane and let L be a line not parallel to C. A **cylinder** is the set of all lines parallel to L that pass through C. The curve C is the **directrix** of the cylinder, and the lines are the **rulings**.

In this text, we consider curves C that lie in planes parallel to one of the coordinate planes, and lines L that are perpendicular to these planes, forming *right cylinders*. Thus the directrix can be defined using equations involving 2 variables, and the rulings will be parallel to the axis of the third variable.

In the example preceding the definition, the curve $x^2 + y^2 = 1$ in the xy-plane is the directrix and the rulings are lines parallel to the z-axis. (Any circle shown in Figure 12.1.17 can be considered a directrix; we simply choose the one where $z = 0$.) Sample rulings can also be viewed in Figure 12.1.17(b). More examples will help us understand this definition.

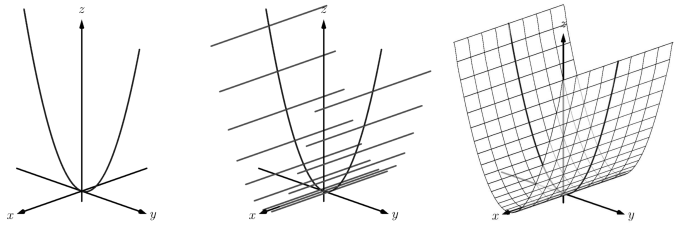
Example 12.1.19 Graphing cylinders.

Graph the following cylinders.

1. $z = y^2$
2. $x = \sin(z)$

Solution.

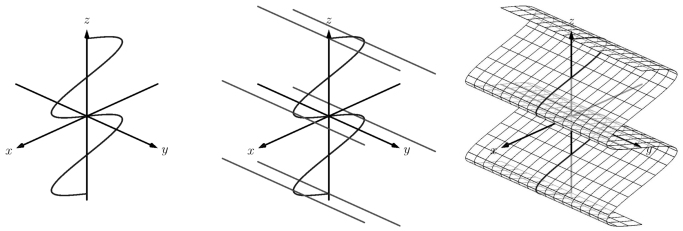
1. We can view the equation $z = y^2$ as a parabola in the yz-plane, as illustrated in Figure 12.1.20(a). As x does not appear in the equation, the rulings are lines through this parabola parallel to the x-axis, shown in Figure 12.1.20(b). These rulings give a general idea as to what the surface looks like, drawn in Figure 12.1.20(c).



(a) **(b)** **(c)**

Figure 12.1.20 Sketching the cylinder defined by $z = y^2$

2. We can view the equation $x = \sin(z)$ as a sine curve that exists in the xz-plane, as shown in Figure 12.1.21(a). The rules are parallel to the y axis as the variable y does not appear in the equation $x = \sin(z)$; some of these are shown in Figure 12.1.21(b). The surface is shown in Figure 12.1.21(c).



(a) **(b)** **(c)**

Figure 12.1.21 Sketching the cylinder defined by $x = \sin(z)$

Video Solution

youtu.be/watch?v=n7Yp4aWqeFc

12.1.5 Surfaces of Revolution

One of the applications of integration we learned previously was to find the volume of solids of revolution — solids formed by revolving a curve about a horizontal or vertical axis. We now consider how to find the equation of the surface of such a solid.

Consider the surface formed by revolving $y = \sqrt{x}$ about the x-axis. Cross-sections of this surface parallel to the yz-plane are circles, as shown in Figure 12.1.23(a). Each circle has equation of the form $y^2 + z^2 = r^2$ for some radius r. The radius is a function of x; in fact, it is $r(x) = \sqrt{x}$. Thus the equation of the surface shown in Figure 12.1.23(b) is $y^2 + z^2 = (\sqrt{x})^2$.

youtu.be/watch?v=gxnpRE0b68U

Figure 12.1.22 Video presentation of Subsection 12.1.5

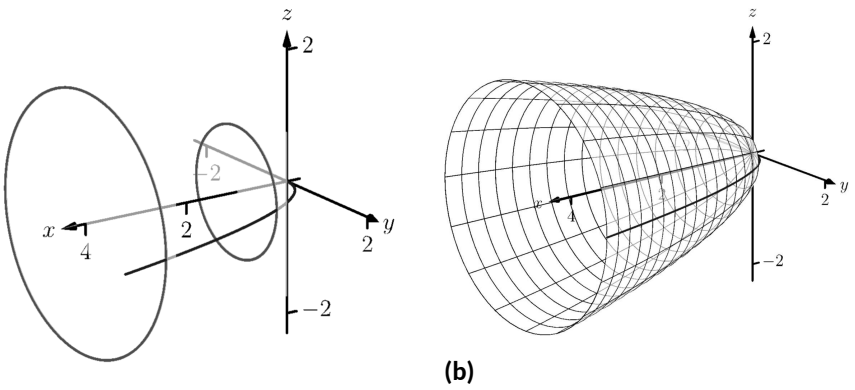


Figure 12.1.23 Introducing surfaces of revolution

We generalize the above principles to give the equations of surfaces formed by revolving curves about the coordinate axes.

Key Idea 12.1.24 Surfaces of Revolution, Part 1.

Let r be a radius function.

1. The equation of the surface formed by revolving $y = r(x)$ or $z = r(x)$ about the x-axis is $y^2 + z^2 = r(x)^2$.
2. The equation of the surface formed by revolving $x = r(y)$ or $z = r(y)$ about the y-axis is $x^2 + z^2 = r(y)^2$.
3. The equation of the surface formed by revolving $x = r(z)$ or $y = r(z)$ about the z-axis is $x^2 + y^2 = r(z)^2$.

Example 12.1.25 Finding equation of a surface of revolution.

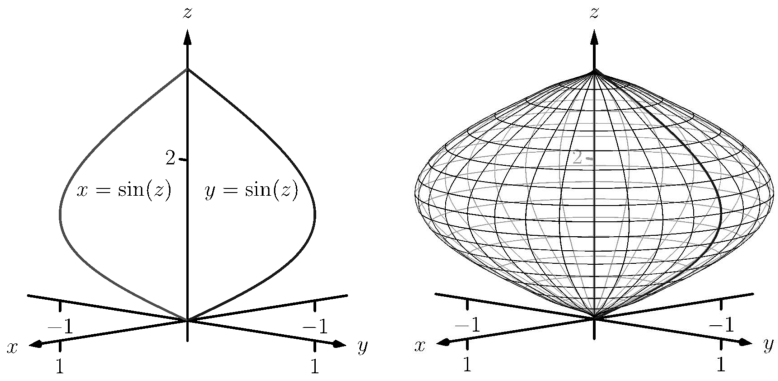
Let $y = \sin(z)$ on $[0, \pi]$. Find the equation of the surface of revolution formed by revolving $y = \sin(z)$ about the z-axis.

Solution. Using Key Idea 12.1.24, we find the surface has equation $x^2 + y^2 = \sin^2(z)$. The curve is sketched in Figure 12.1.26(a) and the surface is drawn in Figure 12.1.26(b).

Note how the surface (and hence the resulting equation) is the same if we began with the curve $x = \sin(z)$, which is also drawn in Figure 12.1.26(a).

Video Solution

youtu.be/watch?v=jrJXSnHU-0o



(a) **(b)**

Figure 12.1.26 Revolving $y = \sin(z)$ about the z-axis in Example 12.1.25

This particular method of creating surfaces of revolution is limited. For instance, in Example 7.3.11 of Section 7.3 we found the volume of the solid formed by revolving $y = \sin(x)$ about the y-axis. Our current method of forming surfaces can only rotate $y = \sin(x)$ about the x-axis. Trying to rewrite $y = \sin(x)$ as a function of y is not trivial, as simply writing $x = \sin^{-1}(y)$ only gives part of the region we desire.

What we desire is a way of writing the surface of revolution formed by rotating $y = f(x)$ about the y-axis. We start by first recognizing this surface is the same as revolving $z = f(x)$ about the z-axis. This will give us a more natural way of viewing the surface.

A value of x is a measurement of distance from the z-axis. At the distance r, we plot a z-height of $f(r)$. When rotating $f(x)$ about the z-axis, we want all points a distance of r from the z-axis in the xy-plane to have a z-height of $f(r)$. All such points satisfy the equation $r^2 = x^2 + y^2$; hence $r = \sqrt{x^2 + y^2}$. Replacing r with $\sqrt{x^2 + y^2}$ in $f(r)$ gives $z = f(\sqrt{x^2 + y^2})$. This is the equation of the surface.

Key Idea 12.1.27 Surfaces of Revolution, Part 2.

Let $z = f(x)$, $x \geq 0$, be a curve in the xz-plane. The surface formed by revolving this curve about the z-axis has equation $z = f\left(\sqrt{x^2 + y^2}\right)$.

Example 12.1.28 Finding equation of surface of revolution.

Find the equation of the surface found by revolving $z = \sin(x)$ about the z-axis.

Solution. Using Key Idea 12.1.27, the surface has equation $z = \sin\left(\sqrt{x^2 + y^2}\right)$. The curve and surface are graphed in Figure 12.1.29.

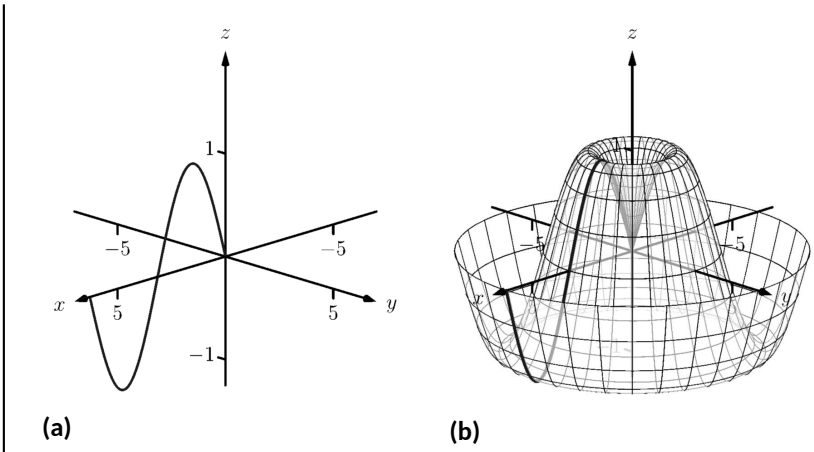


Figure 12.1.29 Revolving $z = \sin(x)$ about the z-axis in Example 12.1.28

youtu.be/watch?v=Ax_QfsRTwrc

Figure 12.1.30 Video introduction to Subsection 12.1.6

12.1.6 Quadric Surfaces

Spheres, planes and cylinders are important surfaces to understand. We now consider one last type of surface, a *quadric surface*. The definition may look intimidating, but we will show how to analyze these surfaces in an illuminating way.

Definition 12.1.31 Quadric Surface.

A **quadric surface** is the graph of the general second-degree equation in three variables:

$$Ax^2 + By^2 + Cz^2 + Dxy + Exz + Fyz + Gx + Hy + Iz + J = 0.$$

When the coefficients D, E or F are not zero, the basic shapes of the quadric surfaces are rotated in space. We will focus on quadric surfaces where these coefficients are 0; we will not consider rotations. There are six basic quadric surfaces: the elliptic paraboloid, elliptic cone, ellipsoid, hyperboloid of one sheet, hyperboloid of two sheets, and the hyperbolic paraboloid.

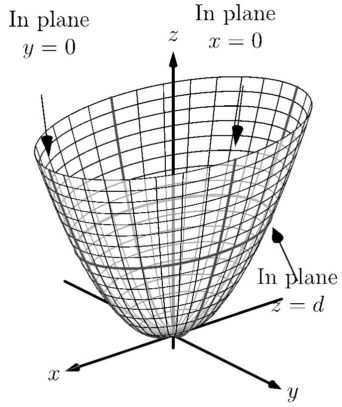


Figure 12.1.32 The elliptic paraboloid $z = x^2/4 + y^2$

We study each shape by considering *traces*, that is, intersections of each surface with a plane parallel to a coordinate plane. For instance, consider the elliptic paraboloid $z = x^2/4 + y^2$, shown in Figure 12.1.32. If we intersect this shape with the plane $z = d$ (i.e., replace z with d), we have the equation:

$$d = \frac{x^2}{4} + y^2.$$

Divide both sides by d:

$$1 = \frac{x^2}{4d} + \frac{y^2}{d}.$$

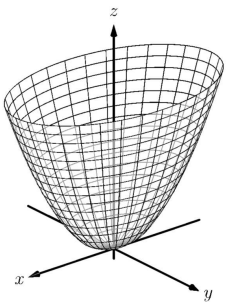
This describes an ellipse — so cross sections parallel to the xy-coordinate plane are ellipses. This ellipse is drawn in the figure.

Now consider cross sections parallel to the xz-plane. For instance, letting $y = 0$ gives the equation $z = x^2/4$, clearly a parabola. Intersecting with the plane $x = 0$ gives a cross section defined by $z = y^2$, another parabola. These parabolas are also sketched in the figure.

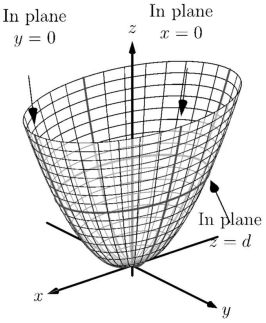
Thus we see where the elliptic paraboloid gets its name: some cross sections are ellipses, and others are parabolas.

Such an analysis can be made with each of the quadric surfaces. We give a sample equation of each, provide a sketch with representative traces, and describe these traces.

Elliptic Paraboloid $z = \dfrac{x^2}{a^2} + \dfrac{y^2}{b^2}$

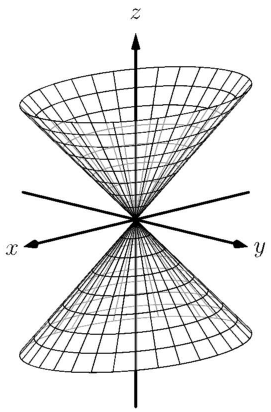


Plane	*Trace*
$x = d$	Parab
$y = d$	Parab
$z = d$	Ellips

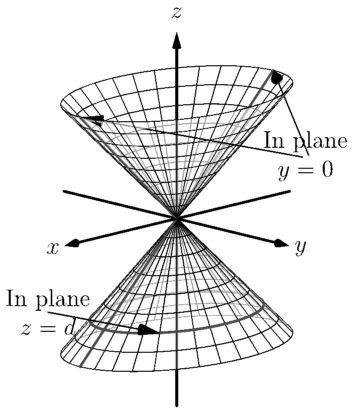


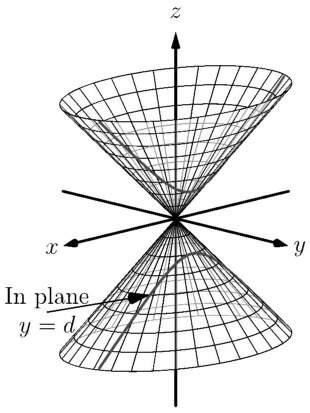
One variable in the equation of the elliptic paraboloid will be raised to the first power; above, this is the z variable. The paraboloid will "open" in the direction of this variable's axis. Thus $x = y^2/a^2 + z^2/b^2$ is an elliptic paraboloid that opens along the x-axis. Multiplying the right hand side by (-1) defines an elliptic paraboloid that "opens" in the opposite direction.

Elliptic Cone $z^2 = \dfrac{x^2}{a^2} + \dfrac{y^2}{b^2}$



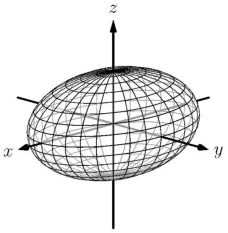
Plane	*Trace*
$x = 0$	Crossed Lines
$y = 0$	Crossed Lines
$x = d$	Hyperbola
$y = d$	Hyperbola
$z = d$	Ellipse



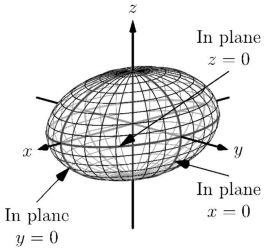


One can rewrite the equation as $z^2 - x^2/a^2 - y^2/b^2 = 0$. The one variable with a positive coefficient corresponds to the axis that the cones "open" along.

Ellipsoid $\dfrac{x^2}{a^2} + \dfrac{y^2}{b^2} + \dfrac{z^2}{c^2} = 1$

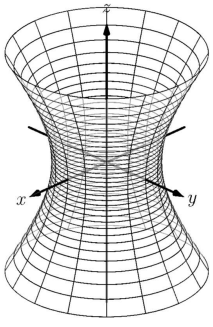


Plane	*Trace*
$x = d$	Ellips
$y = d$	Ellips
$z = d$	Ellips

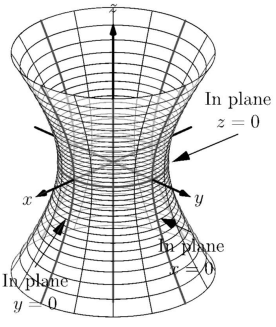


If $a = b = c \neq 0$, the ellipsoid is a sphere with radius a; compare to Key Idea 12.1.8.

Hyperboloid of One Sheet $\dfrac{x^2}{a^2} + \dfrac{y^2}{b^2} - \dfrac{z^2}{c^2} = 1$

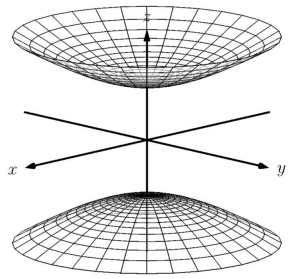


Plane	*Trace*
$x = d$	Hype
$y = d$	Hype
$z = d$	Ellips

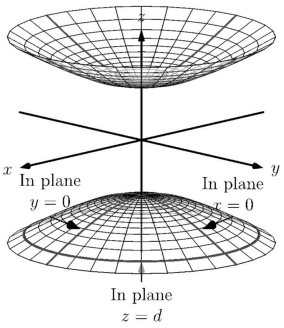


The one variable with a negative coefficient corresponds to the axis that the hyperboloid "opens" along.

Hyperboloid of Two Sheets $\dfrac{z^2}{c^2} - \dfrac{x^2}{a^2} - \dfrac{y^2}{b^2} = 1$

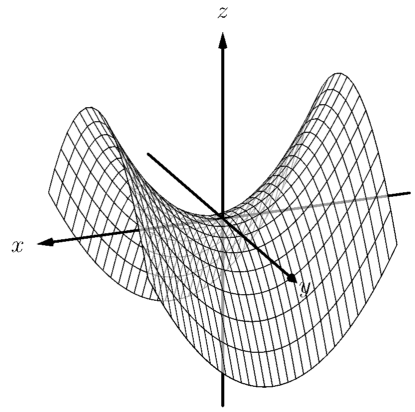


Plane	*Trace*
$x = d$	Hype
$y = d$	Hype
$z = d$	Ellips

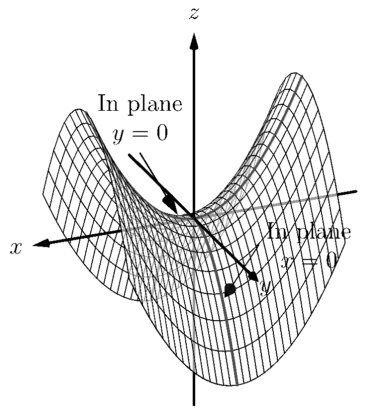


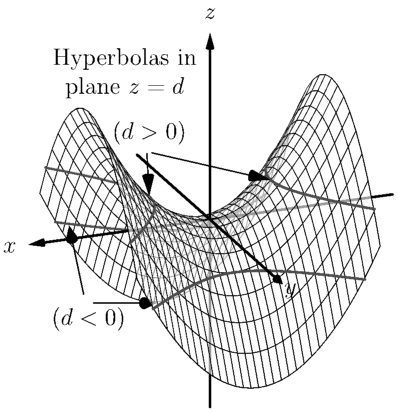
The one variable with a positive coefficient corresponds to the axis that the hyperboloid "opens" along. In the case illustrated, when $|d| < |c|$, there is no trace.

Hyperbolic Paraboloid $z = \dfrac{x^2}{a^2} - \dfrac{y^2}{b^2}$



Plane	*Trace*
$x = d$	Parabola
$y = d$	Parabola
$z = d$	Hyperbola





The parabolic traces will open along the axis of the one variable that is raised to the first power.

Example 12.1.33 Sketching quadric surfaces.

Sketch the quadric surface defined by the given equation.

Video Solution

youtu.be/watch?v=7BPClg70Lmg

1. $y = \dfrac{x^2}{4} + \dfrac{z^2}{16}$
2. $x^2 + \dfrac{y^2}{9} + \dfrac{z^2}{4} = 1$
3. $z = y^2 - x^2$

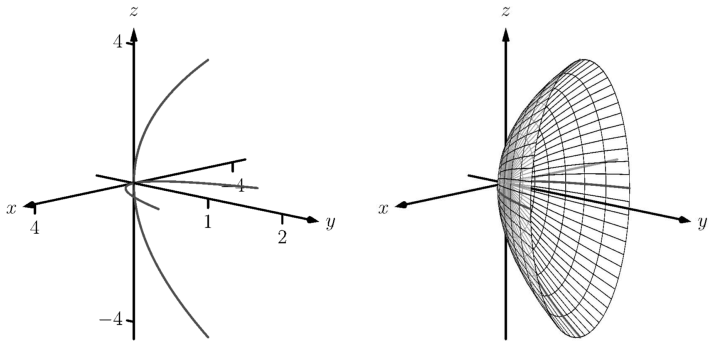
Solution.

1. $y = \dfrac{x^2}{4} + \dfrac{z^2}{16}$: We first identify the quadric by pattern-matching with the equations given previously. Only two surfaces have equations where one variable is raised to the first power, the elliptic paraboloid and the hyperbolic paraboloid. In the latter case, the other variables have different signs, so we conclude that this describes a hyperbolic paraboloid. As the variable with the first power is y, we note the paraboloid opens along the y-axis. To make a decent sketch by hand, we need only draw a few traces. In this case, the traces $x = 0$ and $z = 0$ form parabolas that outline the shape.

   $x = 0$: The trace is the parabola $y = z^2/16$

   $z = 0$: The trace is the parabola $y = x^2/4$.

   Graphing each trace in the respective plane creates a sketch as shown in Figure 12.1.34(a). This is enough to give an idea of what the paraboloid looks like. The surface is filled in in Figure 12.1.34(b).



(a) **(b)**

Figure 12.1.34 Sketching an elliptic paraboloid

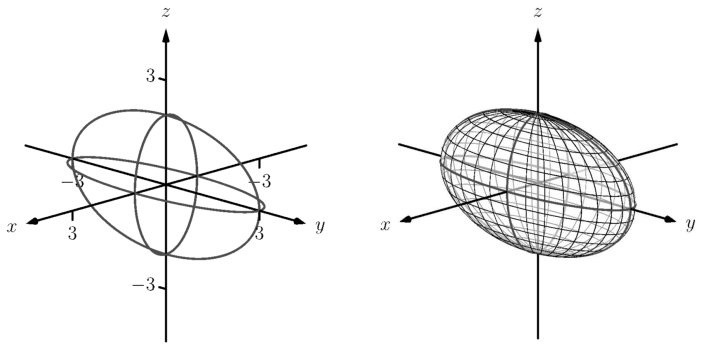
2. $x^2 + \dfrac{y^2}{9} + \dfrac{z^2}{4} = 1$: This is an ellipsoid. We can get a good idea of its shape by drawing the traces in the coordinate planes.

   $x = 0$: The trace is the ellipse $\dfrac{y^2}{9} + \dfrac{z^2}{4} = 1$. The major axis is along the y-axis with length 6 (as $b = 3$, the length of the axis is 6); the minor axis is along the z-axis with length 4.

   $y = 0$: The trace is the ellipse $x^2 + \dfrac{z^2}{4} = 1$. The major axis is along the z-axis, and the minor axis has length 2 along the x-axis.

   $z = 0$: The trace is the ellipse $x^2 + \dfrac{y^2}{9} = 1$, with major axis along the y-axis.

Graphing each trace in the respective plane creates a sketch as shown in Figure 12.1.35(a). Filling in the surface gives Figure 12.1.35(b).



(a) **(b)**

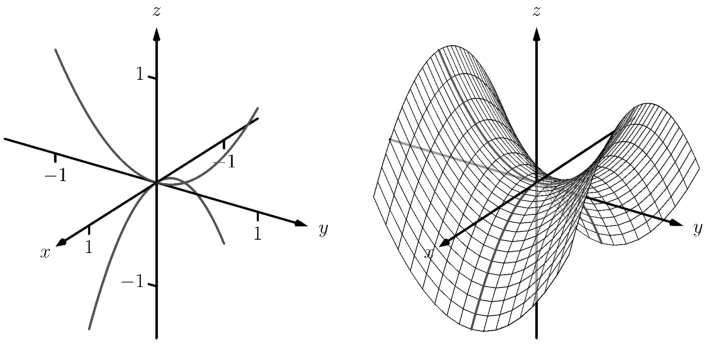
Figure 12.1.35 Sketching an ellipsoid

3. $z = y^2 - x^2$: This defines a hyperbolic paraboloid, very similar to the one shown in the gallery of quadric sections. Consider the traces in the $y - z$ and $x - z$ planes:

   $x = 0$: The trace is $z = y^2$, a parabola opening up in the $y - z$ plane.

   $y = 0$: The trace is $z = -x^2$, a parabola opening down in the $x - z$ plane.

   Sketching these two parabolas gives a sketch like that in Figure 12.1.36(a), and filling in the surface gives a sketch like Figure 12.1.36(b).



(a) **(b)**

Figure 12.1.36 Sketching a hyperbolic paraboloid

Example 12.1.37 Identifying quadric surfaces.

Consider the quadric surface shown in Figure 12.1.38. Which of the following equations best fits this surface?

(a) $x^2 - y^2 - \dfrac{z^2}{9} = 0$

(b) $x^2 - y^2 - z^2 = 1$

(c) $z^2 - x^2 - y^2 = 1$

(d) $4x^2 - y^2 - \dfrac{z^2}{9} = 1$

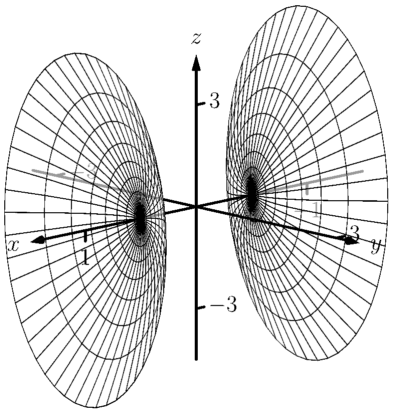


Figure 12.1.38 A possible equation of this quadric surface is found in Example 12.1.37

Solution. The image clearly displays a hyperboloid of two sheets. The gallery informs us that the equation will have a form similar to $\frac{z^2}{c^2} - \frac{x^2}{a^2} - \frac{y^2}{b^2} = 1$.

We can immediately eliminate option (a), as the constant in that equation is not 1.

The hyperboloid "opens" along the x-axis, meaning x must be the only variable with a positive coefficient, eliminating (c).

The hyperboloid is wider in the z-direction than in the y-direction, so we need an equation where $c > b$. This eliminates (b), leaving us with (d). We should verify that the equation given in (d), $4x^2 - y^2 - \frac{z^2}{9} = 1$, fits.

We already established that this equation describes a hyperboloid of two sheets that opens in the x-direction and is wider in the z-direction than in the y. Now note the coefficient of the x-term. Rewriting $4x^2$ in standard form, we have: $4x^2 = \dfrac{x^2}{(1/2)^2}$. Thus when $y = 0$ and $z = 0$, x must be $1/2$; i.e., each hyperboloid "starts" at $x = 1/2$. This matches our figure.

We conclude that $4x^2 - y^2 - \dfrac{z^2}{9} = 1$ best fits the graph.

This section has introduced points in space and shown how equations can describe surfaces. The next sections explore *vectors*, an important mathematical object that we'll use to explore curves in space.

12.1.7 Exercises

Terms and Concepts

1. Axes drawn in space must conform to the _____ _____ rule.

2. In the plane, the equation $x = 2$ defines a _____;
in space, $x = 2$ defines a _____.

3. In the plane, the equation $y = x^2$ defines a _____; in space, $y = x^2$ defines a _____.

4. Which quadric surface looks like a Pringles(TM) chip? (☐ Elliptic Paraboloid ☐ Elliptic Cone ☐ Ellipsoid ☐ Hyperboloid of One Sheet ☐ Hyperboloid of Two Sheets ☐ Hyperbolic Paraboloid)

5. Consider the hyperbola $x^2 - y^2 = 1$ in the plane. If this hyperbola is rotated about the x-axis, what quadric surface is formed?

6. Consider the hyperbola $x^2 - y^2 = 1$ in the plane. If this hyperbola is rotated about the y-axis, what quadric surface is formed?

Problems

7. The points $A = (1, 4, 2)$, $B = (2, 6, 3)$ and $C = (4, 3, 1)$ form a triangle in space. Find the distances between each pair of points and determine if the triangle is a right triangle.
$\|\overline{AB}\| =$_____
$\|\overline{BC}\| =$_____
$\|\overline{CA}\| =$_____
The three points (☐ do ☐ do not) form a right triangle.

8. The points $A = (1, 1, 3)$, $B = (3, 2, 7)$, $C = (2, 0, 8)$ and $D = (0, -1, 4)$ form a quadrilateral $ABCD$ in space. Is this a parallelogram?

9. Find the center and radius of the sphere defined by

$$x^2 - 8x + y^2 + 2y + z^2 + 8 = 0 :$$

Center at _____.
Radius of _____.

10. Find the center and radius of the sphere defined by

$$x^2 + y^2 + z^2 + 4x - 2y - 4z + 4 = 0 :$$

Center at _____.
Radius of _____.

Exercise Group. In the following exercises, describe the region in space defined by the inequalities.

11. $x^2 + y^2 + z^2 < 1$

12. $0 \le x \le 3$

13. $x \ge 0,\ y \ge 0,\ z \ge 0$

14. $y \ge 3$

Exercise Group. In the following exercises, sketch the cylinder in space.

15. $z = x^3$

16. $y = \cos(z)$

17. $\dfrac{x^2}{4} + \dfrac{y^2}{9} = 1$

18. $y = \dfrac{1}{x}$

Exercise Group. In the following exercises, give the equation of the surface of revolution described.

19. Give the equation of the surface formed by revolving $z = \frac{1}{1+y^2}$ in the yz-plane about the y-axis.

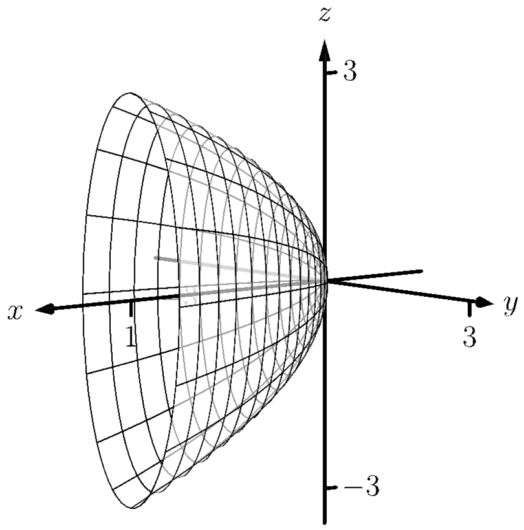
20. Give the equation of the surface formed by revolving $y = x^2$ in the xy-plane about the x-axis.

21. Give the equation of the surface formed by revolving $z = x^2$ in the xz-plane about the z-axis.

22. Give the equation of the surface formed by revolving $z = 1/x$ in the xz-plane about the z-axis.

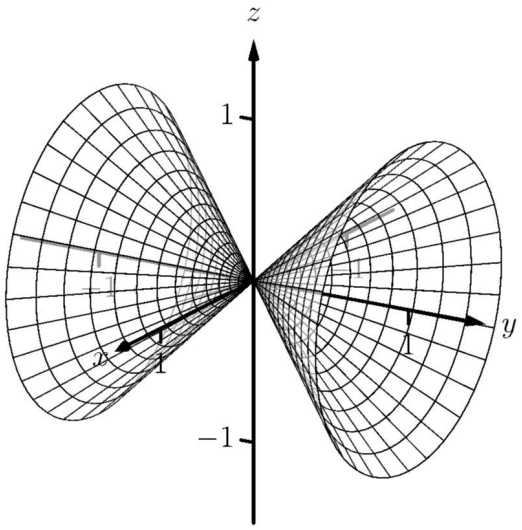
Exercise Group. In the following exercises, a quadric surface is sketched. Determine which of the given equations best fits the graph.

23.



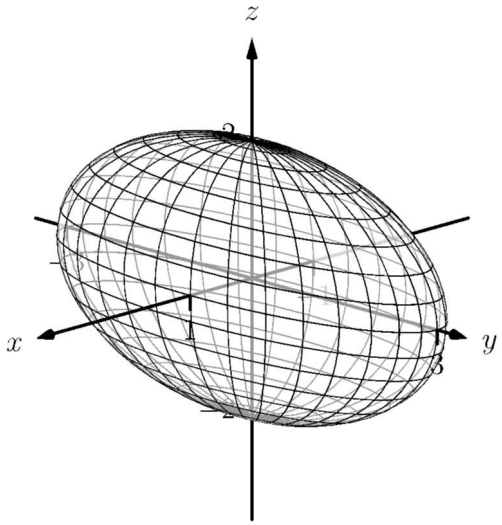
(a) $x = y^2 + \dfrac{z^2}{9}$ (b) $x = y^2 + \dfrac{z^2}{3}$

24.



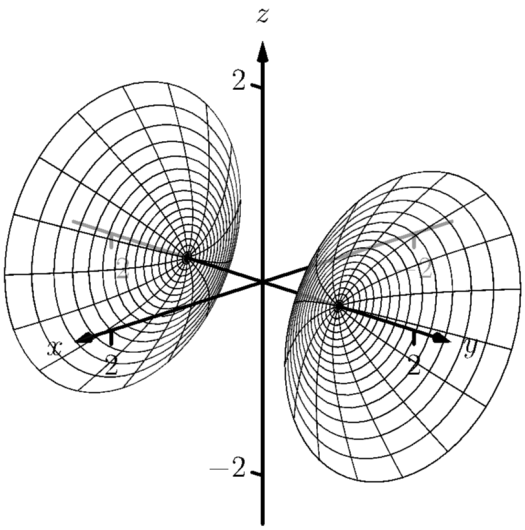
(a) $x^2 - y^2 - z^2 = 0$ (b) $x^2 - y^2 + z^2 = 0$

25.



(a) $x^2 + \dfrac{y^2}{3} + \dfrac{z^2}{2} = 1$ (b) $x^2 + \dfrac{y^2}{9} + \dfrac{z^2}{4} = 1$

26.



(a) $y^2 - x^2 - z^2 = 1$ (b) $y^2 + x^2 - z^2 = 1$

Exercise Group. In the following exercises, sketch the quadric surface.

27. $z - y^2 + x^2 = 0$

28. $z^2 = x^2 + \dfrac{y^2}{4}$

29. $x = -y^2 - z^2$

30. $16x^2 - 16y^2 - 16z^2 = 1$

31. $\dfrac{x^2}{9} - y^2 + \dfrac{z^2}{25} = 1$

32. $4x^2 + 2y^2 + z^2 = 4$

12.2 An Introduction to Vectors

Many quantities we think about daily can be described by a single number: temperature, speed, cost, weight and height. There are also many other concepts we encounter daily that cannot be described with just one number. For instance, a weather forecaster often describes wind with its speed and its direction ("... with winds from the southeast gusting up to 30 mph ..."). When applying a force, we are concerned with both the magnitude and direction of that force. In both of these examples, *direction* is important. Because of this, we study *vectors*, mathematical objects that convey both magnitude and direction information.

One "bare-bones" definition of a vector is based on what we wrote above: "a vector is a mathematical object with magnitude and direction parameters." This definition leaves much to be desired, as it gives no indication as to how such an object is to be used. Several other definitions exist; we choose here a definition rooted in a geometric visualization of vectors. It is very simplistic but readily permits further investigation.

youtu.be/watch?v=_1mYILsTF2U

Figure 12.2.1 Video introduction to Section 12.2

Definition 12.2.2 Vector.

A **vector** is a directed line segment.

Given points P and Q (either in the plane or in space), we denote with $\overrightarrow{PQ}$ the vector **from** P **to** Q. The point P is said to be the **initial point** of the vector, and the point Q is the **terminal point**.

The **magnitude**, **length** or **norm** of $\overrightarrow{PQ}$ is the length of the line segment $\overline{PQ}$: $\left\|\overrightarrow{PQ}\right\| = \|\overline{PQ}\|$.

Two vectors are **equal** if they have the same magnitude and direction.

Figure 12.2.3 shows multiple instances of the same vector. Each directed line segment has the same direction and length (magnitude), hence each is the same vector.

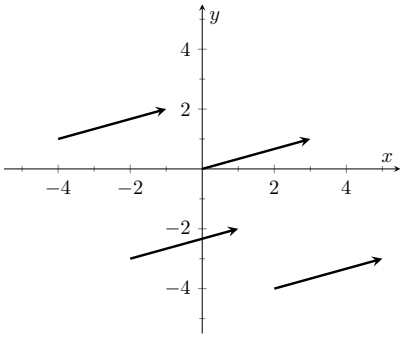


Figure 12.2.3 Drawing the same vector with different initial points

We use $\mathbb{R}^2$ (pronounced "r two") to represent all the vectors in the plane, and use $\mathbb{R}^3$ (pronounced "r three") to represent all the vectors in space.

Consider the vectors $\overrightarrow{PQ}$ and $\overrightarrow{RS}$ as shown in Figure 12.2.4. The vectors look to be equal; that is, they seem to have the same length and direction. Indeed, they are. Both vectors move 2 units to the right and 1 unit up from the initial point to reach the terminal point. One can analyze this movement to measure the magnitude of the vector, and the movement itself gives direction information (one could also measure the slope of the line passing through P and Q or R and S). Since they have the same length and direction, these two vectors are equal.

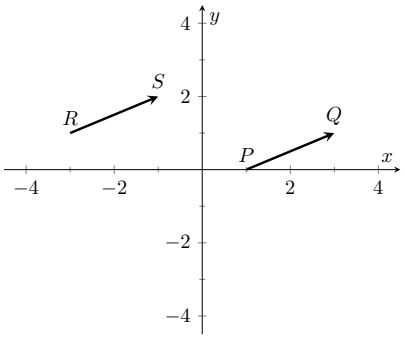


Figure 12.2.4 Illustrating how equal vectors have the same displacement

This demonstrates that inherently all we care about is *displacement*; that is, how far in the x, y and possibly z directions the terminal point is from the initial point. Both the vectors $\overrightarrow{PQ}$ and $\overrightarrow{RS}$ in Figure 12.2.4 have an x-displacement of 2 and a y-displacement of 1. This suggests a standard way of describing vectors in the plane. A vector whose x-displacement is a and whose y-displacement is b will have terminal point (a, b) when the initial point is the origin, $(0, 0)$. This leads us to a definition of a standard and concise way of referring to vectors.

Definition 12.2.5 Component Form of a Vector.

1. The *component form* of a vector $\vec{v}$ in $\mathbb{R}^2$, whose terminal point is (a, b) when its initial point is $(0, 0)$, is $\langle a, b\rangle$.
2. The *component form* of a vector $\vec{v}$ in $\mathbb{R}^3$, whose terminal point is (a, b, c) when its initial point is $(0, 0, 0)$, is $\langle a, b, c\rangle$.

The numbers a, b (and c, respectively) are the **components** of $\vec{v}$.

It follows from the definition that the component form of the vector $\overrightarrow{PQ}$, where $P = (x_1, y_1)$ and $Q = (x_2, y_2)$ is

$$\overrightarrow{PQ} = \langle x_2 - x_1, y_2 - y_1\rangle ;$$

in space, where $P = (x_1, y_1, z_1)$ and $Q = (x_2, y_2, z_2)$, the component form of $\overrightarrow{PQ}$ is

$$\overrightarrow{PQ} = \langle x_2 - x_1, y_2 - y_1, z_2 - z_1\rangle .$$

We practice using this notation in the following example.

Example 12.2.6 Using component form notation for vectors.

1. Sketch the vector $\vec{v} = \langle 2, -1\rangle$ starting at $P = (3, 2)$ and find its magnitude.
2. Find the component form of the vector $\vec{w}$ whose initial point is $R = (-3, -2)$ and whose terminal point is $S = (-1, 2)$.
3. Sketch the vector $\vec{u} = \langle 2, -1, 3\rangle$ starting at the point $Q = (1, 1, 1)$ and find its magnitude.

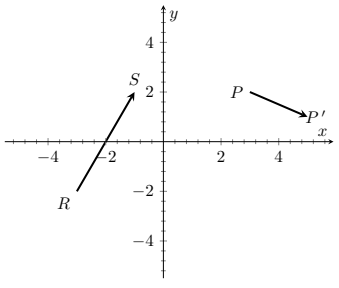
Video Solution

youtu.be/watch?v=w1MzWJzoyGg

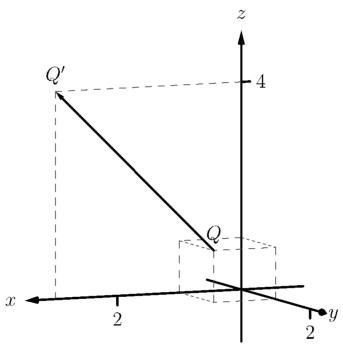
Solution.

1. Using P as the initial point, we move 2 units in the positive x-direction and -1 units in the positive y-direction to arrive at the terminal point $P' = (5, 1)$, as drawn in Figure 12.2.7(a). The magnitude of $\vec{v}$ is determined directly from the component form:

$$\|\vec{v}\| = \sqrt{2^2 + (-1)^2} = \sqrt{5}.$$



(a)



(b)

Figure 12.2.7 Graphing vectors in Example 12.2.6

2. Using the note following Definition 12.2.5, we have

$$\overrightarrow{RS} = \langle -1-(-3), 2-(-2)\rangle = \langle 2, 4\rangle .$$

One can readily see from Figure 12.2.7(a) that the x- and y-displacement of $\overrightarrow{RS}$ is 2 and 4, respectively, as the component form suggests.

3. Using Q as the initial point, we move 2 units in the positive x-direction, -1 unit in the positive y-direction, and 3 units in the positive z-direction to arrive at the terminal point $Q' = (3, 0, 4)$, illustrated in Figure 12.2.7(b). The magnitude of $\vec{u}$ is:

$$\|\vec{u}\| = \sqrt{2^2 + (-1)^2 + 3^2} = \sqrt{14}.$$

Now that we have defined vectors, and have created a nice notation by which to describe them, we start considering how vectors interact with each other. That is, we define an *algebra* on vectors.

Definition 12.2.8 Vector Algebra.

1. Let $\vec{u} = \langle u_1, u_2\rangle$ and $\vec{v} = \langle v_1, v_2\rangle$ be vectors in $\mathbb{R}^2$, and let c be a scalar.
   (a) The addition, or sum, of the vectors $\vec{u}$ and $\vec{v}$ is the vector
   $$\vec{u} + \vec{v} = \langle u_1 + v_1, u_2 + v_2\rangle .$$
   (b) The multiplication of a scalar c and a vector $\vec{v}$ is the vector
   $$c\vec{v} = c\langle v_1, v_2\rangle = \langle cv_1, cv_2\rangle .$$
2. Let $\vec{u} = \langle u_1, u_2, u_3\rangle$ and $\vec{v} = \langle v_1, v_2, v_3\rangle$ be vectors in $\mathbb{R}^3$, and let c be a scalar.
   (a) The addition, or sum, of the vectors $\vec{u}$ and $\vec{v}$ is the vector
   $$\vec{u} + \vec{v} = \langle u_1 + v_1, u_2 + v_2, u_3 + v_3\rangle .$$
   (b) The multiplication of a scalar c and a vector $\vec{v}$ is the vector
   $$c\vec{v} = c\langle v_1, v_2, v_3\rangle = \langle cv_1, cv_2, cv_3\rangle .$$

In short, we say addition and scalar multiplication are computed "componentwise."

youtu.be/watch?v=tOpeq-_K0ME

Figure 12.2.9 Video presentation of Definition 12.2.8 (2 videos)

Video Solution

youtu.be/watch?v=3MSwQJq_3s0

Example 12.2.10 Adding vectors.

Sketch the vectors $\vec{u} = \langle 1, 3\rangle$, $\vec{v} = \langle 2, 1\rangle$ and $\vec{u} + \vec{v}$ all with initial point at the origin.

Solution. We first compute $\vec{u} + \vec{v}$.

$$\begin{aligned}\vec{u} + \vec{v} &= \langle 1, 3\rangle + \langle 2, 1\rangle \\ &= \langle 3, 4\rangle .\end{aligned}$$

These are all sketched in Figure 12.2.11.

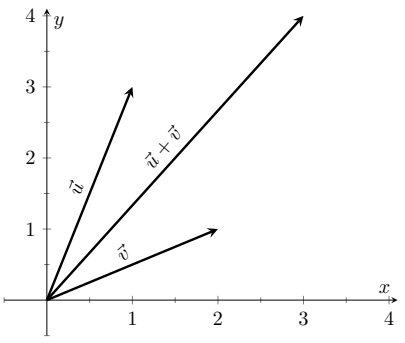


Figure 12.2.11 Graphing the sum of vectors in Example 12.2.10

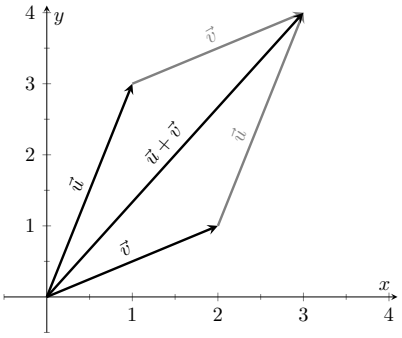


Figure 12.2.12 Illustrating how to add vectors using the Head to Tail Rule and Parallelogram Law

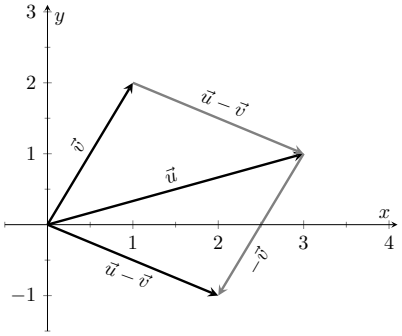


Figure 12.2.14 Illustrating how to subtract vectors graphically

Video Solution

youtu.be/watch?v=a_qRwTkQXYQ

As vectors convey magnitude and direction information, the sum of vectors also convey length and magnitude information. Adding $\vec{u} + \vec{v}$ suggests the following idea:

> "Starting at an initial point, go out $\vec{u}$, then go out $\vec{v}$."

This idea is sketched in Figure 12.2.12, where the initial point of $\vec{v}$ is the terminal point of $\vec{u}$. This is known as the "Head to Tail Rule" of adding vectors. Vector addition is very important. For instance, if the vectors $\vec{u}$ and $\vec{v}$ represent forces acting on a body, the sum $\vec{u} + \vec{v}$ gives the resulting force. Because of various physical applications of vector addition, the sum $\vec{u}+\vec{v}$ is often referred to as the *resultant vector*, or just the "resultant."

Analytically, it is easy to see that $\vec{u} + \vec{v} = \vec{v} + \vec{u}$. Figure 12.2.12 also gives a graphical representation of this, using gray vectors. Note that the vectors $\vec{u}$ and $\vec{v}$, when arranged as in the figure, form a parallelogram. Because of this, the Head to Tail Rule is also known as the Parallelogram Law: the vector $\vec{u}+\vec{v}$ is defined by forming the parallelogram defined by the vectors $\vec{u}$ and $\vec{v}$; the initial point of $\vec{u} + \vec{v}$ is the common initial point of parallelogram, and the terminal point of the sum is the common terminal point of the parallelogram.

While not illustrated here, the Head to Tail Rule and Parallelogram Law hold for vectors in $\mathbb{R}^3$ as well.

It follows from the properties of the real numbers and Definition 12.2.8 that

$$\vec{u} - \vec{v} = \vec{u} + (-1)\vec{v}.$$

The Parallelogram Law gives us a good way to visualize this subtraction. We demonstrate this in the following example.

Example 12.2.13 Vector Subtraction.

Let $\vec{u} = \langle 3, 1\rangle$ and $\vec{v} = \langle 1, 2\rangle$. Compute and sketch $\vec{u} - \vec{v}$.

Solution. The computation of $\vec{u}-\vec{v}$ is straightforward, and we show all steps below. Usually the formal step of multiplying by (-1) is omitted and we "just subtract."

$$\begin{aligned}\vec{u} - \vec{v} &= \vec{u} + (-1)\vec{v}\\ &= \langle 3, 1\rangle + \langle -1, -2\rangle\\ &= \langle 2, -1\rangle\,.\end{aligned}$$

Figure 12.2.14 illustrates, using the Head to Tail Rule, how the subtraction can be viewed as the sum $\vec{u} + (-\vec{v})$. The figure also illustrates how $\vec{u} - \vec{v}$ can be obtained by looking only at the terminal points of $\vec{u}$ and $\vec{v}$ (when their initial points are the same).

Example 12.2.15 Scaling vectors.

1. Sketch the vectors $\vec{v} = \langle 2, 1\rangle$ and $2\vec{v}$ with initial point at the origin.
2. Compute the magnitudes of $\vec{v}$ and $2\vec{v}$.

Solution.

1. We compute $2\vec{v}$:

$$\begin{aligned} 2\vec{v} &= 2\langle 2, 1\rangle \\ &= \langle 4, 2\rangle . \end{aligned}$$

   Both $\vec{v}$ and $2\vec{v}$ are sketched in Figure 12.2.16. Make note that $2\vec{v}$ does not start at the terminal point of $\vec{v}$; rather, its initial point is also the origin.

2. The figure suggests that $2\vec{v}$ is twice as long as $\vec{v}$. We compute their magnitudes to confirm this.

$$\begin{aligned} \|\vec{v}\| &= \sqrt{2^2 + 1^2} \\ &= \sqrt{5}. \\ \|2\vec{v}\| &= \sqrt{4^2 + 2^2} \\ &= \sqrt{20} \\ &= \sqrt{4 \cdot 5} = 2\sqrt{5}. \end{aligned}$$

   As we suspected, $2\vec{v}$ is twice as long as $\vec{v}$.

Video Solution

youtu.be/watch?v=-zvqs7hveXc

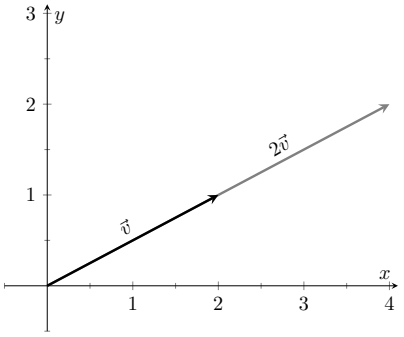


Figure 12.2.16 Graphing vectors $\vec{v}$ and $2\vec{v}$ in Example 12.2.15

The *zero vector* is the vector whose initial point is also its terminal point. It is denoted by $\vec{0}$. Its component form, in $\mathbb{R}^2$, is $\langle 0, 0\rangle$; in $\mathbb{R}^3$, it is $\langle 0, 0, 0\rangle$. Usually the context makes is clear whether $\vec{0}$ is referring to a vector in the plane or in space.

Our examples have illustrated key principles in vector algebra: how to add and subtract vectors and how to multiply vectors by a scalar. The following theorem states formally the properties of these operations.

Theorem 12.2.17 Properties of Vector Operations.

The following are true for all scalars c and d, and for all vectors $\vec{u}$, $\vec{v}$ and $\vec{w}$, where $\vec{u}$, $\vec{v}$ and $\vec{w}$ are all in $\mathbb{R}^2$ or where $\vec{u}$, $\vec{v}$ and $\vec{w}$ are all in $\mathbb{R}^3$:

1. $\vec{u} + \vec{v} = \vec{v} + \vec{u}$ *Commutative Property*
2. $(\vec{u} + \vec{v}) + \vec{w} = \vec{u} + (\vec{v} + \vec{w})$ *Associative Property*
3. $\vec{v} + \vec{0} = \vec{v}$ *Additive Identity*
4. $(cd)\vec{v} = c(d\vec{v})$
5. $c(\vec{u} + \vec{v}) = c\vec{u} + c\vec{v}$ *Distributive Property*
6. $(c + d)\vec{v} = c\vec{v} + d\vec{v}$ *Distributive Property*
7. $0\vec{v} = \vec{0}$

youtu.be/watch?v=oixzQ2EGYPM

Figure 12.2.18 Video presentation of Part of Theorem 12.2.17

youtu.be/watch?v=6kKUB3OSurs

Figure 12.2.20 Video presentation of Definition 12.2.19

Video Solution

youtu.be/watch?v=B9OV0Dja6IY

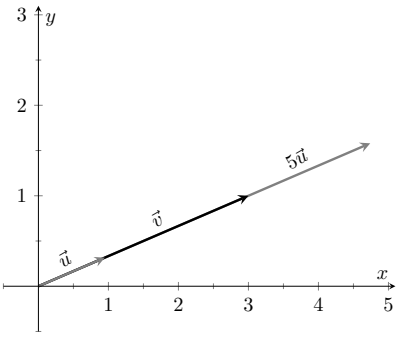


Figure 12.2.22 Graphing vectors in Example 12.2.21. All vectors shown have their initial point at the origin

8. $\|c\vec{v}\| = |c| \cdot \|\vec{v}\|$
9. $\|\vec{u}\| = 0$ *if, and only if,* $\vec{u} = \vec{0}$.

As stated before, each nonzero vector $\vec{v}$ conveys magnitude and direction information. We have a method of extracting the magnitude, which we write as $\|\vec{v}\|$. *Unit vectors* are a way of extracting just the direction information from a vector.

Definition 12.2.19 Unit Vector.

A **unit vector** is a vector $\vec{v}$ with a magnitude of 1; that is,

$$\|\vec{v}\| = 1.$$

Consider this scenario: you are given a vector $\vec{v}$ and are told to create a vector of length 10 in the direction of $\vec{v}$. How does one do that? If we knew that $\vec{u}$ was the unit vector in the direction of $\vec{v}$, the answer would be easy: $10\vec{u}$. So how do we find $\vec{u}$?

Property 8 of Theorem 12.2.17 holds the key. If we divide $\vec{v}$ by its magnitude, it becomes a vector of length 1. Consider:

$$\left\| \frac{1}{\|\vec{v}\|}\vec{v} \right\| = \frac{1}{\|\vec{v}\|}\|\vec{v}\| \qquad \text{(we can pull out } \frac{1}{\|\vec{v}\|} \text{ as it is a positive scalar)}$$
$$= 1.$$

So the vector of length 10 in the direction of $\vec{v}$ is $10\dfrac{1}{\|\vec{v}\|}\vec{v}$. An example will make this more clear.

Example 12.2.21 Using Unit Vectors.

Let $\vec{v} = \langle 3, 1\rangle$ and let $\vec{w} = \langle 1, 2, 2\rangle$.

1. Find the unit vector in the direction of $\vec{v}$.
2. Find the unit vector in the direction of $\vec{w}$.
3. Find the vector in the direction of $\vec{v}$ with magnitude 5.

Solution.

1. We find $\|\vec{v}\| = \sqrt{10}$. So the unit vector $\vec{u}$ in the direction of $\vec{v}$ is
$$\vec{u} = \frac{1}{\sqrt{10}}\vec{v} = \left\langle \frac{3}{\sqrt{10}}, \frac{1}{\sqrt{10}} \right\rangle.$$
2. We find $\|\vec{w}\| = 3$, so the unit vector $\vec{z}$ in the direction of $\vec{w}$ is
$$\vec{u} = \frac{1}{3}\vec{w} = \left\langle \frac{1}{3}, \frac{2}{3}, \frac{2}{3} \right\rangle.$$
3. To create a vector with magnitude 5 in the direction of $\vec{v}$, we multiply the unit vector $\vec{u}$ by 5. Thus $5\vec{u} = \langle 15/\sqrt{10}, 5/\sqrt{10}\rangle$ is the vector we seek. This is sketched in Figure 12.2.22.

The basic formation of the unit vector $\vec{u}$ in the direction of a vector $\vec{v}$ leads to a interesting equation. It is:

$$\vec{v} = \|\vec{v}\| \frac{1}{\|\vec{v}\|}\vec{v}.$$

We rewrite the equation with parentheses to make a point:

$$\vec{v} = \underbrace{\|\vec{v}\|}_{\text{magnitude}} \cdot \underbrace{\left(\frac{1}{\|\vec{v}\|}\vec{v}\right)}_{\text{direction}}.$$

This equation illustrates the fact that a nonzero vector has both magnitude and direction, where we view a unit vector as supplying *only* direction information. Identifying unit vectors with direction allows us to define *parallel vectors*.

Definition 12.2.23 Parallel Vectors.

1. Unit vectors $\vec{u}_1$ and $\vec{u}_2$ are *parallel* if $\vec{u}_1 = \pm\vec{u}_2$.
2. Nonzero vectors $\vec{v}_1$ and $\vec{v}_2$ are *parallel* if their respective unit vectors are parallel.

It is equivalent to say that vectors $\vec{v}_1$ and $\vec{v}_2$ are parallel if there is a scalar $c \neq 0$ such that $\vec{v}_1 = c\vec{v}_2$ (see marginal note).

If one graphed all unit vectors in $\mathbb{R}^2$ with the initial point at the origin, then the terminal points would all lie on the unit circle. Based on what we know from trigonometry, we can then say that the component form of all unit vectors in $\mathbb{R}^2$ is $\langle \cos(\theta), \sin(\theta)\rangle$ for some angle θ.

A similar construction in $\mathbb{R}^3$ shows that the terminal points all lie on the unit sphere. These vectors also have a particular component form, but its derivation is not as straightforward as the one for unit vectors in $\mathbb{R}^2$. Important concepts about unit vectors are given in the following Key Idea.

Key Idea 12.2.25 Unit Vectors.

1. The unit vector in the direction of a nonzero vector $\vec{v}$ is
$$\vec{u} = \frac{1}{\|\vec{v}\|}\vec{v}.$$
2. A vector $\vec{u}$ in $\mathbb{R}^2$ is a unit vector if, and only if, its component form is $\langle \cos\theta, \sin\theta\rangle$ for some angle θ.
3. A vector $\vec{u}$ in $\mathbb{R}^3$ is a unit vector if, and only if, its component form is $\langle \sin(\theta)\cos(\varphi), \sin(\theta)\sin(\varphi), \cos(\theta)\rangle$ for some angles θ and φ.

These formulas can come in handy in a variety of situations, especially the formula for unit vectors in the plane.

Example 12.2.26 Finding Component Forces.

Consider a weight of 50lb hanging from two chains, as shown in Figure 12.2.27. One chain makes an angle of 30° with the vertical, and the other an angle of 45°. Find the force applied to each chain.

Solution. Knowing that gravity is pulling the 50lb weight straight down,

Direction and the zero vector. $\vec{0}$ is directionless; because $\left\|\vec{0}\right\| = 0$, there is no unit vector in the "direction" of $\vec{0}$.

Some texts define two vectors as being parallel if one is a scalar multiple of the other. By this definition, $\vec{0}$ is parallel to all vectors as $\vec{0} = 0\vec{v}$ for all $\vec{v}$.

We define what it means for two vectors to be perpendicular in Definition 12.3.14, which is written to exclude $\vec{0}$. It could be written to include $\vec{0}$; by such a definition, $\vec{0}$ is perpendicular to all vectors. While counter-intuitive, it is mathematically sound to allow $\vec{0}$ to be both parallel and perpendicular to all vectors.

We prefer the given definition of parallel as it is grounded in the fact that unit vectors provide direction information. One may adopt the convention that $\vec{0}$ is parallel to all vectors if they desire. (See also the aside in Section 12.4.)

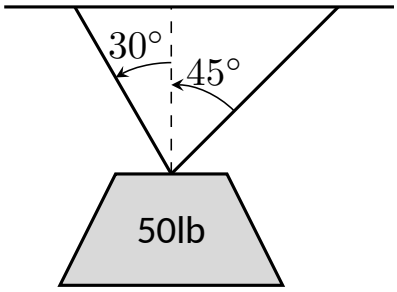


Figure 12.2.27 A diagram of a weight hanging from 2 chains in Example 12.2.26

Video Solution

youtu.be/watch?v=XxSQSU4Nvig

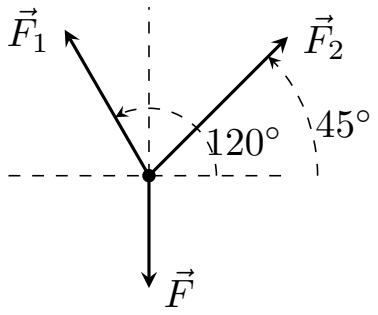


Figure 12.2.28 A diagram of the force vectors from Example 12.2.26

we can create a vector $\vec{F}$ to represent this force.

$$\vec{F} = 50\langle 0, -1\rangle = \langle 0, -50\rangle .$$

We can view each chain as "pulling" the weight up, preventing it from falling. We can represent the force from each chain with a vector. Let $\vec{F}_1$ represent the force from the chain making an angle of $30°$ with the vertical, and let $\vec{F}_2$ represent the force form the other chain. Convert all angles to be measured from the horizontal (as shown in Figure 12.2.28), and apply Key Idea 12.2.25. As we do not yet know the magnitudes of these vectors, (that is the problem at hand), we use m_1 and m_2 to represent them.

$$\vec{F}_1 = m_1\langle \cos(120°), \sin(120°)\rangle$$
$$\vec{F}_2 = m_2\langle \cos(45°), \sin(45°)\rangle$$

As the weight is not moving, we know the sum of the forces is $\vec{0}$. This gives:

$$\vec{F} + \vec{F}_1 + \vec{F}_2 = \vec{0}$$
$$\langle 0, -50\rangle + m_1\langle \cos(120°), \sin(120°)\rangle + m_2\langle \cos(45°), \sin(45°)\rangle = \vec{0}$$

The sum of the entries in the first component is 0, and the sum of the entries in the second component is also 0. This leads us to the following two equations:

$$m_1 \cos(120°) + m_2 \cos(45°) = 0$$
$$m_1 \sin(120°) + m_2 \sin(45°) = 50$$

This is a simple 2-equation, 2-unknown system of linear equations. We leave it to the reader to verify that the solution is

$$m_1 = 50(\sqrt{3} - 1) \approx 36.6; \qquad m_2 = \frac{50\sqrt{2}}{1 + \sqrt{3}} \approx 25.88.$$

It might seem odd that the sum of the forces applied to the chains is more than 50lb. We leave it to a physics class to discuss the full details, but offer this short explanation. Our equations were established so that the *vertical* components of each force sums to 50lb, thus supporting the weight. Since the chains are at an angle, they also pull against each other, creating an "additional" horizontal force while holding the weight in place.

Unit vectors were very important in the previous calculation; they allowed us to define a vector in the proper direction but with an unknown magnitude. Our computations were then computed component-wise. Because such calculations are often necessary, the *standard unit vectors* can be useful.

Definition 12.2.29 Standard Unit Vectors.

1. In $\mathbb{R}^2$, the standard unit vectors are

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

2. In $\mathbb{R}^3$, the standard unit vectors are

$$\vec{i} = \langle 1, 0, 0 \rangle \text{ and } \vec{j} = \langle 0, 1, 0 \rangle \text{ and } \vec{k} = \langle 0, 0, 1 \rangle .$$

Example 12.2.30 Using standard unit vectors.

1. Rewrite $\vec{v} = \langle 2, -3 \rangle$ using the standard unit vectors.
2. Rewrite $\vec{w} = 4\vec{i} - 5\vec{j} + 2\vec{k}$ in component form.

Solution.

1.
$$\begin{aligned} \vec{v} &= \langle 2, -3 \rangle \\ &= \langle 2, 0 \rangle + \langle 0, -3 \rangle \\ &= 2\langle 1, 0 \rangle - 3\langle 0, 1 \rangle \\ &= 2\vec{i} - 3\vec{j} \end{aligned}$$

2.
$$\begin{aligned} \vec{w} &= 4\vec{i} - 5\vec{j} + 2\vec{k} \\ &= \langle 4, 0, 0 \rangle + \langle 0, -5, 0 \rangle + \langle 0, 0, 2 \rangle \\ &= \langle 4, -5, 2 \rangle \end{aligned}$$

These two examples demonstrate that converting between component form and the standard unit vectors is rather straightforward. Many mathematicians prefer component form, and it is the preferred notation in this text. Many engineers prefer using the standard unit vectors, and many engineering text use that notation.

Example 12.2.31 Finding Component Force.

A weight of 25lb is suspended from a chain of length 2ft while a wind pushes the weight to the right with constant force of 5lb as shown in Figure 12.2.32. What angle will the chain make with the vertical as a result of the wind's pushing? How much higher will the weight be?

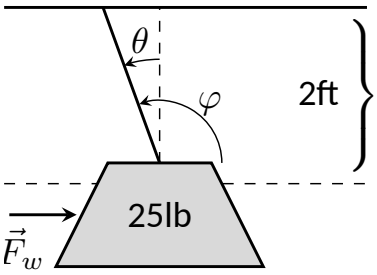


Figure 12.2.32 A figure of a weight being pushed by the wind in Example 12.2.31

Solution. The force of the wind is represented by the vector $\vec{F}_w = 5\vec{i}$. The force of gravity on the weight is represented by $\vec{F}_g = -25\vec{j}$. The direction and magnitude of the vector representing the force on the chain are both unknown. We represent this force with

$$\vec{F}_c = m\langle \cos(\varphi), \sin(\varphi) \rangle = m\cos(\varphi)\,\vec{i} + m\sin(\varphi)\,\vec{j}$$

for some magnitude m and some angle with the horizontal φ. (Note: θ is the angle the chain makes with the *vertical*; φ is the angle with the *horizontal*.)

As the weight is at equilibrium, the sum of the forces is $\vec{0}$:

$$\begin{aligned} \vec{F}_c + \vec{F}_w + \vec{F}_g &= \vec{0} \\ m\cos(\varphi)\,\vec{i} + m\sin(\varphi)\,\vec{j} + 5\vec{i} - 25\vec{j} &= \vec{0} \end{aligned}$$

Thus the sum of the $\vec{i}$ and $\vec{j}$ components are 0, leading us to the

following system of equations:

$$\begin{aligned} 5 + m\cos\varphi &= 0 \\ -25 + m\sin\varphi &= 0 \end{aligned} \tag{12.2.1}$$

This is enough to determine $\vec{F_c}$ already, as we know $m\cos(\varphi) = -5$ and $m\sin(\varphi) = 25$. Thus $F_c = \langle -5, 25\rangle$. We can use this to find the magnitude m:

$$m = \sqrt{(-5)^2 + 25^2} = 5\sqrt{26} \approx 25.5 \text{ lb}.$$

We can then use either equality from Equation (12.2.1) to solve for φ. We choose the first equality as using arccosine will return an angle in the 2nd quadrant:

$$5 + 5\sqrt{26}\cos(\varphi) = 0 \Rightarrow \varphi = \cos^{-1}\left(\frac{-5}{5\sqrt{26}}\right) \approx 1.7682 \approx 101.31^\circ.$$

Subtracting 90° from this angle gives us an angle of 11.31° with the vertical.

We can now use trigonometry to find out how high the weight is lifted. The diagram shows that a right triangle is formed with the 2ft chain as the hypotenuse with an interior angle of 11.31°. The length of the adjacent side (in the diagram, the dashed vertical line) is $2\cos(11.31^\circ) \approx 1.96$ft. Thus the weight is lifted by about 0.04ft, almost 1/2in.

The algebra we have applied to vectors is already demonstrating itself to be very useful. There are two more fundamental operations we can perform with vectors, the *dot product* and the *cross product*. The next two sections explore each in turn.

12.2.1 Exercises

Terms and Concepts

1. Name two different things that cannot be described with just one number, but rather need 2 or more numbers to fully describe them.

2. What is the difference between $(1, 2)$ and $\langle 1, 2\rangle$?

3. What is a unit vector?

4. Unit vectors can be thought of as conveying what type of information?

5. What does it mean for two vectors to be parallel?

6. What effect does multiplying a vector by -2 have?

Problems

Exercise Group. In the following exercises, points P and Q are given. Write the vector $\overrightarrow{PQ}$ in component form and using the standard unit vectors.

7. If $P = (2, -1)$ and $Q = (3, 5)$, write the vector $\overrightarrow{PQ}$:

(a) in component form.

(b) using the standard unit vectors.

8. If $P = (3, 2)$ and $Q = (7, -2)$, write the vector $\overrightarrow{PQ}$:

(a) in component form.

(b) using the standard unit vectors.

9. If $P = (0, 3, -1)$ and $Q = (6, 2, 5)$, write the vector $\overrightarrow{PQ}$:

(a) in component form.

(b) using the standard unit vectors.

10. If $P = (2, 1, 2)$ and $Q = (4, 3, 2)$, write the vector $\overrightarrow{PQ}$:

(a) in component form.

(b) using the standard unit vectors.

11. Let $\vec{u} = \langle 1, -2\rangle$ and $\vec{v} = \langle 1, 1\rangle$.

(a) Find $\vec{u} + \vec{v}$, $\vec{u} - \vec{v}$, $2\vec{u} - 3\vec{v}$.

(b) Sketch the above vectors on the same axes, along with $\vec{u}$ and $\vec{v}$.

(c) Find $\vec{x}$ where $\vec{u} + \vec{x} = 2\vec{v} - \vec{x}$.

12. Let $\vec{u} = \langle 1, 1, -1\rangle$ and $\vec{v} = \langle 2, 1, 2\rangle$.

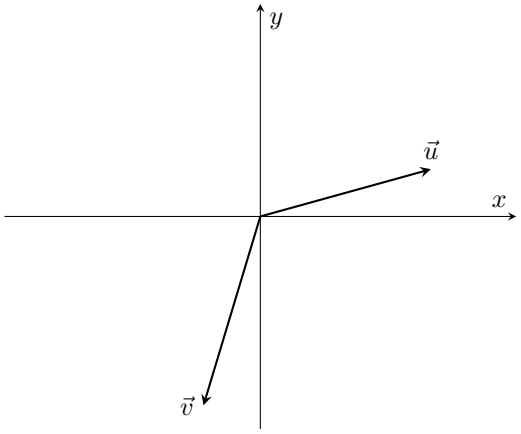
(a) Find $\vec{u} + \vec{v}$, $\vec{u} - \vec{v}$, $\pi\vec{u} - \sqrt{2}\vec{v}$.

(b) Sketch the above vectors on the same axes, along with $\vec{u}$ and $\vec{v}$.

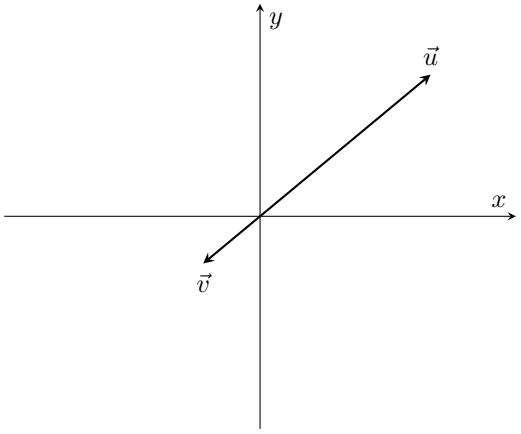
(c) Find $\vec{x}$ where $\vec{u} + \vec{x} = \vec{v} + 2\vec{x}$.

Exercise Group. In the following exercises, sketch $\vec{u}$, $\vec{v}$, $\vec{u} + \vec{v}$ and $\vec{u} - \vec{v}$ on the same axes.

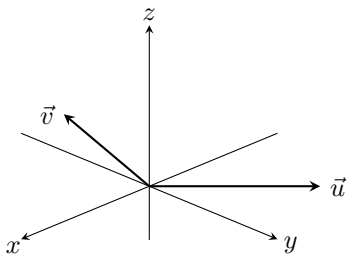
13.



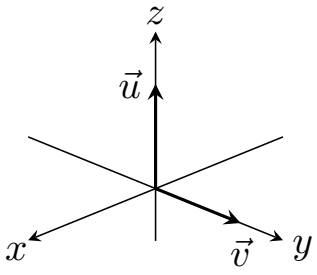
14.



15.



16.



Exercise Group. In the following exercises, find $\|\vec{u}\|$, $\|\vec{v}\|$, $\|\vec{u}+\vec{v}\|$ and $\|\vec{u}-\vec{v}\|$.

17. Given $\vec{u}=\langle 2,1\rangle$ and $\vec{v}=\langle 3,-2\rangle$, find:
$\|\vec{u}\|=$_____
$\|\vec{v}\|=$_____
$\|\vec{u}+\vec{v}\|=$_____
$\|\vec{u}-\vec{v}\|=$_____

18. Given $\vec{u}=\langle -3,2,2\rangle$ and $\vec{v}=\langle 1,-1,1\rangle$, find:
$\|\vec{u}\|=$_____
$\|\vec{v}\|=$_____
$\|\vec{u}+\vec{v}\|=$_____
$\|\vec{u}-\vec{v}\|=$_____

19. Given $\vec{u}=\langle 1,2\rangle$ and $\vec{v}=\langle -3,-6\rangle$, find:
$\|\vec{u}\|=$_____
$\|\vec{v}\|=$_____
$\|\vec{u}+\vec{v}\|=$_____
$\|\vec{u}-\vec{v}\|=$_____

20. Given $\vec{u}=\langle 2,-3,6\rangle$ and $\vec{v}=\langle 10,-15,30\rangle$, find:
$\|\vec{u}\|=$_____
$\|\vec{v}\|=$_____
$\|\vec{u}+\vec{v}\|=$_____
$\|\vec{u}-\vec{v}\|=$_____

21. Under what conditions is $\|\vec{u}\|+\|\vec{v}\|=\|\vec{u}+\vec{v}\|$?

Exercise Group. In the following exercises, find the unit vector $\vec{u}$ in the direction of $\vec{v}$.

22. Find the unit vector $\vec{u}$ in the direction of $\vec{v}=\langle 3,7\rangle$.
$\vec{u}=$_____

23. Find the unit vector $\vec{u}$ in the direction of $\vec{v}=\langle 6,8\rangle$.
$\vec{u}=$_____

24. Find the unit vector $\vec{u}$ in the direction of $\vec{v}=\langle 1,-2,2\rangle$.
$\vec{u}=$_____

25. Find the unit vector $\vec{u}$ in the direction of $\vec{v}=\langle 2,-2,2\rangle$.
$\vec{u}=$_____

26. Find the unit vector in the first quadrant of $\mathbb{R}^2$ that makes a 50° angle with the x-axis.

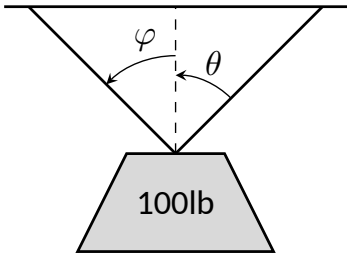
27. Find the unit vector in the second quadrant of $\mathbb{R}^2$ that makes a 30° angle with the y-axis.

28. Verify, from Key Idea 12.2.25, that

$$\vec{u} = \langle \sin(\theta)\cos(\varphi), \sin(\theta)\sin(\varphi), \cos(\theta)\rangle$$

is a unit vector for all angles θ and φ.

Exercise Group. A weight of 100lb is suspended from two chains, making angles with the vertical of θ and φ as shown in the figure below.



In the following exercises, the angles θ and φ are given. Find the magnitude of the force applied to each chain.

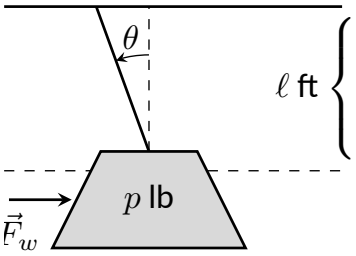
29. $\theta = 30°, \varphi = 30°$

30. $\theta = 60°, \varphi = 60°$

31. $\theta = 20°, \varphi = 15°$

32. $\theta = 0°, \varphi = 0°$

Exercise Group. A weight of $p$lb is suspended from a chain of length ℓ while a constant force of $\vec{F}_w$ pushes the weight to the right, making an angle of θ with the vertical, as shown in the figure below.



In the following exercises, a force $\vec{F}_w$ and length ℓ are given. Find the angle θ and the height the weight is lifted as it moves to the right.

33. $\vec{F}_w = 1$lb, $\ell = 1$ft, $p = 1$lb

34. $\vec{F}_w = 1$lb, $\ell = 1$ft, $p = 10$lb

35. $\vec{F}_w = 1$lb, $\ell = 10$ft, $p = 1$lb

36. $\vec{F}_w = 10$lb, $\ell = 10$ft, $p = 1$lb

youtu.be/watch?v=I996h2jxUhw

Figure 12.3.1 Video introduction to Section 12.3

12.3 The Dot Product

The previous section introduced vectors and described how to add them together and how to multiply them by scalars. This section introduces *a* multiplication on vectors called the *dot product*.

Definition 12.3.2 Dot Product.

1. Let $\vec{u} = \langle u_1, u_2 \rangle$ and $\vec{v} = \langle v_1, v_2 \rangle$ in $\mathbb{R}^2$. The *dot product* of $\vec{u}$ and $\vec{v}$, denoted $\vec{u} \cdot \vec{v}$, is
$$\vec{u} \cdot \vec{v} = u_1 v_1 + u_2 v_2.$$
2. Let $\vec{u} = \langle u_1, u_2, u_3 \rangle$ and $\vec{v} = \langle v_1, v_2, v_3 \rangle$ in $\mathbb{R}^3$. The *dot product* of $\vec{u}$ and $\vec{v}$, denoted $\vec{u} \cdot \vec{v}$, is
$$\vec{u} \cdot \vec{v} = u_1 v_1 + u_2 v_2 + u_3 v_3.$$

Note how this product of vectors returns a *scalar*, not another vector. We practice evaluating a dot product in the following example, then we will discuss why this product is useful.

Example 12.3.3 Evaluating dot products.

1. Let $\vec{u} = \langle 1, 2 \rangle$, $\vec{v} = \langle 3, -1 \rangle$ in $\mathbb{R}^2$. Find $\vec{u} \cdot \vec{v}$.
2. Let $\vec{x} = \langle 2, -2, 5 \rangle$ and $\vec{y} = \langle -1, 0, 3 \rangle$ in $\mathbb{R}^3$. Find $\vec{x} \cdot \vec{y}$.

Solution.

1. Using Definition 12.3.2, we have
$$\vec{u} \cdot \vec{v} = 1(3) + 2(-1) = 1.$$
2. Using the definition, we have
$$\vec{x} \cdot \vec{y} = 2(-1) - 2(0) + 5(3) = 13.$$

Video Solution

youtu.be/watch?v=NXZz_TGvFZ0

The dot product, as shown by the preceding example, is very simple to evaluate. It is only the sum of products. While the definition gives no hint as to why we would care about this operation, there is an amazing connection between the dot product and angles formed by the vectors. Before stating this connection, we give a theorem stating some of the properties of the dot product.

Theorem 12.3.4 Properties of the Dot Product.

Let $\vec{u}$, $\vec{v}$ and $\vec{w}$ be vectors in $\mathbb{R}^2$ or $\mathbb{R}^3$ and let c be a scalar.

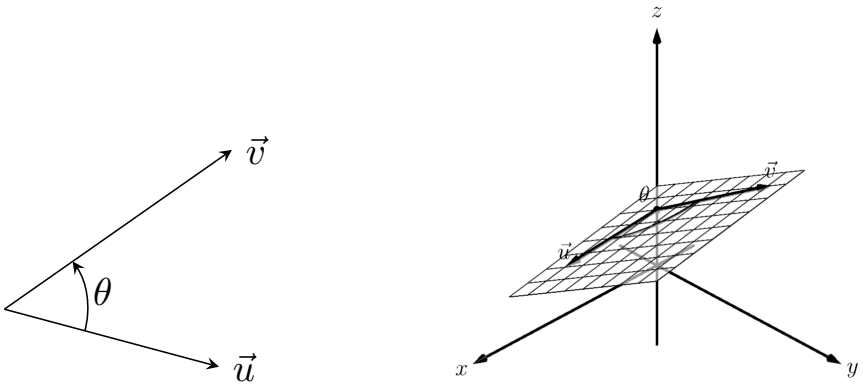
1. $\vec{u} \cdot \vec{v} = \vec{v} \cdot \vec{u}$ *{Commutative Property}*
2. $\vec{u} \cdot (\vec{v} + \vec{w}) = \vec{u} \cdot \vec{v} + \vec{u} \cdot \vec{w}$ *{Distributive Property}*
3. $c(\vec{u} \cdot \vec{v}) = (c\vec{u}) \cdot \vec{v} = \vec{u} \cdot (c\vec{v})$
4. $\vec{0} \cdot \vec{v} = 0$
5. $\vec{v} \cdot \vec{v} = \|\vec{v}\|^2$

The last statement of the theorem makes a handy connection between the magnitude of a vector and the dot product with itself. Our definition and theorem give properties of the dot product, but we are still likely wondering "What does the dot product *mean*?" It is helpful to understand that the dot product of a vector with itself is connected to its magnitude.

The next theorem extends this understanding by connecting the dot product to magnitudes and angles. Given vectors $\vec{u}$ and $\vec{v}$ in the plane, an angle θ is clearly formed when $\vec{u}$ and $\vec{v}$ are drawn with the same initial point as illustrated in Figure 12.3.6(a). (We always take θ to be the angle in $[0, \pi]$ as two angles are actually created.)

youtu.be/watch?v=7mZ-eK8fBCk

Figure 12.3.5 Video presentation of Theorem 12.3.4



(a) **(b)**

Figure 12.3.6 Illustrating the angle formed by two vectors with the same initial point

The same is also true of 2 vectors in space: given $\vec{u}$ and $\vec{v}$ in $\mathbb{R}^3$ with the same initial point, there is a plane that contains both $\vec{u}$ and $\vec{v}$. (When $\vec{u}$ and $\vec{v}$ are co-linear, there are infinitely many planes that contain both vectors.) In that plane, we can again find an angle θ between them (and again, $0 \leq \theta \leq \pi$). This is illustrated in Figure 12.3.6(b).

The following theorem connects this angle θ to the dot product of $\vec{u}$ and $\vec{v}$.

Theorem 12.3.7 The Dot Product and Angles.

Let $\vec{u}$ and $\vec{v}$ be nonzero vectors in $\mathbb{R}^2$ or $\mathbb{R}^3$. Then

$$\vec{u} \cdot \vec{v} = \|\vec{u}\| \, \|\vec{v}\| \cos(\theta),$$

where θ, $0 \leq \theta \leq \pi$, is the angle between $\vec{u}$ and $\vec{v}$.

Using Theorem 12.3.4, we can rewrite this theorem as

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

youtu.be/watch?v=d8mzoznYsJo

Figure 12.3.8 Video presentation of Theorem 12.3.7

Note how on the left hand side of the equation, we are computing the dot product of two unit vectors. Recalling that unit vectors essentially only provide direction information, we can informally restate Theorem 12.3.7 as saying "The dot product of two directions gives the cosine of the angle between them."

When θ is an acute angle (i.e., $0 \leq \theta < \pi/2$), $\cos(\theta)$ is positive; when $\theta = \pi/2$, $\cos(\theta) = 0$; when θ is an obtuse angle ($\pi/2 < \theta \leq \pi$), $\cos(\theta)$ is

negative. Thus the sign of the dot product gives a general indication of the angle between the vectors, illustrated in Figure 12.3.9.

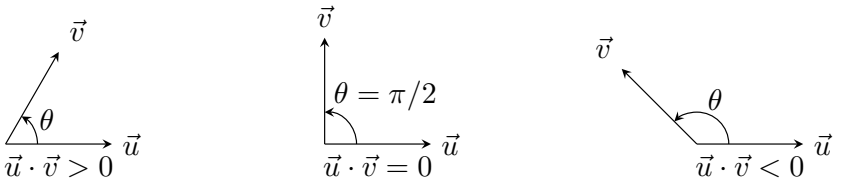


Figure 12.3.9 Illustrating the relationship between the angle between vectors and the sign of their dot product

We *can* use Theorem 12.3.7 to compute the dot product, but generally this theorem is used to find the angle between known vectors (since the dot product is generally easy to compute). To this end, we rewrite the theorem's equation as

$$\cos(\theta) = \frac{\vec{u}\cdot\vec{v}}{\|\vec{u}\|\,\|\vec{v}\|} \Leftrightarrow \theta = \cos^{-1}\left(\frac{\vec{u}\cdot\vec{v}}{\|\vec{u}\|\,\|\vec{v}\|}\right).$$

We practice using this theorem in the following example.

Example 12.3.10 Using the dot product to find angles.

Let $\vec{u} = \langle 3, 1\rangle$, $\vec{v} = \langle -2, 6\rangle$ and $\vec{w} = \langle -4, 3\rangle$, as shown in Figure 12.3.11. Find the angles α, β and θ.

Solution. We start by computing the magnitude of each vector.

$$\|\vec{u}\| = \sqrt{10}; \|\vec{v}\| = 2\sqrt{10}; \|\vec{w}\| = 5.$$

We now apply Theorem 12.3.7 to find the angles.

$$\begin{aligned}\alpha &= \cos^{-1}\left(\frac{\vec{u}\cdot\vec{v}}{(\sqrt{10})(2\sqrt{10})}\right)\\ &= \cos^{-1}(0) = \frac{\pi}{2} = 90^\circ.\end{aligned}$$

$$\begin{aligned}\beta &= \cos^{-1}\left(\frac{\vec{v}\cdot\vec{w}}{(2\sqrt{10})(5)}\right)\\ &= \cos^{-1}\left(\frac{26}{10\sqrt{10}}\right)\\ &\approx 0.6055 \approx 34.7^\circ.\\ \theta &= \cos^{-1}\left(\frac{\vec{u}\cdot\vec{w}}{(\sqrt{10})(5)}\right)\\ &= \cos^{-1}\left(\frac{-9}{5\sqrt{10}}\right)\\ &\approx 2.1763 \approx 124.7^\circ\end{aligned}$$

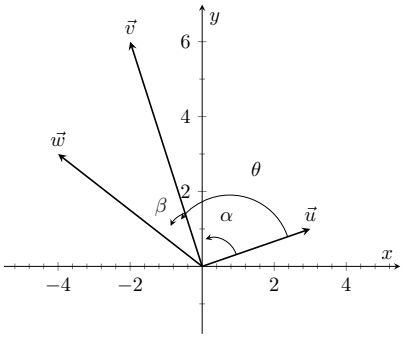


Figure 12.3.11 Vectors used in Example 12.3.10

Video Solution

youtu.be/watch?v=7bkKYObUxFo

We see from our computation that $\alpha+\beta = \theta$, as indicated by Figure 12.3.11. While we knew this should be the case, it is nice to see that this non-intuitive formula indeed returns the results we expected.

We do a similar example next in the context of vectors in space.

Example 12.3.12 Using the dot product to find angles.

Let $\vec{u} = \langle 1, 1, 1\rangle$, $\vec{v} = \langle -1, 3, -2\rangle$ and $\vec{w} = \langle -5, 1, 4\rangle$, as illustrated in Figure 12.3.13. Find the angle between each pair of vectors.

Solution.

1. Between $\vec{u}$ and $\vec{v}$:
$$\begin{aligned}\theta &= \cos^{-1}\left(\frac{\vec{u}\cdot\vec{v}}{\|\vec{u}\|\,\|\vec{v}\|}\right)\\ &= \cos^{-1}\left(\frac{0}{\sqrt{3}\sqrt{14}}\right)\\ &= \frac{\pi}{2}.\end{aligned}$$

2. Between $\vec{u}$ and $\vec{w}$:
$$\begin{aligned}\theta &= \cos^{-1}\left(\frac{\vec{u}\cdot\vec{w}}{\|\vec{u}\|\,\|\vec{w}\|}\right)\\ &= \cos^{-1}\left(\frac{0}{\sqrt{3}\sqrt{42}}\right)\\ &= \frac{\pi}{2}.\end{aligned}$$

3. Between $\vec{v}$ and $\vec{w}$:
$$\begin{aligned}\theta &= \cos^{-1}\left(\frac{\vec{v}\cdot\vec{w}}{\|\vec{v}\|\,\|\vec{w}\|}\right)\\ &= \cos^{-1}\left(\frac{0}{\sqrt{14}\sqrt{42}}\right)\\ &= \frac{\pi}{2}.\end{aligned}$$

While our work shows that each angle is $\pi/2$, i.e., $90°$, none of these angles looks to be a right angle in Figure 12.3.13. Such is the case when drawing three-dimensional objects on the page.

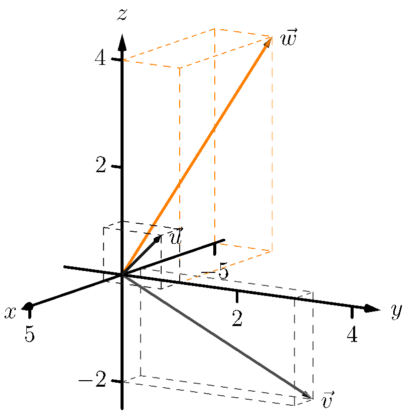


Figure 12.3.13 Vectors used in Example 12.3.12

All three angles between these vectors was $\pi/2$, or $90°$. We know from geometry and everyday life that $90°$ angles are "nice" for a variety of reasons, so it should seem significant that these angles are all $\pi/2$. Notice the common feature in each calculation (and also the calculation of α in Example 12.3.10): the dot products of each pair of angles was 0. We use this as a basis for a definition of the term *orthogonal*, which is essentially synonymous to *perpendicular*.

Definition 12.3.14 Orthogonal.

Nonzero vectors $\vec{u}$ and $\vec{v}$ are **orthogonal** if their dot product is 0.

The term *perpendicular* originally referred to lines. As mathematics progressed, the concept of "being at right angles to" was applied to other objects, such as vectors and planes, and the term *orthogonal* was introduced. It is especially used when discussing objects that are hard, or impossible, to visualize: two vectors in 5-dimensional space are orthogonal if their dot product is 0. It is not wrong to say they are *perpendicular*, but common convention gives preference to the word *orthogonal*.

Example 12.3.15 Finding orthogonal vectors.

Let $\vec{u} = \langle 3, 5\rangle$ and $\vec{v} = \langle 1, 2, 3\rangle$.

1. Find two vectors in $\mathbb{R}^2$ that are orthogonal to $\vec{u}$.

Video Solution

youtu.be/watch?v=eXkGPeTjohM

2. Find two non-parallel vectors in $\mathbb{R}^3$ that are orthogonal to $\vec{v}$.

Solution.

1. Recall that a line perpendicular to a line with slope m has slope $-1/m$, the "opposite reciprocal slope." We can think of the slope of $\vec{u}$ as $5/3$, its "rise over run." A vector orthogonal to $\vec{u}$ will have slope $-3/5$. There are many such choices, though all parallel:

$$\langle -5, 3\rangle \text{ or } \langle 5, -3\rangle \text{ or } \langle -10, 6\rangle \text{ or } \langle 15, -9\rangle, \text{ etc.}$$

2. There are infinitely many directions in space orthogonal to any given direction, so there are an infinite number of non-parallel vectors orthogonal to $\vec{v}$. Since there are so many, we have great leeway in finding some. One way is to arbitrarily pick values for the first two components, leaving the third unknown. For instance, let $\vec{v}_1 = \langle 2, 7, z\rangle$. If $\vec{v}_1$ is to be orthogonal to $\vec{v}$, then $\vec{v}_1 \cdot \vec{v} = 0$, so

$$2 + 14 + 3z = 0 \Rightarrow z = \frac{-16}{3}.$$

So $\vec{v}_1 = \langle 2, 7, -16/3\rangle$ is orthogonal to $\vec{v}$. We can apply a similar technique by leaving the first or second component unknown. Another method of finding a vector orthogonal to $\vec{v}$ mirrors what we did in part 1. Let $\vec{v}_2 = \langle -2, 1, 0\rangle$. Here we switched the first two components of $\vec{v}$, changing the sign of one of them (similar to the "opposite reciprocal" concept before). Letting the third component be 0 effectively ignores the third component of $\vec{v}$, and it is easy to see that

$$\vec{v}_2 \cdot \vec{v} = \langle -2, 1, 0\rangle \cdot \langle 1, 2, 3\rangle = 0.$$

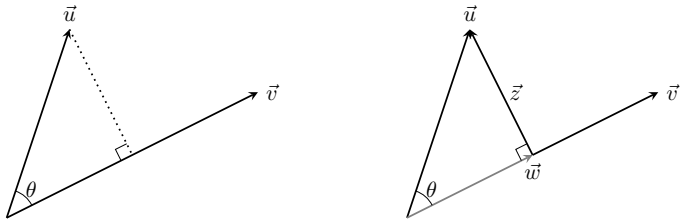
Clearly $\vec{v}_1$ and $\vec{v}_2$ are not parallel.

An important construction is illustrated in Figure 12.3.16, where vectors $\vec{u}$ and $\vec{v}$ are sketched. In Figure 12.3.16(a), a dotted line is drawn from the tip of $\vec{u}$ to the line containing $\vec{v}$, where the dotted line is orthogonal to $\vec{v}$. In Figure 12.3.16(b), the dotted line is replaced with the vector $\vec{z}$ and $\vec{w}$ is formed, parallel to $\vec{v}$. It is clear by the diagram that $\vec{u} = \vec{w} + \vec{z}$. What is important about this construction is this: $\vec{u}$ is *decomposed* as the sum of two vectors, one of which is parallel to $\vec{v}$ and one that is perpendicular to $\vec{v}$. It is hard to overstate the importance of this construction (as we'll see in upcoming examples).

The vectors $\vec{w}$, $\vec{z}$ and $\vec{u}$ as shown in Figure 12.3.16(b) form a right triangle, where the angle between $\vec{v}$ and $\vec{u}$ is labeled θ. We can find $\vec{w}$ in terms of $\vec{v}$ and $\vec{u}$.

Using trigonometry, we can state that

$$\|\vec{w}\| = \|\vec{u}\| \cos(\theta). \tag{12.3.1}$$



(a) **(b)**

Figure 12.3.16 Developing the construction of the *orthogonal projection*

We also know that $\vec{w}$ is parallel to to $\vec{v}$; that is, the direction of $\vec{w}$ is the direction of $\vec{v}$, described by the unit vector $\vec{v}/\left\|\vec{v}\right\|$. The vector $\vec{w}$ is the vector in the direction $\vec{v}/\left\|\vec{v}\right\|$ with magnitude $\left\|\vec{u}\right\| \cos(\theta)$:

$$\begin{aligned}
\vec{w} &= \Big(\left\|\vec{u}\right\| \cos(\theta) \Big) \frac{1}{\left\|\vec{v}\right\|} \vec{v}. \\
&= \left(\left\|\vec{u}\right\| \frac{\vec{u} \cdot \vec{v}}{\left\|\vec{u}\right\| \left\|\vec{v}\right\|} \right) \frac{1}{\left\|\vec{v}\right\|} \vec{v} \text{ (Replacing } \cos(\theta) \text{ using Theorem 12.3.7)} \\
&= \frac{\vec{u} \cdot \vec{v}}{\left\|\vec{v}\right\|^2} \vec{v}. \\
&= \frac{\vec{u} \cdot \vec{v}}{\vec{v} \cdot \vec{v}} \vec{v} \text{ (Applying Theorem 12.3.4).}
\end{aligned}$$

Since this construction is so important, it is given a special name.

Definition 12.3.17 Orthogonal Projection.

Let nonzero vectors $\vec{u}$ and $\vec{v}$ be given. The **orthogonal projection** of $\vec{u}$ onto $\vec{v}$, denoted $\text{proj}_{\vec{v}}\, \vec{u}$, is

$$\text{proj}_{\vec{v}}\, \vec{u} = \frac{\vec{u} \cdot \vec{v}}{\vec{v} \cdot \vec{v}} \vec{v}.$$

youtu.be/watch?v=h8FnrYGOpMg

Figure 12.3.18 Video presentation of Definition 12.3.17

Example 12.3.19 Computing the orthogonal projection.

1. Let $\vec{u} = \langle -2, 1 \rangle$ and $\vec{v} = \langle 3, 1 \rangle$. Find $\text{proj}_{\vec{v}}\, \vec{u}$, and sketch all three vectors with initial points at the origin.
2. Let $\vec{w} = \langle 2, 1, 3 \rangle$ and $\vec{x} = \langle 1, 1, 1 \rangle$. Find $\text{proj}_{\vec{x}}\, \vec{w}$, and sketch all three vectors with initial points at the origin.

Video Solution

youtu.be/watch?v=Pcd-bORHB7o

Solution.

1. Applying Definition 12.3.17, we have

$$\begin{aligned}
\text{proj}_{\vec{v}}\, \vec{u} &= \frac{\vec{u} \cdot \vec{v}}{\vec{v} \cdot \vec{v}} \vec{v} \\
&= \frac{-5}{10} \langle 3, 1 \rangle \\
&= \left\langle -\frac{3}{2}, -\frac{1}{2} \right\rangle.
\end{aligned}$$

Vectors $\vec{u}$, $\vec{v}$ and $\text{proj}_{\vec{v}}\, \vec{u}$ are sketched in Figure 12.3.20. Note how the projection is parallel to $\vec{v}$; that is, it lies on the same line

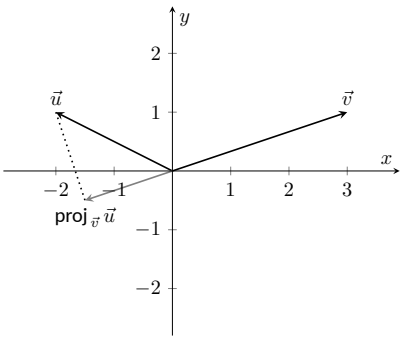


Figure 12.3.20 Sketching the three vectors in Part 1 of Example 12.3.19

through the origin as $\vec{v}$, although it points in the opposite direction. That is because the angle between $\vec{u}$ and $\vec{v}$ is obtuse (i.e., greater than $90°$).

2. Apply the definition:

$$\begin{aligned}\text{proj}_{\vec{x}}\,\vec{w} &= \frac{\vec{w}\cdot\vec{x}}{\vec{x}\cdot\vec{x}}\vec{x}\\ &= \frac{6}{3}\langle 1,1,1\rangle\\ &= \langle 2,2,2\rangle\,.\end{aligned}$$

These vectors are sketched in Figure 12.3.21(a), and again in Figure 12.3.21(b) from a different perspective. Because of the nature of graphing these vectors, the sketch in Figure 12.3.21(a) makes it difficult to recognize that the drawn projection has the geometric properties it should. The graph shown in Figure 12.3.21(b) illustrates these properties better.

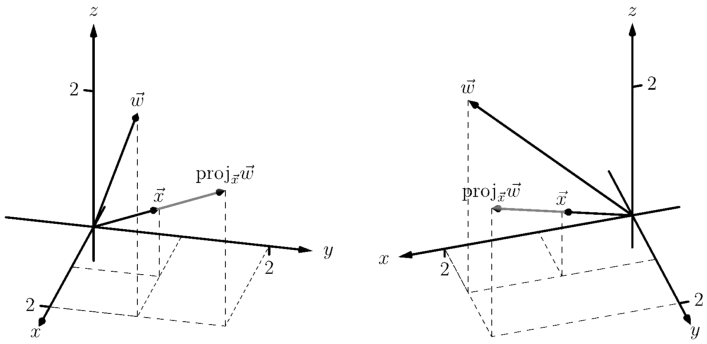


Figure 12.3.21 Sketching the three vectors in Part 2 of Example 12.3.19

We can use the properties of the dot product found in Theorem 12.3.4 to rearrange the formula found in Definition 12.3.17:

$$\begin{aligned}\text{proj}_{\vec{v}}\,\vec{u} &= \frac{\vec{u}\cdot\vec{v}}{\vec{v}\cdot\vec{v}}\vec{v}\\ &= \frac{\vec{u}\cdot\vec{v}}{\|\vec{v}\|^2}\vec{v}\\ &= \left(\vec{u}\cdot\frac{\vec{v}}{\|\vec{v}\|}\right)\frac{\vec{v}}{\|\vec{v}\|}.\end{aligned}$$

The above formula shows that the orthogonal projection of $\vec{u}$ onto $\vec{v}$ is only concerned with the *direction* of $\vec{v}$, as both instances of $\vec{v}$ in the formula come in the form $\vec{v}/\|\vec{v}\|$, the unit vector in the direction of $\vec{v}$.

A special case of orthogonal projection occurs when $\vec{v}$ is a unit vector. In this situation, the formula for the orthogonal projection of a vector $\vec{u}$ onto $\vec{v}$ reduces to just $\text{proj}_{\vec{v}}\,\vec{u} = (\vec{u}\cdot\vec{v})\vec{v}$, as $\vec{v}\cdot\vec{v} = 1$.

This gives us a new understanding of the dot product. When $\vec{v}$ is a unit vector, essentially providing only direction information, the dot product of $\vec{u}$ and $\vec{v}$ gives "how much of $\vec{u}$ is in the direction of $\vec{v}$." This use of the dot product will be very useful in future sections.

Now consider Figure 12.3.22 where the concept of the orthogonal projection is again illustrated. It is clear that

$$\vec{u} = \text{proj}_{\vec{v}}\,\vec{u} + \vec{z}. \tag{12.3.2}$$

As we know what $\vec{u}$ and $\text{proj}_{\vec{v}}\,\vec{u}$ are, we can solve for $\vec{z}$ and state that

$$\vec{z} = \vec{u} - \text{proj}_{\vec{v}}\,\vec{u}.$$

This leads us to rewrite Equation (12.3.2) in a seemingly silly way:

$$\vec{u} = \text{proj}_{\vec{v}}\,\vec{u} + (\vec{u} - \text{proj}_{\vec{v}}\,\vec{u}).$$

This is not nonsense, as pointed out in the following Key Idea. (Notation note: the expression "$\| \vec{y}$" means "is parallel to $\vec{y}$." We can use this notation to state "$\vec{x} \| \vec{y}$" which means "$\vec{x}$ is parallel to $\vec{y}$." The expression "$\perp \vec{y}$" means "is orthogonal to $\vec{y}$," and is used similarly.)

Key Idea 12.3.23 Orthogonal Decomposition of Vectors.

Let nonzero vectors $\vec{u}$ and $\vec{v}$ be given. Then $\vec{u}$ can be written as the sum of two vectors, one of which is parallel to $\vec{v}$, and one of which is orthogonal to $\vec{v}$:

$$\vec{u} = \underbrace{\text{proj}_{\vec{v}}\,\vec{u}}_{\| \vec{v}} + \underbrace{(\vec{u} - \text{proj}_{\vec{v}}\,\vec{u})}_{\perp \vec{v}}.$$

We illustrate the use of this equality in the following example.

Example 12.3.24 Orthogonal decomposition of vectors.

1. Let $\vec{u} = \langle -2, 1\rangle$ and $\vec{v} = \langle 3, 1\rangle$ as in Example 12.3.19. Decompose $\vec{u}$ as the sum of a vector parallel to $\vec{v}$ and a vector orthogonal to $\vec{v}$.
2. Let $\vec{w} = \langle 2, 1, 3\rangle$ and $\vec{x} = \langle 1, 1, 1\rangle$ as in Example 12.3.19. Decompose $\vec{w}$ as the sum of a vector parallel to $\vec{x}$ and a vector orthogonal to $\vec{x}$.

Solution.

1. In Example 12.3.19, we found that $\text{proj}_{\vec{v}}\,\vec{u} = \langle -1.5, -0.5\rangle$. Let

   $$\vec{z} = \vec{u} - \text{proj}_{\vec{v}}\,\vec{u} = \langle -2, 1\rangle - \langle -1.5, -0.5\rangle = \langle -0.5, 1.5\rangle\,.$$

   Is $\vec{z}$ orthogonal to $\vec{v}$? (i.e., is $\vec{z} \perp \vec{v}$?) We check for orthogonality with the dot product:

   $$\vec{z} \cdot \vec{v} = \langle -0.5, 1.5\rangle \cdot \langle 3, 1\rangle = 0.$$

   Since the dot product is 0, we know $\vec{z} \perp \vec{v}$. Thus:

   $$\begin{aligned} \vec{u} &= \text{proj}_{\vec{v}}\,\vec{u} \;+\; (\vec{u} - \text{proj}_{\vec{v}}\,\vec{u}) \\ \langle -2, 1\rangle &= \underbrace{\langle -1.5, -0.5\rangle}_{\| \vec{v}} + \underbrace{\langle -0.5, 1.5\rangle}_{\perp \vec{v}}\,. \end{aligned}$$

2. We found in Example 12.3.19 that $\text{proj}_{\vec{x}}\,\vec{w} = \langle 2, 2, 2\rangle$. Applying the Key Idea, we have:

   $$\vec{z} = \vec{w} - \text{proj}_{\vec{x}}\,\vec{w} = \langle 2, 1, 3\rangle - \langle 2, 2, 2\rangle = \langle 0, -1, 1\rangle\,.$$

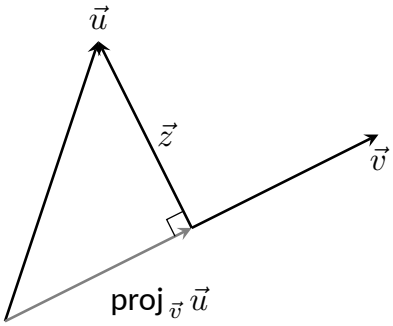


Figure 12.3.22 Illustrating the orthogonal projection

Video Solution

youtu.be/watch?v=9IxUaJ1M4Jk

We check to see if $\vec{z} \perp \vec{x}$:

$$\vec{z} \cdot \vec{x} = \langle 0, -1, 1 \rangle \cdot \langle 1, 1, 1 \rangle = 0.$$

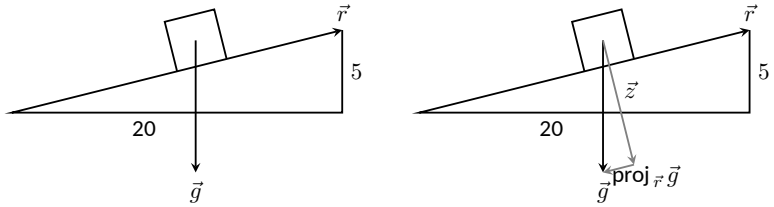
Since the dot product is 0, we know the two vectors are orthogonal. We now write $\vec{w}$ as the sum of two vectors, one parallel and one orthogonal to $\vec{x}$:

$$\begin{aligned} \vec{w} &= \text{proj}_{\vec{x}}\, \vec{w} \ + \ (\vec{w} - \text{proj}_{\vec{x}}\, \vec{w}) \\ \langle 2, 1, 3 \rangle &= \underbrace{\langle 2, 2, 2 \rangle}_{\parallel \vec{x}} + \underbrace{\langle 0, -1, 1 \rangle}_{\perp \vec{x}} \end{aligned}$$

We give an example of where this decomposition is useful.

Example 12.3.25 Orthogonally decomposing a force vector.

Consider Figure 12.3.26(a), showing a box weighing 50lb on a ramp that rises 5ft over a span of 20ft. Find the components of force, and their magnitudes, acting on the box (as sketched in Figure 12.3.26(b)):



(a) **(b)**

Figure 12.3.26 Sketching the ramp and box in Example 12.3.25. Note: *The vectors are not drawn to scale.*

1. in the direction of the ramp, and
2. orthogonal to the ramp.

Video Solution

youtu.be/watch?v=85RquaXgutc

Solution. As the ramp rises 5ft over a horizontal distance of 20ft, we can represent the direction of the ramp with the vector $\vec{r} = \langle 20, 5 \rangle$. Gravity pulls down with a force of 50lb, which we represent with $\vec{g} = \langle 0, -50 \rangle$.

1. To find the force of gravity in the direction of the ramp, we compute $\text{proj}_{\vec{r}}\, \vec{g}$:

$$\begin{aligned} \text{proj}_{\vec{r}}\, \vec{g} &= \frac{\vec{g} \cdot \vec{r}}{\vec{r} \cdot \vec{r}} \vec{r} \\ &= \frac{-250}{425} \langle 20, 5 \rangle \\ &= \left\langle -\frac{200}{17}, -\frac{50}{17} \right\rangle \approx \langle -11.76, -2.94 \rangle \,. \end{aligned}$$

The magnitude of $\text{proj}_{\vec{r}}\, \vec{g}$ is $\|\text{proj}_{\vec{r}}\, \vec{g}\| = 50/\sqrt{17} \approx 12.13$ lb. Though the box weighs 50lb, a force of about 12lb is enough to keep the box from sliding down the ramp.

2. To find the component $\vec{z}$ of gravity orthogonal to the ramp, we use Key Idea 12.3.23.

$$\begin{aligned}\vec{z} &= \vec{g} - \text{proj}_{\vec{r}}\,\vec{g} \\ &= \left\langle \frac{200}{17}, -\frac{800}{17} \right\rangle \approx \langle 11.76, -47.06 \rangle .\end{aligned}$$

The magnitude of this force is $\|\vec{z}\| \approx 48.51$lb. In physics and engineering, knowing this force is important when computing things like static frictional force. (For instance, we could easily compute if the static frictional force alone was enough to keep the box from sliding down the ramp.)

12.3.1 Application to Work

In physics, the application of a force F to move an object in a straight line a distance d produces *work*; the amount of work W is $W = Fd$, (where F is in the direction of travel). The orthogonal projection allows us to compute work when the force is not in the direction of travel.

Consider Figure 12.3.27, where a force $\vec{F}$ is being applied to an object moving in the direction of $\vec{d}$. (The distance the object travels is the magnitude of $\vec{d}$.) The work done is the amount of force in the direction of $\vec{d}$, $\left\|\text{proj}_{\vec{d}}\,\vec{F}\right\|$, times $\left\|\vec{d}\right\|$:

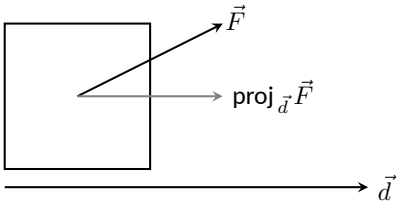


Figure 12.3.27 Finding work when the force and direction of travel are given as vectors

$$\begin{aligned}\left\|\text{proj}_{\vec{d}}\,\vec{F}\right\| \cdot \left\|\vec{d}\right\| &= \left\| \frac{\vec{F}\cdot\vec{d}}{\vec{d}\cdot\vec{d}}\vec{d} \right\| \cdot \left\|\vec{d}\right\| \\ &= \left| \frac{\vec{F}\cdot\vec{d}}{\left\|\vec{d}\right\|^2} \right| \cdot \left\|\vec{d}\right\| \cdot \left\|\vec{d}\right\| \\ &= \frac{\left|\vec{F}\cdot\vec{d}\right|}{\left\|\vec{d}\right\|^2} \left\|\vec{d}\right\|^2 \\ &= \left|\vec{F}\cdot\vec{d}\right| .\end{aligned}$$

The expression $\vec{F}\cdot\vec{d}$ will be positive if the angle between $\vec{F}$ and $\vec{d}$ is acute; when the angle is obtuse (hence $\vec{F}\cdot\vec{d}$ is negative), the force is causing motion in the opposite direction of $\vec{d}$, resulting in "negative work." We want to capture this sign, so we drop the absolute value and find that $W = \vec{F}\cdot\vec{d}$.

Definition 12.3.28 Work.

Let $\vec{F}$ be a constant force that moves an object in a straight line from point P to point Q. Let $\vec{d} = \overrightarrow{PQ}$. The **work** W done by $\vec{F}$ along $\vec{d}$ is $W = \vec{F}\cdot\vec{d}$.

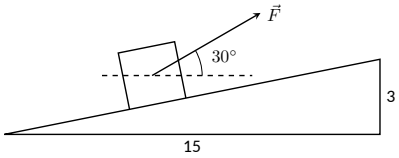


Figure 12.3.30 Computing work when sliding a box up a ramp in Example 12.3.29

Example 12.3.29 Computing work.

A man slides a box along a ramp that rises 3ft over a distance of 15ft by applying 50lb of force as shown in Figure 12.3.30. Compute the work done.

Solution. The figure indicates that the force applied makes a $30°$ angle with the horizontal, so $\vec{F} = 50\,\langle \cos(30°), \sin(30°)\rangle \approx \langle 43.3, 25\rangle$. The ramp is represented by $\vec{d} = \langle 15, 3\rangle$. The work done is simply

$$\vec{F} \cdot \vec{d} = 50\,\langle \cos(30°), \sin(30°)\rangle \cdot \langle 15, 3\rangle \approx 724.5 \text{ ft–lb}.$$

Note how we did not actually compute the distance the object traveled, nor the magnitude of the force in the direction of travel; this is all inherently computed by the dot product!

The dot product is a powerful way of evaluating computations that depend on angles without actually using angles. The next section explores another "product" on vectors, the *cross product*. Once again, angles play an important role, though in a much different way.

12.3.2 Exercises

Terms and Concepts

1. The dot product of two vectors is a _____, not a vector.

2. How are the concepts of the dot product and vector magnitude related?

3. How can one quickly tell if the angle between two vectors is acute or obtuse?

4. Give a synonym for "orthogonal."

Problems

Exercise Group. In the following exercises, find the dot product of the given vectors.

5. $\vec{u} = \langle 2, -4 \rangle$, $\vec{v} = \langle 3, 7 \rangle$
$\vec{u} \cdot \vec{v} =$_____

6. $\vec{u} = \langle 5, 3 \rangle$, $\vec{v} = \langle 6, 1 \rangle$
$\vec{u} \cdot \vec{v} =$_____

7. $\vec{u} = \langle 1, -1, 2 \rangle$, $\vec{v} = \langle 2, 5, 3 \rangle$
$\vec{u} \cdot \vec{v} =$_____

8. $\vec{u} = \langle 3, 5, -1 \rangle$, $\vec{v} = \langle 4, -1, 7 \rangle$
$\vec{u} \cdot \vec{v} =$_____

9. $\vec{u} = \langle 1, 1 \rangle$, $\vec{v} = \langle 1, 2, 3 \rangle$

10. $\vec{u} = \langle 1, 2, 3 \rangle$, $\vec{v} = \langle 0, 0, 0 \rangle$
$\vec{u} \cdot \vec{v} =$_____

11. Create your own vectors $\vec{u}$, $\vec{v}$ and $\vec{w}$ in $\mathbb{R}^2$ and show that $\vec{u} \cdot (\vec{v} + \vec{w}) = \vec{u} \cdot \vec{v} + \vec{u} \cdot \vec{w}$.

12. Create your own vectors $\vec{u}$ and $\vec{v}$ in $\mathbb{R}^3$ and scalar c and show that $c(\vec{u} \cdot \vec{v}) = \vec{u} \cdot (c\vec{v})$.

Exercise Group. In the following exercises, find the measure of the angle between the two vectors in both radians and degrees.

13. The angle between $\vec{u} = \langle 1, 1 \rangle$ and $\vec{v} = \langle 1, 2 \rangle$ is _____.

14. The angle between $\vec{u} = \langle -2, 1 \rangle$ and $\vec{v} = \langle 3, 5 \rangle$ is _____.

15. The angle between $\vec{u} = \langle 8, 1, -4 \rangle$ and $\vec{v} = \langle 2, 2, 0 \rangle$ is _____.

16. The angle between $\vec{u} = \langle 1, 7, 2 \rangle$ and $\vec{v} = \langle 4, -2, 5 \rangle$ is _____.

Exercise Group. In the following exercises, a vector $\vec{v}$ is given. Give two vectors that are orthogonal to $\vec{v}$.

17. Find two nonzero vectors orthogonal to $\vec{v} = \langle 4, 7 \rangle$.
_____ and _____

18. Find two nonzero vectors orthogonal to $\vec{v} = \langle -3, 5 \rangle$.
_____ and _____

19. Find two nonzero vectors orthogonal to $\vec{v} = \langle 1, 1, 1 \rangle$.
_____ and _____

20. Find two nonzero vectors orthogonal to $\vec{v} = \langle 1, -2, 3 \rangle$.
_____ and _____

Exercise Group. In the following exercises, vectors $\vec{u}$ and $\vec{v}$ are given. Find $\text{proj}_{\vec{v}}\, \vec{u}$, the orthogonal projection of $\vec{u}$ onto $\vec{v}$, and sketch all three vectors with the same initial point.

21. If $\vec{u} = \langle 1, 2 \rangle$ and $\vec{v} = \langle -1, 3 \rangle$, then $\text{proj}_{\vec{v}}\, \vec{u} =$_____.
Sketch all three vectors on the same axes.

22. If $\vec{u} = \langle 5, 5 \rangle$ and $\vec{v} = \langle 1, 3 \rangle$, then $\text{proj}_{\vec{v}}\, \vec{u} =$_____.
Sketch all three vectors on the same axes.

23. If $\vec{u} = \langle -3, 2 \rangle$ and $\vec{v} = \langle 1, 1 \rangle$, then $\text{proj}_{\vec{v}}\, \vec{u} =$_____.
Sketch all three vectors on the same axes.

24. If $\vec{u} = \langle -3, 2 \rangle$ and $\vec{v} = \langle 2, 3 \rangle$, then $\text{proj}_{\vec{v}}\, \vec{u} =$_____.
Sketch all three vectors on the same axes.

25. If $\vec{u} = \langle 1, 5, 1 \rangle$ and $\vec{v} = \langle 1, 2, 3 \rangle$, then $\text{proj}_{\vec{v}}\, \vec{u} =$_____.
Sketch all three vectors on the same axes.

26. If $\vec{u} = \langle 3, -1, 2 \rangle$ and $\vec{v} = \langle 2, 2, 1 \rangle$, then $\text{proj}_{\vec{v}}\, \vec{u} =$_____.
Sketch all three vectors on the same axes.

Exercise Group. In the following exercises, vectors $\vec{u}$ and $\vec{v}$ are given. Write $\vec{u}$ as the sum of two vectors, one of which is parallel to $\vec{v}$ (or is zero) and one of which is orthogonal to $\vec{v}$. Note: these are the same pairs of vectors as found in Exercises 12.3.21 — Exercise 12.3.26.

27. Write $\vec{u} = \langle 1, 2 \rangle$ as the sum of two vectors, one parallel to $\vec{v} = \langle -1, 3 \rangle$ (or zero) and the other perpendicular.
$\vec{u} =$_____ + _____

28. Write $\vec{u} = \langle 5, 5 \rangle$ as the sum of two vectors, one parallel to $\vec{v} = \langle 1, 3 \rangle$ (or zero) and the other perpendicular.
$\vec{u} =$_____ + _____

29. Write $\vec{u} = \langle -3, 2 \rangle$ as the sum of two vectors, one parallel to $\vec{v} = \langle 1, 1 \rangle$ (or zero) and the other perpendicular.
$\vec{u} =$_____ + _____

30. Write $\vec{u} = \langle -3, 2 \rangle$ as the sum of two vectors, one parallel to $\vec{v} = \langle 2, 3 \rangle$ (or zero) and the other perpendicular.
$\vec{u} =$_____ + _____

31. Write $\vec{u} = \langle 1, 5, 1 \rangle$ as the sum of two vectors, one parallel to $\vec{v} = \langle 1, 2, 3 \rangle$ (or zero) and the other perpendicular.
$\vec{u} =$_____ + _____

32. Write $\vec{u} = \langle 3, -1, 2 \rangle$ as the sum of two vectors, one parallel to $\vec{v} = \langle 2, 2, 1 \rangle$ (or zero) and the other perpendicular.
$\vec{u} =$_____ + _____

33. A 10lb box sits on a ramp that rises 4ft over a distance of 20ft. How much force is required to keep the box from sliding down the ramp?

34. A 10lb box sits on a 15ft ramp that makes a 30° angle with the horizontal. How much force is required to keep the box from sliding down the ramp?

35. How much work is performed in moving a box horizontally 10ft with a force of 20lb applied at an angle of 45° to the horizontal?

36. How much work is performed in moving a box horizontally 10ft with a force of 20lb applied at an angle of 10° to the horizontal?

37. How much work is performed in moving a box up the length of a ramp that rises 2ft over a distance of 10ft, with a force of 50lb applied horizontally?

38. How much work is performed in moving a box up the length of a ramp that rises 2ft over a distance of 10ft, with a force of 50lb applied at an angle of 45° to the horizontal?

39. How much work is performed in moving a box up the length of a 10ft ramp that makes a 5° angle with the horizontal, with 50lb of force applied in the direction of the ramp?

12.4 The Cross Product

youtu.be/watch?v=63syRlBjyh0

Figure 12.4.1 Video introduction to Section 12.4

"Orthogonality" is immensely important. A quick scan of your current environment will undoubtedly reveal numerous surfaces and edges that are perpendicular to each other (including the edges of this page). The dot product provides a quick test for orthogonality: vectors $\vec{u}$ and $\vec{v}$ are perpendicular if, and only if, $\vec{u} \cdot \vec{v} = 0$.

Given two non-parallel, nonzero vectors $\vec{u}$ and $\vec{v}$ in space, it is very useful to find a vector $\vec{w}$ that is perpendicular to both $\vec{u}$ and $\vec{v}$. There is an operation, called the *cross product*, that creates such a vector. This section defines the cross product, then explores its properties and applications.

Definition 12.4.2 Cross Product.

Let $\vec{u} = \langle u_1, u_2, u_3 \rangle$ and $\vec{v} = \langle v_1, v_2, v_3 \rangle$ be vectors in $\mathbb{R}^3$. The **cross product** of $\vec{u}$ and $\vec{v}$, denoted $\vec{u} \times \vec{v}$, is the vector

$$\vec{u} \times \vec{v} = \langle u_2v_3 - u_3v_2, -(u_1v_3 - u_3v_1), u_1v_2 - u_2v_1 \rangle .$$

This definition can be a bit cumbersome to remember. After an example we will give a convenient method for computing the cross product. For now, careful examination of the products and differences given in the definition should reveal a pattern that is not too difficult to remember. (For instance, in the first component only 2 and 3 appear as subscripts; in the second component, only 1 and 3 appear as subscripts. Further study reveals the order in which they appear.)

Let's practice using this definition by computing a cross product.

Example 12.4.3 Computing a cross product.

Let $\vec{u} = \langle 2, -1, 4 \rangle$ and $\vec{v} = \langle 3, 2, 5 \rangle$. Find $\vec{u} \times \vec{v}$, and verify that it is orthogonal to both $\vec{u}$ and $\vec{v}$.

Solution. Using Definition 12.4.2, we have

$$\vec{u}\times\vec{v} = \big\langle (-1)5 - (4)2, -\big((2)5 - (4)3\big), (2)2 - (-1)3 \big\rangle = \langle -13, 2, 7 \rangle .$$

(We encourage the reader to compute this product on their own, then verify their result.)

We test whether or not $\vec{u} \times \vec{v}$ is orthogonal to $\vec{u}$ and $\vec{v}$ using the dot product:

$$\big(\vec{u} \times \vec{v}\big) \cdot \vec{u} = \langle -13, 2, 7 \rangle \cdot \langle 2, -1, 4 \rangle = 0,$$

$$\big(\vec{u} \times \vec{v}\big) \cdot \vec{v} = \langle -13, 2, 7 \rangle \cdot \langle 3, 2, 5 \rangle = 0.$$

Since both dot products are zero, $\vec{u} \times \vec{v}$ is indeed orthogonal to both $\vec{u}$ and $\vec{v}$.

Video Solution

youtu.be/watch?v=kFpmnYBVQt4

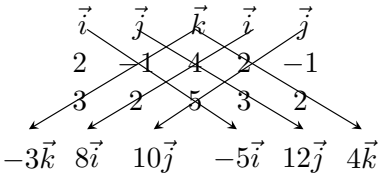
A convenient method of computing the cross product starts with forming a particular 3×3 *matrix*, or rectangular array. The first row comprises the standard unit vectors $\vec{i}$, $\vec{j}$, and $\vec{k}$. The second and third rows are the vectors $\vec{u}$ and $\vec{v}$, respectively. Using $\vec{u}$ and $\vec{v}$ from Example 12.4.3, we begin with:

$$\begin{matrix} \vec{i} & \vec{j} & \vec{k} \\ 2 & -1 & 4 \\ 3 & 2 & 5 \end{matrix}$$

Now repeat the first two columns after the original three:

$$\begin{array}{ccccc} \vec{i} & \vec{j} & \vec{k} & \vec{i} & \vec{j} \\ 2 & -1 & 4 & 2 & -1 \\ 3 & 2 & 5 & 3 & 2 \end{array}$$

This gives three full "upper left to lower right" diagonals, and three full "upper right to lower left" diagonals, as shown. Compute the products along each diagonal, then add the products on the right and subtract the products on the left:



$$\vec{u}\times\vec{v} = \left(-5\vec{i}+12\vec{j}+4\vec{k}\right)-\left(-3\vec{k}+8\vec{i}+10\vec{j}\right) = -13\vec{i}+2\vec{j}+7\vec{k} = \langle -13, 2, 7\rangle\,.$$

youtu.be/watch?v=geWK0Qpl6_Y

Figure 12.4.4 Video presentation of the determinant formula for the cross product

We practice using this method.

Example 12.4.5 Computing a cross product.

Let $\vec{u} = \langle 1, 3, 6\rangle$ and $\vec{v} = \langle -1, 2, 1\rangle$. Compute both $\vec{u}\times\vec{v}$ and $\vec{v}\times\vec{u}$.

Solution. To compute $\vec{u}\times\vec{v}$, we form the matrix as prescribed above, complete with repeated first columns:

$$\begin{array}{ccccc} \vec{i} & \vec{j} & \vec{k} & \vec{i} & \vec{j} \\ 1 & 3 & 6 & 1 & 3 \\ -1 & 2 & 1 & -1 & 2 \end{array}$$

We let the reader compute the products of the diagonals; we give the result:

$$\vec{u}\times\vec{v} = \left(3\vec{i}-6\vec{j}+2\vec{k}\right)-\left(-3\vec{k}+12\vec{i}+\vec{j}\right) = \langle -9, -7, 5\rangle\,.$$

To compute $\vec{v}\times\vec{u}$, we switch the second and third rows of the above matrix, then multiply along diagonals and subtract:

$$\begin{array}{ccccc} \vec{i} & \vec{j} & \vec{k} & \vec{i} & \vec{j} \\ -1 & 2 & 1 & -1 & 2 \\ 1 & 3 & 6 & 1 & 3 \end{array}$$

Note how with the rows being switched, the products that once appeared on the right now appear on the left, and vice-versa. Thus the result is:

$$\vec{v}\times\vec{u} = \left(12\vec{i}+\vec{j}-3\vec{k}\right)-\left(2\vec{k}+3\vec{i}-6\vec{j}\right) = \langle 9, 7, -5\rangle\,,$$

which is the opposite of $\vec{u}\times\vec{v}$. We leave it to the reader to verify that each of these vectors is orthogonal to $\vec{u}$ and $\vec{v}$.

Video Solution

youtu.be/watch?v=6-PGDOhhYbg

12.4.1 Properties of the Cross Product

It is not coincidence that $\vec{v} \times \vec{u} = -(\vec{u} \times \vec{v})$ in the preceding example; one can show using Definition 12.4.2 that this will always be the case. The following theorem states several useful properties of the cross product, each of which can be verified by referring to the definition.

Theorem 12.4.6 Properties of the Cross Product.

Let $\vec{u}$, $\vec{v}$ and $\vec{w}$ be vectors in $\mathbb{R}^3$ and let c be a scalar. The following identities hold:

1. $\vec{u} \times \vec{v} = -(\vec{v} \times \vec{u})$ *Anticommutative Property*
2. *(a)* $(\vec{u} + \vec{v}) \times \vec{w} = \vec{u} \times \vec{w} + \vec{v} \times \vec{w}$ *Distributive Properties*
   (b) $\vec{u} \times (\vec{v} + \vec{w}) = \vec{u} \times \vec{v} + \vec{u} \times \vec{w}$
3. $c(\vec{u} \times \vec{v}) = (c\vec{u}) \times \vec{v} = \vec{u} \times (c\vec{v})$
4. *(a)* $(\vec{u} \times \vec{v}) \cdot \vec{u} = 0$ *Orthogonality Properties*
   (b) $(\vec{u} \times \vec{v}) \cdot \vec{v} = 0$
5. $\vec{u} \times \vec{u} = \vec{0}$
6. $\vec{u} \times \vec{0} = \vec{0}$
7. $\vec{u} \cdot (\vec{v} \times \vec{w}) = (\vec{u} \times \vec{v}) \cdot \vec{w}$ *Triple Scalar Product*

youtu.be/watch?v=QOCSD4qA_Wg

Figure 12.4.7 Video presentation of Theorem 12.4.6

We introduced the cross product as a way to find a vector orthogonal to two given vectors, but we did not give a proof that the construction given in Definition 12.4.2 satisfies this property. Theorem 12.4.6 asserts this property holds; we leave it as a problem in the Exercise section to verify this.

Property 5 from the theorem is also left to the reader to prove in the Exercise section, but it reveals something more interesting than "the cross product of a vector with itself is $\vec{0}$." Let $\vec{u}$ and $\vec{v}$ be parallel vectors; that is, let there be a scalar c such that $\vec{v} = c\vec{u}$. Consider their cross product:

$$\begin{aligned}\vec{u} \times \vec{v} &= \vec{u} \times (c\vec{u}) \\ &= c(\vec{u} \times \vec{u}) \text{ (by Property 3 of Theorem 12.4.6)} \\ &= \vec{0} \text{ (by Property 5 of Theorem 12.4.6).}\end{aligned}$$

We have just shown that the cross product of parallel vectors is $\vec{0}$. This hints at something deeper. Theorem 12.3.7 related the angle between two vectors and their dot product; there is a similar relationship relating the cross product of two vectors and the angle between them, given by the following theorem.

Parallel vectors and the cross product. We could rewrite Definition 12.3.14 and Theorem 12.4.8 to include $\vec{0}$, then define that $\vec{u}$ and $\vec{v}$ are parallel if $\vec{u} \times \vec{v} = \vec{0}$. Since $\vec{0} \cdot \vec{v} = 0$ and $\vec{0} \times \vec{v} = \vec{0}$, this would mean that $\vec{0}$ is both parallel *and* orthogonal to all vectors. Apparent paradoxes such as this are not uncommon in mathematics and can be very useful. (See also the aside in Section 12.2.)

Theorem 12.4.8 The Cross Product and Angles.

Let $\vec{u}$ and $\vec{v}$ be nonzero vectors in $\mathbb{R}^3$. Then

$$\|\vec{u} \times \vec{v}\| = \|\vec{u}\| \, \|\vec{v}\| \sin(\theta),$$

where θ, $0 \leq \theta \leq \pi$, is the angle between $\vec{u}$ and $\vec{v}$.

Note that this theorem makes a statement about the *magnitude* of the cross product. When the angle between $\vec{u}$ and $\vec{v}$ is 0 or π (i.e., the vectors are parallel), the magnitude of the cross product is 0. The only vector with a magnitude of

youtu.be/watch?v=_DjAuiDG2kA

Figure 12.4.9 Video presentation of Theorem 12.4.8 and Example 12.4.10

0 is $\vec{0}$ (see Property 9 of Theorem 12.2.17), hence the cross product of parallel vectors is $\vec{0}$.

We demonstrate the truth of this theorem in the following example.

Example 12.4.10 The cross product and angles.

Let $\vec{u} = \langle 1, 3, 6\rangle$ and $\vec{v} = \langle -1, 2, 1\rangle$ as in Example 12.4.5. Verify Theorem 12.4.8 by finding θ, the angle between $\vec{u}$ and $\vec{v}$, and the magnitude of $\vec{u} \times \vec{v}$.

Solution. We use Theorem 12.3.7 to find the angle between $\vec{u}$ and $\vec{v}$.

$$\begin{aligned}\theta &= \cos^{-1}\left(\frac{\vec{u}\cdot\vec{v}}{\|\vec{u}\|\,\|\vec{v}\|}\right)\\ &= \cos^{-1}\left(\frac{11}{\sqrt{46}\sqrt{6}}\right)\\ &\approx 0.8471 = 48.54^\circ.\end{aligned}$$

Our work in Example 12.4.5 showed that $\vec{u}\times\vec{v} = \langle -9, -7, 5\rangle$, hence $\|\vec{u}\times\vec{v}\| = \sqrt{155}$. Is $\|\vec{u}\times\vec{v}\| = \|\vec{u}\|\,\|\vec{v}\|\sin(\theta)$? Using numerical approximations, we find:

$$\begin{aligned}\|\vec{u}\times\vec{v}\| &= \sqrt{155} & \|\vec{u}\|\,\|\vec{v}\|\sin(\theta) &= \sqrt{46}\sqrt{6}\sin(0.8471)\\ &\approx 12.45. & &\approx 12.45.\end{aligned}$$

Numerically, they seem equal. Using a right triangle, one can show that

$$\sin\left(\cos^{-1}\left(\frac{11}{\sqrt{46}\sqrt{6}}\right)\right) = \frac{\sqrt{155}}{\sqrt{46}\sqrt{6}},$$

which allows us to verify the theorem exactly.

Right Hand Rule. The anticommutative property of the cross product demonstrates that $\vec{u}\times\vec{v}$ and $\vec{v}\times\vec{u}$ differ only by a sign — these vectors have the same magnitude but point in the opposite direction. When seeking a vector perpendicular to $\vec{u}$ and $\vec{v}$, we essentially have two directions to choose from, one in the direction of $\vec{u}\times\vec{v}$ and one in the direction of $\vec{v}\times\vec{u}$. Does it matter which we choose? How can we tell which one we will get without graphing, etc.?

Another wonderful property of the cross product, as defined, is that it follows the *right hand rule*. Given $\vec{u}$ and $\vec{v}$ in $\mathbb{R}^3$ with the same initial point, point the index finger of your right hand in the direction of $\vec{u}$ and let your middle finger point in the direction of $\vec{v}$ (much as we did when establishing the right hand rule for the 3-dimensional coordinate system). Your thumb will naturally extend in the direction of $\vec{u}\times\vec{v}$. One can "practice" this using Figure 12.4.11. If you switch, and point the index finder in the direction of $\vec{v}$ and the middle finger in the direction of $\vec{u}$, your thumb will now point in the opposite direction, allowing you to "visualize" the anticommutative property of the cross product.

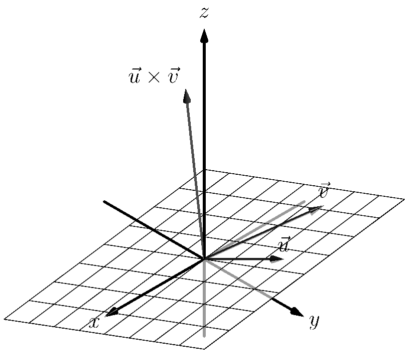


Figure 12.4.11 Illustrating the Right Hand Rule of the cross product

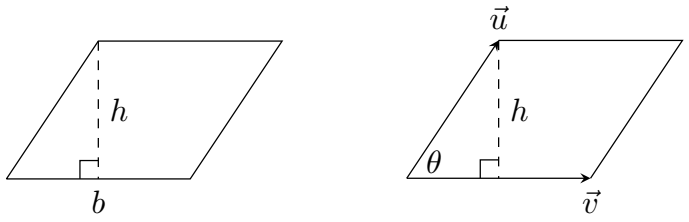
12.4.2 Applications of the Cross Product

There are a number of ways in which the cross product is useful in mathematics, physics and other areas of science beyond "just" finding a vector perpendicular to two others. We highlight a few here.

Area of a Parallelogram. It is a standard geometry fact that the area of a parallelogram is $A = bh$, where b is the length of the base and h is the height of the parallelogram, as illustrated in Figure 12.4.12(a). As shown when defining the Parallelogram Law of vector addition, two vectors $\vec{u}$ and $\vec{v}$ define a parallelogram when drawn from the same initial point, as illustrated in Figure 12.4.12(b). Trigonometry tells us that $h = \|\vec{u}\| \sin(\theta)$, hence the area of the parallelogram is

$$A = \|\vec{u}\| \ \|\vec{v}\| \sin(\theta) = \|\vec{u} \times \vec{v}\|\,, \tag{12.4.1}$$

where the second equality comes from Theorem 12.4.8.



(a) **(b)**

Figure 12.4.12 Using the cross product to find the area of a parallelogram

We illustrate using Equation (12.4.1) in the following example.

Example 12.4.13 Finding the area of a parallelogram.

1. Find the area of the parallelogram defined by the vectors $\vec{u} = \langle 2, 1\rangle$ and $\vec{v} = \langle 1, 3\rangle$.
2. Verify that the points $A = (1, 1, 1)$, $B = (2, 3, 2)$, $C = (4, 5, 3)$ and $D = (3, 3, 2)$ are the vertices of a parallelogram. Find the area of the parallelogram.

Video Solution

youtu.be/watch?v=CYboFA8zYBs

Solution.

1. Figure 12.4.14(a) sketches the parallelogram defined by the vectors $\vec{u}$ and $\vec{v}$. We have a slight problem in that our vectors exist in $\mathbb{R}^2$, not $\mathbb{R}^3$, and the cross product is only defined on vectors in $\mathbb{R}^3$. We skirt this issue by viewing $\vec{u}$ and $\vec{v}$ as vectors in the $x - y$ plane of $\mathbb{R}^3$, and rewrite them as $\vec{u} = \langle 2, 1, 0\rangle$ and $\vec{v} = \langle 1, 3, 0\rangle$. We can now compute the cross product. It is easy to show that $\vec{u} \times \vec{v} = \langle 0, 0, 5\rangle$; therefore the area of the parallelogram is $A = \|\vec{u} \times \vec{v}\| = 5$.

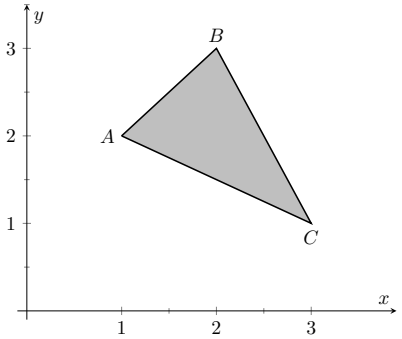


Figure 12.4.16 Finding the area of a triangle in Example 12.4.15

Video Solution

youtu.be/watch?v=GuEn2qOey2U

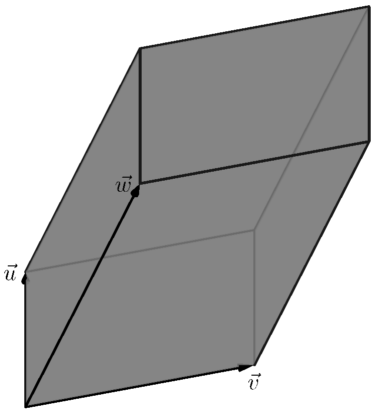
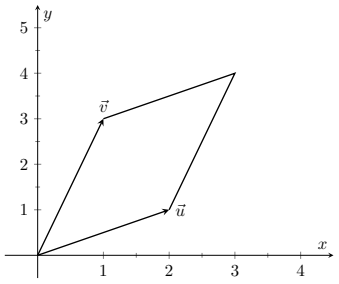
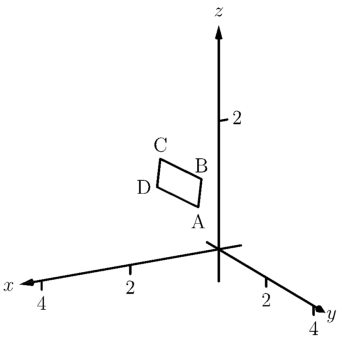


Figure 12.4.17 A parallelepiped is the three dimensional analogue to the parallelogram



(a)



(b)

Figure 12.4.14 Sketching the parallelograms in Example 12.4.13

2. To show that the quadrilateral $ABCD$ is a parallelogram (shown in Figure 12.4.14(b)), we need to show that the opposite sides are parallel. We can quickly show that $\overrightarrow{AB} = \overrightarrow{DC} = \langle 1, 2, 1\rangle$ and $\overrightarrow{BC} = \overrightarrow{AD} = \langle 2, 2, 1\rangle$. We find the area by computing the magnitude of the cross product of $\overrightarrow{AB}$ and $\overrightarrow{BC}$:

$$\overrightarrow{AB} \times \overrightarrow{BC} = \langle 0, 1, -2\rangle \Rightarrow \left\|\overrightarrow{AB} \times \overrightarrow{BC}\right\| = \sqrt{5} \approx 2.236.$$

This application is perhaps more useful in finding the area of a triangle (in short, triangles are used more often than parallelograms). We illustrate this in the following example.

Example 12.4.15 Area of a triangle.

Find the area of the triangle with vertices $A = (1, 2)$, $B = (2, 3)$ and $C = (3, 1)$, as pictured in Figure 12.4.16.

Solution. We found the area of this triangle in Example 7.1.11 to be 1.5 using integration. There we discussed the fact that finding the area of a triangle can be inconvenient using the "$\frac{1}{2}bh$" formula as one has to compute the height, which generally involves finding angles, etc. Using a cross product is much more direct.

We can choose any two sides of the triangle to use to form vectors; we choose $\overrightarrow{AB} = \langle 1, 1\rangle$ and $\overrightarrow{AC} = \langle 2, -1\rangle$. As in the previous example, we will rewrite these vectors with a third component of 0 so that we can apply the cross product. The area of the triangle is

$$\frac{1}{2}\left\|\overrightarrow{AB} \times \overrightarrow{AC}\right\| = \frac{1}{2}\left\|\langle 1, 1, 0\rangle \times \langle 2, -1, 0\rangle\right\| = \frac{1}{2}\left\|\langle 0, 0, -3\rangle\right\| = \frac{3}{2}.$$

We arrive at the same answer as before with less work.

Volume of a Parallelepiped. The three dimensional analogue to the parallelogram is the *parallelepiped*. Each face is parallel to the opposite face, as illustrated in Figure 12.4.17. By crossing $\vec{v}$ and $\vec{w}$, one gets a vector whose magnitude is the area of the base. Dotting this vector with $\vec{u}$ computes the volume of the parallelepiped! (Up to a sign; take the absolute value.)

Thus the volume of a parallelepiped defined by vectors $\vec{u}$, $\vec{v}$ and $\vec{w}$ is

$$V = |\vec{u} \cdot (\vec{v} \times \vec{w})|. \tag{12.4.2}$$

Note how this is the Triple Scalar Product, first seen in Theorem 12.4.6. Applying the identities given in the theorem shows that we can apply the Triple Scalar Product in any "order" we choose to find the volume. That is,

$$V = |\vec{u} \cdot (\vec{v} \times \vec{w})| = |\vec{u} \cdot (\vec{w} \times \vec{v})| = |(\vec{u} \times \vec{v}) \cdot \vec{w}|, \text{ etc.}$$

Example 12.4.18 Finding the volume of parallelepiped.

Find the volume of the parallelepiped defined by the vectors $\vec{u} = \langle 1, 1, 0 \rangle$, $\vec{v} = \langle -1, 1, 0 \rangle$ and $\vec{w} = \langle 0, 1, 1 \rangle$.

Solution. We apply Equation (12.4.2). We first find $\vec{v} \times \vec{w} = \langle 1, 1, -1 \rangle$. Then

$$|\vec{u} \cdot (\vec{v} \times \vec{w})| = |\langle 1, 1, 0 \rangle \cdot \langle 1, 1, -1 \rangle| = 2.$$

So the volume of the parallelepiped is 2 cubic units.

While this application of the Triple Scalar Product is interesting, it is not used all that often: parallelepipeds are not a common shape in physics and engineering. The last application of the cross product is very applicable in engineering.

Video Solution

youtu.be/watch?v=r2_dGc1KWEc

Torque. *Torque* is a measure of the turning force applied to an object. A classic scenario involving torque is the application of a wrench to a bolt. When a force is applied to the wrench, the bolt turns. When we represent the force and wrench with vectors $\vec{F}$ and $\vec{\ell}$, we see that the bolt moves (because of the threads) in a direction orthogonal to $\vec{F}$ and $\vec{\ell}$. Torque is usually represented by the Greek letter τ, or tau, and has units of N·m, a Newton–meter, or ft·lb, a foot–pound.

While a full understanding of torque is beyond the purposes of this book, when a force $\vec{F}$ is applied to a lever arm $\vec{\ell}$, the resulting torque is

$$\vec{\tau} = \vec{\ell} \times \vec{F}. \tag{12.4.3}$$

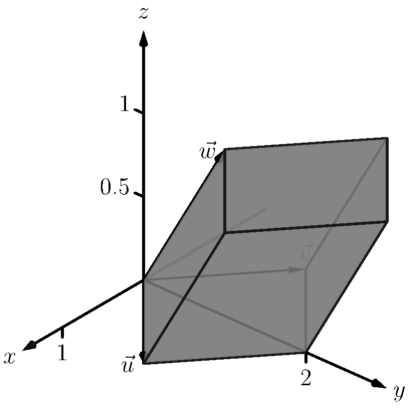


Figure 12.4.19 A parallelepiped in Example 12.4.18

Example 12.4.20 Computing torque.

A lever of length 2ft makes an angle with the horizontal of $45°$. Find the resulting torque when a force of 10lb is applied to the end of the level where:

1. the force is perpendicular to the lever, and
2. the force makes an angle of $60°$ with the lever, as shown in Figure 12.4.21.

Solution.

1. We start by determining vectors for the force and lever arm. Since the lever arm makes a $45°$ angle with the horizontal and is 2ft long, we can state that $\vec{\ell} = 2\langle \cos(45°), \sin(45°) \rangle = \langle \sqrt{2}, \sqrt{2} \rangle$. Since the force vector is perpendicular to the lever arm (as seen in the left hand side of Figure 12.4.21), we can conclude it is making an angle of $-45°$ with the horizontal. As it has a magnitude of 10lb, we can state $\vec{F} = 10\langle \cos(-45°), \sin(-45°) \rangle = \langle 5\sqrt{2}, -5\sqrt{2} \rangle$.

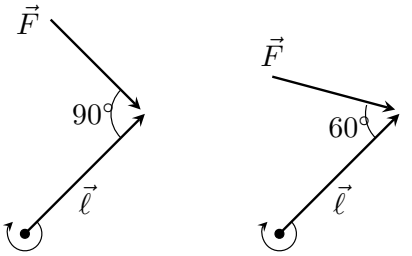


Figure 12.4.21 Showing a force being applied to a lever in Example 12.4.20

Using Equation (12.4.3) to find the torque requires a cross product. We again let the third component of each vector be 0 and compute the cross product:

$$\begin{aligned}\vec{\tau} &= \vec{\ell} \times \vec{F} \\ &= \left\langle \sqrt{2}, \sqrt{2}, 0 \right\rangle \times \left\langle 5\sqrt{2}, -5\sqrt{2}, 0 \right\rangle \\ &= \langle 0, 0, -20 \rangle\end{aligned}$$

This clearly has a magnitude of 20 ft-lb. We can view the force and lever arm vectors as lying "on the page"; our computation of $\vec{\tau}$ shows that the torque goes "into the page." This follows the Right Hand Rule of the cross product, and it also matches well with the example of the wrench turning the bolt. Turning a bolt clockwise moves it in.

2. Our lever arm can still be represented by $\vec{\ell} = \langle \sqrt{2}, \sqrt{2} \rangle$. As our force vector makes a $60°$ angle with $\vec{\ell}$, we can see (referencing the right hand side of the figure) that $\vec{F}$ makes a $-15°$ angle with the horizontal. Thus

$$\begin{aligned}\vec{F} = 10 \langle \cos -15°, \sin -15° \rangle &= \left\langle \frac{5(1+\sqrt{3})}{\sqrt{2}}, \frac{5(-1+\sqrt{3})}{\sqrt{2}} \right\rangle \\ &\approx \langle 9.659, -2.588 \rangle .\end{aligned}$$

We again make the third component 0 and take the cross product to find the torque:

$$\begin{aligned}\vec{\tau} &= \vec{\ell} \times \vec{F} \\ &= \left\langle \sqrt{2}, \sqrt{2}, 0 \right\rangle \times \left\langle \frac{5(1+\sqrt{3})}{\sqrt{2}}, \frac{5(-1+\sqrt{3})}{\sqrt{2}}, 0 \right\rangle \\ &= \left\langle 0, 0, -10\sqrt{3} \right\rangle \\ &\approx \langle 0, 0, -17.321 \rangle .\end{aligned}$$

As one might expect, when the force and lever arm vectors *are* orthogonal, the magnitude of force is greater than when the vectors *are not* orthogonal.

While the cross product has a variety of applications (as noted in this chapter), its fundamental use is finding a vector perpendicular to two others. Knowing a vector is orthogonal to two others is of incredible importance, as it allows us to find the equations of lines and planes in a variety of contexts. The importance of the cross product, in some sense, relies on the importance of lines and planes, which see widespread use throughout engineering, physics and mathematics. We study lines and planes in the next two sections.

12.4.3 Exercises

Terms and Concepts

1. The cross product of two vectors is a _____, not a scalar.

2. One can visualize the direction of $\vec{u} \times \vec{v}$ using the _____ _____ _____.

3. Give a synonym for "orthogonal."

4. True or False? A fundamental principle of the cross product is that $\vec{u} \times \vec{v}$ is orthogonal to $\vec{u}$ and $\vec{v}$. (☐ True ☐ False)

5. _____ is a measure of the turning force applied to an object.

6. T/F: If $\vec{u}$ and $\vec{v}$ are parallel, then $\vec{u} \times \vec{v} = \vec{0}$.

Problems

Exercise Group. In the following exercises, vectors $\vec{u}$ and $\vec{v}$ are given. Compute $\vec{u} \times \vec{v}$ and check this is orthogonal to both $\vec{u}$ and $\vec{v}$.

7. Let $\vec{u} = \langle 3, 2, -2 \rangle$, $\vec{v} = \langle 0, 1, 5 \rangle$.
$\vec{u} \times \vec{v} =$_____
Check this is orthogonal to both $\vec{u}$ and $\vec{v}$.

8. Let $\vec{u} = \langle 5, -4, 3 \rangle$, $\vec{v} = \langle 2, -5, 1 \rangle$.
$\vec{u} \times \vec{v} =$_____
Check this is orthogonal to both $\vec{u}$ and $\vec{v}$.

9. Let $\vec{u} = \langle 4, -5, -5 \rangle$, $\vec{v} = \langle 3, 3, 4 \rangle$.
$\vec{u} \times \vec{v} =$_____
Check this is orthogonal to both $\vec{u}$ and $\vec{v}$.

10. Let $\vec{u} = \langle -4, 7, -10 \rangle$, $\vec{v} = \langle 4, 4, 1 \rangle$.
$\vec{u} \times \vec{v} =$_____
Check this is orthogonal to both $\vec{u}$ and $\vec{v}$.

11. Let $\vec{u} = \langle 1, 0, 1 \rangle$, $\vec{v} = \langle 5, 0, 7 \rangle$.
$\vec{u} \times \vec{v} =$_____
Check this is orthogonal to both $\vec{u}$ and $\vec{v}$.

12. Let $\vec{u} = \langle 1, 5, -4 \rangle$, $\vec{v} = \langle -2, -10, 8 \rangle$.
$\vec{u} \times \vec{v} =$_____
Check this is orthogonal to both $\vec{u}$ and $\vec{v}$.

13. $\vec{u} = \langle a, b, 0 \rangle$, $\vec{v} = \langle c, d, 0 \rangle$

14. Let $\vec{u} = \hat{\imath}$, $\vec{v} = \hat{\jmath}$.
$\vec{u} \times \vec{v} =$_____
Check this is orthogonal to both $\vec{u}$ and $\vec{v}$.

15. Let $\vec{u} = \hat{\imath}$, $\vec{v} = \hat{k}$.
$\vec{u} \times \vec{v} =$_____
Check this is orthogonal to both $\vec{u}$ and $\vec{v}$.

16. Let $\vec{u} = \hat{\jmath}$, $\vec{v} = \hat{k}$.
$\vec{u} \times \vec{v} =$_____
Check this is orthogonal to both $\vec{u}$ and $\vec{v}$.

17. Pick any vectors $\vec{u}$, $\vec{v}$ and $\vec{w}$ in $\mathbb{R}^3$ and show that $\vec{u} \times (\vec{v} + \vec{w}) = \vec{u} \times \vec{v} + \vec{u} \times \vec{w}$.

18. Pick any vectors $\vec{u}$, $\vec{v}$ and $\vec{w}$ in $\mathbb{R}^3$ and show that $\vec{u} \cdot (\vec{v} \times \vec{w}) = (\vec{u} \times \vec{v}) \cdot \vec{w}$.

Exercise Group. In the following exercises, the magnitudes of vectors $\vec{u}$ and $\vec{v}$ in $\mathbb{R}^3$ are given, along with the angle θ between them. Use this information to find the magnitude of $\vec{u} \times \vec{v}$.

19. If $\|\vec{u}\| = 2$, $\|\vec{v}\| = 5$, and $\theta = 30°$ is the angle between $\vec{u}$ and $\vec{v}$, then $\|\vec{u} \times \vec{v}\| =$_____

20. If $\|\vec{u}\| = 3$, $\|\vec{v}\| = 7$, and $\theta = \pi/2$ is the angle between $\vec{u}$ and $\vec{v}$, then $\|\vec{u} \times \vec{v}\| =$_____

21. If $\|\vec{u}\| = 3$, $\|\vec{v}\| = 4$, and $\theta = \pi$ is the angle between $\vec{u}$ and $\vec{v}$, then $\|\vec{u} \times \vec{v}\| =$_____

22. If $\|\vec{u}\| = 2$, $\|\vec{v}\| = 5$, and $\theta = 5\pi/6$ is the angle between $\vec{u}$ and $\vec{v}$, then $\|\vec{u} \times \vec{v}\| =$_____

Exercise Group. In the following exercises, find the area of the parallelogram defined by the given vectors.

23. Find the area of the parallelogram defined by $\vec{u} = \langle 1, 1, 2 \rangle$, and $\vec{v} = \langle 2, 0, 3 \rangle$.

24. Find the area of the parallelogram defined by $\vec{u} = \langle -2, 1, 5 \rangle$, and $\vec{v} = \langle -1, 3, 1 \rangle$.

25. Find the area of the parallelogram defined by $\vec{u} = \langle 1, 2 \rangle$, and $\vec{v} = \langle 2, 1 \rangle$.

26. Find the area of the parallelogram defined by $\vec{u} = \langle 2, 0 \rangle$, and $\vec{v} = \langle 0, 3 \rangle$.

Exercise Group. In the following exercises, find the area of the triangle with the given vertices.

27. Find the area of the triangle with vertices $(0,0,0)$, $(1,3,-1)$ and $(2,1,1)$.

28. Find the area of the triangle with vertices $(5,2,-1)$, $(3,6,2)$ and $(1,0,4)$.

29. Find the area of the triangle with vertices $(1,1)$, $(1,3)$ and $(2,2)$.

30. Find the area of the triangle with vertices $(3,1)$, $(1,2)$ and $(4,3)$.

Exercise Group. In the following exercises, find the area of the quadrilateral with the given vertices. (Hint: break the quadrilateral into two triangles.)

31. Find the area of the quadrilateral with vertices $(0,0)$, $(1,2)$, $(3,0)$, and $(4,3)$.

32. Find the area of the quadrilateral with vertices $(0,0,0)$, $(2,1,1)$, $(-1,2,-8)$, and $(1,-1,5)$.

Exercise Group. In the following exercises, find the volume of the parallelepiped defined by the given vectors.

33. Find the volume of the parallelepiped defined by $\vec{u} = \langle 1,1,1 \rangle$, $\vec{v} = \langle 1,2,3 \rangle$, and $\vec{w} = \langle 1,0,1 \rangle$.

34. Find the volume of the parallelepiped defined by $\vec{u} = \langle -1,2,1 \rangle$, $\vec{v} = \langle 2,2,1 \rangle$, and $\vec{w} = \langle 3,1,3 \rangle$.

Exercise Group. In the following exercises, find a unit vector orthogonal to both $\vec{u}$ and $\vec{v}$.

35. Find a unit vector orthogonal to both $\vec{u} = \langle 1,1,1 \rangle$, and $\vec{v} = \langle 2,0,1 \rangle$.

36. Find a unit vector orthogonal to both $\vec{u} = \langle 1,-2,1 \rangle$, and $\vec{v} = \langle 3,2,1 \rangle$.

37. Find a unit vector orthogonal to both $\vec{u} = \langle 5,0,2 \rangle$, and $\vec{v} = \langle -3,0,7 \rangle$.

38. Find a unit vector orthogonal to both $\vec{u} = \langle 1,-2,1 \rangle$, and $\vec{v} = \langle -2,4,-2 \rangle$.

39. A bicycle rider applies 150lb of force, straight down, onto a pedal that extends 7in horizontally from the crankshaft. Find the magnitude of the torque applied to the crankshaft.

40. A bicycle rider applies 150lb of force, straight down, onto a pedal that extends 7in from the crankshaft, making a 30° angle with the horizontal. Find the magnitude of the torque applied to the crankshaft.

41. To turn a stubborn bolt, 80lb of force is applied to a 10in wrench. What is the maximum amount of torque that can be applied to the bolt?

42. To turn a stubborn bolt, 80lb of force is applied to a 10in wrench in a confined space, where the direction of applied force makes a 10° angle with the wrench. How much torque is subsequently applied to the wrench?

43. Show, using the definition of the Cross Product, that $\vec{u} \cdot (\vec{u} \times \vec{v}) = 0$; that is, that $\vec{u}$ is orthogonal to the cross product of $\vec{u}$ and $\vec{v}$.

44. Show, using the definition of the Cross Product, that $\vec{u} \times \vec{u} = \vec{0}$.

12.5 Lines

To find the equation of a line in the xy-plane, we need two pieces of information: a point and the slope. The slope conveys *direction* information. As vertical lines have an undefined slope, the following statement is more accurate:

To define a line, one needs a point on the line and the direction of the line.

This holds true for lines in space.

youtu.be/watch?v=rA3i_j9MGEM

Figure 12.5.1 Video introduction to Section 12.5

12.5.1 Lines in space

Let P be a point in space, let $\vec{p}$ be the vector with initial point at the origin and terminal point at P (i.e., $\vec{p}$ "points" to P), and let $\vec{d}$ be a vector. Consider the points on the line through P in the direction of $\vec{d}$.

Clearly one point on the line is P; we can say that the *vector* $\vec{p}$ lies at this point on the line. To find another point on the line, we can start at $\vec{p}$ and move in a direction parallel to $\vec{d}$. For instance, starting at $\vec{p}$ and traveling one length of $\vec{d}$ places one at another point on the line. Consider Figure 12.5.2 where certain points along the line are indicated.

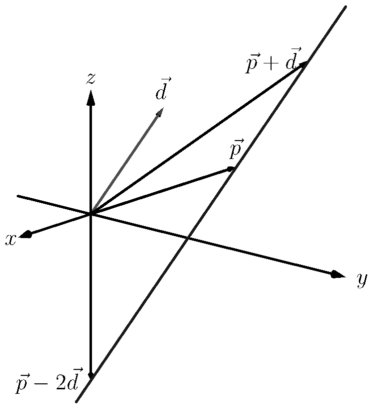


Figure 12.5.2 Defining a line in space

The figure illustrates how every point on the line can be obtained by starting with $\vec{p}$ and moving a certain distance in the direction of $\vec{d}$. That is, we can define the line as a function of t:

$$\vec{\ell}(t) = \vec{p} + t\,\vec{d}. \tag{12.5.1}$$

In many ways, this is *not* a new concept. Compare Equation (12.5.1) to the familiar "$y = mx + b$" equation of a line:

The equations exhibit the same structure: they give a starting point, define a direction, and state how far in that direction to travel.

Equation (12.5.1) is an example of a *vector-valued function*; the input of the function is a real number and the output is a vector. We will cover vector-valued functions extensively in the next chapter.

There are other ways to represent a line. Let $P = (x_0, y_0, z_0)$, $\vec{p} = \langle x_0, y_0, z_0\rangle$, and let $\vec{d} = \langle a, b, c\rangle$. Then the equation of the line through P in the direction of $\vec{d}$ is:

$$\begin{aligned}\vec{\ell}(t) &= \vec{p} + t\vec{d}\\ &= \langle x_0, y_0, z_0\rangle + t\,\langle a, b, c\rangle\\ &= \langle x_0 + at, y_0 + bt, z_0 + ct\rangle\,.\end{aligned}$$

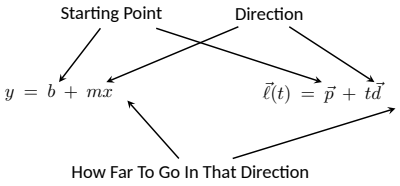


Figure 12.5.3 Understanding the vector equation of a line

The last line states that the x values of the line are given by $x = x_0 + at$, the y values are given by $y = y_0 + bt$, and the z values are given by $z = z_0 + ct$. These three equations, taken together, are the *parametric equations of the line* through $\vec{p}$ in the direction of $\vec{d}$.

Finally, each of the equations for x, y and z above contain the variable t. We can solve for t in each equation:

$$\begin{aligned}x = x_0 + at \Rightarrow t &= \frac{x - x_0}{a},\\ y = y_0 + bt \Rightarrow t &= \frac{y - y_0}{b},\\ z = z_0 + ct \Rightarrow t &= \frac{z - z_0}{c},\end{aligned}$$

assuming $a, b, c \neq 0$. Since t is equal to each expression on the right, we can set these equal to each other, forming the *symmetric equations of the line* through $\vec{p}$ in the direction of $\vec{d}$:

$$\frac{x - x_0}{a} = \frac{y - y_0}{b} = \frac{z - z_0}{c}.$$

Each representation has its own advantages, depending on the context. We summarize these three forms in the following definition, then give examples of their use.

Definition 12.5.4 Equations of Lines in Space.

Let $P = (x_0, y_0, z_0)$ and let $\vec{p} = \langle x_0, y_0, z_0 \rangle$. Consider the line in space that passes through P in the direction of $\vec{d} = \langle a, b, c \rangle$.

1. The *vector equation* of the line is
$$\vec{\ell}(t) = \vec{p} + t\vec{d}.$$
2. The *parametric equations* of the line are
$$x = x_0 + at, y = y_0 + bt, z = z_0 + ct.$$
3. The *symmetric equations* of the line are
$$\frac{x - x_0}{a} = \frac{y - y_0}{b} = \frac{z - z_0}{c}.$$

Example 12.5.5 Finding the equation of a line.

Give all three equations, as given in Definition 12.5.4, of the line through $P = (2, 3, 1)$ in the direction of $\vec{d} = \langle -1, 1, 2 \rangle$. Does the point $Q = (-1, 6, 6)$ lie on this line?

Solution. We identify the point $P = (2, 3, 1)$ with the vector $\vec{p} = \langle 2, 3, 1 \rangle$. Following the definition, we have

- the vector equation of the line is $\vec{\ell}(t) = \langle 2, 3, 1 \rangle + t \langle -1, 1, 2 \rangle$;
- the parametric equations of the line are
$$x = 2 - t, y = 3 + t, z = 1 + 2t; \text{ and}$$
- the symmetric equations of the line are
$$\frac{x - 2}{-1} = \frac{y - 3}{1} = \frac{z - 1}{2}.$$

The first two equations of the line are useful when a t value is given: one can immediately find the corresponding point on the line. These forms are good when calculating with a computer; most software programs easily handle equations in these formats. (For instance, the graphics program that made Figure 12.5.6 can be given the input "`(2-t,3+t,1+2*t)`" for $-1 \le t \le 3$.).

Does the point $Q = (-1, 6, 6)$ lie on the line? The graph in Figure 12.5.6 makes it clear that it does not. We can answer this question without the graph using any of the three equation forms. Of the three, the symmetric equations are probably best suited for this task. Simply plug in the values of x, y and z and see if equality is maintained:

$$\frac{-1 - 2}{-1} \stackrel{?}{=} \frac{6 - 3}{1} \stackrel{?}{=} \frac{6 - 1}{2} \Rightarrow 3 = 3 \neq 2.5.$$

We see that Q does not lie on the line as it did not satisfy the symmetric equations.

Video Solution

youtu.be/watch?v=i307pfanflE

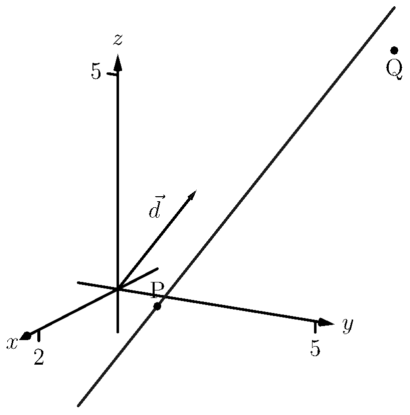


Figure 12.5.6 Graphing a line in Example 12.5.5

Example 12.5.7 Finding the equation of a line through two points.

Find the parametric equations of the line through the points $P = (2, -1, 2)$ and $Q = (1, 3, -1)$.

Solution. Recall the statement made at the beginning of this section: to find the equation of a line, we need a point and a direction. We have *two* points; either one will suffice. The direction of the line can be found by the vector with initial point P and terminal point Q: $\overrightarrow{PQ} = \langle -1, 4, -3\rangle$.

The parametric equations of the line ℓ through P in the direction of $\overrightarrow{PQ}$ are:

$$\ell : x = 2 - ty = -1 + 4tz = 2 - 3t.$$

A graph of the points and line are given in Figure 12.5.8. Note how in the given parametrization of the line, $t = 0$ corresponds to the point P, and $t = 1$ corresponds to the point Q. This relates to the understanding of the vector equation of a line described in Figure 12.5.3. The parametric equations "start" at the point P, and t determines how far in the direction of $\overrightarrow{PQ}$ to travel. When $t = 0$, we travel 0 lengths of $\overrightarrow{PQ}$; when $t = 1$, we travel one length of $\overrightarrow{PQ}$, resulting in the point Q.

Video Solution

youtu.be/watch?v=3ROgMsmz0J0

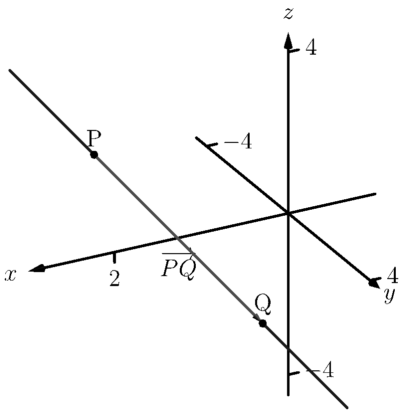


Figure 12.5.8 A graph of the line in Example 12.5.7

12.5.2 Parallel, Intersecting and Skew Lines

In the plane, two *distinct* lines can either be parallel or they will intersect at exactly one point. In space, given equations of two lines, it can sometimes be difficult to tell whether the lines are distinct or not (i.e., the same line can be represented in different ways). Given lines $\vec{\ell}_1(t) = \vec{p}_1 + t\vec{d}_1$ and $\vec{\ell}_2(t) = \vec{p}_2 + t\vec{d}_2$, we have four possibilities: $\vec{\ell}_1$ and $\vec{\ell}_2$ are

the same line	they share all points
intersecting lines	they share only 1 point;
parallel lines	$\vec{d}_1 \parallel \vec{d}_2$, no points in common;
skew lines	$\vec{d}_1 \nparallel \vec{d}_2$, no points in common.

The next two examples investigate these possibilities.

Example 12.5.9 Comparing lines.

Consider lines ℓ_1 and ℓ_2, given in parametric equation form:

$$\ell_1 : \begin{aligned} x &= 1 + 3t \\ y &= 2 - t \\ z &= t \end{aligned} \qquad \ell_2 : \begin{aligned} x &= -2 + 4s \\ y &= 3 + s \\ z &= 5 + 2s \end{aligned}.$$

Determine whether ℓ_1 and ℓ_2 are the same line, intersect, are parallel, or skew.

Solution. We start by looking at the directions of each line. Line ℓ_1 has the direction given by $\vec{d}_1 = \langle 3, -1, 1\rangle$ and line ℓ_2 has the direction given by $\vec{d}_2 = \langle 4, 1, 2\rangle$. It should be clear that $\vec{d}_1$ and $\vec{d}_2$ are not parallel, hence ℓ_1 and ℓ_2 are not the same line, nor are they parallel. Figure 12.5.10 verifies this fact (where the points and directions indicated by the equations of each line are identified).

Video Solution

youtu.be/watch?v=hEYfincwMX0

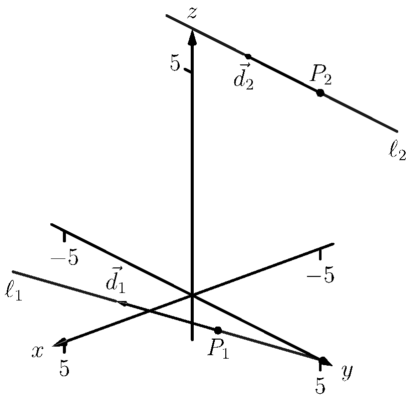


Figure 12.5.10 Sketching the lines from Example 12.5.9

We next check to see if they intersect (if they do not, they are skew lines). To find if they intersect, we look for t and s values such that the respective x, y and z values are the same. That is, we want s and t such that:

$$\begin{aligned} 1+3t &= -2+4s \\ 2-t &= 3+s \\ t &= 5+2s \end{aligned}.$$

This is a relatively simple system of linear equations. Since the last equation is already solved for t, substitute that value of t into the equation above it:

$$2-(5+2s) = 3+s \Rightarrow s=-2,\ t=1.$$

A key to remember is that we have *three* equations; we need to check if $s=-2,\ t=1$ satisfies the first equation as well:

$$1+3(1) \neq -2+4(-2).$$

It does not. Therefore, we conclude that the lines ℓ_1 and ℓ_2 are skew.

Example 12.5.11 Comparing lines.

Consider lines ℓ_1 and ℓ_2, given in parametric equation form:

$$\ell_1: \begin{aligned} x &= -0.7+1.6t \\ y &= 4.2+2.72t \\ z &= 2.3-3.36t \end{aligned} \qquad \ell_2: \begin{aligned} x &= 2.8-2.9s \\ y &= 10.15-4.93s \\ z &= -5.05+6.09s. \end{aligned}$$

Determine whether ℓ_1 and ℓ_2 are the same line, intersect, are parallel, or skew.

Solution. It is obviously very difficult to simply look at these equations and discern anything. This is done intentionally. In the "real world," most equations that are used do not have nice, integer coefficients. Rather, there are lots of digits after the decimal and the equations can look "messy."

We again start by deciding whether or not each line has the same direction. The direction of ℓ_1 is given by $\vec{d}_1 = \langle 1.6, 2.72, -3.36\rangle$ and the direction of ℓ_2 is given by $\vec{d}_2 = \langle -2.9, -4.93, 6.09\rangle$. When it is not clear through observation whether two vectors are parallel or not, the standard way of determining this is by comparing their respective unit vectors. Using a calculator, we find:

$$\vec{u}_1 = \frac{\vec{d}_1}{\left\|\vec{d}_1\right\|} = \langle 0.3471, 0.5901, -0.7289\rangle$$

$$\vec{u}_2 = \frac{\vec{d}_2}{\left\|\vec{d}_2\right\|} = \langle -0.3471, -0.5901, 0.7289\rangle.$$

The two vectors seem to be parallel (at least, their components are equal to 4 decimal places). In most situations, it would suffice to conclude that the lines are at least parallel, if not the same. One way to be

sure is to rewrite $\vec{d}_1$ and $\vec{d}_2$ in terms of fractions, not decimals. We have

$$\vec{d}_1 = \left\langle \frac{16}{10}, \frac{272}{100}, -\frac{336}{100} \right\rangle \qquad \vec{d}_2 = \left\langle -\frac{29}{10}, -\frac{493}{100}, \frac{609}{100} \right\rangle.$$

One can then find the magnitudes of each vector in terms of fractions, then compute the unit vectors likewise. After a lot of manual arithmetic (or after briefly using a computer algebra system), one finds that

$$\vec{u}_1 = \left\langle \sqrt{\frac{10}{83}}, \frac{17}{\sqrt{830}}, -\frac{21}{\sqrt{830}} \right\rangle \qquad \vec{u}_2 = \left\langle -\sqrt{\frac{10}{83}}, -\frac{17}{\sqrt{830}}, \frac{21}{\sqrt{830}} \right\rangle.$$

We can now say without equivocation that these lines are parallel.

Are they the same line? The parametric equations for a line describe one point that lies on the line, so we know that the point $P_1 = (-0.7, 4.2, 2.3)$ lies on ℓ_1. To determine if this point also lies on ℓ_2, plug in the x, y and z values of P_1 into the symmetric equations for ℓ_2:

$$\frac{(-0.7) - 2.8}{-2.9} \stackrel{?}{=} \frac{(4.2) - 10.15}{-4.93} \stackrel{?}{=} \frac{(2.3) - (-5.05)}{6.09}$$
$$1.2069 = 1.2069 = 1.2069.$$

The point P_1 lies on both lines, so we conclude they are the same line, just parametrized differently. Figure 12.5.12 graphs this line along with the points and vectors described by the parametric equations. Note how $\vec{d}_1$ and $\vec{d}_2$ are parallel, though point in opposite directions (as indicated by their unit vectors above).

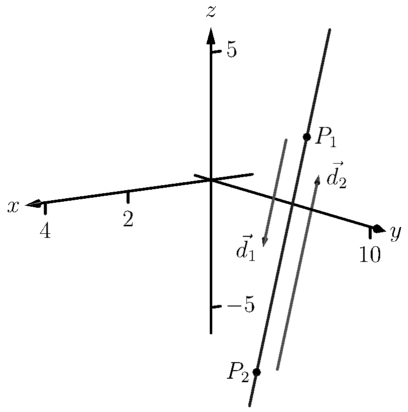


Figure 12.5.12 Graphing the lines in Example 12.5.11

12.5.3 Distances

Given a point Q and a line $\vec{\ell}(t) = \vec{p} + t\vec{d}$ in space, it is often useful to know the distance from the point to the line. (Here we use the standard definition of "distance," i.e., the length of the shortest line segment from the point to the line.) Identifying $\vec{p}$ with the point P, Figure 12.5.13 will help establish a general method of computing this distance h.

From trigonometry, we know $h = \left\|\overrightarrow{PQ}\right\| \sin(\theta)$. We have a similar identity involving the cross product: $\left\|\overrightarrow{PQ} \times \vec{d}\right\| = \left\|\overrightarrow{PQ}\right\| \left\|\vec{d}\right\| \sin(\theta)$. Divide both sides of this latter equation by $\left\|\vec{d}\right\|$ to obtain h:

$$h = \frac{\left\|\overrightarrow{PQ} \times \vec{d}\right\|}{\left\|\vec{d}\right\|}. \tag{12.5.2}$$

It is also useful to determine the distance between lines, which we define as the length of the shortest line segment that connects the two lines (an argument from geometry shows that this line segments is perpendicular to both lines). Let lines $\vec{\ell}_1(t) = \vec{p}_1 + t\vec{d}_1$ and $\vec{\ell}_2(t) = \vec{p}_2 + t\vec{d}_2$ be given, as shown in Figure 12.5.15. To find the direction orthogonal to both $\vec{d}_1$ and $\vec{d}_2$, we take the cross product: $\vec{c} = \vec{d}_1 \times \vec{d}_2$. The magnitude of the orthogonal projection of $\overrightarrow{P_1P_2}$ onto $\vec{c}$ is the distance h we seek:

$$h = \left\|\text{proj}_{\vec{c}}\, \overrightarrow{P_1P_2}\right\|$$

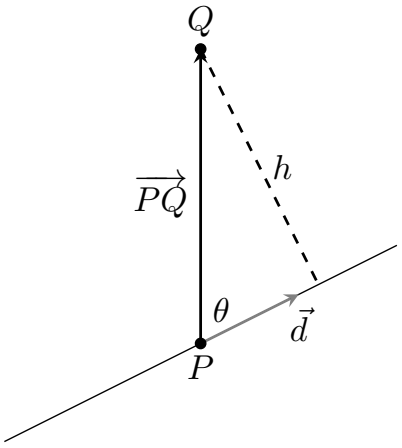


Figure 12.5.13 Establishing the distance from a point to a line

youtu.be/watch?v=ZJ_e_0s2s2M

Figure 12.5.14 Determining distance from a line to a point

$$= \left\| \frac{\overrightarrow{P_1P_2} \cdot \vec{c}}{\vec{c} \cdot \vec{c}} \vec{c} \right\|$$
$$= \frac{\left| \overrightarrow{P_1P_2} \cdot \vec{c} \right|}{\|\vec{c}\|^2} \|\vec{c}\|$$
$$= \frac{\left| \overrightarrow{P_1P_2} \cdot \vec{c} \right|}{\|\vec{c}\|}.$$

A problem in the Exercise section is to show that this distance is 0 when the lines intersect. Note the use of the Triple Scalar Product: $\overrightarrow{P_1P_2} \cdot \vec{c} = \overrightarrow{P_1P_2} \cdot (\vec{d}_1 \times \vec{d}_2)$.

The following Key Idea restates these two distance formulas.

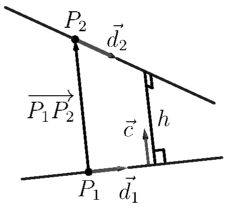


Figure 12.5.15 Establishing the distance between lines

youtu.be/watch?v=h0x9SfUVGVc

Figure 12.5.16 Determining distance between skew lines

Key Idea 12.5.17 Distances to Lines.

1. Let P be a point on a line ℓ that is parallel to $\vec{d}$. The distance h from a point Q to the line ℓ is:
$$h = \frac{\left\| \overrightarrow{PQ} \times \vec{d} \right\|}{\left\| \vec{d} \right\|}.$$
2. Let P_1 be a point on line ℓ_1 that is parallel to $\vec{d}_1$, and let P_2 be a point on line ℓ_2 parallel to $\vec{d}_2$, and let $\vec{c} = \vec{d}_1 \times \vec{d}_2$, where lines ℓ_1 and ℓ_2 are not parallel. The distance h between the two lines is:
$$h = \frac{\left| \overrightarrow{P_1P_2} \cdot \vec{c} \right|}{\|\vec{c}\|}.$$

Example 12.5.18 Finding the distance from a point to a line.

Find the distance from the point $Q = (1, 1, 3)$ to the line $\vec{\ell}(t) = \langle 1, -1, 1 \rangle + t \langle 2, 3, 1 \rangle$.

Solution. The equation of the line gives us the point $P = (1, -1, 1)$ that lies on the line, hence $\overrightarrow{PQ} = \langle 0, 2, 2 \rangle$. The equation also gives $\vec{d} = \langle 2, 3, 1 \rangle$. Following Key Idea 12.5.17, we have the distance as

$$h = \frac{\left\| \overrightarrow{PQ} \times \vec{d} \right\|}{\left\| \vec{d} \right\|}$$
$$= \frac{\|\langle -4, 4, -4 \rangle\|}{\sqrt{14}}$$
$$= \frac{4\sqrt{3}}{\sqrt{14}} \approx 1.852.$$

The point Q is approximately 1.852 units from the line $\vec{\ell}(t)$.

Video Solution

youtu.be/watch?v=xDB7li9crEs

Example 12.5.19 Finding the distance between lines.

Find the distance between the lines

$$\begin{aligned} x &= 1+3t \\ \ell_1: y &= 2-t \\ z &= t \end{aligned} \qquad \begin{aligned} x &= -2+4s \\ \ell_2: y &= 3+s \\ z &= 5+2s. \end{aligned}$$

Video Solution

youtu.be/watch?v=h0x9SfUVGVc

Solution. These are the sames lines as given in Example 12.5.9, where we showed them to be skew. The equations allow us to identify the following points and vectors:

$$P_1 = (1,2,0) P_2 = (-2,3,5) \Rightarrow \overrightarrow{P_1P_2} = \langle -3,1,5\rangle.$$

$$\vec{d_1} = \langle 3,-1,1\rangle \, \vec{d_2} = \langle 4,1,2\rangle \Rightarrow \vec{c} = \vec{d_1} \times \vec{d_2} = \langle -3,-2,7\rangle.$$

From Key Idea 12.5.17 we have the distance h between the two lines is

$$\begin{aligned} h &= \frac{\left|\overrightarrow{P_1P_2} \cdot \vec{c}\right|}{\|\vec{c}\|} \\ &= \frac{42}{\sqrt{62}} \approx 5.334. \end{aligned}$$

The lines are approximately 5.334 units apart.

One of the key points to understand from this section is this: to describe a line, we need a point and a direction. Whenever a problem is posed concerning a line, one needs to take whatever information is offered and glean point and direction information. Many questions can be asked (and *are* asked in the Exercise section) whose answer immediately follows from this understanding.

Lines are one of two fundamental objects of study in space. The other fundamental object is the *plane*, which we study in detail in the next section. Many complex three dimensional objects are studied by approximating their surfaces with lines and planes.

12.5.4 Exercises

Terms and Concepts

1. To find an equation of a line, what two pieces of information are needed?

2. Two distinct lines in the plane can intersect or be _____.

3. Two distinct lines in space can intersect, be _____ or be _____.

4. Use your own words to describe what it means for two lines in space to be skew.

Problems

Exercise Group. Write the vector, parametric and symmetric equations of the lines described.

5. Passes through $P = (2, -4, 1)$, parallel to $\vec{d} = \langle 9, 2, 5 \rangle$.

6. ℓ is a line that passes through $P = (6, 1, 7)$, parallel to $\vec{d} = \langle -3, 2, 5 \rangle$.
$\ell(t) =$_____.

7. Passes through $P = (2, 1, 5)$ and $Q = (7, -2, 4)$.

8. ℓ is a line that passes through $P = (1, -2, 3)$ and $Q = (5, 5, 5)$.
$\ell(t) =$_____.

9. Passes through $P = (0, 1, 2)$ and orthogonal to both $\vec{d_1} = \langle 2, -1, 7 \rangle$ and $\vec{d_2} = \langle 7, 1, 3 \rangle$.

10. ℓ is a line that passes through $P = (5, 1, 9)$ and is orthogonal to both $\vec{d_1} = \langle 1, 0, 1 \rangle$ and $\vec{d_2} = \langle 2, 0, 3 \rangle$.
$\ell(t) =$_____.

11. ℓ is a line that passes through the intersection of $\vec{\ell_1}(t) = \langle 2, 1, 1 \rangle + t \langle 5, 1, -2 \rangle$ and $\vec{\ell_2}(t) = \langle -2, -1, 2 \rangle + t \langle 3, 1, -1 \rangle$, and is orthogonal to both lines.
$\ell(t) =$_____.

12. ℓ is a line that passes through the intersection of $\vec{\ell_1}(t) = \begin{cases} x &= t \\ y &= -2 + 2t \\ z &= 1 + t \end{cases}$ and $\vec{\ell_2}(t) = \begin{cases} x &= 2 + t \\ y &= 2 - t \\ z &= 3 + 2t \end{cases}$, and is orthogonal to both lines.
$\ell(t) =$_____.

13. Passes through $P = (1, 1)$, parallel to $\vec{d} = \langle 2, 3 \rangle$.

14. ℓ is a line that passes through $P = (-2, 5)$, parallel to $\vec{d} = \langle 0, 1 \rangle$.
$\ell(t) =$_____.

Exercise Group. Determine if the described lines are the same line, parallel lines, intersecting or skew lines. If intersecting, give the point of intersection.

15. Determine if $\vec{\ell_1}(t) = \langle 1, 2, 1 \rangle + t \langle 2, -1, 1 \rangle$ and $\vec{\ell_2}(t) = \langle 3, 3, 3 \rangle + t \langle -4, 2, -2 \rangle$ describe the same line, parallel lines, intersecting or skew lines. You can answer with "`same`", "`parallel`", "`skew`", or give the point of intersection.

16. Determine if $\vec{\ell_1}(t) = \langle 2, 1, 1 \rangle + t \langle 5, 1, 3 \rangle$ and $\vec{\ell_2}(t) = \langle 14, 5, 9 \rangle + t \langle 1, 1, 1 \rangle$ describe the same line, parallel lines, intersecting or skew lines. You can answer with "`same`", "`parallel`", "`skew`", or give the point of intersection.

17. $\vec{\ell_1}(t) = \langle 3, 4, 1 \rangle + t \langle 2, -3, 4 \rangle$, $\vec{\ell_2}(t) = \langle -3, 3, -3 \rangle + t \langle 3, -2, 4 \rangle$.

18. Determine if $\vec{\ell_1}(t) = \langle 1, 1, 1 \rangle + t \langle 3, 1, 3 \rangle$ and $\vec{\ell_2}(t) = \langle 7, 3, 7 \rangle + t \langle 6, 2, 6 \rangle$ describe the same line, parallel lines, intersecting or skew lines. You can answer with "`same`", "`parallel`", "`skew`", or give the point of intersection.

19. Determine if $\vec{\ell}_1(t) = \begin{cases} x &= 1+2t \\ y &= 3-2t \\ z &= t \end{cases}$ and $\vec{\ell}_2(t) = \begin{cases} x &= 3-t \\ y &= 3+5t \\ z &= 2+7t \end{cases}$ describe the same line, parallel lines, intersecting or skew lines. You can answer with "`same`", "`parallel`", "`skew`", or give the point of intersection.

20. Determine if $\vec{\ell}_1(t) = \begin{cases} x &= 1.1+0.6t \\ y &= 3.77+0.9t \\ z &= -2.3+1.5t \end{cases}$ and $\vec{\ell}_2(t) = \begin{cases} x &= 3.11+3.4t \\ y &= 2+5.1t \\ z &= 2.5+8.5t \end{cases}$ describe the same line, parallel lines, intersecting or skew lines. You can answer with "`same`", "`parallel`", "`skew`", or give the point of intersection.

21. $\ell_1 = \begin{cases} x = 0.2+0.6t \\ y = 1.33-0.45t \\ z = -4.2+1.05t \end{cases}$ and $\ell_2 = \begin{cases} x = 0.86+9.2t \\ y = 0.835-6.9t \\ z = -3.045+16.1t \end{cases}$

22. Determine if $\vec{\ell}_1(t) = \begin{cases} x &= 0.1+1.1t \\ y &= 2.9-1.5t \\ z &= 3.2+1.6t \end{cases}$ and $\vec{\ell}_2(t) = \begin{cases} x &= 4-2.1t \\ y &= 1.8+7.2t \\ z &= 3.1+1.1t \end{cases}$ describe the same line, parallel lines, intersecting or skew lines. You can answer with "`same`", "`parallel`", "`skew`", or give the point of intersection.

Exercise Group. Find the distance from the point to the line.

23. $Q = (1,1,1)$, $\vec{\ell}(t) = \langle 2,1,3\rangle + t\langle 2,1,-2\rangle$

24. Find the distance from the point $Q = (2,5,6)$ to the line $\vec{\ell}(t) = \langle -1,1,1\rangle + t\langle 1,0,1\rangle$.

25. $Q = (0,3)$, $\vec{\ell}(t) = \langle 2,0\rangle + t\langle 1,1\rangle$

26. Find the distance from the point $Q = (1,1)$ to the line $\vec{\ell}(t) = \langle 4,5\rangle + t\langle -4,3\rangle$.

Exercise Group. Find the distance between the two lines.

27. $\vec{\ell}_1(t) = \langle 1,2,1\rangle + t\langle 2,-1,1\rangle$, $\vec{\ell}_2(t) = \langle 3,3,3\rangle + t\langle 4,2,-2\rangle$.

28. Find the distance between the line $\vec{\ell}_1(t) = \langle 0,0,1\rangle + t\langle 1,0,0\rangle$ and the line $\vec{\ell}_2(t) = \langle 0,0,3\rangle + t\langle 0,1,0\rangle$.

Exercise Group. The following exercises explore special cases of the distance formulas found in Key Idea 12.5.17.

29. Let Q be a point on the line $\vec{\ell}(t)$. Show why the distance formula correctly gives the distance from the point to the line as 0.

30. Let lines $\vec{\ell}_1(t)$ and $\vec{\ell}_2(t)$ be intersecting lines. Show why the distance formula correctly gives the distance between these lines as 0.

31. Let lines $\vec{\ell}_1(t)$ and $\vec{\ell}_2(t)$ be parallel.

(a) Show why the distance formula for distance between lines cannot be used as stated to find the distance between the lines.

(b) Show why letting $\vec{c} = (\overrightarrow{P_1P_2} \times \vec{d}_2) \times \vec{d}_2$ allows one to the use the formula.

(c) Show how one can use the formula for the distance between a point and a line to find the distance between parallel lines.

12.6 Planes

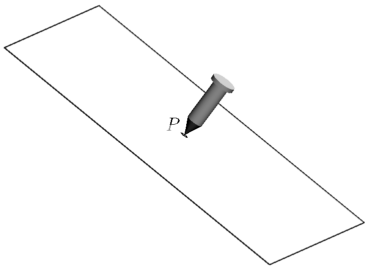


Figure 12.6.1 Illustrating defining a plane with a sheet of cardboard and a nail

youtu.be/watch?v=aP-fcbArvtA

Figure 12.6.2 Video inroduction to Section 12.6

Any flat surface, such as a wall, table top or stiff piece of cardboard can be thought of as representing part of a plane. Consider a piece of cardboard with a point P marked on it. One can take a nail and stick it into the cardboard at P such that the nail is perpendicular to the cardboard; see Figure 12.6.1.

This nail provides a "handle" for the cardboard. Moving the cardboard around moves P to different locations in space. Tilting the nail (but keeping P fixed) tilts the cardboard. Both moving and tilting the cardboard defines a different plane in space. In fact, we can define a plane by: 1) the location of P in space, and 2) the direction of the nail.

The previous section showed that one can define a line given a point on the line and the direction of the line (usually given by a vector). One can make a similar statement about planes: we can define a plane in space given a point on the plane and the direction the plane "faces" (using the description above, the direction of the nail). Once again, the direction information will be supplied by a vector, called a *normal vector*, that is orthogonal to the plane.

What exactly does "orthogonal to the plane" mean? Choose any two points P and Q in the plane, and consider the vector $\overrightarrow{PQ}$. We say a vector $\vec{n}$ is orthogonal to the plane if $\vec{n}$ is perpendicular to $\overrightarrow{PQ}$ for all choices of P and Q; that is, if $\vec{n} \cdot \overrightarrow{PQ} = 0$ for all P and Q.

This gives us way of writing an equation describing the plane. Let $P = (x_0, y_0, z_0)$ be a point in the plane and let $\vec{n} = \langle a, b, c\rangle$ be a normal vector to the plane. A point $Q = (x, y, z)$ lies in the plane defined by P and $\vec{n}$ if, and only if, $\overrightarrow{PQ}$ is orthogonal to $\vec{n}$. Knowing $\overrightarrow{PQ} = \langle x - x_0, y - y_0, z - z_0\rangle$, consider:

$$\begin{aligned} \overrightarrow{PQ} \cdot \vec{n} &= 0 \\ \langle x - x_0, y - y_0, z - z_0\rangle \cdot \langle a, b, c\rangle &= 0 \\ a(x - x_0) + b(y - y_0) + c(z - z_0) &= 0. \end{aligned} \tag{12.6.1}$$

Equation (12.6.1) defines an *implicit* function describing the plane. More algebra produces:

$$ax + by + cz = ax_0 + by_0 + cz_0.$$

The right hand side is just a number, so we replace it with d:

$$ax + by + cz = d. \tag{12.6.2}$$

As long as $c \neq 0$, we can solve for z:

$$z = \frac{1}{c}(d - ax - by). \tag{12.6.3}$$

Equation (12.6.3) is especially useful as many computer programs can graph functions in this form. Equations (12.6.1) and (12.6.2) have specific names, given next.

Definition 12.6.3 Equations of a Plane in Standard and General Forms.

The plane passing through the point $P = (x_0, y_0, z_0)$ with normal vector $\vec{n} = \langle a, b, c\rangle$ can be described by an equation with **standard form**

$$a(x - x_0) + b(y - y_0) + c(z - z_0) = 0;$$

the equation's **general form** is

$$ax + by + cz = d.$$

A key to remember throughout this section is this: to find the equation of a plane, we need a point and a normal vector. We will give several examples of finding the equation of a plane, and in each one different types of information are given. In each case, we need to use the given information to find a point on the plane and a normal vector.

Example 12.6.4 Finding the equation of a plane.

Write the equation of the plane that passes through the points $P = (1,1,0)$, $Q = (1,2,-1)$ and $R = (0,1,2)$ in standard form.

Solution. We need a vector $\vec{n}$ that is orthogonal to the plane. Since P, Q and R are in the plane, so are the vectors $\overrightarrow{PQ}$ and $\overrightarrow{PR}$; $\overrightarrow{PQ} \times \overrightarrow{PR}$ is orthogonal to $\overrightarrow{PQ}$ and $\overrightarrow{PR}$ and hence the plane itself.

It is straightforward to compute $\vec{n} = \overrightarrow{PQ} \times \overrightarrow{PR} = \langle 2,1,1 \rangle$. We can use any point we wish in the plane (any of P, Q or R will do) and we arbitrarily choose P. Following Definition 12.6.3, the equation of the plane in standard form is

$$2(x-1) + (y-1) + z = 0.$$

The plane is sketched in Figure 12.6.5.

Video Solution

youtu.be/watch?v=oc77R0am1vk

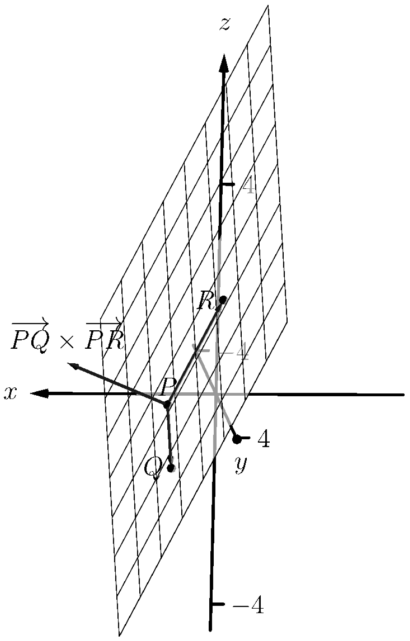


Figure 12.6.5 Sketching the plane in Example 12.6.4

We have just demonstrated the fact that any three non-collinear points define a plane. (This is why a three-legged stool does not "rock;" it's three feet always lie in a plane. A four-legged stool will rock unless all four feet lie in the same plane.)

Example 12.6.6 Finding the equation of a plane.

Verify that lines ℓ_1 and ℓ_2, whose parametric equations are given below, intersect, then give the equation of the plane that contains these two lines in general form.

$$\ell_1: \begin{aligned} x &= -5+2s \\ y &= 1+s \\ z &= -4+2s \end{aligned} \qquad \ell_2: \begin{aligned} x &= 2+3t \\ y &= 1-2t \\ z &= 1+t \end{aligned}$$

Solution. The lines clearly are not parallel. If they do not intersect, they are skew, meaning there is not a plane that contains them both. If they do intersect, there is such a plane.

To find their point of intersection, we set the x, y and z equations equal to each other and solve for s and t:

$$\begin{aligned} -5+2s &= 2+3t \\ 1+s &= 1-2t \Rightarrow s = 2, t = -1. \\ -4+2s &= 1+t \end{aligned}$$

When $s = 2$ and $t = -1$, the lines intersect at the point $P = (-1,3,0)$.

Let $\vec{d_1} = \langle 2,1,2 \rangle$ and $\vec{d_2} = \langle 3,-2,1 \rangle$ be the directions of lines ℓ_1 and ℓ_2, respectively. A normal vector to the plane containing these the two lines will also be orthogonal to $\vec{d_1}$ and $\vec{d_2}$. Thus we find a normal vector $\vec{n}$ by computing $\vec{n} = \vec{d_1} \times \vec{d_2} = \langle 5,4,-7 \rangle$.

We can pick any point in the plane with which to write our equation; each line gives us infinite choices of points. We choose P, the point of

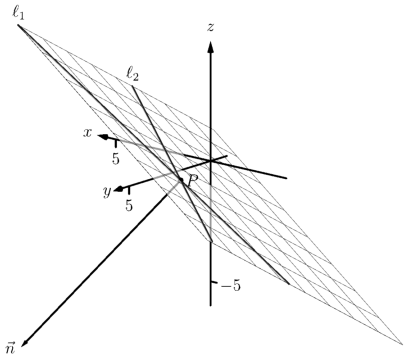


Figure 12.6.7 Sketching the plane in Example 12.6.6

intersection. We follow Definition 12.6.3 to write the plane's equation in general form:

$$\begin{aligned} 5(x+1)+4(y-3)-7z &= 0 \\ 5x+5+4y-12-7z &= 0 \\ 5x+4y-7z &= 7. \end{aligned}$$

The plane's equation in general form is $5x + 4y - 7z = 7$; it is sketched in Figure 12.6.7.

Example 12.6.8 Finding the equation of a plane.

Give the equation, in standard form, of the plane that passes through the point $P = (-1, 0, 1)$ and is orthogonal to the line with vector equation $\vec{\ell}(t) = \langle -1, 0, 1\rangle + t\,\langle 1, 2, 2\rangle$.

Solution. As the plane is to be orthogonal to the line, the plane must be orthogonal to the direction of the line given by $\vec{d} = \langle 1, 2, 2\rangle$. We use this as our normal vector. Thus the plane's equation, in standard form, is

$$(x+1)+2y+2(z-1)=0.$$

The line and plane are sketched in Figure 12.6.9.

Video Solution

youtu.be/watch?v=2diP-H-QLoI

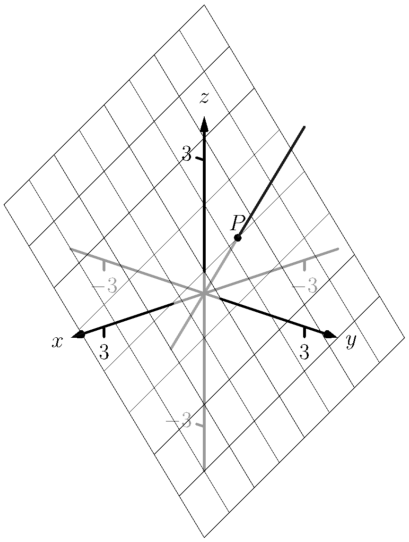


Figure 12.6.9 The line and plane in Example 12.6.8

Example 12.6.10 Finding the intersection of two planes.

Give the parametric equations of the line that is the intersection of the planes p_1 and p_2, where:

$$\begin{aligned} p_1 &: x-(y-2)+(z-1)=0 \\ p_2 &: -2(x-2)+(y+1)+(z-3)=0 \end{aligned}$$

Solution. To find an equation of a line, we need a point on the line and the direction of the line.

We can find a point on the line by solving each equation of the planes for z:

$$\begin{aligned} p_1 &: z = -x+y-1 \\ p_2 &: z = 2x-y-2 \end{aligned}$$

We can now set these two equations equal to each other (i.e., we are finding values of x and y where the planes have the same z value):

$$\begin{aligned} -x+y-1 &= 2x-y-2 \\ 2y &= 3x-1 \\ y &= \frac{1}{2}(3x-1) \end{aligned}$$

We can choose any value for x; we choose $x = 1$. This determines that $y = 1$. We can now use the equations of either plane to find z: when $x = 1$ and $y = 1$, $z = -1$ on both planes. We have found a point P on the line: $P = (1, 1, -1)$.

Video Solution

youtu.be/watch?v=x2HB6huEEmQ

We now need the direction of the line. Since the line lies in each plane, its direction is orthogonal to a normal vector for each plane. Considering the equations for p_1 and p_2, we can quickly determine their normal vectors. For p_1, $\vec{n}_1 = \langle 1, -1, 1\rangle$ and for p_2, $\vec{n}_2 = \langle -2, 1, 1\rangle$. A direction orthogonal to both of these directions is their cross product: $\vec{d} = \vec{n}_1 \times \vec{n}_2 = \langle -2, -3, -1\rangle$.

The parametric equations of the line through $P = (1, 1, -1)$ in the direction of $d = \langle -2, -3, -1\rangle$ is:

$$\ell : x = -2t + 1y = -3t + 1z = -t - 1.$$

The planes and line are graphed in Figure 12.6.11.

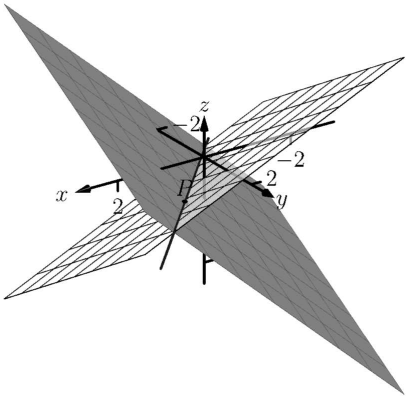


Figure 12.6.11 Graphing the planes and their line of intersection in Example 12.6.10

youtu.be/watch?v=tthyDV7cjR8

Figure 12.6.12 An alternative method for solving Example 12.6.10

Example 12.6.13 Finding the intersection of a plane and a line.

Find the point of intersection, if any, of the line $\ell(t) = \langle 3, -3, -1\rangle + t\langle -1, 2, 1\rangle$ and the plane with equation in general form $2x + y + z = 4$.

Solution. The equation of the plane shows that the vector $\vec{n} = \langle 2, 1, 1\rangle$ is a normal vector to the plane, and the equation of the line shows that the line moves parallel to $\vec{d} = \langle -1, 2, 1\rangle$. Since these are not orthogonal, we know there is a point of intersection. (If there were orthogonal, it would mean that the plane and line were parallel to each other, either never intersecting or the line was in the plane itself.)

To find the point of intersection, we need to find a t value such that $\ell(t)$ satisfies the equation of the plane. Rewriting the equation of the line with parametric equations will help:

$$\ell(t) = \begin{cases} x = 3 - t \\ y = -3 + 2t \\ z = -1 + t \end{cases}.$$

Replacing x, y and z in the equation of the plane with the expressions containing t found in the equation of the line allows us to determine a t value that indicates the point of intersection:

$$\begin{aligned} 2x + y + z &= 4 \\ 2(3 - t) + (-3 + 2t) + (-1 + t) &= 4 \\ t &= 2. \end{aligned}$$

When $t = 2$, the point on the line satisfies the equation of the plane; that point is $\ell(2) = \langle 1, 1, 1\rangle$. Thus the point $(1, 1, 1)$ is the point of intersection between the plane and the line, illustrated in Figure 12.6.14.

Video Solution

youtu.be/watch?v=DjuaixH0ayo

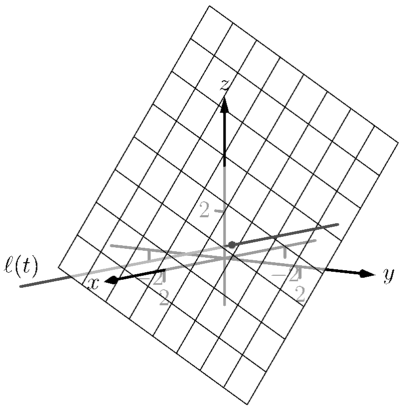


Figure 12.6.14 Illustrating the intersection of a line and a plane in Example 12.6.13

youtu.be/watch?v=DT4JMy5Ak54

Figure 12.6.15 Video introduction to Subsection 12.6.1

12.6.1 Distances

Just as it was useful to find distances between points and lines in the previous section, it is also often necessary to find the distance from a point to a plane.

Consider Figure 12.6.16, where a plane with normal vector $\vec{n}$ is sketched containing a point P and a point Q, not on the plane, is given. We measure the distance from Q to the plane by measuring the length of the projection of $\overrightarrow{PQ}$ onto $\vec{n}$. That is, we want:

$$\left|\left| \text{ proj }_{\vec{n}} \overrightarrow{PQ} \right|\right| = \left|\left| \frac{\vec{n}\cdot\overrightarrow{PQ}}{\|\vec{n}\|^2}\vec{n} \right|\right| = \frac{\left|\vec{n}\cdot\overrightarrow{PQ}\right|}{\|\vec{n}\|} \tag{12.6.4}$$

Equation (12.6.4) is important as it does more than just give the distance between a point and a plane. We will see how it allows us to find several other distances as well: the distance between parallel planes and the distance from a line and a plane. Because Equation (12.6.4) is important, we restate it as a Key Idea.

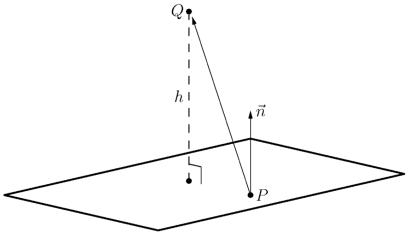


Figure 12.6.16 Illustrating finding the distance from a point to a plane

Key Idea 12.6.17 Distance from a Point to a Plane.

Let a plane with normal vector $\vec{n}$ be given, and let Q be a point. The distance h from Q to the plane is

$$h = \frac{\left|\vec{n}\cdot\overrightarrow{PQ}\right|}{\|\vec{n}\|},$$

where P is any point in the plane.

Example 12.6.18 Distance between a point and a plane.

Find the distance between the point $Q = (2, 1, 4)$ and the plane with equation $2x - 5y + 6z = 9$.

Solution. Using the equation of the plane, we find the normal vector $\vec{n} = \langle 2, -5, 6\rangle$. To find a point on the plane, we can let x and y be anything we choose, then let z be whatever satisfies the equation. Letting x and y be 0 seems simple; this makes $z = 1.5$. Thus we let $P = \langle 0, 0, 1.5\rangle$, and $\overrightarrow{PQ} = \langle 2, 1, 2.5\rangle$.

The distance h from Q to the plane is given by Key Idea 12.6.17:

$$\begin{aligned} h &= \frac{\left|\vec{n}\cdot\overrightarrow{PQ}\right|}{\|\vec{n}\|} \\ &= \frac{|\langle 2,-5,6\rangle\cdot\langle 2,1,2.5\rangle|}{\|\langle 2,-5,6\rangle\|} \\ &= \frac{|14|}{\sqrt{65}} \\ &\approx 1.74. \end{aligned}$$

Video Solution

youtu.be/watch?v=SqA4OLUmoFk

We can use Key Idea 12.6.17 to find other distances. Given two parallel planes, we can find the distance between these planes by letting P be a point on one plane and Q a point on the other. If ℓ is a line parallel to a plane, we can use the Key Idea to find the distance between them as well: again, let P be a point in the plane and let Q be any point on the line. (One can also use Key Idea 12.5.17.) The Exercise section contains problems of these types.

These past two sections have not explored lines and planes in space as an exercise of mathematical curiosity. However, there are many, many applications of these fundamental concepts. Complex shapes can be modeled (or, *approximated*) using planes. For instance, part of the exterior of an aircraft may have a complex, yet smooth, shape, and engineers will want to know how air flows across this piece as well as how heat might build up due to air friction. Many equations that help determine air flow and heat dissipation are difficult to apply to arbitrary surfaces, but simple to apply to planes. By approximating a surface with millions of small planes one can more readily model the needed behavior.

12.6.2 Exercises

Terms and Concepts

1. In order to find the equation of a plane, what two pieces of information must one have?

2. What is the relationship between a plane and one of its normal vectors?

Problems

Exercise Group. In the following exercises, give any two points in the given plane.

3. $2x - 4y + 7z = 2$

4. List any two points in the plane with equation $3(x+2) + 5(y-9) - 4z = 0$.

5. $x = 2$

6. List any two points in the plane with equation $4(y+2) - (z-6) = 0$.

Exercise Group. In the following exercises, give the equation of the described plane in standard and general forms.

7. Passes through $(2,3,4)$ and has normal vector $\vec{n} = \langle 3, -1, 7\rangle$.

8. A plane passes through $(1,3,5)$ and has normal vector $\vec{n} = \langle 0, 2, 4\rangle$.

9. Passes through the points $(1,2,3)$, $(3,-1,4)$ and $(1,0,1)$.

10. A plane passes through the points $(5,3,8)$, $(6,4,9)$ and $(3,3,3)$.

11. Contains the intersecting lines $\vec{\ell}_1(t) = \langle 2,1,2\rangle + t\langle 1,2,3\rangle$ and $\vec{\ell}_2(t) = \langle 2,1,2\rangle + t\langle 2,5,4\rangle$.

12. A plane contains the intersecting lines $\vec{\ell}_1(t) = \langle 5,0,3\rangle + t\langle -1,1,1\rangle$ and $\vec{\ell}_2(t) = \langle 1,4,7\rangle + t\langle 3,0,-3\rangle$.

13. Contains the parallel lines $\vec{\ell}_1(t) = \langle 1,1,1\rangle + t\langle 1,2,3\rangle$ and $\vec{\ell}_2(t) = \langle 1,1,2\rangle + t\langle 1,2,3\rangle$.

14. A plane contains the parallel lines $\vec{\ell}_1(t) = \langle 1,1,1\rangle + t\langle 4,1,3\rangle$ and $\vec{\ell}_2(t) = \langle 2,2,2\rangle + t\langle 4,1,3\rangle$.

15. Contains the point $(2,-6,1)$ and the line $\vec{\ell}(t) = \begin{cases} x = 2 + 5t \\ y = 2 + 2t \\ z = -1 + 2t \end{cases}$

16. A plane contains the point $(5,7,3)$ and the line $\vec{\ell}(t) = \begin{cases} x &= t \\ y &= t \\ z &= t \end{cases}$.

17. A plane contains the point $(5,7,3)$ and is orthogonal to the line $\vec{\ell}(t) = \langle 4,5,6\rangle + t\langle 1,1,1\rangle$.

18. A plane contains the point $(4,1,1)$ and is orthogonal to the line $\begin{cases} x &= 4 + 4t \\ y &= 1 + t \\ z &= 1 + t \end{cases}$.

19. A plane contains the point $(-4,7,2)$ and is parallel to the plane $3(x-2) + 8(y+1) - 10z = 0$.

20. A plane contains the point $(1,2,3)$ and is parallel to the plane $x = 5$.

Exercise Group. In the following exercises, give the equation of the line that is the intersection of the given planes.

21. $p1:\ 3(x-2) + (y-1) + 4z = 0$, and $p2:\ 2(x-1) - 2(y+3) + 6(z-1) = 0$.

22. Give the equation of the line (in vector form) that is the intersection of the planes $5(x-5) + 2(y+2) + 4(z-1) = 0$, and $3x - 4(y-1) + 2(z-1) = 0$.

Exercise Group. Find the point of intersection between the line and the plane.

23. line: $\langle 5,1,-1\rangle + t\langle 2,2,1\rangle$, plane: $5x - y - z = -3$

24. Find the point of intersection between the line $\langle 4,1,0\rangle + t\langle 1,0,-1\rangle$ and the plane $3x + y - 2z = 8$. (If the plane contains the line, enter "the entire line". If the line does not intersect the plane at all, enter "none".)

25. line: $\langle 1,2,3\rangle + t\,\langle 3,5,-1\rangle$,
plane: $3x-2y-z=4$

26. Find the point of intersection between the line $\langle 1,2,3\rangle + t\,\langle 3,5,-1\rangle$ and the plane $3x-2y-z=-4$. (If the plane contains the line, enter "the entire line". If the line does not intersect the plane at all, enter "none".)

Exercise Group. Find the indicated distance.

27. The distance from the point $(1,2,3)$ to the plane $3(x-1)+(y-2)+5(z-2)=0$.

28. Find the distance from the point $(2,6,2)$ to the plane $2(x-1)-y+4(z+1)=0$.

29. The distance between the parallel planes $x+y+z=0$ and $(x-2)+(y-3)+(z+4)=0$

30. Find the distance between the parallel planes $2(x-1)+2(y+1)+(z-2)=0$ and $2(x-3)+2(y-1)+(z-3)=0$.

31. Show why if the point Q lies in a plane, then the distance formula correctly gives the distance from the point to the plane as 0.

32. How is Exercise 12.5.30 in Section 12.5 easier to answer once we have an understanding of planes?

Chapter 13

Vector Valued Functions

In the previous chapter, we learned about vectors and were introduced to the power of vectors within mathematics. In this chapter, we'll build on this foundation to define functions whose input is a real number and whose output is a vector. We'll see how to graph these functions and apply calculus techniques to analyze their behavior. Most importantly, we'll see *why* we are interested in doing this: we'll see beautiful applications to the study of moving objects.

13.1 Vector-Valued Functions

We are very familiar with *real valued functions*, that is, functions whose output is a real number. This section introduces *vector-valued functions* — functions whose output is a vector.

youtu.be/watch?v=MJBP_8ZxiwM

Figure 13.1.1 Video introduction to Section 13.1

Definition 13.1.2 Vector-Valued Functions.

A **vector-valued function** is a function of the form

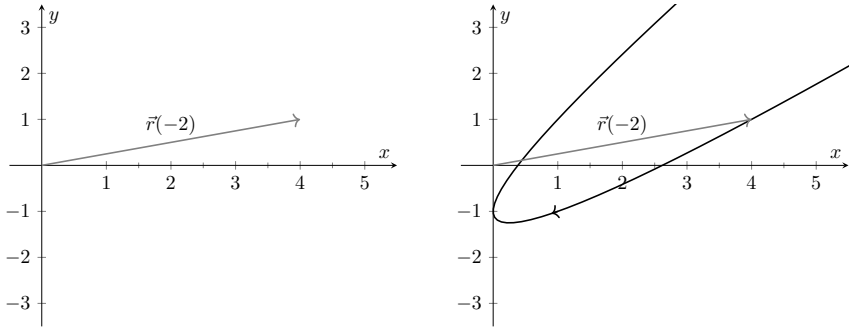
$$\vec{r}(t) = \langle\, f(t), g(t) \,\rangle \text{ or } \vec{r}(t) = \langle\, f(t), g(t), h(t) \,\rangle\,,$$

where f, g and h are real valued functions.

The **domain** of $\vec{r}$ is the set of all values of t for which $\vec{r}(t)$ is defined. The **range** of $\vec{r}$ is the set of all possible output vectors $\vec{r}(t)$.

13.1.1 Evaluating and Graphing Vector-Valued Functions

Evaluating a vector-valued function at a specific value of t is straightforward; simply evaluate each component function at that value of t. For instance, if $\vec{r}(t) = \left\langle t^2, t^2 + t - 1 \right\rangle$, then $\vec{r}(-2) = \langle 4, 1 \rangle$. We can sketch this vector, as is done in Figure 13.1.3(a). Plotting lots of vectors is cumbersome, though, so generally we do not sketch the whole vector but just the terminal point. The *graph* of a vector-valued function is the set of all terminal points of $\vec{r}(t)$, where the initial point of each vector is always the origin. In Figure 13.1.3(b) we sketch the graph of $\vec{r}$; we can indicate individual points on the graph with their respective vector, as shown.



(a) **(b)**

Figure 13.1.3 Sketching the graph of a vector-valued function

Vector-valued functions are closely related to parametric equations of graphs. While in both methods we plot points $\big(x(t), y(t)\big)$ or $\big(x(t), y(t), z(t)\big)$ to produce a graph, in the context of vector-valued functions each such point represents a vector. The implications of this will be more fully realized in the next section as we apply calculus ideas to these functions.

Video Solution

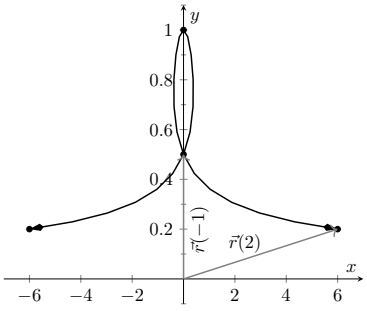
youtu.be/watch?v=wXzQo0ce3Us

Example 13.1.4 Graphing vector-valued functions.

Graph $\vec{r}(t) = \left\langle t^3 - t, \dfrac{1}{t^2+1} \right\rangle$, for $-2 \le t \le 2$. Sketch $\vec{r}(-1)$ and $\vec{r}(2)$.

Solution. We start by making a table of t, x and y values as shown in Figure 13.1.5(a). Plotting these points gives an indication of what the graph looks like. In Figure 13.1.5(b), we indicate these points and sketch the full graph. We also highlight $\vec{r}(-1)$ and $\vec{r}(2)$ on the graph.

t	$t^3 - t$	$\dfrac{1}{t^2+1}$
-2	-6	1/5
-1	0	1/2
0	0	1
1	0	1/2
2	6	1/5



(a) **(b)**

Figure 13.1.5 Sketching the vector-valued function of Example 13.1.4

Video Solution

youtu.be/watch?v=N1qqqHWR6yw

Example 13.1.6 Graphing vector-valued functions.

Graph $\vec{r}(t) = \langle \cos(t), \sin(t), t \rangle$ for $0 \le t \le 4\pi$.

Solution. We can again plot points, but careful consideration of this function is very revealing. Momentarily ignoring the third component, we see the x and y components trace out a circle of radius 1 centered at the origin. Noticing that the z component is t, we see that as the graph winds around the z-axis, it is also increasing at a constant rate in

the positive z direction, forming a spiral. This is graphed in Figure 13.1.7. In the graph $\vec{r}(7\pi/4) \approx (0.707, -0.707, 5.498)$ is highlighted to help us understand the graph.

13.1.2 Algebra of Vector-Valued Functions

Definition 13.1.8 Operations on Vector-Valued Functions.

Let $\vec{r}_1(t) = \langle f_1(t), g_1(t)\rangle$ and $\vec{r}_2(t) = \langle f_2(t), g_2(t)\rangle$ be vector-valued functions in $\mathbb{R}^2$ and let c be a scalar. Then:

1. $\vec{r}_1(t) \pm \vec{r}_2(t) = \langle\, f_1(t) \pm f_2(t), g_1(t) \pm g_2(t)\,\rangle$.
2. $c\vec{r}_1(t) = \langle\, cf_1(t), cg_1(t)\,\rangle$.

A similar definition holds for vector-valued functions in $\mathbb{R}^3$.

This definition states that we add, subtract and scale vector-valued functions component-wise. Combining vector-valued functions in this way can be very useful (as well as create interesting graphs).

youtu.be/watch?v=5jygsmDFjZw

Figure 13.1.9 Video presentation of Definition 13.1.8

Example 13.1.10 Adding and scaling vector-valued functions.

Let $\vec{r}_1(t) = \langle\, 0.2t, 0.3t\,\rangle$, $\vec{r}_2(t) = \langle\, \cos(t), \sin(t)\,\rangle$ and $\vec{r}(t) = \vec{r}_1(t) + \vec{r}_2(t)$. Graph $\vec{r}_1(t)$, $\vec{r}_2(t)$, $\vec{r}(t)$ and $5\vec{r}(t)$ on $-10 \le t \le 10$.

Solution. We can graph $\vec{r}_1$ and $\vec{r}_2$ easily by plotting points (or just using technology). Let's think about each for a moment to better understand how vector-valued functions work.

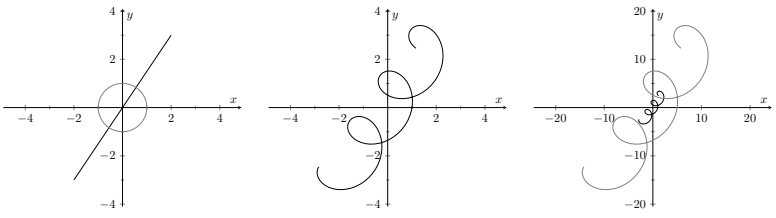
We can rewrite $\vec{r}_1(t) = \langle\, 0.2t, 0.3t\,\rangle$ as $\vec{r}_1(t) = t\,\langle 0.2, 0.3\rangle$. That is, the function $\vec{r}_1$ scales the vector $\langle 0.2, 0.3\rangle$ by t. This scaling of a vector produces a line in the direction of $\langle 0.2, 0.3\rangle$.

We are familiar with $\vec{r}_2(t) = \langle\, \cos(t), \sin(t)\,\rangle$; it traces out a circle, centered at the origin, of radius 1. Figure 13.1.11(a) graphs $\vec{r}_1(t)$ and $\vec{r}_2(t)$.

Adding $\vec{r}_1(t)$ to $\vec{r}_2(t)$ produces $\vec{r}(t) = \langle\, \cos(t) + 0.2t, \sin(t) + 0.3t\,\rangle$, graphed in Figure 13.1.11(b). The linear movement of the line combines with the circle to create loops that move in the direction of $\langle 0.2, 0.3\rangle$. (We encourage the reader to experiment by changing $\vec{r}_1(t)$ to $\langle 2t, 3t\rangle$, etc., and observe the effects on the loops.)

Video Solution

youtu.be/watch?v=6A3op7KYppw



(a) **(b)** **(c)**

Figure 13.1.11 Graphing the functions in Example 13.1.10

Multiplying $\vec{r}(t)$ by 5 scales the function by 5, producing $5\vec{r}(t) = \langle 5\cos(t) + t, 5\sin(t) + 1.5t\rangle$, which is graphed in Figure 13.1.11(c) along with $\vec{r}(t)$. The new function is "5 times bigger" than $\vec{r}(t)$. Note how the

graph of $5\vec{r}(t)$ in Figure 13.1.11(c) looks identical to the graph of $\vec{r}(t)$ in Figure 13.1.11(b). This is due to the fact that the x and y bounds of the plot in Figure 13.1.11(c) are exactly 5 times larger than the bounds in Figure 13.1.11(b).

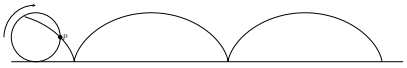


Figure 13.1.13 Tracing a cycloid

Example 13.1.12 Adding and scaling vector-valued functions.

A *cycloid* is a graph traced by a point p on a rolling circle, as shown in Figure 13.1.13. Find an equation describing the cycloid, where the circle has radius 1.

Solution. This problem is not very difficult if we approach it in a clever way. We start by letting $\vec{p}(t)$ describe the position of the point p on the circle, where the circle is centered at the origin and only rotates clockwise (i.e., it does not roll). This is relatively simple given our previous experiences with parametric equations; $\vec{p}(t) = \langle \cos(t), -\sin(t)\rangle$.

We now want the circle to roll. We represent this by letting $\vec{c}(t)$ represent the location of the center of the circle. It should be clear that the y component of $\vec{c}(t)$ should be 1; the center of the circle is always going to be 1 if it rolls on a horizontal surface.

The x component of $\vec{c}(t)$ is a linear function of t: $f(t) = mt$ for some scalar m. When $t = 0$, $f(t) = 0$ (the circle starts centered on the y-axis). When $t = 2\pi$, the circle has made one complete revolution, traveling a distance equal to its circumference, which is also 2π. This gives us a point on our line $f(t) = mt$, the point $(2\pi, 2\pi)$. It should be clear that $m = 1$ and $f(t) = t$. So $\vec{c}(t) = \langle t, 1\rangle$.

We now combine $\vec{p}$ and $\vec{c}$ together to form the equation of the cycloid: $\vec{r}(t) = \vec{p}(t) + \vec{c}(t) = \langle \cos(t) + t, -\sin(t) + 1\rangle$, which is graphed in Figure 13.1.14.

Video Solution

youtu.be/watch?v=BH9FZMTdgQw

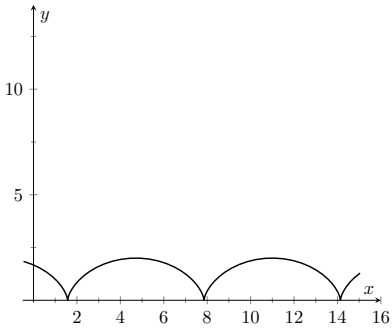


Figure 13.1.14 The cycloid in Example 13.1.12

13.1.3 Displacement

A vector-valued function $\vec{r}(t)$ is often used to describe the position of a moving object at time t. At $t = t_0$, the object is at $\vec{r}(t_0)$; at $t = t_1$, the object is at $\vec{r}(t_1)$. Knowing the locations $\vec{r}(t_0)$ and $\vec{r}(t_1)$ give no indication of the path taken between them, but often we only care about the difference of the locations, $\vec{r}(t_1) - \vec{r}(t_0)$, the *displacement*.

Definition 13.1.15 Displacement.

Let $\vec{r}(t)$ be a vector-valued function and let $t_0 < t_1$ be values in the domain. The **displacement** $\vec{d}$ of $\vec{r}$, from $t = t_0$ to $t = t_1$, is

$$\vec{d} = \vec{r}(t_1) - \vec{r}(t_0).$$

When the displacement vector is drawn with initial point at $\vec{r}(t_0)$, its terminal point is $\vec{r}(t_1)$. We think of it as the vector which points from a starting position to an ending position.

Example 13.1.17 Finding and graphing displacement vectors.

Let $\vec{r}(t) = \left\langle \cos(\frac{\pi}{2}t), \sin(\frac{\pi}{2}t)\right\rangle$. Graph $\vec{r}(t)$ on $-1 \le t \le 1$, and find the displacement of $\vec{r}(t)$ on this interval.

youtu.be/watch?v=LfOpC91t7H8

Figure 13.1.16 Video presentation of Definition 13.1.15

Solution. The function $\vec{r}(t)$ traces out the unit circle, though at a different rate than the "usual" $\langle \cos(t), \sin(t) \rangle$ parametrization. At $t_0 = -1$, we have $\vec{r}(t_0) = \langle 0, -1 \rangle$; at $t_1 = 1$, we have $\vec{r}(t_1) = \langle 0, 1 \rangle$. The displacement of $\vec{r}(t)$ on $[-1, 1]$ is thus $\vec{d} = \langle 0, 1 \rangle - \langle 0, -1 \rangle = \langle 0, 2 \rangle$.

A graph of $\vec{r}(t)$ on $[-1, 1]$ is given in Figure 13.1.18, along with the displacement vector $\vec{d}$ on this interval.

Measuring displacement makes us contemplate related, yet very different, concepts. Considering the semi-circular path the object in Example 13.1.17 took, we can quickly verify that the object ended up a distance of 2 units from its initial location. That is, we can compute $\left\| \vec{d} \right\| = 2$. However, measuring *distance from the starting point* is different from measuring *distance traveled*. Being a semi-circle, we can measure the distance traveled by this object as $\pi \approx 3.14$ units. Knowing *distance from the starting point* allows us to compute *average rate of change*.

Definition 13.1.19 Average Rate of Change.

Let $\vec{r}(t)$ be a vector-valued function, where each of its component functions is continuous on its domain, and let $t_0 < t_1$. The **average rate of change** of $\vec{r}(t)$ on $[t_0, t_1]$ is

$$\text{average rate of change } = \frac{\vec{r}(t_1) - \vec{r}(t_0)}{t_1 - t_0}.$$

Example 13.1.20 Average rate of change.

Let $\vec{r}(t) = \left\langle \cos(\frac{\pi}{2}t), \sin(\frac{\pi}{2}t) \right\rangle$ as in Example 13.1.17. Find the average rate of change of $\vec{r}(t)$ on $[-1, 1]$ and on $[-1, 5]$.

Solution. We computed in Example 13.1.17 that the displacement of $\vec{r}(t)$ on $[-1, 1]$ was $\vec{d} = \langle 0, 2 \rangle$. Thus the average rate of change of $\vec{r}(t)$ on $[-1, 1]$ is:

$$\frac{\vec{r}(1) - \vec{r}(-1)}{1 - (-1)} = \frac{\langle 0, 2 \rangle}{2} = \langle 0, 1 \rangle .$$

We interpret this as follows: the object followed a semi-circular path, meaning it moved towards the right then moved back to the left, while climbing slowly, then quickly, then slowly again. *On average*, however, it progressed straight up at a constant rate of $\langle 0, 1 \rangle$ per unit of time.

We can quickly see that the displacement on $[-1, 5]$ is the same as on $[-1, 1]$, so $\vec{d} = \langle 0, 2 \rangle$. The average rate of change is different, though:

$$\frac{\vec{r}(5) - \vec{r}(-1)}{5 - (-1)} = \frac{\langle 0, 2 \rangle}{6} = \langle 0, 1/3 \rangle .$$

As it took "3 times as long" to arrive at the same place, this average rate of change on $[-1, 5]$ is $1/3$ the average rate of change on $[-1, 1]$.

We considered average rates of change in Sections 1.1 and Section 2.1 as we studied limits and derivatives. The same is true here; in the following section we apply calculus concepts to vector-valued functions as we find limits, derivatives, and integrals. Understanding the average rate of change will give us an understanding of the derivative; displacement gives us one application of integration.

Video Solution

youtu.be/watch?v=eGsIv3hYlak

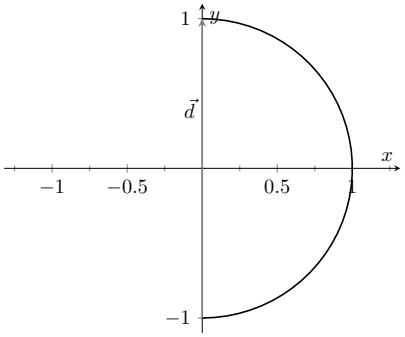


Figure 13.1.18 Graphing the displacement of a position function in Example 13.1.17

Video Solution

youtu.be/watch?v=REA-v1g68bo

13.1.4 Exercises

Terms and Concepts

1. Vector-valued functions are closely related to _____ _____ of graphs.

2. When sketching vector-valued functions, technically one isn't graphing points, but rather _____.

3. It can be useful to think of _____ as a vector that points from a starting position to an ending position.

4. In the context of vector-valued functions, average rate of change is _____ divided by time.

Problems

Exercise Group. In the following exercises, sketch the vector-valued function on the given interval.

5. $\vec{r}(t) = \langle t^2, t^2 - 1\rangle$, for $-2 \le t \le 2$.

6. $\vec{r}(t) = \langle t^2, t^3\rangle$, for $-2 \le t \le 2$.

7. $\vec{r}(t) = \langle 1/t, 1/t^2\rangle$, for $-2 \le t \le 2$.

8. $\vec{r}(t) = \langle \frac{1}{10}t^2, \sin(t)\rangle$, for $-2\pi \le t \le 2\pi$.

9. $\vec{r}(t) = \langle \frac{1}{10}t^2, \sin(t)\rangle$, for $-2\pi \le t \le 2\pi$.

10. $\vec{r}(t) = \langle 3\sin(\pi t), 2\cos(\pi t)\rangle$, on $[0, 2]$.

11. $\vec{r}(t) = \langle 3\cos(t), 2\sin(2t)\rangle$, on $[0, 2\pi]$.

12. $\vec{r}(t) = \langle 2\sec(t), \tan(t)\rangle$, on $[-\pi, \pi]$.

Exercise Group. In the following exercises, sketch the vector-valued function on the given interval in $\mathbb{R}^3$. Technology may be useful in creating the sketch.

13. $\vec{r}(t) = \langle 2\cos(t), t, 2\sin(t)\rangle$, on $[0, 2\pi]$.

14. $\vec{r}(t) = \langle 3\cos(t), \sin(t), t/\pi\rangle$ on $[0, 2\pi]$.

15. $\vec{r}(t) = \langle \cos(t), \sin(t), \sin(t)\rangle$ on $[0, 2\pi]$.

16. $\vec{r}(t) = \langle \cos(t), \sin(t), \sin(2t)\rangle$ on $[0, 2\pi]$.

Exercise Group. In the following exercises, find $\|\vec{r}(t)\|$.

17. If $\vec{r}(t) = \langle t, t^2\rangle$, then $\|\vec{r}(t)\| =$_____.

18. $\vec{r}(t) = \langle 5\cos(t), 3\sin(t)\rangle$.

19. If $\vec{r}(t) = \langle 2\cos(t), 2\sin(t), t\rangle$, then $\|\vec{r}(t)\| =$_____.

20. $\vec{r}(t) = \langle \cos(t), t, t^2\rangle$.

Exercise Group. Create a vector-valued function whose graph matches the given description.

21. A circle of radius 2, centered at $(1, 2)$, traced counter-clockwise once at constant speed on $[0, 2\pi)$.

22. A circle of radius 3, centered at $(5, 5)$, traced clockwise once on $[0, 2\pi]$.

23. An ellipse, centered at $(0, 0)$ with vertical major axis of length 10 and minor axis of length 3, traced once counter-clockwise on $[0, 2\pi]$.

24. An ellipse, centered at $(3, -2)$ with horizontal major axis of length 6 and minor axis of length 4, traced once clockwise on $[0, 2\pi]$.

25. A line through $(2, 3)$ with a slope of 5.

26. A line through $(1, 5)$ with a slope of $-1/2$.

27. The line through points $(1, 2, 3)$ and $(4, 5, 6)$, where $\vec{r}(0) = \langle 1, 2, 3\rangle$ and $\vec{r}(1) = \langle 4, 5, 6\rangle$.

28. The line through points $(1, 2)$ and $(4, 4)$, where $\vec{r}(0) = \langle 1, 2\rangle$ and $\vec{r}(1) = \langle 4, 4\rangle$.

29. A vertically oriented helix with radius of 2 that starts at $(2, 0, 0)$ and ends at $(2, 0, 4\pi)$ after one revolution on $[0, 2\pi]$.

30. A vertically oriented helix with radius of 3 that starts at $(3, 0, 0)$ and ends at $(3, 0, 3)$ after 2 revolutions on $[0, 1]$.

Exercise Group. Find the average rate of change of $\vec{r}(t)$ on the given interval.

31. $\vec{r}(t) = \langle t, t^2\rangle$ on $[-2, 2]$.

32. $\vec{r}(t) = \langle t, t + \sin(t)\rangle$ on $[0, 2\pi]$.

33. $\vec{r}(t) = \langle 3\cos(t), 2\sin(t), t\rangle$ on $[0, 2\pi]$.

34. $\vec{r}(t) = \langle t, t^2, t^3\rangle$ on $[-1, 3]$.

13.2 Calculus and Vector-Valued Functions

The previous section introduced us to a new mathematical object, the vector-valued function. We now apply calculus concepts to these functions. We start with the limit, then work our way through derivatives to integrals.

13.2.1 Limits of Vector-Valued Functions

The initial definition of the limit of a vector-valued function is a bit intimidating, as was the definition of the limit in Definition 1.2.2. The theorem following the definition shows that in practice, taking limits of vector-valued functions is no more difficult than taking limits of real-valued functions.

We can define one-sided limits in a manner very similar to Definition 13.2.1.

Definition 13.2.1 Limits of Vector-Valued Functions.

Let I be an open interval containing c, and let $\vec{r}(t)$ be a vector-valued function defined on I, except possibly at c. The **limit of $\vec{r}(t)$, as t approaches c, is $\vec{L}$**, expressed as

$$\lim_{t\to c} \vec{r}(t) = \vec{L},$$

means that given any $\varepsilon > 0$, there exists a $\delta > 0$ such that for all $t \neq c$, if $|t-c| < \delta$, we have $\left\|\vec{r}(t) - \vec{L}\right\| < \varepsilon$.

Note how the measurement of distance between real numbers is the absolute value of their difference; the measure of distance between vectors is the vector norm, or magnitude, of their difference.

Theorem 13.2.2 states that we can compute limits of vector-valued functions component-wise.

Theorem 13.2.2 Limits of Vector-Valued Functions.

1. *Let $\vec{r}(t) = \langle\, f(t), g(t) \,\rangle$ be a vector-valued function in $\mathbb{R}^2$ defined on an open interval I containing c, except possibly at c. Then*

$$\lim_{t\to c} \vec{r}(t) = \left\langle \lim_{t\to c} f(t)\,,\ \lim_{t\to c} g(t) \right\rangle.$$

2. *Let $\vec{r}(t) = \langle\, f(t), g(t), h(t) \,\rangle$ be a vector-valued function in $\mathbb{R}^3$ defined on an open interval I containing c, except possibly at c. Then*

$$\lim_{t\to c} \vec{r}(t) = \left\langle \lim_{t\to c} f(t)\,,\ \lim_{t\to c} g(t)\,,\ \lim_{t\to c} h(t) \right\rangle$$

youtu.be/watch?v=MS6iGW1AQ2c

Figure 13.2.3 Video presentation of Definition 13.2.1 and Theorem 13.2.2

Example 13.2.4 Finding limits of vector-valued functions.

Let $\vec{r}(t) = \left\langle \dfrac{\sin(t)}{t},\ t^2 - 3t + 3,\ \cos(t) \right\rangle$. Find $\lim\limits_{t\to 0} \vec{r}(t)$.

Solution. We apply the theorem and compute limits component-wise.

$$\begin{aligned}\lim_{t\to 0} \vec{r}(t) &= \left\langle \lim_{t\to 0} \frac{\sin(t)}{t}\,,\ \lim_{t\to 0} t^2 - 3t + 3\,,\ \lim_{t\to 0} \cos(t) \right\rangle \\ &= \langle 1, 3, 1 \rangle\,.\end{aligned}$$

Video Solution

youtu.be/watch?v=FCFNyv2V8yk

13.2.2 Continuity

Definition 13.2.5 Continuity of Vector-Valued Functions.

Let $\vec{r}(t)$ be a vector-valued function defined on an open interval I containing c.

1. $\vec{r}(t)$ is *continuous at* c if $\lim_{t\to c} \vec{r}(t) = r(c)$.
2. If $\vec{r}(t)$ is continuous at all c in I, then $\vec{r}(t)$ is *continuous on* I.

Using one-sided limits, we can also define continuity on closed intervals as done before.

We again have a theorem that lets us evaluate continuity component-wise.

Theorem 13.2.6 Continuity of Vector-Valued Functions.

Let $\vec{r}(t)$ be a vector-valued function defined on an open interval I containing c. Then $\vec{r}(t)$ is continuous at c if, and only if, each of its component functions is continuous at c.

youtu.be/watch?v=5sX-TKcgnA0

Figure 13.2.7 Video presentation of Subsection 13.2.2

Example 13.2.8 Evaluating continuity of vector-valued functions.

Let $\vec{r}(t) = \left\langle \dfrac{\sin(t)}{t},\ t^2 - 3t + 3,\ \cos(t) \right\rangle$. Determine whether $\vec{r}$ is continuous at $t = 0$ and $t = 1$.

Solution. While the second and third components of $\vec{r}(t)$ are defined at $t = 0$, the first component, $(\sin(t))/t$, is not. Since the first component is not even defined at $t = 0$, $\vec{r}(t)$ is not defined at $t = 0$, and hence it is not continuous at $t = 0$.

At $t = 1$ each of the component functions is continuous. Therefore $\vec{r}(t)$ is continuous at $t = 1$.

13.2.3 Derivatives

Consider a vector-valued function $\vec{r}$ defined on an open interval I containing t_0 and t_1. We can compute the displacement of $\vec{r}$ on $[t_0, t_1]$, as shown in Figure 13.2.9(a). Recall that dividing the displacement vector by $t_1 - t_0$ gives the average rate of change on $[t_0, t_1]$, as shown in Figure 13.2.9(b).

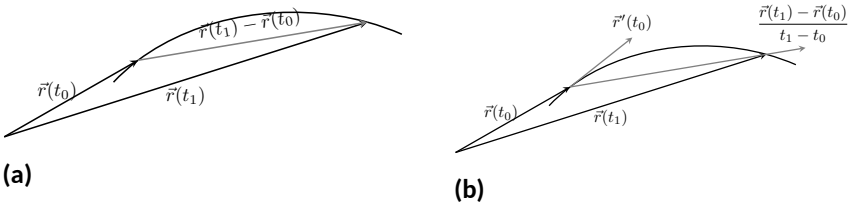


Figure 13.2.9 Illustrating displacement, leading to an understanding of the derivative of vector-valued functions

The *derivative* of a vector-valued function is a measure of the *instantaneous* rate of change, measured by taking the limit as the length of $[t_0, t_1]$ goes to 0. Instead of thinking of an interval as $[t_0, t_1]$, we think of it as $[c, c + h]$ for some value of h (hence the interval has length h). The *average* rate of change is

$$\frac{\vec{r}(c + h) - \vec{r}(c)}{h}$$

for any value of $h \neq 0$. We take the limit as $h \to 0$ to measure the instantaneous rate of change; this is the derivative of $\vec{r}$.

Definition 13.2.10 Derivative of a Vector-Valued Function.

Let $\vec{r}(t)$ be continuous on an open interval I containing c.

1. The **derivative** of $\vec{r}$ at $t = c$ is
$$\vec{r}\,'(c) = \lim_{h\to 0} \frac{\vec{r}(c+h) - \vec{r}(c)}{h}.$$
2. The **derivative of** $\vec{r}$ is
$$\vec{r}\,'(t) = \lim_{h\to 0} \frac{\vec{r}(t+h) - \vec{r}(t)}{h}.$$

If a vector-valued function has a derivative for all c in an open interval I, we say that $\vec{r}(t)$ is **differentiable** on I.

Once again we might view this definition as intimidating, but recall that we can evaluate limits component-wise. The following theorem verifies that this means we can compute derivatives component-wise as well, making the task not too difficult.

Theorem 13.2.11 Derivatives of Vector-Valued Functions.

1. *Let* $\vec{r}(t) = \langle\, f(t), g(t) \,\rangle$. *Then*
$$\vec{r}\,'(t) = \langle\, f'(t), g'(t) \,\rangle.$$
2. *Let* $\vec{r}(t) = \langle\, f(t), g(t), h(t) \,\rangle$. *Then*
$$\vec{r}\,'(t) = \langle\, f'(t), g'(t), h'(t) \,\rangle.$$

Alternate notations for the derivative of $\vec{r}$ include:
$$\vec{r}\,'(t) = \frac{d}{dt}\big(\,\vec{r}(t)\,\big) = \frac{d\vec{r}}{dt}.$$

Again, using one-sided limits, we can define differentiability on closed intervals. We'll make use of this a few times in this chapter.

youtu.be/watch?v=-o7FIEjwkQs

Figure 13.2.12 Video presentation of Definition 13.2.10 and Theorem 13.2.11

Example 13.2.13 Derivatives of vector-valued functions.

Let $\vec{r}(t) = \langle t^2, t \rangle$.

1. Sketch $\vec{r}(t)$ and $\vec{r}\,'(t)$ on the same axes.
2. Compute $\vec{r}\,'(1)$ and sketch this vector with its initial point at the origin and at $\vec{r}(1)$.

Solution.

1. Theorem 13.2.11 allows us to compute derivatives component-wise, so
$$\vec{r}\,'(t) = \langle 2t, 1 \rangle.$$
$\vec{r}(t)$ and $\vec{r}\,'(t)$ are graphed together in Figure 13.2.14(a). Note how plotting the two of these together, in this way, is not very illuminating. When dealing with real-valued functions, plotting $f(x)$ with $f'(x)$ gave us useful information as we were able to compare f and f' at the same x-values. When dealing with vector-valued functions, it is hard to tell which points on the graph of $\vec{r}\,'$ correspond to which points on the graph of $\vec{r}$.

Video Solution

youtu.be/watch?v=vcTn9uy2Fi8

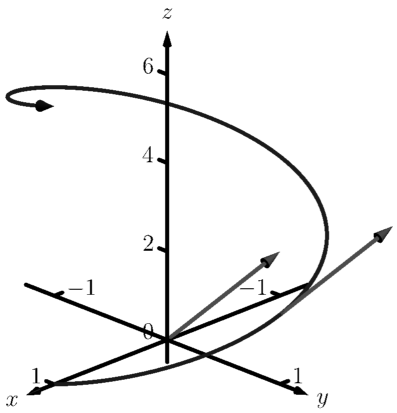
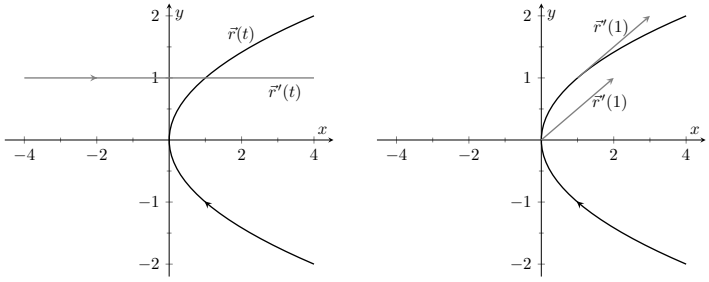


Figure 13.2.16 Viewing a vector-valued function and its derivative at one point

2. We easily compute $\vec{r}'(1) = \langle 2, 1\rangle$, which is drawn in Figure 13.2.14 with its initial point at the origin, as well as at $\vec{r}(1) = \langle 1, 1\rangle$. These are sketched in Figure 13.2.14(b).



(a) **(b)**

Figure 13.2.14 Graphing the derivative of a vector-valued function in Example 13.2.13

Example 13.2.15 Derivatives of vector-valued functions.

Let $\vec{r}(t) = \langle \cos(t), \sin(t), t\rangle$. Compute $\vec{r}'(t)$ and $\vec{r}'(\pi/2)$. Sketch $\vec{r}'(\pi/2)$ with its initial point at the origin and at $\vec{r}(\pi/2)$.

Solution. We compute $\vec{r}'$ as $\vec{r}'(t) = \langle -\sin(t), \cos(t), 1\rangle$. At $t = \pi/2$, we have $\vec{r}'(\pi/2) = \langle -1, 0, 1\rangle$. Figure 13.2.16 shows a graph of $\vec{r}(t)$, with $\vec{r}'(\pi/2)$ plotted with its initial point at the origin and at $\vec{r}(\pi/2)$.

In Examples 13.2.13 and Example 13.2.15, sketching a particular derivative with its initial point at the origin did not seem to reveal anything significant. However, when we sketched the vector with its initial point on the corresponding point on the graph, we did see something significant: the vector appeared to be *tangent* to the graph. We have not yet defined what "tangent" means in terms of curves in space; in fact, we use the derivative to define this term.

youtu.be/watch?v=WwKvTWkgetI

Figure 13.2.18 Video presentation of Definition 13.2.17

Definition 13.2.17 Tangent Vector, Tangent Line.

Let $\vec{r}(t)$ be a differentiable vector-valued function on an open interval I containing c, where $\vec{r}'(c) \neq \vec{0}$.

1. A vector $\vec{v}$ is *tangent to the graph of* $\vec{r}(t)$ *at* $t = c$ if $\vec{v}$ is parallel to $\vec{r}'(c)$.
2. The *tangent line* to the graph of $\vec{r}(t)$ at $t = c$ is the line through $\vec{r}(c)$ with direction parallel to $\vec{r}'(c)$. An equation of the tangent line is
$$\vec{\ell}(t) = \vec{r}(c) + t\,\vec{r}'(c).$$

Video Solution

youtu.be/watch?v=mG-r7_uCzSk

Example 13.2.19 Finding tangent lines to curves in space.

Let $\vec{r}(t) = \langle t, t^2, t^3\rangle$ on $[-1.5, 1.5]$. Find the vector equation of the line tangent to the graph of $\vec{r}$ at $t = -1$.

Solution. To find the equation of a line, we need a point on the line and the line's direction. The point is given by $\vec{r}(-1) = \langle -1, 1, -1\rangle$. (To be

clear, $\langle -1, 1, -1\rangle$ is a *vector*, not a point, but we use the point "pointed to" by this vector.)

The direction comes from $\vec{r}\,'(-1)$. We compute, component-wise, $\vec{r}\,'(t) = \langle 1, 2t, 3t^2\rangle$. Thus $\vec{r}\,'(-1) = \langle 1, -2, 3\rangle$.

The vector equation of the line is $\ell(t) = \langle -1, 1, -1\rangle + t\langle 1, -2, 3\rangle$. This line and $\vec{r}(t)$ are sketched in Figure 13.2.20.

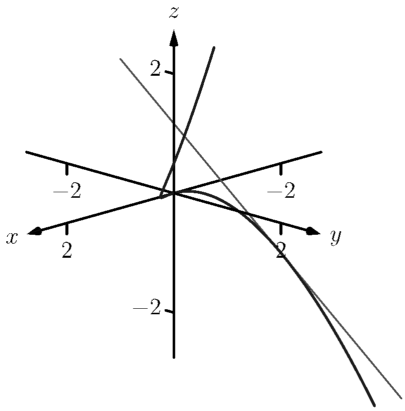


Figure 13.2.20 Graphing a curve in space with its tangent line

Example 13.2.21 Finding tangent lines to curves.

Find the equations of the lines tangent to $\vec{r}(t) = \langle t^3, t^2\rangle$ at $t = -1$ and $t = 0$.

Solution. We find that $\vec{r}\,'(t) = \langle 3t^2, 2t\rangle$. At $t = -1$, we have

$$\vec{r}(-1) = \langle -1, 1\rangle \text{ and } \vec{r}\,'(-1) = \langle 3, -2\rangle,$$

so the equation of the line tangent to the graph of $\vec{r}(t)$ at $t = -1$ is

$$\ell(t) = \langle -1, 1\rangle + t\langle 3, -2\rangle.$$

This line is graphed with $\vec{r}(t)$ in Figure 13.2.22.

At $t = 0$, we have $\vec{r}\,'(0) = \langle 0, 0\rangle = \vec{0}$! This implies that the tangent line "has no direction." We cannot apply Definition 13.2.17, hence cannot find the equation of the tangent line.

Video Solution

youtu.be/watch?v=czvJ5AtLpa4

We were unable to compute the equation of the tangent line to $\vec{r}(t) = \langle t^3, t^2\rangle$ at $t = 0$ because $\vec{r}\,'(0) = \vec{0}$. The graph in Figure 13.2.22 shows that there is a cusp at this point. This leads us to another definition of *smooth*, previously defined by Definition 10.2.22 in Section 10.2.

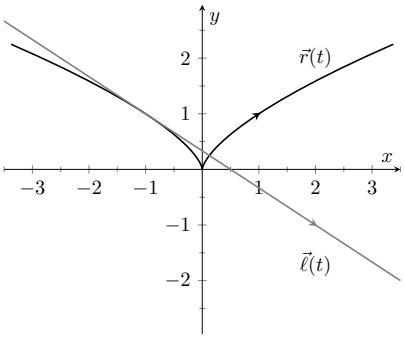


Figure 13.2.22 Graphing $\vec{r}(t)$ and its tangent line in Example 13.2.21

Definition 13.2.23 Smooth Vector-Valued Functions.

Let $\vec{r}(t)$ be a differentiable vector-valued function on an open interval I where $\vec{r}\,'(t)$ is continuous on I. $\vec{r}(t)$ is **smooth** on I if $\vec{r}\,'(t) \neq \vec{0}$ on I.

Having established derivatives of vector-valued functions, we now explore the relationships between the derivative and other vector operations. The following theorem states how the derivative interacts with vector addition and the various vector products.

Theorem 13.2.25 Properties of Derivatives of Vector-Valued Functions.

Let $\vec{r}$ and $\vec{s}$ be differentiable vector-valued functions, let f be a differentiable real-valued function, and let c be a real number.

1. $\dfrac{d}{dt}\Big(\vec{r}(t) \pm \vec{s}(t)\Big) = \vec{r}\,'(t) \pm \vec{s}\,'(t)$
2. $\dfrac{d}{dt}\Big(c\vec{r}(t)\Big) = c\vec{r}\,'(t)$
3. $\dfrac{d}{dt}\Big(f(t)\vec{r}(t)\Big) = f'(t)\vec{r}(t) + f(t)\vec{r}\,'(t)$ Product Rule
4. $\dfrac{d}{dt}\Big(\vec{r}(t) \cdot \vec{s}(t)\Big) = \vec{r}\,'(t) \cdot \vec{s}(t) + \vec{r}(t) \cdot \vec{s}\,'(t)$ Product Rule

youtu.be/watch?v=sSPWKKcGno4

Figure 13.2.24 Video presentation of Definition 13.2.23

youtu.be/watch?v=8ulgeUUc0rY

Figure 13.2.26 Video presentation of Theorem 13.2.25

Video Solution

youtu.be/watch?v=sD1mgVOWQFo

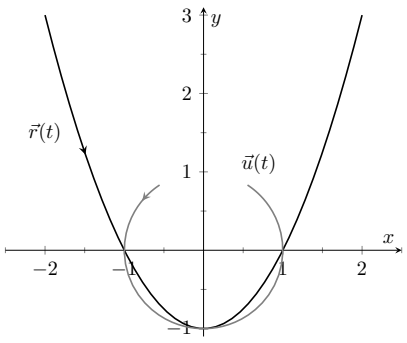


Figure 13.2.28 Graphing $\vec{r}(t)$ and $\vec{u}(t)$ in Example 13.2.27

5. $\dfrac{d}{dt}\Big(\vec{r}(t)\times\vec{s}(t)\Big)=\vec{r}\,'(t)\times\vec{s}(t)+\vec{r}(t)\times\vec{s}\,'(t)$ Product Rule
6. $\dfrac{d}{dt}\Big(\vec{r}\big(f(t)\big)\Big)=\vec{r}\,'\big(f(t)\big)f'(t)$ Chain Rule

Example 13.2.27 Using derivative properties of vector-valued functions.

Let $\vec{r}(t)=\langle t,t^2-1\rangle$ and let $\vec{u}(t)$ be the unit vector that points in the direction of $\vec{r}(t)$.

1. Graph $\vec{r}(t)$ and $\vec{u}(t)$ on the same axes, on $[-2,2]$.
2. Find $\vec{u}\,'(t)$ and sketch $\vec{u}\,'(-2)$, $\vec{u}\,'(-1)$ and $\vec{u}\,'(0)$. Sketch each with initial point the corresponding point on the graph of $\vec{u}$.

Solution.

1. To form the unit vector that points in the direction of $\vec{r}$, we need to divide $\vec{r}(t)$ by its magnitude.

   $$\|\vec{r}(t)\|=\sqrt{t^2+(t^2-1)^2}\Rightarrow\vec{u}(t)=\frac{1}{\sqrt{t^2+(t^2-1)^2}}\langle t,t^2-1\rangle.$$

   $\vec{r}(t)$ and $\vec{u}(t)$ are graphed in Figure 13.2.28. Note how the graph of $\vec{u}(t)$ forms part of a circle; this must be the case, as the length of $\vec{u}(t)$ is 1 for all t.

2. To compute $\vec{u}\,'(t)$, we use Theorem 13.2.25, writing

   $$\vec{u}(t)=f(t)\vec{r}(t),\text{ where }f(t)=\frac{1}{\sqrt{t^2+(t^2-1)^2}}=\big(t^2+(t^2-1)^2\big)^{-1/2}.$$

   (We *could* write

   $$\vec{u}(t)=\left\langle\frac{t}{\sqrt{t^2+(t^2-1)^2}},\frac{t^2-1}{\sqrt{t^2+(t^2-1)^2}}\right\rangle$$

   and then take the derivative. It is a matter of preference; this latter method requires two applications of the Quotient Rule where our method uses the Product and Chain Rules.) We find $f'(t)$ using the Chain Rule:

   $$\begin{aligned}f'(t)&=-\frac{1}{2}\big(t^2+(t^2-1)^2\big)^{-3/2}\big(2t+2(t^2-1)(2t)\big)\\&=-\frac{2t(2t^2-1)}{2\big(\sqrt{t^2+(t^2-1)^2}\big)^3}\end{aligned}$$

   We now find $\vec{u}\,'(t)$ using part 3 of Theorem 13.2.25:

   $$\begin{aligned}\vec{u}\,'(t)&=f'(t)\vec{u}(t)+f(t)\vec{u}\,'(t)\\&=-\frac{2t(2t^2-1)}{2\big(\sqrt{t^2+(t^2-1)^2}\big)^3}\langle t,t^2-1\rangle+\frac{1}{\sqrt{t^2+(t^2-1)^2}}\langle 1,2t\rangle.\end{aligned}$$

This is admittedly very "messy;" such is usually the case when we deal with unit vectors. We can use this formula to compute $\vec{u}\,'(-2)$, $\vec{u}\,'(-1)$ and $\vec{u}\,'(0)$:

$$\vec{u}\,'(-2) = \left\langle -\frac{15}{13\sqrt{13}}, -\frac{10}{13\sqrt{13}} \right\rangle \approx \langle -0.320, -0.213 \rangle$$
$$\vec{u}\,'(-1) = \langle 0, -2 \rangle$$
$$\vec{u}\,'(0) = \langle 1, 0 \rangle$$

Each of these is sketched in Figure 13.2.29. Note how the length of the vector gives an indication of how quickly the circle is being traced at that point. When $t = -2$, the circle is being drawn relatively slow; when $t = -1$, the circle is being traced much more quickly.

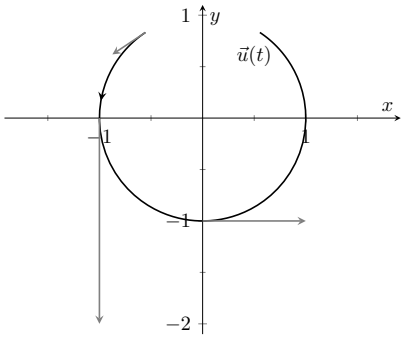


Figure 13.2.29 Graphing some of the derivatives of $\vec{u}(t)$ in Example 13.2.27

It is a basic geometric fact that a line tangent to a circle at a point P is perpendicular to the line passing through the center of the circle and P. This is illustrated in Figure 13.2.29; each tangent vector is perpendicular to the line that passes through its initial point and the center of the circle. Since the center of the circle is the origin, we can state this another way: $\vec{u}\,'(t)$ is orthogonal to $\vec{u}(t)$.

Recall that the dot product serves as a test for orthogonality: if $\vec{u} \cdot \vec{v} = 0$, then $\vec{u}$ is orthogonal to $\vec{v}$. Thus in the above example, $\vec{u}(t) \cdot \vec{u}\,'(t) = 0$.

This is true of any vector-valued function that has a constant length, that is, that traces out part of a circle. It has important implications later on, so we state it as a theorem (and leave its formal proof as an Exercise.)

Theorem 13.2.30 Vector-Valued Functions of Constant Length.

Let $\vec{r}(t)$ be a vector-valued function of constant length that is differentiable on an open interval I. That is, $\|\vec{r}(t)\| = c$ for all t in I; equivalently, $\vec{r}(t) \cdot \vec{r}(t) = c^2$ for all t in I. Then $\vec{r}(t) \cdot \vec{r}\,'(t) = 0$ for all t in I.

13.2.4 Integration

Before formally defining integrals of vector-valued functions, consider the following equation that our calculus experience tells us *should* be true:

$$\int_a^b \vec{r}\,'(t)\,dt = \vec{r}(b) - \vec{r}(a).$$

That is, the integral of a rate of change function should give total change. In the context of vector-valued functions, this total change is displacement. The above equation *is* true; we now develop the theory to show why.

We can define antiderivatives and the indefinite integral of vector-valued functions in the same manner we defined indefinite integrals in Definition 5.1.2. However, we cannot define the definite integral of a vector-valued function as we did in Definition 5.2.6. That definition was based on the signed area between a function $y = f(x)$ and the x-axis. An area-based definition will not be useful in the context of vector-valued functions. Instead, we define the definite integral of a vector-valued function in a manner similar to that of Theorem 5.3.26, utilizing Riemann sums.

Definition 13.2.31 Antiderivatives, Indefinite and Definite Integrals of Vector-Valued Functions.

Let $\vec{r}(t)$ be a continuous vector-valued function on $[a, b]$. An **antiderivative** of $\vec{r}(t)$ is a function $\vec{R}(t)$ such that $\vec{R}'(t) = \vec{r}(t)$.

The set of all antiderivatives of $\vec{r}(t)$ is the **indefinite integral** of $\vec{r}(t)$, denoted by

$$\int \vec{r}(t)\,dt.$$

The **definite integral** of $\vec{r}(t)$ on $[a, b]$ is

$$\int_a^b \vec{r}(t)\,dt = \lim_{\|\Delta t\|\to 0} \sum_{i=1}^{n} \vec{r}(c_i)\Delta t_i,$$

where Δt_i is the length of the ith subinterval of a partition of $[a, b]$, $\|\Delta t\|$ is the length of the largest subinterval in the partition, and c_i is any value in the ith subinterval of the partition.

It is probably difficult to infer meaning from the definition of the definite integral. The important thing to realize from the definition is that it is built upon limits, which we can evaluate component-wise.

The following theorem simplifies the computation of definite integrals; the rest of this section and the following section will give meaning and application to these integrals.

Theorem 13.2.32 Indefinite and Definite Integrals of Vector-Valued Functions.

Let $\vec{r}(t) = \langle f(t), g(t)\rangle$ be a vector-valued function in $\mathbb{R}^2$ that is continuous on $[a, b]$.

1. $\displaystyle\int \vec{r}(t)\,dt = \left\langle \int f(t)\,dt, \int g(t)\,dt \right\rangle$
2. $\displaystyle\int_a^b \vec{r}(t)\,dt = \left\langle \int_a^b f(t)\,dt, \int_a^b g(t)\,dt \right\rangle$

A similar statement holds for vector-valued functions in $\mathbb{R}^3$.

youtu.be/watch?v=GQTMVNKQusc

Figure 13.2.33 Video presentation of Definition 13.2.31 and Theorem 13.2.32

Example 13.2.34 Evaluating a definite integral of a vector-valued function.

Let $\vec{r}(t) = \langle e^{2t}, \sin(t)\rangle$. Evaluate $\displaystyle\int_0^1 \vec{r}(t)\,dt$.

Solution. We follow Theorem 13.2.32.

$$\begin{aligned}\int_0^1 \vec{r}(t)\,dt &= \int_0^1 \langle e^{2t}, \sin(t)\rangle\,dt \\ &= \left\langle \int_0^1 e^{2t}\,dt\,, \int_0^1 \sin(t)\,dt \right\rangle \\ &= \left\langle \frac{1}{2}e^{2t}\Big|_0^1\,, -\cos(t)\Big|_0^1 \right\rangle\end{aligned}$$

Video Solution

youtu.be/watch?v=57pkOzkF7TE

$$= \left\langle \frac{1}{2}(e^2-1)\,, -\cos(1)+1 \right\rangle$$
$$\approx \langle 3.19, 0.460 \rangle\,.$$

Example 13.2.35 Solving an initial value problem.

Let $\vec{r}\,''(t) = \langle 2, \cos(t), 12t\rangle$. Find $\vec{r}(t)$ where:

- $\vec{r}(0) = \langle -7, -1, 2\rangle$ and
- $\vec{r}\,'(0) = \langle 5, 3, 0\rangle$.

Video Solution

youtu.be/watch?v=0-1B6C9jj9k

Solution. Knowing $\vec{r}\,''(t) = \langle 2, \cos(t), 12t\rangle$, we find $\vec{r}\,'(t)$ by evaluating the indefinite integral.

$$\begin{aligned}\int \vec{r}\,''(t)\,dt &= \left\langle \int 2\,dt\,, \int \cos(t)\,dt\,, \int 12t\,dt \right\rangle \\ &= \left\langle 2t + C_1, \sin(t) + C_2, 6t^2 + C_3 \right\rangle \\ &= \left\langle 2t, \sin(t), 6t^2 \right\rangle + \langle C_1, C_2, C_3\rangle \\ &= \left\langle 2t, \sin(t), 6t^2 \right\rangle + \vec{C}.\end{aligned}$$

Note how each indefinite integral creates its own constant which we collect as one constant vector $\vec{C}$. Knowing $\vec{r}\,'(0) = \langle 5, 3, 0\rangle$ allows us to solve for $\vec{C}$:

$$\begin{aligned}\vec{r}\,'(t) &= \left\langle 2t, \sin(t), 6t^2 \right\rangle + \vec{C} \\ \vec{r}\,'(0) &= \langle 0, 0, 0\rangle + \vec{C} \\ \langle 5, 3, 0\rangle &= \vec{C}.\end{aligned}$$

So $\vec{r}\,'(t) = \left\langle 2t, \sin(t), 6t^2 \right\rangle + \langle 5, 3, 0\rangle = \left\langle 2t+5, \sin(t)+3, 6t^2 \right\rangle$. To find $\vec{r}(t)$, we integrate once more.

$$\begin{aligned}\int \vec{r}\,'(t)\,dt &= \left\langle \int 2t+5\,dt, \int \sin(t)+3\,dt, \int 6t^2\,dt \right\rangle \\ &= \left\langle t^2+5t, -\cos(t)+3t, 2t^3 \right\rangle + \vec{C}.\end{aligned}$$

With $\vec{r}(0) = \langle -7, -1, 2\rangle$, we solve for $\vec{C}$:

$$\begin{aligned}\vec{r}(t) &= \left\langle t^2+5t, -\cos(t)+3t, 2t^3 \right\rangle + \vec{C} \\ \vec{r}(0) &= \langle 0, -1, 0\rangle + \vec{C} \\ \langle -7, -1, 2\rangle &= \langle 0, -1, 0\rangle + \vec{C} \\ \langle -7, 0, 2\rangle &= \vec{C}.\end{aligned}$$

So

$$\begin{aligned}\vec{r}(t) &= \left\langle t^2+5t, -\cos(t)+3t, 2t^3 \right\rangle + \langle -7, 0, 2\rangle \\ &= \left\langle t^2+5t-7, -\cos(t)+3t, 2t^3+2 \right\rangle.\end{aligned}$$

What does the integration of a vector-valued function *mean*? There are

many applications, but none as direct as "the area under the curve" that we used in understanding the integral of a real-valued function.

A key understanding for us comes from considering the integral of a derivative:

$$\int_a^b \vec{r}\,'(t)\,dt = \vec{r}(t)\Big|_a^b = \vec{r}(b) - \vec{r}(a).$$

Integrating a *rate of change* function gives *displacement*.

Noting that vector-valued functions are closely related to parametric equations, we can describe the arc length of the graph of a vector-valued function as an integral. Given parametric equations $x = f(t)$, $y = g(t)$, the arc length on $[a, b]$ of the graph is

$$\text{Arc Length } = \int_a^b \sqrt{f'(t)^2 + g'(t)^2}\,dt,$$

as stated in Theorem 10.3.17 in Section 10.3. If $\vec{r}(t) = \langle f(t), g(t)\rangle$, note that $\sqrt{f'(t)^2 + g'(t)^2} = \|\vec{r}\,'(t)\|$. Therefore we can express the arc length of the graph of a vector-valued function as an integral of the magnitude of its derivative.

Theorem 13.2.36 Arc Length of a Vector-Valued Function.

Let $\vec{r}(t)$ be a vector-valued function where $\vec{r}\,'(t)$ is continuous on $[a, b]$. The arc length L of the graph of $\vec{r}(t)$ is

$$L = \int_a^b \|\vec{r}\,'(t)\|\,dt.$$

Note that we are actually integrating a scalar-function here, not a vector-valued function.

The next section takes what we have established thus far and applies it to objects in motion. We will let $\vec{r}(t)$ describe the path of an object in the plane or in space and will discover the information provided by $\vec{r}\,'(t)$ and $\vec{r}\,''(t)$.

13.2.5 Exercises

Terms and Concepts

1. Limits, derivatives and integrals of vector-valued functions are all evaluated _____-wise.

2. The definite integral of a rate of change function gives _____.

3. Why is it generally not useful to graph both $\vec{r}(t)$ and $\vec{r}\,'(t)$ on the same axes?

4. Theorem 13.2.25 contains three product rules. What are the three different types of products used in these rules?

Problems

Exercise Group. Evaluate the given limit.

5. $\lim\limits_{t\to 5}\left\langle 2t+1, 3t^2-1, \sin(t)\right\rangle$

6. $\lim\limits_{t\to 3}\left\langle e^t, \frac{t^2-9}{t+3}\right\rangle$

7. $\lim\limits_{t\to 0}\left\langle \frac{t}{\sin(t)}, (1+t)^{\frac{1}{t}}\right\rangle$

8. $\lim\limits_{h\to 0}\frac{\vec{r}(t+h)-\vec{r}(t)}{h}$, where $\vec{r}(t)=\left\langle t^2, t, 1\right\rangle$.

Exercise Group. Identify the interval or union of intervals on which $\vec{r}(t)$ is continuous.

9. $\vec{r}(t)=\left\langle t^2, 1/t\right\rangle$

10. $\vec{r}(t)=\langle\cos(t), e^t, \ln(t)\rangle$

Exercise Group. Find the derivative of the given function.

11. $\vec{r}(t)=\langle\cos(t), e^t, \ln(t)\rangle$

12. $\vec{r}(t)=\left\langle \frac{1}{t}, \frac{2t-1}{3t+1}, \tan(t)\right\rangle$

13. $\vec{r}(t)=(t^2)\,\langle\sin(t), 2t+5\rangle$

14. $\vec{r}(t)=\left\langle t^2+1, t-1\right\rangle\cdot\langle\sin(t), 2t+5\rangle$

15. $\vec{r}(t)=\left\langle t^2+1, t-1, 1\right\rangle\times\langle\sin(t), 2t+5, 1\rangle$

16. $\vec{r}(t)=\langle\cosh t, \sinh t\rangle$

Exercise Group. First, find $\vec{r}\,'(t)$. Then sketch $\vec{r}(t)$ and $\vec{r}\,'(1)$, with the initial point of $\vec{r}\,'(1)$ at $\vec{r}(1)$.

17. $\vec{r}(t)=\left\langle t^2+t, t^2-t\right\rangle$

18. $\vec{r}(t)=\left\langle t^2-2t+2, t^3-3t^2+2t\right\rangle$

19. $\vec{r}(t)=\left\langle t^2+1, t^3-t\right\rangle$

20. $\vec{r}(t)=\left\langle t^2-4t+5, t^3-6t^2+11t-6\right\rangle$

Exercise Group. Give the equation of the line tangent to the graph of $\vec{r}(t)$ at the given t value.

21. $\vec{r}(t)=\left\langle t^2+t, t^2-t\right\rangle$, at $t=1$

22. $\vec{r}(t)=\langle 3\cos(t), \sin(t)\rangle$, at $t=\pi/4$

23. $\vec{r}(t)=\langle 3\cos(t), 3\sin(t), t\rangle$ at $t=\pi$.

24. $\vec{r}(t)=\langle e^t, \tan(t), t\rangle$, at $t=0$.

Exercise Group. Find the value(s) of t for which $\vec{r}(t)$ is not smooth.

25. $\vec{r}(t)=\langle\cos(t), \sin(t)-t\rangle$

26. $\vec{r}(t)=\left\langle t^2-2t+1, t^3+t^2-5t+3\right\rangle$

27. $\vec{r}(t)=\langle\cos(t)-\sin(t), \sin(t)-\cos(t), \cos(4t)\rangle$

28. $\vec{r}(t)=\left\langle t^3-3t+2, -\cos(\pi t), \sin^2(\pi t)\right\rangle$

Exercise Group. The following exercises ask you to verify parts of Theorem 13.2.25. In each let $f(t)=t^3$, $\vec{r}(t)=\left\langle t^2, t-1, 1\right\rangle$ and $\vec{s}(t)=\langle\sin(t), e^t, t\rangle$. Compute the various derivatives as indicated.

29. Simplify $f(t)\vec{r}(t)$, then find its derivative; show this is the same as $f'(t)\vec{r}(t)+f(t)\vec{r}\,'(t)$.

30. Simplify $\vec{r}(t)\cdot\vec{s}(t)$, then find its derivative; show this is the same as $\vec{r}\,'(t)\cdot\vec{s}(t)+\vec{r}(t)\cdot\vec{s}\,'(t)$.

31. Simplify $\vec{r}(t)\times\vec{s}(t)$, then find its derivative; show this is the same as $\vec{r}\,'(t)\times\vec{s}(t)+\vec{r}(t)\times\vec{s}\,'(t)$.

32. Simplify $\vec{r}\big(f(t)\big)$, then find its derivative; show this is the same as $\vec{r}\,'\big(f(t)\big)f'(t)$.

Exercise Group. In the following exercises, evaluate the given definite or indefinite integral.

33. $\int \langle t^3, \cos(t), te^t \rangle \, dt$

34. $\int \left\langle \frac{1}{1+t^2}, \sec^2(t) \right\rangle \, dt$

35. $\int_0^{\pi} \langle -\sin(t), \cos(t) \rangle \, dt =$ _____

36. $\int_{-2}^{2} \langle 2t+1, 2t-1 \rangle \, dt$

Exercise Group. Solve the given initial value problems.

37. Find $\vec{r}(t)$, given that $\vec{r}\,'(t) = \langle t, \sin(t) \rangle$ and $\vec{r}(0) = \langle 2, 2 \rangle$.
$\vec{r}(t) =$ _____

38. Find $\vec{r}(t)$, given that $\vec{r}\,'(t) = \langle 1/(t+1), \tan(t) \rangle$ and $\vec{r}(0) = \langle 1, 2 \rangle$.

39. Find $\vec{r}(t)$, given that $\vec{r}\,''(t) = \langle t^2, t, 1 \rangle$, $\vec{r}\,'(0) = \langle 1, 2, 3 \rangle$ and $\vec{r}(0) = \langle 4, 5, 6 \rangle$.
$\vec{r}(t) =$ _____

40. Find $\vec{r}(t)$, given that $\vec{r}\,''(t) = \langle \cos(t), \sin(t), e^t \rangle$, $\vec{r}\,'(0) = \langle 0, 0, 0 \rangle$ and $\vec{r}(0) = \langle 0, 0, 0 \rangle$.

Exercise Group. Find the arc length of $\vec{r}(t)$ on the indicated interval.

41. $\vec{r}(t) = \langle 2\cos(t), 2\sin(t), 3t \rangle$ on $[0, 2\pi]$.

42. $\vec{r}(t) = \langle 5\cos(t), 3\sin(t), 4\sin(t) \rangle$ on $[0, 2\pi]$.

43. $\vec{r}(t) = \langle t^3, t^2, t^3 \rangle$ on $[0, 1]$.

44. $\vec{r}(t) = \langle e^{-t}\cos(t), e^{-t}\sin(t) \rangle$ on $[0, 1]$.

45. Prove Theorem 13.2.30; that is, show if $\vec{r}(t)$ has constant length and is differentiable, then $\vec{r}(t) \cdot \vec{r}\,'(t) = 0$. (Hint: use the Product Rule to compute $\frac{d}{dt}\big(\vec{r}(t) \cdot \vec{r}(t)\big)$.)

13.3 The Calculus of Motion

A common use of vector-valued functions is to describe the motion of an object in the plane or in space. A *position function* $\vec{r}(t)$ gives the position of an object at *time* t. More formally, let $O = \vec{0}$ (either in the plane or in space) and suppose an object is at point P_c at time $t = t_c$. Then $\vec{r}(t_c) = \overrightarrow{OP_c}$; that is, the vector $\vec{r}(t_c)$ "points" to the location of the object at a given time. This section explores how derivatives and integrals are used to study the motion described by such a function.

Definition 13.3.2 Velocity, Speed and Acceleration.

Let $\vec{r}(t)$ be a position function in $\mathbb{R}^2$ or $\mathbb{R}^3$.

Velocity	The instantaneous rate of position change, denoted $\vec{v}(t)$; that is, $\vec{v}(t) = \vec{r}\,'(t)$.
Speed	The magnitude of velocity: $\|\vec{v}(t)\|$.
Acceleration	The instantaneous rate of velocity change, denoted $\vec{a}(t)$; that is, $\vec{a}(t) = \vec{v}\,'(t) = \vec{r}\,''(t)$.

youtu.be/watch?v=UAGqV-r6jBI

Figure 13.3.1 Video introduction to Section 13.3

Example 13.3.3 Finding velocity and acceleration.

An object is moving with position function $\vec{r}(t) = \langle t^2 - t, t^2 + t\rangle$, $-3 \le t \le 3$, where distances are measured in feet and time is measured in seconds.

1. Find $\vec{v}(t)$ and $\vec{a}(t)$.
2. Sketch $\vec{r}(t)$; plot $\vec{v}(-1)$, $\vec{a}(-1)$, $\vec{v}(1)$ and $\vec{a}(1)$, each with their initial point at their corresponding point on the graph of $\vec{r}(t)$.
3. When is the object's speed minimized?

Video Solution

youtu.be/watch?v=qpQidOR7TQQ

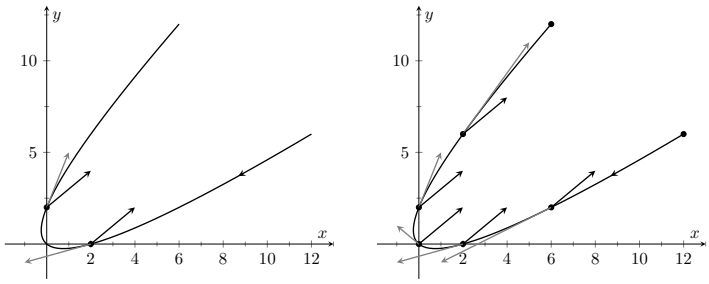
Solution.

1. Taking derivatives, we find

$$\vec{v}(t) = \vec{r}\,'(t) = \langle 2t - 1, 2t + 1\rangle \text{ and } \vec{a}(t) = \vec{r}\,''(t) = \langle 2, 2\rangle\,.$$

   Note that acceleration is constant.

2. $\vec{v}(-1) = \langle -3, -1\rangle$, $\vec{a}(-1) = \langle 2, 2\rangle$; $\vec{v}(1) = \langle 1, 3\rangle$, $\vec{a}(1) = \langle 2, 2\rangle$. These are plotted with $\vec{r}(t)$ in Figure 13.3.4(a). We can think of acceleration as "pulling" the velocity vector in a certain direction. At $t = -1$, the velocity vector points down and to the left; at $t = 1$, the velocity vector has been pulled in the $\langle 2, 2\rangle$ direction and is now pointing up and to the right. In Figure 13.3.4(b) we plot more velocity/acceleration vectors, making more clear the effect acceleration has on velocity.



(a) **(b)**

Figure 13.3.4 Graphing the position, velocity and acceleration of an object in Example 13.3.3

Since $\vec{a}(t)$ is constant in this example, as t grows large $\vec{v}(t)$ becomes almost parallel to $\vec{a}(t)$. For instance, when $t = 10$, $\vec{v}(10) = \langle 19, 21\rangle$, which is nearly parallel to $\langle 2, 2\rangle$.

3. The object's speed is given by

$$\|\vec{v}(t)\| = \sqrt{(2t-1)^2 + (2t+1)^2} = \sqrt{8t^2+2}.$$

To find the minimal speed, we could apply calculus techniques (such as set the derivative equal to 0 and solve for t, etc.) but we can find it by inspection. Inside the square root we have a quadratic which is minimized when $t = 0$. Thus the speed is minimized at $t = 0$, with a speed of $\sqrt{2}\ \frac{\text{ft}}{\text{s}}$. The graph in Figure 13.3.4(b) also implies speed is minimized here. The filled dots on the graph are located at integer values of t between -3 and 3. Dots that are far apart imply the object traveled a far distance in 1 second, indicating high speed; dots that are close together imply the object did not travel far in 1 second, indicating a low speed. The dots are closest together near $t = 0$, implying the speed is minimized near that value.

Example 13.3.5 Analyzing Motion.

Two objects follow an identical path at different rates on $[-1, 1]$. The position function for Object 1 is $\vec{r}_1(t) = \langle t, t^2\rangle$; the position function for Object 2 is $\vec{r}_2(t) = \langle t^3, t^6\rangle$, where distances are measured in feet and time is measured in seconds. Compare the velocity, speed and acceleration of the two objects on the path.

Solution. We begin by computing the velocity and acceleration function for each object:

$$\vec{v}_1(t) = \langle 1, 2t\rangle \qquad \vec{v}_2(t) = \langle 3t^2, 6t^5\rangle$$
$$\vec{a}_1(t) = \langle 0, 2\rangle \qquad \vec{a}_2(t) = \langle 6t, 30t^4\rangle$$

We immediately see that Object 1 has constant acceleration, whereas Object 2 does not.

At $t = -1$, we have $\vec{v}_1(-1) = \langle 1, -2\rangle$ and $\vec{v}_2(-1) = \langle 3, -6\rangle$; the velocity of Object 2 is three times that of Object 1 and so it follows that

the speed of Object 2 is three times that of Object 1 ($3\sqrt{5}$ ft/s compared to $\sqrt{5}$ ft/s.)

At $t = 0$, the velocity of Object 1 is $\vec{v}(1) = \langle 1, 0\rangle$ and the velocity of Object 2 is $\vec{0}$! This tells us that Object 2 comes to a complete stop at $t = 0$.

In Figure 13.3.6, we see the velocity and acceleration vectors for Object 1 plotted for $t = -1, -1/2, 0, 1/2$ and $t = 1$. Note again how the constant acceleration vector seems to "pull" the velocity vector from pointing down, right to up, right. We could plot the analogous picture for Object 2, but the velocity and acceleration vectors are rather large ($\vec{a}_2(-1) = \langle -6, 30\rangle$!)

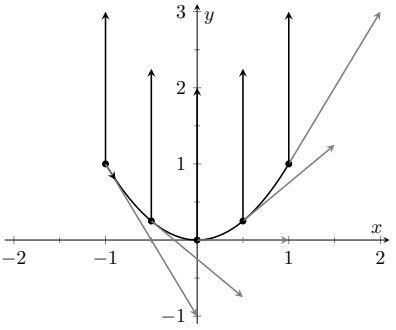
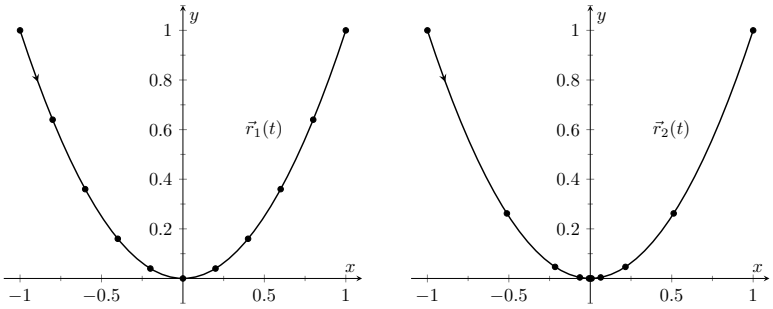


Figure 13.3.6 Plotting velocity and acceleration vectors for Object 1 in Example 13.3.5

Instead, we simply plot the locations of Object 1 and 2 on intervals of $1/10^{\text{th}}$ of a second, shown in Figure 13.3.7(a) and Figure 13.3.7(b) . Note how the x-values of Object 1 increase at a steady rate. This is because the x-component of $\vec{a}(t)$ is 0; there is no acceleration in the x-component. The dots are not evenly spaced; the object is moving faster near $t = -1$ and $t = 1$ than near $t = 0$.



(a) **(b)**

Figure 13.3.7 Comparing the positions of Objects 1 and 2 in Example 13.3.5

In Figure 13.3.7(b), we see the points plotted for Object 2. Note the large change in position from $t = -1$ to $t = -0.8$; the object starts moving very quickly. However, it slows considerably at it approaches the origin, and comes to a complete stop at $t = 0$. While it looks like there are 3 points near the origin, there are in reality 5 points there.

Since the objects begin and end at the same location, they have the same displacement. Since they begin and end at the same time, with the same displacement, they have the same average rate of change (i.e., they have the same average velocity). Since they follow the same path, they have the same distance traveled. Even though these three measurements are the same, the objects obviously travel the path in very different ways.

Example 13.3.8 Analyzing the motion of a whirling ball on a string.

A young boy whirls a ball, attached to a string, above his head in a counterclockwise circle. The ball follows a circular path and makes 2 revolutions per second. The string has length 2 ft.

1. Find the position function $\vec{r}(t)$ that describes this situation.

Video Solution

youtu.be/watch?v=qrNYh5hmyLE

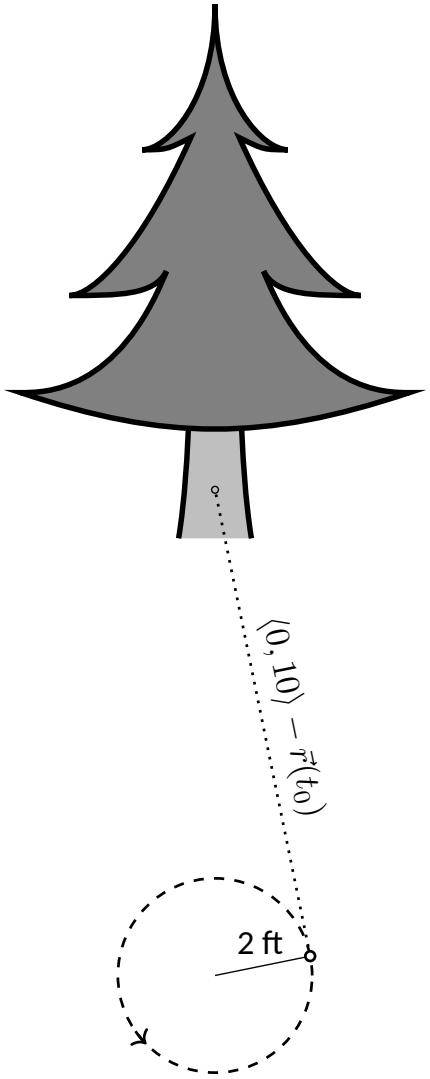


Figure 13.3.9 Modeling the flight of a ball in Example 13.3.8

2. Find the acceleration of the ball and give a physical interpretation of it.

3. A tree stands 10 ft in front of the boy. At what t-values should the boy release the string so that the ball hits the tree?

Solution.

1. The ball whirls in a circle. Since the string is 2ft long, the radius of the circle is 2. The position function $\vec{r}(t) = \langle 2\cos(t), 2\sin(t)\rangle$ describes a circle with radius 2, centered at the origin, but makes a full revolution every 2π seconds, not two revolutions per second. We modify the period of the trigonometric functions to be 1/2 by multiplying t by 4π. The final position function is thus

$$\vec{r}(t) = \langle 2\cos(4\pi t), 2\sin(4\pi t)\rangle\,.$$

   (Plot this for $0 \le t \le 1/2$ to verify that one revolution is made in 1/2 a second.)

2. To find $\vec{a}(t)$, we take the derivative of $\vec{r}(t)$ twice.

$$\begin{aligned}\vec{v}(t) = \vec{r}\,'(t) &= \langle -8\pi\sin(4\pi t), 8\pi\cos(4\pi t)\rangle\\ \vec{a}(t) = \vec{r}\,''(t) &= \left\langle -32\pi^2\cos(4\pi t), -32\pi^2\sin(4\pi t)\right\rangle\\ &= -32\pi^2\left\langle\cos(4\pi t), \sin(4\pi t)\right\rangle.\end{aligned}$$

   Note how $\vec{a}(t)$ is parallel to $\vec{r}(t)$, but has a different magnitude and points in the opposite direction. Why is this?

   Recall the classic physics equation, "Force $=$ mass $\times$ acceleration." A force acting on a mass induces acceleration (i.e., the mass moves); a mass that is accelerating is being acted upon by a force. Thus force and acceleration are closely related. A moving ball "wants" to travel in a straight line. Why does the ball in our example move in a circle? It is attached to the boy's hand by a string. The string applies a force to the ball, affecting its motion: the string *accelerates* the ball. This is not acceleration in the sense of "it travels faster;" rather, this acceleration is changing the velocity of the ball. In what direction is this force/acceleration being applied? In the direction of the string, towards the boy's hand.

   The magnitude of the acceleration is related to the speed at which the ball is traveling. A ball whirling quickly is rapidly changing direction/velocity. When velocity is changing rapidly, the acceleration must be "large."

3. When the boy releases the string, the string no longer applies a force to the ball, meaning acceleration is $\vec{0}$ and the ball can now move in a straight line in the direction of $\vec{v}(t)$.

   Let $t = t_0$ be the time when the boy lets go of the string. The ball will be at $\vec{r}(t_0)$, traveling in the direction of $\vec{v}(t_0)$. We want to find t_0 so that this line contains the point $(0, 10)$ (since the tree is 10 ft directly in front of the boy).

   There are many ways to find this time value. We choose one that is relatively simple computationally. As shown in Figure 13.3.9,

the vector from the release point to the tree is $\langle 0, 10\rangle - \vec{r}(t_0)$. This line segment is tangent to the circle, which means it is also perpendicular to $\vec{r}(t_0)$ itself, so their dot product is 0.

$$\begin{aligned}
&\vec{r}(t_0) \cdot \big(\langle 0, 10\rangle - \vec{r}(t_0)\big) = 0\\
&\langle 2\cos(4\pi t_0), 2\sin(4\pi t_0)\rangle \cdot \langle -2\cos(4\pi t_0), 10 - 2\sin(4\pi t_0)\rangle = 0\\
&-4\cos^2(4\pi t_0) + 20\sin(4\pi t_0) - 4\sin^2(4\pi t_0) = 0\\
&20\sin(4\pi t_0) - 4 = 0\\
&\sin(4\pi t_0) = 1/5\\
&4\pi t_0 = \sin^{-1}(1/5)\\
&4\pi t_0 \approx 0.2 + 2\pi n,
\end{aligned}$$

where n is an integer. Solving for t_0 we have:

$$t_0 \approx 0.016 + n/2$$

This is a wonderful formula. Every 1/2 second after $t = 0.016$ s the boy can release the string (since the ball makes 2 revolutions per second, he has two chances each second to release the ball).

Example 13.3.10 Analyzing motion in space.

An object moves in a helix with position function $\vec{r}(t) = \langle \cos(t), \sin(t), t\rangle$, where distances are measured in meters and time is in minutes. Describe the object's speed and acceleration at time t.

Solution. With $\vec{r}(t) = \langle \cos(t), \sin(t), t\rangle$, we have:

$$\begin{aligned}
\vec{v}(t) &= \langle -\sin(t), \cos(t), 1\rangle \text{ and}\\
\vec{a}(t) &= \langle -\cos(t), -\sin(t), 0\rangle .
\end{aligned}$$

The speed of the object is $\|\vec{v}(t)\| = \sqrt{(-\sin(t))^2 + \cos^2(t) + 1} = \sqrt{2}\ \frac{\text{m}}{\text{min}}$; it moves at a constant speed. Note that the object does not accelerate in the z-direction, but rather moves up at a constant rate of $1\ \frac{\text{m}}{\text{min}}$.

The objects in Examples 13.3.8 and Example 13.3.10 traveled at a constant speed. That is, $\|\vec{v}(t)\| = c$ for some constant c. Recall Theorem 13.2.30, which states that if a vector-valued function $\vec{r}(t)$ has constant length, then $\vec{r}(t)$ is perpendicular to its derivative: $\vec{r}(t) \cdot \vec{r}\,'(t) = 0$. In these examples, the velocity function has constant length, therefore we can conclude that the velocity is perpendicular to the acceleration: $\vec{v}(t) \cdot \vec{a}(t) = 0$. A quick check verifies this.

There is an intuitive understanding of this. If acceleration is parallel to velocity, then it is only affecting the object's speed; it does not change the direction of travel. (For example, consider a dropped stone. Acceleration and velocity are parallel — straight down — and the direction of velocity never changes, though speed does increase.) If acceleration is not perpendicular to velocity, then there is some acceleration in the direction of travel, influencing the speed. If speed is constant, then acceleration must be orthogonal to velocity, as it then only affects direction, and not speed.

Key Idea 13.3.11 Objects With Constant Speed.

If an object moves with constant speed, then its velocity and acceleration vectors are orthogonal. That is, $\vec{v}(t) \cdot \vec{a}(t) = 0$.

13.3.1 Projectile Motion

youtu.be/watch?v=-n1VJ3ngrTw

Figure 13.3.12 Video presentation of Subsection 13.3.1

This text takes g to be $32\ \frac{\text{ft}}{\text{s}^2}$ when using Imperial units, and $9.8\ \frac{\text{m}}{\text{s}^2}$ when using SI units.

An important application of vector-valued position functions is *projectile motion*: the motion of objects under only the influence of gravity. We will measure time in seconds, and distances will either be in meters or feet. We will show that we can completely describe the path of such an object knowing its initial position and initial velocity (i.e., where it *is* and where it *is going*.)

Suppose an object has initial position $\vec{r}(0) = \langle x_0, y_0 \rangle$ and initial velocity $\vec{v}(0) = \langle v_x, v_y \rangle$. It is customary to rewrite $\vec{v}(0)$ in terms of its speed v_0 and direction $\vec{u}$, where $\vec{u}$ is a unit vector. Recall all unit vectors in $\mathbb{R}^2$ can be written as $\langle \cos(\theta), \sin(\theta) \rangle$, where θ is an angle measure counter-clockwise from the x-axis. (We refer to θ as the *angle of elevation*.) Thus $\vec{v}(0) = v_0 \langle \cos(\theta), \sin(\theta) \rangle$.

Since the acceleration of the object is known, namely $\vec{a}(t) = \langle 0, -g \rangle$, where g is the gravitational constant, we can find $\vec{r}(t)$ knowing our two initial conditions. We first find $\vec{v}(t)$:

$$\begin{aligned} \vec{v}(t) &= \int \vec{a}(t)\, dt \\ \vec{v}(t) &= \int \langle 0, -g \rangle \ dt \\ \vec{v}(t) &= \langle 0, -gt \rangle + \vec{C}. \end{aligned}$$

Knowing $\vec{v}(0) = v_0 \langle \cos(\theta), \sin(\theta) \rangle$, we have $\vec{C} = v_0 \langle \cos(t), \sin(t) \rangle$ and so

$$\vec{v}(t) = \langle v_0 \cos(\theta), -gt + v_0 \sin(\theta) \rangle .$$

We integrate once more to find $\vec{r}(t)$:

$$\begin{aligned} \vec{r}(t) &= \int \vec{v}(t)\, dt \\ \vec{r}(t) &= \int \langle v_0 \cos(\theta), -gt + v_0 \sin(\theta) \rangle \ dt \\ \vec{r}(t) &= \left\langle \big(v_0 \cos(\theta)\big)t, -\frac{1}{2}gt^2 + \big(v_0 \sin(\theta)\big)t \right\rangle + \vec{C}. \end{aligned}$$

Knowing $\vec{r}(0) = \langle x_0, y_0 \rangle$, we conclude $\vec{C} = \langle x_0, y_0 \rangle$ and

$$\vec{r}(t) = \left\langle \big(v_0 \cos(\theta)\big)t + x_0\ , -\frac{1}{2}gt^2 + \big(v_0 \sin(\theta)\big)t + y_0 \right\rangle .$$

Key Idea 13.3.13 Projectile Motion.

The position function of a projectile propelled from an initial position of $\vec{r}_0 = \langle x_0, y_0 \rangle$, with initial speed v_0, with angle of elevation θ and neglecting all accelerations but gravity is

$$\vec{r}(t) = \left\langle \big(v_0 \cos(\theta)\big)t + x_0\ , -\frac{1}{2}gt^2 + \big(v_0 \sin(\theta)\big)t + y_0 \right\rangle .$$

Letting $\vec{v}_0 = v_0 \langle \cos(\theta), \sin(\theta) \rangle$, $\vec{r}(t)$ can be written as

$$\vec{r}(t) = \left\langle 0, -\frac{1}{2}gt^2 \right\rangle + \vec{v}_0 t + \vec{r}_0.$$

We demonstrate how to use this position function in the next two examples.

Example 13.3.14 Projectile Motion.

Sydney shoots her Red Ryder® bb gun across level ground from an elevation of 4 ft, where the barrel of the gun makes a 5° angle with the horizontal. Find how far the bb travels before landing, assuming the bb is fired at the advertised rate of 350 $\frac{\text{ft}}{\text{s}}$ and ignoring air resistance.

Video Solution

youtu.be/watch?v=9_Ev_uu0Y74

Solution. A direct application of Key Idea 13.3.13 gives

$$\begin{aligned}\vec{r}(t) &= \left\langle (350\cos(5^\circ))t, -16t^2 + (350\sin(5^\circ))t + 4 \right\rangle \\ &\approx \left\langle 346.67t, -16t^2 + 30.50t + 4 \right\rangle,\end{aligned}$$

where we set her initial position to be $\langle 0, 4 \rangle$. We need to find *when* the bb lands, then we can find *where*. We accomplish this by setting the y-component equal to 0 and solving for t:

$$\begin{aligned}-16t^2 + 30.50t + 4 &= 0 \\ t &= \frac{-30.50 \pm \sqrt{30.50^2 - 4(-16)(4)}}{-32} \\ t &\approx 2.03 \text{ s}.\end{aligned}$$

(We discarded a negative solution that resulted from our quadratic equation.)

We have found that the bb lands 2.03 s after firing; with $t = 2.03$, we find the x-component of our position function is $346.67(2.03) = 703.74$ ft. The bb lands about 704 feet away.

Example 13.3.15 Projectile Motion.

Alex holds his sister's bb gun at a height of 3 ft and wants to shoot a target that is 6 ft above the ground, 25 ft away. At what angle should he hold the gun to hit his target? (We still assume the muzzle velocity is 350 $\frac{\text{ft}}{\text{s}}$.)

Solution. The position function for the path of Alex's bb is

$$\vec{r}(t) = \left\langle (350\cos(\theta))t, -16t^2 + (350\sin(\theta))t + 3 \right\rangle.$$

We need to find θ so that $\vec{r}(t) = \langle 25, 6 \rangle$ for some value of t. That is, we want to find θ and t such that

$$(350\cos(\theta))t = 25 \text{ and } -16t^2 + (350\sin(\theta))t + 3 = 6.$$

This is not trivial (though not "hard"). We start by solving each equation for $\cos(\theta)$ and $\sin(\theta)$, respectively.

$$\cos(\theta) = \frac{25}{350t} \text{ and } \sin(\theta) = \frac{3 + 16t^2}{350t}.$$

Using the Pythagorean Identity $\cos^2(\theta) + \sin^2(\theta) = 1$, we have

$$\left(\frac{25}{350t}\right)^2 + \left(\frac{3+16t^2}{350t}\right)^2 = 1$$

Multiply both sides by $(350t)^2$:

$$25^2 + (3+16t^2)^2 = 350^2 t^2$$
$$256t^4 - 122,404t^2 + 634 = 0.$$

This is a quadratic *in* t^2. That is, we can apply the quadratic formula to find t^2, then solve for t itself.

$$t^2 = \frac{122,404 \pm \sqrt{122,404^2 - 4(256)(634)}}{512}$$
$$t^2 = 0.0052,\ 478.135$$
$$t = \pm 0.072,\ \pm 21.866$$

Clearly the negative t values do not fit our context, so we have $t = 0.072$ and $t = 21.866$. Using $\cos(\theta) = 25/(350t)$, we can solve for θ:

$$\theta = \cos^{-1}\left(\frac{25}{350 \cdot 0.072}\right) \text{ and } \cos^{-1}\left(\frac{25}{350 \cdot 21.866}\right)$$
$$\theta = 7.03^\circ \text{ and } 89.8^\circ.$$

Alex has two choices of angle. He can hold the rifle at an angle of about 7° with the horizontal and hit his target 0.07 s after firing, or he can hold his rifle almost straight up, with an angle of 89.8°, where he'll hit his target about 22 s later. The first option is clearly the option he should choose.

13.3.2 Distance Traveled

Consider a driver who sets her cruise-control to 60 mph, and travels at this speed for an hour. We can ask:

1. How far did the driver travel?
2. How far from her starting position is the driver?

The first is easy to answer: she traveled 60 miles. The second is impossible to answer with the given information. We do not know if she traveled in a straight line, on an oval racetrack, or along a slowly-winding highway.

This highlights an important fact: to compute distance traveled, we need only to know the speed, given by $\|\vec{v}(t)\|$.

Theorem 13.3.16 Distance Traveled.

Let $\vec{v}(t)$ be a velocity function for a moving object. The distance traveled by the object on $[a, b]$ is:

$$\textit{distance traveled} = \int_a^b \|\vec{v}(t)\|\, dt.$$

Note that this is just a restatement of Theorem 13.2.36: arc length is the same as distance traveled, just viewed in a different context.

Example 13.3.18 Distance Traveled, Displacement, and Average Speed.

A particle moves in space with position function $\vec{r}(t) = \langle t, t^2, \sin(\pi t)\rangle$ on $[-2, 2]$, where t is measured in seconds and distances are in meters. Find:

1. The distance traveled by the particle on $[-2, 2]$.
2. The displacement of the particle on $[-2, 2]$.
3. The particle's average speed.

Solution.

1. We use Theorem 13.3.16 to establish the integral:

$$\begin{aligned}\text{distance traveled} &= \int_{-2}^{2} \|\vec{v}(t)\|\, dt \\ &= \int_{-2}^{2} \sqrt{1 + (2t)^2 + \pi^2 \cos^2(\pi t)}\, dt.\end{aligned}$$

   This cannot be solved in terms of elementary functions so we turn to numerical integration, finding the distance to be 12.88 m.

2. The displacement is the vector

$$\vec{r}(2) - \vec{r}(-2) = \langle 2, 4, 0\rangle - \langle -2, 4, 0\rangle = \langle 4, 0, 0\rangle\,.$$

   That is, the particle ends with an x-value increased by 4 and with y- and z-values the same (see Figure 13.3.19).

3. We found above that the particle traveled 12.88 m over 4 seconds. We can compute average speed by dividing: $12.88/4 = 3.22$ m/s.

   We should also consider Definition 5.4.34 of Section 5.4, which says that the average value of a function f on $[a, b]$ is $\frac{1}{b-a}\int_a^b f(x)\, dx$. In our context, the average value of the speed is

$$\text{average speed} = \frac{1}{2-(-2)}\int_{-2}^{2} \|\vec{v}(t)\|\, dt \approx \frac{1}{4}12.88 = 3.22 \text{ m/s}.$$

   Note how the physical context of a particle traveling gives meaning to a more abstract concept learned earlier.

In Definition 5.4.34 of Chapter 5 we defined the average value of a function $f(x)$ on $[a, b]$ to be

$$\frac{1}{b-a}\int_a^b f(x)\, dx.$$

Note how in Example 13.3.18 we computed the average speed as

$$\frac{\text{distance traveled}}{\text{travel time}} = \frac{1}{2-(-2)}\int_{-2}^{2} \|\vec{v}(t)\|\, dt;$$

that is, we just found the average value of $\|\vec{v}(t)\|$ on $[-2, 2]$.

youtu.be/watch?v=vTtjBp2SC1I

Figure 13.3.17 Video presentation of Theorem 13.3.16

Video Solution

youtu.be/watch?v=vLSocZ7XLQM

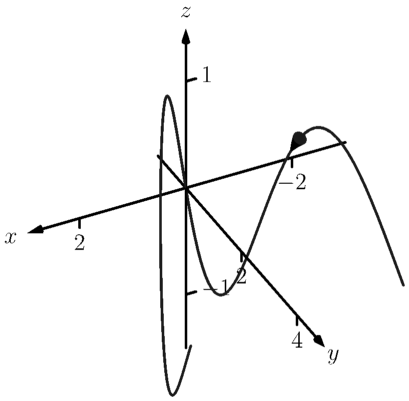


Figure 13.3.19 The path of the particle in Example 13.3.18

Likewise, given position function $\vec{r}(t)$, the average velocity on $[a, b]$ is

$$\frac{\text{displacement}}{\text{travel time}} = \frac{1}{b-a}\int_a^b \vec{r}\,'(t)\,dt = \frac{\vec{r}(b)-\vec{r}(a)}{b-a};$$

that is, it is the average value of $\vec{r}\,'(t)$, or $\vec{v}(t)$, on $[a, b]$.

Key Idea 13.3.20 Average Speed, Average Velocity.

Let $\vec{r}(t)$ be a differentiable position function on $[a, b]$.

The **average speed** is:

$$\frac{\text{distance traveled}}{\text{travel time}} = \frac{\int_a^b \|\vec{r}\,'(t)\|\,dt}{b-a} = \frac{1}{b-a}\int_a^b \|\vec{v}(t)\|\,dt.$$

The **average velocity** is:

$$\frac{\text{displacement}}{\text{travel time}} = \frac{\int_a^b \vec{r}\,'(t)\,dt}{b-a} = \frac{1}{b-a}\int_a^b \vec{r}\,'(t)\,dt.$$

The next two sections investigate more properties of the graphs of vector-valued functions and we'll apply these new ideas to what we just learned about motion.

13.3.3 Exercises

Terms and Concepts

1. How is *velocity* different from *speed*?

2. What is the difference between *displacement* and *distance traveled*?

3. What is the difference between *average velocity* and *average speed*?

4. *Distance traveled* is the same as _____ _____, just viewed in a different context.

5. Describe a scenario where an object's average speed is a large number, but the magnitude of the average velocity is not a large number.

6. Explain why it is not possible to have an average velocity with a large magnitude but a small average speed.

Problems

Exercise Group. In the following exercises, a position function $\vec{r}(t)$ is given. Find $\vec{v}(t)$ and $\vec{a}(t)$.

7. $\vec{r}(t) = \langle 2t+1, 5t-2, 7\rangle$

8. $\vec{r}(t) = \langle 3t^2-2t+1, -t^2+t+14\rangle$

9. $\vec{r}(t) = \langle \cos(t), \sin(t)\rangle$

10. $\vec{r}(t) = \langle t/10, -\cos(t), \sin(t)\rangle$

Exercise Group. In the following exercises, a position function $\vec{r}(t)$ is given. Sketch $\vec{r}(t)$ on the indicated interval. Find $\vec{v}(t)$ and $\vec{a}(t)$, then add $\vec{v}(t_0)$ and $\vec{a}(t_0)$ to your sketch, with their initial points at $\vec{r}(t_0)$, for the given value of t_0.

11. $\vec{r}(t) = \langle t, \sin(t)\rangle$ on $[0, \pi/2]$; $t_0 = \pi/4$

12. $\vec{r}(t) = \langle t^2, \sin(t^2)\rangle$ on $[0, \pi/2]$; $t_0 = \sqrt{\pi/4}$

13. $\vec{r}(t) = \langle t^2+t, -t^2+2t\rangle$ on $[-2, 2]$; $t_0 = 1$

14. $\vec{r}(t) = \left\langle \dfrac{2t+3}{t^2+1}, t^2 \right\rangle$ on $[-1, 1]$; $t_0 = 0$

Exercise Group. In the following exercises, a position function $\vec{r}(t)$ of an object is given. Find the speed of the object in terms of t, and find where the speed is minimized/maximized on the indicated interval.

15. $\vec{r}(t) = \langle t^2, t\rangle$ on $[-1, 1]$

16. $\vec{r}(t) = \langle t^2, t^2-t^3\rangle$ on $[-1, 1]$

17. $\vec{r}(t) = \langle 5\cos(t), 5\sin(t)\rangle$ on $[0, 2\pi]$

18. $\vec{r}(t) = \langle 2\cos(t), 5\sin(t)\rangle$ on $[0, 2\pi]$

19. $\vec{r}(t) = \langle \sec(t), \tan(t)\rangle$ on $[0, \pi/4]$.

20. $\vec{r}(t) = \langle t+\cos(t), 1-\sin(t)\rangle$ on $[0, 2\pi]$

21. $\vec{r}(t) = \langle 12t, 5\cos(t), 5\sin(t)\rangle$ on $[0, 4\pi]$

22. $\vec{r}(t) = \langle t^2-t, t^2+t, t\rangle$ on $[0, 1]$.

23. $\vec{r}(t) = \left\langle t, t^2, \sqrt{1-t^2}\right\rangle$ on $[-1, 1]$

24. *Projectile Motion:*
$\vec{r}(t) = \left\langle (v_0\cos(\theta))t, -\dfrac{1}{2}gt^2 + (v_0\sin(\theta))t \right\rangle$
on $\left[0, \dfrac{2v_0\sin(\theta)}{g}\right]$

Exercise Group. In the following exercises, position functions $\vec{r}_1(t)$ and $\vec{r}_2(s)$ for two objects are given that follow the same path on the respective intervals.

(a) Show that the positions are the same at the indicated t_0 and s_0 values; i.e., show $\vec{r}_1(t_0) = \vec{r}_2(s_0)$.

(b) Find the velocity, speed and acceleration of the two objects at t_0 and s_0, respectively.

25. $\vec{r}_1(t) = \langle t, t^2\rangle$ on $[0, 1]$; $t_0 = 1$
$\vec{r}_2(s) = \langle s^2, s^4\rangle$ on $[0, 1]$; $s_0 = 1$

26. $\vec{r}_1(t) = \langle 3\cos(t), 3\sin(t)\rangle$ on $[0, 2\pi]$; $t_0 = \pi/2$
$\vec{r}_2(s) = \langle 3\cos(4s), 3\sin(4s)\rangle$ on $[0, \pi/2]$;
$s_0 = \pi/8$

27. $\vec{r}_1(t) = \langle 3t, 2t\rangle$ on $[0, 2]$; $t_0 = 2$
$\vec{r}_2(s) = \langle 6s-6, 4s-4\rangle$ on $[1, 2]$; $s_0 = 2$

28. $\vec{r}_1(t) = \langle t, \sqrt{t}\rangle$ on $[0, 1]$; $t_0 = 1$
$\vec{r}_2(s) = \left\langle \sin(s), \sqrt{\sin(s)}\right\rangle$ on $[0, \pi/2]$;
$s_0 = \pi/2$

Exercise Group. In the following exercises, find the position function of an object given its acceleration and initial velocity and position.

29. $\vec{a}(t) = \langle 2, 3\rangle; \vec{v}(0) = \langle 1, 2\rangle, \vec{r}(0) = \langle 5, -2\rangle$

30. Given $\vec{a}(t) = \langle 2, 3\rangle$, $\vec{v}(1) = \langle 1, 2\rangle$, and $\vec{r}(1) = \langle 5, -2\rangle$, find the position function $\vec{r}(t)$.

31. $\vec{a}(t) = \langle \cos(t), -\sin(t)\rangle; \vec{v}(0) = \langle 0, 1\rangle, \vec{r}(0) = \langle 0, 0\rangle$

32. Given $\vec{a}(t) = \langle 0, -32\rangle$, $\vec{v}(0) = \langle 10, 50\rangle$, and $\vec{r}(0) = \langle 0, 0\rangle$, find the position function $\vec{r}(t)$.

Exercise Group. In the following exercises, find the displacement, distance traveled, average velocity and average speed of the described object on the given interval.

33. An object with position function $\vec{r}(t) = \langle 2\cos(t), 2\sin(t), 3t\rangle$, where distances are measured in feet and time is in seconds, on $[0, 2\pi]$.

34. An object has position function $\vec{r}(t) = \langle 5\cos(t), -5\sin(t)\rangle$, where distances are measured in feet and time is in seconds. Over $[0, \pi]$.

35. An object with velocity function $\vec{v}(t) = \langle \cos(t), \sin(t)\rangle$, where distances are measured in feet and time is in seconds, on $[0, 2\pi]$.

36. An object has velocity function $\vec{v}(t) = \langle 1, 2, -1\rangle$, where distances are measured in feet and time is in seconds. Over $[0, 10]$.

Exercise Group. The following exercises ask you to solve a variety of problems based on the principles of projectile motion.

37. A boy whirls a ball, attached to a 3 ft string, above his head in a counter-clockwise circle. The ball makes 2 revolutions per second. At what t-values should the boy release the string so that the ball heads directly for a tree standing 10 ft in front of him?

38. David faces Goliath with only a stone in a 3 ft sling, which he whirls above his head at 4 revolutions per second. They stand 20 ft apart.

(a) At what t-values must David release the stone in his sling in order to hit Goliath?

(b) What is the speed at which the stone is traveling when released?

(c) Assume David releases the stone from a height of 6ft and Goliath's forehead is 9 ft above the ground. What angle of elevation must David apply to the stone to hit Goliath's head?

39. A hunter aims at a deer which is 40 yards away. Her crossbow is at a height of 5 ft, and she aims for a spot on the deer 4 ft above the ground. The crossbow fires her arrows at 300 ft/s.

(a) At what angle of elevation should she hold the crossbow to hit her target?

(b) If the deer is moving perpendicularly to her line of sight at a rate of 20 mph, by approximately how much should she lead the deer in order to hit it in the desired location? (How far ahead of the deer should she aim?)

40. A baseball player hits a ball at 100 mph, with an initial height of 3 ft and an angle of elevation of $20°$, at Boston's Fenway Park. The ball flies towards the famed "Green Monster," a wall 37 ft high located 310 ft from home plate.

(a) Show that as hit, the ball hits the wall.

(b) Show that if the angle of elevation is $21°$, the ball clears the Green Monster.

41. A Cessna flies at 1000 ft at 150 mph and drops a box of supplies to the professor (and his wife) on an island. Ignoring wind resistance, how far horizontally will the supplies travel before they land?

42. A football quarterback throws a pass from a height of 6 ft, intending to hit his receiver 20 yds away at a height of 5 ft.

(a) If the ball is thrown at a rate of 50mph, what angle of elevation is needed to hit his intended target?

(b) If the ball is thrown at with an angle of elevation of 8°, what initial ball speed is needed to hit his target?

13.4 Unit Tangent and Normal Vectors

13.4.1 Unit Tangent Vector

Given a smooth vector-valued function $\vec{r}(t)$, we defined in Definition 13.2.17 that any vector parallel to $\vec{r}\,'(t_0)$ is *tangent* to the graph of $\vec{r}(t)$ at $t = t_0$. It is often useful to consider just the *direction* of $\vec{r}\,'(t)$ and not its magnitude. Therefore we are interested in the unit vector in the direction of $\vec{r}\,'(t)$. This leads to a definition.

youtu.be/watch?v=tmKmAKUdSZ8

Figure 13.4.2 Video presentation of Definition 13.4.1

Definition 13.4.1 Unit Tangent Vector.

Let $\vec{r}(t)$ be a smooth function on an open interval I. The unit tangent vector $\vec{T}(t)$ is

$$\vec{T}(t) = \frac{1}{\|\vec{r}\,'(t)\|}\vec{r}\,'(t).$$

Example 13.4.3 Computing the unit tangent vector.

Let $\vec{r}(t) = \langle 3\cos(t), 3\sin(t), 4t\rangle$. Find $\vec{T}(t)$ and compute $\vec{T}(0)$ and $\vec{T}(1)$.

Solution. We apply Definition 13.4.1 to find $\vec{T}(t)$.

$$\begin{aligned}\vec{T}(t) &= \frac{1}{\|\vec{r}\,'(t)\|}\vec{r}\,'(t)\\ &= \frac{1}{\sqrt{\big(-3\sin(t)\big)^2 + \big(3\cos(t)\big)^2 + 4^2}}\langle -3\sin(t), 3\cos(t), 4\rangle\\ &= \left\langle -\frac{3}{5}\sin(t), \frac{3}{5}\cos(t), \frac{4}{5}\right\rangle.\end{aligned}$$

We can now easily compute $\vec{T}(0)$ and $\vec{T}(1)$:

$$\vec{T}(0) = \left\langle 0, \frac{3}{5}, \frac{4}{5}\right\rangle; \vec{T}(1) = \left\langle -\frac{3}{5}\sin(1), \frac{3}{5}\cos(1), \frac{4}{5}\right\rangle \approx \langle -0.505, 0.324, 0.8\rangle.$$

These are plotted in Figure 13.4.4 with their initial points at $\vec{r}(0)$ and $\vec{r}(1)$, respectively. (They look rather "short" since they are only length 1.)

The unit tangent vector $\vec{T}(t)$ always has a magnitude of 1, though it is sometimes easy to doubt that is true. We can help solidify this thought in our minds by computing $\left\|\vec{T}(1)\right\|$:

$$\left\|\vec{T}(1)\right\| \approx \sqrt{(-0.505)^2 + 0.324^2 + 0.8^2} = 1.000001.$$

We have rounded in our computation of $\vec{T}(1)$, so we don't get 1 exactly. We leave it to the reader to use the exact representation of $\vec{T}(1)$ to verify it has length 1.

Video Solution

youtu.be/watch?v=aGEMbQrBpPI

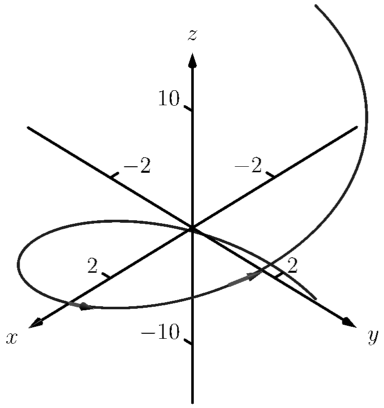


Figure 13.4.4 Plotting unit tangent vectors in Example 13.4.3

In many ways, the previous example was "too nice." It turned out that $\vec{r}\,'(t)$ was always of length 5. In the next example the length of $\vec{r}\,'(t)$ is variable, leaving us with a formula that is not as clean.

Example 13.4.5 Computing the unit tangent vector.

Let $\vec{r}(t) = \langle t^2 - t, t^2 + t\rangle$. Find $\vec{T}(t)$ and compute $\vec{T}(0)$ and $\vec{T}(1)$.

Solution. We find $\vec{r}\,'(t) = \langle 2t - 1, 2t + 1\rangle$, and

$$\|\vec{r}\,'(t)\| = \sqrt{(2t-1)^2 + (2t+1)^2} = \sqrt{8t^2 + 2}.$$

Therefore

$$\vec{T}(t) = \frac{1}{\sqrt{8t^2+2}}\langle 2t - 1, 2t + 1\rangle = \left\langle \frac{2t-1}{\sqrt{8t^2+2}}, \frac{2t+1}{\sqrt{8t^2+2}} \right\rangle.$$

When $t = 0$, we have $\vec{T}(0) = \langle -1/\sqrt{2}, 1/\sqrt{2}\rangle$; when $t = 1$, we have $\vec{T}(1) = \langle 1/\sqrt{10}, 3/\sqrt{10}\rangle$. We leave it to the reader to verify each of these is a unit vector. They are plotted in Figure 13.4.6

Video Solution

youtu.be/watch?v=zdfMTWdXzJs

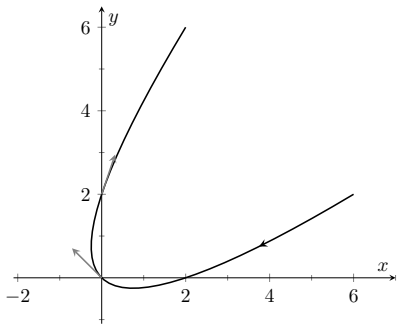


Figure 13.4.6 Plotting unit tangent vectors in Example 13.4.5

13.4.2 Unit Normal Vector

Just as knowing the direction tangent to a path is important, knowing a direction orthogonal to a path is important. When dealing with real-valued functions, we defined the normal line at a point to the be the line through the point that was perpendicular to the tangent line at that point. We can do a similar thing with vector-valued functions. Given $\vec{r}(t)$ in $\mathbb{R}^2$, we have 2 directions perpendicular to the tangent vector, as shown in Figure 13.4.7. It is good to wonder "Is one of these two directions preferable over the other?"

Given $\vec{r}(t)$ in $\mathbb{R}^3$, there are infinitely many vectors orthogonal to the tangent vector at a given point. Again, we might wonder "Is one out of this infinite number of choices preferable over the others? Is one of these the 'right' choice?"

The answer in both $\mathbb{R}^2$ and $\mathbb{R}^3$ is "Yes, there is one vector that is not only preferable, it is the 'right' one to choose." Recall Theorem 13.2.30, which states that if $\vec{r}(t)$ has constant length, then $\vec{r}(t)$ is orthogonal to $\vec{r}\,'(t)$ for all t. We know $\vec{T}(t)$, the unit tangent vector, has constant length. Therefore $\vec{T}(t)$ is orthogonal to $\vec{T}\,'(t)$.

We'll see that $\vec{T}\,'(t)$ is more than just a convenient choice of vector that is orthogonal to $\vec{r}\,'(t)$; rather, it is the "right" choice. Since all we care about is the direction, we define this newly found vector to be a unit vector.

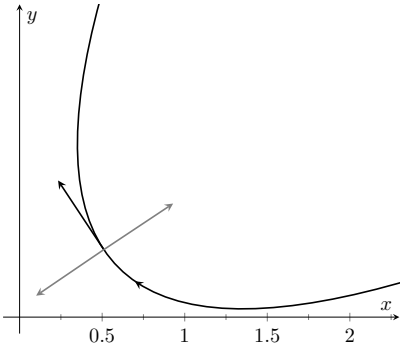


Figure 13.4.7 Given a direction in the plane, there are always two directions orthogonal to it

Definition 13.4.8 Unit Normal Vector.

Let $\vec{r}(t)$ be a vector-valued function where the unit tangent vector, $\vec{T}(t)$, is smooth on an open interval I. The **unit normal vector** $\vec{N}(t)$ is

$$\vec{N}(t) = \frac{1}{\left\|\vec{T}\,'(t)\right\|}\vec{T}\,'(t).$$

$\vec{T}(t)$ is a unit vector, by definition. This *does not* imply that $\vec{T}\,'(t)$ is also a unit vector.

youtu.be/watch?v=HWX5uPUCFJs

Figure 13.4.9 Video presentation of Definition 13.4.8

Video Solution

youtu.be/watch?v=cZkvyF0t5P4

Example 13.4.10 Computing the unit normal vector.

Let $\vec{r}(t) = \langle 3\cos(t), 3\sin(t), 4t\rangle$ as in Example 13.4.3. Sketch both $\vec{T}(\pi/2)$ and $\vec{N}(\pi/2)$ with initial points at $\vec{r}(\pi/2)$.

Solution. In Example 13.4.3, we found

$$\vec{T}(t) = \langle(-3/5)\sin(t), (3/5)\cos(t), 4/5\rangle\,.$$

Therefore

$$\vec{T}\,'(t) = \left\langle -\frac{3}{5}\cos(t), -\frac{3}{5}\sin(t), 0\right\rangle \text{ and } \left\|\vec{T}\,'(t)\right\| = \frac{3}{5}.$$

Thus

$$\vec{N}(t) = \frac{\vec{T}\,'(t)}{3/5} = \langle -\cos(t), -\sin(t), 0\rangle\,.$$

We compute $\vec{T}(\pi/2) = \langle -3/5, 0, 4/5\rangle$ and $\vec{N}(\pi/2) = \langle 0, -1, 0\rangle$. These are sketched in Figure 13.4.11.

There is one flaw in our definition of $\vec{N}(t)$: it is possible that we could have $\vec{T}\,'(t) = 0$! Indeed, this is the case for any line of the form $\vec{\ell}(t) = \vec{r}_0 + t\vec{v}$. For straight lines in the plane, it is most common to orient the normal vector 90° counterclockwise from the tangent vector, but for lines in three dimensions, there is no preferred choice of normal vector.

Figure 13.4.11 Plotting unit tangent and normal vectors in Example 13.4.11

The previous example was once again "too nice." In general, the expression for $\vec{T}(t)$ contains fractions of square roots, hence the expression of $\vec{T}\,'(t)$ is very messy. We demonstrate this in the next example.

Example 13.4.12 Computing the unit normal vector.

Let $\vec{r}(t) = \left\langle t^2 - t, t^2 + t\right\rangle$ as in Example 13.4.5. Find $\vec{N}(t)$ and sketch $\vec{r}(t)$ with the unit tangent and normal vectors at $t = -1, 0$ and 1.

Solution. In Example 13.4.5, we found

$$\vec{T}(t) = \left\langle \frac{2t-1}{\sqrt{8t^2+2}}, \frac{2t+1}{\sqrt{8t^2+2}}\right\rangle.$$

Finding $\vec{T}\,'(t)$ requires two applications of the Quotient Rule:

$$\begin{aligned} T\,'(t) = \Bigg\langle & \frac{\sqrt{8t^2+2}(2) - (2t-1)\left(\frac{1}{2}(8t^2+2)^{-1/2}(16t)\right)}{8t^2+2}, \\ & \frac{\sqrt{8t^2+2}(2) - (2t+1)\left(\frac{1}{2}(8t^2+2)^{-1/2}(16t)\right)}{8t^2+2}\Bigg\rangle \\ = & \left\langle \frac{4(2t+1)}{(8t^2+2)^{3/2}}, \frac{4(1-2t)}{(8t^2+2)^{3/2}}\right\rangle \end{aligned}$$

Video Solution

youtu.be/watch?v=VExazWvgGlQ

This is not a unit vector; to find $\vec{N}(t)$, we need to divide $\vec{T}\,'(t)$ by its magnitude.

$$\begin{aligned} \left\|\vec{T}\,'(t)\right\| &= \sqrt{\frac{16(2t+1)^2}{(8t^2+2)^3} + \frac{16(1-2t)^2}{(8t^2+2)^3}} \\ &= \sqrt{\frac{16(8t^2+2)}{(8t^2+2)^3}} \\ &= \frac{4}{8t^2+2}. \end{aligned}$$

Finally,

$$\begin{aligned}\vec{N}(t) &= \frac{1}{4/(8t^2+2)}\left\langle \frac{4(2t+1)}{(8t^2+2)^{3/2}}, \frac{4(1-2t)}{(8t^2+2)^{3/2}} \right\rangle \\ &= \left\langle \frac{2t+1}{\sqrt{8t^2+2}}, -\frac{2t-1}{\sqrt{8t^2+2}} \right\rangle.\end{aligned}$$

Using this formula for $\vec{N}(t)$, we compute the unit tangent and normal vectors for $t = -1, 0$ and 1 and sketch them in Figure 13.4.13.

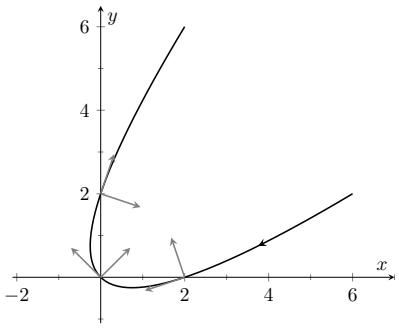


Figure 13.4.13 Plotting unit tangent and normal vectors in Example 13.4.12

The final result for $\vec{N}(t)$ in Example 13.4.12 is suspiciously similar to $\vec{T}(t)$. There is a clear reason for this. If $\vec{u} = \langle u_1, u_2 \rangle$ is a unit vector in $\mathbb{R}^2$, then the *only* unit vectors orthogonal to $\vec{u}$ are $\langle -u_2, u_1 \rangle$ and $\langle u_2, -u_1 \rangle$. Given $\vec{T}(t)$, we can quickly determine $\vec{N}(t)$ if we know which term to multiply by (-1).

Consider again Figure 13.4.13, where we have plotted some unit tangent and normal vectors. Note how $\vec{N}(t)$ always points "inside" the curve, or to the concave side of the curve. This is not a coincidence; this is true in general. Knowing the direction that $\vec{r}(t)$ "turns" allows us to quickly find $\vec{N}(t)$.

Theorem 13.4.14 Unit Normal Vectors in $\mathbb{R}^2$.

Let $\vec{r}(t)$ be a vector-valued function in $\mathbb{R}^2$ where $\vec{T}'(t)$ is smooth on an open interval I. Let t_0 be in I and $\vec{T}(t_0) = \langle t_1, t_2 \rangle$ Then $\vec{N}(t_0)$ is either

$$\vec{N}(t_0) = \langle -t_2, t_1 \rangle \text{ or } \vec{N}(t_0) = \langle t_2, -t_1 \rangle,$$

whichever is the vector that points to the concave side of the graph of $\vec{r}$.

13.4.3 Application to Acceleration

Let $\vec{r}(t)$ be a position function. It is a fact (stated later in Theorem 13.4.15) that acceleration, $\vec{a}(t)$, lies in the plane defined by $\vec{T}$ and $\vec{N}$. That is, there are scalar functions $a_\mathsf{T}(t)$ and $a_\mathsf{N}(t)$ such that

$$\vec{a}(t) = a_\mathsf{T}(t)\vec{T}(t) + a_\mathsf{N}(t)\vec{N}(t).$$

We generally drop the "of t" part of the notation and just write a_T and a_N.

The scalar a_T measures "how much" acceleration is in the direction of travel, that is, it measures the component of acceleration that affects the speed. The scalar a_N measures "how much" acceleration is perpendicular to the direction of travel, that is, it measures the component of acceleration that affects the direction of travel.

We can find a_T using the orthogonal projection of $\vec{a}(t)$ onto $\vec{T}(t)$ (review Definition 12.3.17 in Section 12.3 if needed). Recalling that since $\vec{T}(t)$ is a unit vector, $\vec{T}(t) \cdot \vec{T}(t) = 1$, so we have

$$\text{proj}_{\vec{T}(t)}\,\vec{a}(t) = \frac{\vec{a}(t)\cdot\vec{T}(t)}{\vec{T}(t)\cdot\vec{T}(t)}\vec{T}(t) = \underbrace{\left(\vec{a}(t)\cdot\vec{T}(t)\right)}_{a_\mathsf{T}}\vec{T}(t).$$

Thus the amount of $\vec{a}(t)$ in the direction of $\vec{T}(t)$ is $a_\mathsf{T} = \vec{a}(t)\cdot\vec{T}(t)$. The same logic gives $a_\mathsf{N} = \vec{a}(t)\cdot\vec{N}(t)$.

Keep in mind that both a_T and a_N are functions of t; that is, the scalar changes depending on t. It is convention to drop the "(t)" notation from $a_T(t)$ and simply write a_T.

youtu.be/watch?v=mRKDo_rot7Y

Figure 13.4.16 Video presentation of Theorem 13.4.15

While this is a fine way of computing a_T, there are simpler ways of finding a_N (as finding $\vec{N}$ itself can be complicated). The following theorem gives alternate formulas for a_T and a_N.

Theorem 13.4.15 Acceleration in the Plane Defined by $\vec{T}$ and $\vec{N}$.

Let $\vec{r}(t)$ be a position function with acceleration $\vec{a}(t)$ and unit tangent and normal vectors $\vec{T}(t)$ and $\vec{N}(t)$. Then $\vec{a}(t)$ lies in the plane defined by $\vec{T}(t)$ and $\vec{N}(t)$; that is, there exists scalars a_T and a_N such that

$$\vec{a}(t) = a_T\vec{T}(t) + a_N\vec{N}(t).$$

Moreover,

$$a_T = \vec{a}(t)\cdot\vec{T}(t) = \frac{d}{dt}\Big(\,\|\vec{v}(t)\|\,\Big)$$
$$a_N = \vec{a}(t)\cdot\vec{N}(t) = \sqrt{\|\vec{a}(t)\|^2 - a_T^2} = \frac{\|\vec{a}(t)\times\vec{v}(t)\|}{\|\vec{v}(t)\|} = \|\vec{v}(t)\|\,\left\|\vec{T}\,'(t)\right\|$$

Note the second formula for a_T: $\dfrac{d}{dt}\Big(\,\|\vec{v}(t)\|\,\Big)$. This measures the rate of change of speed, which again is the amount of acceleration in the direction of travel.

Example 13.4.17 Computing a_T and a_N.

Let $\vec{r}(t) = \langle 3\cos(t), 3\sin(t), 4t\rangle$ as in Examples 13.4.3 and Example 13.4.10. Find a_T and a_N.

Solution. The previous examples give $\vec{a}(t) = \langle -3\cos(t), -3\sin(t), 0\rangle$ and

$$\vec{T}(t) = \left\langle -\frac{3}{5}\sin(t), \frac{3}{5}\cos(t), \frac{4}{5}\right\rangle \text{ and } \vec{N}(t) = \langle -\cos(t), -\sin(t), 0\rangle.$$

We can find a_T and a_N directly with dot products:

$$a_T = \vec{a}(t)\cdot\vec{T}(t) = \frac{9}{5}\cos(t)\sin(t) - \frac{9}{5}\cos(t)\sin(t) + 0 = 0.$$
$$a_N = \vec{a}(t)\cdot\vec{N}(t) = 3\cos^2(t) + 3\sin^2(t) + 0 = 3.$$

Thus $\vec{a}(t) = 0\vec{T}(t) + 3\vec{N}(t) = 3\vec{N}(t)$, which is clearly the case.

What is the practical interpretation of these numbers? $a_T = 0$ means the object is moving at a constant speed, and hence all acceleration comes in the form of direction change.

Video Solution

youtu.be/watch?v=qA776vEHYM0

Example 13.4.18 Computing a_T and a_N.

Let $\vec{r}(t) = \langle t^2 - t, t^2 + t\rangle$ as in Examples 13.4.5 and Example 13.4.12. Find a_T and a_N.

Solution. The previous examples give $\vec{a}(t) = \langle 2, 2\rangle$ and

$$\vec{T}(t) = \left\langle \frac{2t-1}{\sqrt{8t^2+2}}, \frac{2t+1}{\sqrt{8t^2+2}}\right\rangle \text{ and } \vec{N}(t) = \left\langle \frac{2t+1}{\sqrt{8t^2+2}}, -\frac{2t-1}{\sqrt{8t^2+2}}\right\rangle.$$

While we can compute a_N using $\vec{N}(t)$, we instead demonstrate using

Video Solution

youtu.be/watch?v=8LTqjQPvh-A

another formula from Theorem 13.4.15.

$$a_T = \vec{a}(t) \cdot \vec{T}(t) = \frac{4t-2}{\sqrt{8t^2+2}} + \frac{4t+2}{\sqrt{8t^2+2}} = \frac{8t}{\sqrt{8t^2+2}}.$$

$$a_N = \sqrt{\|\vec{a}(t)\|^2 - a_T^2} = \sqrt{8 - \left(\frac{8t}{\sqrt{8t^2+2}}\right)^2} = \frac{4}{\sqrt{8t^2+2}}.$$

When $t = 2$, $a_T = \dfrac{16}{\sqrt{34}} \approx 2.74$ and $a_N = \dfrac{4}{\sqrt{34}} \approx 0.69$. We interpret this to mean that at $t = 2$, the particle is accelerating mostly by increasing speed, not by changing direction. As the path near $t = 2$ is relatively straight, this should make intuitive sense. Figure 13.4.19 gives a graph of the path for reference.

Contrast this with $t = 0$, where $a_T = 0$ and $a_N = 4/\sqrt{2} \approx 2.82$. Here the particle's speed is not changing and all acceleration is in the form of direction change.

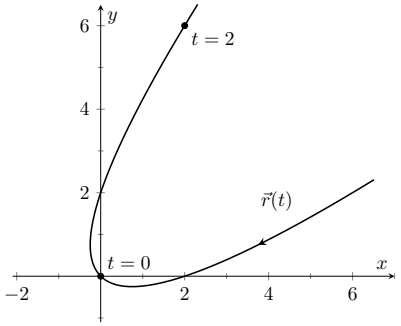


Figure 13.4.19 Graphing $\vec{r}(t)$ in Example 13.4.18

Example 13.4.20 Analyzing projectile motion.

A ball is thrown from a height of 240 ft with an initial speed of 64 $\frac{\text{ft}}{\text{s}}$ and an angle of elevation of $30°$. Find the position function $\vec{r}(t)$ of the ball and analyze a_T and a_N.

Solution. Using Key Idea 13.3.13 of Section 13.3 we form the position function of the ball:

$$\vec{r}(t) = \left\langle \big(64\cos(30°)\big)t, -16t^2 + \big(64\sin(30°)\big)t + 240 \right\rangle,$$

which we plot in Figure 13.4.21.

From this we find $\vec{v}(t) = \langle 64\cos(30°), -32t + 64\sin(30°)\rangle$ and $\vec{a}(t) = \langle 0, -32\rangle$. Computing $\vec{T}(t)$ is not difficult, and with some simplification we find

$$\vec{T}(t) = \left\langle \frac{\sqrt{3}}{\sqrt{t^2-2t+4}}, \frac{1-t}{\sqrt{t^2-2t+4}} \right\rangle.$$

With $\vec{a}(t)$ as simple as it is, finding a_T is also simple:

$$a_T = \vec{a}(t) \cdot \vec{T}(t) = \frac{32t-32}{\sqrt{t^2-2t+4}}.$$

We choose to not find $\vec{N}(t)$ and find a_N through the formula $a_N = \sqrt{\|\vec{a}(t)\|^2 - a_T^2}$:

$$a_N = \sqrt{32^2 - \left(\frac{32t-32}{\sqrt{t^2-2t+4}}\right)^2} = \frac{32\sqrt{3}}{\sqrt{t^2-2t+4}}.$$

Figure 13.4.22 gives a table of values of a_T and a_N. When $t = 0$, we see the ball's speed is decreasing; when $t = 1$ the speed of the ball is unchanged. This corresponds to the fact that at $t = 1$ the ball reaches its highest point.

After $t = 1$ we see that a_N is decreasing in value. This is because as the ball falls, its path becomes straighter and most of the acceleration is in the form of speeding up the ball, and not in changing its direction.

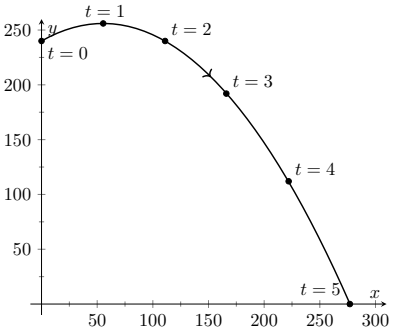


Figure 13.4.21 Plotting the position of a thrown ball, with 1s increments shown

t	a_T	a_N
0	−16	27.7
1	0	32
2	16	27.7
3	24.2	20.9
4	27.7	16
5	29.4	12.7

Figure 13.4.22 A table of values of a_T and a_N in Example 13.4.20

Our understanding of the unit tangent and normal vectors is aiding our understanding of motion. The work in Example 13.4.20 gave quantitative analysis of what we intuitively knew.

The next section provides two more important steps towards this analysis. We currently describe position only in terms of time. In everyday life, though, we often describe position in terms of distance ("The gas station is about 2 miles ahead, on the left."). The *arc length parameter* allows us to reference position in terms of distance traveled.

We also intuitively know that some paths are straighter than others — and some are curvier than others, but we lack a measurement of "curviness." The arc length parameter provides a way for us to compute *curvature*, a quantitative measurement of how curvy a curve is.

13.4.4 Exercises

Terms and Concepts

1. If $\vec{T}(t)$ is a unit tangent vector, what is $\left\|\vec{T}(t)\right\|$?

2. If $\vec{N}(t)$ is a unit normal vector, what is $\vec{N}(t) \cdot \vec{r}\,'(t)$?

3. The acceleration vector $\vec{a}(t)$ lies in the plane defined by what two vectors?

4. a_{T} measures how much the acceleration is affecting the _____ of an object.

Problems

Exercise Group. Given $\vec{r}(t)$, find $\vec{T}(t)$ and evaluate it at the indicated value of t.

5. $\vec{r}(t) = \left\langle 2t^2, t^2 - t \right\rangle$,$t = 1$

6. $\vec{r}(t) = \langle t, \cos(t) \rangle$, $t = \pi/4$.

7. $\vec{r}(t) = \left\langle \cos^3(t), \sin^3(t) \right\rangle$,$t = \pi/4$

8. $\vec{r}(t) = \langle \cos(t), \sin(t) \rangle$, $t = \pi$.

Exercise Group. Find the equation of the line tangent to the curve at the indicated t-value using the unit tangent vector. Note: these are the same problems as in Exercises 13.4.5 — Exercise 13.4.8.

9. Find the vector equation of the line tangent to $\vec{r}(t) = \left\langle 2t^2, t^2 - t \right\rangle$ at $t = 1$ using the unit tangent vector.

10. Find the vector equation of the line tangent to $\vec{r}(t) = \langle t, \cos(t) \rangle$ at $t = \pi/4$ using the unit tangent vector.

11. $\vec{r}(t) = \left\langle \cos^3(t), \sin^3(t) \right\rangle$,$t = \pi/4$

12. Find the vector equation of the line tangent to $\vec{r}(t) = \langle \cos(t), \sin(t) \rangle$ at $t = \pi$ using the unit tangent vector.

Exercise Group. In the following exercises, find $\vec{N}(t)$ using Definition 13.4.8. Confirm the result using Theorem 13.4.14.

13. $\vec{r}(t) = \langle 3\cos(t), 3\sin(t) \rangle$

14. $\vec{r}(t) = \left\langle t, t^2 \right\rangle$

15. $\vec{r}(t) = \langle \cos(t), 2\sin(t) \rangle$

16. $\vec{r}(t) = \langle e^t, e^{-t} \rangle$

Exercise Group. In the following exercises, a position function $\vec{r}(t)$ is given along with its unit tangent vector $\vec{T}(t)$ evaluated at $t = a$, for some value of a.

(a) Confirm that $\vec{T}(a)$ is as stated.

(b) Using a graph of $\vec{r}(t)$ and Theorem 13.4.14, find $\vec{N}(a)$.

17. $\vec{r}(t) = \langle 3\cos(t), 5\sin(t) \rangle$;$\vec{T}(\pi/4) = \left\langle -\frac{3}{\sqrt{34}}, \frac{5}{\sqrt{34}} \right\rangle$.

18. $\vec{r}(t) = \left\langle t, \frac{1}{t^2+1} \right\rangle$;$\vec{T}(1) = \left\langle \frac{2}{\sqrt{5}}, -\frac{1}{\sqrt{5}} \right\rangle$.

19. $\vec{r}(t) = (1 + 2\sin(t)) \langle \cos(t), \sin(t) \rangle$;$\vec{T}(0) = \left\langle \frac{2}{\sqrt{5}}, \frac{1}{\sqrt{5}} \right\rangle$.

20. $\vec{r}(t) = \left\langle \cos^3(t), \sin^3(t) \right\rangle$;$\vec{T}(\pi/4) = \left\langle -\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}} \right\rangle$.

Exercise Group. In the following exercises, find $\vec{N}(t)$.

21. $\vec{r}(t) = \langle 4t, 2\sin(t), 2\cos(t) \rangle$

22. If $\vec{r}(t) = \langle 5\cos(t), 3\sin(t), 4\sin(t) \rangle$, find $\vec{N}(t)$.

23. $\vec{r}(t) = \langle a\cos(t), a\sin(t), bt \rangle$; $a > 0$

24. If $\vec{r}(t) = \langle \cos(at), \sin(at), t \rangle$, find $\vec{N}(t)$.

Exercise Group. In the following exercises, find a_{T} and a_{N} given $\vec{r}(t)$. Be sure you can sketch $\vec{r}(t)$ on the indicated interval, and comment on the relative sizes of a_{T} and a_{N} at the indicated t values.

25. $\vec{r}(t) = \langle t, t^2 \rangle$ on $[-1, 1]$; consider $t = 0$ and $t = 1$.

26. $\vec{r}(t) = \langle t, 1/t \rangle$ on $(0, 4]$; consider $t = 1$ and $t = 2$.

27. $\vec{r}(t) = \langle 2\cos(t), 2\sin(t) \rangle$ on $[0, 2\pi]$; consider $t = 0$ and $t = \pi/2$.

28. $\vec{r}(t) = \langle \cos(t^2), \sin(t^2) \rangle$ on $(0, 2\pi]$; consider $t = \pi/2$ and $t = \pi$.

29. $\vec{r}(t) = \langle a\cos(t), a\sin(t), bt \rangle$ on $[0, 2\pi]$, where $a, b > 0$; consider $t = 0$ and $t = \pi/2$.

30. $\vec{r}(t) = \langle 5\cos(t), 4\sin(t), 3\sin(t) \rangle$ on $[0, 2\pi]$; consider $t = 0$ and $t = \pi/2$.

13.5 The Arc Length Parameter and Curvature

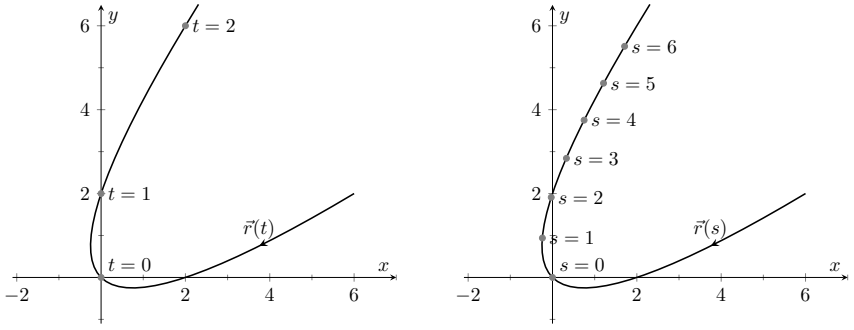
13.5.1 The Arc Length Paramtrization

In normal conversation we describe position in terms of both *time* and *distance*. For instance, imagine driving to visit a friend. If she calls and asks where you are, you might answer "I am 20 minutes from your house," or you might say "I am 10 miles from your house." Both answers provide your friend with a general idea of where you are.

Currently, our vector-valued functions have defined points with a parameter t, which we often take to represent time. Consider Figure 13.5.2(a), where $\vec{r}(t) = \langle t^2 - t, t^2 + t \rangle$ is graphed and the points corresponding to $t = 0,\ 1$ and 2 are shown. Note how the arc length between $t = 0$ and $t = 1$ is smaller than the arc length between $t = 1$ and $t = 2$; if the parameter t is time and $\vec{r}$ is position, we can say that the particle traveled faster on $[1, 2]$ than on $[0, 1]$.

youtu.be/watch?v=VHcu7bcQjsk

Figure 13.5.1 Video introduction to Section 13.5



(a) **(b)**

Figure 13.5.2 Introducing the arc length parameter

Now consider Figure 13.5.2(b), where the same graph is parametrized by a different variable s. Points corresponding to $s = 0$ through $s = 6$ are plotted. The arc length of the graph between each adjacent pair of points is 1. We can view this parameter s as distance; that is, the arc length of the graph from $s = 0$ to $s = 3$ is 3, the arc length from $s = 2$ to $s = 6$ is 4, etc. If one wants to find the point 2.5 units from an initial location (i.e., $s = 0$), one would compute $\vec{r}(2.5)$. This parameter s is very useful, and is called the *arc length parameter*.

How do we find the arc length parameter?

Start with any parametrization of $\vec{r}$. We can compute the arc length of the graph of $\vec{r}$ on the interval $[0, t]$ with

$$\text{arc length } = \int_0^t \|\vec{r}\,'(u)\|\ du.$$

We can turn this into a function: as t varies, we find the arc length s from 0 to t. This function is

$$s(t) = \int_0^t \|\vec{r}\,'(u)\|\ du. \tag{13.5.1}$$

This establishes a relationship between s and t. Knowing this relationship explicitly, we can rewrite $\vec{r}(t)$ as a function of s: $\vec{r}(s)$. We demonstrate this in an example.

Example 13.5.3 Finding the arc length parameter.

Let $\vec{r}(t) = \langle 3t - 1, 4t + 2 \rangle$. Parametrize $\vec{r}$ with the arc length parameter s.

Video Solution

youtu.be/watch?v=UZkbOv5zW1U

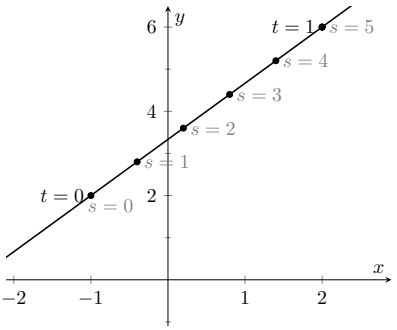


Figure 13.5.4 Graphing $\vec{r}$ in Example 13.5.3 with parameters t and s

Solution. Using Equation (13.5.1), we write

$$s(t) = \int_0^t \|\vec{r}\,'(u)\|\, du.$$

We can integrate this, explicitly finding a relationship between s and t:

$$\begin{aligned} s(t) &= \int_0^t \|\vec{r}\,'(u)\|\, du \\ &= \int_0^t \sqrt{3^2+4^2}\, du \\ &= \int_0^t 5\, du \\ &= 5t. \end{aligned}$$

Since $s = 5t$, we can write $t = s/5$ and replace t in $\vec{r}(t)$ with $s/5$:

$$\vec{r}(s) = \langle 3(s/5) - 1, 4(s/5) + 2\rangle = \left\langle \frac{3}{5}s - 1, \frac{4}{5}s + 2 \right\rangle.$$

Clearly, as shown in Figure 13.5.4, the graph of $\vec{r}$ is a line, where $t = 0$ corresponds to the point $(-1, 2)$. What point on the line is 2 units away from this initial point? We find it with $\vec{r}(2) = \langle 1/5, 18/5\rangle$.

Is the point $(1/5, 18/5)$ really 2 units away from $(-1, 2)$? We use the Distance Formula to check:

$$d = \sqrt{\left(\frac{1}{5} - (-1)\right)^2 + \left(\frac{18}{5} - 2\right)^2} = \sqrt{\frac{36}{25} + \frac{64}{25}} = \sqrt{4} = 2.$$

Yes, $\vec{r}(2)$ is indeed 2 units away, in the direction of travel, from the initial point.

Things worked out very nicely in Example 13.5.3; we were able to establish directly that $s = 5t$. Usually, the arc length parameter is much more difficult to describe in terms of t, a result of integrating a square root. There are a number of things that we can learn about the arc length parameter from Equation (13.5.1), though, that are incredibly useful.

First, take the derivative of s with respect to t. The Fundamental Theorem of Calculus (see Theorem 5.4.7) states that

$$\frac{ds}{dt} = s'(t) = \|\vec{r}\,'(t)\|. \tag{13.5.2}$$

Letting t represent time and $\vec{r}(t)$ represent position, we see that the rate of change of s with respect to t is speed; that is, the rate of change of "distance traveled" is speed, which should match our intuition.

The Chain Rule states that

$$\begin{aligned} \frac{d\vec{r}}{dt} &= \frac{d\vec{r}}{ds}\cdot\frac{ds}{dt} \\ \vec{r}\,'(t) &= \vec{r}\,'(s)\cdot\|\vec{r}\,'(t)\|. \end{aligned}$$

Solving for $\vec{r}\,'(s)$, we have

$$\vec{r}\,'(s) = \frac{\vec{r}\,'(t)}{\|\vec{r}\,'(t)\|} = \vec{T}(t), \tag{13.5.3}$$

where $\vec{T}(t)$ is the unit tangent vector. Equation (13.5.3) is often misinterpreted, as one is tempted to think it states $\vec{r}\,'(t) = \vec{T}(t)$, but there is a big difference between $\vec{r}\,'(s)$ and $\vec{r}\,'(t)$. The key to take from it is that $\vec{r}\,'(s)$ is a unit vector. In fact, the following theorem states that this characterizes the arc length parameter.

Theorem 13.5.5 Arc Length Parameter.

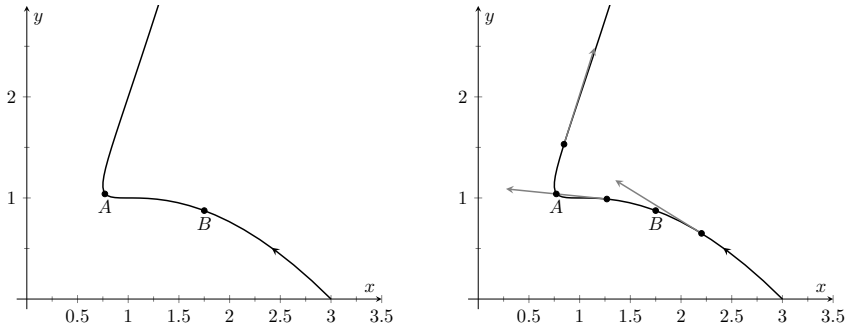
Let $\vec{r}(s)$ be a vector-valued function. The parameter s is the arc length parameter if, and only if, $\|\vec{r}\,'(s)\| = 1$.

youtu.be/watch?v=FLydDuy3r2I

Figure 13.5.6 Video presentation of Theorem 13.5.5

13.5.2 Curvature

Consider points A and B on the curve graphed in Figure 13.5.7(a). One can readily argue that the curve curves more sharply at A than at B. It is useful to use a number to describe how sharply the curve bends; that number is the *curvature* of the curve.



(a) **(b)**

Figure 13.5.7 Establishing the concept of curvature

We derive this number in the following way. Consider Figure 13.5.7(b), where unit tangent vectors are graphed around points A and B. Notice how the direction of the unit tangent vector changes quite a bit near A, whereas it does not change as much around B. This leads to an important concept: measuring the rate of change of the unit tangent vector with respect to arc length gives us a measurement of curvature.

Definition 13.5.8 Curvature.

Let $\vec{r}(s)$ be a vector-valued function where s is the arc length parameter. The curvature κ of the graph of $\vec{r}(s)$ is

$$\kappa = \left\| \frac{d\vec{T}}{ds} \right\| = \left\| \vec{T}\,'(s) \right\|.$$

youtu.be/watch?v=sjz_eHxgbLg

Figure 13.5.9 Video presentation of Definition 13.5.8

If $\vec{r}(s)$ is parametrized by the arc length parameter, then

$$\vec{T}(s) = \frac{\vec{r}\,'(s)}{\|\vec{r}\,'(s)\|} \text{ and } \vec{N}(s) = \frac{\vec{T}\,'(s)}{\left\|\vec{T}\,'(s)\right\|}.$$

Having defined $\left\|\vec{T}\,'(s)\right\| = \kappa$, we can rewrite the second equation as

$$\vec{T}\,'(s) = \kappa \vec{N}(s). \tag{13.5.4}$$

We already knew that $\vec{T}'(s)$ is in the same direction as $\vec{N}(s)$; that is, we can think of $\vec{T}(s)$ as being "pulled" in the direction of $\vec{N}(s)$. How "hard" is it being pulled? By a factor of κ. When the curvature is large, $\vec{T}(s)$ is being "pulled hard" and the direction of $\vec{T}(s)$ changes rapidly. When κ is small, $T(s)$ is not being pulled hard and hence its direction is not changing rapidly.

We use Definition 13.5.8 to find the curvature of the line in Example 13.5.3.

Video Solution

youtu.be/watch?v=SZigON0uhqU

Example 13.5.10 Finding the curvature of a line.

Use Definition 13.5.8 to find the curvature of $\vec{r}(t) = \langle 3t-1, 4t+2\rangle$.

Solution. In Example 13.5.3, we found that the arc length parameter was defined by $s = 5t$, so $\vec{r}(s) = \langle 3s/5 - 1, 4s/5 + 2\rangle$ parametrized $\vec{r}$ with the arc length parameter. To find κ, we need to find $\vec{T}'(s)$.

$$\begin{aligned}\vec{T}(s) &= \vec{r}'(s) \text{ (recall this is a unit vector)}\\ &= \langle 3/5, 4/5\rangle\,.\end{aligned}$$

Therefore

$$\vec{T}'(s) = \langle 0, 0\rangle$$

and

$$\kappa = \left|\left| \vec{T}'(s) \right|\right| = 0.$$

It probably comes as no surprise that the curvature of a line is 0. (How "curvy" is a line? It is not curvy at all.)

While the definition of curvature is a beautiful mathematical concept, it is nearly impossible to use most of the time; writing $\vec{r}$ in terms of the arc length parameter is generally very hard. Fortunately, there are other methods of calculating this value that are much easier. There is a tradeoff: the definition is "easy" to understand though hard to compute, whereas these other formulas are easy to compute though it may be hard to understand why they work.

Theorem 13.5.11 Formulas for Curvature.

Let C be a smooth curve in the plane or in space.

1. *If C is defined by $y = f(x)$, then*

$$\kappa = \frac{|f''(x)|}{\left(1 + \left(f'(x)\right)^2\right)^{3/2}}.$$

2. *If C is defined as a vector-valued function in the plane, $\vec{r}(t) = \langle x(t), y(t)\rangle$, then*

$$\kappa = \frac{|x'y'' - x''y'|}{\left((x')^2 + (y')^2\right)^{3/2}}.$$

3. *If C is defined in space by a vector-valued function $\vec{r}(t)$, then*

$$\kappa = \frac{\left\|\vec{T}'(t)\right\|}{\|\vec{r}'(t)\|} = \frac{\|\vec{r}'(t) \times \vec{r}''(t)\|}{\|\vec{r}'(t)\|^3} = \frac{\vec{a}(t) \cdot \vec{N}(t)}{\|\vec{v}(t)\|^2}.$$

We practice using these formulas.

Example 13.5.12 Finding the curvature of a circle.

Find the curvature of a circle with radius r, defined by $\vec{c}(t) = \langle r\cos(t), r\sin(t)\rangle$.

Solution. Before we start, we should expect the curvature of a circle to be constant, and not dependent on t. (Why?)

We compute κ using the second part of Theorem 13.5.11.

$$\begin{aligned}\kappa &= \frac{|(-r\sin(t))(-r\sin(t)) - (-r\cos(t))(r\cos(t))|}{\left((-r\sin(t))^2 + (r\cos(t))^2\right)^{3/2}} \\ &= \frac{r^2(\sin^2(t) + \cos^2(t))}{\left(r^2(\sin^2(t) + \cos^2(t))\right)^{3/2}} \\ &= \frac{r^2}{r^3} = \frac{1}{r}.\end{aligned}$$

We have found that a circle with radius r has curvature $\kappa = 1/r$.

Video Solution

youtu.be/watch?v=NgffyzJBrTc

Example 13.5.12 gives a great result. Before this example, if we were told "The curve has a curvature of 5 at point A," we would have no idea what this really meant. Is 5 "big" — does is correspond to a really sharp turn, or a not-so-sharp turn? Now we can think of 5 in terms of a circle with radius 1/5. Knowing the units (inches vs. miles, for instance) allows us to determine how sharply the curve is curving.

Let a point P on a smooth curve C be given, and let κ be the curvature of the curve at P. A circle that:

- passes through P,
- lies on the concave side of C,
- has a common tangent line as C at P and
- has radius $r = 1/\kappa$ (hence has curvature κ)

is the *osculating circle*, or *circle of curvature*, to C at P, and r is the *radius of curvature*. Figure 13.5.13 shows the graph of the curve seen earlier in Figure 13.5.7 and its osculating circles at A and B. A sharp turn corresponds to a circle with a small radius; a gradual turn corresponds to a circle with a large radius. Being able to think of curvature in terms of the radius of a circle is very useful.

(The word "osculating" comes from a Latin word related to kissing; an osculating circle "kisses" the graph at a particular point. Many beautiful ideas in mathematics have come from studying the osculating circles to a curve.)

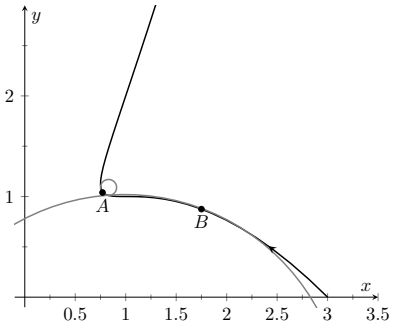


Figure 13.5.13 Illustrating the osculating circles for the curve seen in Figure 13.5.7

youtu.be/watch?v=x8TWhPo42QE

Figure 13.5.14 The osculating circle

Example 13.5.15 Finding curvature.

Find the curvature of the parabola defined by $y = x^2$ at the vertex and at $x = 1$.

Solution. We use the first formula found in Theorem 13.5.11.

$$\kappa(x) = \frac{|2|}{\left(1 + (2x)^2\right)^{3/2}}$$

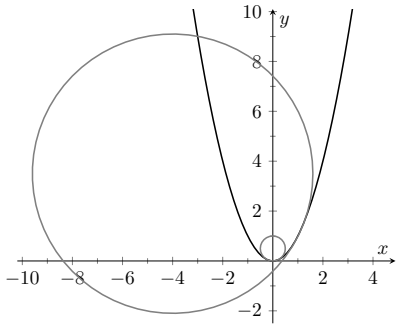


Figure 13.5.16 Examining the curvature of $y = x^2$

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

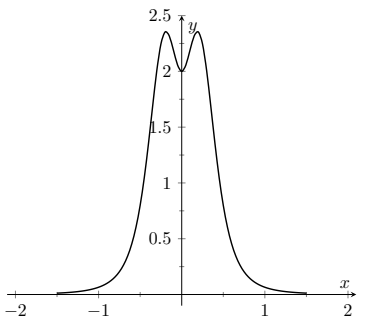
At the vertex ($x = 0$), the curvature is $\kappa = 2$. At $x = 1$, the curvature is $\kappa = 2/(5)^{3/2} \approx 0.179$. So at $x = 0$, the curvature of $y = x^2$ is that of a circle of radius $1/2$; at $x = 1$, the curvature is that of a circle with radius $\approx 1/0.179 \approx 5.59$. This is illustrated in Figure 13.5.16. At $x = 3$, the curvature is 0.009; the graph is nearly straight as the curvature is very close to 0.

Example 13.5.17 Finding curvature.

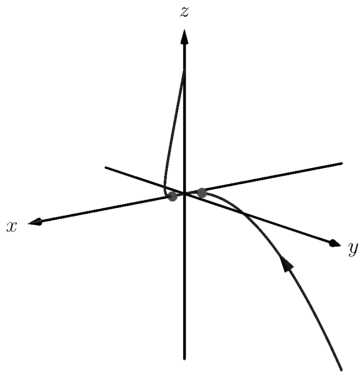
Find where the curvature of $\vec{r}(t) = \langle t, t^2, 2t^3\rangle$ is maximized.

Solution. We use the third formula in Theorem 13.5.11 as $\vec{r}(t)$ is defined in space. We leave it to the reader to verify that

$$\vec{r}'(t) = \langle 1, 2t, 6t^2\rangle, \vec{r}''(t) = \langle 0, 2, 12t\rangle, \text{ and } \vec{r}'(t)\times\vec{r}''(t) = \langle 12t^2, -12t, 2\rangle.$$



(a)



(b)

Figure 13.5.18 Understanding the curvature of a curve in space

Thus

$$\begin{aligned}\kappa(t) &= \frac{\|\vec{r}'(t)\times\vec{r}''(t)\|}{\|\vec{r}'(t)\|^3}\\ &= \frac{\left\|\langle 12t^2, -12t, 2\rangle\right\|}{\|\langle 1, 2t, 6t^2\rangle\|^3}\\ &= \frac{\sqrt{144t^4+144t^2+4}}{\left(\sqrt{1+4t^2+36t^4}\right)^3}\end{aligned}$$

While this is not a particularly "nice" formula, it does explicitly tell us what the curvature is at a given t value. To maximize $\kappa(t)$, we should solve $\kappa'(t) = 0$ for t. This is doable, but *very* time consuming. Instead, consider the graph of $\kappa(t)$ as given in Figure 13.5.18(a). We see that κ is maximized at two t values; using a numerical solver, we find these values are $t \approx \pm 0.189$. In Figure 13.5.18(b) we graph $\vec{r}(t)$ and indicate the points where curvature is maximized.

13.5.3 Curvature and Motion

Let $\vec{r}(t)$ be a position function of an object, with velocity $\vec{v}(t) = \vec{r}\,'(t)$ and acceleration $\vec{a}(t) = \vec{r}\,''(t)$. In Section 13.4 we established that acceleration is in the plane formed by $\vec{T}(t)$ and $\vec{N}(t)$, and that we can find scalars a_T and a_N such that

$$\vec{a}(t) = a_T\vec{T}(t) + a_N\vec{N}(t).$$

Theorem 13.4.15 gives formulas for a_T and a_N:

$$a_T = \frac{d}{dt}\Big(\|\vec{v}(t)\| \Big) \text{ and } a_N = \frac{\|\vec{v}(t) \times \vec{a}(t)\|}{\|\vec{v}(t)\|}.$$

We understood that the amount of acceleration in the direction of $\vec{T}$ relates only to how the speed of the object is changing, and that the amount of acceleration in the direction of $\vec{N}$ relates to how the direction of travel of the object is changing. (That is, if the object travels at constant speed, $a_T = 0$; if the object travels in a constant direction, $a_N = 0$.)

In Equation (13.5.2) at the beginning of this section, we found $s'(t) = \|\vec{v}(t)\|$. We can combine this fact with the above formula for a_T to write

$$a_T = \frac{d}{dt}\Big(\|\vec{v}(t)\| \Big) = \frac{d}{dt}\big(s'(t)\big) = s''(t).$$

Since $s'(t)$ is speed, $s''(t)$ is the rate at which speed is changing with respect to time. We see once more that the component of acceleration in the direction of travel relates only to speed, not to a change in direction.

Now compare the formula for a_N above to the formula for curvature in Theorem 13.5.11:

$$a_N = \frac{\|\vec{v}(t) \times \vec{a}(t)\|}{\|\vec{v}(t)\|} \text{ and } \kappa = \frac{\|\vec{r}\,'(t) \times \vec{r}\,''(t)\|}{\|\vec{r}\,'(t)\|^3} = \frac{\|\vec{v}(t) \times \vec{a}(t)\|}{\|\vec{v}(t)\|^3}.$$

Thus

$$\begin{aligned} a_N &= \kappa \|\vec{v}(t)\|^2 \\ &= \kappa\Big(s'(t)\Big)^2 \end{aligned}$$

This last equation shows that the component of acceleration that changes the object's direction is dependent on two things: the curvature of the path and the speed of the object.

Imagine driving a car in a clockwise circle. You will naturally feel a force pushing you towards the door (more accurately, the door is pushing you as the car is turning and you want to travel in a straight line). If you keep the radius of the circle constant but speed up (i.e., increasing $s'(t)$), the door pushes harder against you (a_N has increased). If you keep your speed constant but tighten the turn (i.e., increase κ), once again the door will push harder against you.

Putting our new formulas for a_T and a_N together, we have

$$\vec{a}(t) = s''(t)\vec{T}(t) + \kappa \|\vec{v}(t)\|^2 \vec{N}(t).$$

This is not a particularly practical way of finding a_T and a_N, but it reveals some great concepts about how acceleration interacts with speed and the shape of a curve.

Example 13.5.19 Curvature and road design.

The minimum radius of the curve in a highway cloverleaf is determined by the operating speed, as given in the table in Table 13.5.20. For each curve and speed, compute a_N.

Solution.

Table 13.5.20 Operating speed and minimum radius in highway cloverleaf design

Operating Speed (mph)	Minimum Radius (ft)
35	310
40	430
45	540

Using Equation (13.5.4), we can compute the acceleration normal to the curve in each case. We start by converting each speed from "miles per hour" to "feet per second" by multiplying by $5280/3600$.

$$\begin{aligned} \text{35 mph, 310 ft} \Rightarrow 51.33\,\text{ft/s},\ \kappa = 1/310 \\ a_N &= \kappa \left\|\vec{v}(t)\right\|^2 \\ &= \frac{1}{310}\left(51.33\right)^2 \\ &= 8.50\,\text{ft/s}^2. \end{aligned}$$

$$\begin{aligned} \text{40 mph, 430 ft} \Rightarrow 58.67\,\text{ft/s},\ \kappa = 1/430 \\ a_N &= \frac{1}{430}\left(58.67\right)^2 \\ &= 8.00\,\text{ft/s}^2. \end{aligned}$$

$$\begin{aligned} \text{45 mph, 540 ft} \Rightarrow 66\,\text{ft/s},\ \kappa = 1/540 \\ a_N &= \frac{1}{540}\left(66\right)^2 \\ &= 8.07\,\text{ft/s}^2. \end{aligned}$$

Note that each acceleration is similar; this is by design. Considering the classic "Force = mass × acceleration" formula, this acceleration must be kept small in order for the tires of a vehicle to keep a "grip" on the road. If one travels on a turn of radius 310 ft at a rate of 50 mph, the acceleration is double, at $17.35\ \frac{\text{ft}}{\text{s}^2}$. If the acceleration is too high, the frictional force created by the tires may not be enough to keep the car from sliding. Civil engineers routinely compute a "safe" design speed, then subtract 5-10 mph to create the posted speed limit for additional safety.

We end this chapter with a reflection on what we've covered. We started with vector-valued functions, which may have seemed at the time to be just another way of writing parametric equations. However, we have seen that the vector perspective has given us great insight into the behavior of functions and the study of motion. Vector-valued position functions convey displacement, dis-

tance traveled, speed, velocity, acceleration and curvature information, each of which has great importance in science and engineering.

13.5.4 Exercises

Terms and Concepts

1. It is common to describe position in terms of both _____ and/or _____.

2. A measure of the "curviness" of a curve is _____.

3. Give two shapes with constant curvature.

4. Describe in your own words what an "osculating circle" is.

5. Complete the identity: $\vec{T}'(s) =$ _____ $\vec{N}(s)$.

6. Given a position function $\vec{r}(t)$, how are a_T and a_N affected by the curvature?

Problems

Exercise Group. In the following exercises, a position function $\vec{r}(t)$ is given, where $t = 0$ corresponds to the initial position. Find the arc length parameter s, and rewrite $\vec{r}(t)$ in terms of s; that is, find $\vec{r}(s)$.

7. $\vec{r}(t) = \langle 2t, t, -2t \rangle$

8. $\vec{r}(t) = \langle 7\cos(t), 7\sin(t) \rangle$.

9. $\vec{r}(t) = \langle 3\cos(t), 3\sin(t), 2t \rangle$

10. $\vec{r}(t) = \vec{r}(t) = \langle 5\cos(t), 13\sin(t), 12\cos(t) \rangle$.

Exercise Group. In the following exercises, a curve C is described along with 2 points on C.

(a) Using a sketch, determine at which of these points the curvature is greater.

(b) Find the curvature κ of C, and evaluate κ at each of the 2 given points.

11. C is defined by $y = x^3 - x$; points given at $x = 0$ and $x = 1/2$.

12. C is defined by $y = \frac{1}{x^2+1}$; points given at $x = 0$ and $x = 2$.
The curvature at $x = 0$ is (☐ greater than ☐ equal to ☐ less than) the curvature at $x = 2$.
$\kappa(0) =$_____

13. C is defined by $y = \cos(x)$; points given at $x = 0$ and $x = \pi/2$.

14. C is defined by $y = \sqrt{1 - x^2}$ on $(-1, 1)$; points given at $x = 0$ and $x = 1/2$.

15. C is defined by $\vec{r}(t) = \langle \cos(t), \sin(2t) \rangle$; points given at $t = 0$ and $t = \pi/4$.
The curvature at $t = 0$ is (☐ greater than ☐ equal to ☐ less than) the curvature at $t = \pi/4$.
$\kappa(0) =$_____

16. C is defined by $\vec{r}(t) = \left\langle \cos^2(t), \sin(t)\cos(t) \right\rangle$; points given at $t = 0$ and $t = \pi/3$.

17. C is defined by $\vec{r}(t) = \left\langle t^2 - 1, t^3 - t \right\rangle$; points given at $t = 0$ and $t = 5$.

18. C is defined by $\vec{r}(t) = \langle \tan(t), \sec(t) \rangle$; points given at $t = 0$ and $t = \pi/6$.
The curvature at $t = 0$ is (☐ greater than ☐ equal to ☐ less than) the curvature at $t = \pi/6$.
$\kappa(0) =$_____

19. C is defined by $\vec{r}(t) = \langle 4t + 2, 3t - 1, 2t + 5 \rangle$; points given at $t = 0$ and $t = 1$.

20. C is defined by $\vec{r}(t) = \left\langle t^3 - t, t^3 - 4, t^2 - 1 \right\rangle$; points given at $t = 0$ and $t = 1$.
The curvature at $t = 0$ is (☐ greater than ☐ equal to ☐ less than) the curvature at $t = 1$.
$\kappa(0) =$_____

21. C is defined by $\vec{r}(t) = \langle 3\cos(t), 3\sin(t), 2t\rangle$; points given at $t = 0$ and $t = \pi/2$.

22. C is defined by $\vec{r}(t) = \langle 5\cos(t), 13\sin(t), 12\cos(t)\rangle$. Points given at $t = 0$ and $t = \pi/2$.
The curvature at $t = 0$ is (☐ greater than ☐ equal to ☐ less than) the curvature at $t = \pi/2$.
$\kappa(0) =$_____

Exercise Group. Find the value of x or t where curvature is maximized.

23. $y = \frac{1}{6}x^3$

24. $y = \sin(x)$

25. $\vec{r}(t) = \left\langle t^2 + 2t, 3t - t^2\right\rangle$

26. $\vec{r}(t) = \langle t, 4/t, 3/t\rangle$

Exercise Group. Find the radius of curvature at the indicated value.

27. $y = \tan(x)$, at $x = \pi/4$

28. $y = x^2 + x - 3$ at $x = 1$

29. $\vec{r}(t) = \langle\cos(t), \sin(3t)\rangle$, at $t = 0$

30. $\vec{r}(t) = \langle 5\cos(3t), t\rangle$ at $t = 0$

Exercise Group. Find the equation of the osculating circle to the curve at the indicated t-value.

31. $\vec{r}(t) = \left\langle t, t^2\right\rangle$, at $t = 0$

32. $\vec{r}(t) = \langle 3\cos(t), \sin(t)\rangle$ at $t = 0$

33. $\vec{r}(t) = \langle 3\cos(t), \sin(t)\rangle$, at $t = \pi/2$

34. $\vec{r}(t) = \left\langle t^2 - t, t^2 + t\right\rangle$ at $t = 0$

Chapter 14

Functions of Several Variables, Continued

In Chapter 11 we introduced functions of several variables, and considered limits, continuity, and partial derivatives. This chapter continues the study of functions of several variables in more detail. We begin in Section 14.1 with what it means for a multivariable function to be *differentiable*. We then continue with multivariable analogues of elements from single variable calculus, such as the chain rule and extreme values.

14.1 Differentiability and the Total Differential

We studied *differentials* in Section 4.4, where Definition 4.4.5 states that if $y = f(x)$ and f is differentiable, then $dy = f'(x)dx$. One important use of this differential is in Integration by Substitution. Another important application is approximation. Let $\Delta x = dx$ represent a change in x. When dx is small, $dy \approx \Delta y$, the change in y resulting from the change in x. Fundamental in this understanding is this: as dx gets small, the difference between Δy and dy goes to 0. Another way of stating this: as dx goes to 0, the *error* in approximating Δy with dy goes to 0.

We extend this idea to functions of two variables. Let $z = f(x, y)$, and let $\Delta x = dx$ and $\Delta y = dy$ represent changes in x and y, respectively. Let $\Delta z = f(x + dx, y + dy) - f(x, y)$ be the change in z over the change in x and y. Recalling that f_x and f_y give the instantaneous rates of z-change in the x- and y-directions, respectively, we can approximate Δz with $dz = f_x dx + f_y dy$; in words, the total change in z is approximately the change caused by changing x plus the change caused by changing y. In a moment we give an indication of whether or not this approximation is any good. First we give a name to dz.

14.1.1 The Total Differential

Definition 14.1.1 Total Differential.

Let $z = f(x, y)$ be continuous on a set S. Let dx and dy represent changes in x and y, respectively. Where the partial derivatives f_x and

f_y exist, the **total differential** of z is

$$dz = f_x(x, y)\,dx + f_y(x, y)\,dy.$$

Example 14.1.2 Finding the total differential.

Let $z = x^4e^{3y}$. Find dz.

Solution. We compute the partial derivatives: $f_x = 4x^3e^{3y}$ and $f_y = 3x^4e^{3y}$. Following Definition 14.1.1, we have

$$dz = 4x^3e^{3y}dx + 3x^4e^{3y}dy.$$

From Definition 14.1.1, we can write

$$dz = \langle f_x, f_y\rangle \cdot \langle dx, dy\rangle.$$

While not explored in this section, the vector $\langle f_x, f_y\rangle$ is seen again in the next section and fully defined in Section 14.3.

We *can* approximate Δz with dz, but as with all approximations, there is error involved. A good approximation is one in which the error is small. At a given point (x_0, y_0), let E_x and E_y be functions of dx and dy such that $E_xdx + E_ydy$ describes this error. Then

$$\begin{aligned}\Delta z &= dz + E_xdx + E_ydy\\ &= f_x(x_0, y_0)dx + f_y(x_0, y_0)dy + E_xdx + E_ydy.\end{aligned}$$

If the approximation of Δz by dz is good, then as dx and dy get small, so does $E_xdx + E_ydy$. The approximation of Δz by dz is even better if, as dx and dy go to 0, so do E_x and E_y. This leads us to our definition of differentiability.

Definition 14.1.3 Multivariable Differentiability.

Let $z = f(x, y)$ be defined on a set S containing (x_0, y_0) where $f_x(x_0, y_0)$ and $f_y(x_0, y_0)$ exist. Let dz be the total differential of z at (x_0, y_0), let $\Delta z = f(x_0 + dx, y_0 + dy) - f(x_0, y_0)$, and let E_x and E_y be functions of dx and dy such that

$$\Delta z = dz + E_xdx + E_ydy.$$

1. We say f is *differentiable at* (x_0, y_0) if, given $\varepsilon > 0$, there is a $\delta > 0$ such that if $\|\langle dx, dy\rangle\| < \delta$, then $\|\langle E_x, E_y\rangle\| < \varepsilon$. That is, as dx and dy go to 0, so do E_x and E_y.
2. We say f is *differentiable on* S if f is differentiable at every point in S. If f is differentiable on $\mathbb{R}^2$, we say that f is *differentiable everywhere*.

youtu.be/watch?v=hJaLxYbPgdM

Figure 14.1.4 Another approach to defining differentiability

Example 14.1.5 Showing a function is differentiable.

Show $f(x, y) = xy + 3y^2$ is differentiable using Definition 14.1.3.

Solution. We begin by finding $f(x + dx, y + dy)$, Δz, f_x and f_y.

$$\begin{aligned}f(x + dx, y + dy) &= (x + dx)(y + dy) + 3(y + dy)^2\\ &= xy + xdy + ydx + dxdy + 3y^2 + 6ydy + 3dy^2.\end{aligned}$$

$\Delta z = f(x + dx, y + dy) - f(x, y)$, so

$$\Delta z = xdy + ydx + dxdy + 6ydy + 3dy^2.$$

It is straightforward to compute $f_x = y$ and $f_y = x + 6y$. Consider

once more Δz:

$$\begin{aligned}\Delta z &= xdy + ydx + dxdy + 6ydy + 3dy^2 \qquad \text{(now reorder)}\\ &= ydx + xdy + 6ydy + dxdy + 3dy^2\\ &= \underbrace{(y)}_{f_x}\,dx + \underbrace{(x+6y)}_{f_y}\,dy + \underbrace{(dy)}_{E_x}\,dx + \underbrace{(3dy)}_{E_y}\,dy\\ &= f_x dx + f_y dy + E_x dx + E_y dy.\end{aligned}$$

With $E_x = dy$ and $E_y = 3dy$, it is clear that as dx and dy go to 0, E_x and E_y also go to 0. Since this did not depend on a specific point (x_0, y_0), we can say that $f(x, y)$ is differentiable for all pairs (x, y) in $\mathbb{R}^2$, or, equivalently, that f is differentiable everywhere.

youtu.be/watch?v=rYvDghlxlUA

Figure 14.1.6 Establishing differentiability using the definition in Figure 14.1.4

Our intuitive understanding of differentiability of functions $y = f(x)$ of one variable was that the graph of f was "smooth." A similar intuitive understanding of functions $z = f(x, y)$ of two variables is that the surface defined by f is also "smooth," not containing cusps, edges, breaks, etc. The following theorem states that differentiable functions are continuous, followed by another theorem that provides a more tangible way of determining whether a great number of functions are differentiable or not.

Theorem 14.1.7 Continuity and Differentiability of Multivariable Functions.

Let $z = f(x, y)$ be defined on a set S containing (x_0, y_0). If f is differentiable at (x_0, y_0), then f is continuous at (x_0, y_0).

Theorem 14.1.8 Differentiability of Multivariable Functions.

Let $z = f(x, y)$ be defined on a set S. If f_x and f_y are both continuous on S, then f is differentiable on S.

The theorems assure us that essentially all functions that we see in the course of our studies here are differentiable (and hence continuous) on their natural domains. There is a difference between Definition 14.1.3 and Theorem 14.1.8, though: it is possible for a function f to be differentiable yet f_x and/or f_y is *not* continuous. Such strange behavior of functions is a source of delight for many mathematicians.

When f_x and f_y exist at a point but are not continuous at that point, we need to use other methods to determine whether or not f is differentiable at that point.

For instance, consider the function

$$f(x,y) = \begin{cases} \frac{xy}{x^2+y^2} & (x,y) \neq (0,0) \\ 0 & (x,y) = (0,0) \end{cases}.$$

We can find $f_x(0,0)$ and $f_y(0,0)$ using Definition 11.3.3:

$$\begin{aligned} f_x(0,0) &= \lim_{h\to 0} \frac{f(0+h,0) - f(0,0)}{h} \\ &= \lim_{h\to 0} \frac{0}{h^2} = 0; \\ f_y(0,0) &= \lim_{h\to 0} \frac{f(0,0+h) - f(0,0)}{h} \end{aligned}$$

$$= \lim_{h\to 0} \frac{0}{h^2} = 0.$$

Both f_x and f_y *exist* at $(0,0)$, but they are not continuous at $(0,0)$, as

$$f_x(x,y) = \frac{y(y^2-x^2)}{(x^2+y^2)^2} \qquad \text{and} \qquad f_y(x,y) = \frac{x(x^2-y^2)}{(x^2+y^2)^2}$$

are not continuous at $(0,0)$. (Take the limit of f_x as $(x,y) \to (0,0)$ along the x- and y-axes; they give different results.) So even though f_x and f_y *exist* at every point in the xy-plane, they are not continuous. Therefore it is possible, by Theorem 14.1.8, for f to not be differentiable.

Indeed, it is not. One can show that f is not continuous at $(0,0)$ (see Example 11.2.11), and by Theorem 14.1.7, this means f is not differentiable at $(0,0)$.

14.1.2 Approximating with the Total Differential

By the definition, when f is differentiable dz is a good approximation for Δz when dx and dy are small. We give some simple examples of how this is used here.

Example 14.1.9 Approximating with the total differential.

Let $z = \sqrt{x}\sin(y)$. Approximate $f(4.1, 0.8)$.

Solution. Recognizing that $\pi/4 \approx 0.785 \approx 0.8$, we can approximate $f(4.1, 0.8)$ using $f(4, \pi/4)$. We can easily compute $f(4,\pi/4) = \sqrt{4}\sin(\pi/4) = 2\left(\frac{\sqrt{2}}{2}\right) = \sqrt{2} \approx 1.414$. Without calculus, this is the best approximation we could reasonably come up with. The total differential gives us a way of adjusting this initial approximation to hopefully get a more accurate answer.

We let $\Delta z = f(4.1, 0.8) - f(4, \pi/4)$. The total differential dz is approximately equal to Δz, so

$$f(4.1,0.8) - f(4,\pi/4) \approx dz \Rightarrow f(4.1,0.8) \approx dz + f(4,\pi/4). \quad (14.1.1)$$

To find dz, we need f_x and f_y.

$$f_x(x,y) = \frac{\sin(y)}{2\sqrt{x}} \Rightarrow \qquad f_x(4,\pi/4) = \frac{\sin(\pi)/4}{2\sqrt{4}} = \frac{\sqrt{2}/2}{4} = \sqrt{2}/8.$$

$$f_y(x,y) = \sqrt{x}\cos(y) \Rightarrow \qquad f_y(4,\pi/4) = \sqrt{4}\frac{\sqrt{2}}{2} = \sqrt{2}.$$

Approximating 4.1 with 4 gives $dx = 0.1$; approximating 0.8 with $\pi/4$ gives $dy \approx 0.015$. Thus

$$\begin{aligned} dz(4,\pi/4) &= f_x(4,\pi/4)(0.1) + f_y(4,\pi/4)(0.015) \\ &= \frac{\sqrt{2}}{8}(0.1) + \sqrt{2}(0.015) \\ &\approx 0.039. \end{aligned}$$

Returning to Equation (14.1.1), we have

$$f(4.1, 0.8) \approx 0.039 + 1.414 = 1.4531.$$

We, of course, can compute the actual value of $f(4.1, 0.8)$ with a calculator; the actual value, accurate to 5 places after the decimal, is 1.45254. Obviously our approximation is quite good.

The point of the previous example was *not* to develop an approximation method for known functions. After all, we can very easily compute $f(4.1, 0.8)$ using readily available technology. Rather, it serves to illustrate how well this method of approximation works, and to reinforce the following concept:

"New position = old position $+$ amount of change," so

"New position $\approx$ old position + approximate amount of change."

In the previous example, we could easily compute $f(4, \pi/4)$ and could approximate the amount of z-change when computing $f(4.1, 0.8)$, letting us approximate the new z-value.

It may be surprising to learn that it is not uncommon to know the values of f, f_x and f_y at a particular point without actually knowing the function f. The total differential gives a good method of approximating f at nearby points.

Example 14.1.10 Approximating an unknown function.

Given that $f(2, -3) = 6$, $f_x(2, -3) = 1.3$ and $f_y(2, -3) = -0.6$, approximate $f(2.1, -3.03)$.

Solution. The total differential approximates how much f changes from the point $(2, -3)$ to the point $(2.1, -3.03)$. With $dx = 0.1$ and $dy = -0.03$, we have

$$\begin{aligned} dz &= f_x(2, -3)dx + f_y(2, -3)dy \\ &= 1.3(0.1) + (-0.6)(-0.03) \\ &= 0.148. \end{aligned}$$

The change in z is approximately 0.148, so we approximate $f(2.1, -3.03) \approx 6.148$.

14.1.3 Tangent Plane Approximation

Recall from Chapter 2 that in one variable, the essence of differentiability is the tangent line approximation. This idea is emphasized in Section 4.4, where we first introduced the differential.

In Section 11.3.2 we saw that the partial derivatives of a function $f(x, y)$ can be used to define the tangent *plane* to a graph $z = f(x, y)$. We will now see that this plane plays the same role for functions of two variables as the tangent line to a graph $y = f(x)$ for a function of one variable.

Recall from Definition 4.4.4 that for a function $f(x)$, when x is near c we have the linear approximation $f(x) \approx \ell(x)$, where

$$\ell(x) = f(c) + f'(c)(x - c)$$

is the linearization of f at c. If we set $dx = \Delta x = x - c$, and evaluate the differential $dy = f'(x)\,dx$ at c, then we have

$$\begin{aligned} \Delta y &= f(x) - f(c) \\ dy &= \ell(x) - f(c). \end{aligned}$$

Given the graph $y = f(x)$, we know that $y = \ell(x)$ gives the tangent line to the graph at c. For the graph $z = f(x, y)$ of a function of two variables, we

similarly have the tangent plane

$$z = f(a,b) + f_x(a,b)(x-a) + f_y(a,b)(y-b)$$

defined in Definition 11.3.10, suggesting that we define the two variable linearization

$$\ell(x,y) = f(a,b) + f_x(a,b)(x-a) + f_y(a,b)(y-b).$$

Consider the total differential dz at (a,b):

$$dz = f_x(a,b)\,dx + f_y(a,b)\,dy.$$

If we assume that (x,y) is "close" to (a,b), and set $dx = x-a$, $dy = y-b$, then we have

$$dz = f_x(a,b)\,dx + f_y(a,b)\,dy = f_x(a,b)(x-a) + f_y(a,b)(y-b).$$

Since $\ell(a,b) = f(a,b)$, we have $\ell(x,y) - \ell(a,b) = dz$, which agrees with the one-variable situation, and reinforces the concept of the differential as the "linear change" in a function.

If we recast Definition 14.1.3 in the language of tangent planes, we can more easily see the analogy with functions of a single variable. We can now say that $f(x,y)$ is differentiable at (a,b) if it has a valid tangent plane approximation at (a,b). Note that $f(x,y) - \ell(x,y)$ is equal to the error term $E_x\,dx + E_y\,dy$.

By Theorem 14.1.7, we know that the tangent plane at $(a,b,f(a,b))$ exists, and gives a good approximation to the graph $z = f(x,y)$, as long as the partial derivatives of f exist and are continuous at (a,b).

14.1.4 Error/Sensitivity Analysis

The total differential gives an approximation of the change in z given small changes in x and y. We can use this to approximate error propagation; that is, if the input is a little off from what it should be, how far from correct will the output be? We demonstrate this in an example.

Example 14.1.11 Sensitivity analysis.

A cylindrical steel storage tank is to be built that is 10ft tall and 4ft across in diameter. It is known that the steel will expand/contract with temperature changes; is the overall volume of the tank more sensitive to changes in the diameter or in the height of the tank?

Solution. A cylindrical solid with height h and radius r has volume $V = \pi r^2 h$. We can view V as a function of two variables, r and h. We can compute partial derivatives of V:

$$\frac{\partial V}{\partial r} = V_r(r,h) = 2\pi r h \qquad \text{and} \qquad \frac{\partial V}{\partial h} = V_h(r,h) = \pi r^2.$$

The total differential is $dV = (2\pi r h)dr + (\pi r^2)dh$. When $h = 10$ and $r = 2$, we have $dV = 40\pi dr + 4\pi dh$. Note that the coefficient of dr is $40\pi \approx 125.7$; the coefficient of dh is a tenth of that, approximately 12.57. A small change in radius will be multiplied by 125.7, whereas a small change in height will be multiplied by 12.57. Thus the volume of the tank is more sensitive to changes in radius than in height.

The previous example showed that the volume of a particular tank was more sensitive to changes in radius than in height. Keep in mind that this analysis only applies to a tank of those dimensions. A tank with a height of 1ft and radius of 5ft would be more sensitive to changes in height than in radius.

One could make a chart of small changes in radius and height and find exact changes in volume given specific changes. While this provides exact numbers, it does not give as much insight as the error analysis using the total differential.

14.1.5 Differentiability of Functions of Three Variables

The definition of differentiability for functions of three variables is very similar to that of functions of two variables. We again start with the total differential.

Definition 14.1.12 Total Differential.

Let $w = f(x, y, z)$ be continuous on a set D. Let dx, dy and dz represent changes in x, y and z, respectively. Where the partial derivatives f_x, f_y and f_z exist, the **total differential** of w is

$$dw = f_x(x, y, z)\,dx + f_y(x, y, z)\,dy + f_z(x, y, z)\,dz.$$

This differential can be a good approximation of the change in w when $w = f(x, y, z)$ is *differentiable*.

Definition 14.1.13 Multivariable Differentiability.

Let $w = f(x, y, z)$ be defined on a set D containing (x_0, y_0, z_0) where $f_x(x_0, y_0, z_0)$, $f_y(x_0, y_0, z_0)$ and $f_z(x_0, y_0, z_0)$ exist. Let dw be the total differential of w at (x_0, y_0, z_0), let $\Delta w = f(x_0 + dx, y_0 + dy, z_0 + dz) - f(x_0, y_0, z_0)$, and let E_x, E_y and E_z be functions of dx, dy and dz such that

$$\Delta w = dw + E_x dx + E_y dy + E_z dz.$$

1. We say f is *differentiable at* (x_0, y_0, z_0) if, given $\varepsilon > 0$, there is a $\delta > 0$ such that if $\|\langle dx, dy, dz\rangle\| < \delta$, then $\|\langle E_x, E_y, E_z\rangle\| < \varepsilon$.
2. We say f is *differentiable on* B if f is differentiable at every point in B. If f is differentiable on $\mathbb{R}^3$, we say that f is *differentiable everywhere*.

Just as before, this definition gives a rigorous statement about what it means to be differentiable that is not very intuitive. We follow it with a theorem similar to Theorem 14.1.8.

Theorem 14.1.14 Continuity and Differentiability of Functions of Three Variables.

Let $w = f(x, y, z)$ be defined on a set D containing (x_0, y_0, z_0).

1. *If f is differentiable at (x_0, y_0, z_0), then f is continuous at (x_0, y_0, z_0).*
2. *If f_x, f_y and f_z are continuous on B, then f is differentiable on B.*

This set of definition and theorem extends to functions of any number of variables. The theorem again gives us a simple way of verifying that most functions that we encounter are differentiable on their natural domains.

This section has given us a formal definition of what it means for a functions to be "differentiable," along with a theorem that gives a more accessible understanding. The following sections return to notions prompted by our study of partial derivatives that make use of the fact that most functions we encounter are differentiable.

14.1.6 Exercises

Terms and Concepts

1. T/F: If $f(x, y)$ is differentiable on S, the f is continuous on S.

2. T/F: If f_x and f_y are continuous on S, then f is differentiable on S.

3. T/F: If $z = f(x, y)$ is differentiable, then the change in z over small changes dx and dy in x and y is approximately dz.

4. Finish the sentence: "The new z-value is approximately the old z-value plus the approximate _____."

Problems

Exercise Group. In the following exercises, find the total differential dz.

5. $z = x\sin(y) + x^2$

6. $z = (2x^2 + 3y)^2$

7. $z = 5x - 7y$

8. $z = xe^{x+y}$

Exercise Group. In the following exercises, a function $f(x, y)$ is given. Give the indicated approximation using the total differential.

9. $f(x, y) = \sqrt{x^2 + y}$. Approximate $f(2.95, 7.1)$ knowing $f(3, 7) = 4$.

10. $f(x, y) = \sin(x)\cos(y)$. Approximate $f(0.1, -0.1)$ knowing $f(0, 0) = 0$.

11. $f(x, y) = x^2y - xy^2$. Approximate $f(2.04, 3.06)$ knowing $f(2, 3) = -6$.

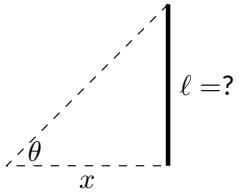
12. $f(x, y) = \ln(x - y)$. Approximate $f(5.1, 3.98)$ knowing $f(5, 4) = 0$.

Exercise Group. The following exercises ask a variety of questions dealing with approximating error and sensitivity analysis.

13. A cylindrical storage tank is to be 2ft tall with a radius of 1ft. Is the volume of the tank more sensitive to changes in the radius or the height?

14. *Projectile Motion:* The x-value of an object moving under the principles of projectile motion is $x(\theta, v_0, t) = (v_0\cos(\theta))t$. A particular projectile is fired with an initial velocity of $v_0 = 250$ft/s and an angle of elevation of $\theta = 60°$. It travels a distance of 375ft in 3 seconds.
Is the projectile more sensitive to errors in initial speed or angle of elevation?

15. The length ℓ of a long wall is to be approximated. The angle θ, as shown in the diagram (not to scale), is measured to be $85°$, and the distance x is measured to be 30'. Assume that the triangle formed is a right triangle.
Is the measurement of the length of ℓ more sensitive to errors in the measurement of x or in θ?



16. It is "common sense" that it is far better to measure a long distance with a long measuring tape rather than a short one. A measured distance D can be viewed as the product of the length ℓ of a measuring tape times the number n of times it was used. For instance, using a 3' tape 10 times gives a length of 30'. To measure the same distance with a 12' tape, we would use the tape 2.5 times. (i.e., $30 = 12 \times 2.5$.) Thus $D = n\ell$.
Suppose each time a measurement is taken with the tape, the recorded distance is within 1/16'' of the actual distance. (i.e., $d\ell = 1/16'' \approx 0.005$ft). Using differentials, show why common sense proves correct in that it is better to use a long tape to measure long distances.

Exercise Group. In the following exercises, find the total differential dw.

17. $w = x^2yz^3$

18. $w = e^x \sin(y) \ln(z)$

Exercise Group. In the following exercises, use the information provided and the total differential to make the given approximation.

19. $f(3,1) = 7,\ f_x(3,1) = 9, f_y(3,1) = -2.$ Approximate $f(3.05, 0.9)$.

20. $f(-4,2) = 13,\ f_x(-4,2) = 2.6,$ $f_y(-4,2) = 5.1$. Approximate $f(-4.12, 2.07)$.

21. $f(2,4,5) = -1,\ f_x(2,4,5) = 2,\ f_y(2,4,5) = -3,\ f_z(2,4,5) = 3.7$. Approximate $f(2.5, 4.1, 4.8)$.

22. $f(3,3,3) = 5,\ f_x(3,3,3) = 2,\ f_y(3,3,3) = 0,\ f_z(3,3,3) = -2$. Approximate $f(3.1, 3.1, 3.1)$.

14.2 The Multivariable Chain Rule

Consider driving an off-road vehicle along a dirt road. As you drive, your elevation likely changes. What factors determine how quickly your elevation rises and falls? After some thought, generally one recognizes that one's *velocity* (speed and direction) and the *terrain* influence your rise and fall.

One can represent the terrain as the surface defined by a multivariable function $f(x, y)$; one can represent the path of the off-road vehicle, as seen from above, with a vector-valued function $\vec{r}(t) = \langle x(t), y(t)\rangle$; the velocity of the vehicle is thus $\vec{r}\,'(t) = \langle x'(t), y'(t)\rangle$.

Consider Figure 14.2.1 in which a surface $z = f(x, y)$ is drawn, along with a dashed curve in the xy-plane. Restricting f to just the points on this circle gives the curve shown on the surface (i.e., "the path of the off-road vehicle.") The derivative $\frac{df}{dt}$ gives the instantaneous rate of change of f with respect to t. If we consider an object traveling along this path, $\frac{df}{dt} = \frac{dz}{dt}$ gives the rate at which the object rises/falls (i.e., "the rate of elevation change" of the vehicle.) Conceptually, the Multivariable Chain Rule combines terrain and velocity information properly to compute this rate of elevation change.

Abstractly, let z be a function of x and y; that is, $z = f(x, y)$ for some function f, and let x and y each be functions of t. By choosing a t-value, x- and y-values are determined, which in turn determine z: this defines z as a function of t. The Multivariable Chain Rule gives a method of computing $\frac{dz}{dt}$.

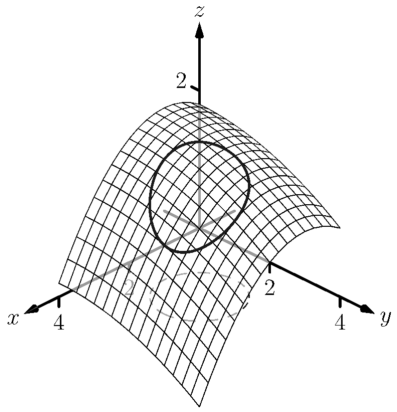


Figure 14.2.1 Understanding the application of the Multivariable Chain Rule

14.2.1 Multivariable Chain Rule, Part I

Theorem 14.2.3 Multivariable Chain Rule, Part I.

Let $z = f(x, y)$, $x = g(t)$ and $y = h(t)$, where f, g and h are differentiable functions. Then $z = f(x, y) = f\big(g(t), h(t)\big)$ is a function of t, and

$$\begin{aligned}\frac{dz}{dt} = \frac{df}{dt} &= f_x(x, y)\frac{dx}{dt} + f_y(x, y)\frac{dy}{dt}\\ &= \frac{\partial f}{\partial x}\frac{dx}{dt} + \frac{\partial f}{\partial y}\frac{dy}{dt}\\ &= \langle\, f_x, f_y\rangle \cdot \langle x', y'\rangle.\end{aligned}$$

The Chain Rule of Section 2.5 states that

$$\frac{d}{dx}\Big(f\big(g(x)\big)\Big) = f'\big(g(x)\big)g'(x).$$

If $t = g(x)$, we can express the Chain Rule as

$$\frac{df}{dx} = \frac{df}{dt}\frac{dt}{dx};$$

recall that the derivative notation is deliberately chosen to reflect their fraction-like properties. A similar effect is seen in Theorem 14.2.3. In the second line of equations, one can think of the dx and ∂x as "sort of" canceling out, and likewise with dy and ∂y.

Notice, too, the third line of equations in Theorem 14.2.3. The vector $\langle\, f_x, f_y\rangle$ contains information about the surface (terrain); the vector $\langle x', y'\rangle$ can represent velocity. In the context measuring the rate of elevation change of the off-road vehicle, the Multivariable Chain Rule states it can be found through a product of terrain and velocity information.

We now practice applying the Multivariable Chain Rule.

youtu.be/watch?v=e4uf2cA7YyA

Figure 14.2.2 Introducing the chain rule in several variables

youtu.be/watch?v=bagsQndc8nw

Figure 14.2.4 Examples involving the chain rule

Example 14.2.5 Using the Multivariable Chain Rule.

Let $z = x^2y + x$, where $x = \sin(t)$ and $y = e^{5t}$. Find $\frac{dz}{dt}$ using the Chain Rule.

Solution. Following Theorem 14.2.3, we find

$$f_x(x,y) = 2xy+1, \qquad f_y(x,y) = x^2, \qquad \frac{dx}{dt} = \cos(t), \qquad \frac{dy}{dt} = 5e^{5t}.$$

Applying the theorem, we have

$$\frac{dz}{dt} = (2xy+1)\cos(t) + 5x^2e^{5t}.$$

This may look odd, as it seems that $\frac{dz}{dt}$ is a function of x, y and t. Since x and y are functions of t, $\frac{dz}{dt}$ is really just a function of t, and we can replace x with $\sin(t)$ and y with e^{5t}:

$$\frac{dz}{dt} = (2xy+1)\cos(t)+5x^2e^{5t} = (2\sin(t)e^{5t}+1)\cos(t)+5e^{5t}\sin^2(t).$$

The previous example can make us wonder: if we substituted for x and y at the end to show that $\frac{dz}{dt}$ is really just a function of t, why not substitute before differentiating, showing clearly that z is a function of t?

That is, $z = x^2y+x = (\sin(t))^2e^{5t}+\sin(t)$. Applying the Chain and Product Rules, we have

$$\frac{dz}{dt} = 2\sin(t)\cos(t)\,e^{5t} + 5\sin^2(t)\,e^{5t} + \cos(t),$$

which matches the result from the example.

This may now make one wonder "What's the point? If we could already find the derivative, why learn another way of finding it?" In some cases, applying this rule makes deriving simpler, but this is hardly the power of the Chain Rule. Rather, in the case where $z = f(x,y)$, $x = g(t)$ and $y = h(t)$, the Chain Rule is extremely powerful when *we do not know what f, g and/or h are*. It may be hard to believe, but often in "the real world" we know rate-of-change information (i.e., information about derivatives) without explicitly knowing the underlying functions. The Chain Rule allows us to combine several rates of change to find another rate of change. The Chain Rule also has theoretic use, giving us insight into the behavior of certain constructions (as we'll see in the next section).

We demonstrate this in the next example.

Example 14.2.6 Applying the Multivariable Chain Rule.

An object travels along a path on a surface. The exact path and surface are not known, but at time $t = t_0$ it is known that :

$$\frac{\partial z}{\partial x} = 5, \qquad \frac{\partial z}{\partial y} = -2, \qquad \frac{dx}{dt} = 3 \qquad \text{and} \qquad \frac{dy}{dt} = 7.$$

Find $\frac{dz}{dt}$ at time t_0.

Solution. The Multivariable Chain Rule states that

$$\begin{aligned}\frac{dz}{dt} &= \frac{\partial z}{\partial x}\frac{dx}{dt} + \frac{\partial z}{\partial y}\frac{dy}{dt}\\ &= 5(3) + (-2)(7)\end{aligned}$$

$$= 1.$$

By knowing certain rates-of-change information about the surface and about the path of the particle in the xy-plane, we can determine how quickly the object is rising/falling.

We next apply the Chain Rule to solve a max/min problem.

Example 14.2.7 Applying the Multivariable Chain Rule.

Consider the surface $z = x^2 + y^2 - xy$, a paraboloid, on which a particle moves with x and y coordinates given by $x = \cos(t)$ and $y = \sin(t)$. Find $\frac{dz}{dt}$ when $t = 0$, and find where the particle reaches its maximum/minimum z-values.

Solution. It is straightforward to compute

$$f_x(x,y) = 2x - y \qquad f_y(x,y) = 2y - x$$
$$\frac{dx}{dt} = -\sin(t) \qquad \frac{dy}{dt} = \cos(t).$$

Combining these according to the Chain Rule gives:

$$\frac{dz}{dt} = -(2x - y)\sin(t) + (2y - x)\cos(t).$$

When $t = 0$, $x = 1$ and $y = 0$. Thus $\dfrac{dz}{dt} = -(2)(0) + (-1)(1) = -1$. When $t = 0$, the particle is moving down, as shown in Figure 14.2.8.

To find where z-value is maximized/minimized on the particle's path, we set $\frac{dz}{dt} = 0$ and solve for t:

$$\begin{aligned}
\frac{dz}{dt} = 0 &= -(2x - y)\sin(t) + (2y - x)\cos(t) \\
0 &= -(2\cos(t) - \sin(t))\sin(t) + (2\sin(t) - \cos(t))\cos(t) \\
0 &= \sin^2(t) - \cos^2(t) \\
\cos^2(t) &= \sin^2(t) \\
t &= n\frac{\pi}{4} \text{ (for odd } n)
\end{aligned}$$

We can use the First Derivative Test to find that on $[0, 2\pi]$, z has reaches its absolute minimum at $t = \pi/4$ and $5\pi/4$; it reaches its absolute maximum at $t = 3\pi/4$ and $7\pi/4$, as shown in Figure 14.2.8.

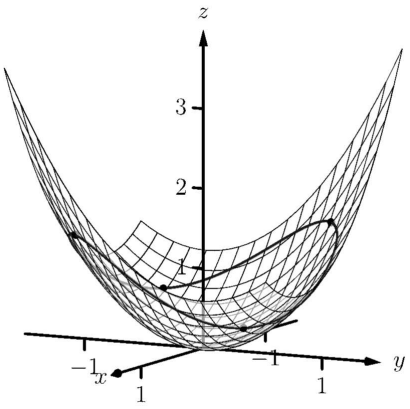


Figure 14.2.8 Plotting the path of a particle on a surface in Example 14.2.7

We can extend the Chain Rule to include the situation where z is a function of more than one variable, and each of these variables is also a function of more than one variable. The basic case of this is where $z = f(x, y)$, and x and y are functions of two variables, say s and t.

Theorem 14.2.9 Multivariable Chain Rule, Part II.

1. *Let $z = f(x, y)$, $x = g(s, t)$ and $y = h(s, t)$, where f, g and h are differentiable functions. Then z is a function of s and t, and*
   - $\dfrac{\partial z}{\partial s} = \dfrac{\partial f}{\partial x}\dfrac{\partial x}{\partial s} + \dfrac{\partial f}{\partial y}\dfrac{\partial y}{\partial s}$

- $\dfrac{\partial z}{\partial t} = \dfrac{\partial f}{\partial x}\dfrac{\partial x}{\partial t} + \dfrac{\partial f}{\partial y}\dfrac{\partial y}{\partial t}$

2. *Let $z = f(x_1, x_2, \ldots, x_m)$ be a differentiable function of m variables, where each of the x_i is a differentiable function of the variables $t_1, t_2, \ldots, t_n$. Then z is a function of the t_i, and*

$$\frac{\partial z}{\partial t_i} = \frac{\partial f}{\partial x_1}\frac{\partial x_1}{\partial t_i} + \frac{\partial f}{\partial x_2}\frac{\partial x_2}{\partial t_i} + \cdots + \frac{\partial f}{\partial x_m}\frac{\partial x_m}{\partial t_i}.$$

Example 14.2.10 Using the Multivariable Chain Rule, Part II.

Let $z = x^2y + x$, $x = s^2 + 3t$ and $y = 2s - t$. Find $\frac{\partial z}{\partial s}$ and $\frac{\partial z}{\partial t}$, and evaluate each when $s = 1$ and $t = 2$.

Solution. Following Theorem 14.2.9, we compute the following partial derivatives:

$$\frac{\partial f}{\partial x} = 2xy + 1 \qquad \frac{\partial f}{\partial y} = x^2,$$

$$\frac{\partial x}{\partial s} = 2s \qquad \frac{\partial x}{\partial t} = 3 \qquad \frac{\partial y}{\partial s} = 2 \qquad \frac{\partial y}{\partial t} = -1.$$

Thus

$$\frac{\partial z}{\partial s} = (2xy + 1)(2s) + (x^2)(2) = 4xys + 2s + 2x^2, \text{ and}$$

$$\frac{\partial z}{\partial t} = (2xy + 1)(3) + (x^2)(-1) = 6xy - x^2 + 3.$$

When $s = 1$ and $t = 2$, $x = 7$ and $y = 0$, so

$$\frac{\partial z}{\partial s} = 100 \qquad \text{and} \qquad \frac{\partial z}{\partial t} = -46.$$

Example 14.2.11 Using the Multivariable Chain Rule, Part II.

Let $w = xy + z^2$, where $x = t^2e^s$, $y = t\cos(s)$, and $z = s\sin(t)$. Find $\frac{\partial w}{\partial t}$ when $s = 0$ and $t = \pi$.

Solution. Following Theorem 14.2.9, we compute the following partial derivatives:

$$\frac{\partial f}{\partial x} = y \qquad \frac{\partial f}{\partial y} = x \qquad \frac{\partial f}{\partial z} = 2z$$

$$\frac{\partial x}{\partial t} = 2te^s \qquad \frac{\partial y}{\partial t} = \cos(s) \qquad \frac{\partial z}{\partial t} = s\cos(t).$$

Thus

$$\frac{\partial w}{\partial t} = y(2te^s) + x(\cos(s)) + 2z(s\cos(t)).$$

When $s = 0$ and $t = \pi$, we have $x = \pi^2$, $y = \pi$ and $z = 0$. Thus

$$\frac{\partial w}{\partial t} = \pi(2\pi) + \pi^2 = 3\pi^2.$$

14.2.2 Implicit Differentiation

We studied finding $\frac{dy}{dx}$ when y is given as an implicit function of x in detail in Section 2.6. We find here that the Multivariable Chain Rule gives a simpler method of finding $\frac{dy}{dx}$.

For instance, consider the implicit function $x^2y - xy^3 = 3$. We learned to use the following steps to find $\frac{dy}{dx}$:

$$\begin{aligned} \frac{d}{dx}\Big(x^2y - xy^3\Big) &= \frac{d}{dx}\Big(3\Big) \\ 2xy + x^2\frac{dy}{dx} - y^3 - 3xy^2\frac{dy}{dx} &= 0 \\ \frac{dy}{dx} &= -\frac{2xy - y^3}{x^2 - 3xy^2}. \end{aligned}$$

Instead of using this method, consider $z = x^2y - xy^3$. The implicit function above describes the level curve $z = 3$. Considering x and y as functions of x, the Multivariable Chain Rule states that

$$\frac{dz}{dx} = \frac{\partial z}{\partial x}\frac{dx}{dx} + \frac{\partial z}{\partial y}\frac{dy}{dx}. \tag{14.2.1}$$

Since z is constant (in our example, $z = 3$), $\frac{dz}{dx} = 0$. We also know $\frac{dx}{dx} = 1$. Equation (14.2.1) becomes

$$\begin{aligned} 0 &= \frac{\partial z}{\partial x}(1) + \frac{\partial z}{\partial y}\frac{dy}{dx} \Rightarrow \\ \frac{dy}{dx} &= -\frac{\partial z}{\partial x}\Big/\frac{\partial z}{\partial y} \\ &= -\frac{f_x}{f_y}. \end{aligned} \tag{14.2.2}$$

Note how our solution for $\frac{dy}{dx}$ in Equation (14.2.2) is just the partial derivative of z with respect to x, divided by the partial derivative of z with respect to y, all multiplied by (-1).

We state the above as a theorem.

Theorem 14.2.12 Implicit Differentiation.

Let f be a differentiable function of x and y, where $f(x, y) = c$ defines y as an implicit function of x, for some constant c. Then

$$\frac{dy}{dx} = -\frac{f_x(x, y)}{f_y(x, y)}.$$

We practice using Theorem 14.2.12 by applying it to a problem from Section 2.6.

Example 14.2.13 Implicit Differentiation.

Given the implicitly defined function $\sin(x^2y^2) + y^3 = x + y$, find y'. Note: this is the same problem as given in Example 2.6.8 of Section 2.6, where the solution took about a full page to find.

Solution. Let $f(x, y) = \sin(x^2y^2) + y^3 - x - y$; the implicitly defined function above is equivalent to $f(x, y) = 0$. We find $\frac{dy}{dx}$ by applying

Theorem 14.2.12. We find

$$f_x(x,y) = 2xy^2\cos(x^2y^2) - 1$$
$$f_y(x,y) = 2x^2y\cos(x^2y^2) + 3y^2 - 1,$$

so

$$\frac{dy}{dx} = -\frac{2xy^2\cos(x^2y^2) - 1}{2x^2y\cos(x^2y^2) + 3y^2 - 1},$$

which matches our solution from Example 2.6.8.

youtu.be/watch?v=o9Zyg-QY8Qo

Figure 14.2.14 Examples with implicit differentiation

We can also do implicit differentiation for functions of three variables. In the same way that a level curve $f(x,y) = c$ is used to implicitly define y as a function of x, a level surface $f(x,y,z) = c$ can be viewed as implicitly defining z as a function of x and y.

Suppose the equation $f(x,y,z) = c$, where c is a constant, defines the function $z = g(x,y)$. Then we can use the chain rule to compute the derivatives of $f(x,y,z)$ with respect to x and y, where we set $x = x$, $y = y$, and $z = g(x,y)$. Since $f(x,y,z)$ is constant, we have

$$\begin{aligned} 0 &= \frac{\partial}{\partial x} f(x,y,z) \\ &= f_x(x,y,z)\frac{\partial x}{\partial x} + f_y(x,y,z)\frac{\partial y}{\partial x} + f_z(x,y,z)\frac{\partial z}{\partial x} \\ &= f_x(x,y,z)(1) + f_y(x,y,z)(0) + f_z(x,y,z)\frac{\partial z}{\partial x}. \end{aligned}$$

There is a subtlety in this calculation that can be quite confusing. It appears to be chain rule, but shouldn't the derivative of $f(x,y,z)$ with respect to x be simply $f_x(x,y,z)$? The catch here is that we are considering the variables x, y, z (viewed as coordinates in $\mathbb{R}^3$) as functions of x and y, viewed as coordinates in $\mathbb{R}^2$.

In other words, the x and y in $g(x,y)$ are not the same as the ones in $f(x,y,z)$! If this is still confusing, try setting $x = s, y = t$, and $z = g(s,t)$, and then applying the chain rule as usual, for derivatives of $f(x,y,z)$ with respect to s and t. At the end, we can relabel s and t as x and y.

Solving for $\frac{\partial z}{\partial x}$ gives us

$$\frac{\partial z}{\partial x} = -\frac{f_x(x,y,z)}{f_z(x,y,z)},$$

and similarly,

$$\frac{\partial z}{\partial y} = -\frac{f_y(x,y,z)}{f_z(x,y,z)}.$$

In Subsection 11.3.2 we saw that we can use partial derivatives to determine the equation of the tangent plane to a graph $z = f(x,y)$. Using implicit differentiation, we can do the same for a level surface $f(x,y,z) = c$.

Example 14.2.15 Implicit Differentiation with three variables.

Given that the equation

$$x^2yz^3 - \sin(x - 3z) + 4xy^2 - 3yz = 0 \qquad (14.2.3)$$

defines z implicitly as a function of x and y, compute $\frac{\partial z}{\partial x}$ and $\frac{\partial z}{\partial y}$ using implicit differentiation. Then, determine the equation of the tangent plane to the surface at the point $(3,0,1)$.

Solution. There are two ways to proceed. One is to use implicit differentiation as before, but using partial derivatives. Whenever we differentiate a function of z, we multiply by the appropriate partial derivative of z. The other option is to use the formula derived above. We will use the first method for the x derivative, and the second for y.

We first take the partial derivative of both sides of (14.2.3) with respect to x:

$$\frac{\partial}{\partial x}(x^2yz^3 - \sin(x - 3z) + 4xy^2 - 3yz) = 0$$

$$2xyz^3 + x^2y(3z^2)\frac{\partial z}{\partial x} - \cos(x-3z)\left(1 - 3\frac{\partial z}{\partial x}\right) + 4y^2 - 3y\frac{\partial z}{\partial x} = 0.$$

Note that we treated y as a constant, since the derivative is with respect to x. Next, we collect terms:

$$\frac{\partial z}{\partial x}\left(3x^2yz^2 + 3\cos(x-3z) - 3y\right) = -2xyz^3 + \cos(x-3z) - 4y^2.$$

Lastly, we solve for $\frac{\partial z}{\partial x}$:

$$\frac{\partial z}{\partial x} = \frac{-2xyz^3 + \cos(x-3z) - 4y^2}{3x^2yz^2 + 3\cos(x-3z) - 3y}.$$

For the y derivative, we will use the result given above. Setting $f(x,y,z) = x^2yz^3 - \sin(x-3z) + 4xy^2 - 3yz$, we have $\frac{\partial z}{\partial y} = -\frac{f_y(x,y,z)}{f_z(x,y,z)}$. Therefore,

$$\frac{\partial z}{\partial y} = -\frac{x^2z^3 + 8xy - 3z}{3x^2yz^2 + 3\cos(x-3z) - 3y}.$$

The second method certainly seems simpler! The reader is invited to try each part with the other method, and compare answers.

Finally, we consider the problem of the tangent plane. First, we check that the point $(3,0,1)$ is indeed on the surface: $f(3,0,1) = 0$, as required. Next we note that $z = 1$ is given to us from this point. So if $f(x,y,z) = c$ implicitly defines the graph $z = g(x,y)$, then we must have $g(3,0) = 1$. Next, we have

$$g_x(3,0) = \left.\frac{\partial z}{\partial x}\right|_{(3,0)} = \frac{0+1-0}{0+3-0} = \frac{1}{3}$$
$$g_y(3,0) = \left.\frac{\partial z}{\partial y}\right|_{(3,0)} = -\frac{9+0-3}{0+3-0} = -2.$$

The equation of the tangent plane is therefore

$$z = g(3,0) + g_x(3,0)(x-3) + g_y(3,0)(y-0) = 1 + \frac{1}{3}(x-3) - 2y.$$

In Section 11.3 we learned how partial derivatives give certain instantaneous rate of change information about a function $f(x,y)$. In that section, we measured the rate of change of f by holding one variable constant and letting the other vary (such as, holding y constant and letting x vary gives f_x). We can visualize this change by considering the surface defined by f at a point and moving parallel to the x-axis.

What if we want to move in a direction that is not parallel to a coordinate axis? Can we still measure instantaneous rates of change? Yes; we find out how in Section 14.3. In doing so, we'll see how the Multivariable Chain Rule informs our understanding of these *directional derivatives*.

14.2.3 Exercises

Terms and Concepts

1. Let a level curve of $z = f(x, y)$ be described by $x = g(t)$, $y = h(t)$. Explain why $\frac{dz}{dt} = 0$.

2. Fill in the blank: The single variable Chain Rule states $\dfrac{d}{dx}\Big(f\big(g(x)\big)\Big) = f'\big(g(x)\big) \cdot$ _____.

3. Fill in the blank: The Multivariable Chain Rule states
$\dfrac{df}{dt} = \dfrac{\partial f}{\partial x} \cdot$ _____ $+$ _____ $\cdot \dfrac{dy}{dt}$.

4. If $z = f(x, y)$, where $x = g(t)$ and $y = h(t)$, we can substitute and write z as an explicit function of t.
T/F: Using the Multivariable Chain Rule to find $\frac{dz}{dt}$ is sometimes easier than first substituting and then taking the derivative.

5. T/F: The Multivariable Chain Rule is only useful when all the related functions are known explicitly.

6. The Multivariable Chain Rule allows us to compute implicit derivatives easily by just computing two _____ derivatives.

Problems

Exercise Group. Given the functions $z = f(x, y)$, $x = g(t)$ and $y = h(t)$:

(a) Use the Multivariable Chain Rule to compute $\frac{dz}{dt}$.

(b) Evaluate $\frac{dz}{dt}$ at the indicated t-value.

7. $z = 3x + 4y$, $x = t^2$, $y = 2t$; $t = 1$

8. $z = x^2 - y^2$, $x = t$, and $y = t^2 - 1$; $t = 1$

9. $z = 5x + 2y$, $x = 2\cos(t) + 1$, $y = \sin(t) - 3$; $t = \pi/4$

10. $z = \frac{x}{y^2+1}$, $x = \cos(t)$, and $y = \sin(t)$; $t = \pi/2$

11. $z = x^2 + 2y^2$, $x = \sin(t)$, $y = 3\sin(t)$; $t = \pi/4$

12. $z = \cos(x)\sin(y)$, $x = \pi t$, and $y = 2\pi t + \pi/2$; $t = 3$

Exercise Group. In the following exercises, functions $z = f(x, y)$, $x = g(t)$ and $y = h(t)$ are given. Find the values of t where $\frac{dz}{dt} = 0$. Note: these are the same surfaces/curves as found in Exercises 14.2.7–14.2.12.

13. $z = 3x + 4y$, $x = t^2$, $y = 2t$

14. Given $z = x^2 - y^2$, $x = t$, and $y = t^2 - 1$, at what values of t does $\frac{dz}{dt} = 0$?

15. $z = 5x + 2y$, $x = 2\cos(t) + 1$, $y = \sin(t) - 3$

16. Given $z = \frac{x}{y^2+1}$, $x = \cos(t)$, and $y = \sin(t)$, at what values of t in $[0, 2\pi)$ does $\frac{dz}{dt} = 0$?

17. $z = x^2 + 2y^2$, $x = \sin(t)$, $y = 3\sin(t)$

18. Given $z = \cos(x)\sin(y)$, $x = \pi t$, and $y = 2\pi t + \pi/2$, at what values of t in $[0, 2)$ does $\frac{dz}{dt} = 0$?

Exercise Group. Given the functions $z = f(x, y)$, $x = g(s, t)$ and $y = h(s, t)$:

(a) Use the Multivariable Chain Rule to compute $\frac{\partial z}{\partial s}$ and $\frac{\partial z}{\partial t}$.

(b) Evaluate $\frac{\partial z}{\partial s}$ and $\frac{\partial z}{\partial t}$ at the indicated s and t values.

19. $z = x^2 y$, $x = s - t$, $y = 2s + 4t$; $s = 1, t = 0$

20. $z = \cos\big(\pi x + \frac{\pi}{2} y\big)$, $x = st^2$, and $y = s^2 t$; $s = 1, t = 0$

21. $z = x^2 + y^2$, $x = s\cos(t)$, and $y = s\sin(t)$; $s = 2, t = \pi/4$

22. $z = e^{-(x^2+y^2)}$, $x = t$, and $y = st^2$, $s = 1, t = 1$

Exercise Group. The given equation defines y implicitly as a function of x. Find $\frac{dy}{dx}$ using Implicit Differentiation and Theorem 14.2.12.

23. $x^2 \tan(y) = 50$

24. $\left(3x^2 + 2y^3\right)^4 = 2$

25. $\dfrac{x^2 + y}{x + y^2} = 17$

26. $\ln\left(x^2 + xy + y^2\right) = 1$

Exercise Group. Find $\frac{dz}{dt}$, or $\frac{\partial z}{\partial s}$ and $\frac{\partial z}{\partial t}$, using the supplied information.

27. $\dfrac{\partial z}{\partial x} = 2, \dfrac{\partial z}{\partial y} = 1, \dfrac{dx}{dt} = 4, \dfrac{dy}{dt} = -5$

28. $\frac{\partial z}{\partial x} = 1$, $\frac{\partial z}{\partial y} = -3$, $\frac{dx}{dt} = 6$, and $\frac{dy}{dt} = 2$.

29. $\dfrac{\partial z}{\partial x} = -4, \dfrac{\partial z}{\partial y} = 9,$
$\dfrac{\partial x}{\partial s} = 5, \dfrac{\partial x}{\partial t} = 7, \dfrac{\partial y}{\partial s} = -2, \dfrac{\partial y}{\partial t} = 6$

30. $\frac{\partial z}{\partial x} = 2$, $\frac{\partial z}{\partial y} = 1$, $\frac{\partial x}{\partial s} = -2$, $\frac{\partial x}{\partial t} = 3$, $\frac{\partial y}{\partial s} = 2$ and $\frac{\partial y}{\partial t} = -1$

14.3 Directional Derivatives

Partial derivatives give us an understanding of how a surface changes when we move in the x and y directions. We made the comparison to standing in a rolling meadow and heading due east: the amount of rise/fall in doing so is comparable to f_x. Likewise, the rise/fall in moving due north is comparable to f_y. The steeper the slope, the greater in magnitude f_y.

But what if we didn't move due north or east? What if we needed to move northeast and wanted to measure the amount of rise/fall? Partial derivatives alone cannot measure this. This section investigates *directional derivatives*, which do measure this rate of change.

youtu.be/watch?v=ZQPnmL5CM6s

Figure 14.3.1 Introducing directional derivatives

14.3.1 Functions of Two Variables

We begin with a definition.

Definition 14.3.2 Directional Derivatives.

Let $z = f(x, y)$ be continuous on a set S and let $\vec{u} = \langle u_1, u_2 \rangle$ be a unit vector. For all points (x, y), the **directional derivative** of f at (x, y) in the direction of $\vec{u}$ is

$$D_{\vec{u}} f(x, y) = \lim_{h \to 0} \frac{f(x + hu_1, y + hu_2) - f(x, y)}{h}.$$

The partial derivatives f_x and f_y are defined with similar limits, but only x or y varies with h, not both. Here both x and y vary with a weighted h, determined by a particular unit vector $\vec{u}$. This may look a bit intimidating but in reality it is not too difficult to deal with; it often just requires extra algebra. However, the following theorem reduces this algebraic load.

Theorem 14.3.3 Directional Derivatives.

Let $z = f(x, y)$ be differentiable on a set S containing (x_0, y_0), and let $\vec{u} = \langle u_1, u_2 \rangle$ be a unit vector. The directional derivative of f at (x_0, y_0) in the direction of $\vec{u}$ is

$$D_{\vec{u}} f(x_0, y_0) = f_x(x_0, y_0)u_1 + f_y(x_0, y_0)u_2.$$

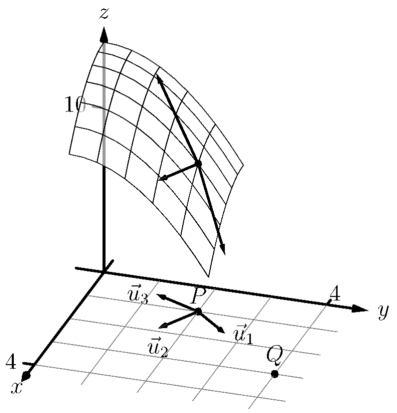


Figure 14.3.5 Understanding the directional derivative in Example 14.3.4

Example 14.3.4 Computing directional derivatives.

Let $z = 14 - x^2 - y^2$ and let $P = (1, 2)$. Find the directional derivative of f, at P, in the following directions:

1. toward the point $Q = (3, 4)$,
2. in the direction of $\langle 2, -1 \rangle$, and
3. toward the origin.

Solution. The surface is plotted in Figure 14.3.5, where the point $P = (1, 2)$ is indicated in the x, y-plane as well as the point $(1, 2, 9)$ which lies on the surface of f. We find that $f_x(x, y) = -2x$ and $f_x(1, 2) = -2$; $f_y(x, y) = -2y$ and $f_y(1, 2) = -4$.

1. Let $\vec{u}_1$ be the unit vector that points from the point $(1, 2)$ to the point $Q = (3, 4)$, as shown in the figure. The vector $\overrightarrow{PQ} = \langle 2, 2 \rangle$;

the unit vector in this direction is $\vec{u}_1 = \langle 1/\sqrt{2}, 1/\sqrt{2}\rangle$. Thus the directional derivative of f at $(1,2)$ in the direction of $\vec{u}_1$ is

$$D_{\vec{u}_1} f(1,2) = -2(1/\sqrt{2}) + (-4)(1/\sqrt{2}) = -6/\sqrt{2} \approx -4.24.$$

Thus the instantaneous rate of change in moving from the point $(1,2,9)$ on the surface in the direction of $\vec{u}_1$ (which points toward the point Q) is about -4.24. Moving in this direction moves one steeply downward.

2. We seek the directional derivative in the direction of $\langle 2,-1\rangle$. The unit vector in this direction is $\vec{u}_2 = \langle 2/\sqrt{5}, -1/\sqrt{5}\rangle$. Thus the directional derivative of f at $(1,2)$ in the direction of $\vec{u}_2$ is

$$D_{\vec{u}_2} f(1,2) = -2(2/\sqrt{5}) + (-4)(-1/\sqrt{5}) = 0.$$

Starting on the surface of f at $(1,2)$ and moving in the direction of $\langle 2,-1\rangle$ (or $\vec{u}_2$) results in no instantaneous change in z-value. This is analogous to standing on the side of a hill and choosing a direction to walk that does not change the elevation. One neither walks up nor down, rather just "along the side" of the hill. Finding these directions of "no elevation change" is important.

3. At $P = (1,2)$, the direction towards the origin is given by the vector $\langle -1,-2\rangle$; the unit vector in this direction is $\vec{u}_3 = \langle -1/\sqrt{5}, -2/\sqrt{5}\rangle$. The directional derivative of f at P in the direction of the origin is

$$D_{\vec{u}_3} f(1,2) = -2(-1/\sqrt{5}) + (-4)(-2/\sqrt{5}) = 10/\sqrt{5} \approx 4.47.$$

Moving towards the origin means "walking uphill" quite steeply, with an initial slope of about 4.47.

As we study directional derivatives, it will help to make an important connection between the unit vector $\vec{u} = \langle u_1, u_2\rangle$ that describes the direction and the partial derivatives f_x and f_y. We start with a definition and follow this with a Key Idea.

youtu.be/watch?v=aoR9ANpc334

Figure 14.3.6 Computing directional derivatives

Definition 14.3.7 Gradient.

Let $z = f(x,y)$ be differentiable on a set S that contains the point (x_0, y_0).

1. The *gradient of* f is $\nabla f(x,y) = \langle f_x(x,y), f_y(x,y)\rangle$.
2. The *gradient of* f *at* (x_0,y_0) *is* $\nabla f(x_0,y_0) = \langle f_x(x_0,y_0), f_y(x_0,y_0)\rangle$.

The symbol "∇" is named "nabla," derived from the Greek name of a Jewish harp. Oddly enough, in mathematics the expression ∇f is pronounced "del f."

To simplify notation, we often express the gradient as $\nabla f = \langle f_x, f_y\rangle$. The gradient allows us to compute directional derivatives in terms of a dot product.

Key Idea 14.3.8 The Gradient and Directional Derivatives.

The directional derivative of $z = f(x,y)$ in the direction of $\vec{u}$ is

$$D_{\vec{u}} f = \nabla f \cdot \vec{u}.$$

The properties of the dot product previously studied allow us to investigate the properties of the directional derivative. Given that the directional derivative

gives the instantaneous rate of change of z when moving in the direction of $\vec{u}$, three questions naturally arise:

1. In what direction(s) is the change in z the greatest (i.e., the "steepest uphill")?
2. In what direction(s) is the change in z the least (i.e., the "steepest downhill")?
3. In what direction(s) is there no change in z?

Using the key property of the dot product, we have

$$\nabla f \cdot \vec{u} = \|\nabla f\| \, \|\vec{u}\| \cos(\theta) = \|\nabla f\| \cos(\theta), \tag{14.3.1}$$

where θ is the angle between the gradient and $\vec{u}$. (Since $\vec{u}$ is a unit vector, $\|\vec{u}\| = 1$.) This equation allows us to answer the three questions stated previously.

1. Equation (14.3.1) is maximized when $\cos(\theta) = 1$, i.e., when the gradient and $\vec{u}$ have the same direction. We conclude the gradient points in the direction of greatest z change.
2. Equation (14.3.1) is minimized when $\cos(\theta) = -1$, i.e., when the gradient and $\vec{u}$ have opposite directions. We conclude the gradient points in the opposite direction of the least z change.
3. Equation (14.3.1) is 0 when $\cos(\theta) = 0$, i.e., when the gradient and $\vec{u}$ are orthogonal to each other. We conclude the gradient is orthogonal to directions of no z change.

This result is rather amazing. Once again imagine standing in a rolling meadow and face the direction that leads you steepest uphill. Then the direction that leads steepest downhill is directly behind you, and side-stepping either left or right (i.e., moving perpendicularly to the direction you face) does not change your elevation at all.

Recall that a level curve is defined as a curve in the xy-plane along which the z-values of a function do not change. Let a surface $z = f(x, y)$ be given, and let's represent one such level curve as a vector-valued function, $\vec{r}(t) = \langle x(t), y(t) \rangle$. As the output of f does not change along this curve, $f\big(x(t), y(t)\big) = c$ for all t, for some constant c.

Since f is constant for all t, $\frac{df}{dt} = 0$. By the Multivariable Chain Rule, we also know

$$\begin{aligned} \frac{df}{dt} &= f_x(x, y)x'(t) + f_y(x, y)y'(t) \\ &= \langle f_x(x, y), f_y(x, y) \rangle \cdot \langle x'(t), y'(t) \rangle \\ &= \nabla f \cdot \vec{r}\,'(t) \\ &= 0. \end{aligned}$$

This last equality states $\nabla f \cdot \vec{r}\,'(t) = 0$: the gradient is orthogonal to the derivative of $\vec{r}$, meaning the gradient is orthogonal to the graph of $\vec{r}$. Our conclusion: at any point on a surface, the gradient at that point is orthogonal to the level curve that passes through that point.

We restate these ideas in a theorem, then use them in an example.

Theorem 14.3.9 The Gradient and Directional Derivatives.

Let $z = f(x, y)$ be differentiable on a set S with gradient ∇f, let $P = (x_0, y_0)$ be a point in S and let $\vec{u}$ be a unit vector.

1. *The maximum value of $D_{\vec{u}} f(x_0, y_0)$ is $\|\nabla f(x_0, y_0)\|$; the direction of maximal z increase is $\nabla f(x_0, y_0)$.*
2. *The minimum value of $D_{\vec{u}} f(x_0, y_0)$ is $-\|\nabla f(x_0, y_0)\|$; the direction of minimal z increase is $-\nabla f(x_0, y_0)$.*
3. *At P, $\nabla f(x_0, y_0)$ is orthogonal to the level curve passing through (x_0, y_0).*

Example 14.3.10 Finding directions of maximal and minimal increase.

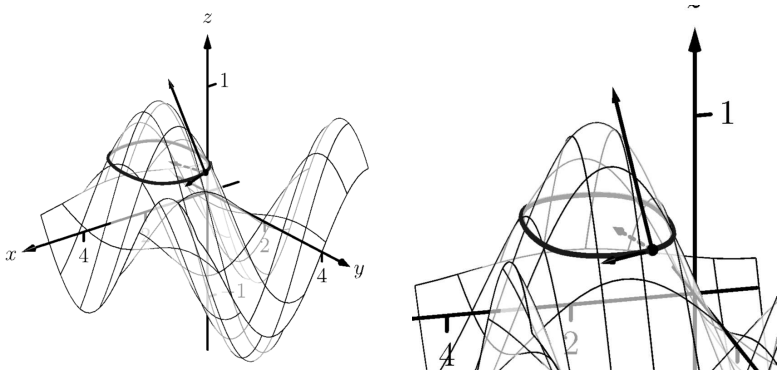
Let $f(x, y) = \sin(x)\cos(y)$ and let $P = (\pi/3, \pi/3)$. Find the directions of maximal/minimal increase, and find a direction where the instantaneous rate of z change is 0.

Solution. We begin by finding the gradient. $f_x = \cos(x)\cos(y)$ and $f_y = -\sin(x)\sin(y)$, thus

$$\nabla f = \langle \cos(x)\cos(y), -\sin(x)\sin(y)\rangle \text{ and, at } P, \nabla f\left(\frac{\pi}{3}, \frac{\pi}{3}\right) = \left\langle \frac{1}{4}, -\frac{3}{4}\right\rangle.$$

Thus the direction of maximal increase is $\langle 1/4, -3/4\rangle$. In this direction, the instantaneous rate of z change is $\|\langle 1/4, -3/4\rangle\| = \sqrt{10}/4 \approx 0.79$.

Figure 14.3.11 shows the surface plotted from two different perspectives. In each, the gradient is drawn at P with a dashed line (because of the nature of this surface, the gradient points "into" the surface). Let $\vec{u} = \langle u_1, u_2\rangle$ be the unit vector in the direction of ∇f at P. Each graph of the figure also contains the vector $\langle u_1, u_2, \|\nabla f\|\rangle$. This vector has a "run" of 1 (because in the xy-plane it moves 1 unit) and a "rise" of $\|\nabla f\|$, hence we can think of it as a vector with slope of $\|\nabla f\|$ in the direction of ∇f, helping us visualize how "steep" the surface is in its steepest direction.



(a) **(b)**

Figure 14.3.11 Graphing the surface and important directions in Example 14.3.10

The direction of minimal increase is $\langle -1/4, 3/4\rangle$; in this direction the instantaneous rate of z change is $-\sqrt{10}/4 \approx -0.79$.

Any direction orthogonal to ∇f is a direction of no z change. We have two choices: the direction of $\langle 3, 1\rangle$ and the direction of $\langle -3, -1\rangle$. The unit vector in the direction of $\langle 3, 1\rangle$ is shown in each graph of the figure as well. The level curve at $z = \sqrt{3}/4$ is drawn: recall that along this curve the z-values do not change. Since $\langle 3, 1\rangle$ is a direction of no z-change, this vector is tangent to the level curve at P.

Example 14.3.12 Understanding when $\nabla f = \vec{0}$.

Let $f(x, y) = -x^2 + 2x - y^2 + 2y + 1$. Find the directional derivative of f in any direction at $P = (1, 1)$.

Solution. We find $\nabla f = \langle -2x + 2, -2y + 2\rangle$. At P, we have $\nabla f(1, 1) = \langle 0, 0\rangle$. According to Theorem 14.3.9, this is the direction of maximal increase. However, $\langle 0, 0\rangle$ is directionless; it has no displacement. And regardless of the unit vector $\vec{u}$ chosen, $D_{\vec{u}} f = 0$.

Figure 14.3.13 helps us understand what this means. We can see that P lies at the top of a paraboloid. In all directions, the instantaneous rate of change is 0.

So what is the direction of maximal increase? It is fine to give an answer of $\vec{0} = \langle 0, 0\rangle$, as this indicates that all directional derivatives are 0.

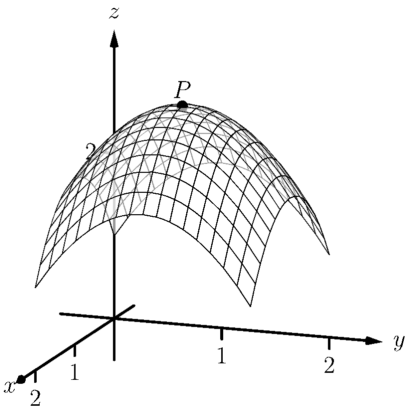


Figure 14.3.13 At the top of a paraboloid, all directional derivatives are 0

The fact that the gradient of a surface always points in the direction of steepest increase/decrease is very useful, as illustrated in the following example.

Example 14.3.14 The flow of water downhill.

Consider the surface given by the graph of $f(x, y) = 20 - x^2 - 2y^2$. Water is poured on the surface at $(1, 1/4)$. What path does it take as it flows downhill?

Solution. Let $\vec{r}(t) = \langle x(t), y(t)\rangle$ be the vector-valued function describing the path of the water in the xy-plane; we seek $x(t)$ and $y(t)$. We know that water will always flow downhill in the steepest direction; therefore, at any point on its path, it will be moving in the direction of $-\nabla f$. (We ignore the physical effects of momentum on the water.) Thus $\vec{r}'(t)$ will be parallel to ∇f, and there is some constant c such that $c\nabla f = \vec{r}'(t) = \langle x'(t), y'(t)\rangle$.

We find $\nabla f = \langle -2x, -4y\rangle$ and write $x'(t)$ as $\frac{dx}{dt}$ and $y'(t)$ as $\frac{dy}{dt}$. Then

$$c\nabla f = \langle x'(t), y'(t)\rangle$$
$$\langle -2cx, -4cy\rangle = \left\langle \frac{dx}{dt}, \frac{dy}{dt}\right\rangle.$$

This implies

$$-2cx = \frac{dx}{dt} \text{ and } -4cy = \frac{dy}{dt}, \text{ i.e.,}$$
$$c = -\frac{1}{2x}\frac{dx}{dt} \text{ and } c = -\frac{1}{4y}\frac{dy}{dt}.$$

As c equals both expressions, we have

$$\frac{1}{2x}\frac{dx}{dt} = \frac{1}{4y}\frac{dy}{dt}.$$

To find an explicit relationship between x and y, we can integrate both sides with respect to t. Recall from our study of differentials that $\frac{dx}{dt}dt = dx$. Thus:

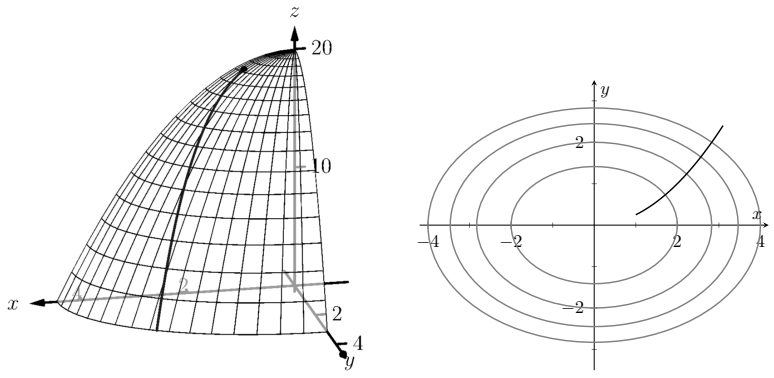
$$\begin{aligned}
\int \frac{1}{2x}\frac{dx}{dt}\,dt &= \int \frac{1}{4y}\frac{dy}{dt}\,dt \\
\int \frac{1}{2x}\,dx &= \int \frac{1}{4y}\,dy \\
\frac{1}{2}\ln|x| &= \frac{1}{4}\ln|y| + C_1 \\
2\ln|x| &= \ln|y| + C_1 \\
\ln\left|x^2\right| &= \ln|y| + C_1
\end{aligned}$$

Now raise both sides as a power of e:

$$\begin{aligned}
x^2 &= e^{\ln|y|+C_1} \\
x^2 &= e^{\ln|y|}e^{C_1} \qquad \text{(Note that } e^{C_1} \text{ is just a constant.)} \\
x^2 &= yC_2 \\
\frac{1}{C_2}x^2 &= y \qquad \text{(Note that } 1/C_2 \text{ is just a constant.)} \\
Cx^2 &= y.
\end{aligned}$$

As the water started at the point $(1, 1/4)$, we can solve for C:

$$C(1)^2 = \frac{1}{4} \Rightarrow C = \frac{1}{4}.$$



(a) **(b)**

Figure 14.3.15 A sketch of the surface described in Example 14.3.14 along with the path in the xy-plane with the level curves

Thus the water follows the curve $y = x^2/4$ in the xy-plane. The surface and the path of the water is graphed in Figure 14.3.15(a). In Figure 14.3.15(b), the level curves of the surface are plotted in the xy-plane, along with the curve $y = x^2/4$. Notice how the path intersects the level curves at right angles. As the path follows the gradient downhill, this reinforces the fact that the gradient is orthogonal to level curves.

youtu.be/watch?v=Re0amVb6QdQ

Figure 14.3.16 Discussing the significance of directional derivatives

14.3.2 Functions of Three Variables

The concepts of directional derivatives and the gradient are easily extended to three (and more) variables. We combine the concepts behind Definitions 14.3.2 and Definition 14.3.7 and Theorem 14.3.3 into one set of definitions.

Definition 14.3.17 Directional Derivatives and Gradient with Three Variables.

Let $w = F(x, y, z)$ be differentiable on a set D and let $\vec{u}$ be a unit vector in $\mathbb{R}^3$.

1. The *gradient* of F is $\nabla F = \langle F_x, F_y, F_z \rangle$.
2. The *directional derivative of F in the direction of $\vec{u}$* is

$$D_{\vec{u}} F = \nabla F \cdot \vec{u}.$$

The same properties of the gradient given in Theorem 14.3.9, when f is a function of two variables, hold for F, a function of three variables.

Theorem 14.3.18 The Gradient and Directional Derivatives with Three Variables.

Let $w = F(x, y, z)$ be differentiable on a set D, let ∇F be the gradient of F, and let $\vec{u}$ be a unit vector.

1. *The maximum value of $D_{\vec{u}} F$ is $\|\nabla F\|$, obtained when the angle between ∇F and $\vec{u}$ is 0, i.e., the direction of maximal increase is ∇F.*
2. *The minimum value of $D_{\vec{u}} F$ is $-\|\nabla F\|$, obtained when the angle between ∇F and $\vec{u}$ is π, i.e., the direction of minimal increase is $-\nabla F$.*
3. *$D_{\vec{u}} F = 0$ when ∇F and $\vec{u}$ are orthogonal.*

We interpret the third statement of the theorem as "the gradient is orthogonal to level surfaces," the three-variable analogue to level curves.

Example 14.3.19 Finding directional derivatives with functions of three variables.

If a point source S is radiating energy, the intensity I at a given point P in space is inversely proportional to the square of the distance between S and P. That is, when $S = (0, 0, 0)$, $I(x, y, z) = \frac{k}{x^2+y^2+z^2}$ for some constant k.

Let $k = 1$, let $\vec{u} = \langle 2/3, 2/3, 1/3 \rangle$ be a unit vector, and let $P = (2, 5, 3)$. Measure distances in inches. Find the directional derivative of I at P in the direction of $\vec{u}$, and find the direction of greatest intensity increase at P.

Solution. We need the gradient ∇I, meaning we need I_x, I_y and I_z. Each partial derivative requires a simple application of the Quotient Rule,

giving

$$\begin{aligned}
\nabla I &= \left\langle \frac{-2x}{(x^2+y^2+z^2)^2}, \frac{-2y}{(x^2+y^2+z^2)^2}, \frac{-2z}{(x^2+y^2+z^2)^2} \right\rangle \\
\nabla I(2,5,3) &= \left\langle \frac{-4}{1444}, \frac{-10}{1444}, \frac{-6}{1444} \right\rangle \approx \langle -0.003, -0.007, -0.004 \rangle \\
D_{\vec{u}}\, I &= \nabla I(2,5,3) \cdot \vec{u} \\
&= -\frac{17}{2166} \approx -0.0078.
\end{aligned}$$

The directional derivative tells us that moving in the direction of $\vec{u}$ from P results in a decrease in intensity of about -0.008 units per inch. (The intensity is decreasing as $\vec{u}$ moves one farther from the origin than P.)

The gradient gives the direction of greatest intensity increase. Notice that

$$\begin{aligned}
\nabla I(2,5,3) &= \left\langle \frac{-4}{1444}, \frac{-10}{1444}, \frac{-6}{1444} \right\rangle \\
&= \frac{2}{1444} \langle -2, -5, -3 \rangle .
\end{aligned}$$

That is, the gradient at $(2,5,3)$ is pointing in the direction of $\langle -2, -5, -3 \rangle$, that is, towards the origin. That should make intuitive sense: the greatest increase in intensity is found by moving towards to source of the energy.

The directional derivative allows us to find the instantaneous rate of z change in any direction at a point. We can use these instantaneous rates of change to define lines and planes that are *tangent* to a surface at a point, which is the topic of the next section.

14.3.3 Exercises

Terms and Concepts

1. What is the difference between a directional derivative and a partial derivative?

2. For $f(x,y)$, for what $\vec{u}$ is $D_{\vec{u}}\, f = f_x$?

3. For $f(x,y)$, for what $\vec{u}$ is $D_{\vec{u}}\, f = f_y$?

4. The gradient is _____ to level curves.

5. The gradient points in the direction of _____ increase.

6. It is generally more informative to view the directional derivative not as the result of a limit, but rather as the result of a _____ product.

Problems

Exercise Group. In the following exercises, a function $f(x,y)$ is given. Find ∇f.

7. $f(x,y) = -x^2y + xy^2 + xy$

8. Find ∇f, where $f(x,y) = \sin(x)\cos(y)$.

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

10. Find ∇f, where $f(x,y) = -4x + 3y$.

11. $f(x,y) = x^2 + 2y^2 - xy - 7x$

12. Find ∇f, where $f(x,y) = x^2y^3 - 2x$.

Exercise Group. In the following exercises, a function $f(x,y)$ and a point P are given. Find the directional derivative of f in the indicated directions. Note: these are the same functions as in Exercises 14.3.7–14.3.12.

13. $f(x,y) = -x^2y + xy^2 + xy$, $P = (2,1)$

(a) In the direction of $\vec{v} = \langle 3,4\rangle$

(b) In the direction toward the point $Q = (1,-1)$.

14. Consider $f(x,y) = \sin(x)\cos(y)$, at $P = \left(\frac{\pi}{4}, \frac{\pi}{3}\right)$.

(a) In the direction of $\vec{v} = \langle 1,1\rangle$.

(b) In the direction toward the point $Q = (0,0)$.

15. $f(x,y) = \dfrac{1}{x^2+y^2+1}$, $P = (1,1)$.

(a) In the direction of $\vec{v} = \langle 1,-1\rangle$.

(b) In the direction toward the point $Q = (-2,-2)$.

16. Consider $f(x,y) = -4x + 3y$, at $P = (5,2)$.

(a) In the direction of $\vec{v} = \langle 3,1\rangle$.

(b) In the direction toward the point $Q = (2,7)$.

17. $f(x,y) = x^2 + 2y^2 - xy - 7x$, $P = (4,1)$

(a) In the direction of $\vec{v} = \langle -2,5\rangle$

(b) In the direction toward the point $Q = (4,0)$.

18. Consider $f(x,y) = x^2y^3 - 2x$, at $P = (1,1)$.

(a) Find the directional derivative in the direction of $\vec{v} = \langle 3,3\rangle$.

(b) Find the directional derivative in the direction toward the point $Q = (1,2)$.

Exercise Group. In the following exercises, a function $f(x,y)$ and a point P are given. Investigate the directions of maximal increase and decrease, as indicated.

Note: these are the same functions and points as in Exercises 14.3.13 through Exercise 14.3.18.

19. $f(x,y) = -x^2y + xy^2 + xy$, $P = (2,1)$

(a) Find the direction of maximal increase of f at P.

(b) What is the maximal value of $D_{\vec{u}}\, f$ at P?

(c) Find the direction of maximal decrease in f at P.

(d) Give a direction $\vec{u}$ such that $D_{\vec{u}}\, f = 0$ at P.

20. $f(x,y) = \sin(x)\cos(y)$, $P = \left(\frac{\pi}{4}, \frac{\pi}{3}\right)$:

(a) Find the direction of maximal increase of f at P.

(b) What is the maximal value of $D_{\vec{u}}\, f$ at P?

(c) Find the direction of maximal decrease in f at P.

(d) Give a direction $\vec{u}$ such that $D_{\vec{u}}\, f = 0$ at P.

21. $f(x,y) = \dfrac{1}{x^2+y^2+1}$, $P = (1,1)$.

(a) Find the direction of maximal increase of f at P.

(b) What is the maximal value of $D_{\vec{u}}\, f$ at P?

(c) Find the direction of maximal decrease in f at P.

(d) Give a direction $\vec{u}$ such that $D_{\vec{u}}\, f = 0$ at P.

22. $f(x,y) = -4x + 3y$, $P = (5,4)$

(a) Find the direction of maximal increase of f at P.

(b) What is the maximal value of $D_{\vec{u}}\, f$ at P?

(c) Find the direction of maximal decrease in f at P.

(d) Give a direction $\vec{u}$ such that $D_{\vec{u}}\, f = 0$ at P.

23. $f(x,y) = x^2 + 2y^2 - xy - 7x$, $P = (4,1)$

(a) Find the direction of maximal increase of f at P.

(b) What is the maximal value of $D_{\vec{u}}\, f$ at P?

(c) Find the direction of maximal decrease in f at P.

(d) Give a direction $\vec{u}$ such that $D_{\vec{u}}\, f = 0$ at P.

24. Given $f(x,y) = x^2y^3 - 2x$, $P = (1,1)$:

(a) Find the direction of maximal increase of f at P.

(b) What is the maximal value of $D_{\vec{u}}\, f$ at P?

(c) Find the direction of maximal decrease in f at P.

(d) Give a direction $\vec{u}$ such that $D_{\vec{u}}\, f = 0$ at P.

Exercise Group. In the following exercises, a function $F(x,y,z)$, a vector $\vec{v}$ and a point P are given. Compute the gradient of F, and the derivative of F in the direction of $\vec{v}$ at P.

25. $F(x,y,z) = 3x^2z^3 + 4xy - 3z^2$, $\vec{v} = \langle 1,1,1\rangle$, $P = (3,2,1)$

(a) Compute the gradient of F.

(b) Find the derivative of F at P in the direction of $\vec{v}$.

26. $F(x,y,z) = \sin(x)\cos(y)e^z$, $\vec{v} = \langle 2,2,1\rangle$, $P = (0,0,0)$.

(a) Find $\nabla F(x,y,z)$.

(b) Find $D_{\vec{u}}\, F$ at P.

27. $F(x,y,z) = x^2y^2 - y^2z^2$, $\vec{v} = \langle -1,7,3\rangle$, $P = (1,0,-1)$

(a) Compute the gradient of F.

(b) Find the derivative of F at P in the direction of $\vec{v}$.

28. Given $F(x,y,z) = \frac{2}{x^2+y^2+z^2}$, $\vec{v} = \langle 1,1,-2\rangle$, $P = (1,1,1)$:

(a) Find $\nabla F(x,y,z)$.

(b) Find $D_{\vec{u}}\, F$ at P.

14.4 Tangent Lines, Normal Lines, and Tangent Planes

14.4.1 Tangent Lines

Derivatives and tangent lines go hand-in-hand. Given $y = f(x)$, the line tangent to the graph of f at $x = x_0$ is the line through $\big(x_0, f(x_0)\big)$ with slope $f'(x_0)$; that is, the slope of the tangent line is the instantaneous rate of change of f at x_0.

When dealing with functions of two variables, the graph is no longer a curve but a surface. At a given point on the surface, it seems there are many lines that fit our intuition of being "tangent" to the surface.

In Section 11.3.2 we introduced the concept of the tangent plane, which could be thought of as consisting of all possible lines tangent to the surface at a given point. In this section, we explore this idea in more detail.

In Figure 14.4.1 we see lines that are tangent to curves in space. Since each curve lies on a surface, it makes sense to say that the lines are also tangent to the surface. The next definition formally defines what it means to be "tangent to a surface."

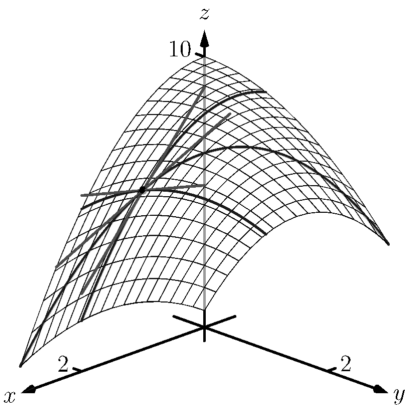


Figure 14.4.1 Showing various lines tangent to a surface

Definition 14.4.2 Directional Tangent Line.

Let $z = f(x, y)$ be differentiable on a set S containing (x_0, y_0) and let $\vec{u} = \langle u_1, u_2 \rangle$ be a unit vector.

1. The line ℓ_x through $\big(x_0, y_0, f(x_0, y_0)\big)$ parallel to $\langle 1, 0, f_x(x_0, y_0)\rangle$ is the **tangent line to f in the direction of x at** (x_0, y_0).
2. The line ℓ_y through $\big(x_0, y_0, f(x_0, y_0)\big)$ parallel to $\langle 0, 1, f_y(x_0, y_0)\rangle$ is the **tangent line to f in the direction of y at** (x_0, y_0).
3. The line $\ell_{\vec{u}}$ through $\big(x_0, y_0, f(x_0, y_0)\big)$ parallel to $\langle u_1, u_2, D_{\vec{u}} f(x_0, y_0)\rangle$ is the **tangent line to f in the direction of $\vec{u}$ at** (x_0, y_0).

It is instructive to consider each of three directions given in the definition in terms of "slope." The direction of ℓ_x is $\langle 1, 0, f_x(x_0, y_0)\rangle$; that is, the "run" is one unit in the x-direction and the "rise" is $f_x(x_0, y_0)$ units in the z-direction. Note how the slope is just the partial derivative with respect to x. A similar statement can be made for ℓ_y. The direction of $\ell_{\vec{u}}$ is $\langle u_1, u_2, D_{\vec{u}} f(x_0, y_0)\rangle$; the "run" is one unit in the $\vec{u}$ direction (where $\vec{u}$ is a unit vector) and the "rise" is the directional derivative of z in that direction.

Definition 14.4.2 leads to the following parametric equations of directional tangent lines:

$$\ell_x(t) = \begin{cases} x = x_0 + t \\ y = y_0 \\ z = z_0 + f_x(x_0, y_0)t \end{cases}$$

$$\ell_y(t) = \begin{cases} x = x_0 \\ y = y_0 + t \\ z = z_0 + f_y(x_0, y_0)t \end{cases}$$

$$\ell_{\vec{u}}(t) = \begin{cases} x = x_0 + u_1 t \\ y = y_0 + u_2 t \\ z = z_0 + D_{\vec{u}} f(x_0, y_0)t \end{cases}.$$

Example 14.4.3 Finding directional tangent lines.

Find the lines tangent to the surface $z = \sin(x)\cos(y)$ at $(\pi/2, \pi/2)$ in the x and y directions and also in the direction of $\vec{v} = \langle -1, 1\rangle$.

Solution. The partial derivatives with respect to x and y are:

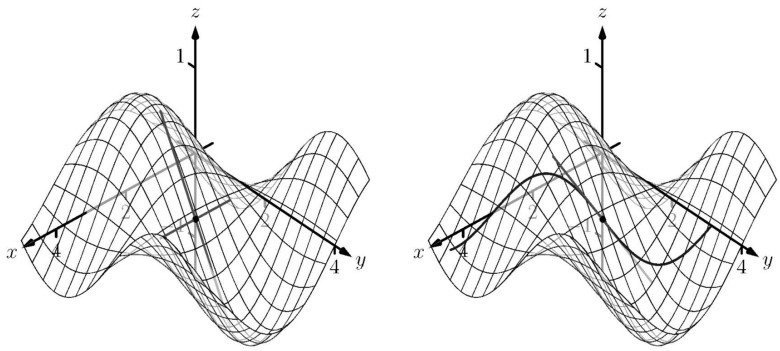
$$f_x(x,y) = \cos(x)\cos(y) \Rightarrow f_x(\pi/2, \pi/2) = 0$$
$$f_y(x,y) = -\sin(x)\sin(y) \Rightarrow f_y(\pi/2, \pi/2) = -1.$$

At $(\pi/2, \pi/2)$, the z-value is 0.

Thus the parametric equations of the line tangent to f at $(\pi/2, \pi/2)$ in the directions of x and y are:

$$\ell_x(t) = \begin{cases} x = \pi/2 + t \\ y = \pi/2 \\ z = 0 \end{cases} \quad \text{and } \ell_y(t) = \begin{cases} x = \pi/2 \\ y = \pi/2 + t \\ z = -t \end{cases}.$$

The two lines are shown with the surface in Figure 14.4.4(a).



(a) **(b)**

Figure 14.4.4 A surface and directional tangent lines in Example 14.4.3

To find the equation of the tangent line in the direction of $\vec{v}$, we first find the unit vector in the direction of $\vec{v}$: $\vec{u} = \left\langle -1/\sqrt{2}, 1/\sqrt{2}\right\rangle$. The directional derivative at $(\pi/2, \pi, 2)$ in the direction of $\vec{u}$ is

$$D_{\vec{u}}\, f(\pi/2, \pi, 2) = \langle 0, -1\rangle \cdot \left\langle -1/\sqrt{2}, 1/\sqrt{2}\right\rangle = -1/\sqrt{2}.$$

Thus the directional tangent line is

$$\ell_{\vec{u}}(t) = \begin{cases} x = \pi/2 - t/\sqrt{2} \\ y = \pi/2 + t/\sqrt{2} \\ z = -t/\sqrt{2} \end{cases}.$$

The curve through $(\pi/2, \pi/2, 0)$ in the direction of $\vec{v}$ is shown in Figure 14.4.4(b) along with $\ell_{\vec{u}}(t)$.

Example 14.4.5 Finding directional tangent lines.

Let $f(x,y) = 4xy - x^4 - y^4$. Find the equations of *all* directional tangent lines to f at $(1,1)$.

Solution. First note that $f(1,1) = 2$. We need to compute directional derivatives, so we need ∇f. We begin by computing partial derivatives.

$$f_x = 4y - 4x^3 \Rightarrow f_x(1,1) = 0; f_y = 4x - 4y^3 \Rightarrow f_y(1,1) = 0.$$

Thus $\nabla f(1,1) = \langle 0,0\rangle$. Let $\vec{u} = \langle u_1, u_2\rangle$ be any unit vector. The directional derivative of f at $(1,1)$ will be $D_{\vec{u}} f(1,1) = \langle 0,0\rangle \cdot \langle u_1, u_2\rangle = 0$. It does not matter what direction we choose; the directional derivative is always 0. Therefore

$$\ell_{\vec{u}}(t) = \begin{cases} x = 1 + u_1 t \\ y = 1 + u_2 t \\ z = 2 \end{cases}.$$

Figure 14.4.6 shows a graph of f and the point $(1,1,2)$. Note that this point comes at the top of a "hill," and therefore every tangent line through this point will have a "slope" of 0.

That is, consider any curve on the surface that goes through this point. Each curve will have a relative maximum at this point, hence its tangent line will have a slope of 0. The following section investigates the points on surfaces where all tangent lines have a slope of 0.

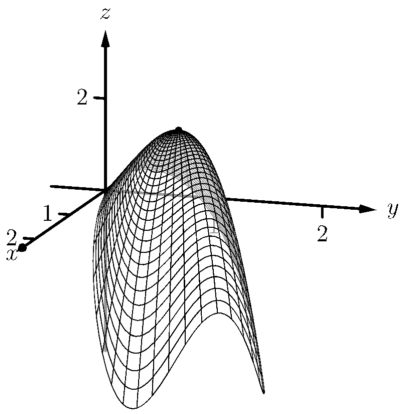


Figure 14.4.6 Graphing f in Example 14.4.5

14.4.2 Normal Lines

When dealing with a function $y = f(x)$ of one variable, we stated that a line through $(c, f(c))$ was *tangent* to f if the line had a slope of $f'(c)$ and was *normal* (or, *perpendicular, orthogonal*) to f if it had a slope of $-1/f'(c)$. We extend the concept of normal, or orthogonal, to functions of two variables.

Let $z = f(x,y)$ be a differentiable function of two variables. By Definition 14.4.2, at (x_0, y_0), $\ell_x(t)$ is a line parallel to the vector $\vec{d}_x = \langle 1, 0, f_x(x_0,y_0)\rangle$ and $\ell_y(t)$ is a line parallel to $\vec{d}_y = \langle 0, 1, f_y(x_0,y_0)\rangle$. Since lines in these directions through $\big(x_0, y_0, f(x_0,y_0)\big)$ are *tangent* to the surface, a line through this point and orthogonal to these directions would be *orthogonal*, or *normal*, to the surface. We can use this direction to create a normal line.

The direction of the normal line is orthogonal to $\vec{d}_x$ and $\vec{d}_y$, hence the direction is parallel to $\vec{d}_n = \vec{d}_x \times \vec{d}_y$. It turns out this cross product has a very simple form:

$$\vec{d}_x \times \vec{d}_y = \langle 1, 0, f_x\rangle \times \langle 0, 1, f_y\rangle = \langle -f_x, -f_y, 1\rangle.$$

It is often more convenient to refer to the opposite of this direction, namely $\langle f_x, f_y, -1\rangle$. This leads to a definition.

Definition 14.4.7 Normal Line.

Let $z = f(x,y)$ be differentiable on a set S containing (x_0, y_0) where

$$a = f_x(x_0,y_0) \text{ and } b = f_y(x_0,y_0)$$

are defined.

1. A nonzero vector parallel to $\vec{n} = \langle a, b, -1\rangle$ is **orthogonal** to f at

$P = \big(x_0, y_0, f(x_0, y_0)\big)$.

2. The line ℓ_n through P with direction parallel to $\vec{n}$ is the **normal line** to f at P.

Thus the parametric equations of the normal line to a surface $z = f(x, y)$ at $\big(x_0, y_0, f(x_0, y_0)\big)$ is:

$$\ell_n(t) = \begin{cases} x = x_0 + at \\ y = y_0 + bt \\ z = f(x_0, y_0) - t \end{cases}.$$

Example 14.4.8 Finding a normal line.

Find the equation of the normal line to $z = -x^2 - y^2 + 2$ at $(0, 1)$.

Solution. We find $z_x(x, y) = -2x$ and $z_y(x, y) = -2y$; at $(0, 1)$, we have $z_x = 0$ and $z_y = -2$. We take the direction of the normal line, following Definition 14.4.7, to be $\vec{n} = \langle 0, -2, -1\rangle$. The line with this direction going through the point $(0, 1, 1)$ is

$$\ell_n(t) = \begin{cases} x = 0 \\ y = -2t + 1 \\ z = -t + 1 \end{cases} \quad \text{or } \ell_n(t) = \langle 0, -2, -1\rangle\, t + \langle 0, 1, 1\rangle\,.$$

The surface $z = -x^2 - y^2 + 2$, along with the found normal line, is graphed in Figure 14.4.9.

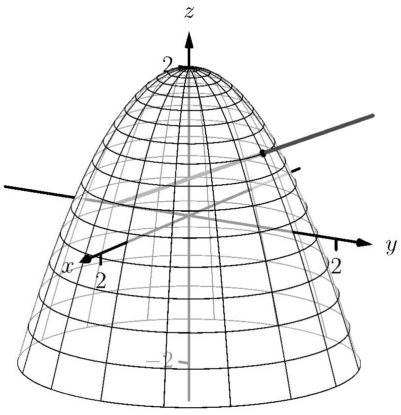


Figure 14.4.9 Graphing a surface with a normal line from Example 14.4.8

The direction of the normal line has many uses, one of which is the definition of the *tangent plane* which we define shortly. Another use is in measuring distances from the surface to a point. Given a point Q in space, it is a general geometric concept to define the distance from Q to the surface as being the length of the shortest line segment $\overline{PQ}$ over all points P on the surface. This, in turn, implies that $\overrightarrow{PQ}$ will be orthogonal to the surface at P. Therefore we can measure the distance from Q to the surface $z = f(x, y)$ by finding a point P on the surface such that $\overrightarrow{PQ}$ is parallel to the normal line to f at P.

Example 14.4.10 Finding the distance from a point to a surface.

Let $f(x, y) = 2 - x^2 - y^2$ and let $Q = (2, 2, 2)$. Find the distance from Q to the surface defined by f.

Solution. This surface is used in Example 14.4.5, so we know that at (x, y), the direction of the normal line will be $\vec{d}_n = \langle -2x, -2y, -1\rangle$. A point P on the surface will have coordinates $(x, y, 2 - x^2 - y^2)$, so $\overrightarrow{PQ} = \big\langle 2 - x, 2 - y, x^2 + y^2\big\rangle$. To find where $\overrightarrow{PQ}$ is parallel to $\vec{d}_n$, we need to find x, y and c such that $c\overrightarrow{PQ} = \vec{d}_n$.

$$c\overrightarrow{PQ} = \vec{d}_n$$
$$c\big\langle 2 - x, 2 - y, x^2 + y^2\big\rangle = \langle -2x, -2y, -1\rangle\,.$$

This implies

$$c(2 - x) = -2x$$
$$c(2 - y) = -2y$$

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

In each equation, we can solve for c:

$$c = \frac{-2x}{2-x} = \frac{-2y}{2-y} = \frac{-1}{x^2+y^2}.$$

The first two fractions imply $x = y$, and so the last fraction can be rewritten as $c = -1/(2x^2)$. Then

$$\begin{aligned} \frac{-2x}{2-x} &= \frac{-1}{2x^2} \\ -2x(2x^2) &= -1(2-x) \\ 4x^3 &= 2-x \\ 4x^3 + x - 2 &= 0. \end{aligned}$$

This last equation is a cubic, which is not difficult to solve with a numeric solver. We find that $x = 0.689$, hence $P = (0.689, 0.689, 1.051)$. We find the distance from Q to the graph of f is

$$\left\|\overrightarrow{PQ}\right\| = \sqrt{(2-0.689)^2 + (2-0.689)^2 + (2-1.051)^2} = 2.083.$$

We can take the concept of measuring the distance from a point to a surface to find a point Q a particular distance from a surface at a given point P on the surface.

Example 14.4.11 Finding a point a set distance from a surface.

Let $f(x,y) = x - y^2 + 3$. Let $P = \big(2, 1, f(2,1)\big) = (2,1,4)$. Find points Q in space that are 4 units from the graph of f at P. That is, find Q such that $\left\|\overrightarrow{PQ}\right\| = 4$ and $\overrightarrow{PQ}$ is orthogonal to f at P.

Solution. We begin by finding partial derivatives:

$$\begin{aligned} f_x(x,y) = 1 \quad &\Rightarrow \quad f_x(2,1) = 1 \\ f_y(x,y) = -2y \quad &\Rightarrow \quad f_y(2,1) = -2 \end{aligned}$$

The vector $\vec{n} = \langle 1, -2, -1\rangle$ is orthogonal to f at P. For reasons that will become more clear in a moment, we find the unit vector in the direction of $\vec{n}$:

$$\vec{u} = \frac{\vec{n}}{\|\vec{n}\|} = \left\langle 1/\sqrt{6}, -2/\sqrt{6}, -1/\sqrt{6}\right\rangle \approx \langle 0.408, -0.816, -0.408\rangle.$$

Thus a the normal line to f at P can be written as

$$\ell_n(t) = \langle 2,1,4\rangle + t\,\langle 0.408, -0.816, -0.408\rangle.$$

An advantage of this parametrization of the line is that letting $t = t_0$ gives a point on the line that is $|t_0|$ units from P. (This is because the direction of the line is given in terms of a unit vector.) There are thus two points in space 4 units from P:

$$Q_1 = \ell_n(4) \qquad\qquad Q_2 = \ell_n(-4)$$

$$\approx \langle 3.63, -2.27, 2.37\rangle \qquad \approx \langle 0.37, 4.27, 5.63\rangle$$

The surface is graphed along with points P, Q_1, Q_2 and a portion of the normal line to f at P.

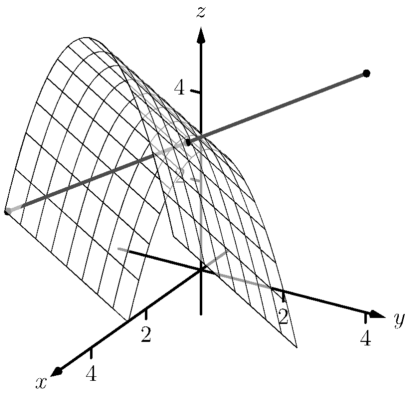


Figure 14.4.12 Graphing the surface in Example 14.4.11 along with points 4 units from the surface

14.4.3 Tangent Planes

We can use the direction of the normal line to define a plane. With $a = f_x(x_0, y_0)$, $b = f_y(x_0, y_0)$ and $P = \big(x_0, y_0, f(x_0, y_0)\big)$, the vector $\vec{n} = \langle a, b, -1\rangle$ is orthogonal to f at P. (See Definition 11.3.10.) The plane through P with normal vector $\vec{n}$ is therefore *tangent* to f at P.

Definition 14.4.13 Tangent Plane.

Let $z = f(x, y)$ be differentiable on a set S containing (x_0, y_0), where $a = f_x(x_0, y_0), b = f_y(x_0, y_0), \vec{n} = \langle a, b, -1\rangle$ and $P = \big(x_0, y_0, f(x_0, y_0)\big)$.

The plane through P with normal vector $\vec{n}$ is the **tangent plane** to f at P. The standard form of this plane is

$$a(x - x_0) + b(y - y_0) - \big(z - f(x_0, y_0)\big) = 0.$$

When we introduced the tangent plane in Section 11.3, we computed the normal vector to be $\vec{n} = \langle -f_x(x_0, y_0), -f_y(x_0, y_0), 1\rangle$. Here, for convenience, we take the negative of this vector, and use $\vec{n} = \langle f_x(x_0, y_0), f_y(x_0, y_0), -1\rangle$.

Example 14.4.14 Finding tangent planes.

Find the equation of the tangent plane to $z = -x^2 - y^2 + 2$ at $(0, 1)$.

Solution. Note that this is the same surface and point used in Example 14.4.8. There we found $\vec{n} = \langle 0, -2, -1\rangle$ and $P = (0, 1, 1)$. Therefore the equation of the tangent plane is

$$-2(y - 1) - (z - 1) = 0.$$

The surface $z = -x^2 - y^2 + 2$ and tangent plane are graphed in Figure 14.4.15.

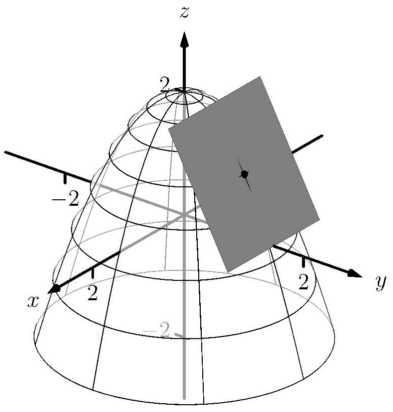


Figure 14.4.15 Graphing a surface with tangent plane from Example 14.4.14

Example 14.4.16 Using the tangent plane to approximate function values.

The point $(3, -1, 4)$ lies on the graph of an unknown differentiable function f where $f_x(3, -1) = 2$ and $f_y(3, -1) = -1/2$. Find the equation of the tangent plane to f at P, and use this to approximate the value of $f(2.9, -0.8)$.

Solution. Knowing the partial derivatives at $(3, -1)$ allows us to form the normal vector to the tangent plane, $\vec{n} = \langle 2, -1/2, -1\rangle$. Thus the equation of the tangent line to f at P is:

$$2(x-3) - 1/2(y+1) - (z-4) = 0 \Rightarrow z = 2(x-3) - 1/2(y+1) + 4. \tag{14.4.1}$$

Just as tangent lines provide excellent approximations of curves near their point of intersection, tangent planes provide excellent approximations of surfaces near their point of intersection. So $f(2.9, -0.8) \approx z(2.9, -0.8) = 3.7$.

This is not a new method of approximation. Compare the right hand

expression for z in Equation (14.4.1) to the total differential:

$$dz = f_x dx + f_y dy \text{ and } z = \underbrace{\underbrace{2}_{f_x}\underbrace{(x-3)}_{dx} + \underbrace{-1/2}_{f_y}\underbrace{(y+1)}_{dy}}_{dz} + 4.$$

Thus the "new z-value" is the sum of the change in z (i.e., dz) and the old z-value (4). As mentioned when studying the total differential, it is not uncommon to know partial derivative information about a unknown function, and tangent planes are used to give accurate approximations of the function.

14.4.4 The Gradient and Normal Lines, Tangent Planes

The methods developed in this section so far give a straightforward method of finding equations of normal lines and tangent planes for surfaces with explicit equations of the form $z = f(x, y)$. However, they do not handle implicit equations well, such as $x^2 + y^2 + z^2 = 1$. There is a technique that allows us to find vectors orthogonal to these surfaces based on the *gradient*.

Definition 14.4.17 Gradient.

Let $w = F(x, y, z)$ be differentiable on a set D that contains the point (x_0, y_0, z_0).

1. The **gradient** of F is $\nabla F(x, y, z) = \langle f_x(x, y, z), f_y(x, y, z), f_z(x, y, z)\rangle$.
2. The **gradient** of F at (x_0, y_0, z_0) is

$$\nabla F(x_0, y_0, z_0) = \langle f_x(x_0, y_0, z_0), f_y(x_0, y_0, z_0), f_z(x_0, y_0, z_0)\rangle .$$

Recall that when $z = f(x, y)$, the gradient $\nabla f = \langle f_x, f_y\rangle$ is orthogonal to level curves of f. An analogous statement can be made about the gradient ∇F, where $w = F(x, y, z)$. Given a point (x_0, y_0, z_0), let $c = F(x_0, y_0, z_0)$. Then $F(x, y, z) = c$ is a **level surface** that contains the point (x_0, y_0, z_0). The following theorem states that $\nabla F(x_0, y_0, z_0)$ is orthogonal to this level surface.

Theorem 14.4.18 The Gradient and Level Surfaces.

Let $w = F(x, y, z)$ be differentiable on a set D containing (x_0, y_0, z_0) with gradient ∇F, where $F(x_0, y_0, z_0) = c$.

The vector $\nabla F(x_0, y_0, z_0)$ is orthogonal to the level surface $F(x, y, z) = c$ at (x_0, y_0, z_0).

The gradient at a point gives a vector orthogonal to the surface at that point. This direction can be used to find tangent planes and normal lines.

Example 14.4.19 Using the gradient to find a tangent plane.

Find the equation of the plane tangent to the ellipsoid $\frac{x^2}{12} + \frac{y^2}{6} + \frac{z^2}{4} = 1$ at $P = (1, 2, 1)$.

Solution. We consider the equation of the ellipsoid as a level surface of a function F of three variables, where $F(x, y, z) = \frac{x^2}{12} + \frac{y^2}{6} + \frac{z^2}{4}$.

The gradient is:

$$\begin{aligned}\nabla F(x,y,z) &= \langle F_x, F_y, F_z\rangle \\ &= \left\langle \frac{x}{6}, \frac{y}{3}, \frac{z}{2}\right\rangle.\end{aligned}$$

At P, the gradient is $\nabla F(1,2,1) = \langle 1/6, 2/3, 1/2\rangle$. Thus the equation of the plane tangent to the ellipsoid at P is

$$\frac{1}{6}(x-1) + \frac{2}{3}(y-2) + \frac{1}{2}(z-1) = 0.$$

The ellipsoid and tangent plane are graphed in Figure 14.4.20.

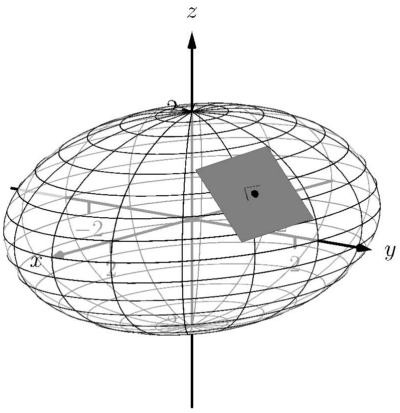


Figure 14.4.20 An ellipsoid and its tangent plane at a point

To understand why Theorem 14.4.18 is true, recall the method of implicit differentiation given in Subsection 14.2.2. A level surface $f(x,y,z) = 0$ can be viewed as defining $z = g(x,y)$ implicitly. We found that the partial derivatives of z with respect to x and y are then given by

$$\frac{\partial z}{\partial x} = g_x(x,y) = -\frac{f_x(x,y,z)}{f_z(x,y,z)} \quad \frac{\partial z}{\partial y} = g_y(x,y) = -\frac{f_y(x,y,z)}{f_z(x,y,z)}.$$

If we plug these values into the tangent plane equation

$$z = g(a,b) + g_x(a,b)(x-a) + g_y(a,b)(y-b)$$

we get, with $c = g(a,b)$,

$$z = c - \frac{f_x(a,b,c)}{f_z(a,b,c)}(x-a) - \frac{f_y(a,b,c)}{f_z(a,b,c)}(y-b).$$

If we move everything to the left-hand side of the equation and multiply through by $f_z(a,b,c)$, we get

$$f_x(a,b,c)(x-a) + f_y(a,b,c)(y-b) + f_z(a,b,c)(z-c) = 0,$$

which is the equation of a plane with normal vector $\nabla f(a,b,c)$.

Example 14.4.21 Finding the tangent plane of a level surface.

Determine the equation of the tangent plane to the level surface $x^2yz^3 - \sin(x-3z) + 4xy^2 - 3yz = 0$ at the point $(3,0,1)$. (Note that this is the same problem as Example 14.2.15.)

Solution. With $f(x,y,z) = x^2yz^3 - \sin(x-3z) + 4xy^2 - 3yz$ we have

$$\begin{aligned} f_x(x,y,z) &= 2xyz^3 - \cos(x-3z) + 4y^2 & f_x(3,0,1) &= -1 \\ f_y(x,y,z) &= x^2z^3 + 8xy - 3z & f_y(3,0,1) &= 6 \\ f_z(x,y,z) &= 3x^2yz^2 + 3\cos(x-3z) - 3y & f_z(3,0,1) &= 3. \end{aligned}$$

The equation of the tangent plane is therefore

$$-1(x-3) + 6y + 3(z-1) = 0.$$

Note that solving for z gives $z = 1 + \frac{1}{3}(x-3) - 2y$, which is the same result as Example 14.2.15.

Tangent lines and planes to surfaces have many uses, including the study of instantaneous rates of changes and making approximations. Normal lines also have many uses. In this section we focused on using them to measure distances from a surface. Another interesting application is in computer graphics, where the effects of light on a surface are determined using normal vectors.

The next section investigates another use of partial derivatives: determining relative extrema. When dealing with functions of the form $y = f(x)$, we found relative extrema by finding x where $f'(x) = 0$. We can *start* finding relative extrema of $z = f(x, y)$ by setting f_x and f_y to 0, but it turns out that there is more to consider.

14.4.5 Exercises

Terms and Concepts

1. Explain how the vector $\vec{v} = \langle 1, 0, 3\rangle$ can be thought of as having a "slope" of 3.

2. Explain how the vector $\vec{v} = \langle 0.6, 0.8, -2\rangle$ can be thought of as having a "slope" of -2.

3. True or False? Let $z = f(x, y)$ be differentiable at P. If $\vec{n}$ is a normal vector to the tangent plane of f at P, then $\vec{n}$ is orthogonal to ℓ_x and ℓ_y at P. (☐ True ☐ False)

4. Explain in your own words why we do not refer to *the* tangent line to a surface at a point, but rather to *directional* tangent lines to a surface at a point.

Problems

Exercise Group. A function $f(x, y)$, a vector $\vec{v}$ and a point P are given. Give the parametric equations of the following directional tangent lines to $z = f(x, y)$ at P:

(a) $\ell_x(t)$

(b) $\ell_y(t)$

(c) $\ell_{\vec{u}}(t)$, where $\vec{u}$ is the unit vector in the direction of $\vec{v}$.

5. $f(x, y) = 2x^2y - 4xy^2$, $\vec{v} = \langle 1, 3\rangle$, $P = (2, 3)$.

6. $f(x, y) = 3\cos(x)\sin(y)$, $\vec{v} = \langle 1, 2\rangle$, $P = (\pi/3, \pi/6)$

7. $f(x, y) = 3x - 5y$, $\vec{v} = \langle 1, 1\rangle$, $P = (4, 2)$.

8. $f(x, y) = x^2 - 2x - y^2 + 4y$, $\vec{v} = \langle 1, 1\rangle$, $P = (1, 2)$

Exercise Group. A function $f(x, y)$ and a point P are given. Find the equation of the normal line to $z = f(x, y)$ at P. Note: these are the same functions as in Exercises 14.4.5 — Exercise 14.4.8.

9. $f(x, y) = 2x^2y - 4xy^2$, $P = (2, 3)$.

10. $f(x, y) = 3\cos(x)\sin(y)$ and $P = (\pi/3, \pi/6)$

11. $f(x, y) = 3x - 5y$, $P = (4, 2)$.

12. $f(x, y) = x^2 - 2x - y^2 + 4y$ and $P = (1, 2)$

Exercise Group. A function $f(x, y)$ and a point P are given. Find the two points that are 2 units from the surface $z = f(x, y)$ at P. Note: these are the same functions as in Exercises 14.4.5–14.4.8.

13. $f(x, y) = 2x^2y - 4xy^2$, $P = (2, 3)$.

14. $f(x, y) = 3\cos(x)\sin(y)$, $P = (\pi/3, \pi/6)$.

15. $f(x, y) = 3x - 5y$, $P = (4, 2)$.

16. $f(x, y) = x^2 - 2x - y^2 + 4y$, $P = (1, 2)$.

Exercise Group. A function $f(x, y)$ and a point P are given. Find an equation of the tangent plane to $z = f(x, y)$ at P. Note: these are the same functions as in Exercises 14.4.5–14.4.8.

17. $f(x, y) = 2x^2y - 4xy^2$, $P = (2, 3)$.

18. $f(x, y) = 3\cos(x)\sin(y)$, $P = (\pi/3, \pi/6)$.

19. $f(x, y) = 3x - 5y$, $P = (4, 2)$.

20. $f(x, y) = x^2 - 2x - y^2 + 4y$, $P = (1, 2)$

Exercise Group. An implicitly defined function of x, y and z is given, along with a point P that lies on the surface. Use the gradient ∇F to:

(a) find the equation of the normal line to the surface at P, and

(b) find the equation of the plane tangent to the surface at P.

21. $\dfrac{x^2}{8} + \dfrac{y^2}{4} + \dfrac{z^2}{16} = 1$, at $P = (1, \sqrt{2}, \sqrt{6})$

22. $z^2 - \frac{x^2}{4} - \frac{y^2}{9} = 0$, at $P = (4, -3, \sqrt{5})$

23. $xy^2 - xz^2 = 0$, at $P = (2, 1, -1)$

24. $\sin(xy) + \cos(yz) = 1$, at $P = (2, \pi/12, 4)$

youtu.be/watch?v=smAvod2dFsw

Figure 14.5.1 Introducing extreme values for functions of several variables

14.5 Extreme Values

14.5.1 Critical Points of Functions of Two Variables

Given a function $f(x, y)$, we are often interested in points where $z = f(x, y)$ takes on the largest or smallest values. For instance, if f represents a cost function, we would likely want to know what (x, y) values minimize the cost. If f represents the ratio of a volume to surface area, we would likely want to know where f is greatest. This leads to the following definition.

Definition 14.5.2 Relative and Absolute Extrema.

Let $z = f(x, y)$ be defined on a set S containing the point $P = (x_0, y_0)$.

1. If $f(x_0, y_0) \geq f(x, y)$ for all (x, y) in S, then f has an **absolute maximum** at P If $f(x_0, y_0) \leq f(x, y)$ for all (x, y) in S, then f has an **absolute minimum** at P.
2. If there is an open disk D containing P such that $f(x_0, y_0) \geq f(x, y)$ for all points (x, y) that are in both D and S, then f has a **relative maximum** at P. If there is an open disk D containing P such that $f(x_0, y_0) \leq f(x, y)$ for all points (x, y) that are in both D and S, then f has a **relative minimum** at P.
3. If f has an absolute maximum or minimum at P, then f has an **absolute extremum** at P. If f has a relative maximum or minimum at P, then f has a **relative extremum** at P.

If f has a relative or absolute maximum at (x_0, y_0), it means every curve on the graph of f through $(x_0, y_0, f(x_0, y_0))$ will also have a relative or absolute maximum at P. Recalling what we learned in Section 3.1, the slopes of the tangent lines to these curves at P must be 0 or undefined. Since directional derivatives are computed using f_x and f_y, we are led to the following definition and theorem.

Definition 14.5.3 Critical Point.

Let $z = f(x, y)$ be continuous on a set S. A **critical point** $P = (x_0, y_0)$ of f is a point in S such that, at P,

- $f_x(x_0, y_0) = 0$ and $f_y(x_0, y_0) = 0$, or
- $f_x(x_0, y_0)$ and/or $f_y(x_0, y_0)$ is undefined.

Theorem 14.5.4 Critical Points and Relative Extrema.

Let $z = f(x, y)$ be defined on an open set S containing $P = (x_0, y_0)$. If f has a relative extrema at P, then P is a critical point of f.

Therefore, to find relative extrema, we find the critical points of f and determine which correspond to relative maxima, relative minima, or neither. The following examples demonstrate this process.

youtu.be/watch?v=smAvod2dFsw

Figure 14.5.5 Finding critical points

Example 14.5.6 Finding critical points and relative extrema.

Let $f(x,y) = x^2 + y^2 - xy - x - 2$. Find the relative extrema of f.

Solution. We start by computing the partial derivatives of f:

$$f_x(x,y) = 2x - y - 1 \qquad \text{and} \qquad f_y(x,y) = 2y - x.$$

Each is never undefined. A critical point occurs when f_x and f_y are simultaneously 0, leading us to solve the following system of linear equations:

$$2x - y - 1 = 0 \qquad \text{and} \qquad -x + 2y = 0.$$

This solution to this system is $x = 2/3$, $y = 1/3$. (Check that at $(2/3, 1/3)$, both f_x and f_y are 0.)

The graph in Figure 14.5.7 shows f along with this critical point. It is clear from the graph that this is a relative minimum; further consideration of the function shows that this is actually the absolute minimum.

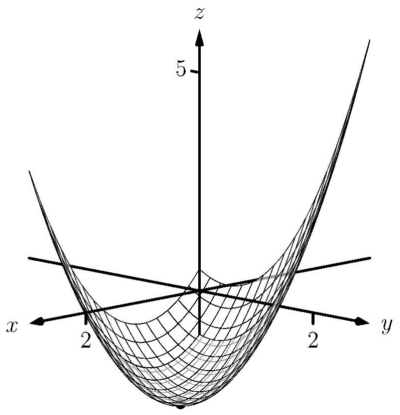


Figure 14.5.7 The surface in Example 14.5.6 with its absolute minimum indicated

Example 14.5.8 Finding critical points and relative extrema.

Let $f(x,y) = -\sqrt{x^2 + y^2} + 2$. Find the relative extrema of f.

Solution. We start by computing the partial derivatives of f:

$$f_x(x,y) = \frac{-x}{\sqrt{x^2+y^2}} \qquad \text{and} \qquad f_y(x,y) = \frac{-y}{\sqrt{x^2+y^2}}.$$

It is clear that $f_x = 0$ when $x = 0$ & $y \neq 0$, and that $f_y = 0$ when $y = 0$ & $x \neq 0$. At $(0,0)$, both f_x and f_y are *not* 0, but rather undefined. The point $(0,0)$ is still a critical point, though, because the partial derivatives are undefined. This is the only critical point of f.

The graph of f is plotted in Figure 14.5.9 along with the point $(0,0,2)$. The graph shows that this point is the absolute maximum of f.

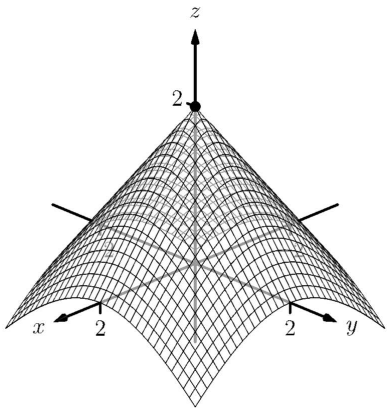


Figure 14.5.9 The surface in Example 14.5.8 with its absolute maximum indicated

In each of the previous two examples, we found a critical point of f and then determined whether or not it was a relative (or absolute) maximum or minimum by graphing. It would be nice to be able to determine whether a critical point corresponded to a max or a min without a graph. Before we develop such a test, we do one more example that sheds more light on the issues our test needs to consider.

Example 14.5.10 Finding critical points and relative extrema.

Let $f(x,y) = x^3 - 3x - y^2 + 4y$. Find the relative extrema of f.

Solution. Once again we start by finding the partial derivatives of f:

$$f_x(x,y) = 3x^2 - 3 \qquad \text{and} \qquad f_y(x,y) = -2y + 4.$$

Each is always defined. Setting each equal to 0 and solving for x and y, we find

$$f_x(x,y) = 0 \Rightarrow x = \pm 1$$
$$f_y(x,y) = 0 \Rightarrow y = 2.$$

We have two critical points: $(-1,2)$ and $(1,2)$. To determine if they

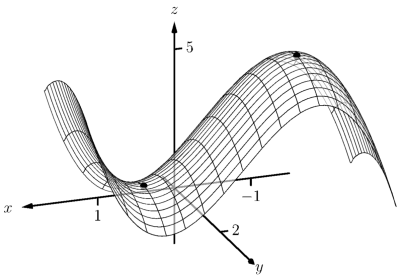


Figure 14.5.11 The surface in Example 14.5.10 with both critical points marked

correspond to a relative maximum or minimum, we consider the graph of f in Figure 14.5.11.

The critical point $(-1, 2)$ clearly corresponds to a relative maximum. However, the critical point at $(1, 2)$ is neither a maximum nor a minimum, displaying a different, interesting characteristic.

If one walks parallel to the y-axis towards this critical point, then this point becomes a relative maximum along this path. But if one walks towards this point parallel to the x-axis, this point becomes a relative minimum along this path. A point that seems to act as both a max and a min is a *saddle point*. A formal definition follows.

Definition 14.5.12 Saddle Point.

Let $P = (x_0, y_0)$ be in the domain of f where $f_x = 0$ and $f_y = 0$ at P. We say P is a **saddle point** of f if, for every open disk D containing P, there are points (x_1, y_1) and (x_2, y_2) in D such that $f(x_0, y_0) > f(x_1, y_1)$ and $f(x_0, y_0) < f(x_2, y_2)$.

At a saddle point, the instantaneous rate of change in all directions is 0 and there are points nearby with z-values both less than and greater than the z-value of the saddle point.

Before Example 14.5.10 we mentioned the need for a test to differentiate between relative maxima and minima. We now recognize that our test also needs to account for saddle points. To do so, we consider the second partial derivatives of f.

Recall that with single variable functions, such as $y = f(x)$, if $f''(c) > 0$, then f is concave up at c, and if $f'(c) = 0$, then f has a relative minimum at $x = c$. (We called this the Second Derivative Test.) Note that at a saddle point, it seems the graph is "both" concave up and concave down, depending on which direction you are considering.

It would be nice if the following were true:

f_{xx} and $f_{yy} > 0$	$\Rightarrow$	relative minimum
f_{xx} and $f_{yy} < 0$	$\Rightarrow$	relative maximum
f_{xx} and f_{yy} have opposite signs	$\Rightarrow$	saddle point.

However, this is not the case. Functions f exist where f_{xx} and f_{yy} are both positive but a saddle point still exists. In such a case, while the concavity in the x-direction is up (i.e., $f_{xx} > 0$) and the concavity in the y-direction is also up (i.e., $f_{yy} > 0$), the concavity switches somewhere in between the x- and y-directions.

To account for this, consider $D = f_{xx}f_{yy} - f_{xy}f_{yx}$. Since f_{xy} and f_{yx} are equal when continuous (refer back to Theorem 11.3.15), we can rewrite this as $D = f_{xx}f_{yy} - f_{xy}^2$. D can be used to test whether the concavity at a point changes depending on direction. If $D > 0$, the concavity does not switch (i.e., at that point, the graph is concave up or down in all directions). If $D < 0$, the concavity does switch. If $D = 0$, our test fails to determine whether concavity switches or not. We state the use of D in the following theorem.

Theorem 14.5.13 Second Derivative Test.

Let R be an open set on which a function $z = f(x, y)$ and all its first and second partial derivatives are defined, let $P = (x_0, y_0)$ be a critical

point of f in R, and let

$$D = f_{xx}(x_0, y_0)f_{yy}(x_0, y_0) - f_{xy}^2(x_0, y_0).$$

1. *If $D > 0$ and $f_{xx}(x_0, y_0) > 0$, then f has a relative minimum at P.*
2. *If $D > 0$ and $f_{xx}(x_0, y_0) < 0$, then f has a relative maximum at P.*
3. *If $D < 0$, then f has a saddle point at P.*
4. *If $D = 0$, the test is inconclusive.*

We first practice using this test with the function in the previous example, where we visually determined we had a relative maximum and a saddle point.

youtu.be/watch?v=xap1euIjzww

Figure 14.5.14 Presenting the second derivative test

Example 14.5.15 Using the Second Derivative Test.

Let $f(x, y) = x^3 - 3x - y^2 + 4y$ as in Example 14.5.10. Determine whether the function has a relative minimum, maximum, or saddle point at each critical point.

Solution. We determined previously that the critical points of f are $(-1, 2)$ and $(1, 2)$. To use the Second Derivative Test, we must find the second partial derivatives of f:

$$f_{xx} = 6x; \qquad f_{yy} = -2; \qquad f_{xy} = 0.$$

Thus $D(x, y) = -12x$.

At $(-1, 2)$: $D(-1, 2) = 12 > 0$, and $f_{xx}(-1, 2) = -6$. By the Second Derivative Test, f has a relative maximum at $(-1, 2)$.

At $(1, 2)$: $D(1, 2) = -12 < 0$. The Second Derivative Test states that f has a saddle point at $(1, 2)$.

The Second Derivative Test confirmed what we determined visually.

Example 14.5.16 Using the Second Derivative Test.

Find the relative extrema of $f(x, y) = x^2y + y^2 + xy$.

Solution. We start by finding the first and second partial derivatives of f:

$$\begin{aligned} f_x &= 2xy + y & f_y &= x^2 + 2y + x \\ f_{xx} &= 2y & f_{yy} &= 2 \\ f_{xy} &= 2x + 1 & f_{yx} &= 2x + 1. \end{aligned}$$

We find the critical points by finding where f_x and f_y are simultaneously 0 (they are both never undefined). Setting $f_x = 0$, we have:

$$f_x = 0 \Rightarrow 2xy + y = 0 \Rightarrow y(2x + 1) = 0.$$

This implies that for $f_x = 0$, either $y = 0$ or $2x + 1 = 0$.

Assume $y = 0$ then consider $f_y = 0$:

$$f_y = 0$$

$$x^2 + 2y + x = 0, \qquad \text{and since } y = 0, \text{ we have}$$
$$x^2 + x = 0$$
$$x(x+1) = 0.$$

Thus if $y = 0$, we have either $x = 0$ or $x = -1$, giving two critical points: $(-1, 0)$ and $(0, 0)$.

Going back to f_x, now assume $2x + 1 = 0$, i.e., that $x = -1/2$, then consider $f_y = 0$:

$$f_y = 0$$
$$x^2 + 2y + x = 0, \qquad \text{and since } x = -1/2, \text{ we have}$$
$$1/4 + 2y - 1/2 = 0$$
$$y = 1/8.$$

Thus if $x = -1/2$, $y = 1/8$ giving the critical point $(-1/2, 1/8)$.

With $D = 4y - (2x + 1)^2$, we apply the Second Derivative Test to each critical point.

At $(-1, 0)$, $D < 0$, so $(-1, 0)$ is a saddle point.

At $(0, 0)$, $D < 0$, so $(0, 0)$ is also a saddle point.

At $(-1/2, 1/8)$, $D > 0$ and $f_{xx} > 0$, so $(-1/2, 1/8)$ is a relative minimum.

Figure 14.5.17 shows a graph of f and the three critical points. Note how this function does not vary much near the critical points — that is, visually it is difficult to determine whether a point is a saddle point or relative minimum (or even a critical point at all!). This is one reason why the Second Derivative Test is so important to have.

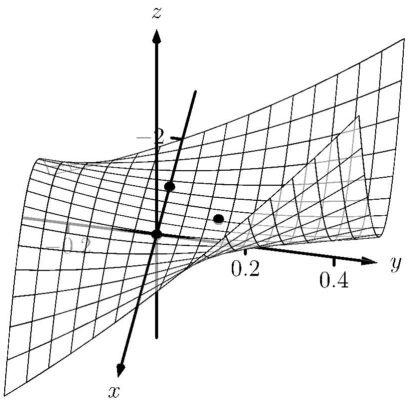


Figure 14.5.17 Graphing f from Example 14.5.16 and its relative extrema

youtu.be/watch?v=XysapWrA7eA

Figure 14.5.18 Two more examples with classification of critical points

14.5.2 Constrained Optimization

When optimizing functions of one variable such as $y = f(x)$, we made use of Theorem 3.1.4, the Extreme Value Theorem, that said that over a closed interval $I = [a, b]$, a continuous function has both a maximum and minimum value. To find these maximum and minimum values, we evaluated f at all critical points in the interval, as well as at the endpoints (the "boundary") of the interval.

A similar theorem and procedure applies to functions of two variables. A continuous function over a closed set also attains a maximum and minimum value (see the following theorem). We can find these values by evaluating the function at the critical values in the set and over the boundary of the set. After formally stating this extreme value theorem, we give examples.

Theorem 14.5.19 Extreme Value Theorem.

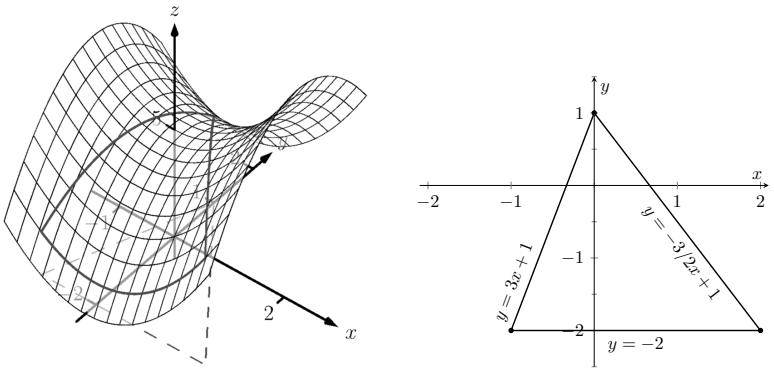
Let $z = f(x, y)$ be a continuous function on a closed, bounded set S. Then f has a maximum and minimum value on S.

Example 14.5.20 Finding extrema on a closed set.

Let $f(x, y) = x^2 - y^2 + 5$ and let S be the triangle with vertices $(-1, -2)$, $(0, 1)$ and $(2, -2)$. Find the maximum and minimum values of f on S.

Solution. It can help to see a graph of f along with the set S. In Fig-

ure 14.5.21(a) the triangle defining S is shown in the xy-plane in a dashed line. Above it is the graph of f; we are only concerned with the portion of the surface $z = f(x, y)$ enclosed by the "triangle".



(a) **(b)**

Figure 14.5.21 Plotting the graph of f along with the restricted domain S in Example 14.5.20

We begin by finding the critical points of f. With $f_x = 2x$ and $f_y = -2y$, we find only one critical point, at $(0, 0)$.

We now find the maximum and minimum values that f attains along the boundary of S, that is, along the edges of the triangle. In Figure 14.5.21(b) we see the triangle sketched in the plane with the equations of the lines forming its edges labeled.

Start with the bottom edge, along the line $y = -2$. If y is -2, then on the surface, we are considering points $f(x, -2)$; that is, our function reduces to $f(x, -2) = x^2 - (-2)^2 + 5 = x^2 + 1 = f_1(x)$. We want to maximize/minimize $f_1(x) = x^2 + 1$ on the interval $[-1, 2]$. To do so, we evaluate $f_1(x)$ at its critical points and at the endpoints.

The critical points of f_1 are found by setting its derivative equal to 0:

$$f_1'(x) = 0 \qquad \Rightarrow x = 0.$$

Evaluating f_1 at this critical point, and at the endpoints of $[-1, 2]$ gives:

$$\begin{aligned} f_1(-1) = 2 \quad &\Rightarrow \quad f(-1, -2) = 2 \\ f_1(0) = 1 \quad &\Rightarrow \quad f(0, -2) = 1 \\ f_1(2) = 5 \quad &\Rightarrow \quad f(2, -2) = 5. \end{aligned}$$

Notice how evaluating f_1 at a point is the same as evaluating f at its corresponding point.

We need to do this process twice more, for the other two edges of the triangle.

Along the left edge, along the line $y = 3x + 1$, we substitute $3x + 1$ in for y in $f(x, y)$:

$$f(x, y) = f(x, 3x+1) = x^2 - (3x+1)^2 + 5 = -8x^2 - 6x + 4 = f_2(x).$$

We want the maximum and minimum values of f_2 on the interval $[-1, 0]$, so we evaluate f_2 at its critical points and the endpoints of the

interval. We find the critical points:

$$f_2'(x) = -16x - 6 = 0 \qquad \Rightarrow \qquad x = -3/8.$$

Evaluate f_2 at its critical point and the endpoints of $[-1, 0]$:

$$\begin{aligned} f_2(-1) = 2 \qquad &\Rightarrow \qquad f(-1,-2) = 2 \\ f_2(-3/8) = 41/8 = 5.125 \qquad &\Rightarrow \qquad f(-3/8, -0.125) = 5.125 \\ f_2(0) = 4 \qquad &\Rightarrow \qquad f(0,1) = 4. \end{aligned}$$

Finally, we evaluate f along the right edge of the triangle, where $y = -3/2x + 1$.

$$f(x,y) = f(x, -3/2x+1) = x^2 - (-3/2x+1)^2 + 5 = -\frac{5}{4}x^2 + 3x + 4 = f_3(x).$$

The critical points of $f_3(x)$ are:

$$f_3'(x) = 0 \qquad \Rightarrow \qquad x = 6/5 = 1.2.$$

We evaluate f_3 at this critical point and at the endpoints of the interval $[0, 2]$:

$$\begin{aligned} f_3(0) = 4 \qquad &\Rightarrow \qquad f(0,1) = 4 \\ f_3(1.2) = 5.8 \qquad &\Rightarrow \qquad f(1.2, -0.8) = 5.8 \\ f_3(2) = 5 \qquad &\Rightarrow \qquad f(2,-2) = 5. \end{aligned}$$

One last point to test: the critical point of f, $(0, 0)$. We find $f(0, 0) = 5$.

We have evaluated f at a total of 7 different places, all shown in Figure 14.5.21(b). We checked each vertex of the triangle twice, as each showed up as the endpoint of an interval twice. Of all the z-values found, the maximum is 5.8, found at $(1.2, -0.8)$; the minimum is 1, found at $(0, -2)$.

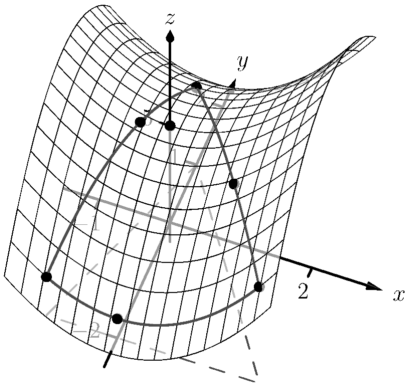


Figure 14.5.22 The graph of f along with important points along the boundary of S and the interior in Example 14.5.20.

youtu.be/watch?v=diIso_CFgCQ

Figure 14.5.23 Finding extreme values

This portion of the text is entitled "Constrained Optimization" because we want to optimize a function (i.e., find its maximum and/or minimum values) subject to a *constraint* — some limit to what values the function can attain. In the previous example, we constrained ourselves by considering a function only within the boundary of a triangle. This was largely arbitrary; the function and the boundary were chosen just as an example, with no real "meaning" behind the function or the chosen constraint.

However, solving constrained optimization problems is a very important topic in applied mathematics. The techniques developed here are the basis for solving larger problems, where more than two variables are involved.

We illustrate the technique once more with a classic problem.

Example 14.5.24 Constrained Optimization.

The U.S. Postal Service states that the girth+length of Standard Post Package must not exceed 130''. Given a rectangular box, the "length" is the longest side, and the "girth" is twice the width+height.

Given a rectangular box where the width and height are equal, what are the dimensions of the box that give the maximum volume subject to

the constraint of the size of a Standard Post Package?

Solution. Let w, h and ℓ denote the width, height and length of a rectangular box; we assume here that $w = h$. The girth is then $2(w+h) = 4w$. The volume of the box is $V(w,\ell) = wh\ell = w^2\ell$. We wish to maximize this volume subject to the constraint $4w+\ell \le 130$, or $\ell \le 130-4w$. (Common sense also indicates that $\ell > 0, w > 0$.)

We begin by finding the critical values of V. We find that $V_w = 2w\ell$ and $V_\ell = w^2$; these are simultaneously 0 at points of the form $(0,\ell)$. These give a volume of 0, so we can ignore these critical points.

We now consider the volume along the constraint $\ell = 130 - 4w$. Along this line, we have:

$$V(w,\ell) = V(w, 130-4w) = w^2(130-4w) = 130w^2 - 4w^3 = V_1(w).$$

The constraint is applicable on the w-interval $[0, 32.5]$ as indicated in the figure. Thus we want to maximize V_1 on $[0, 32.5]$.

Finding the critical values of V_1, we take the derivative and set it equal to 0:

$$V_1'(w) = 260w - 12w^2 = 0 \Rightarrow w(260-12w) = 0 \Rightarrow w = 0, \frac{260}{12} \approx 21.67.$$

We found two critical values: when $w = 0$ and when $w = 21.67$. We again ignore the $w = 0$ solution; the maximum volume, subject to the constraint, comes at $w = h = 21.67$, $\ell = 130 - 4(21.6) = 43.33$. This gives a volume of $V(21.67, 43.33) \approx 20,343 \text{in}^3$.

The volume function $V(w,\ell)$ is shown in Figure 14.5.25 along with the constraint $\ell = 130-4w$. As done previously, the constraint is drawn dashed in the xy-plane and also along the graph of the function. The point where the volume is maximized is indicated.

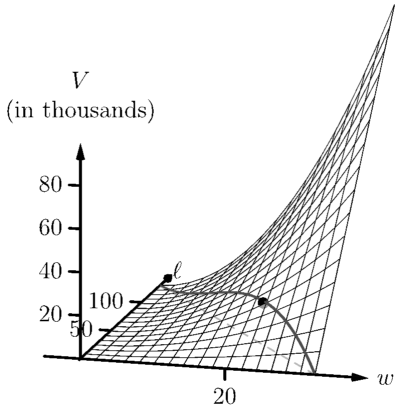


Figure 14.5.25 Graphing the volume of a box with girth $4w$ and length ℓ, subject to a size constraint

It is hard to overemphasize the importance of optimization. In "the real world," we routinely seek to make *something* better. By expressing the *something* as a mathematical function, "making *something* better" means "optimize *some function*."

The techniques shown here are only the beginning of an incredibly important field. Many functions that we seek to optimize are incredibly complex, making the step of "find the gradient and set it equal to $\vec{0}$" highly nontrivial. Mastery of the principles here are key to being able to tackle these more complicated problems.

14.5.3 Exercises

Terms and Concepts

1. True or False? Theorem 14.5.4 states that if f has a critical point at P, then f has a relative extrema at P. (☐ True ☐ False)

2. True or False? A point P is a critical point of f if f_x and f_y are both 0 at P. (☐ True ☐ False)

3. True or False? A point P is a critical point of f if f_x or f_y are undefined at P. (☐ True ☐ False)

4. Explain what it means to "solve a constrained optimization" problem.

Problems

Exercise Group. Find the critical points of the given function. Use the Second Derivative Test to determine if each critical point corresponds to a relative maximum, minimum, or saddle point.

5. $f(x,y) = \frac{1}{2}x^2 + 2y^2 - 8y + 4x$

6. $f(x,y) = x^2 + 4x + y^2 - 9y + 3xy$

7. $f(x,y) = x^2 + 3y^2 - 6y + 4xy$

8. $f(x,y) = \frac{1}{x^2+y^2+1}$

9. $f(x,y) = x^2 + y^3 - 3y + 1$

10. $f(x,y) = \frac{1}{3}x^3 - x + \frac{1}{3}y^3 - 4y$

11. $f(x,y) = x^2y^2$

12. $f(x,y) = x^4 - 2x^2 + y^3 - 27y - 15$

13. $f(x,y) = \sqrt{16 - (x-3)^2 - y^2}$

14. $f(x,y) = \sqrt{x^2 + y^2}$

Exercise Group. Find the absolute maximum and minimum of the function subject to the given constraint.

15. Let $f(x,y) = x^2 + y^2 + y + 1$, constrained to the triangle with vertices $(0,1)$, $(-1,-1)$ and $(1,-1)$.
_____.
_____.

16. Let $f(x,y) = 5x - 7y$, constrained to the region bounded by $y = x^2$ and $y = 1$.
_____.
_____.

17. $f(x,y) = x^2 + 2x + y^2 + 2y$, constrained to the region bounded by the circle $x^2 + y^2 = 4$.

18. $f(x,y) = 3y - 2x^2$, constrained to the region bounded by the parabola $y = x^2 + x - 1$ and the line $y = x$.

14.6 The Derivative as a Linear Transformation

We defined what it means for a real-valued function of two variables to be *differentiable* in Definition Definition 14.1.3 in Section 14.1.

The definition there easily extends to real-valued functions of three or more variables, but it leaves unanswered a couple of natural questions:

1. What about *vector*-valued functions of several variables? (That is, functions f with a domain $D \subseteq \mathbb{R}^n$ and range in $\mathbb{R}^m$ for some $m > 1$.)
2. What is *the* derivative of a function of several variables? After all, we know how to define $f'(x)$ and $\vec{r}\,'(t)$ for real or vector-valued functions of one variable.

One might be tempted at first to simply mimic the definition of the derivative from Chapter 2, but we quickly run into trouble, for a reason that is immediately obvious.

Let $\vec{a}$ be a fixed point in $\mathbb{R}^n$, and let $\vec{h}$ represent a point $(h_1, h_2, \ldots, h_n)$. Since we're treating $\vec{h}$ and $\vec{a}$ as vectors, we can add them, and write down the limit

$$\lim_{\vec{h}\to\vec{0}} \frac{f(\vec{a}+\vec{h}) - f(\vec{a})}{\left\|\vec{h}\right\|}.$$

(Note that division by a vector is nonsense, so we must divide by $\left\|\vec{h}\right\|$, not $\vec{h}$.) But of course, we know that this limit does not exist, because it depends on the direction in which $\vec{h}$ approaches $\vec{0}$! Indeed, if $\vec{h} = h\vec{i}$ or $h\vec{j}$, we get a partial derivative, and for any unit vector $\vec{u}$, setting $\vec{h} = h\vec{u}$ gives us a directional derivative, and we know from Section 14.3 that a directional derivative depends on $\vec{u}$. It seems this approach is doomed to failure. What can we try instead?

To simplify notation, we shift focus slightly and represent points in $\mathbb{R}^n$ by their position vectors, and think of functions of several variables as functions of a vector variable. For example, we'll write $f(\vec{x})$ instead of $f(x_1, x_2, \ldots, x_n)$.

youtu.be/watch?v=Ss4wvGRb6L8

Figure 14.6.1 Defining differentiability in general

14.6.1 The Definition of the Derivative

The key to generalizing the definition of the derivative given in Definition 2.1.7 in Chapter 2 is remembering the following essential property of the derivative: the derivative $f'(a)$ is used to compute the *best linear approximation* to f at a. Indeed, the *linearization* of f at a is the linear function

$$L_a(x) = f(a) + f'(a)(x-a). \tag{14.6.1}$$

That this is the best linear approximation of f at a can be understood as follows: first, note that the graph $y = L_a(x)$ is simply the equation of the tangent line to $y = f(x)$ at a. Second, note that the difference between $f(x)$ and $L_a(x)$ vanishes faster than the difference $x - a$ as x approaches a:

$$\begin{aligned}\lim_{x\to a} \frac{f(x) - L_a(x)}{x-a} &= \lim_{x\to a} \frac{f(x) - (f(a) + f'(a)(x-a))}{x-a} \\ &= \lim_{x\to a} \left(\frac{f(x)-f(a)}{x-a} - f'(a)\frac{x-a}{x-a} \right) \\ &= f'(a) - f'(a) = 0.\end{aligned}$$

While the definition of the derivative doesn't generalize well to several variables, the notion of linear approximation does. Recall from your first course in linear algebra that, given any $m \times n$ matrix A, we can define a function T, called

In general, we say that two real-valued functions of one variable f and g *agree to first order* at a if

$$\lim_{x\to a} \frac{|f(x) - g(x)|}{x-a} = 0.$$

The linearization of f at a is the unique linear function that agrees with f to first order at a. Going further, we can say that f and g *agree to order* k at a if

$$\lim_{x\to a} \frac{|f(x) - g(x)|}{(x-a)^k} = 0.$$

For example one could define the degree n Taylor polynomial of a function f at a to be the unique polynomial of degree n that agrees with f to order k at a.

a *linear transformation*, that takes an $n \times 1$ column vector as input, and produces an $m \times 1$ column vector as output:

$$T(\vec{x}) = A\vec{x} = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{bmatrix} = \vec{y}.$$

In the above definition, the product $A\vec{x}$ is the usual matrix product of the $m \times n$ matrix A with the $n \times 1$ matrix $\vec{x}$. In this text, we generally do not write our vectors as columns, so for a vector $\vec{x} = \langle x_1, \ldots, x_n \rangle$ we will use the notation

$$A \cdot \vec{x} = \langle a_{11}x_1 + \cdots + a_{1n}x_n, \ldots, a_{m1}x_1 + \cdots + a_{mn}x_n \rangle$$

to represent the same product in our notation. (And yes, the dot in this product is intended to remind you of the dot product between vectors: recall that the (i, j)-entry of a matrix product AB is the dot product of the ith row of A with the jth column of B.) We can now make the following definition.

To avoid confusion between the meaning of *linear function* in Calculus, and *linear transformation* in Linear Algebra, we will use ℓ to denote the former, and T to denote the latter. Notice that if $\vec{b} = \vec{0}$ in Definition 14.6.2, then a linear function *is* a linear transformation.

Definition 14.6.2 Linear function.

A function ℓ from $\mathbb{R}^n$ to $\mathbb{R}^m$ will be called a **linear function** if ℓ is of the form

$$\ell(\vec{x}) = M \cdot \vec{x} + \vec{b}$$

for some $m \times n$ matrix M and vector $\vec{b}\,\mathbb{R}^m$.

If we apply the convention of representing points in terms of their position vectors to the codomain as well as the domain, we can express such a function as $f = \langle f_1, \ldots, f_n \rangle$, where each function f_i is a real-valued function of n variables. We want differentiability of f to mean that f has a linear approximation ℓ that agrees with f to first order at a. Since $f(\vec{x})$ and $\ell(\vec{x})$ are now vectors, saying that ℓ is a good approximation of f requires that the magnitude $\|f(\vec{x}) - \ell(\vec{x})\|$ is small relative to the size of $\|\vec{x} - \vec{a}\|$.

Definition 14.6.3 General definition of differentiability.

Let D be an open subset of $\mathbb{R}^n$ and let f be a function with domain D and values in $\mathbb{R}^m$. We say that f is **differentiable** at a point $\vec{a} \in D$ if there exists a linear function $\ell : \mathbb{R}^n \to \mathbb{R}^m$ that agrees with f to first order at $\vec{a}$; that is, if

$$\lim_{\vec{x} \to \vec{a}} \frac{\|f(\vec{x}) - \ell(\vec{x})\|}{\|\vec{x} - \vec{a}\|} = 0.$$

This definition is going to take a lot of unpacking. First of all, what is this function ℓ? How do we compute it? Does this definition include Definition 14.1.3 from Section 14.1 as a special case? What about differentiability for vector-valued functions of one variable, or real-valued functions of one variable?

We will answer the first two questions in due course. The answer to the rest is, "Yes." The above definition generalizes all the definitions of differentiability we've encountered so far. As a first step, let us note that for $\ell(\vec{x}) = M \cdot \vec{x} + \vec{b}$, we must have $\ell(\vec{a}) = f(\vec{a})$, or the limit above will not exist. Thus $M \cdot \vec{a} + \vec{b} = f(\vec{a})$, so $\vec{b} = f(\vec{a}) - M \cdot \vec{a}$. This tells us that ℓ must have the following form:

$$\ell(\vec{x}) = M \cdot \vec{x} + \vec{b}$$

$$\begin{aligned} &= M \cdot \vec{x} + (f(\vec{a}) - M \cdot \vec{a}) \\ &= f(\vec{a}) + M \cdot (\vec{x} - \vec{a}). \end{aligned} \tag{14.6.2}$$

This should ring some bells: the form of ℓ is very similar to that of the linearization given for a function of one variable in (14.6.1) above, with the matrix M playing the role of $f'(a)$. Perhaps this matrix is the derivative we seek?

14.6.2 Real-valued functions of several variables

Let $f : D \subseteq \mathbb{R}^n \to \mathbb{R}$ be a given function of n variables (you can assume $n = 1, 2$ or 3 if you prefer). Let us denote a point $(x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$ using the vector $\vec{x} = \langle x_1, x_2, \ldots, x_n \rangle$, so that $f(\vec{x}) = f(x_1, x_2, \ldots, x_n)$. Let $\vec{a} = \langle a_1, a_2, \ldots, a_n \rangle$ denote a fixed point $(a_1, a_2, \ldots, a_n) \in D$.

In Section 14.1, we saw that differentiability means that the difference $\Delta z = f(x + dx, y + dy) - f(x, y)$ can be approximated by the differential $dz = f_x(x, y)\, dx + f_y(x, y)\, dy$. Differentiability was defined to mean that the error functions E_x and E_y, defined by

$$E_x\, dx + E_y\, dy = \Delta z - dz,$$

go to zero as $\langle dx, dy \rangle$ goes to zero. Let's rephrase this so that it works for any number of variables. Recall that the *gradient* of f at $\vec{a} \in D$ is the vector $\nabla f(\vec{a})$ defined by

$$\nabla f(\vec{a}) = \left\langle \frac{\partial f}{\partial x_1}(\vec{a}), \frac{\partial f}{\partial x_2}(\vec{a}), \ldots, \frac{\partial f}{\partial x_n}(\vec{a}) \right\rangle.$$

Definition 14.6.4 The linearization of a function of several variables.

Let f be continuously differentiable on some open set $D \subseteq \mathbb{R}^n$, and let $\vec{a} \in D$. The **linearization** of f at $\vec{a}$ is the function $L_{\vec{a}}(\vec{x})$ defined by

$$L_{\vec{a}}(\vec{x}) = f(\vec{a}) + \nabla f(\vec{a}) \cdot (\vec{x} - \vec{a}).$$

If the point $\vec{a} \in D$ at which we are considering the linearization is fixed, or clear from context in a given problem, we can drop the subscript in the notation, and simply write $L(\vec{x})$ instead of $L_{\vec{a}}(\vec{x})$.

When $n = 1$, we get the linearization $L_a(x) = f(a) + f'(a)(x - a)$, which is the usual linearization from Calculus I. (You might also notice that $L_a(x)$ is the first-degree Taylor polynomial of f about $x = a$. The same is true of the linearization of f for more than one variable, although we will not be considering Taylor polynomials in several variables.)

For $n = 2$, we get the linear approximation associated to the total differential:

$$\begin{aligned} L_{(a,b)}(x, y) &= f(a, b) + \langle f_x(a, b), f_y(a, b) \rangle \cdot \langle x - a, y - b \rangle \\ &= f(a, b) + f_x(a, b)(x - a) + f_y(a, b)(y - b). \end{aligned}$$

Compare this with Equation (14.6.2) above. It seems that the gradient $\nabla f(\vec{a})$ is our matrix M in this case: for a real-valued function, $m = 1$, so we expect a $1 \times n$ row matrix, and the gradient certainly can be interpreted to fit that description.

Viewing the gradient ∇f as a $1 \times n$ matrix M, the product $M \cdot \vec{x}$ defined above is indeed exactly the same as the usual dot product $\nabla f(\vec{a}) \cdot \vec{x}$.

For real-valued functions, Definition 14.6.3 becomes the following:

Definition 14.6.5 Differentiability of real-valued functions.

We say that f is **differentiable** at $\vec{a} \in D$ if $\nabla f(\vec{a})$ exists, and $f(\vec{x})$ and

$L_{\vec{a}}(\vec{x})$ agree to first order at $\vec{a}$; that is, if

$$\lim_{\vec{x}\to\vec{a}} \frac{|f(\vec{x}) - L_{\vec{a}}(\vec{x})|}{\|\vec{x}-\vec{a}\|} = \lim_{\vec{x}\to\vec{a}} \frac{|f(\vec{x}) - f(\vec{a}) - \nabla f(\vec{a})\cdot\langle\vec{x}-\vec{a}\rangle|}{\|\vec{x}-\vec{a}\|} = 0.$$

What this definition says is that the linearization $L_{\vec{a}}(\vec{x})$ is a good linear approximation to f at $\vec{a}$. In fact, it's the *only* (and hence, best) linear approximation: if a linear approximation exists, it has to be $L_{\vec{a}}(\vec{x})$.

If you want to see why this has to be true, recall that since the above limit exists, we have to be able to evaluate it along any path we like. Suppose we chose the path

$$\vec{r}(t) = \langle h, a_2, \ldots, a_n\rangle.$$

Then $\vec{x}-\vec{a} = \langle h, 0, \ldots, 0\rangle = h\vec{i}$, and our definition becomes:

$$\lim_{h\to 0}\left|\frac{f(a_1+h, a_2, \ldots, a_n) - f(a_1, a_2, \ldots, a_n)}{h} - \frac{\partial f}{\partial x_1}(a_1, a_2, \ldots, a_n)\right| = 0,$$

which is just another way of stating the definition of the partial derivative with respect to x_1. Of course, approaching along any of the other coordinate directions will similarly produce the other partial derivatives.

Recall that in one variable, the derivative is often written instead in terms of $h = x - a$, so that

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

Recall that

$$\|\vec{x}-\vec{a}\| = \sqrt{(x_1-a_1)^2 + \cdots + (x_n-a_n)^2}$$

is the distance from $\vec{x}$ to $\vec{a}$. In general, we would say that two functions $f(\vec{x})$ and $g(\vec{x})$ "agree up to order k" at $\vec{a}$ if

$$\lim_{\vec{x}\to\vec{a}} \frac{f(\vec{x}) - g(\vec{x})}{\|\vec{x}-\vec{a}\|^k} = 0.$$

As an exercise, check that, for $n = 1$, two functions f and g agree up to order k at a if and only if their degree k Taylor polynomials are equal. (A similar statement is true in more than one variable.)

In more than one variable, we can define $h_i = x_i - a_i$, for $i = 1, \ldots, n$, or the corresponding vector $\vec{h} = \vec{x} - \vec{a}$. The definition of differentiability then can be written as

$$\lim_{\vec{h}\to\mathbf{0}} \frac{|f(\vec{a}+\vec{h}) - f(\vec{a}) - \nabla f(\vec{a})\cdot\vec{h}|}{\|\vec{h}\|} = 0. \qquad (14.6.3)$$

Note that we want the difference between $f(\vec{a}+\vec{h})$ and $L_{\vec{a}}(\vec{h})$ to go to zero faster than $\|\vec{h}\|$ goes to zero, and that it only makes sense to divide by the length of $\vec{h}$, since division by a vector (or the corresponding point) is not defined.

Let's return to $n = 2$ and Definition 14.1.3 from Section 14.1. If we write $\vec{h} = \langle dx, dy\rangle$, then $f(\vec{a}+\vec{h}) - f(\vec{a}) = \Delta z$, and $\nabla f(\vec{a})\cdot\vec{h} = dz$, and Equation (14.6.3) becomes

$$\lim_{\vec{h}\to\vec{0}} \frac{|\Delta z - dz|}{\left\|\vec{h}\right\|} = \lim_{\vec{h}\to\vec{0}} \frac{|E_x\,dx + E_y\,dz|}{\|\langle dx, dy\rangle\|} = 0,$$

which is another way of saying that the error terms E_x, E_y must vanish as dx and dy approach zero. Success! Definition 14.6.3 is indeed a generalization of Definition 14.1.3.

Note that we've also generalized Definition 2.1.7 for functions of one variable as well: Equation (14.6.3) becomes

$$\lim_{h\to 0}\left|\frac{f(a+h)-f(a)}{h} - f'(a)\right| = 0,$$

which is just another way of re-writing the usual definition of the derivative. In fact, we've also generalized Definition 13.2.10 from Chapter 13 for differentiability of vector-valued functions: all we have to do is write our vector-valued function as a column matrix.

For

$$\vec{r}(t) = \begin{bmatrix} x_1(t) \\ x_2(t) \\ \vdots \\ x_m(t) \end{bmatrix} \quad \text{and} \quad \vec{r}'(t) = \begin{bmatrix} {x_1}'(t) \\ {x_2}'(t) \\ \vdots \\ {x_m}'(t) \end{bmatrix},$$

we have

$$\lim_{h\to 0} \left\| \frac{1}{h}(\vec{r}(a+h) - \vec{r}(a)) - \vec{r}'(a) \right\| = 0,$$

which again reproduces the definition of $\vec{r}'(a)$.

One of the results we learn in Calculus I is that differentiability implies continuity. The situation is no different in general, and with our new definition of differentiability, an easy proof is possible.

Theorem 14.6.6 Differentiability implies continuity.

If $f : D \subseteq \mathbb{R}^n \to \mathbb{R}$ is differentiable at $\vec{a} \in D$, then f is continuous at $\vec{a}$.

Proof. Suppose that f is differentiable at $\vec{a}$. Then we know that

$$\lim_{\vec{x}\to\vec{a}} \frac{f(\vec{x}) - L_{\vec{a}}(\vec{x})}{\|\vec{x} - \vec{a}\|} = \lim_{\vec{x}\to\vec{a}} \frac{f(\vec{x}) - f(\vec{a}) - \nabla f(\vec{a}) \cdot \langle \vec{x} - \vec{a} \rangle}{\|\vec{x} - \vec{a}\|} = 0.$$

By the definition of continuity, we need to show that $\lim_{\vec{x}\to\vec{a}} f(\vec{x}) = f(\vec{a})$. We have that

$$\begin{aligned} f(\vec{x}) &= f(\vec{a}) + (f(\vec{x}) - f(\vec{a})) \\ &= f(\vec{a}) + (f(\vec{x}) - f(\vec{a}) - \nabla f(\vec{a}) \cdot (\vec{x} - \vec{a})) + \nabla f(\vec{a}) \cdot (\vec{x} - \vec{a}) \\ &= f(\vec{a}) + \left(\frac{f(\vec{x}) - f(\vec{a}) - \nabla f(\vec{a}) \cdot (\vec{x} - \vec{a})}{\|\vec{x} - \vec{a}\|} \right) (\|\vec{x} - \vec{a}\|) + \nabla f(\vec{a}) \cdot (\vec{x} - \vec{a}). \end{aligned}$$

Thus, taking limits of the above as $\vec{x} \to \vec{a}$, we find $\lim_{\vec{x}\to\vec{a}} f(\vec{x}) = f(\vec{a})$, since the first term is a constant ($f(\vec{a})$), the second is the product of two terms that both go to zero (the first term is zero by the definition of differentiability, and clearly $\lim_{\vec{x}\to\vec{a}} \|\vec{x} - \vec{a}\| = 0$), and the last term vanishes since it's linear (and thus continuous) in $\vec{x}$, and so, by direct substitution,

$$\lim_{\vec{x}\to\vec{a}} \nabla f(\vec{a}) \cdot (\vec{x} - \vec{a}) = \nabla f(\vec{a}) \cdot (\vec{a} - \vec{a}) = 0.$$

■

14.6.3 Vector-valued functions of several variables

Let us now consider Definition 14.6.3 for general functions $f : D \subseteq \mathbb{R}^n \to \mathbb{R}^m$. If f is differentiable at $\vec{a}$, then we must have

$$\lim_{\vec{x}\to\vec{a}} \frac{\|f(\vec{x}) - \ell(\vec{x})\|}{\|\vec{x} - \vec{a}\|} = 0$$

for some linear function $\ell(\vec{x})$. Moreover, we'll see below that (a) the matrix M is uniquely defined, and (b) M is deserving of the title of "the" derivative of f.

We saw in Equation (14.6.2) above that T must have the form of a linear approximation:

$$\ell(\vec{x}) = L_{\vec{a}}(\vec{x}) = f(\vec{a}) + M \cdot (\vec{x} - \vec{a}).$$

To keep everything straight, in this exposition we are going to write our vectors as column matrices rather than using angle bracket notation.

Let's compare again to the one variable case: $L_a(x) = f(a) + f'(a)(x-a)$. With this in mind, the matrix M, whatever it is, certainly seems to play the role of the derivative for general functions from $\mathbb{R}^n$ to $\mathbb{R}^m$. It remains to determine the matrix M, and see that there can only be one possibility. To that end, let us write

$$M = \begin{bmatrix} c_{11} & c_{12} & \cdots & c_{1n} \\ c_{21} & c_{22} & \cdots & c_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ c_{m1} & c_{m2} & \cdots & c_{mn} \end{bmatrix},$$

and consider what happens when we let $\vec{x} \to \vec{a}$ along different paths.

If we consider the path $x_1 = a_1 + t, x_2 = a_2, \ldots, x_n = a_n$ (that is, varying x_1 while holding the other variables constant) then

$$\vec{x} - \vec{a} = \langle a_1 + t, a_2, \ldots, a_n \rangle - \langle a_1, a_2, \vdots, a_n \rangle = \langle t, 0, \ldots, 0 \rangle,$$

so $M \cdot (\vec{x} - \vec{a})$ gives us t times the first column of M, since for each row of M, the first entry is multiplied by t, and the remaining entries are multiplied by zero. Thus,

$$M \cdot (\vec{x} - \vec{a}) = \langle c_{11}, c_{21}, \ldots, c_{m1} \rangle$$

along this path.

Now we consider the limit as $t \to 0$.

$$\lim_{t \to 0} \left| \frac{f(a_1 + t, a_2, \ldots, a_n) - f(a_1, a_2, \ldots, a_n)}{t} - \langle c_{11}, c_{21}, \ldots, c_{m1} \rangle \right| = 0.$$

Since $\langle c_{11}, c_{21}, \ldots, c_{m1} \rangle$ is a constant vector, from differentiability of f, together with Definition 14.6.3, we get

$$\lim_{t \to 0} \frac{f(a_1 + t, a_2, \ldots, a_n) - f(a_1, a_2, \ldots, a_n)}{t} = \langle c_{11}, c_{21}, \ldots, c_{m1} \rangle.$$

But this limit on the left is just the partial derivative of f with respect to x_1! If we write $f(\vec{x}) = \langle f_1(\vec{x}), f_2(\vec{x}), \ldots, f_m(\vec{x}) \langle$, then we have

$$\lim_{t \to 0} \frac{f(a_1 + t, a_2, \ldots, a_n) - f(a_1, a_2, \ldots, a_n)}{t} = \left\langle \frac{\partial f_1}{\partial x_1}(\vec{a}), \frac{\partial f_2}{\partial x_1}(\vec{a}), \ldots, \frac{\partial f_m}{\partial x_1}(\vec{a}) \right\rangle,$$

and this gives us the first column of M! Repeating this for each variable, we see that the matrix M is exactly the matrix of all the partial derivatives of f. This matrix is important enough to have a name:

Definition 14.6.7 The Jacobian matrix of a differentiable function.

Let $D \subseteq \mathbb{R}^n$ be an open subset, and let $f : D \to \mathbb{R}^m$ be a differentiable function. At any point $\vec{a} \in D$, the **Jacobian matrix** of f at $\vec{a}$, denoted $Df(\vec{a})$, is the $m \times n$ matrix defined by

$$Df(\vec{a}) = \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial x_1} & \cdots & \frac{\partial f_m}{x_n} \end{bmatrix}.$$

The linear transformation $T_{f,\vec{a}} : \mathbb{R}^n \to \mathbb{R}^m$ defined by $T_{f,\vec{a}}(\vec{x}) = Df(\vec{a}) \cdot \vec{x}$ is defined to be the **derivative** of f at $\vec{a}$.

Notice that if f is differentiable, the Jacobian matrix is the only matrix that can fit the definition: the fact that the limit must be zero along a path parallel to one of the coordinate axes forces the matrix M to contain the partial derivatives of f.

In particular, note that for a function $f : \mathbb{R}^n \to \mathbb{R}$, we recover the gradient vector. Technically, the derivative in this sense is a *row* vector (some might say *dual vector*), not a column vector. Note that multiplying a row vector by a column vector is the same as taking the dot product of two column vectors.

This definition also accounts for parametric curves, viewed as vector-valued functions of one variable. If $\mathbf{r} : \mathbb{R} \to \mathbb{R}^n$ defines a parametric curve, then the derivative $\mathbf{r}'(t) = \begin{bmatrix} x_1'(t) \\ x_2'(t) \\ \vdots \\ x_n'(t) \end{bmatrix}$ as introduced in Chapter 13 is the same as the one obtained using this definition.

14.6.4 The General Chain Rule

One of the big advantages of representing the derivative of a function of several variables in terms of its Jacobian matrix is that the Chain Rule becomes completely transparent. Arguably, the version of the Chain Rule we're about to present is even more intuitive than the single-variable version!

Recall that the Chain Rule is all about derivatives of composite functions. In one variable, given $h = f \circ g$, if $b = g(a)$, we have

$$h'(a) = f'(g(a))g'(a) = f'(b)g'(a).$$

The derivative of the composition is the product of the derivatives of the functions being composed, as long as we take care to evaluate them at the appropriate points.

In Section 14.2 we saw that in several variables, the Chain Rule comes in various flavours, depending on the number of variables involved in each function being composed. If we think of derivatives in terms of the Jacobian matrix, then each of these flavours says exactly the same thing as the original Chain Rule above!

Theorem 14.6.8 The general Chain Rule (matrix form).

Let $f : U \subseteq \mathbb{R}^m \to \mathbb{R}^p$ and $g : V \subseteq \mathbb{R}^n \to \mathbb{R}^m$ be differentiable functions, such that the range of g is contained in the domain U of f. Then the composite function $h = f \circ g$ is differentiable on V, and for each $\vec{a} \in V$, we have

$$Dh(\vec{a}) = D(f \circ g)(\vec{a}) = Df(\vec{b})Dg(\vec{a}),$$

where $\vec{b} = g(\vec{a})$, and the product on the right is the usual matrix product of the two Jacobian matrices.

This is a remarkable result. Let's unpack it in a couple of examples.

Example 14.6.9 Applying the general chain rule.

Let $f : U \subseteq \mathbb{R}^3 \to \mathbb{R}$ be a differentiable function of three variables, and let $\vec{r}(t) = \langle x(t), y(t), z(t) \rangle$ be a vector-valued function of one variable. Use Theorem 14.6.8 to determine a formula for the derivative of $h(t) =$

$f(\vec{r}(t))$.

Solution. We already know what this derivative should look like from Section 14.2. The point is to confirm that this is a special case of Theorem 14.6.8. The Jacobian matrix of f is a 1×3 matrix and Jacobian matrix of $\vec{r}$ is a 3×1 matrix. They are given, respectively, by

$$Df(\vec{x}) = \begin{bmatrix} f_x(\vec{x}) & f_y(\vec{x}) & f_z(\vec{x}) \end{bmatrix} \quad \text{and} \quad D\vec{r}(t) = \begin{bmatrix} x'(t) \\ y'(t) \\ z'(t) \end{bmatrix}.$$

Theorem 14.6.8 then gives us

$$h'(t) = Df(\vec{r}(t))D\vec{r}(t) = f_x(r(\vec{t}))x'(t) + f_y(r(\vec{t}))y'(t) + f_z(r(\vec{t}))z'(t),$$

as before. Of course, in this context we usually write $Df(\vec{x})$ as $\nabla f(\vec{x})$ and $D\vec{r}(t)$ as $\vec{r}\,'(t)$, and instead of a matrix product, we write a dot product. But this is simply a shift in notation — the quantities involved are no different than before.

Example 14.6.10 Applying the general chain rule.

Let $f : U \subseteq \mathbb{R}^2 \to \mathbb{R}$ be a function of 2 variables, and let $g : V \subseteq \mathbb{R}^2 \to \mathbb{R}^2$ be given by

$$g(u, v) = (x(u, v), y(u, v)).$$

Given $h = f \circ g$, use Theorem 14.6.8 to determine h_u and h_v.

Solution. First we compute the Jacobian matrices for f and g. We have

$$Df(x, y) = \begin{bmatrix} f_x(x, y) & f_y(x, y) \end{bmatrix} \quad \text{and} \quad Dg(u, v) = \begin{bmatrix} x_u(u, v) & x_v(u, v) \\ y_u(u, v) & y_v(u, v) \end{bmatrix}.$$

The Chain Rule then gives

$$\begin{aligned} Dh(u, v) &= \begin{bmatrix} h_u(u, v) & h_v(u, v) \end{bmatrix} = Df(h(u, v))Dh(u, v) \\ &= \begin{bmatrix} f_x(h(u, v)) & f_y(h(u, v)) \end{bmatrix} \begin{bmatrix} x_u(u, v) & x_v(u, v) \\ y_u(u, v) & y_v(u, v) \end{bmatrix} \\ &= [f_x(h(u, v)x_u(u, v) + f_y(h(u, v))y_u(u, v) \\ &\qquad f_x(h(u, v))x_v(u, v) + f_y(h(u, v))y_v(u, v)]. \end{aligned}$$

Equating coefficients of the first and last matrices, we have, in Leibniz notation,

$$\begin{aligned} \frac{\partial h}{\partial u} &= \frac{\partial f}{\partial x}\frac{\partial x}{\partial u} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial u} \\ \frac{\partial h}{\partial v} &= \frac{\partial f}{\partial x}\frac{\partial x}{\partial v} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial v}. \end{aligned}$$

Again, this reproduces another instance of the Chain Rule from Section 14.2.

With additional experimentation, you will find that every instance of the Chain Rule you have previously encountered can be interpreted as a special case of Theorem 14.6.8. Moreover, a slight shift in interpretation makes this version of the Chain Rule even more obvious! (There's another detour coming, but stick

with us.)

Let us digress briefly and discuss the progression of mathematics from Calculus to higher math. If you continue on to upper-level undergraduate mathematics, you will encounter courses in Analysis and Topology. Analysis deals with the theoretical underpinnings of Calculus: this is where you see all the careful proofs of theorems that have been omitted from this text. Topology is a further abstraction of Analysis. In Topology, one studies continuity (and its consequences) at its most fundamental, abstract level.

The corresponding successors to Calculus in several variables are known as *differential geometry* and *differential topology*. You probably won't encounter these unless you continue on to graduate studies in mathematics. One of the core philosophies in these two (closely related) subjects is the following:

Functions map points. Derivatives map tangent vectors.

This can be understood in our context. At any point $\vec{a}$ in $\mathbb{R}^n$, we can attach a copy of the *vector space* $\mathbb{R}^n$, thought of as all the possible tangent vectors to curves passing through that point.

Let $\vec{r} : (a, b) \to \mathbb{R}^n$ be such a curve, and let $f : \mathbb{R}^n \to \mathbb{R}^m$ be a differentiable function. The composite function $\vec{s} = f \circ \vec{r}$ is then a curve in $\mathbb{R}^m$. The point $\vec{a} = \vec{r}(t_0)$ on our first curve in $\mathbb{R}^n$ becomes a point

$$\vec{b} = f(\vec{a}) = f(\vec{r}(t_0)) = \vec{s}(t_0)$$

on our new curve in $\mathbb{R}^m$. What about tangent vectors?

At the point $\vec{a}$, we have the tangent vector $\vec{v} = \vec{r}\,'(t_0)$. What is the tangent vector to $\vec{s}(t)$ at the point $\vec{b}$? On the one hand, by definition, we have the tangent vector

$$\vec{w} = \vec{s}\,'(t_0).$$

On the other hand, the Chain Rule gives us

$$\vec{s}\,'(t_0) = (f \circ \vec{r})'(t_0) = Df(\vec{r}(t_0))\vec{r}\,'(t_0).$$

But $\vec{r}\,'(t_0) = \vec{v}$, so we have

$$\vec{w} = Df(\vec{a}) \cdot \vec{v}.$$

Multiplying the original tangent vector by the derivative of f gives us the new tangent vector. Cool!

What's more, we can view this as a *linear transformation*. Let V denote the vector space of all tangent vectors at the point $\vec{a}$ in $\mathbb{R}^n$ (this is just a copy of $\mathbb{R}^n$) and let W denote the space of all tangent vectors in $\mathbb{R}^m$ at the point $\vec{b}$. Then we have the linear transformation $T : V \to W$ given by

$$T(\vec{v}) = Df(\vec{a}) \cdot \vec{v}.$$

In more advanced Calculus, or Differential Geometry, we view this linear transformation as *the* derivative of f at $\vec{a}$. Now, recall from Linear Algebra that matrix multiplication corresponds to the composition of the corresponding linear transformations: if $S(\vec{v}) = A\vec{v}$ and $T(\vec{w}) = B\vec{w}$, and the matrices A and B are of the appropriate sizes, then

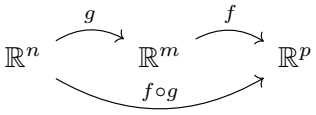
$$S \circ T(\vec{w}) = S(T(\vec{w})) = A(B\vec{w}) = (AB)\vec{w}.$$

Suppose we have differentiable functions $f : \mathbb{R}^n \to \mathbb{R}^m$ and $g : \mathbb{R}^m \to \mathbb{R}^p$. Let $T_f : \mathbb{R}^n \to \mathbb{R}^m$ be the linear function given by the derivative of f, and

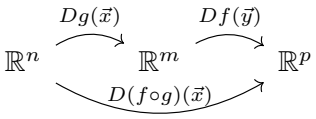
let $T_g : \mathbb{R}^m \to \mathbb{R}^p$ be the linear function given by the derivative of g. The chain rule is then essentially telling us that *the derivative of a composition is the composition of the derivatives*: we have

$$T_f \circ T_g(\vec{v}) = T_f(T_g(\vec{v})) = Df(\vec{y})(Dg(\vec{x})\vec{v}) = (Df(g(\vec{x}))Dg(\vec{x}))\vec{v} = T_{f\circ g}(\vec{v}).$$

In other words, given the composition



we have the corresponding composition



(But beware of the dual usage of $\mathbb{R}^n$ here. In the first composition, we're thinking of it as a set of points in the domain of a function. In the second composition, we're thinking of it as the set of tangent vectors at a point.)

This turns out to be an extremely powerful way of looking at derivatives and the Chain Rule. You may want to keep this in mind in later sections, such as when we consider change of variables in multiple integrals at the end of Chapter 15, and when we define integrals over curves and surfaces in Chapter 16. We won't use this language when we get there, but many of the results in those sections (for example, the formula for surface area of a parametric surface) can be understood according to the two principles we have just seen: *functions map points*, *while derivatives map tangent vectors*, and *the derivative of a composition is the composition of the derivatives*.

14.7 Constrained Optimization and Lagrange Multipliers

Let us continue our discussion of constrained optimization begun in Section 14.5. Theorem 14.5.19 tells us that the Extreme Value Theorem extends to functions of two variables; in fact, this is true for a function of any number of variables: if a real-valued function f is continuous on a closed, bounded subset of $\mathbb{R}^n$, then it is guaranteed to have an absolute maximum and minimum.

However, as the number of variables increases, the job of finding these absolute extrema becomes more and more complicated. We saw one approach in Section 14.5: given a continuous function on a closed, bounded region D, we first consider critical values on the interior of D. We then restrict our function f to points on the boundary of D, and attempt to reduce the problem to optimization in one variable.

In many cases, this approach is best accomplished by parametrizing the boundary. We learned how to define parametric curves in the plane in Section 10.2.

Example 14.7.1 Constrained optimization by parametrization.

Find the absolute maximum and minimum values of $f(x, y) = x^2 - 8x - 3y^2$ on the disc $x^2 + y^2 \leq 4$.

Solution. First, we check for critical points: We have

$$\nabla f(x,y) = \langle 2x-8, -6y\rangle\,,$$

which vanishes when $(x,y) = (4,0)$. This critical point is outside our region, so we do not consider it.

Next, we look for extreme values on the boundary. The boundary of our region is the circle $x^2+y^2=4$, which we can parametrize using $x = 2\cos t$, $y = 2\sin t$, for $t \in [0,2\pi]$. For (x,y) on the boundary, we have

$$f(x,y) = x^2 - 8x - 3y^2 = 4\cos^2 t - 16\cos t - 12\sin^2 t = h(t),$$

a function of one variable, with domain $[0,2\pi]$.

We learned how to find the extreme values of such a function back in our first course in calculus: see Section 3.1. We have $h(0) = h(2\pi) = -12$, and

$$h'(t) = -8\cos t\sin t + 16\sin t - 24\sin t\cos t = 16\sin t(1-2\cos t).$$

Thus, $h'(t)=0$ if $\sin t = 0$ ($t = 0,\pi,2\pi$) or $\cos = \frac{1}{2}$ ($t = \pi/3, 5\pi/3$). We have already checked that $h(0) = h(2\pi) = -12$, so we check the remaining points:

$$h(\pi) = 4(-1)^2 - 16(-1) = 20$$
$$h(\pi/3) = h(5\pi/3) = 4\left(\frac{1}{4}\right) - 16\left(\frac{1}{2}\right) - 12\left(\frac{3}{4}\right) = -16.$$

We see that the absolute maximum is when $t=\pi$: $h(\pi) = f(-2,0) = 20$, and the absolute minimum is -16, which occurs when $t = \pi/3$ and $t = 5\pi/3$, corresponding to the points $(1,\sqrt{3})$ and $(1,-\sqrt{3})$, respectively.

The above method works well, when it's straightforward to set up. The advantage is that it reduces the problem of optimization along the boundary to an optimization problem in one variable, which is something we mastered long ago.

One downside is that it is not always easy to come up with a parametrization for a curve. In the above example, the boundary $x^2+y^2=4$ is a *level curve*: it's of the form $g(x,y)=c$. When we're trying to optimize subject to a constraint of this form, there is another approach, called the **method of Lagrange multipliers**.

youtu.be/watch?v=W5ykqB7261c

Figure 14.7.2 Introducing Lagrange multipliers

Suppose we are trying to maximize a function $f(x,y)$ subject to a constrain $g(x,y)=c$. We could follow the approach given above: find a function $\vec{r}:[a,b]\to\mathbb{R}^2$ that parametrizes the curve $g(x,y)=c$. As we saw above, the maximum (or minimum) should occur at some point t_0 that is a critical number of $h(t) = f(\vec{r}(t))$; that is, such that

$$h'(t_0) = \nabla f(\vec{r}(t_0))\cdot\vec{r}\,'(t_0) = 0.$$

This tells us that the gradient ∇f should be orthogonal to the constraint curve $g(x,y)=c$ at the point $(x_0,y_0) = (x(t_0),y(t_0))$. But we know another gradient that is orthogonal to this curve: ∇g! Recall from Theorem 14.3.9 that $\nabla g(x,y)$ is always orthogonal to the level curve $g(x,y)=c$ at points along the curve.

If (x_0, y_0) is a critical point of a function g; that is, if $\nabla g(x_0, y_0) = \vec{0}$, and $g(x_0, y_0) = c$, we call the number c a **critical value** of g. Any number that is not a critical value is called a regular value. Often, if c is a critical value, the level set defined by $g(x, y) = c$ is not a smooth curve, or even a curve at all.

For example, $g(x, y) = x^2 + y^2$ has the critical point $(0, 0)$, and critical value 0. The set of (x, y) with $x^2 + y^2 = 0$ is not a curve; it's a single point. Because of this, it's usually a safe assumption that when a level curve $g(x, y) = c$ is given, the value c is a regular value.

Let's sum up: the vectors $\nabla f(x_0, y_0)$ and $\nabla g(x_0, y_0)$ are both orthogonal to the vector $\vec{r}'(t_0)$. We assume that $\nabla f(x_0, y_0) \neq \vec{0}$, since critical points of f have already been checked. We also assume that c is a **regular value** of g, meaning that there are no critical points of g along the curve $g(x, y) = c$, so $\nabla g(x_0, y_0) \neq \vec{0}$ as well.

But the only way that two non-zero vectors in the plane can both be orthogonal to a third vector is if they're parallel! This means that there must be some scalar λ such that

$$\nabla f(x_0, y_0) = \lambda \nabla g(x_0, y_0).$$

We have demonstrated the following:

Theorem 14.7.3 Location of constrained extrema.

Let f and g be continuously differentiable functions of two variables, and suppose c is a regular value for g. If the function f, when constrained to the level curve $g(x, y) = c$ has a local maximum or minimum value at a point (x_0, y_0), then

$$\nabla f(x_0, y_0) = \lambda \nabla g(x_0, y_0)$$

for some scalar λ.

Note that there are two possibilities: either $\lambda = 0$, in which case (x_0, y_0) is a critical point of f, or $\lambda \neq 0$, in which case the level curve of f that passes through (x_0, y_0) must be *tangent* to the curve $g(x, y) = c$ at that point. Putting Theorem 14.7.3 to use is a matter of solving a system of equations.

Key Idea 14.7.4 Method of Lagrange Multipliers.

To find the maximum and minimum values of a function f of two variables subject to a constraint $g(x, y) = c$, we must find the simultaneous solutions to the following equations, where λ is an unknown constant (called a Lagrange multiplier):

$$\begin{aligned} f_x(x, y) &= \lambda g_x(x, y) \\ f_y(x, y) &= \lambda g_y(x, y) \\ g(x, y) &= c. \end{aligned}$$

Example 14.7.5 Using Lagrange multipliers.

Find the absolute maximum and minimum values of $f(x, y) = x^2 - 8x - 3y^2$ on the disc $x^2 + y^2 \leq 4$.

Solution. This is the same problem as Example 14.7.1, but this time, we will attempt to solve it using the method of Lagrange multipliers. Again, since $\nabla f(x, y) = \langle 2x - 8, -6y \rangle$, the only critical point for f is outside the given disc. It remains to find the maximum and minimum of f subject to the constraint $x^2 + y^2 = 4$, so our constraint function is $g(x, y) = x^2 + y^2$. We have

$$\nabla f(x, y) = \langle 2x - 8, -6y \rangle = \lambda \langle 2x, 2y \rangle = \lambda \nabla g(x, y).$$

Together with the constraint, we have three equations:

$$2x - 8 = 2\lambda x \quad \Rightarrow (1 - \lambda)x = 4$$

$$-6y = 2\lambda y \quad \Rightarrow\ y = 0 \text{ or } \lambda = -3$$
$$x^2 + y^2 = 4.$$

Now we encounter the primary difficulty with Lagrange multipliers. While the idea is simple, the equations it leads to frequently are not. The equations are rarely linear, so there is no systematic method for solving them: solving a Lagrange multiplier problem requires a certain amount of patience and creativity!

One of the possibilities we see above is $y = 0$. If $y = 0$, the constraint equation requires $x = \pm 2$, and in either case we can choose a value for λ (-1 and 3, respectively) that solves the equation $(1 - \lambda)x = 4$.

We find $f(2, 0) = -12$, and $f(-2, 0) = 20$. If $y \neq 0$, then we must have $\lambda = -3$. Putting this into the equation $(1 - \lambda)x = 4$ gives us $4x = 4$, or $x = 1$. If $x = 1$, the constraint equation gives us $1 + y^2 = 4$, so $y = \pm\sqrt{3}$.

We find $f(1, \sqrt{3}) = f(1, -\sqrt{3}) = -16$. There are no other points that satisfy all three equations, so we compare values to complete the problem: the maximum is $f(-2, 0) = 20$, and the minimum is $f(1, \pm\sqrt{3}) = -16$, as before.

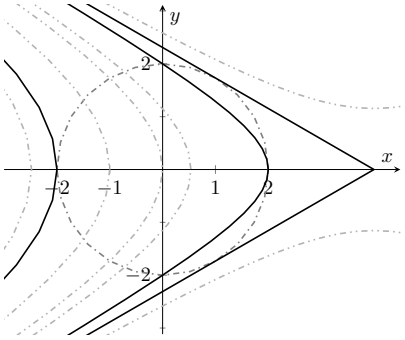


Figure 14.7.7 The constraint curve and several level curves in Example 14.7.5

The method of Lagrange multipliers seems rather arcane at first glance, but it's actually not hard to understand geometrically why it works.

Consider Figure 14.7.7. The constraint curve $x^2 + y^2 = 4$ is the dashed circle. We also see the three level curves (solid) that were obtained as solutions to the Lagrange multiplier equations:

- $f(x, y) = -12$: passing through $(2, 0)$
- $f(x, y) = 20$: passing through $(-2, 0)$
- $f(x, y) = -16$: this curve is actually a pair of lines, $\sqrt{3}y = \pm(x - 4)$, passing through $(1, \pm\sqrt{3})$, respectively.

We see that all three curves are tangent to the constraint curve, as we expect from the requirement that the gradients ∇f and ∇g be parallel where the curves intersect.

Additional level curves $f(x, y) = c$ are plotted as well, with dashed-dotted lines. For values of c with $c > 20$ (greater than the maximum) or $c < -16$ (less than the minimum), the level curve $f(x, y) = c$ does not intersect the constraint curve at all.

For values of c with $-16 < c < 20$, the curve $f(x, y) = c$ intersects the constraint curve, but the intersection is what's called *transversal*: at these points of intersection, the two curves are not tangent, and the gradients are not parallel.

In Figure 14.7.7, you can imagine that increasing or decreasing the value of c has the effect of shifting the level curve one way or the other, until it *just* touches the circle. Any bigger than the maximum, or smaller than the minimum, and the curves no longer intersect. Of course, saying that the curves "just touch" amounts to saying that they are tangent at their point of intersection, just as Theorem 14.7.3 predicts.

youtu.be/watch?v=WgzwyBEDTzA

Figure 14.7.6 Using the method of Lagrange multipliers

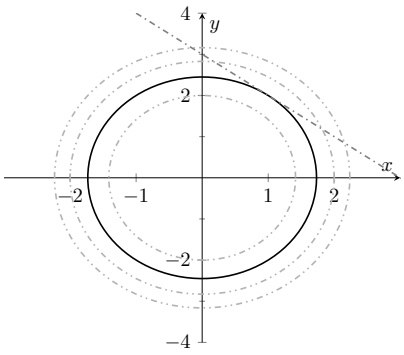


Figure 14.7.9 The constraint and some level curves in Example 14.7.8

Example 14.7.8 Exploring Lagrange Multipliers Geometrically.

Use Lagrange multipliers to locate the extrema of $f(x, y) = 2x^2 + y^2$, subject to the constraint $x + y = 3$.

Solution. Let's see what happens if we dive right in and apply our machinery. With $g(x, y) = x + y$, we need to have

$$\nabla f(x, y) = \langle 4x, 2y\rangle = \lambda \langle 1, 1\rangle = \lambda \nabla g(x, y),$$

so $x+y = 4$, from our constraint, and $4x = \lambda = 2y$, giving us $y = 3-x$ and $y = 2x$, so $2x = 3 - x$, giving $x = 1$, and $y = 2$.

We get only one solution: the value $f(1, 2) = 6$. But is this a maximum or a minimum? And shouldn't we get both?

Rather than blindly attacking the equations, perhaps it would do to take a step back and think about the problem. First, consider the constraint equation: $x+y = 3$. This is a line; it certainly is not the boundary of a closed, bounded reason in the plane. Thus, we haven't satisfied the conditions of the Extreme Value Theorem, and have no reason to expect both an absolute maximum and an absolute minimum.

Now, since the line $x + y = 3$ extends without bound, it's clear that there can be no maximum value c beyond which the ellipse $2x^2+y^2 = c$ does not intersect the line. There is, however, a minimum value: when $c = 6$, the ellipse $2x^2 + y^2 = 6$ has gradient $\nabla f(x, y) = \langle 4, 4\rangle$, giving us the tangent line

$$4(x - 1) + 4(y - 2) = 0, \text{ or } x + y = 3,$$

the equation of our constraint. For value of c less than 3, the ellipse $2x^2 + y^2 = c$ does not intersect the line $x + y = 3$.

The method of Lagrange multipliers is not restricted to functions of two variables or to single constraints. We can similarly determine the extrema of a function f of three variables on a closed bounded subset of $\mathbb{R}^3$.

Example 14.7.10 Determining constrained extrema for a function of three variables.

Determine the maximum and minimum values of the function $f(x, y, z) = x^4 + y^4 + z^4$, subject to the constraint $x^2 + y^2 + z^2 = 1$.

Solution. With $g(x, y, z) = x^2 + y^2 + z^2$, the equation $\nabla f(x, y, z) = \lambda \nabla g(x, y, z)$ gives us

$$\left\langle 4x^3, 4y^3, 4z^3\right\rangle = \lambda \langle 2x, 2y, 2z\rangle .$$

Equating first components, we have $2x^3 = \lambda x$. One possibility is $x = 0$; the other is $\lambda = 2x^2$. Similar results hold for the other two variables, leaving us with several possibilities to consider.

- We take the solution $x = 0$, $y = 0$, and $z = 0$ from the vector equation above. But this result cannot satisfy our constraint, so we rule it out.
- We have $x = 0$ and $y = 0$, but $z \neq 0$. The constraint equation forces $z = \pm 1$. Similarly, we can have $x = 0$, $y = \pm 1$, and $z = 0$,

or $x = \pm 1$, $y = 0$, and $z = 0$. This gives us six points, and they all give the same value for f:

$$f(\pm 1, 0, 0) = f(0, \pm 1, 0) = f(0, 0, \pm 1) = 1.$$

- One of the three variables is zero. If $x = 0$, with y and z nonzero, then we have $2y^2 = \lambda = 2z^2$, and since $x^2 + y^2 + z^2 = 1$, we must have $y^2 = z^2 = \frac{1}{2}$. This gives us $f(x, y, z) = 0 + \frac{1}{4} + \frac{1}{4} = \frac{1}{2}$.

  There are twelve possibilities here: one variable zero, and the other two can be $\pm\frac{1}{\sqrt{2}}$. Each one gives a value of $\frac{1}{2}$ for f.

- Finally, we could have all three variables nonzero. In this case the Lagrange multiplier equations give us

  $$2x^2 = 2y^2 = 2z^2 = \lambda,$$

  and putting these into the constraint equation gives us $x^2 = y^2 = z^2 = \frac{1}{3}$. There are eight different points satisfying this requirement, but all of them give us a value of

  $$f(x, y, z) = \frac{1}{9} + \frac{1}{9} + \frac{1}{9} = \frac{1}{3}.$$

Comparing values, we see that the maximum value for f, when constrained to unit sphere, is 1, and there are 6 points on the sphere with this value. The minimum value is $\frac{1}{3}$, and this occurs at 8 different points.

As the above examples show, Lagrange multiplier problems are often easy to set up, but hard to solve by hand. So why is the method useful? One reason is that it can be used to establish useful theoretical results. But more practically, the method of Lagrange multipliers is useful because it is easy to program into a computer: we simply provide the function and the constraint(s), and the computer solves the resulting equations. There is no need for the same degree of problem-solving employed when we first tackled optimization problems in one variable, back in Chapter 4. To emphasize this, we consider one more example: a reprise of one of the optimization problems from Section 4.3.

For functions of three or more variables, it is also possible to consider two or more constraints. If we wished to optimize a function $f(x, y, z)$ subject to constraints $g(x, y, z) = a$ and $h(x, y, z) = b$, we would have to solve the vector equation

$$\nabla f(x, y, z) = \lambda g(x, y, z) + \mu h(x, y, z),$$

along with the two constraint equations. Problems involving large (or even arbitrary!) numbers of variables and constraints are encountered everywhere from Economics to Quantum Mechanics.

Example 14.7.11 Solving an optimization problem with Lagrange multipliers.

Find the dimensions of a cylindrical can of volume 206 in^3 that minimize the can's surface area.

Solution. This was one of the exercises at the end of Section 4.3. The surface area of a cylinder of radius r and height h is given by

$$s(r, h) = 2r^2 + 2\pi rh.$$

This is the function we wish to minimize, subject to the volume constraint $\pi r^2 h = 206$.

In Section 4.3, our next step would have been to solve the constraint equation for one of the two variables (likely h) in terms of the other, so we could rewrite $s(r, h)$ as a function of one variable and apply the techniques of Section 3.1.

Instead, we introduce the constraint function $v(r, h) = \pi r^2 h$. The

Lagrange multiplier equation $\nabla s = \lambda \nabla v$ gives us

$$\langle 4r + 2\pi h, 2\pi r \rangle = \lambda \left\langle 2\pi r h, \pi r^2 \right\rangle .$$

Equating the second components gives us $2\pi r = \lambda \pi r^2$. Since the constraint ensures that $r \neq 0$, we have $\lambda r = 2$.

Now, we equate the first components:

$$4r + 2\pi h = \lambda \cdot 2\pi r,$$

but $\lambda r = 2$, so we have simply $4r + 2\pi h = 4\pi h$, or $\pi h = 2r$. Putting this into the constraint equation gives us

$$\pi r^2 h = 2r^2 = 206,$$

so $r = \sqrt[3]{103} \approx 4.688$, and $h = 2\sqrt[3]{103}/\pi \approx 2.984$. This is, of course, the same result you would have found if you did this exercise back in Section 4.3.

youtu.be/watch?v=uvQku50u_-0

Figure 14.7.12 Another optimization problem using Lagrange multipliers

14.8 Hessians and the General Second Derivative Test

In Section 14.5, we saw that, just as for functions of a single variable (recall Section 3.1), local extreme values occur at critical points. Definition 14.5.3 defined a critical point (a, b) of a function $f(x, y)$ to be one where the gradient vanishes:

$$\nabla f(a, b) = \langle f_x(a, b), f_y(a, b) \rangle = \langle 0, 0 \rangle .$$

Given a critical point for a function f of two variables, Theorem 14.5.13, the Second Derivative Test, tells us how to determine whether that critical point corresponds to a local minimum, local maximum, or saddle point. You might have been left wondering why the second derivative test looks so different in two variables. You might also have been left wondering what this test looks like if we have three or more variables!

The appearance of the quantity

$$D = f_{xx}(a, b) f_{yy}(a, b) - f_{xy}^2(a, b)$$

seems a bit weird at first, but the idea is actually fairly simple, if you're willing to accept Taylor's Theorem without proof for functions of more than one variable. We already know that if $f(x, y)$ is C^1 (continuously differentiable), then we get the linear approximation

$$f(x, y) \approx f(a, b) + \nabla f(a, b) \cdot \langle x - a, y - b \rangle$$

near a point (a, b) in the domain of f. (Multiplying out the dot product above gives us the differential df defined in Definition 14.1.1.)

Taylor's theorem tells us that if f is C^2 (has continuous second-order derivatives), then we get the *quadratic* approximation

$$\begin{aligned} f(x, y) \approx\ & f(a, b) + \nabla f(a, b) \cdot \langle x - a, y - b \rangle \\ & + \frac{1}{2} A(x - a)^2 + B(x - a)(y - b) + \frac{1}{2} C(y - b)^2, \end{aligned}$$

where $A = \dfrac{\partial^2 f}{\partial x^2}(a, b)$, $B = \dfrac{\partial^2 f}{\partial x \partial y}(a, b)$, and $C = \dfrac{\partial^2 f}{\partial y^2}(a, b)$. (Compare this to the single-variable version: $f(x) \approx f(a) + f'(a)(x - a) + \frac{1}{2} f''(a)(x - a)^2$.)

Now, if (a, b) is a critical point, then $\nabla f(a, b) = \vec{0}$, and we get the approximation

$$f(x, y) \approx k + \frac{1}{2}\left(AX^2 + 2BXY + CY^2\right),$$

where $k = f(a, b), X = x - a, Y = y - b$. So it's enough to understand the critical points of the function

$$g(x, y) = Ax^2 + 2Bxy + Cy^2,$$

since f locally looks just like g. (We've basically just done a shift of the graph, and stretched by a factor of 2 to get rid of the 1/2.)

Now, we can re-write g as follows, assuming $A \neq 0$:

$$\begin{aligned} g(x, y) &= Ax^2 + 2Bxy + Cy^2 \\ &= A\left(x^2 + 2\frac{B}{A}xy\right) + Cy^2 \\ &= A\left(x + \frac{B}{A}y\right)^2 - \frac{B^2}{A}y^2 + Cy^2 \\ &= A\left(x + \frac{B}{A}y\right)^2 + \frac{1}{A}(AC - B^2)y^2. \end{aligned}$$

Now we can see that this is basically just a paraboloid, as long as $AC - B^2 \neq 0$. (Otherwise, we end up with a parabolic cylinder.)

If $AC - B^2 > 0$ (note that this is just the discriminant D!), then both the coefficient for both terms has the same sign; if $A > 0$ we get an elliptic paraboloid opening upwards (local minimum), and if $A < 0$ we get an elliptic paraboloid opening downwards (local maximum). If $AC - B^2 < 0$, then the two terms have coefficients with opposite signs, and that gives us a hyperbolic paraboloid (saddle point).

And what if $A = 0$? Well, in that case $AC - B^2 = -B^2 \leq 0$, so there are two cases: if $B \neq 0$, the second derivative test tells us to expect a saddle point, and indeed this is what we get. Either $C = 0$ as well, and $g(x, y) = 2Bxy$, which is just a hyperbolic paraboloid rotated by $\pi/4$ (its contour curves are the hyperbolas $xy = c$), or $C \neq 0$, in which case you can complete the square in y, and check that the result is once again a hyperbolic paraboloid (exercise).

The other case is if $B = 0$, in which case $D = 0$, so we can't make any conclusions from the second derivative test (although we'll have $g(x, y) = Cy^2$, which is again a parabolic cylinder).

We will now explain how to state second derivative test in general, for functions of n variables, where $n = 1, 2, 3, \ldots$. We will also give an outline of the proof of this result. The proof requires the use of Taylor's theorem for a function of several variables, which we will not prove, and a bit of terminology from linear algebra. Our sketch of the proof follows the exposition given in the text *Vector Calculus*, 4th edition, by Marsden and Tromba.

14.8.1 Taylor Polynomials in Several Variables

Before getting to the general result, let's take a brief detour and discuss Taylor polynomials. One way of thinking about differentiability of a function $f : D \subseteq \mathbb{R}^n \to \mathbb{R}$ is to think of the linearization $L(\vec{x})$ as the degree one Taylor polynomial

$$P_1(\vec{x}) = f(\vec{a}) + \nabla f(\vec{a}) \cdot (\vec{x} - \vec{a}) = f(\vec{a}) + \frac{\partial f}{\partial x_1}(\vec{a})(x_1 - a_1) + \cdots + \frac{\partial f}{\partial x_n}(x_n - a_n).$$

We will use the shorthand $f(\vec{x})$ from Section 14.6 for the function $f(x_1, \ldots, x_n)$, where $\vec{x} = \langle x_1, \ldots, x_n \rangle$. We will also write our vectors using angle bracket notation, even when we should really write them as column vectors for the purposes of matrix multiplication. Finally, as in Section 14.6, for an $n \times n$ matrix A, we will use the dot product $A \cdot \vec{x}$ to account for this.

The requirement of differentiability is then that the remainder $R_1(\vec{x}) = f(\vec{x}) - P_1(\vec{x})$ goes to zero faster than $\|\vec{x} - \vec{a}\|$; that is,

$$\lim_{\vec{x}\to\vec{a}} \frac{R_1(\vec{x})}{\|\vec{x} - \vec{a}\|} = 0.$$

Using the terminology from Section 14.6, we say that f and P_1 "agree to first order". From here we can go on and ask for degree k Taylor polynomials $P_k(\vec{x})$ that give a "kth-order approximation" of f near $\vec{a}$.

In other words, we want a polynomial

$$P_k(x_1, \ldots, x_n) = a_0 + a_1x_1 + \cdots + a_nx_n + a_{11}x_1^2 + a_{12}x_1x_2 + \cdots + a_{nn}x_n^2 + \cdots + a_{1\ldots1}x_1^k + a_{1\ldots12}x_1^{k-1}x_2 + \cdots + a_{n\ldots n}x_n^k,$$

in n variables, of degree k, such that the remainder $R_k(\vec{x}) = f(\vec{x}) - P_k(\vec{x})$ satisfies $R_k(\vec{x}) \approx C\|\vec{x} - \vec{a}\|^l$, with $l > k$. In terms of limits, this means

$$\lim_{\vec{x}\to\vec{a}} \frac{R_k(\vec{x})}{\|\vec{x} - \vec{a}\|^k} = 0.$$

You've probably already noticed a problem with talking about higher-order polynomials in several variables: the notation gets really messy, since there are so many more possible terms! For example, even a relatively simple case like a degree 3 polynomial in 3 variables looks like

$$P(x, y, z) = a + bx + cy + dz + ex^2 + fxy + gxz + hy^2 + kyz + lz^2 + mx^3 + nx^2y + oxy^2 + pxyz + qx^2z + rxz^2 + sy^3 + ty^2z + uyz^2 + vz^3$$

for constants $a, b, c, d, e, f, g, h, i, j, k, l, m, n, o, p, q, r, s, t, u, v$!

Usually we get around this using "multi-index" notation We let $\alpha = (a_1, \ldots, a_n)$ denote a n-tuple of non-negative integers, and then we define $\vec{x}^\alpha = x_1^{a_1}x_2^{a_2}\cdots x_n^{a_n}$, $|\alpha| = a_1 + \cdots + a_n$ (so that $\vec{x}^\alpha$ is a monomial of order $|\alpha|$), and we denote a possible coefficient of $\vec{x}^\alpha$ by a_α. A general k^{th}-order polynomial then looks like

$$P_k(\vec{x}) = \sum_{|\alpha|=0}^{k} a_\alpha x^\alpha.$$

For example, in 3 variables, the terms where $|\alpha| = 3$ would involve $\alpha = (3,0,0), (2,1,0), (2,0,1), (1,2,0), (1,0,2), (0,3,0), (0,2,1), (0,1,2), (0,0,3)$, so in the above polynomial $m = a_{(3,0,0)}$, $n = a_{(2,1,0)}$, etc., with $\vec{x}^{(3,0,0)} = x^3$, $\vec{x}^{(2,1,0)} = x^2y$, and so on. (Note that $\alpha = (0, \ldots, 0)$ is the only multi-index with $|\alpha| = 0$).

With all of that notational unpleasantness out of the way, we can say what the kth-order Taylor polynomial for f near $\vec{a}$ should be: Taylor's Theorem, generalized to n variables, states that

$$P_k(\vec{x}) = \sum_{|\alpha|=0}^{k} \frac{f^{(\alpha)}(\vec{a})}{\alpha!}(\vec{x} - \vec{a})^\alpha,$$

where $\alpha! = a_1!a_2!\cdots a_n!$, and

$$f^{(\alpha)}(\vec{a}) = \left(\frac{\partial^{a_1}}{\partial x_1^{a_1}}\frac{\partial^{a_2}}{\partial x_2^{a_n}}\cdots\frac{\partial^{a_n}}{\partial x_n}f\right)(\vec{a}).$$

As an exercise, check that putting $k = 1$ reproduces the linearization $P_1(\vec{x})$ (note that if $|\alpha| = 1$ we have to have $\alpha = (1, 0, \ldots, 0)$, $(0, 1, 0, \ldots, 0)$, etc.), and that putting $k = 2$ gives the quadratic approximation discussed below.

14.8.2 Quadratic Functions in Several Variables

Let $A = [a_{ij}]$ be an $n \times n$ matrix. We say that A is *symmetric* if $A^T = A$, or equivalently, if $a_{ij} = a_{ji}$ for each i, j between 1 and n. To each such A we can associate a *quadratic function* $q : \mathbb{R}^n \to \mathbb{R}$ given by

$$q(\vec{x}) = \vec{x} \cdot (A \cdot \vec{x}),$$

or in terms of components,

$$q(x_1, \ldots, x_n) = \sum_{i,j=1}^{n} a_{ij} x^i x^j.$$

We say that A is *non-degenerate* if $\det A \neq 0$; this is equivalent to saying that A is invertible, or that $A\vec{x} = \vec{0}$ is possible only if $\vec{x} = \vec{0}$. (Note however that the corresponding property does not hold for q: it is possible to have $q(\vec{x}) = 0$ for $\vec{x} \neq \vec{0}$ even if the corresponding matrix A is non-degenerate. For example, the quadratic function $q(x, y) = x^2 - y^2$ has $q(1, 1) = 0$ and corresponds to the non-degenerate matrix $\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$.

A quadratic function q is called *positive-definite* if $q(\vec{x}) \geq 0$ for all $\vec{x} \in \mathbb{R}^n$, and $q(\vec{x}) = 0$ only for $\vec{x} = \vec{0}$. (Note that the quadratic function $q(x, y) = x^2 - y^2$ given above is *not* positive definite; however, $\tilde{q}(x, y) = x^2 + y^2$ is.) Similarly, q is *negative-definite* if $q(\vec{x}) \leq 0$ for all $\vec{x} \in \mathbb{R}^n$ with $q(\vec{x}) = 0$ for $\vec{x} = \vec{0}$ only.

If $q(\vec{x}) = \vec{x} \cdot A\vec{x}$ is positive(negative)-definite, we refer to the corresponding symmetric matrix A as positive(negative)-definite as well. In general it can be difficult to determine when a given quadratic function (or its corresponding matrix) is positive or negative-definite. In the case of a 2×2 matrix $A = \begin{bmatrix} a & b \\ b & c \end{bmatrix}$ we get

$$\begin{aligned} q(x_1, x_2) &= ax_1^2 + 2bx_1x_2 + cx_2^2 \\ &= a\left(x_1 + \frac{b}{a}x_2\right)^2 + \left(c - \frac{b^2}{a}\right)x_2^2, \end{aligned}$$

by completing the square. Since we must have $q(x_1, 0) > 0$ if $x_1 \neq 0$, we get $a > 0$, and since $q(0, x_2) > 0$ for $x_2 \neq 0$, it follows that $ac - b^2 = \det A > 0$. Similarly q is negative-definite if $a < 0$ and $\det A > 0$.

For an $n \times n$ matrix, one test is as follows: consider the sequence of $j \times j$ matrices A_j, for $j = 1, \ldots, n$, given by $A_1 = [a_{11}]$, $A_2 = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}, \ldots, A_n = A$. (i.e. we take upper-left square sub-matrices of increasing size.) Then A is positive-definite if and only if $\det A_j > 0$ for each $j = 1, \ldots n$, and A is negative-definite if the signs of $\det A_j$ alternate between negative and positive. (So $\det A_1 = a_{11} < 0$, $\det A_2 > 0$, $\det A_3 < 0, \ldots$)

Another approach, which is more illuminating but requires more advanced linear algebra, is to use the fact that for any symmetric matrix A, there exists a change of basis such that A becomes a diagonal matrix $\tilde{A}$ with respect to that basis. (i.e. A can be diagonalized.) If the entries $\tilde{a}_{ii}$ of $\tilde{A}$ along the main diagonal (that is, the **eigenvalues** of A) are all non-zero, then A is non-degenerate. If they are all positive, then A is positive-definite. If they are all negative, then A is negative-definite.

We will need the following lemma below, which is a consequence of the Extreme Value Theorem.

Theorem 14.8.1

If $q : \mathbb{R}^n \to \mathbb{R}$ is a positive-definite quadratic function, then there exists a real number $M > 0$ such that

$$q(\vec{x}) \geq M\|\vec{x}\|^2$$

for any $x \in \mathbb{R}^n$.

To see that this is true, consider $q(\vec{x})$ on the set B of all $\vec{x}$ with $\|\vec{x}\| = 1$. The set B is closed and bounded and q is continuous on B, so by the Extreme Value Theorem, q must attain a minimum value M for some $\vec{a} \in B$. Now, for any constant $c \in \mathbb{R}$, the fact that q is quadratic implies that $q(c\vec{x}) = c^2 q(\vec{x})$. For any non-zero $\vec{x} \in \mathbb{R}^n$, we know that $\dfrac{\vec{x}}{\|\vec{x}\|} \in B$, and thus, we have

$$q(\vec{x}) = q\left(\|\vec{x}\|\frac{\vec{x}}{\|\vec{x}\|}\right) = \|\vec{x}\|^2 q\left(\frac{\vec{x}}{\|\vec{x}\|}\right) \geq M\|\vec{x}\|^2.$$

Finally, if $\vec{x} = \vec{0}$ we get $q(\vec{0}) = 0 = M\|\vec{0}\|^2$.

14.8.3 The Hessian Matrix of a Real-Valued Function

Definition 14.8.2 The Hessian Matrix.

Let $f : \mathbb{R}^n \to \mathbb{R}$ be a function with continuous second-order partial derivatives. We define the *Hessian matrix* of f at a point $\vec{a}$ in the domain of f to be the $n \times n$ symmetric matrix

$$\text{Hess } f(\vec{a}) = \frac{1}{2}\begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2}(\vec{a}) & \frac{\partial^2 f}{\partial x_1 \partial x_2}(\vec{a}) & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_n}(\vec{a}) \\ \frac{\partial^2 f}{\partial x_2 \partial x_1}(\vec{a}) & \frac{\partial^2 f}{\partial x_2^2}(\vec{a}) & \cdots & \frac{\partial^2 f}{\partial x_2 \partial x_n}(\vec{a}) \\ \vdots & \vdots & & \vdots \\ \frac{\partial^2 f}{\partial x_n \partial x_1}(\vec{a}) & \frac{\partial^2 f}{\partial x_n \partial x_2}(\vec{a}) & \cdots & \frac{\partial^2 f}{\partial x_n^2}(\vec{a}) \end{bmatrix}.$$

Note that Hess $f(\vec{a})$ is symmetric by Clairaut's theorem (Theorem 11.3.15. The factor of $1/2$ is included for convenience with respect to Taylor's theorem. Recall that for a function of one variable, the second-order Taylor polynomial of f about $x = a$ is

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

For $\vec{x} \in \mathbb{R}^n$, let us define the quadratic function $h_{f,\vec{a}}(\vec{x}) = \vec{x} \cdot (\text{Hess } f(\vec{a}) \cdot \vec{x})$ associated to the Hessian of f at $\vec{a}$. Taylor's theorem for functions of several variables tells us that if all the *third* derivatives of f are continuous, then near $\vec{a} \in \mathbb{R}^n$ we have

$$f(\vec{x}) = f(\vec{a}) + \nabla f(\vec{a}) \cdot (\vec{x} - \vec{a}) + h_{f,\vec{a}}(\vec{x} - \vec{a}) + R(\vec{a}, \vec{x}), \tag{14.8.1}$$

where the remainder term $R(\vec{a}, \vec{x})$ satisfies

$$\lim_{\vec{x} \to \vec{a}} \frac{R(\vec{x}, \vec{a})}{\|\vec{x} - \vec{a}\|^2} = 0. \tag{14.8.2}$$

Finally, let us define a critical point $\vec{a}$ for f to be *non-degenerate* if Hess $f(\vec{a})$ is non-degenerate. Now we're ready to state our result on the second derivative test.

Theorem 14.8.3 The General Second Derivative Test.

Let $f : \mathbb{R}^n \to \mathbb{R}$ be three times continuously differentiable, and suppose that f has a non-degenerate critical point at $\vec{a}$. If Hess $f(\vec{a})$ *is positive-definite, then $\vec{a}$ is a local minimum for f. Similarly, if* Hess $f(\vec{a})$ *is negative-definite, then $\vec{a}$ is a local maximum for f.*

The way to think about this intuitively is the following: the matrix Hess $f(\vec{a})$ is symmetric. We know from Linear Algebra that every symmetric matrix can be diagonalized. Less obvious (but still true) is that we can make a (linear) change of variables $(u_1, \ldots, u_n) = T(x_1, \ldots, x_n)$ so that the vectors in the direction of the u_i coordinate axes are eigenvectors for Hess $f(\vec{a})$. Slightly harder to show (but also true) is that this change of variables can be chosen so that it is **orthogonal**. That is, we simply have to *rotate* our coordinate system: lengths and angles are all preserved.

In this new coordinate system, the Hessian matrix is diagonal:

$$\text{Hess } f(\vec{a}) = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix}.$$

If each of the eigenvalues $\lambda_1, \ldots, \lambda_n$ is positive, the Hessian is positive-definite, and our critical point is a local minimum. If all the eigenvalues are negative, our critical point is a local maximum. If some of the eigenvalues are positive and some are negative, we have a saddle point.

Proving the result is somewhat more technical. Suppose $\vec{a}$ is a critical point for f, and that Hess $f(\vec{a})$ is positive definite. We know that $\nabla f(\vec{a}) = 0$ at a critical point, so from (14.8.1) we get

$$f(\vec{x}) - f(\vec{a}) = h_{f,\vec{a}}(\vec{x} - \vec{a}) + R(\vec{a}, \vec{x}).$$

Theorem 14.8.1 tells us that $h_{f,\vec{a}}(\vec{x} - \vec{a}) \geq M\|\vec{x} - \vec{a}\|^2$ for some M, and by (14.8.2), there exists a $\delta > 0$ such that whenever $0 < \|\vec{x} - \vec{a}\| < \delta$, we get $|R(\vec{a}, \vec{x})| < M\|\vec{x} - \vec{a}\|^2$. (Take $\epsilon = M$ in the definition of the limit.)

If we carefully put all this together, we can show that

$$h_{f,\vec{a}}(\vec{x} - \vec{a}) + R(\vec{a}, \vec{x}) > 0,$$

since

$$h_{f,\vec{a}}(\vec{x} - \vec{a}) \geq M\|\vec{x} - \vec{a}\|^2 > |R(\vec{a}, \vec{x})|.$$

Substituting this into the above equation, we get $f(\vec{x}) - f(\vec{a}) > 0$ for any $\vec{x}$ with $0 < \|\vec{x} - \vec{a}\| < \delta$, and thus f has a local minimum at $\vec{a} \in \mathbb{R}^n$. The case of a local maximum can be handled similarly (or by replacing f with $-f$).

We in fact do slightly better than a local minimum: we get the *strict* inequality $f(\vec{x}) > f(\vec{a})$, and not just $f(\vec{x}) \geq f(\vec{a})$. Often this fact is expressed by saying that non-degenerate critical points are *isolated* — if f has a critical point at $\vec{a}$, then there is some neighbourhood of $\vec{a}$ in which f has no other critical points. (We have only established this for local maxima and minima, but this fact is true for non-degenerate critical points in general.) This observation is the starting point for an important area of differential topology known as Morse Theory.

Chapter 15

Multiple Integration

Chapters 11–14 introduced multivariable functions and we applied concepts of differential calculus to these functions. We learned how we can view a function of two variables as a surface in space, and learned how partial derivatives convey information about how the surface is changing in any direction.

In this chapter we apply techniques of integral calculus to multivariable functions. In Chapter 5 we learned how the definite integral of a single variable function gave us "area under the curve." In this chapter we will see that integration applied to a multivariable function gives us "volume under a surface." And just as we learned applications of integration beyond finding areas, we will find applications of integration in this chapter beyond finding volume.

15.1 Iterated Integrals and Area

In Section 11.3 we found that it was useful to differentiate functions of several variables with respect to one variable, while treating all the other variables as constants or coefficients. We can integrate functions of several variables in a similar way. For instance, if we are told that $f_x(x, y) = 2xy$, we can treat y as staying constant and integrate to obtain $f(x, y)$:

$$\begin{aligned} f(x,y) &= \int f_x(x,y)\,dx \\ &= \int 2xy\,dx \\ &= x^2y + C. \end{aligned}$$

Make a careful note about the constant of integration, C. This "constant" is something with a derivative of 0 with respect to x, so it could be any expression that contains only constants and functions of y. For instance, if $f(x, y) = x^2y + \sin(y) + y^3 + 17$, then $f_x(x, y) = 2xy$. To signify that C is actually a function of y, we write:

$$f(x,y) = \int f_x(x,y)\,dx = x^2y + C(y).$$

Using this process we can even evaluate definite integrals.

15.1.1 Iterated integrals

Example 15.1.1 Integrating functions of more than one variable.

Evaluate the integral $\displaystyle\int_1^{2y} 2xy\,dx$.

Solution. We find the indefinite integral as before, then apply the Fundamental Theorem of Calculus to evaluate the definite integral:

$$\begin{aligned}\int_1^{2y} 2xy\,dx &= x^2y\Big|_1^{2y}\\ &= (2y)^2y - (1)^2y\\ &= 4y^3 - y.\end{aligned}$$

We can also integrate with respect to y. In general,

$$\int_{h_1(y)}^{h_2(y)} f_x(x,y)\,dx = f(x,y)\Big|_{h_1(y)}^{h_2(y)} = f\big(h_2(y),y\big) - f\big(h_1(y),y\big),$$

and

$$\int_{g_1(x)}^{g_2(x)} f_y(x,y)\,dy = f(x,y)\Big|_{g_1(x)}^{g_2(x)} = f\big(x,g_2(x)\big) - f\big(x,g_1(x)\big).$$

Note that when integrating with respect to x, the bounds are functions of y (of the form $x = h_1(y)$ and $x = h_2(y)$) and the final result is also a function of y. When integrating with respect to y, the bounds are functions of x (of the form $y = g_1(x)$ and $y = g_2(x)$) and the final result is a function of x. Another example will help us understand this.

Example 15.1.2 Integrating functions of more than one variable.

Evaluate $\displaystyle\int_1^{x} \left(5x^3y^{-3} + 6y^2\right) dy$.

Solution. We consider x as staying constant and integrate with respect to y:

$$\begin{aligned}\int_1^{x} \left(5x^3y^{-3} + 6y^2\right) dy &= \left(\frac{5x^3y^{-2}}{-2} + \frac{6y^3}{3}\right)\Bigg|_1^{x}\\ &= \left(-\frac{5}{2}x^3x^{-2} + 2x^3\right) - \left(-\frac{5}{2}x^3 + 2\right)\\ &= \frac{9}{2}x^3 - \frac{5}{2}x - 2.\end{aligned}$$

Note how the bounds of the integral are from $y = 1$ to $y = x$ and that the final answer is a function of x.

In the previous example, we integrated a function with respect to y and ended up with a function of x. We can integrate this as well. This process is known as *iterated integration*, or *multiple integration*.

Example 15.1.3 Evaluating an integral.

Evaluate $\displaystyle\int_1^2 \left(\int_1^x \left(5x^3y^{-3} + 6y^2\right) dy \right) dx$.

Solution. We follow a standard "order of operations" and perform the operations inside parentheses first (which is the integral evaluated in Example 15.1.2.)

$$\begin{aligned}
\int_1^2 \left(\int_1^x \left(5x^3y^{-3} + 6y^2\right) dy \right) dx &= \int_1^2 \left(\left[\frac{5x^3y^{-2}}{-2} + \frac{6y^3}{3} \right] \Bigg|_1^x \right) dx \\
&= \int_1^2 \left(\frac{9}{2}x^3 - \frac{5}{2}x - 2 \right) dx \\
&= \left(\frac{9}{8}x^4 - \frac{5}{4}x^2 - 2x \right) \Bigg|_1^2 \\
&= \frac{89}{8}.
\end{aligned}$$

Note how the bounds of x were $x = 1$ to $x = 2$ and the final result was a number.

The previous example showed how we could perform something called an iterated integral; we do not yet know *why* we would be interested in doing so nor what the result, such as the number $89/8$, *means*. Before we investigate these questions, we offer some definitions.

Definition 15.1.4 Iterated Integration.

Iterated integration is the process of repeatedly integrating the results of previous integrations. Evaulating one integral is denoted as follows.

Let a, b, c and d be numbers and let $g_1(x)$, $g_2(x)$, $h_1(y)$ and $h_2(y)$ be functions of x and y, respectively. Then:

1. $\displaystyle\int_c^d \int_{h_1(y)}^{h_2(y)} f(x,y)\,dx\,dy = \int_c^d \left(\int_{h_1(y)}^{h_2(y)} f(x,y)\,dx \right) dy.$
2. $\displaystyle\int_a^b \int_{g_1(x)}^{g_2(x)} f(x,y)\,dy\,dx = \int_a^b \left(\int_{g_1(x)}^{g_2(x)} f(x,y)\,dy \right) dx.$

Again make note of the bounds of these iterated integrals.

With $\displaystyle\int_c^d \int_{h_1(y)}^{h_2(y)} f(x,y)\,dx\,dy$, x varies from $h_1(y)$ to $h_2(y)$, whereas y varies from c to d. That is, the bounds of x are *curves*, the curves $x = h_1(y)$ and $x = h_2(y)$, whereas the bounds of y are *constants*, $y = c$ and $y = d$. It is useful to remember that when setting up and evaluating such iterated integrals, we integrate "from curve to curve, then from point to point."

We now begin to investigate *why* we are interested in iterated integrals and *what* they mean.

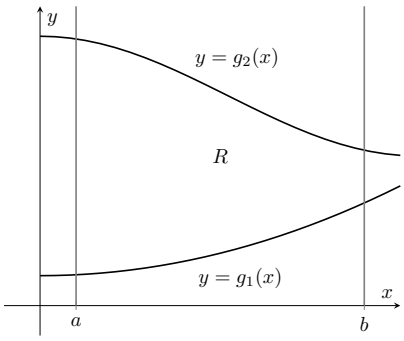


Figure 15.1.5 Calculating the area of a plane region R with an iterated integral

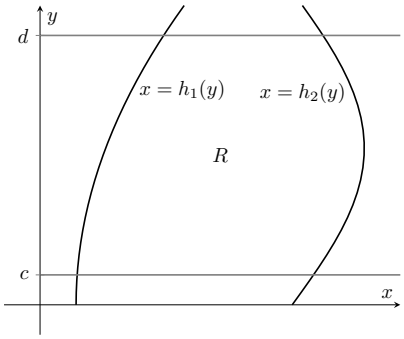


Figure 15.1.6 Calculating the area of a plane region R with an iterated integral

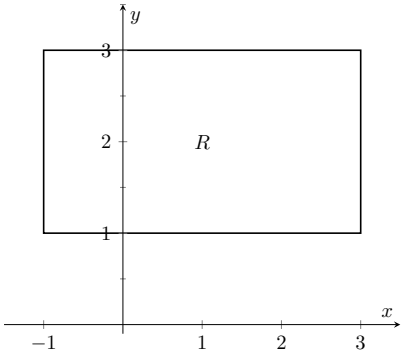


Figure 15.1.9 Calculating the area of a rectangle with an iterated integral in Example 15.1.8

15.1.2 Area of a plane region

Consider the plane region R bounded by $a \le x \le b$ and $g_1(x) \le y \le g_2(x)$, shown in Figure 15.1.5. We learned in Section 7.1 that the area of R is given by

$$\int_a^b \big(g_2(x) - g_1(x)\big)\,dx.$$

We can view the expression $\big(g_2(x) - g_1(x)\big)$ as

$$\big(g_2(x) - g_1(x)\big) = \int_{g_1(x)}^{g_2(x)} 1\,dy = \int_{g_1(x)}^{g_2(x)} dy,$$

meaning we can express the area of R as an iterated integral:

$$\text{area of } R = \int_a^b \big(g_2(x) - g_1(x)\big)\,dx = \int_a^b \left(\int_{g_1(x)}^{g_2(x)} dy\right) dx = \int_a^b \int_{g_1(x)}^{g_2(x)} dy\,dx.$$

In short: a certain iterated integral can be viewed as giving the area of a plane region.

A region R could also be defined by $c \le y \le d$ and $h_1(y) \le x \le h_2(y)$, as shown in Figure 15.1.6. Using a process similar to that above, we have

$$\text{the area of } R = \int_c^d \int_{h_1(y)}^{h_2(y)} dx\,dy.$$

We state this formally in a theorem.

Theorem 15.1.7 Area of a plane region.

1. *Let R be a plane region bounded by $a \le x \le b$ and $g_1(x) \le y \le g_2(x)$, where g_1 and g_2 are continuous functions on $[a, b]$. The area A of R is*
$$A = \int_a^b \int_{g_1(x)}^{g_2(x)} dy\,dx.$$
2. *Let R be a plane region bounded by $c \le y \le d$ and $h_1(y) \le x \le h_2(y)$, where h_1 and h_2 are continuous functions on $[c, d]$. The area A of R is*
$$A = \int_c^d \int_{h_1(y)}^{h_2(y)} dx\,dy.$$

The following examples should help us understand this theorem.

Example 15.1.8 Area of a rectangle.

Find the area A of the rectangle with corners $(-1, 1)$ and $(3, 3)$, as shown in Figure 15.1.9.

Solution. Multiple integration is obviously overkill in this situation, but we proceed to establish its use.

The region R is bounded by $x = -1$, $x = 3$, $y = 1$ and $y = 3$. Choosing to integrate with respect to y first, we have

$$A = \int_{-1}^{3}\int_{1}^{3} 1\,dy\,dx = \int_{-1}^{3} \left(y\Big|_1^3\right) dx = \int_{-1}^{3} 2\,dx = 2x\Big|_{-1}^{3} = 8.$$

We could also integrate with respect to x first, giving:

$$A = \int_1^3 \int_{-1}^3 1\,dx\,dy = \int_1^3 \left(x\Big|_{-1}^3 \right) dy = \int_1^3 4\,dy = 4y\Big|_1^3 = 8.$$

Clearly there are simpler ways to find this area, but it is interesting to note that this method works.

Example 15.1.10 Area of a triangle.

Find the area A of the triangle with vertices at $(1,1)$, $(3,1)$ and $(5,5)$, as shown in Figure 15.1.11.

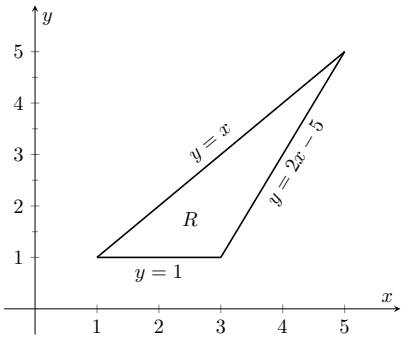


Figure 15.1.11 Calculating the area of a triangle with iterated integrals in Example 15.1.10

Solution. The triangle is bounded by the lines as shown in the figure. Choosing to integrate with respect to x first gives that x is bounded by $x = y$ to $x = \frac{y+5}{2}$, while y is bounded by $y = 1$ to $y = 5$. (Recall that since x-values increase from left to right, the leftmost curve, $x = y$, is the lower bound and the rightmost curve, $x = (y+5)/2$, is the upper bound.) The area is

$$\begin{aligned} A &= \int_1^5 \int_y^{\frac{y+5}{2}} dx\,dy \\ &= \int_1^5 \left(x\Big|_y^{\frac{y+5}{2}} \right) dy \\ &= \int_1^5 \left(-\frac{1}{2}y + \frac{5}{2} \right) dy \\ &= \left(-\frac{1}{4}y^2 + \frac{5}{2}y \right)\Big|_1^5 \\ &= 4. \end{aligned}$$

We can also find the area by integrating with respect to y first. In this situation, though, we have two functions that act as the lower bound for the region R, $y = 1$ and $y = 2x - 5$. This requires us to use two iterated integrals. Note how the x-bounds are different for each integral:

$$\begin{aligned} A &= \int_1^3 \int_1^x 1\,dy\,dx & + & \int_3^5 \int_{2x-5}^x 1\,dy\,dx \\ &= \int_1^3 \big(y\big)\Big|_1^x\,dx & + & \int_3^5 \big(y\big)\Big|_{2x-5}^x\,dx \\ &= \int_1^3 \big(x-1\big)\,dx & + & \int_3^5 \big(-x+5\big)\,dx \\ &= 2 & + & 2 \\ &= 4. \end{aligned}$$

As expected, we get the same answer both ways.

Example 15.1.12 Area of a plane region.

Find the area of the region enclosed by $y = 2x$ and $y = x^2$, as shown in Figure 15.1.13.

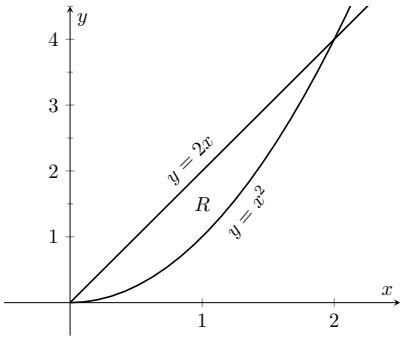


Figure 15.1.13 Calculating the area of a plane region with iterated integrals in Example 15.1.12

Solution. Once again we'll find the area of the region using both orders of integration.

Using $dy\,dx$:

$$\int_0^2 \int_{x^2}^{2x} 1\,dy\,dx = \int_0^2 (2x - x^2)\,dx = \left(x^2 - \frac{1}{3}x^3\right)\Big|_0^2 = \frac{4}{3}.$$

Using $dx\,dy$:

$$\int_0^4 \int_{y/2}^{\sqrt{y}} 1\,dx\,dy = \int_0^4 (\sqrt{y} - y/2)\,dy = \left(\frac{2}{3}y^{3/2} - \frac{1}{4}y^2\right)\Big|_0^4 = \frac{4}{3}.$$

15.1.3 Changing Order of Integration

In each of the previous examples, we have been given a region R and found the bounds needed to find the area of R using both orders of integration. We integrated using both orders of integration to demonstrate their equality.

We now approach the skill of describing a region using both orders of integration from a different perspective. Instead of starting with a region and creating iterated integrals, we will start with an iterated integral and rewrite it in the other integration order. To do so, we'll need to understand the region over which we are integrating.

The simplest of all cases is when both integrals are bound by constants. The region described by these bounds is a rectangle (see Example 15.1.8), and so:

$$\int_a^b \int_c^d 1\,dy\,dx = \int_c^d \int_a^b 1\,dx\,dy.$$

When the inner integral's bounds are not constants, it is generally very useful to sketch the bounds to determine what the region we are integrating over looks like. From the sketch we can then rewrite the integral with the other order of integration.

Examples will help us develop this skill.

Example 15.1.14 Changing the order of integration.

Rewrite the iterated integral $\displaystyle\int_0^6 \int_0^{x/3} 1\,dy\,dx$ with the order of integration $dx\,dy$.

Solution. We need to use the bounds of integration to determine the region we are integrating over.

The bounds tell us that y is bounded by 0 and $x/3$; x is bounded by 0 and 6. We plot these four curves: $y = 0$, $y = x/3$, $x = 0$ and $x = 6$ to find the region described by the bounds. Figure 15.1.15 shows these curves, indicating that R is a triangle.

To change the order of integration, we need to consider the curves that bound the x-values. We see that the lower bound is $x = 3y$ and the upper bound is $x = 6$. The bounds on y are 0 to 2. Thus we can rewrite the integral as $\displaystyle\int_0^2 \int_{3y}^6 1\,dx\,dy$.

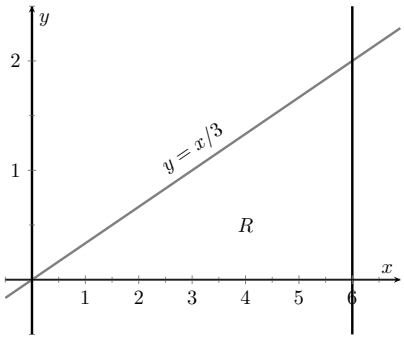


Figure 15.1.15 Sketching the region R described by the iterated integral in Example 15.1.14

Example 15.1.16 Changing the order of integration.

Change the order of integration of $\displaystyle\int_0^4 \int_{y^2/4}^{(y+4)/2} 1\,dx\,dy$.

Solution. We sketch the region described by the bounds to help us change the integration order. x is bounded below and above (i.e., to the left and right) by $x = y^2/4$ and $x = (y+4)/2$ respectively, and y is bounded between 0 and 4. Graphing the previous curves, we find the region R to be that shown in Figure 15.1.17.

To change the order of integration, we need to establish curves that bound y. The figure makes it clear that there are two lower bounds for y: $y = 0$ on $0 \le x \le 2$, and $y = 2x - 4$ on $2 \le x \le 4$. Thus we need two double integrals. The upper bound for each is $y = 2\sqrt{x}$. Thus we have

$$\int_0^4 \int_{y^2/4}^{(y+4)/2} 1\,dx\,dy = \int_0^2 \int_0^{2\sqrt{x}} 1\,dy\,dx + \int_2^4 \int_{2x-4}^{2\sqrt{x}} 1\,dy\,dx.$$

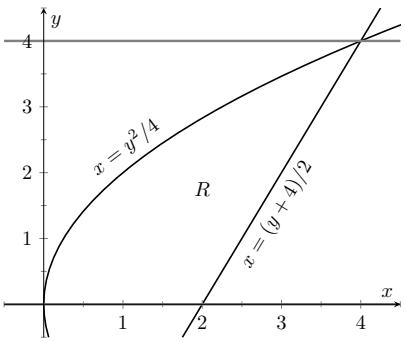


Figure 15.1.17 Drawing the region determined by the bounds of integration in Example 15.1.16

This section has introduced a new concept, the iterated integral. We developed one application for iterated integration: area between curves. However, this is not new, for we already know how to find areas bounded by curves.

In the next section we apply iterated integration to solve problems we currently do not know how to handle. The "real" goal of this section was not to learn a new way of computing area. Rather, our goal was to learn how to define a region in the plane using the bounds of an iterated integral. That skill is very important in the following sections.

15.1.4 Exercises

Terms and Concepts

1. When integrating $f_x(x, y)$ with respect to x, the constant of integration C is really which: $C(x)$ or $C(y)$? What does this mean?

2. Evaluating a double integral in steps is called _____ _____.

3. When evaluating an iterated integral, we integrate from _____ to _____, then from _____ to _____.

4. One understanding of an iterated integral is that $\displaystyle\int_a^b \int_{g_1(x)}^{g_2(x)} dy\,dx$ gives the _____ of a plane region.

Problems

Exercise Group. In the following exercises, evaluate the integral and subsequent iterated integral.

5.

(a) $\displaystyle\int_2^5 \left(6x^2 + 4xy - 3y^2\right) dy$

(b) $\displaystyle\int_{-3}^{-2} \int_2^5 \left(6x^2 + 4xy - 3y^2\right) dy\,dx$

6.

(a) $\displaystyle\int_0^{\pi} (2x\cos(y) + \sin(x))\,dx$

(b) $\displaystyle\int_0^{\pi/2} \int_0^{\pi} (2x\cos(y) + \sin(x))\,dx\,dy$

7.

(a) $\displaystyle\int_1^x \left(x^2y - y + 2\right) dy$

(b) $\displaystyle\int_0^2 \int_1^x \left(x^2y - y + 2\right) dy\,dx$

8.

(a) $\displaystyle\int_y^{y^2} (x - y)\,dx$

(b) $\displaystyle\int_{-1}^1 \int_y^{y^2} (x - y)\,dx\,dy$

9.

(a) $\displaystyle\int_0^y \left(\cos(x)\sin(y)\right) dx$

(b) $\displaystyle\int_0^{\pi} \int_0^y \left(\cos(x)\sin(y)\right) dx\,dy$

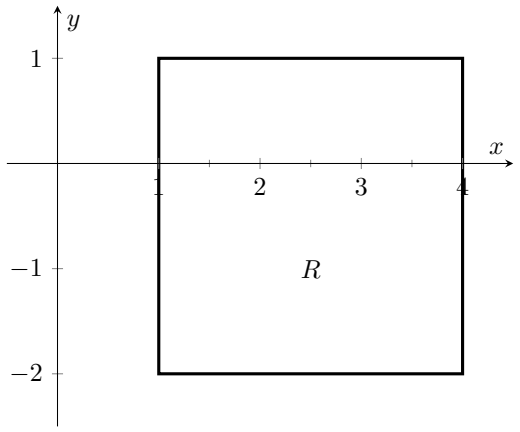
10.

(a) $\displaystyle\int_0^x \left(\frac{1}{1+x^2}\right) dy$

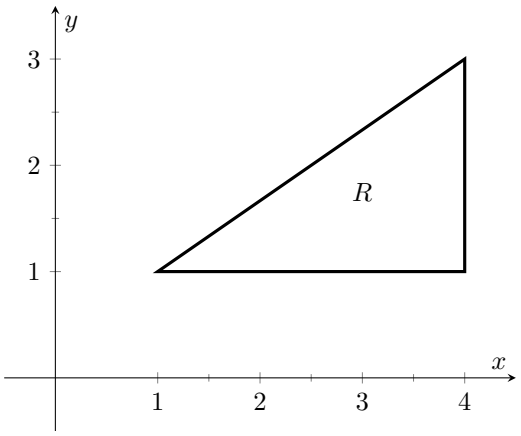
(b) $\displaystyle\int_1^2 \int_0^x \left(\frac{1}{1+x^2}\right) dy\,dx$

Exercise Group. In the following exercises, a graph of a planar region R is given. Give the iterated integrals, with both orders of integration $dy\,dx$ and $dx\,dy$, that give the area of R. Evaluate one of the iterated integrals to find the area.

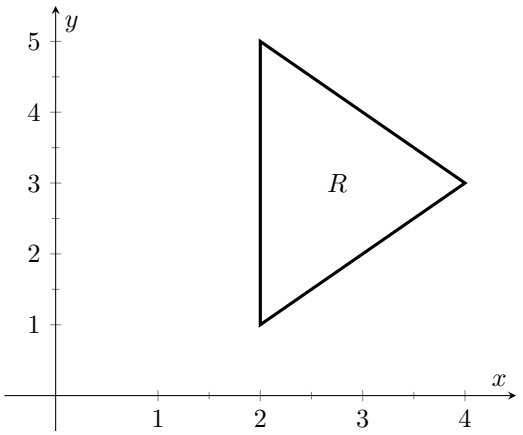
11.



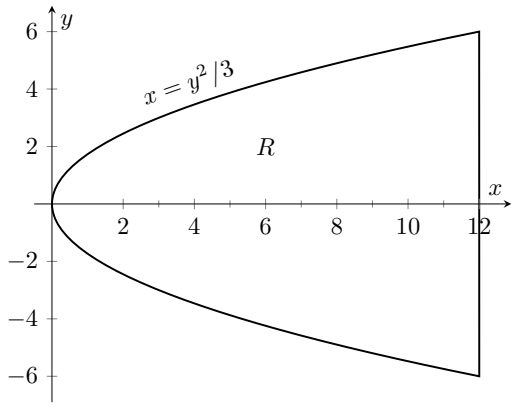
12.



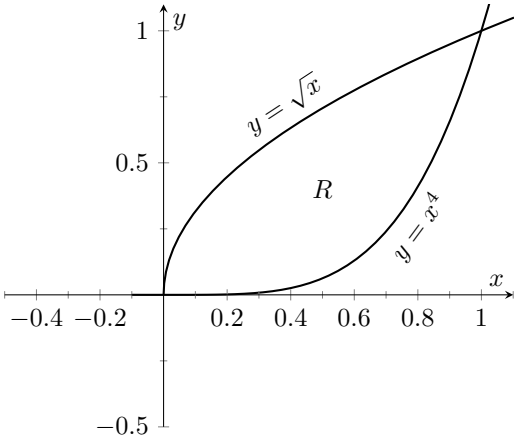
13.



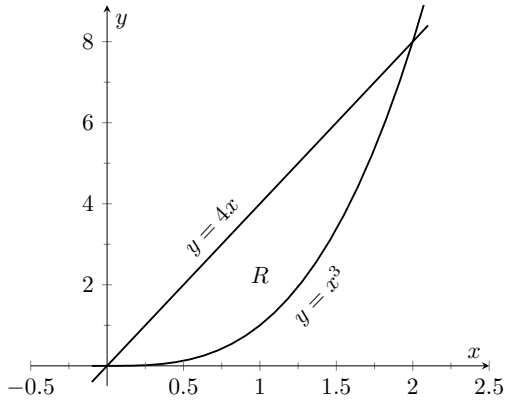
14.



15.



16.



Exercise Group. In the following exercises, iterated integrals are given that compute the area of a region R in the xy-plane. Sketch the region R, and give the iterated integral(s) that give the area of R with the opposite order of integration.

17. $\displaystyle\int_{-2}^{2}\int_{0}^{4-x^2} dy\,dx$

18. $\displaystyle\int_{0}^{1}\int_{5-5x}^{5-5x^2} dy\,dx$

19. $\displaystyle\int_{-2}^{2}\int_{0}^{2\sqrt{4-y^2}} dx\,dy$

20. $\displaystyle\int_{-3}^{3}\int_{-\sqrt{9-x^2}}^{\sqrt{9-x^2}} dy\,dx$

21. $\displaystyle\int_{0}^{1}\int_{-\sqrt{y}}^{\sqrt{y}} dx\,dy + \int_{1}^{4}\int_{y-2}^{\sqrt{y}} dx\,dy$

22. $\displaystyle\int_{-1}^{1}\int_{(x-1)/2}^{(1-x)/2} dy\,dx$

15.2 Double Integration and Volume

The definite integral of f over $[a, b]$, $\int_a^b f(x)\,dx$, was introduced as "the signed area under the curve." We approximated the value of this area by first subdividing $[a, b]$ into n subintervals, where the ith subinterval has length Δx_i, and letting c_i be any value in the ith subinterval. We formed rectangles that approximated part of the region under the curve with width Δx_i, height $f(c_i)$, and hence with area $f(c_i)\Delta x_i$. Summing all the rectangle's areas gave an approximation of the definite integral, and Theorem 5.3.26 stated that

$$\int_a^b f(x)\,dx = \lim_{\|\Delta x\|\to 0} \sum f(c_i)\Delta x_i,$$

connecting the area under the curve with sums of the areas of rectangles.

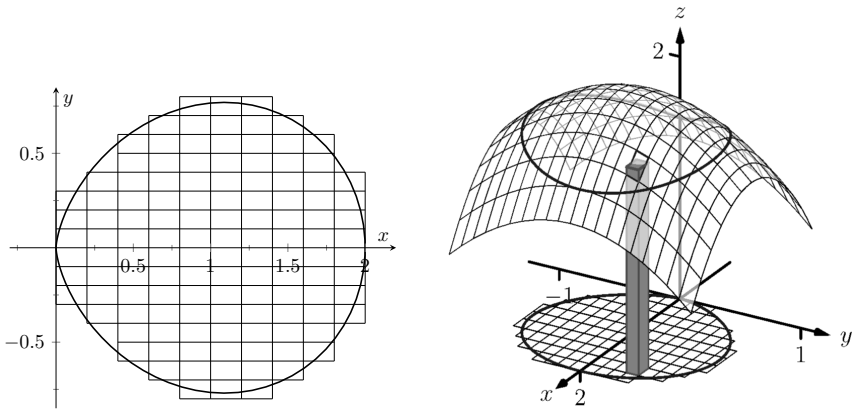
We use a similar approach in this section to find volume under a surface.

Let R be a closed, bounded region in the xy-plane and let $z = f(x, y)$ be a continuous function defined on R. We wish to find the signed volume under the graph of f over R. (We use the term "signed volume" to denote that space above the xy-plane, under f, will have a positive volume; space above f and under the xy-plane will have a "negative" volume, similar to the notion of signed area used before.)

youtu.be/watch?v=5FpKcnG-5vY

Figure 15.2.1 Introducing the double integral

We start by partitioning R into n rectangular subregions as shown in Figure 15.2.2(a). For simplicity's sake, we let all widths be Δx and all heights be Δy. Note that the sum of the areas of the rectangles is not equal to the area of R, but rather is a close approximation. Arbitrarily number the rectangles 1 through n, and pick a point (x_i, y_i) in the ith subregion.



(a) **(b)**

Figure 15.2.2 Developing a method for finding signed volume under a surface

The volume of the rectangular solid whose base is the ith subregion and whose height is $f(x_i, y_i)$ is $V_i = f(x_i, y_i)\Delta x\Delta y$. Such a solid is shown in Figure 15.2.2(b). Note how this rectangular solid only approximates the true volume under the surface; part of the solid is above the surface and part is below.

For each subregion R_i used to approximate R, create the rectangular solid with base area $\Delta x\Delta y$ and height $f(x_i, y_i)$. The sum of all rectangular solids is

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

This approximates the signed volume under f over R. As we have done before, to get a better approximation we can use more rectangles to approximate the region R.

In general, each rectangle could have a different width Δx_j and height Δy_k, giving the ith rectangle an area $\Delta A_i = \Delta x_j \Delta y_k$ and the ith rectangular solid a volume of $f(x_i, y_i)\Delta A_i$. Let $\|\Delta A\|$ denote the length of the longest diagonal of all rectangles in the subdivision of R; $\|\Delta A\| \to 0$ means each rectangle's width and height are both approaching 0. If f is a continuous function, as $\|\Delta A\|$ shrinks (and hence $n \to \infty$) the summation $\sum_{i=1}^{n} f(x_i, y_i)\Delta A_i$ approximates the signed volume better and better. This leads to a definition.

Definition 15.2.3 Double Integral, Signed Volume.

Let $z = f(x, y)$ be a continuous function defined over a closed, bounded region R in the xy-plane. The **signed volume** V under f over R is denoted by the **double integral**

$$V = \iint_R f(x, y)\, dA.$$

Alternate notations for the double integral are

$$\iint_R f(x, y)\, dA = \iint_R f(x, y)\, dx\, dy = \iint_R f(x, y)\, dy\, dx.$$

Definition 15.2.3 does not state how to find the signed volume, though the notation offers a hint. We need the next two theorems to evaluate double integrals to find volume.

Theorem 15.2.5 Double Integrals and Signed Volume.

Let $z = f(x, y)$ be a continuous function defined over a closed , bounded region R in the xy-plane. Then the signed volume V under f over R is

$$V = \iint_R f(x, y)\, dA = \lim_{\|\Delta A\| \to 0} \sum_{i=1}^{n} f(x_i, y_i)\Delta A_i.$$

This theorem states that we can find the exact signed volume using a limit of sums. The partition of the region R is not specified, so any partitioning where the diagonal of each rectangle shrinks to 0 results in the same answer.

This does not offer a very satisfying way of computing volume, though. Our experience has shown that evaluating the limits of sums can be tedious. We seek a more direct method.

Recall Theorem 7.2.3 in Section 7.2. This stated that if $A(x)$ gives the cross-sectional area of a solid at x, then $\int_a^b A(x)\, dx$ gave the volume of that solid over $[a, b]$.

Consider Figure 15.2.6, where a surface $z = f(x, y)$ is drawn over a region R. Fixing a particular x value, we can consider the area under f over R where x has that fixed value. That area can be found with a definite integral, namely

$$A(x) = \int_{g_1(x)}^{g_2(x)} f(x, y)\, dy.$$

Remember that though the integrand contains x, we are viewing x as fixed. Also note that the bounds of integration are functions of x: the bounds depend on the value of x.

Double integrals as limits of double sums. Recall that the integration symbol "$\int$" is an "elongated S," representing the word "sum." We interpreted $\int_a^b f(x)\, dx$ as "take the *sum* of the areas of rectangles over the interval $[a, b]$." The double integral uses two integration symbols to represent a "double sum." When adding up the volumes of rectangular solids over a partition of a region R, as done in Figure 15.2.2, one could first add up the volumes across each row (one type of sum), then add these totals together (another sum), as in

$$\sum_{j=1}^{n} \sum_{i=1}^{m} f(x_i, y_j)\Delta x_i \Delta y_j.$$

One can rewrite this as

$$\sum_{j=1}^{n} \left(\sum_{i=1}^{m} f(x_i, y_j)\Delta x_i \right) \Delta y_j.$$

The summation inside the parenthesis indicates the sum of heights × widths, which gives an area; multiplying these areas by the thickness Δy_j gives a volume. The illustration in Figure 15.2.6 relates to this understanding.

youtu.be/watch?v=l7hGaXcsq9g

Figure 15.2.4 Defining the double integral

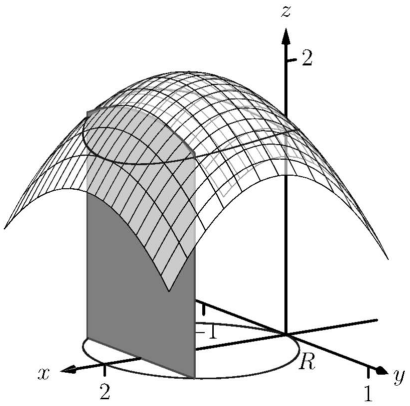


Figure 15.2.6 Finding volume under a surface by sweeping out a cross-sectional area

As $A(x)$ is a cross-sectional area function, we can find the signed volume V under f by integrating it:

$$V = \int_a^b A(x)\,dx = \int_a^b \left(\int_{g_1(x)}^{g_2(x)} f(x,y)\,dy \right) dx = \int_a^b \int_{g_1(x)}^{g_2(x)} f(x,y)\,dy\,dx.$$

This gives a concrete method for finding signed volume under a surface. We could do a similar procedure where we started with y fixed, resulting in an iterated integral with the order of integration $dx\,dy$. The following theorem states that both methods give the same result, which is the value of the double integral. It is such an important theorem it has a name associated with it.

Theorem 15.2.7 Fubini's Theorem.

Let R be a closed, bounded region in the xy-plane and let $z = f(x,y)$ be a continuous function on R.

1. *If R is bounded by $a \le x \le b$ and $g_1(x) \le y \le g_2(x)$, where g_1 and g_2 are continuous functions on $[a,b]$, then*

$$\iint_R f(x,y)\,dA = \int_a^b \int_{g_1(x)}^{g_2(x)} f(x,y)\,dy\,dx.$$

2. *If R is bounded by $c \le y \le d$ and $h_1(y) \le x \le h_2(y)$, where h_1 and h_2 are continuous functions on $[c,d]$, then*

$$\iint_R f(x,y)\,dA = \int_c^d \int_{h_1(y)}^{h_2(y)} f(x,y)\,dx\,dy.$$

youtu.be/watch?v=smdRuI_ztAw

Figure 15.2.8 Fubini's Theorem and iterated integrals

Note that once again the bounds of integration follow the "curve to curve, point to point" pattern discussed in the previous section. In fact, one of the main points of the previous section is developing the skill of describing a region R with the bounds of an iterated integral. Once this skill is developed, we can use double integrals to compute many quantities, not just signed volume under a surface.

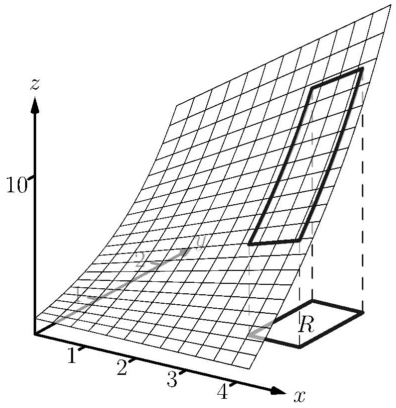


Figure 15.2.10 Finding the signed volume under a surface in Example 15.2.9

Example 15.2.9 Evaluating a double integral.

Let $f(x,y) = xy + e^y$. Find the signed volume under f on the region R, which is the rectangle with corners $(3,1)$ and $(4,2)$ pictured in Figure 15.2.10, using Fubini's Theorem and both orders of integration.

Solution. We wish to evaluate $\iint_R \left(xy + e^y\right) dA$. As R is a rectangle, the bounds are easily described as $3 \le x \le 4$ and $1 \le y \le 2$.

Using the order $dy\,dx$:

$$\begin{aligned}
\iint_R \left(xy + e^y\right) dA &= \int_3^4 \int_1^2 \left(xy + e^y\right) dy\,dx \\
&= \int_3^4 \left(\left[\frac{1}{2}xy^2 + e^y \right]_1^2 \right) dx \\
&= \int_3^4 \left(\frac{3}{2}x + e^2 - e \right) dx
\end{aligned}$$

$$= \left(\frac{3}{4}x^2 + (e^2 - e)x\right)\Big|_3^4$$
$$= \frac{21}{4} + e^2 - e \approx 9.92.$$

Now we check the validity of Fubini's Theorem by using the order $dx\,dy$:

$$\begin{aligned}\iint_R \left(xy + e^y\right) dA &= \int_1^2 \int_3^4 \left(xy + e^y\right) dx\,dy \\ &= \int_1^2 \left(\left[\frac{1}{2}x^2 y + xe^y\right]\Big|_3^4 \right) dy \\ &= \int_1^2 \left(\frac{7}{2}y + e^y\right) dy \\ &= \left(\frac{7}{4}y^2 + e^y\right)\Big|_1^2 \\ &= \frac{21}{4} + e^2 - e \approx 9.92.\end{aligned}$$

Both orders of integration return the same result, as expected.

Example 15.2.11 Evaluating a double integral.

Evaluate $\iint_R \left(3xy - x^2 - y^2 + 6\right) dA$, where R is the triangle bounded by $x = 0$, $y = 0$ and $x/2 + y = 1$, as shown in Figure 15.2.12.

Solution. While it is not specified which order we are to use, we will evaluate the double integral using both orders to help drive home the point that it does not matter which order we use.

Using the order $dy\,dx$: The bounds on y go from "curve to curve," i.e., $0 \le y \le 1 - x/2$, and the bounds on x go from "point to point," i.e., $0 \le x \le 2$.

$$\begin{aligned}\iint_R \left(3xy - x^2 - y^2 + 6\right) dA &= \int_0^2 \int_0^{-\frac{x}{2}+1} \left(3xy - x^2 - y^2 + 6\right) dy\,dx \\ &= \int_0^2 \left(\frac{3}{2}xy^2 - x^2 y - \frac{1}{3}y^3 + 6y\right)\Big|_0^{-\frac{x}{2}+1} dx \\ &= \int_0^2 \left(\frac{11}{12}x^3 - \frac{11}{4}x^2 - x + \frac{17}{3}\right) dx \\ &= \left(\frac{11}{48}x^4 - \frac{11}{12}x^3 - \frac{1}{2}x^2 + \frac{17}{3}x\right)\Big|_0^2 \\ &= \frac{17}{3} = 5.\overline{6}.\end{aligned}$$

Now lets consider the order $dx\,dy$. Here x goes from "curve to curve," $0 \le x \le 2 - 2y$, and y goes from "point to point," $0 \le y \le 1$:

$$\iint_R \left(3xy - x^2 - y^2 + 6\right) dA = \int_0^1 \int_0^{2-2y} \left(3xy - x^2 - y^2 + 6\right) dx\,dy$$

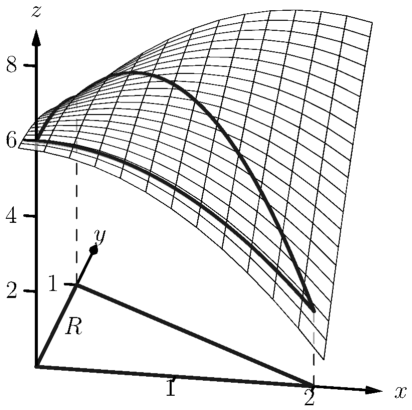


Figure 15.2.12 Finding the signed volume under the surface in Example 15.2.11

youtu.be/watch?v=MMbLzFw-5Rg

Figure 15.2.13 Further examples of double integrals over rectangles

$$= \int_0^1 \left(\frac{3}{2}x^2y - \frac{1}{3}x^3 - xy^2 + 6x\right)\Big|_0^{2-2y} dy$$
$$= \int_0^1 \left(\frac{32}{3}y^3 - 22y^2 + 2y + \frac{28}{3}\right) dy$$
$$= \left(\frac{8}{3}y^4 - \frac{22}{3}y^3 + y^2 + \frac{28}{3}y\right)\Big|_0^1$$
$$= \frac{17}{3} = 5.\overline{6}.$$

We obtained the same result using both orders of integration.

Note how in these examples that the bounds of integration depend only on R; the bounds of integration have nothing to do with $f(x, y)$. This is an important concept, so we include it as a Key Idea.

Key Idea 15.2.14 Double Integration Bounds.

When evaluating $\iint_R f(x, y)\, dA$ using an iterated integral, the bounds of integration depend only on R. The function f does not determine the bounds of integration.

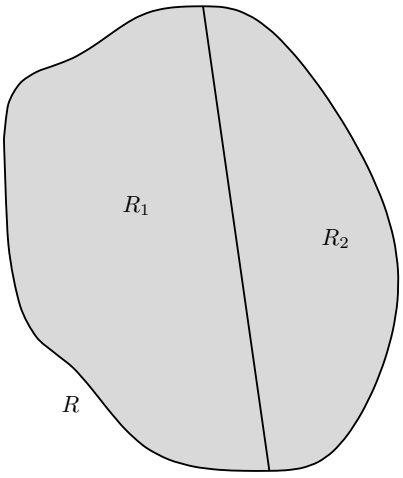


Figure 15.2.16 *R is the union of two nonoverlapping regions, R_1 and R_2*

Before doing another example, we give some properties of double integrals. Each should make sense if we view them in the context of finding signed volume under a surface, over a region.

Theorem 15.2.15 Properties of Double Integrals.

Let f and g be continuous functions over a closed, bounded plane region R, and let c be a constant.

1. $\displaystyle\iint_R c\, f(x, y)\, dA = c \iint_R f(x, y)\, dA.$
2. $\displaystyle\iint_R \big(f(x, y) \pm g(x, y)\big)\, dA = \iint_R f(x, y)\, dA \pm \iint_R g(x, y)\, dA$
3. *If $f(x, y) \geq 0$ on R, then* $\displaystyle\iint_R f(x, y)\, dA \geq 0.$
4. *If $f(x, y) \geq g(x, y)$ on R, then* $\displaystyle\iint_R f(x, y)\, dA \geq \iint_R g(x, y)\, dA.$
5. *Let R be the union of two nonoverlapping regions, $R = R_1 \bigcup R_2$ (see Figure 15.2.16). Then*

$$\iint_R f(x, y)\, dA = \iint_{R_1} f(x, y)\, dA + \iint_{R_2} f(x, y)\, dA.$$

youtu.be/watch?v=FcxnuFSNOfY

Figure 15.2.17 Evaluating a double integral over a more general region

Example 15.2.18 Evaluating a double integral.

Let $f(x, y) = \sin(x)\cos(y)$ and R be the triangle with vertices $(-1, 0)$, $(1, 0)$ and $(0, 1)$ (see Figure 15.2.19). Evaluate the double integral $\iint_R f(x, y)\, dA$.

Solution. If we attempt to integrate using an iterated integral with the

order $dy\,dx$, note how there are two upper bounds on R meaning we'll need to use two iterated integrals. We would need to split the triangle into two regions along the y-axis, then use Theorem 15.2.15, part 5.

Instead, let's use the order $dx\,dy$. The curves bounding x are $y-1 \le x \le 1-y$; the bounds on y are $0 \le y \le 1$. This gives us:

$$\begin{aligned}\iint_R f(x,y)\,dA &= \int_0^1 \int_{y-1}^{1-y} \sin(x)\cos(y)\,dx\,dy \\ &= \int_0^1 \Big(-\cos(x)\cos(y)\Big)\Big|_{y-1}^{1-y}\,dy \\ &= \int_0^1 \cos(y)\Big(-\cos(1-y)+\cos(y-1)\Big)dy.\end{aligned}$$

Recall that the cosine function is an even function; that is, $\cos(x) = \cos(-x)$. Therefore, from the last integral above, we have $\cos(y-1) = \cos(1-y)$. Thus the integrand simplifies to 0, and we have

$$\begin{aligned}\iint_R f(x,y)\,dA &= \int_0^1 0\,dy \\ &= 0.\end{aligned}$$

It turns out that over R, there is just as much volume above the xy-plane as below (look again at Figure 15.2.19), giving a final signed volume of 0.

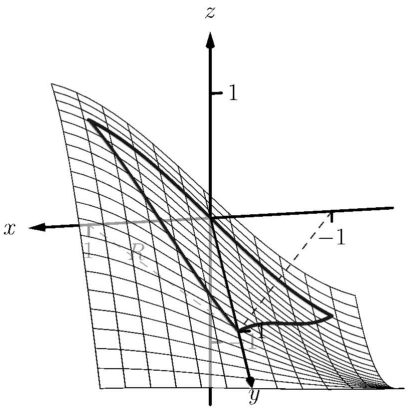


Figure 15.2.19 Finding the signed volume under a surface in Example 15.2.18

youtu.be/watch?v=m_8kBmNVxEM

Figure 15.2.20 Finding the volume of a tetrahedron

Example 15.2.21 Evaluating a double integral.

Evaluate $\iint_R (4-y)\,dA$, where R is the region bounded by the parabolas $y^2 = 4x$ and $x^2 = 4y$, graphed in Figure 15.2.22.

Solution. Graphing each curve can help us find their points of intersection. Solving analytically, the second equation tells us that $y = x^2/4$. Substituting this value in for y in the first equation gives us $x^4/16 = 4x$. Solving for x:

$$\begin{aligned}\frac{x^4}{16} &= 4x \\ x^4 - 64x &= 0 \\ x(x^3-64) &= 0 \\ x &= 0,\ 4.\end{aligned}$$

Thus we've found analytically what was easy to approximate graphically: the regions intersect at $(0,0)$ and $(4,4)$, as shown in Figure 15.2.22.

We now choose an order of integration: $dy\,dx$ or $dx\,dy$? Either order works; since the integrand does not contain x, choosing $dx\,dy$ might be simpler — at least, the first integral is very simple.

Thus we have the following "curve to curve, point to point" bounds:

$$y^2/4 \le x \le 2\sqrt{y}, \text{ and } 0 \le y \le 4.$$

Therefore,

$$\iint_R (4-y)\,dA = \int_0^4 \int_{y^2/4}^{2\sqrt{y}} (4-y)\,dx\,dy$$

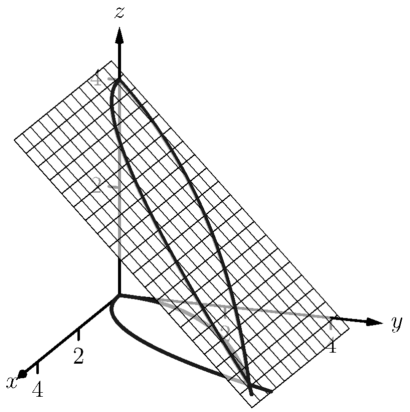


Figure 15.2.22 Finding the volume under the surface in Example 15.2.21

youtu.be/watch?v=a4XiokkzyrU

Figure 15.2.23 Evaluating a double integral over a general region

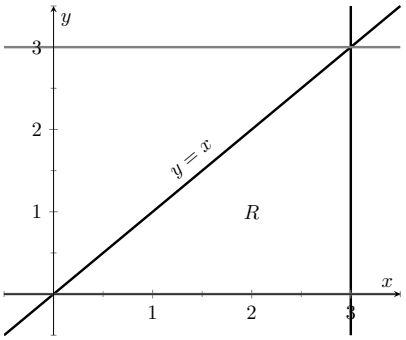


Figure 15.2.25 Determining the region R determined by the bounds of integration in Example 15.2.24

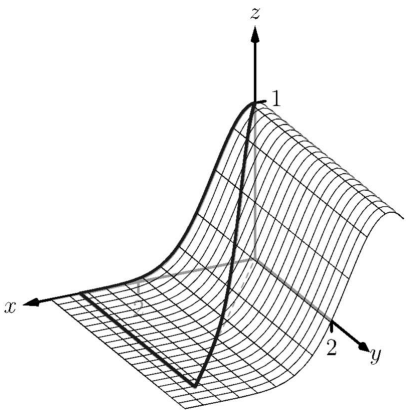


Figure 15.2.26 Showing the surface $z = f(x, y)$ defined in Example 15.2.24 over its region R

$$
\begin{aligned}
&= \int_0^4 \big(x(4-y)\big)\Big|_{y^2/4}^{2\sqrt{y}} dy \\
&= \int_0^4 \Big(\big(2\sqrt{y} - \frac{y^2}{4}\big)\big(4-y\big)\Big)\, dy \\
&= \int_0^4 \Big(\frac{y^3}{4} - y^2 - 2y^{3/2} + 8y^{1/2}\Big)\, dy \\
&= \left(\frac{y^4}{16} - \frac{y^3}{3} - \frac{4y^{5/2}}{5} + \frac{16y^{3/2}}{3}\right)\Big|_0^4 \\
&= \frac{176}{15} = 11.7\overline{3}.
\end{aligned}
$$

The signed volume under the surface $z = f(x, y)$ is about 11.7 cubic units.

In the previous section we practiced changing the order of integration of a given iterated integral, where the region R was not explicitly given. Changing the bounds of an integral is more than just an test of understanding. Rather, there are cases where integrating in one order is really hard, if not impossible, whereas integrating with the other order is feasible.

Example 15.2.24 Changing the order of integration.

Rewrite the iterated integral $\displaystyle\int_0^3 \int_y^3 e^{-x^2}\, dx\, dy$ with the order $dy\, dx$. Comment on the feasibility to evaluate each integral.

Solution. Once again we make a sketch of the region over which we are integrating to facilitate changing the order. The bounds on x are from $x = y$ to $x = 3$; the bounds on y are from $y = 0$ to $y = 3$. These curves are sketched in Figure 15.2.25, enclosing the region R.

To change the bounds, note that the curves bounding y are $y = 0$ up to $y = x$; the triangle is enclosed between $x = 0$ and $x = 3$. Thus the new bounds of integration are $0 \le y \le x$ and $0 \le x \le 3$, giving the iterated integral $\displaystyle\int_0^3 \int_0^x e^{-x^2}\, dy\, dx$.

How easy is it to evaluate each iterated integral? Consider the order of integrating $dx\, dy$, as given in the original problem. The first indefinite integral we need to evaluate is $\int e^{-x^2}\, dx$; we have stated before (see Section 5.5) that this integral cannot be evaluated in terms of elementary functions. We are stuck.

Changing the order of integration makes a big difference here. In the second iterated integral, we are faced with $\int e^{-x^2}\, dy$; integrating with respect to y gives us $ye^{-x^2} + C$, and the first definite integral evaluates to

$$\int_0^x e^{-x^2}\, dy = xe^{-x^2}.$$

Thus

$$\int_0^3 \int_0^x e^{-x^2}\, dy\, dx = \int_0^3 \Big(xe^{-x^2}\Big) dx.$$

This last integral is easy to evaluate with substitution, giving a final answer of $\frac{1}{2}(1 - e^{-9}) \approx 0.5$. Figure 15.2.26 shows the surface over R.

In short, evaluating one iterated integral is impossible; the other iterated integral is relatively simple.

Definition 5.4.34 defines the average value of a single-variable function $f(x)$ on the interval $[a, b]$ as

$$\text{average value of } f(x) \text{ on } [a,b] = \frac{1}{b-a}\int_a^b f(x)\,dx;$$

that is, it is the "area under f over an interval divided by the length of the interval." We make an analogous statement here: the average value of $z = f(x, y)$ over a region R is the volume under f over R divided by the area of R.

Definition 15.2.28 The Average Value of f on R.

Let $z = f(x, y)$ be a continuous function defined over a closed, bounded region R in the xy-plane. The **average value** of f on R is

$$\text{average value of } f \text{ on } R = \frac{\iint_R f(x,y)\,dA}{\iint_R dA}.$$

Example 15.2.29 Finding average value of a function over a region R.

Find the average value of $f(x, y) = 4 - y$ over the region R, which is bounded by the parabolas $y^2 = 4x$ and $x^2 = 4y$. Note: this is the same function and region as used in Example 15.2.21.

Solution. In Example 15.2.21 we found

$$\iint_R f(x,y)\,dA = \int_0^4 \int_{y^2/4}^{2\sqrt{y}} (4-y)\,dx\,dy = \frac{176}{15}.$$

We find the area of R by computing $\iint_R dA$:

$$\iint_R dA = \int_0^4 \int_{y^2/4}^{2\sqrt{y}} dx\,dy = \frac{16}{3}.$$

Dividing the volume under the surface by the area gives the average value:

$$\text{average value of } f \text{ on } R = \frac{176/15}{16/3} = \frac{11}{5} = 2.2.$$

While the surface, as shown in Figure 15.2.30, covers z-values from $z = 0$ to $z = 4$, the "average" z-value on R is 2.2.

The previous section introduced the iterated integral in the context of finding the area of plane regions. This section has extended our understanding of iterated integrals; now we see they can be used to find the signed volume under a surface.

This new understanding allows us to revisit what we did in the previous section. Given a region R in the plane, we computed $\iint_R 1\,dA$; again, our understanding at the time was that we were finding the area of R. However, we can

youtu.be/watch?v=2lQA7pTALGk

Figure 15.2.27 Two more examples of changing the order of integration

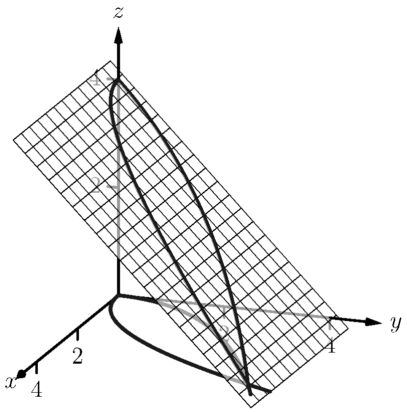


Figure 15.2.30 Finding the average value of f in Example 15.2.29

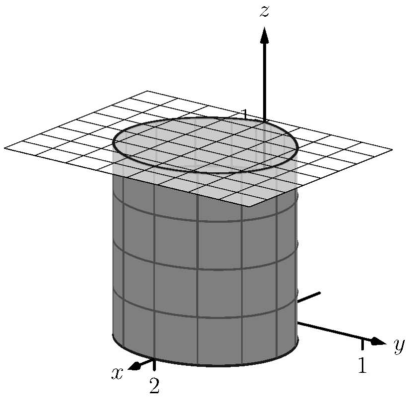


Figure 15.2.31 Showing how an iterated integral used to find area also finds a certain volume

now view the graph $z = 1$ as a surface, a flat surface with constant z-value of 1. The double integral $\iint_R 1\,dA$ finds the volume, under $z = 1$, over R, as shown in Figure 15.2.31. Basic geometry tells us that if the base of a general right cylinder has area A, its volume is $A \cdot h$, where h is the height. In our case, the height is 1. We were "actually" computing the volume of a solid, though we interpreted the number as an area.

The next section extends our abilities to find "volumes under surfaces." Currently, some integrals are hard to compute because either the region R we are integrating over is hard to define with rectangular curves, or the integrand itself is hard to deal with. Some of these problems can be solved by converting everything into polar coordinates.

15.2.1 Exercises

Terms and Concepts

1. An integral can be interpreted as giving the signed area over an interval; a double integral can be interpreted as giving the signed _____ over a region.

2. Explain why the following statement is false: Fubini's Theorem states that

$$\int_a^b \int_{g_1(x)}^{g_2(x)} f(x,y)\,dy\,dx = \int_a^b \int_{g_1(y)}^{g_2(y)} f(x,y)\,dx\,dy.$$

3. Explain why if $f(x,y) > 0$ over a region R, then $\iint_R f(x,y)\,dA > 0$.

4. If $\iint_R f(x,y)\,dA = \iint_R g(x,y)\,dA$, does this imply $f(x,y) = g(x,y)$?

Problems

Exercise Group. For the given integral,

(a) Evaluate the given iterated integral, and

(b) rewrite the integral using the other order of integration.

5. $\displaystyle\int_1^2 \int_{-1}^1 \left(\frac{x}{y} + 3\right)\,dx\,dy$

6. $\displaystyle\int_{-\pi/2}^{\pi/2} \int_0^{\pi} \left(\sin(x)\cos(y)\right)\,dx\,dy$

7. $\displaystyle\int_0^4 \int_0^{-x/2+2} \left(3x^2 - y + 2\right)\,dy\,dx$

8. $\displaystyle\int_1^3 \int_y^3 \left(x^2y - xy^2\right)\,dx\,dy$

9. $\displaystyle\int_0^1 \int_{-\sqrt{1-y}}^{\sqrt{1-y}} (x + y + 2)\,dx\,dy$

10. $\displaystyle\int_0^9 \int_{y/3}^{\sqrt{y}} \left(xy^2\right)\,dx\,dy$

Exercise Group. In the following exercises:

(a) Sketch the region R given by the problem.

(b) Set up the iterated integrals, in both orders, that evaluate the given double integral for the described region R.

(c) Evaluate one of the iterated integrals to find the signed volume under the surface $z = f(x,y)$ over the region R.

11. $\displaystyle\iint_R x^2y\,dA$, where R is bounded by $y = \sqrt{x}$ and $y = x^2$.

12. $\displaystyle\iint_R x^2y\,dA$, where R is bounded by $y = \sqrt[3]{x}$ and $y = x^3$.

13. $\displaystyle\iint_R x^2 - y^2\,dA$, where R is the rectangle with corners $(-1,-1)$, $(1,-1)$, $(1,1)$ and $(-1,1)$.

14. $\displaystyle\iint_R ye^x\,dA$, where R is bounded by $x = 0$, $x = y^2$ and $y = 1$.

15. $\displaystyle\iint_R \left(6 - 3x - 2y\right)\,dA$, where R is bounded by $x = 0$, $y = 0$ and $3x + 2y = 6$.

16. $\displaystyle\iint_R e^y\,dA$, where R is bounded by $y = \ln(x)$ and $y = \dfrac{1}{e-1}(x-1)$.

17. $\displaystyle\iint_R \left(x^3y - x\right)\,dA$, where R is the half of the circle $x^2 + y^2 = 9$ in the first and second quadrants.

18. $\displaystyle\iint_R \left(4 - 3y\right)\,dA$, where R is bounded by $y = 0$, $y = x/e$ and $y = \ln(x)$.

Exercise Group. In the following exercises, state why it is difficult/impossible to integrate the iterated integral in the given order of integration. Change the order of integration and evaluate the new iterated integral.

19. $\displaystyle\int_0^4 \int_{y/2}^2 e^{x^2}\,dx\,dy$

20. $\displaystyle\int_0^{\sqrt{\pi/2}} \int_x^{\sqrt{\pi/2}} \cos\left(y^2\right)\,dy\,dx$

21. $\displaystyle\int_0^1 \int_y^1 \frac{2y}{x^2+y^2}\,dx\,dy$

22. $\displaystyle\int_{-1}^1 \int_1^2 \frac{x\tan^2(y)}{1+\ln(y)}\,dy\,dx$

Exercise Group. In the following exercises, find the average value of f over the region R. Notice how these functions and regions are related to the iterated integrals given in Exercises 15.2.5 — Exercise 15.2.8.

23. $f(x,y)=\dfrac{x}{y}+3$; R is the rectangle with opposite corners $(-1,1)$ and $(1,2)$.

24. $f(x,y)=\sin(x)\cos(y)$; R is bounded by $x=0$, $x=\pi$, $y=-\pi/2$ and $y=\pi/2$.

25. $f(x,y)=3x^2-y+2$; R is bounded by the lines $y=0$, $y=2-x/2$ and $x=0$.

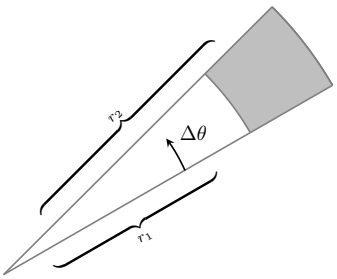
26. $f(x,y)=x^2y-xy^2$; R is bounded by $y=x$, $y=1$ and $x=3$.

15.3 Double Integration with Polar Coordinates

We have used iterated integrals to evaluate double integrals, which give the signed volume under a surface, $z = f(x, y)$, over a region R of the xy-plane. The integrand is simply $f(x, y)$, and the bounds of the integrals are determined by the region R.

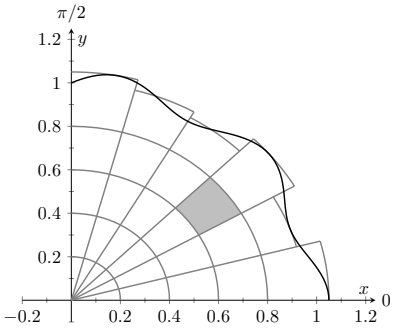
Some regions R are easy to describe using rectangular coordinates — that is, with equations of the form $y = f(x)$, $x = a$, etc. However, some regions are easier to handle if we represent their boundaries with polar equations of the form $r = f(\theta)$, $\theta = \alpha$, etc.

The basic form of the double integral is $\iint_R f(x, y)\, dA$. We interpret this integral as follows: over the region R, sum up lots of products of heights (given by $f(x_i, y_i)$) and areas (given by ΔA_i). That is, dA represents "a little bit of area." In rectangular coordinates, we can describe a small rectangle as having area $dx\, dy$ or $dy\, dx$ — the area of a rectangle is simply length×width — a small change in x times a small change in y. Thus we replace dA in the double integral with $dx\, dy$ or $dy\, dx$.



youtu.be/watch?v=K7hpTI0PNGo

Figure 15.3.1 Introducing the double integral in polar coordinates



(a) **(b)**

Figure 15.3.2 Approximating a region R with portions of sectors of circles

Now consider representing a region R with polar coordinates. Consider Figure 15.3.2(a). Let R be the region in the first quadrant bounded by the curve. We can approximate this region using the natural shape of polar coordinates: portions of sectors of circles. In the figure, one such region is shaded, shown again in Figure 15.3.2(b).

As the area of a sector of a circle with radius r, subtended by an angle θ, is $A = \frac{1}{2}r^2\theta$, we can find the area of the shaded region. The whole sector has area $\frac{1}{2}r_2^2\Delta\theta$, whereas the smaller, unshaded sector has area $\frac{1}{2}r_1^2\Delta\theta$. The area of the shaded region is the difference of these areas:

$$\Delta A_i = \frac{1}{2}r_2^2\Delta\theta - \frac{1}{2}r_1^2\Delta\theta = \frac{1}{2}\big(r_2^2 - r_1^2\big)\big(\Delta\theta\big) = \frac{r_2 + r_1}{2}\big(r_2 - r_1\big)\Delta\theta.$$

Note that $(r_2 + r_1)/2$ is just the average of the two radii.

To approximate the region R, we use many such subregions; doing so shrinks the difference $r_2 - r_1$ between radii to 0 and shrinks the change in angle $\Delta\theta$ also to 0. We represent these infinitesimal changes in radius and angle as dr and $d\theta$, respectively. Finally, as dr is small, $r_2 \approx r_1$, and so $(r_2 + r_1)/2 \approx r_1$. Thus, when dr and $d\theta$ are small,

$$\Delta A_i \approx r_i\, dr\, d\theta.$$

Taking a limit, where the number of subregions goes to infinity and both $r_2 - r_1$ and $\Delta\theta$ go to 0, we get

$$dA = r\, dr\, d\theta.$$

So to evaluate $\iint_R f(x,y)\,dA$, replace dA with $r\,dr\,d\theta$. Convert the function $f(x,y)$ to a function with polar coordinates with the substitutions $x = r\cos(\theta)$, $y = r\sin(\theta)$. Finally, find bounds $g_1(\theta) \le r \le g_2(\theta)$ and $\alpha \le \theta \le \beta$ that describe R. This is the key principle of this section, so we restate it here as a Key Idea.

Key Idea 15.3.3 Evaluating Double Integrals with Polar Coordinates.

Let $z = f(x,y)$ be a continuous function defined over a closed, bounded region R in the xy-plane, where R is bounded by the polar equations $\alpha \le \theta \le \beta$ and $g_1(\theta) \le r \le g_2(\theta)$. Then

$$\iint_R f(x,y)\,dA = \int_\alpha^\beta \int_{g_1(\theta)}^{g_2(\theta)} f\big(r\cos(\theta), r\sin(\theta)\big)\, r\,dr\,d\theta.$$

Examples will help us understand this Key Idea.

Example 15.3.4 Evaluating a double integral with polar coordinates.

Find the signed volume under the plane $z = 4 - x - 2y$ over the disk bounded by the circle with equation $x^2 + y^2 = 1$.

Solution. The bounds of the integral are determined solely by the region R over which we are integrating. In this case, it is a disk with boundary $x^2 + y^2 = 1$. We need to find polar bounds for this region. It may help to review Section 10.4; bounds for this disk are $0 \le r \le 1$ and $0 \le \theta \le 2\pi$.

We replace $f(x,y)$ with $f(r\cos(\theta), r\sin(\theta))$. That means we make the following substitutions:

$$4 - x - 2y \Rightarrow 4 - r\cos(\theta) - 2r\sin(\theta).$$

Finally, we replace dA in the double integral with $r\,dr\,d\theta$. This gives the final iterated integral, which we evaluate:

$$\begin{aligned}
\iint_R f(x,y)\,dA &= \int_0^{2\pi}\int_0^1 \big(4 - r\cos(\theta) - 2r\sin(\theta)\big) r\,dr\,d\theta \\
&= \int_0^{2\pi}\int_0^1 \big(4r - r^2(\cos(\theta) - 2\sin(\theta))\big)\,dr\,d\theta \\
&= \int_0^{2\pi} \left(2r^2 - \frac{1}{3}r^3(\cos(\theta) - 2\sin(\theta))\right)\Big|_0^1 d\theta \\
&= \int_0^{2\pi} \left(2 - \frac{1}{3}\big(\cos(\theta) - 2\sin(\theta)\big)\right) d\theta \\
&= \left(2\theta - \frac{1}{3}\big(\sin(\theta) + 2\cos(\theta)\big)\right)\Big|_0^{2\pi} \\
&= 4\pi \approx 12.566.
\end{aligned}$$

The surface and region R are shown in Figure 15.3.5.

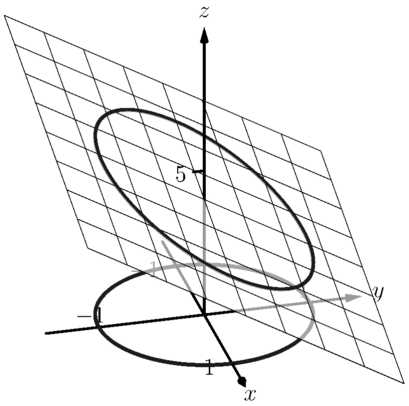


Figure 15.3.5 Evaluating a double integral with polar coordinates in Example 15.3.4

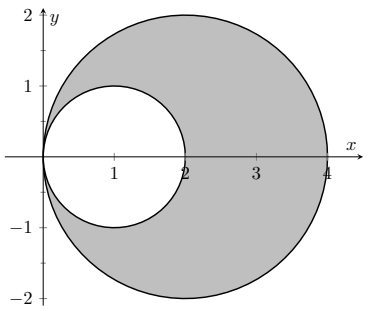
youtu.be/watch?v=Gmt0dak05Lk

Figure 15.3.6 Two simple examples of double integrals in polar coordinates

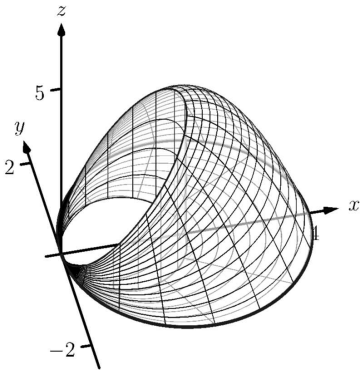
Example 15.3.7 Evaluating a double integral with polar coordinates.

Find the volume under the paraboloid $z = 4 - (x-2)^2 - y^2$ over the region bounded by the circles $(x-1)^2 + y^2 = 1$ and $(x-2)^2 + y^2 = 4$.

Solution. At first glance, this seems like a very hard volume to compute as the region R (shown in Figure 15.3.8(a)) has a hole in it, cutting out a strange portion of the surface, as shown in Figure 15.3.8(b). However, by describing R in terms of polar equations, the volume is not very difficult to compute.



(a)



(b)

Figure 15.3.8 Showing the region R and surface used in Example 15.3.7

It is straightforward to show that the circle $(x-1)^2 + y^2 = 1$ has polar equation $r = 2\cos(\theta)$, and that the circle $(x-2)^2 + y^2 = 4$ has polar equation $r = 4\cos(\theta)$. Each of these circles is traced out on the interval $0 \le \theta \le \pi$. The bounds on r are $2\cos(\theta) \le r \le 4\cos(\theta)$.

Replacing x with $r\cos(\theta)$ in the integrand, along with replacing y with $r\sin(\theta)$, prepares us to evaluate the double integral $\iint_R f(x,y)\,dA$:

$$\begin{aligned}
\iint_R f(x,y)\,dA &= \int_0^\pi \int_{2\cos(\theta)}^{4\cos(\theta)} \Big(4 - \big(r\cos(\theta) - 2\big)^2 - \big(r\sin(\theta)\big)^2\Big) r\,dr\,d\theta \\
&= \int_0^\pi \int_{2\cos(\theta)}^{4\cos(\theta)} \big(-r^3 + 4r^2\cos(\theta)\big)\,dr\,d\theta \\
&= \int_0^\pi \left(-\frac{1}{4}r^4 + \frac{4}{3}r^3\cos(\theta)\right)\Big|_{2\cos(\theta)}^{4\cos(\theta)} d\theta \\
&= \int_0^\pi \left(\left[-\frac{1}{4}(256\cos^4(\theta)) + \frac{4}{3}(64\cos^4(\theta))\right] - \right. \\
&\qquad \left.\left[-\frac{1}{4}(16\cos^4(\theta)) + \frac{4}{3}(8\cos^4(\theta))\right]\right) d\theta \\
&= \int_0^\pi \frac{44}{3}\cos^4(\theta)\,d\theta
\end{aligned}$$

To integrate $\cos^4(\theta)$, rewrite it as $\cos^2(\theta)\cos^2(\theta)$ and employ the power-reducing formula twice:

$$\begin{aligned}
\cos^4(\theta) &= \cos^2(\theta)\cos^2(\theta) \\
&= \frac{1}{2}\big(1 + \cos(2\theta)\big)\frac{1}{2}\big(1 + \cos(2\theta)\big)
\end{aligned}$$

$$= \frac{1}{4}\big(1 + 2\cos(2\theta) + \cos^2(2\theta)\big)$$
$$= \frac{1}{4}\Big(1 + 2\cos(2\theta) + \frac{1}{2}\big(1 + \cos(4\theta)\big)\Big)$$
$$= \frac{3}{8} + \frac{1}{2}\cos(2\theta) + \frac{1}{8}\cos(4\theta).$$

Picking up from where we left off above, we have

$$\begin{aligned}\iint_R f(x,y)\,dA &= \int_0^\pi \frac{44}{3}\cos^4(\theta)\,d\theta \\ &= \int_0^\pi \frac{44}{3}\left(\frac{3}{8} + \frac{1}{2}\cos(2\theta) + \frac{1}{8}\cos(4\theta)\right)d\theta \\ &= \frac{44}{3}\left(\frac{3}{8}\theta + \frac{1}{4}\sin(2\theta) + \frac{1}{32}\sin(4\theta)\right)\Big|_0^\pi \\ &= \frac{11}{2}\pi \approx 17.279.\end{aligned}$$

While this example was not trivial, the double integral would have been *much* harder to evaluate had we used rectangular coordinates.

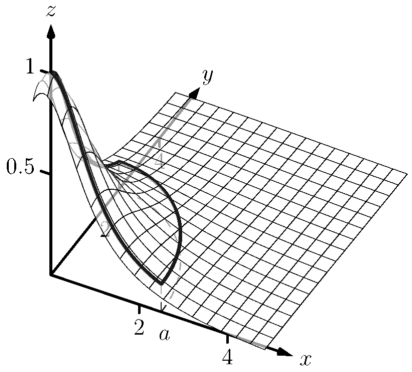


Figure 15.3.10 The surface and region R used in Example 15.3.9

Previous work has shown that there is finite *area* under $\frac{1}{x^2+1}$ over the entire x-axis. However, Example 15.3.9 shows that there is infinite *volume* under $\frac{1}{x^2+y^2+1}$ over the entire xy-plane.

Example 15.3.9 Evaluating a double integral with polar coordinates.

Find the volume under the surface given by the graph of $f(x,y) = \dfrac{1}{x^2+y^2+1}$ over the sector of the circle with radius a centered at the origin in the first quadrant, as shown in Figure 15.3.10.

Solution. The region R we are integrating over is a circle with radius a, restricted to the first quadrant. Thus, in polar, the bounds on R are $0 \le r \le a$, $0 \le \theta \le \pi/2$. The integrand is rewritten in polar as

$$\frac{1}{x^2+y^2+1} \Rightarrow \frac{1}{r^2\cos^2(\theta) + r^2\sin^2(\theta) + 1} = \frac{1}{r^2+1}.$$

We find the volume as follows:

$$\begin{aligned}\iint_R f(x,y)\,dA &= \int_0^{\pi/2}\int_0^a \frac{r}{r^2+1}\,dr\,d\theta \\ &= \int_0^{\pi/2} \frac{1}{2}\big(\ln\left|r^2+1\right|\big)\Big|_0^a\,d\theta \\ &= \int_0^{\pi/2} \frac{1}{2}\ln(a^2+1)\,d\theta \\ &= \left(\frac{1}{2}\ln(a^2+1)\theta\right)\Big|_0^{\pi/2} \\ &= \frac{\pi}{4}\ln(a^2+1).\end{aligned}$$

Figure 15.3.10 shows that f shrinks to near 0 very quickly. Regardless, as a grows, so does the volume, without bound.

Example 15.3.11 Finding the volume of a sphere.

Find the volume of a sphere with radius a.

Solution. The sphere of radius a, centered at the origin, has equation $x^2 + y^2 + z^2 = a^2$; solving for z, we have $z = \sqrt{a^2 - x^2 - y^2}$. This gives the upper half of a sphere. We wish to find the volume under this top half, then double it to find the total volume.

The region we need to integrate over is the disk of radius a, centered at the origin. Polar bounds for this equation are $0 \le r \le a, 0 \le \theta \le 2\pi$.

All together, the volume of a sphere with radius a is:

$$\begin{aligned} 2\iint_R \sqrt{a^2 - x^2 - y^2}\, dA &= 2\int_0^{2\pi}\int_0^a \sqrt{a^2 - (r\cos(\theta))^2 - (r\sin(\theta))^2} r\, dr\, d\theta \\ &= 2\int_0^{2\pi}\int_0^a r\sqrt{a^2 - r^2}\, dr\, d\theta. \end{aligned}$$

We can evaluate this inner integral with substitution. With $u = a^2 - r^2$, $du = -2r\, dr$. The new bounds of integration are $u(0) = a^2$ to $u(a) = 0$. Thus we have:

$$\begin{aligned} &= \int_0^{2\pi}\int_{a^2}^0 \left(-u^{1/2}\right) du\, d\theta \\ &= \int_0^{2\pi} \left(-\frac{2}{3}u^{3/2}\right)\Big|_{a^2}^0 d\theta \\ &= \int_0^{2\pi} \left(\frac{2}{3}a^3\right) d\theta \\ &= \left(\frac{2}{3}a^3\theta\right)\Big|_0^{2\pi} \\ &= \frac{4}{3}\pi a^3. \end{aligned}$$

Generally, the formula for the volume of a sphere with radius r is given as $4/3\pi r^3$; we have justified this formula with our calculation.

youtu.be/watch?v=47Xebo9FsKs

Figure 15.3.12 Finding the volume bounded by two surfaces

Example 15.3.13 Finding the volume of a solid.

A sculptor wants to make a solid bronze cast of the solid shown in Figure 15.3.14, where the base of the solid has boundary, in polar coordinates, $r = \cos(3\theta)$, and the top is defined by the plane $z = 1 - x + 0.1y$. Find the volume of the solid.

Solution. From the outset, we should recognize that knowing *how to set up* this problem is probably more important than knowing *how to compute the integrals*. The iterated integral to come is not "hard" to evaluate, though it is long, requiring lots of algebra. Once the proper iterated integral is determined, one can use readily available technology to help compute the final answer.

The region R that we are integrating over is bound by $0 \le r \le \cos(3\theta)$, for $0 \le \theta \le \pi$ (note that this rose curve is traced out on the interval $[0, \pi]$, not $[0, 2\pi]$). This gives us our bounds of integration. The integrand is $z = 1 - x + 0.1y$; converting to polar, we have that the

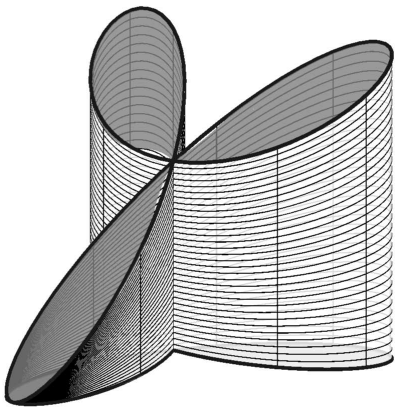


Figure 15.3.14 Visualizing the solid used in Example 15.3.13

volume V is:

$$V = \iint_R f(x,y)\,dA = \int_0^{\pi} \int_0^{\cos(3\theta)} \big(1 - r\cos(\theta) + 0.1r\sin(\theta)\big) r\,dr\,d\theta.$$

Distributing the r, the inner integral is easy to evaluate, leading to

$$\int_0^{\pi} \left(\frac{1}{2}\cos^2(3\theta) - \frac{1}{3}\cos^3(3\theta)\cos(\theta) + \frac{0.1}{3}\cos^3(3\theta)\sin(\theta) \right) d\theta.$$

This integral takes time to compute by hand; it is rather long and cumbersome. The powers of cosine need to be reduced, and products like $\cos(3\theta)\cos(\theta)$ need to be turned to sums using the Product To Sum formulas in the back cover of this text.

We rewrite $\frac{1}{2}\cos^2(3\theta)$ as $\frac{1}{4}(1+\cos(6\theta))$. We can also rewrite $\frac{1}{3}\cos^3(3\theta)\cos(\theta)$ as:

$$\begin{aligned} \frac{1}{3}\cos^3(3\theta)\cos(\theta) &= \frac{1}{3}\cos^2(3\theta)\cos(3\theta)\cos(\theta) \\ &= \frac{1}{3}\frac{1+\cos(6\theta)}{2}\big(\cos(4\theta) + \cos(2\theta)\big). \end{aligned}$$

This last expression still needs simplification, but eventually all terms can be reduced to the form $a\cos(m\theta)$ or $a\sin(m\theta)$ for various values of a and m.

We forgo the algebra and recommend the reader employ technology, such as WolframAlpha®, to compute the numeric answer. Such technology gives:

$$\int_0^{\pi} \int_0^{\cos(3\theta)} \big(1 - r\cos(\theta) + 0.1r\sin(\theta)\big) r\,dr\,d\theta = \frac{\pi}{4} \approx 0.785u^3.$$

Since the units were not specified, we leave the result as almost 0.8 cubic units (meters, feet, etc.) Should the artist want to scale the piece uniformly, so that each rose petal had a length other than 1, she should keep in mind that scaling by a factor of k scales the volume by a factor of k^3.

youtu.be/watch?v=x72_v_zZkI4

Figure 15.3.15 Another integral over a rose curve

We have used iterated integrals to find areas of plane regions and volumes under surfaces. Just as a single integral can be used to compute much more than "area under the curve," iterated integrals can be used to compute much more than we have thus far seen. The next two sections show two, among many, applications of iterated integrals.

15.3.1 Exercises

Terms and Concepts

1. When evaluating $\iint_R f(x,y)\,dA$ using polar coordinates, $f(x,y)$ is replaced with _____ and dA is replaced with _____.

2. Why would one be interested in evaluating a double integral with polar coordinates?

Problems

Exercise Group. A function $f(x,y)$ is given and a region R of the xy-plane is described. Set up and evaluate $\iint_R f(x,y)\,dA$ using polar coordinates.

3. $f(x,y) = 3x - y + 4$ and R is the region enclosed by the circle $x^2 + y^2 = 1$.

4. $f(x,y) = 4x + 4y$; R is the region enclosed by the circle $x^2 + y^2 = 4$.

5. $f(x,y) = 8 - y$ and R is the region enclosed by the circles with polar equations $r = \cos(\theta)$ and $r = 3\cos(\theta)$.

6. $f(x,y) = 4$; R is the region enclosed by the petal of the rose curve $r = \sin(2\theta)$ in the first quadrant.

7. $f(x,y) = \ln(x^2 + y^2)$; R is the annulus enclosed by the circles $x^2 + y^2 = 1$ and $x^2 + y^2 = 4$.

8. $f(x,y) = 1 - x^2 - y^2$ and R is the region enclosed by the circle $x^2 + y^2 = 1$.

9. $f(x,y) = x^2 - y^2$; R is the region enclosed by the circle $x^2 + y^2 = 36$ in the first and fourth quadrants.

10. $f(x,y) = (x-y)/(x+y)$; R is the region enclosed by the lines $y = x$, $y = 0$ and the circle $x^2 + y^2 = 1$ in the first quadrant.

Exercise Group. An iterated integral in rectangular coordinates is given. Rewrite the integral using polar coordinates and evaluate the new double integral.

11. $\displaystyle\int_0^5 \int_{-\sqrt{25-x^2}}^{\sqrt{25-x^2}} \sqrt{x^2+y^2}\,dy\,dx$

12. $\displaystyle\int_{-4}^{4} \int_{-\sqrt{16-y^2}}^{0} \big(2y - x\big)\,dx\,dy$

13. $\displaystyle\int_0^2 \int_{y}^{\sqrt{8-y^2}} \big(x + y\big)\,dx\,dy$

14. $\displaystyle\int_{-2}^{-1} \int_{0}^{\sqrt{4-x^2}} \big(x+5\big)\,dy\,dx + \int_{-1}^{1} \int_{\sqrt{1-x^2}}^{\sqrt{4-x^2}} \big(x+5\big)\,dy\,dx + \int_{1}^{2} \int_{0}^{\sqrt{4-x^2}} \big(x+5\big)\,dy\,dx$

Hint: draw the region of each integral carefully and see how they all connect.

Exercise Group. In the following exercises, special double integrals are presented that are especially well suited for evaluation in polar coordinates.

15. Consider $\iint_R e^{-(x^2+y^2)}\,dA$.

(a) Why is this integral difficult to evaluate in rectangular coordinates, regardless of the region R?

(b) Let R be the region bounded by the circle of radius a centered at the origin. Evaluate the double integral using polar coordinates.

(c) Take the limit of your answer from (b), as $a \to \infty$. What does this imply about the volume under the surface $z = e^{-(x^2+y^2)}$ over the entire xy-plane?

16. The surface of a right circular cone with height h and base radius a can be described by the equation $f(x,y) = h - h\sqrt{\dfrac{x^2}{a^2} + \dfrac{y^2}{a^2}}$, where the tip of the cone lies at $(0,0,h)$ and the circular base lies in the xy-plane, centered at the origin.
Confirm that the volume of a right circular cone with height h and base radius a is $V = \dfrac{1}{3}\pi a^2 h$ by evaluating $\iint_R f(x,y)\,dA$ in polar coordinates.

15.4 Center of Mass

We have used iterated integrals to find areas of plane regions and signed volumes under surfaces. A brief recap of these uses will be useful in this section as we apply iterated integrals to compute the *mass* and *center of mass* of planar regions.

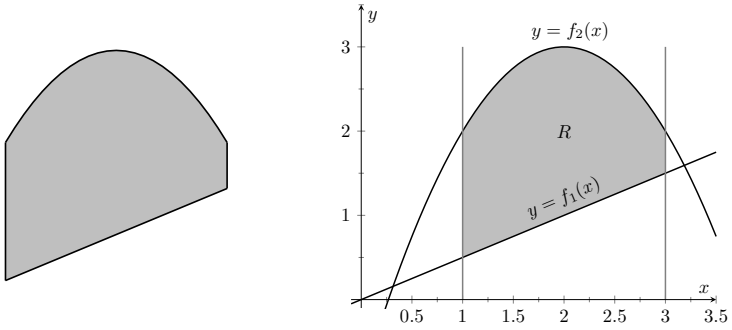
To find the area of a planar region, we evaluated the double integral $\iint_R dA$. That is, summing up the areas of lots of little subregions of R gave us the total area. Informally, we think of $\iint_R dA$ as meaning "sum up lots of little areas over R."

To find the signed volume under a surface, we evaluated the double integral $\iint_R f(x,y)\,dA$. Recall that the "dA" is not just a "bookend" at the end of an integral; rather, it is multiplied by $f(x,y)$. We regard $f(x,y)$ as giving a height, and dA still giving an area: $f(x,y)\,dA$ gives a volume. Thus, informally, $\iint_R f(x,y)\,dA$ means "sum up lots of little volumes over R."

We now extend these ideas to other contexts.

15.4.1 Mass and Weight

Consider a thin sheet of material with constant thickness and finite area. Mathematicians (and physicists and engineers) call such a sheet a *lamina*. So consider a lamina, as shown in Figure 15.4.1(a), with the shape of some planar region R, as shown in Figure 15.4.1(b).



(a) **(b)**

Figure 15.4.1 Illustrating the concept of a lamina

We can write a simple double integral that represents the mass of the lamina: $\iint_R dm$, where "dm" means "a little mass." That is, the double integral states the total mass of the lamina can be found by "summing up lots of little masses over R."

To evaluate this double integral, partition R into n subregions as we have done in the past. The ith subregion has area ΔA_i. A fundamental property of mass is that "mass=density×area." If the lamina has a constant density δ, then the mass of this ith subregion is $\Delta m_i = \delta \Delta A_i$. That is, we can compute a small amount of mass by multiplying a small amount of area by the density.

If density is variable, with density function $\delta = \delta(x,y)$, then we can approximate the mass of the ith subregion of R by multiplying ΔA_i by $\delta(x_i, y_i)$, where (x_i, y_i) is a point in that subregion. That is, for a small enough subregion of R, the density across that region is almost constant.

Mass and *weight* are different measures. Since they are scalar multiples of each other, it is often easy to treat them as the same measure. In this section we effectively treat them as the same, as our technique for finding mass is the same as for finding weight. The density functions used will simply have different units.

The total mass M of the lamina is approximately the sum of approximate masses of subregions:

$$M \approx \sum_{i=1}^{n} \Delta m_i = \sum_{i=1}^{n} \delta(x_i, y_i) \Delta A_i.$$

Taking the limit as the size of the subregions shrinks to 0 gives us the actual mass; that is, integrating $\delta(x, y)$ over R gives the mass of the lamina.

Definition 15.4.2 Mass of a Lamina with Variable Density.

Let $\delta(x, y)$ be a continuous density function of a lamina corresponding to a closed, bounded plane region R. The mass M of the lamina is

$$\text{mass } M = \iint_R dm = \iint_R \delta(x, y)\, dA.$$

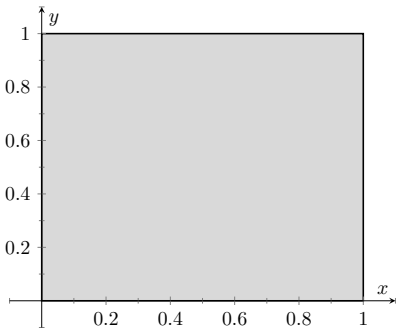


Figure 15.4.4 A region R representing a lamina in Example 15.4.3

Example 15.4.3 Finding the mass of a lamina with constant density.

Find the mass of a square lamina, with side length 1, with a density of $\delta = 3$ g/cm^2.

Solution. We represent the lamina with a square region in the plane as shown in Figure 15.4.4. As the density is constant, it does not matter where we place the square.

Following Definition 15.4.2, the mass M of the lamina is

$$M = \iint_R 3\, dA = \int_0^1 \int_0^1 3\, dx\, dy = 3 \int_0^1 \int_0^1 dx\, dy = 3 \text{ g}.$$

This is all very straightforward; note that all we really did was find the area of the lamina and multiply it by the constant density of $3 \frac{\text{g}}{\text{cm}^2}$.

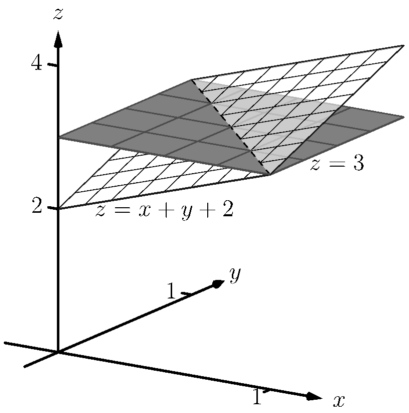


Figure 15.4.6 Graphing the density functions in Examples 15.4.3 and Example 15.4.5

Example 15.4.5 Finding the mass of a lamina with variable density.

Find the mass of a square lamina, represented by the unit square with lower lefthand corner at the origin (see Figure 15.4.4), with variable density $\delta(x, y) = (x + y + 2)$ g/cm^2.

Solution. The variable density δ, in this example, is very uniform, giving a density of 3 in the center of the square and changing linearly. A graph of $\delta(x, y)$ can be seen in Figure 15.4.6; notice how "same amount" of density is above $z = 3$ as below. We'll comment on the significance of this momentarily.

The mass M is found by integrating $\delta(x, y)$ over R. The order of integration is not important; we choose $dx\, dy$ arbitrarily. Thus:

$$\begin{aligned} M = \iint_R (x + y + 2)\, dA &= \int_0^1 \int_0^1 (x + y + 2)\, dx\, dy \\ &= \int_0^1 \left(\frac{1}{2}x^2 + x(y + 2) \right) \Big|_0^1 dy \\ &= \int_0^1 \left(\frac{5}{2} + y \right) dy \end{aligned}$$

$$= \left(\frac{5}{2}y + \frac{1}{2}y^2\right)\Big|_0^1$$
$$= 3\,\text{g}.$$

It turns out that since the density of the lamina is so uniformly distributed "above and below" $z = 3$ that the mass of the lamina is the same as if it had a constant density of 3. The density functions in Examples 15.4.3 and Example 15.4.5 are graphed in Figure 15.4.6, which illustrates this concept.

Example 15.4.7 Finding the weight of a lamina with variable density.

Find the weight of the lamina represented by the disk with radius 2 ft, centered at the origin, with density function $\delta(x, y) = (x^2+y^2+1)$ lb/ft^2. Compare this to the weight of the lamina with the same shape and density $\delta(x, y) = (2\sqrt{x^2 + y^2} + 1)$ lb/ft^2.

Solution. A direct application of Definition 15.4.2 states that the weight of the lamina is $\iint_R \delta(x, y)\, dA$. Since our lamina is in the shape of a circle, it makes sense to approach the double integral using polar coordinates.

The density function $\delta(x, y) = x^2 + y^2 + 1$ becomes

$$\delta(r, \theta) = (r\cos(\theta))^2 + (r\sin(\theta))^2 + 1 = r^2 + 1.$$

The circle is bounded by $0 \le r \le 2$ and $0 \le \theta \le 2\pi$. Thus the weight W is:

$$\begin{aligned} W &= \int_0^{2\pi}\int_0^2 (r^2 + 1)r\, dr\, d\theta \\ &= \int_0^{2\pi} \left(\frac{1}{4}r^4 + \frac{1}{2}r^2\right)\Big|_0^2 d\theta \\ &= \int_0^{2\pi} (6)\, d\theta \\ &= 12\pi \approx 37.70 \text{ lb}. \end{aligned}$$

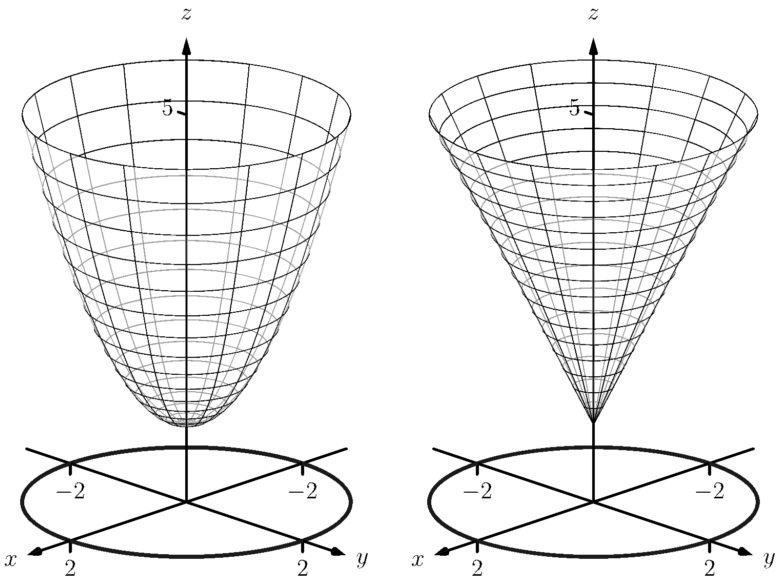
Now compare this with the density function $\delta(x, y) = 2\sqrt{x^2 + y^2} + 1$. Converting this to polar coordinates gives

$$\delta(r, \theta) = 2\sqrt{(r\cos(\theta))^2 + (r\sin(\theta))^2} + 1 = 2r + 1.$$

Thus the weight W is:

$$\begin{aligned} W &= \int_0^{2\pi}\int_0^2 (2r + 1)r\, dr\, d\theta \\ &= \int_0^{2\pi} (\frac{2}{3}r^3 + \frac{1}{2}r^2)\Big|_0^2 d\theta \\ &= \int_0^{2\pi} \left(\frac{22}{3}\right)\, d\theta \\ &= \frac{44}{3}\pi \approx 46.08 \text{ lb}. \end{aligned}$$

One would expect different density functions to return different weights, as we have here. The density functions were chosen, though, to be similar: each gives a density of 1 at the origin and a density of 5 at the outside edge of the circle, as seen in Figure 15.4.8.



(a) **(b)**

Figure 15.4.8 Graphing the density functions in Example 15.4.7. In (a) is the density function $\delta(x,y) = x^2+y^2+1$; in (b) is $\delta(x,y) = 2\sqrt{x^2+y^2}+1$

Notice how $x^2+y^2+1 \leq 2\sqrt{x^2+y^2}+1$ over the circle; this results in less weight.

Plotting the density functions can be useful as our understanding of mass can be related to our understanding of "volume under a surface." We interpreted $\iint_R f(x,y)\,dA$ as giving the volume under f over R; we can understand $\iint_R \delta(x,y)\,dA$ in the same way. The "volume" under δ over R is actually mass; by compressing the "volume" under δ onto the xy-plane, we get "more mass" in some areas than others — i.e., areas of greater density.

Knowing the mass of a lamina is one of several important measures. Another is the *center of mass*, which we discuss next.

15.4.2 Center of Mass

Consider a disk of radius 1 with uniform density. It is common knowledge that the disk will balance on a point if the point is placed at the center of the disk. What if the disk does not have a uniform density? Through trial-and-error, we should still be able to find a spot on the disk at which the disk will balance on a point. This balance point is referred to as the *center of mass*, or *center of gravity*. It is though all the mass is "centered" there. In fact, if the disk has a mass of 3 kg, the disk will behave physically as though it were a point-mass of 3 kg located at its center of mass. For instance, the disk will naturally spin with an axis through its center of mass (which is why it is important to "balance" the tires of your car: if they are "out of balance", their center of mass will be outside of the axle and it will shake terribly).

We find the center of mass based on the principle of a *weighted average*. Consider a college class in which your homework average is 90%, your test average is 73%, and your final exam grade is an 85%. Experience tells us that our final grade *is not* the *average* of these three grades: that is, it is not:

$$\frac{0.9 + 0.73 + 0.85}{3} \approx 0.837 = 83.7$$

That is, you are probably not pulling a B in the course. Rather, your grades are *weighted*. Let's say the homework is worth 10% of the grade, tests are 60% and the exam is 30%. Then your final grade is:

$$(0.1)(0.9) + (0.6)(0.73) + (0.3)(0.85) = 0.783 = 78.3$$

Each grade is multiplied by a *weight*.

In general, given values $x_1, x_2, \ldots, x_n$ and weights $w_1, w_2, \ldots, w_n$, the weighted average of the n values is

$$\sum_{i=1}^{n} w_i x_i \Bigg/ \sum_{i=1}^{n} w_i.$$

In the grading example above, the sum of the weights 0.1, 0.6 and 0.3 is 1, so we don't see the division by the sum of weights in that instance.

How this relates to center of mass is given in the following theorem.

Theorem 15.4.9 Center of Mass of Discrete Linear System.

Let point masses $m_1, m_2, \ldots, m_n$ be distributed along the x-axis at locations $x_1, x_2, \ldots, x_n$, respectively. The center of mass $\overline{x}$ of the system is located at

$$\overline{x} = \sum_{i=1}^{n} m_i x_i \Bigg/ \sum_{i=1}^{n} m_i.$$

Example 15.4.10 Finding the center of mass of a discrete linear system.

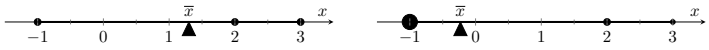
1. Point masses of $2\,\text{g}$ are located at $x = -1$, $x = 2$ and $x = 3$ are connected by a thin rod of negligible weight. Find the center of mass of the system.
2. Point masses of $10\,\text{g}$, $2\,\text{g}$ and $1\,\text{g}$ are located at $x = -1$, $x = 2$ and $x = 3$, respectively, are connected by a thin rod of negligible weight. Find the center of mass of the system.

Solution.

1. Following Theorem 15.4.9, we compute the center of mass as:

$$\overline{x} = \frac{2(-1) + 2(2) + 2(3)}{2 + 2 + 2} = \frac{4}{3} = 1.\overline{3}.$$

So the system would balance on a point placed at $x = 4/3$, as illustrated in Figure 15.4.11(a).



(a) **(b)**

Figure 15.4.11 Illustrating point masses along a thin rod and the center of mass

2. Again following Theorem 15.4.9, we find:

$$\overline{x} = \frac{10(-1) + 2(2) + 1(3)}{10 + 2 + 1} = \frac{-3}{13} \approx -0.23.$$

Placing a large weight at the left hand side of the system moves the center of mass left, as shown in Figure 15.4.11(b).

In a discrete system (i.e., mass is located at individual points, not along a continuum) we find the center of mass by dividing the mass into a *moment* of the system. In general, a moment is a weighted measure of distance from a particular point or line. In the case described by Theorem 15.4.9, we are finding a weighted measure of distances from the y-axis, so we refer to this as *the moment about the y-axis*, represented by M_y. Letting M be the total mass of the system, we have $\overline{x} = M_y/M$.

We can extend the concept of the center of mass of discrete points along a line to the center of mass of discrete points in the plane rather easily. To do so, we define some terms then give a theorem.

Definition 15.4.12 Moments about the x and y Axes.

Let point masses $m_1, m_2, \ldots, m_n$ be located at points

$$(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n),$$

respectively, in the xy-plane.

1. The *moment about the y-axis*, M_y, is $M_y = \sum_{i=1}^{n} m_i x_i$.
2. The *moment about the x-axis*, M_x, is $M_x = \sum_{i=1}^{n} m_i y_i$.

One can think that these definitions are "backwards" as M_y sums up "x" distances. But remember, "x" distances are measurements of distance from the y-axis, hence defining the moment about the y-axis.

We now define the center of mass of discrete points in the plane.

Theorem 15.4.13 Center of Mass of Discrete Planar System.

Let point masses $m_1, m_2, \ldots, m_n$ be located at points

$$(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n),$$

respectively, in the xy-plane, and let $M = \sum_{i=1}^{n} m_i$.

The center of mass of the system is at $(\overline{x}, \overline{y})$, where

$$\overline{x} = \frac{M_y}{M} \text{ and } \overline{y} = \frac{M_x}{M}.$$

Example 15.4.14 Finding the center of mass of a discrete planar system.

Let point masses of 1 kg, 2 kg and 5 kg be located at points $(2, 0)$, $(1, 1)$ and $(3, 1)$, respectively, and are connected by thin rods of negligible weight. Find the center of mass of the system.

Solution. We follow Theorem 15.4.13 and Definition 15.4.12 to find M, M_x and M_y:

$M = 1 + 2 + 5 = 8$ kg.

$$\begin{aligned} M_x &= \sum_{i=1}^{n} m_i y_i \\ &= 1(0) + 2(1) + 5(1) \\ &= 7. \end{aligned} \qquad \begin{aligned} M_y &= \sum_{i=1}^{n} m_i x_i \\ &= 1(2) + 2(1) + 5(3) \\ &= 19. \end{aligned}$$

Thus the center of mass is $(\overline{x}, \overline{y}) = \left(\frac{M_y}{M}, \frac{M_x}{M}\right) = \left(\frac{19}{8}, \frac{7}{8}\right) = (2.375, 0.875)$, illustrated in Figure 15.4.15.

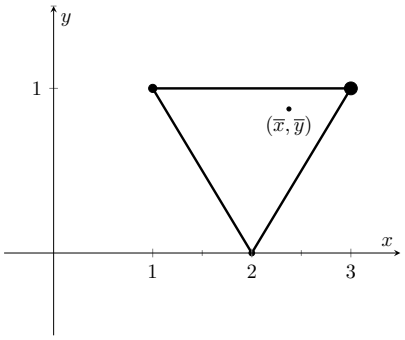


Figure 15.4.15 Illustrating the center of mass of a discrete planar system in Example 15.4.14

We finally arrive at our true goal of this section: finding the center of mass of a lamina with variable density. While the above measurement of center of mass is interesting, it does not directly answer more realistic situations where we need to find the center of mass of a contiguous region. However, understanding the discrete case allows us to approximate the center of mass of a planar lamina; using calculus, we can refine the approximation to an exact value.

We begin by representing a planar lamina with a region R in the xy-plane with density function $\delta(x, y)$. Partition R into n subdivisions, each with area ΔA_i. As done before, we can approximate the mass of the ith subregion with $\delta(x_i, y_i)\Delta A_i$, where (x_i, y_i) is a point inside the ith subregion. We can approximate the moment of this subregion about the y-axis with $x_i\delta(x_i, y_i)\Delta A_i$ — that is, by multiplying the approximate mass of the region by its approximate distance from the y-axis. Similarly, we can approximate the moment about the x-axis with $y_i\delta(x_i, y_i)\Delta A_i$. By summing over all subregions, we have:

$$\text{mass: } M \approx \sum_{i=1}^{n} \delta(x_i, y_i)\Delta A_i \text{ (as seen before)}$$

$$\text{moment about the } x\text{-axis: } M_x \approx \sum_{i=1}^{n} y_i\delta(x_i, y_i)\Delta A_i$$

$$\text{moment about the } y\text{-axis: } M_y \approx \sum_{i=1}^{n} x_i\delta(x_i, y_i)\Delta A_i$$

By taking limits, where size of each subregion shrinks to 0 in both the x and y directions, we arrive at the double integrals given in the following theorem.

Theorem 15.4.16 Center of Mass of a Planar Lamina, Moments.

Let a planar lamina be represented by a closed, bounded region R in the xy-plane with density function $\delta(x, y)$.

1. *mass:* $M = \iint_R \delta(x, y)\, dA$
2. *moment about the x-axis:* $M_x = \iint_R y\delta(x, y)\, dA$
3. *moment about the y-axis:* $M_y = \iint_R x\delta(x, y)\, dA$
4. *The center of mass of the lamina is*

$$(\overline{x}, \overline{y}) = \left(\frac{M_y}{M}, \frac{M_x}{M}\right).$$

We start our practice of finding centers of mass by revisiting some of the lamina used previously in this section when finding mass. We will just set up the integrals needed to compute M, M_x and M_y and leave the details of the integration to the reader.

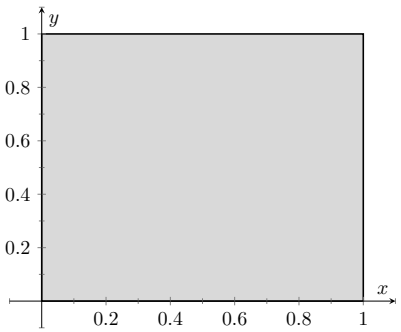


Figure 15.4.18 A region R representing a lamina in Example 15.4.3

Example 15.4.17 Finding the center of mass of a lamina.

Find the center mass of a square lamina, with side length 1, with a density of $\delta = 3$ g/cm^2. (Note: this is the lamina from Example 15.4.3.)

Solution. We represent the lamina with a square region in the plane as shown in Figure 15.4.18 as done previously.

Following Theorem 15.4.16, we find M, M_x and M_y:

$$M = \iint_R 3\, dA = \int_0^1 \int_0^1 3\, dx\, dy = 3 \text{ g}$$
$$M_x = \iint_R 3y\, dA = \int_0^1 \int_0^1 3y\, dx\, dy = 3/2 = 1.5$$
$$M_y = \iint_R 3x\, dA = \int_0^1 \int_0^1 3x\, dx\, dy = 3/2 = 1.5.$$

Thus the center of mass is $(\overline{x}, \overline{y}) = \left(\frac{M_y}{M}, \frac{M_x}{M}\right) = (1.5/3, 1.5/3) = (0.5, 0.5)$. This is what we should have expected: the center of mass of a square with constant density is the center of the square.

Example 15.4.19 Finding the center of mass of a lamina.

Find the center of mass of a square lamina, represented by the unit square with lower lefthand corner at the origin (see Figure 15.4.18), with variable density $\delta(x, y) = (x + y + 2)$ g/cm^2. (Note: this is the lamina from Example 15.4.5.)

Solution. We follow Theorem 15.4.16, to find M, M_x and M_y:

$$M = \iint_R (x+y+2)\,dA = \int_0^1\int_0^1 (x+y+2)\,dx\,dy = 3\,\text{g}$$
$$M_x = \iint_R y(x+y+2)\,dA = \int_0^1\int_0^1 y(x+y+2)\,dx\,dy = \frac{19}{12}$$
$$M_y = \iint_R x(x+y+2)\,dA = \int_0^1\int_0^1 x(x+y+2)\,dx\,dy = \frac{19}{12}.$$

Thus the center of mass is

$$(\overline{x},\overline{y}) = \left(\frac{M_y}{M},\frac{M_x}{M}\right) = \left(\frac{19}{36},\frac{19}{36}\right) \approx (0.528, 0.528).$$

While the mass of this lamina is the same as the lamina in the previous example, the greater density found with greater x and y values pulls the center of mass from the center slightly towards the upper righthand corner.

Example 15.4.20 Finding the center of mass of a lamina.

Find the center of mass of the lamina represented by the circle with radius 2 ft, centered at the origin, with density function $\delta(x,y) = (x^2+y^2+1)$ lb/ft^2. (Note: this is one of the lamina used in Example 15.4.7.)

Solution. As done in Example 15.4.7, it is best to describe R using polar coordinates. Thus when we compute M_y, we will integrate not $x\delta(x,y) = x(x^2+y^2+1)$, but rather $\big(r\cos(\theta)\big)\delta(r\cos(\theta), r\sin(\theta)) = \big(r\cos(\theta)\big)\big(r^2+1\big)$. We compute M, M_x and M_y:

$$M = \int_0^{2\pi}\int_0^2 (r^2+1)r\,dr\,d\theta = 12\pi \approx 37.7\,\text{lb}$$
$$M_x = \int_0^{2\pi}\int_0^2 (r\sin(\theta))(r^2+1)r\,dr\,d\theta = 0$$
$$M_y = \int_0^{2\pi}\int_0^2 (r\cos(\theta))(r^2+1)r\,dr\,d\theta = 0.$$

Since R and the density of R are both symmetric about the x and y axes, it should come as no big surprise that the moments about each axis is 0. Thus the center of mass is $(\overline{x},\overline{y}) = (0,0)$.

Example 15.4.21 Finding the center of mass of a lamina.

Find the center of mass of the lamina represented by the region R shown in Figure 15.4.22, half an annulus with outer radius 6 ft and inner radius 5 ft, with constant density $2\,\frac{\text{lb}}{\text{ft}^2}$.

Solution. Once again it will be useful to represent R in polar coordinates. Using the description of R and/or the illustration, we see that R is bounded by $5 \le r \le 6$ and $0 \le \theta \le \pi$. As the lamina is symmetric about the y-axis, we should expect $M_y = 0$. We compute M, M_x and

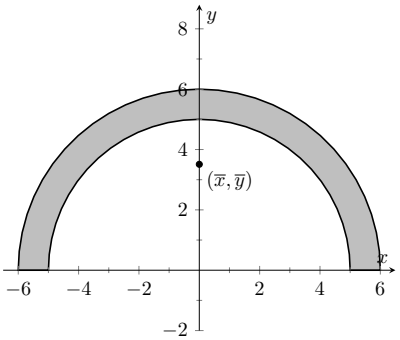


Figure 15.4.22 Illustrating the region R in Example 15.4.21

M_y:

$$M = \int_0^\pi \int_5^6 (2)r\,dr\,d\theta = 11\pi \text{ lb}$$
$$M_x = \int_0^\pi \int_5^6 (r\sin(\theta))(2)r\,dr\,d\theta = \frac{364}{3} \approx 121.33$$
$$M_y = \int_0^\pi \int_5^6 (r\cos(\theta))(2)r\,dr\,d\theta = 0.$$

Thus the center of mass is $(\overline{x}, \overline{y}) = \left(0, \frac{364}{33\pi}\right) \approx (0, 3.51)$. The center of mass is indicated in Figure 15.4.22; note how it lies outside of R!

This section has shown us another use for iterated integrals beyond finding area or signed volume under the curve. While there are many uses for iterated integrals, we give one more application in the following section: computing surface area.

15.4.3 Exercises

Terms and Concepts

1. Why is it easy to use "mass" and "weight" interchangeably, even though they are different measures?

2. Given a point (x, y), the value of x is a measure of distance from the _____-axis.

3. We can think of $\iint_R \, dm$ as meaning "sum up lots of _____"

4. What is a "discrete planar system?"

5. Why does M_x use $\iint_R y\delta(x, y)\, dA$ instead of $\iint_R x\delta(x, y)\, dA$; that is, why do we use "y" and not "x"?

6. Describe a situation where the center of mass of a lamina does not lie within the region of the lamina itself.

Problems

Exercise Group. In the following exercises, point masses are given along a line or in the plane. Find the center of mass $\overline{x}$ or $(\overline{x}, \overline{y})$, as appropriate. (All masses are in grams and distances are in cm.)

7. $m_1 = 4$ at $x = 1$; $m_2 = 3$ at $x = 3$; $m_3 = 5$ at $x = 10$

8. $m_1 = 2$ at $x = -3$; $m_2 = 2$ at $x = -1$; $m_3 = 3$ at $x = 0$;$m_4 = 3$ at $x = 7$

9. $m_1 = 2$ at $(-2, -2)$; $m_2 = 2$ at $(2, -2)$; $m_3 = 20$ at $(0, 4)$

10. $m_1 = 1$ at $(-1, -1)$; $m_2 = 2$ at $(-1, 1)$; $m_3 = 2$ at $(1, 1)$;$m_4 = 1$ at $(1, -1)$

Exercise Group. In the following exercises, find the mass/weight of the lamina described by the region R in the plane and its density function $\delta(x, y)$.

11. R is the rectangle with corners $(1, -3)$, $(1, 2)$, $(7, 2)$ and $(7, -3)$; $\delta(x, y) = 5$ g/cm^2

12. R is the rectangle with corners $(1, -3)$, $(1, 2)$, $(7, 2)$ and $(7, -3)$; $\delta(x, y) = (x + y^2)$ g/cm^2

13. R is the triangle with corners $(-1, 0)$, $(1, 0)$, and $(0, 1)$; $\delta(x, y) = 2$ lb/in^2

14. R is the triangle with corners $(0, 0)$, $(1, 0)$, and $(0, 1)$; $\delta(x, y) = (x^2 + y^2 + 1)$ lb/in^2

15. R is the disk centered at the origin with radius 2; $\delta(x, y) = (x + y + 4)$ kg/m^2

16. R is the circle sector bounded by $x^2 + y^2 = 25$ in the first quadrant; $\delta(x, y) = (\sqrt{x^2 + y^2} + 1)$ kg/m^2

17. R is the annulus in the first and second quadrants bounded by $x^2 + y^2 = 9$ and $x^2 + y^2 = 36$; $\delta(x, y) = 4$ lb/ft^2

18. R is the annulus in the first and second quadrants bounded by $x^2 + y^2 = 9$ and $x^2 + y^2 = 36$; $\delta(x, y) = \sqrt{x^2 + y^2}$ lb/ft^2

Exercise Group. In the following exercises, find the center of mass of the lamina described by the region R in the plane and its density function $\delta(x, y)$.

Note: these are the same lamina as in Exercises 15.4.11 — Exercise 15.4.18.

19. R is the rectangle with corners $(1, -3)$, $(1, 2)$, $(7, 2)$ and $(7, -3)$; $\delta(x, y) = 5$ g/cm^2

20. R is the rectangle with corners $(1, -3)$, $(1, 2)$, $(7, 2)$ and $(7, -3)$; $\delta(x, y) = (x + y^2)$ g/cm^2

21. R is the triangle with corners $(-1, 0)$, $(1, 0)$, and $(0, 1)$; $\delta(x, y) = 2$ lb/in^2

22. R is the triangle with corners $(0, 0)$, $(1, 0)$, and $(0, 1)$; $\delta(x, y) = (x^2 + y^2 + 1)$ lb/in^2

23. R is the disk centered at the origin with radius 2; $\delta(x, y) = (x + y + 4)$ kg/m^2

24. R is the circle sector bounded by $x^2 + y^2 = 25$ in the first quadrant; $\delta(x, y) = (\sqrt{x^2 + y^2} + 1)$ kg/m^2

25. R is the annulus in the first and second quadrants bounded by $x^2 + y^2 = 9$ and $x^2 + y^2 = 36$; $\delta(x, y) = 4$ lb/ft^2

26. R is the annulus in the first and second quadrants bounded by $x^2 + y^2 = 9$ and $x^2 + y^2 = 36$; $\delta(x, y) = \sqrt{x^2 + y^2}$ lb/ft^2

Exercise Group. *The moment of inertia* I is a measure of the tendency of a lamina to resist rotating about an axis or continue to rotate about an axis. I_x is the moment of inertia about the x-axis, I_y is the moment of inertia about the y-axis, and I_O is the moment of inertia about the origin. These are computed as follows:

- $I_x = \iint_R y^2 \, dm$
- $I_y = \iint_R x^2 \, dm$
- $I_O = \iint_R \left(x^2 + y^2\right) dm$

In the following exercises, a lamina corresponding to a planar region R is given with a mass of 16 units. For each, compute I_x, I_y and I_O.

27. R is the 4×4 square with corners at $(-2, -2)$ and $(2, 2)$ with density $\delta(x, y) = 1$.

28. R is the 8×2 rectangle with corners at $(-4, -1)$ and $(4, 1)$ with density $\delta(x, y) = 1$.

29. R is the 4×2 rectangle with corners at $(-2, -1)$ and $(2, 1)$ with density $\delta(x, y) = 2$.

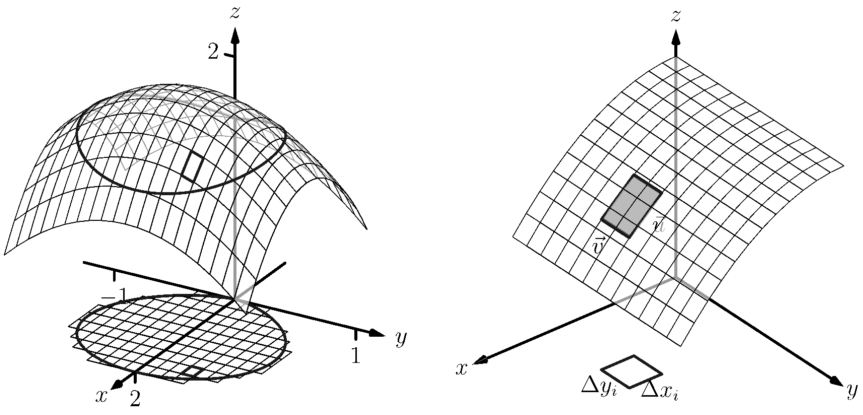
30. R is the disk with radius 2 centered at the origin with density $\delta(x, y) = 4/\pi$.

15.5 Surface Area

In Section 7.4 we used definite integrals to compute the arc length of plane curves of the form $y = f(x)$. We later extended these ideas to compute the arc length of plane curves defined by parametric or polar equations.

The natural extension of the concept of "arc length over an interval" to surfaces is "surface area over a region."

Consider the surface $z = f(x, y)$ over a region R in the xy-plane, shown in Figure 15.5.1(a). Because of the domed shape of the surface, the surface area will be greater than that of the area of the region R. We can find this area using the same basic technique we have used over and over: we'll make an approximation, then using limits, we'll refine the approximation to the exact value.



(a) **(b)**

Figure 15.5.1 Developing a method of computing surface area

As done to find the volume under a surface or the mass of a lamina, we subdivide R into n subregions. Here we subdivide R into rectangles, as shown in the figure. One such subregion is outlined in the figure, where the rectangle has dimensions Δx_i and Δy_i, along with its corresponding region on the surface.

In part Figure 15.5.1(b) of the figure, we zoom in on this portion of the surface. When Δx_i and Δy_i are small, the function is approximated well by the tangent plane at any point (x_i, y_i) in this subregion, which is graphed in part Figure 15.5.1(b). In fact, the tangent plane approximates the function so well that in this figure, it is virtually indistinguishable from the surface itself! Therefore we can approximate the surface area S_i of this region of the surface with the area T_i of the corresponding portion of the tangent plane.

This portion of the tangent plane is a parallelogram, defined by sides $\vec{u}$ and $\vec{v}$, as shown. One of the applications of the cross product from Section 12.4 is that the area of this parallelogram is $\|\vec{u} \times \vec{v}\|$. Once we can determine $\vec{u}$ and $\vec{v}$, we can determine the area.

The vector $\vec{u}$ is tangent to the surface in the direction of x, therefore, from Section 14.4, $\vec{u}$ is parallel to $\langle 1, 0, f_x(x_i, y_i)\rangle$. The x-displacement of $\vec{u}$ is Δx_i, so we know that $\vec{u} = \Delta x_i \langle 1, 0, f_x(x_i, y_i)\rangle$. Similar logic shows that $\vec{v} = \Delta y_i \langle 0, 1, f_y(x_i, y_i)\rangle$. Thus:

$$\begin{aligned}
\text{surface area } S_i &\approx \text{ area of } T_i \\
&= \|\vec{u} \times \vec{v}\| \\
&= \|\Delta x_i \langle 1, 0, f_x(x_i, y_i)\rangle \times \Delta y_i \langle 0, 1, f_y(x_i, y_i)\rangle\| \\
&= \sqrt{1 + f_x(x_i, y_i)^2 + f_y(x_i, y_i)^2}\Delta x_i \Delta y_i.
\end{aligned}$$

Note that $\Delta x_i \Delta y_i = \Delta A_i$, the area of the ith subregion.
Summing up all n of the approximations to the surface area gives

$$\text{surface area over } R \approx \sum_{i=1}^{n} \sqrt{1 + f_x(x_i, y_i)^2 + f_y(x_i, y_i)^2}\,\Delta A_i.$$

Once again take a limit as all of the Δx_i and Δy_i shrink to 0; this leads to a double integral.

As done before, we think of "$\iint_R \, dS$" as meaning "sum up lots of little surface areas over R."

The concept of surface area is *defined* here, for while we already have a notion of the area of a region in the *plane*, we did not yet have a solid grasp of what "the area of a surface in *space*" means.

Definition 15.5.2 Surface Area.

Let $z = f(x, y)$ where f_x and f_y are continuous over a closed, bounded region R. The surface area S over R is

$$\begin{aligned} S &= \iint_R dS \\ &= \iint_R \sqrt{1 + f_x(x, y)^2 + f_y(x, y)^2}\, dA. \end{aligned}$$

We test this definition by using it to compute surface areas of known surfaces. We start with a triangle.

Example 15.5.3 Finding the surface area of a plane over a triangle.

Let $f(x, y) = 4 - x - 2y$, and let R be the region in the plane bounded by $x = 0$, $y = 0$ and $y = 2 - x/2$, as shown in Figure 15.5.4. Find the surface area of f over R.

Solution. We follow Definition 15.5.2. We start by noting that $f_x(x, y) = -1$ and $f_y(x, y) = -2$. To define R, we use bounds $0 \le y \le 2 - x/2$ and $0 \le x \le 4$. Therefore

$$\begin{aligned} S &= \iint_R dS \\ &= \int_0^4 \int_0^{2-x/2} \sqrt{1 + (-1)^2 + (-2)^2}\, dy\, dx \\ &= \int_0^4 \sqrt{6}\left(2 - \frac{x}{2}\right) dx \\ &= 4\sqrt{6}. \end{aligned}$$

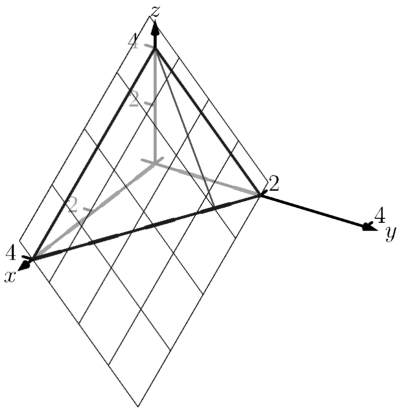


Figure 15.5.4 Finding the area of a triangle in space in Example 15.5.3

Because the surface is a triangle, we can figure out the area using geometry. Considering the base of the triangle to be the side in the xy-plane, we find the length of the base to be $\sqrt{20}$. We can find the height using our knowledge of vectors: let $\vec{u}$ be the side in the xz-plane and let $\vec{v}$ be the side in the xy-plane. The height is then $\|\vec{u} - \text{proj}_{\vec{v}}\, \vec{u}\| = 4\sqrt{6/5}$. Geometry states that the area is thus

$$\frac{1}{2} \cdot 4\sqrt{6/5} \cdot \sqrt{20} = 4\sqrt{6}.$$

We affirm the validity of our formula.

It is "common knowledge" that the surface area of a sphere of radius r is $4\pi r^2$. We confirm this in the following example, which involves using our formula with polar coordinates.

Example 15.5.5 The surface area of a sphere.

Find the surface area of the sphere with radius a centered at the origin, whose top hemisphere has equation $f(x,y) = \sqrt{a^2 - x^2 - y^2}$.

Solution. We start by computing partial derivatives and find

$$f_x(x,y) = \frac{-x}{\sqrt{a^2 - x^2 - y^2}} \text{ and } f_y(x,y) = \frac{-y}{\sqrt{a^2 - x^2 - y^2}}.$$

As our function f only defines the top upper hemisphere of the sphere, we double our surface area result to get the total area:

$$\begin{aligned} S &= 2\iint_R \sqrt{1 + f_x(x,y)^2 + f_y(x,y)^2}\, dA \\ &= 2\iint_R \sqrt{1 + \frac{x^2+y^2}{a^2 - x^2 - y^2}}\, dA. \end{aligned}$$

The region R that we are integrating over is bounded by the circle, centered at the origin, with radius a: $x^2 + y^2 = a^2$. Because of this region, we are likely to have greater success with our integration by converting to polar coordinates. Using the substitutions $x = r\cos(\theta)$, $y = r\sin(\theta)$, $dA = r\,dr\,d\theta$ and bounds $0 \le \theta \le 2\pi$ and $0 \le r \le a$, we have:

$$\begin{aligned} S &= 2\int_0^{2\pi}\int_0^a \sqrt{1 + \frac{r^2\cos^2(\theta) + r^2\sin^2(\theta)}{a^2 - r^2\cos^2(\theta) - r^2\sin^2(\theta)}}\, r\,dr\,d\theta \\ &= 2\int_0^{2\pi}\int_0^a r\sqrt{1 + \frac{r^2}{a^2 - r^2}}\, dr\,d\theta \\ &= 2\int_0^{2\pi}\int_0^a r\sqrt{\frac{a^2}{a^2 - r^2}}\, dr\,d\theta. \end{aligned} \tag{15.5.1}$$

Apply substitution $u = a^2 - r^2$ and integrate the inner integral, giving

$$\begin{aligned} &= 2\int_0^{2\pi} a^2\, d\theta \\ &= 4\pi a^2. \end{aligned}$$

Our work confirms our previous formula.

The inner integral in Equation (15.5.1) is an improper integral, as the integrand of $\int_0^a r\sqrt{\frac{a^2}{a^2-r^2}}\, dr$ is not defined at $r = a$. To properly evaluate this integral, one must use the techniques of Section 6.5.

The reason this need arises is that the function

$$f(x,y) = \sqrt{a^2 - x^2 - y^2}$$

fails the requirements of Definition 15.5.2, as f_x and f_y are not continuous on the boundary of the circle $x^2 + y^2 = a^2$.

The computation of the surface area is still valid. The definition makes stronger requirements than necessary in part to avoid the use of improper integration, as when f_x and/or f_y are not continuous, the resulting improper integral may not converge. Since the improper integral does converge in this example, the surface area is accurately computed.

Example 15.5.6 Finding the surface area of a cone.

The general formula for a right cone with height h and base radius a is

$$f(x,y) = h - \frac{h}{a}\sqrt{x^2 + y^2},$$

shown in Figure 15.5.7. Find the surface area of this cone.

Solution. We begin by computing partial derivatives.

$$f_x(x,y) = -\frac{xh}{a\sqrt{x^2+y^2}} \text{ and } f_y(x,y) = -\frac{yh}{a\sqrt{x^2+y^2}}.$$

Since we are integrating over the disk bounded by $x^2 + y^2 = a^2$, we again use polar coordinates. Using the standard substitutions, our

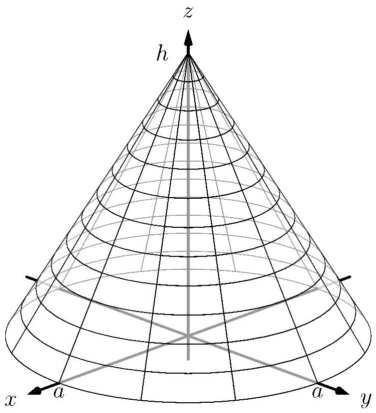


Figure 15.5.7 Finding the surface area of a cone in Example 15.5.6

integrand becomes

$$\sqrt{1+\left(\frac{hr\cos(\theta)}{a\sqrt{r^2}}\right)^2+\left(\frac{hr\sin(\theta)}{a\sqrt{r^2}}\right)^2}.$$

This may look intimidating at first, but there are lots of simple simplifications to be done. It amazingly reduces to just

$$\sqrt{1+\frac{h^2}{a^2}}=\frac{1}{a}\sqrt{a^2+h^2}.$$

Our polar bounds are $0\le\theta\le 2\pi$ and $0\le r\le a$. Thus

$$\begin{aligned}S&=\int_0^{2\pi}\int_0^a r\frac{1}{a}\sqrt{a^2+h^2}\,dr\,d\theta\\&=\int_0^{2\pi}\left(\frac{1}{2}r^2\frac{1}{a}\sqrt{a^2+h^2}\right)\Big|_0^a d\theta\\&=\int_0^{2\pi}\frac{1}{2}a\sqrt{a^2+h^2}\,d\theta\\&=\pi a\sqrt{a^2+h^2}.\end{aligned}$$

This matches the formula found in the back of this text.

Note that once again f_x and f_y are not continuous on the domain of f, as both are undefined at $(0,0)$. (A similar problem occurred in the previous example.) Once again the resulting improper integral converges and the computation of the surface area is valid.

Example 15.5.8 Finding surface area over a region.

Find the area of the surface $f(x,y)=x^2-3y+3$ over the region R bounded by $-x\le y\le x$, $0\le x\le 4$, as pictured in Figure 15.5.9.

Solution. It is straightforward to compute $f_x(x,y)=2x$ and $f_y(x,y)=-3$. Thus the surface area is described by the double integral

$$\iint_R\sqrt{1+(2x)^2+(-3)^2}\,dA=\iint_R\sqrt{10+4x^2}\,dA.$$

As with integrals describing arc length, double integrals describing surface area are in general hard to evaluate directly because of the square-root. This particular integral can be easily evaluated, though, with judicious choice of our order of integration.

Integrating with order $dx\,dy$ requires us to evaluate $\int\sqrt{10+4x^2}\,dx$. This can be done, though it involves Integration By Parts and $\sinh^{-1}(x)$. Integrating with order $dy\,dx$ has as its first integral $\int\sqrt{10+4x^2}\,dy$, which is easy to evaluate: it is simply $y\sqrt{10+4x^2}+C$. So we proceed with the order $dy\,dx$; the bounds are already given in the statement of the problem.

$$\begin{aligned}\iint_R\sqrt{10+4x^2}\,dA&=\int_0^4\int_{-x}^x\sqrt{10+4x^2}\,dy\,dx\\&=\int_0^4\left(y\sqrt{10+4x^2}\right)\Big|_{-x}^x dx\\&=\int_0^4\left(2x\sqrt{10+4x^2}\right)dx.\end{aligned}$$

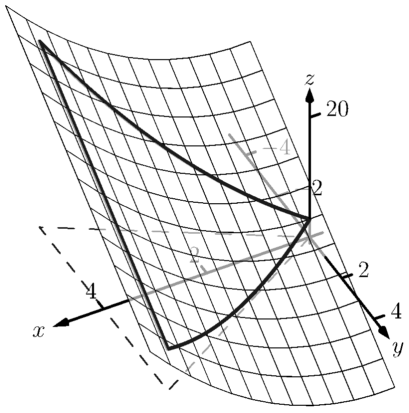


Figure 15.5.9 Graphing the surface in Example 15.5.8

Apply substitution with $u = 10 + 4x^2$:

$$= \left(\frac{1}{6}\left(10 + 4x^2\right)^{3/2}\right)\Big|_0^4$$
$$= \frac{1}{3}\left(37\sqrt{74} - 5\sqrt{10}\right) \approx 100.825 \text{ units}^2.$$

So while the region R over which we integrate has an area of 16 square units, the surface has a much greater area as its z-values change dramatically over R.

In practice, technology helps greatly in the evaluation of such integrals. High powered computer algebra systems can compute integrals that are difficult, or at least time consuming, by hand, and can at the least produce very accurate approximations with numerical methods. In general, just knowing *how* to set up the proper integrals brings one very close to being able to compute the needed value. Most of the work is actually done in just describing the region R in terms of polar or rectangular coordinates. Once this is done, technology can usually provide a good answer.

We have learned how to integrate integrals; that is, we have learned to evaluate double integrals. In the next section, we learn how to integrate double integrals — that is, we learn to evaluate *triple integrals*, along with learning some uses for this operation.

15.5.1 Exercises

Terms and Concepts

1. "Surface area" is analogous to what previously studied concept?

2. To approximate the area of a small portion of a surface, we computed the area of its _____ plane.

3. We interpret $\iint_R dS$ as "sum up lots of little _____ _____."

4. Why is it important to know how to set up a double integral to compute surface area, even if the resulting integral is hard to evaluate?

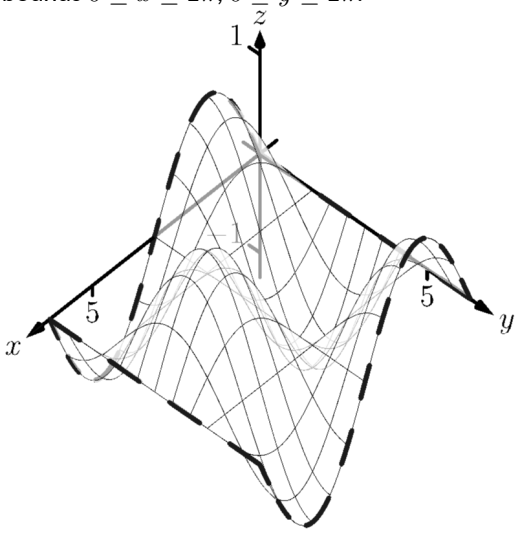
5. Why do $z = f(x, y)$ and $z = g(x, y) = f(x, y) + h$, for some real number h, have the same surface area over a region R?

6. Let $f(x, y)$ be a function defined over a region R and let $g(x, y) = 2f(x, y)$. Why is the surface area of $z = g(x, y)$ over R not twice the surface area of $z = f(x, y)$ over R?

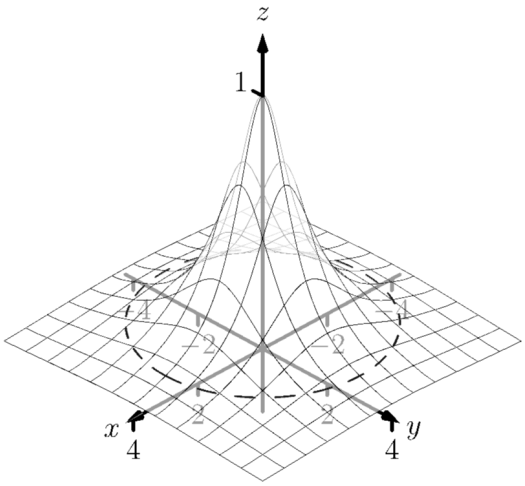
Problems

Exercise Group. In the following exercises, *set up* the iterated integral that computes the surface area of the graph of the given function over the region R.

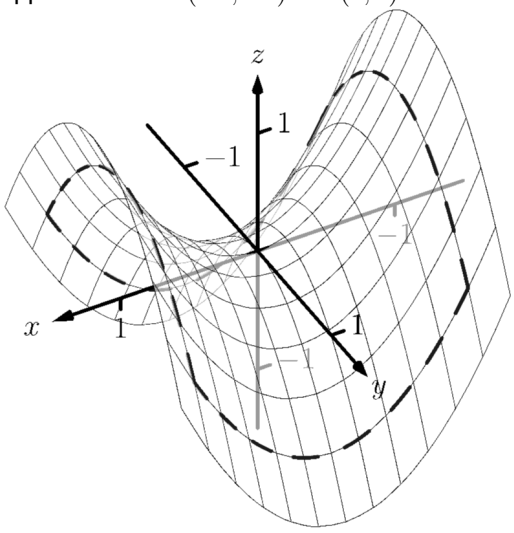
7. $f(x, y) = \sin(x)\cos(y)$; R is the rectangle with bounds $0 \le x \le 2\pi$, $0 \le y \le 2\pi$.



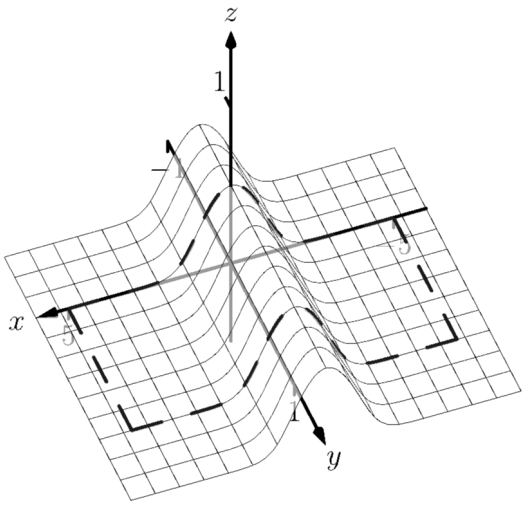
8. $f(x, y) = \dfrac{1}{x^2 + y^2 + 1}$; R is bounded by the circle $x^2 + y^2 = 9$.



9. $f(x,y) = x^2 - y^2$; R is the rectangle with opposite corners $(-1,-1)$ and $(1,1)$.



10. $f(x,y) = \dfrac{1}{e^{x^2}+1}$; R is the rectangle bounded by $-5 \le x \le 5$ and $0 \le y \le 1$.



Exercise Group. In the following exercises, find the area of the given surface over the region R.

11. $z = 3x - 7y + 2$; R is the rectangle with opposite corners $(-1,0)$ and $(1,3)$.

12. $z = 2x + 2y + 2$; R is the triangle with corners $(0,0)$, $(1,0)$ and $(0,1)$.

13. $z = x^2 + y^2 + 10$; R is bounded by the circle $x^2 + y^2 = 16$.

14. $z = -2x + 4y^2 + 7$ over R, the triangle bounded by $y = -x$, $y = x$, $0 \le y \le 1$.

15. $z = x^2 + y$ over R, the triangle bounded by $y = 2x$, $y = 0$ and $x = 2$.

16. $z = \frac{2}{3}x^{3/2} + 2y^{3/2}$ over R, the rectangle with opposite corners $(0,0)$ and $(1,1)$.

17. $z = 10 - 2\sqrt{x^2+y^2}$ over the region R bounded by the circle $x^2 + y^2 = 25$. (This is the cone with height 10 and base radius 5; be sure to compare your result with the known formula.)

18. Find the surface area of the sphere with radius 5 by doubling the surface area of $f(x,y) = \sqrt{25 - x^2 - y^2}$ over the region R bounded by the circle $x^2 + y^2 = 25$. (Be sure to compare your result with the known formula.)

19. Find the surface area of the ellipse formed by restricting the plane $f(x,y) = cx + dy + h$ to the region R bounded by the circle $x^2 + y^2 = 1$, where c, d and h are some constants. Your answer should be given in terms of c and d; why does the value of h not matter?

15.6 Volume Between Surfaces and Triple Integration

15.6.1 Volume between surfaces

We learned in Section 15.2 how to compute the signed volume V under a surface $z = f(x, y)$ over a region R: $V = \iint_R f(x, y)\, dA$. It follows naturally that if $f(x, y) \geq g(x, y)$ on R, then the *volume between* $f(x, y)$ *and* $g(x, y)$ *on* R is

$$V = \iint_R f(x, y)\, dA - \iint_R g(x, y)\, dA = \iint_R \big(f(x, y) - g(x, y)\big)\, dA.$$

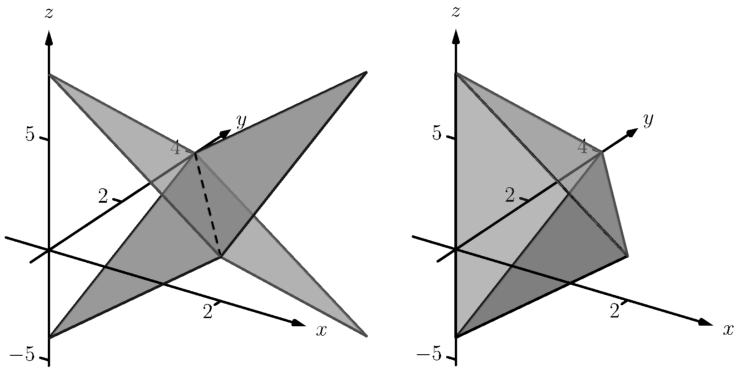
Theorem 15.6.1 Volume Between Surfaces.

Let f and g be continuous functions on a closed, bounded region R, where $f(x, y) \geq g(x, y)$ for all (x, y) in R. The volume V between f and g over R is

$$V = \iint_R \big(f(x, y) - g(x, y)\big)\, dA.$$

Example 15.6.2 Finding volume between surfaces.

Find the volume of the space region bounded by the planes $z = 3x + y - 4$, $z = 8 - 3x - 2y$, $x = 0$ and $y = 0$. In Figure 15.6.3(a) the planes are drawn; in Figure 15.6.3(b), only the defined region is given.



(a) **(b)**

Figure 15.6.3 Finding the volume between the planes given in Example 15.6.2

Solution. We need to determine the region R over which we will integrate. To do so, we need to determine where the planes intersect. They have common z-values when $3x + y - 4 = 8 - 3x - 2y$. Applying a little algebra, we have:

$$\begin{aligned} 3x + y - 4 &= 8 - 3x - 2y \\ 6x + 3y &= 12 \\ 2x + y &= 4 \end{aligned}$$

The planes intersect along the line $2x + y = 4$. Therefore the region R is bounded by $x = 0$, $y = 0$, and $y = 4 - 2x$; we can convert these bounds to integration bounds of $0 \leq x \leq 2$, $0 \leq y \leq 4 - 2x$. Thus

$$V = \iint_R \big(8 - 3x - 2y - (3x + y - 4)\big)\, dA$$

$$= \int_0^2 \int_0^{4-2x} \big(12 - 6x - 3y\big)\, dy\, dx$$
$$= 16 \text{ units}^3.$$

The volume between the surfaces is 16 cubic units.

In the preceding example, we found the volume by evaluating the integral

$$\int_0^2 \int_0^{4-2x} \big(8 - 3x - 2y - (3x + y - 4)\big)\, dy\, dx.$$

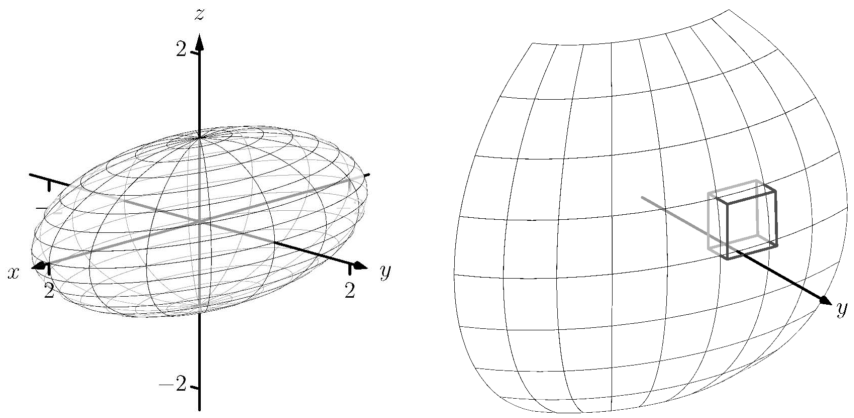
Note how we can rewrite the integrand as an integral, much as we did in Section 15.1:

$$8 - 3x - 2y - (3x + y - 4) = \int_{3x+y-4}^{8-3x-2y} dz.$$

Thus we can rewrite the double integral that finds volume as

$$\int_0^2 \int_0^{4-2x} \big(8-3x-2y-(3x+y-4)\big)\, dy\, dx = \int_0^2 \int_0^{4-2x} \left(\int_{3x+y-4}^{8-3x-2y} dz\right) dy\, dx.$$

This no longer looks like a "double integral," but more like a "triple integral." Just as our first introduction to double integrals was in the context of finding the area of a plane region, our introduction into triple integrals will be in the context of finding the volume of a space region.



(a) **(b)**

Figure 15.6.4 Approximating the volume of a region D in space

To formally find the volume of a closed, bounded region D in space, such as the one shown in Figure 15.6.4(a), we start with an approximation. Break D into n rectangular solids; the solids near the boundary of D may possibly not include portions of D and/or include extra space. In Figure 15.6.4(b), we zoom in on a portion of the boundary of D to show a rectangular solid that contains space not in D; as this is an approximation of the volume, this is acceptable and this error will be reduced as we shrink the size of our solids.

The volume ΔV_i of the ith solid D_i is $\Delta V_i = \Delta x_i \Delta y_i \Delta z_i$, where Δx_i, Δy_i and Δz_i give the dimensions of the rectangular solid in the x, y and z directions,

respectively. By summing up the volumes of all n solids, we get an approximation of the volume V of D:

$$V \approx \sum_{i=1}^{n} \Delta V_i = \sum_{i=1}^{n} \Delta x_i \Delta y_i \Delta z_i.$$

Let $\|\Delta D\|$ represent the length of the longest diagonal of rectangular solids in the subdivision of D. As $\|\Delta D\| \to 0$, the volume of each solid goes to 0, as do each of Δx_i, Δy_i and Δz_i, for all i. Our calculus experience tells us that taking a limit as $\|\Delta D\| \to 0$ turns our approximation of V into an exact calculation of V. Before we state this result in a theorem, we use a definition to define some terms.

Definition 15.6.5 Triple Integrals, Iterated Integration (Part I).

Let D be a closed, bounded region in space. Let a and b be real numbers, let $g_1(x)$ and $g_2(x)$ be continuous functions of x, and let $f_1(x, y)$ and $f_2(x, y)$ be continuous functions of x and y.

1. The volume V of D is denoted by a *triple integral*,
$$V = \iiint_D dV.$$
2. The iterated integral $\displaystyle\int_a^b \int_{g_1(x)}^{g_2(x)} \int_{f_1(x,y)}^{f_2(x,y)} dz\, dy\, dx$ is evaluated as
$$\int_a^b \int_{g_1(x)}^{g_2(x)} \int_{f_1(x,y)}^{f_2(x,y)} dz\, dy\, dx = \int_a^b \int_{g_1(x)}^{g_2(x)} \left(\int_{f_1(x,y)}^{f_2(x,y)} dz \right) dy\, dx.$$
Evaluating the above iterated integral is *triple integration*.

youtu.be/watch?v=EEWiIQfJktI

Figure 15.6.6 Introducing the triple integral

Our informal understanding of the notation $\iiint_D dV$ is "sum up lots of little volumes over D," analogous to our understanding of $\iint_R dA$ and $\iint_R dm$.

We now state the major theorem of this section.

Theorem 15.6.7 Triple Integration (Part I).

Let D be a closed, bounded region in space and let ΔD be any subdivision of D into n rectangular solids, where the ith subregion D_i has dimensions $\Delta x_i \times \Delta y_i \times \Delta z_i$ and volume ΔV_i.

1. *The volume V of D is*
$$V = \iiint_D dV = \lim_{\|\Delta D\| \to 0} \sum_{i=1}^{n} \Delta V_i = \lim_{\|\Delta D\| \to 0} \sum_{i=1}^{n} \Delta x_i \Delta y_i \Delta z_i.$$
2. *If D is defined as the region bounded by the planes $x = a$ and $x = b$, the cylinders $y = g_1(x)$ and $y = g_2(x)$, and the surfaces $z = f_1(x, y)$ and $z = f_2(x, y)$, where $a < b$, $g_1(x) \le g_2(x)$ and $f_1(x, y) \le f_2(x, y)$ on D, then*
$$\iiint_D dV = \int_a^b \int_{g_1(x)}^{g_2(x)} \int_{f_1(x,y)}^{f_2(x,y)} dz\, dy\, dx.$$
3. *V can be determined using iterated integration with other orders of integration (there are 6 total), as long as D is defined by the*

region enclosed by a pair of planes, a pair of cylinders, and a pair of surfaces.

We evaluated the area of a plane region R by iterated integration, where the bounds were "from curve to curve, then from point to point." Theorem 15.6.7 allows us to find the volume of a space region with an iterated integral with bounds "from surface to surface, then from curve to curve, then from point to point." In the iterated integral

$$\int_a^b \int_{g_1(x)}^{g_2(x)} \int_{f_1(x,y)}^{f_2(x,y)} dz\,dy\,dx,$$

the bounds $a \le x \le b$ and $g_1(x) \le y \le g_2(x)$ define a region R in the xy-plane over which the region D exists in space. However, these bounds are also defining surfaces in space; $x = a$ is a plane and $y = g_1(x)$ is a cylinder. The combination of these 6 surfaces enclose, and define, D.

Examples will help us understand triple integration, including integrating with various orders of integration.

youtu.be/watch?v=lswNoU6cAmA

Figure 15.6.8 Finding the volume of a tetrahedron

Example 15.6.9 Finding the volume of a space region with triple integration.

Find the volume of the space region in the first octant bounded by the plane $z = 2 - y/3 - 2x/3$, shown in Figure 15.6.10, using the order of integration $dz\,dy\,dx$. Set up the triple integrals that give the volume in the other 5 orders of integration.

Example 15.6.9 uses the term "first octant." Recall how the x-, y- and z-axes divide space into eight *octants*; the octant in which x, y and z are all positive is called the *first octant*.

Solution. Starting with the order of integration $dz\,dy\,dx$, we need to first find bounds on z. The region D is bounded below by the plane $z = 0$ (because we are restricted to the first octant) and above by $z = 2 - y/3 - 2x/3$; $0 \le z \le 2 - y/3 - 2x/3$.

To find the bounds on y and x, we "collapse" the region onto the xy-plane, giving the triangle shown in Figure 15.6.11. (We know the equation of the line $y = 6 - 2x$ in two ways. First, by setting $z = 0$, we have $0 = 2 - y/3 - 2x/3 \Rightarrow y = 6 - 2x$. Secondly, we know this is going to be a straight line between the points $(3, 0)$ and $(0, 6)$ in the xy-plane.)

We define that region R, in the integration order of $dy\,dx$, with bounds $0 \le y \le 6 - 2x$ and $0 \le x \le 3$. Thus the volume V of the region D is:

$$\begin{aligned} V &= \iiint_D dV \\ &= \int_0^3 \int_0^{6-2x} \int_0^{2-\frac{1}{3}y-\frac{2}{3}x} dz\,dy\,dx \\ &= \int_0^3 \int_0^{6-2x} \left(\int_0^{2-\frac{1}{3}y-\frac{2}{3}x} dz \right) dy\,dx \\ &= \int_0^3 \int_0^{6-2x} z\Big|_0^{2-\frac{1}{3}y-\frac{2}{3}x} dy\,dx \\ &= \int_0^3 \int_0^{6-2x} \left(2 - \frac{1}{3}y - \frac{2}{3}x\right) dy\,dx. \end{aligned}$$

From this step on, we are evaluating a double integral as done many times before. We skip these steps and give the final volume, $V = 6$.

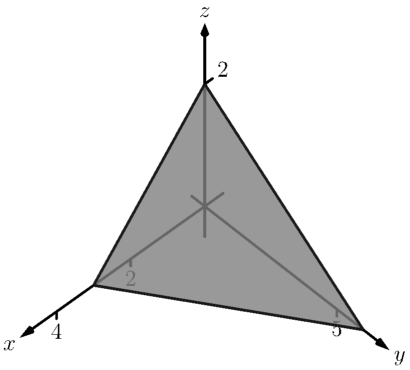


Figure 15.6.10 The region D used in Example 15.6.9

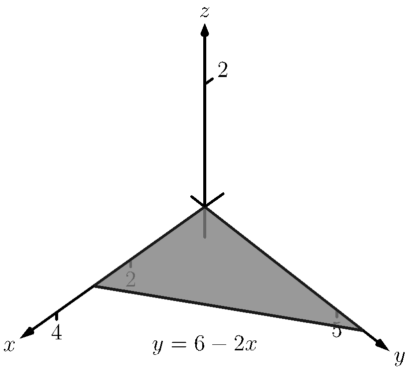


Figure 15.6.11 The region found by collapsing D onto the xy-plane

The order $dz\,dx\,dy$:

Now consider the volume using the order of integration $dz\,dx\,dy$. The bounds on z are the same as before, $0 \le z \le 2 - y/3 - 2x/3$. Collapsing the space region on the xy-plane as shown in Figure 15.6.11, we now describe this triangle with the order of integration $dx\,dy$. This gives bounds $0 \le x \le 3 - y/2$ and $0 \le y \le 6$. Thus the volume is given by the triple integral

$$V = \int_0^6 \int_0^{3-\frac{1}{2}y} \int_0^{2-\frac{1}{3}y-\frac{2}{3}x} dz\,dx\,dy.$$

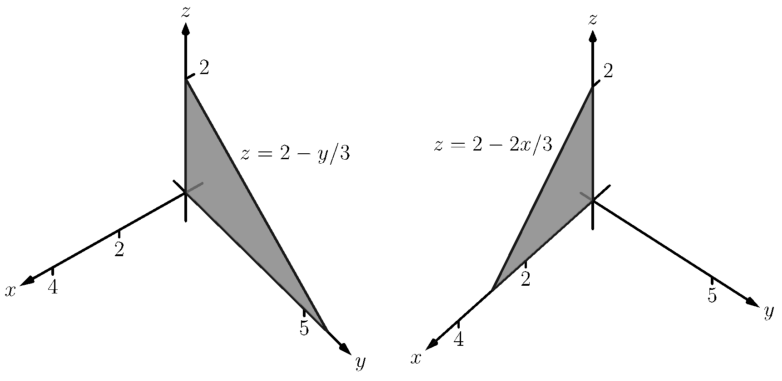
The order $dx\,dy\,dz$:

Following our "surface to surface. . ." strategy, we need to determine the x-*surfaces* that bound our space region. To do so, approach the region "from behind," in the direction of increasing x. The first surface we hit as we enter the region is the yz-plane, defined by $x = 0$. We come out of the region at the plane $z = 2 - y/3 - 2x/3$; solving for x, we have $x = 3 - y/2 - 3z/2$. Thus the bounds on x are: $0 \le x \le 3 - y/2 - 3z/2$.

Now collapse the space region onto the yz-plane, as shown in Figure 15.6.12(a). (Again, we find the equation of the line $z = 2 - y/3$ by setting $x = 0$ in the equation $x = 3 - y/2 - 3z/2$.) We need to find bounds on this region with the order $dy\,dz$. The *curves* that bound y are $y = 0$ and $y = 6 - 3z$; the *points* that bound z are 0 and 2. Thus the triple integral giving volume is:

$$\begin{aligned} &0 \le x \le 3 - y/2 - 3z/2 \\ &0 \le y \le 6 - 3z \\ &0 \le z \le 2 \end{aligned} \quad \Rightarrow \quad \int_0^2 \int_0^{6-3z} \int_0^{3-y/2-3z/2} dx\,dy\,dz.$$

The order $dx\,dz\,dy$:



(a) **(b)**

Figure 15.6.12 The region D in Example 15.6.9 is collapsed onto the yz-plane in (a); in (b), the region is collapsed onto the xz-plane

The x-bounds are the same as the order above. We now consider the triangle in Figure 15.6.12(a) and describe it with the order $dz\,dy$: $0 \le z \le 2 - y/3$ and $0 \le y \le 6$. Thus the volume is given by:

$$\begin{aligned} 0 \le x \le 3 - y/2 - 3z/2 \\ 0 \le z \le 2 - y/3 \\ 0 \le y \le 6 \end{aligned} \quad \Rightarrow \quad \int_0^6 \int_0^{2-y/3} \int_0^{3-y/2-3z/2} dx\, dz\, dy.$$

The order $dy\, dz\, dx$:

We now need to determine the y-surfaces that determine our region. Approaching the space region from "behind" and moving in the direction of increasing y, we first enter the region at $y = 0$, and exit along the plane $z = 2 - y/3 - 2x/3$. Solving for y, this plane has equation $y = 6 - 2x - 3z$. Thus y has bounds $0 \le y \le 6 - 2x - 3z$.

Now collapse the region onto the xz-plane, as shown in Figure 15.6.12(b). The curves bounding this triangle are $z = 0$ and $z = 2 - 2x/3$; x is bounded by the points $x = 0$ to $x = 3$. Thus the triple integral giving volume is:

$$\begin{aligned} 0 \le y \le 6 - 2x - 3z \\ 0 \le z \le 2 - 2x/3 \\ 0 \le x \le 3 \end{aligned} \quad \Rightarrow \quad \int_0^3 \int_0^{2-2x/3} \int_0^{6-2x-3z} dy\, dz\, dx.$$

The order $dy\, dx\, dz$:

The y-bounds are the same as in the order above. We now determine the bounds of the triangle in Figure 15.6.12(b) using the order $dy\, dx\, dz$. x is bounded by $x = 0$ and $x = 3 - 3z/2$; z is bounded between $z = 0$ and $z = 2$. This leads to the triple integral:

$$\begin{aligned} 0 \le y \le 6 - 2x - 3z \\ 0 \le x \le 3 - 3z/2 \\ 0 \le z \le 2 \end{aligned} \quad \Rightarrow \quad \int_0^2 \int_0^{3-3z/2} \int_0^{6-2x-3z} dy\, dx\, dz.$$

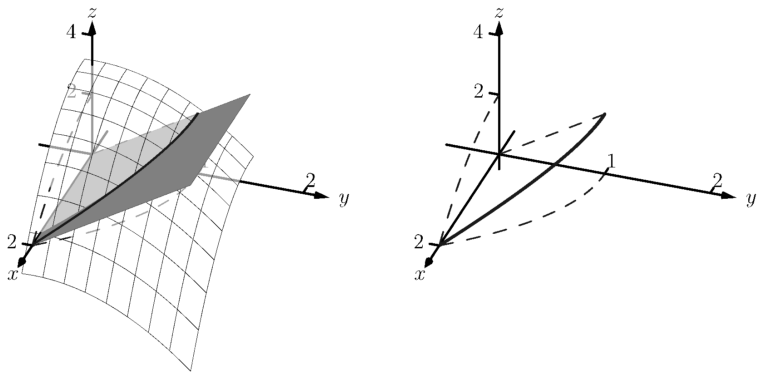
This problem was long, but hopefully useful, demonstrating how to determine bounds with every order of integration to describe the region D. In practice, we only need 1, but being able to do them all gives us flexibility to choose the order that suits us best.

In the previous example, we collapsed the surface into the x-y, x-z, and yz-planes as we determined the "curve to curve, point to point" bounds of integration. Since the surface was a triangular portion of a plane, this collapsing, or *projecting*, was simple: the *projection* of a straight line in space onto a coordinate plane is a line.

The following example shows us how to do this when dealing with more complicated surfaces and curves.

Example 15.6.13 Finding the projection of a curve in space onto the coordinate planes.

Consider the surfaces $z = 3 - x^2 - y^2$ and $z = 2y$, as shown in Figure 15.6.14(a). The curve of their intersection is shown, along with the projection of this curve into the coordinate planes, shown dashed. Find the equations of the projections into the coordinate planes.



(a) **(b)**

Figure 15.6.14 Finding the projections of the curve of intersection in Example 15.6.13

Solution. The two surfaces are $z = 3 - x^2 - y^2$ and $z = 2y$. To find where they intersect, it is natural to set them equal to each other: $3 - x^2 - y^2 = 2y$. This is an implicit function of x and y that gives all points (x, y) in the xy-plane where the z values of the two surfaces are equal.

We can rewrite this implicit function by completing the square:

$$3 - x^2 - y^2 = 2y \Rightarrow y^2 + 2y + x^2 = 3 \Rightarrow (y+1)^2 + x^2 = 4.$$

Thus in the xy-plane the projection of the intersection is a circle with radius 2, centered at $(0, -1)$.

To project onto the xz-plane, we do a similar procedure: find the x and z values where the y values on the surface are the same. We start by solving the equation of each surface for y. In this particular case, it works well to actually solve for y^2:

$z = 3 - x^2 - y^2 \Rightarrow y^2 = 3 - x^2 - z$

$z = 2y \Rightarrow y^2 = z^2/4$.

Thus we have (after again completing the square):

$$3 - x^2 - z = z^2/4 \Rightarrow \frac{(z+2)^2}{16} + \frac{x^2}{4} = 1,$$

and ellipse centered at $(0, -2)$ in the xz-plane with a major axis of length 8 and a minor axis of length 4.

Finally, to project the curve of intersection into the yz-plane, we solve equation for x. Since $z = 2y$ is a cylinder that lacks the variable x, it becomes our equation of the projection in the yz-plane.

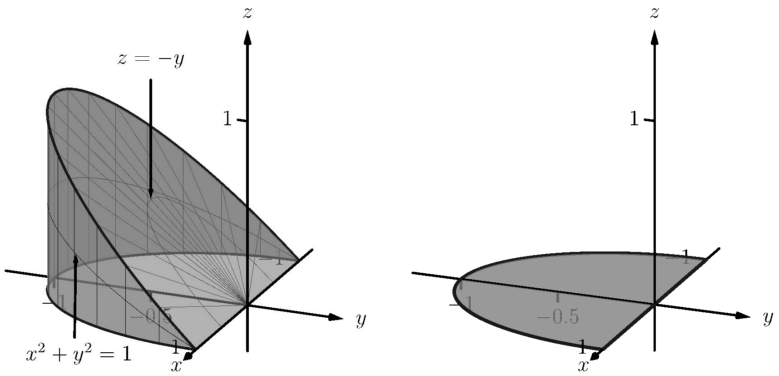
All three projections are shown in Figure 15.6.14(b).

youtu.be/watch?v=PA9fkl8bLaI

Figure 15.6.15 Changing the order of integration in a triple integral

Example 15.6.16 Finding the volume of a space region with triple integration.

Set up the triple integrals that find the volume of the space region D bounded by the surfaces $x^2 + y^2 = 1$, $z = 0$ and $z = -y$, as shown in Figure 15.6.17(a), with the orders of integration $dz\,dy\,dx$, $dy\,dx\,dz$ and $dx\,dz\,dy$.



(a) **(b)**

Figure 15.6.17 The region D in Example 15.6.16 is shown in (a); in (b), it is collapsed onto the xy-plane

Solution. The order $dz\,dy\,dx$:

The region D is bounded below by the plane $z = 0$ and above by the plane $z = -y$. The cylinder $x^2+y^2 = 1$ does not offer any bounds in the z-direction, as that surface is parallel to the z-axis. Thus $0 \le z \le -y$.

Collapsing the region into the xy-plane, we get part of the circle with equation $x^2 + y^2 = 1$ as shown in Figure 15.6.17(b). As a function of x, this half circle has equation $y = -\sqrt{1-x^2}$. Thus y is bounded below by $-\sqrt{1-x^2}$ and above by $y = 0$: $-\sqrt{1-x^2} \le y \le 0$. The x bounds of the half circle are $-1 \le x \le 1$. All together, the bounds of integration and triple integral are as follows:

$$\begin{aligned} 0 \le z &\le -y \\ -\sqrt{1-x^2} \le y &\le 0 \\ -1 \le x &\le 1 \end{aligned} \quad \Rightarrow \quad \int_{-1}^{1}\int_{-\sqrt{1-x^2}}^{0}\int_{0}^{-y} dz\,dy\,dx.$$

We evaluate this triple integral:

$$\begin{aligned} \int_{-1}^{1}\int_{-\sqrt{1-x^2}}^{0}\int_{0}^{-y} dz\,dy\,dx &= \int_{-1}^{1}\int_{-\sqrt{1-x^2}}^{0} \big(-y\big)\,dy\,dx \\ &= \int_{-1}^{1} \big(-\frac{1}{2}y^2\big)\Big|_{-\sqrt{1-x^2}}^{0} dx \\ &= \int_{-1}^{1} \frac{1}{2}\big(1-x^2\big)\,dx \\ &= \left(\frac{1}{2}\left(x - \frac{1}{3}x^3\right)\right)\Big|_{-1}^{1} \\ &= \frac{2}{3} \text{ units}^3. \end{aligned}$$

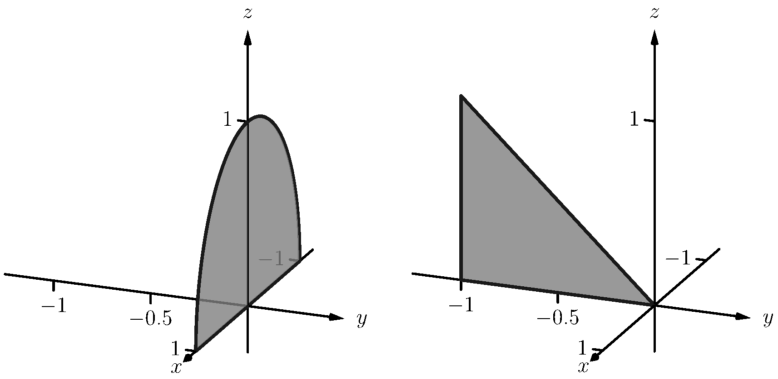
With the order $dy\,dx\,dz$:

The region is bounded "below" in the y-direction by the surface $x^2 + y^2 = 1 \Rightarrow y = -\sqrt{1-x^2}$ and "above" by the surface $y = -z$. Thus the y bounds are $-\sqrt{1-x^2} \le y \le -z$.

Collapsing the region onto the xz-plane gives the region shown in Figure 15.6.18(a); this half disk is bounded by $z = 0$ and $x^2 + z^2 =$

1. (We find this curve by solving each surface for y^2, then setting them equal to each other. We have $y^2 = 1 - x^2$ and $y = -z \Rightarrow y^2 = z^2$. Thus $x^2 + z^2 = 1$.) It is bounded below by $x = -\sqrt{1 - z^2}$ and above by $x = \sqrt{1 - z^2}$, where z is bounded by $0 \le z \le 1$. All together, we have:

$$\begin{aligned} -\sqrt{1-x^2} \le y \le -z \\ -\sqrt{1-z^2} \le x \le \sqrt{1-z^2} \\ 0 \le z \le 1 \end{aligned} \quad \Rightarrow \quad \int_0^1 \int_{-\sqrt{1-z^2}}^{\sqrt{1-z^2}} \int_{-\sqrt{1-x^2}}^{-z} dy\,dx\,dz.$$



(a) **(b)**

Figure 15.6.18 The region D in Example 15.6.16 is shown collapsed onto the xz-plane in (a); in (b), it is collapsed onto the yz-plane

With the order $dx\,dz\,dy$:

D is bounded below by the surface $x = -\sqrt{1 - y^2}$ and above by $\sqrt{1 - y^2}$. We then collapse the region onto the yz-plane and get the triangle shown in Figure 15.6.18(b). (The hypotenuse is the line $z = -y$, just as the plane.) Thus z is bounded by $0 \le z \le -y$ and y is bounded by $-1 \le y \le 0$. This gives:

$$\begin{aligned} -\sqrt{1-y^2} \le x \le \sqrt{1-y^2} \\ 0 \le z \le -y \\ -1 \le y \le 0 \end{aligned} \quad \Rightarrow \quad \int_{-1}^0 \int_0^{-y} \int_{-\sqrt{1-y^2}}^{\sqrt{1-y^2}} dx\,dz\,dy.$$

The following theorem states two things that should make "common sense" to us. First, using the triple integral to find volume of a region D should always return a positive number; we are computing *volume* here, not *signed volume*. Secondly, to compute the volume of a "complicated" region, we could break it up into subregions and compute the volumes of each subregion separately, summing them later to find the total volume.

Theorem 15.6.19 Properties of Triple Integrals.

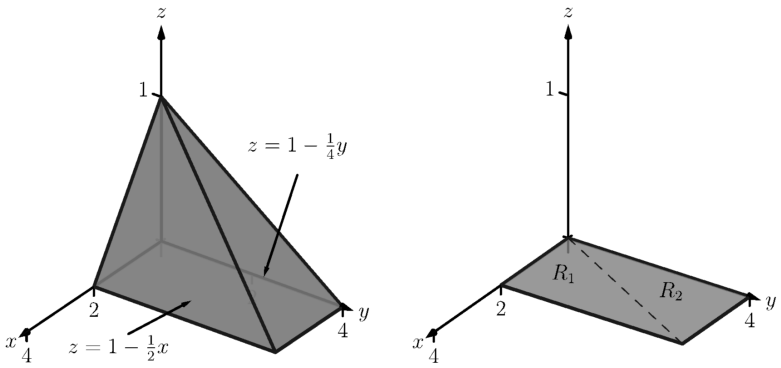
Let D be a closed, bounded region in space, and let D_1 and D_2 be non-overlapping regions such that $D = D_1 \bigcup D_2$.

1. $\iiint_D dV \geq 0$
2. $\iiint_D dV = \iiint_{D_1} dV + \iiint_{D_2} dV.$

We use this latter property in the next example.

Example 15.6.20 Finding the volume of a space region with triple integration.

Find the volume of the space region D bounded by the coordinate planes, $z = 1 - x/2$ and $z = 1 - y/4$, as shown in Figure 15.6.21(a). Set up the triple integrals that find the volume of D in all 6 orders of integration.



(a) **(b)**

Figure 15.6.21 The region D in Example 15.6.20 is shown in (a); in (b), it is collapsed onto the xy-plane

Solution. Following the bounds-determining strategy of "surface to surface, curve to curve, and point to point," we can see that the most difficult orders of integration are the two in which we integrate with respect to z first, for there are two "upper" surfaces that bound D in the z-direction. So we start by noting that we have

$$0 \leq z \leq 1 - \frac{1}{2}x \text{ and } 0 \leq z \leq 1 - \frac{1}{4}y.$$

We now collapse the region D onto the xy-axis, as shown in Figure 15.6.21(b). The boundary of D, the line from $(0, 0, 1)$ to $(2, 4, 0)$, is shown in Figure 15.6.21(b) as a dashed line; it has equation $y = 2x$. (We can recognize this in two ways: one, in collapsing the line from $(0, 0, 1)$ to $(2, 4, 0)$ onto the xy-plane, we simply ignore the z-values, meaning the line now goes from $(0, 0)$ to $(2, 4)$. Secondly, the two surfaces meet where $z = 1 - x/2$ is equal to $z = 1 - y/4$: thus $1 - x/2 = 1 - y/4 \Rightarrow y = 2x$.)

We use the second property of Theorem 15.6.19 to state that

$$\iiint_D dV = \iiint_{D_1} dV + \iiint_{D_2} dV,$$

where D_1 and D_2 are the space regions above the plane regions R_1 and R_2, respectively. Thus we can say

$$\iiint_D dV = \iint_{R_1} \left(\int_0^{1-x/2} dz \right) dA + \iint_{R_2} \left(\int_0^{1-y/4} dz \right) dA.$$

All that is left is to determine bounds of R_1 and R_2, depending on whether we are integrating with order $dx\,dy$ or $dy\,dx$. We give the final integrals here, leaving it to the reader to confirm these results.

$dz\,dy\,dx$:

$$\begin{aligned} &0 \le z \le 1 - x/2 \\ &0 \le y \le 2x \\ &0 \le x \le 2 \end{aligned} \qquad \begin{aligned} &0 \le z \le 1 - y/4 \\ &2x \le y \le 4 \\ &0 \le x \le 2 \end{aligned}$$

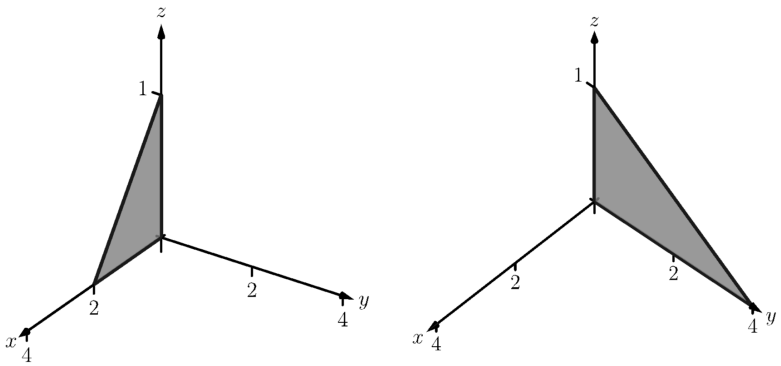
$$\iiint_D dV = \int_0^2 \int_0^{2x} \int_0^{1-x/2} dz\,dy\,dx \quad + \quad \int_0^2 \int_{2x}^4 \int_0^{1-y/4} dz\,dy\,dx$$

$dz\,dx\,dy$:

$$\begin{aligned} &0 \le z \le 1 - x/2 \\ &y/2 \le x \le 2 \\ &0 \le y \le 4 \end{aligned} \qquad \begin{aligned} &0 \le z \le 1 - y/4 \\ &0 \le x \le y/2 \\ &0 \le y \le 4 \end{aligned}$$

$$\iiint_D dV = \int_0^4 \int_{y/2}^2 \int_0^{1-x/2} dz\,dx\,dy \quad + \quad \int_0^4 \int_0^{y/2} \int_0^{1-y/4} dz\,dx\,dy$$

The remaining four orders of integration do not require a sum of triple integrals. In Figure 15.6.22 we show D collapsed onto the other two coordinate planes. Using these graphs, we give the final orders of integration here, again leaving it to the reader to confirm these results.



(a) **(b)**

Figure 15.6.22 The region D in Example 15.6.20 is shown collapsed onto the xz-plane in (a); in (b), it is collapsed onto the yz-plane

$dy\,dx\,dz$:

$$\begin{aligned} &0 \le y \le 4-4z \\ &0 \le x \le 2-2z \\ &0 \le z \le 1 \end{aligned} \quad \Rightarrow \quad \int_0^1 \int_0^{2-2z} \int_0^{4-4z} dy\,dx\,dz.$$

$dy\,dz\,dx$:

$$\begin{aligned} &0 \le y \le 4-4z \\ &0 \le z \le 1-x/2 \\ &0 \le x \le 2 \end{aligned} \quad \Rightarrow \quad \int_0^2 \int_0^{1-x/2} \int_0^{4-4z} dy\,dx\,dz.$$

$dx\,dy\,dz$:

$$\begin{aligned} &0 \le x \le 2-2z \\ &0 \le y \le 4-4z \\ &0 \le z \le 1 \end{aligned} \quad \Rightarrow \quad \int_0^1 \int_0^{4-4z} \int_0^{2-2z} dx\,dy\,dz.$$

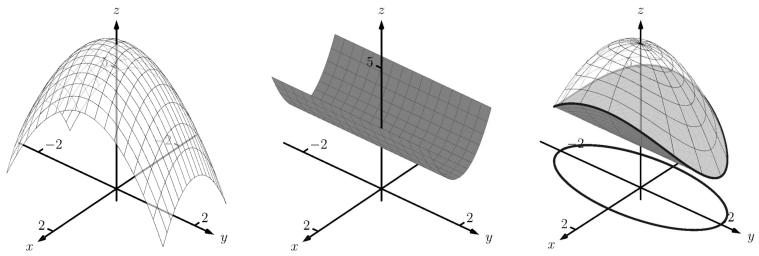
$dx\,dz\,dy$:

$$\begin{aligned} &0 \le x \le 2-2z \\ &0 \le z \le 1-y/4 \\ &0 \le y \le 4 \end{aligned} \quad \Rightarrow \quad \int_0^4 \int_0^{1-y/4} \int_0^{2-2z} dx\,dz\,dy.$$

We give one more example of finding the volume of a space region.

Example 15.6.23 Finding the volume of a space region.

Set up a triple integral that gives the volume of the space region D bounded by $z = 2x^2 + 2$ and $z = 6 - 2x^2 - y^2$. These surfaces are plotted in Figure 15.6.24(a) and Figure 15.6.24(b), respectively; the region D is shown in Figure 15.6.24(c).



(a) **(b)** **(c)**

Figure 15.6.24 The region D is bounded by the surfaces shown in (a) and (b); D is shown in (c)

Solution. The main point of this example is this: integrating with respect to z first is rather straightforward; integrating with respect to x first is not.

The order $dz\,dy\,dx$:

The bounds on z are clearly $2x^2+2 \le z \le 6-2x^2-y^2$. Collapsing D onto the xy-plane gives the ellipse shown in Figure 15.6.24(c). The equation of this ellipse is found by setting the two surfaces equal to each other:

$$2x^2+2 = 6-2x^2-y^2 \Rightarrow 4x^2+y^2 = 4 \Rightarrow x^2+\frac{y^2}{4} = 1.$$

We can describe this ellipse with the bounds

$$-\sqrt{4-4x^2} \le y \le \sqrt{4-4x^2} \text{ and } -1 \le x \le 1.$$

Thus we find volume as

$$\begin{array}{c} 2x^2+2 \le z \le 6-2x^2-y^2 \\ -\sqrt{4-4x^2} \le y \le \sqrt{4-4x^2} \\ -1 \le x \le 1 \end{array} \Rightarrow \qquad \int_{-1}^{1}\int_{-\sqrt{4-4x^2}}^{\sqrt{4-4x^2}}\int_{2x^2+2}^{6-2x^2-y^2} dz\,dy\,dx.$$

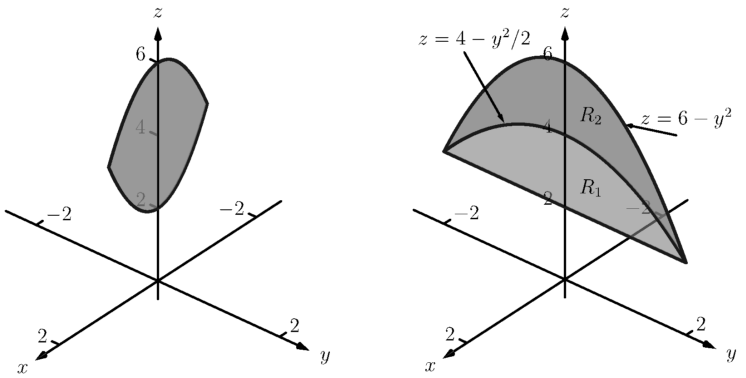
The order $dy\,dz\,dx$:

Integrating with respect to y is not too difficult. Since the surface $z = 2x^2+2$ is a cylinder whose directrix is the y-axis, it does not create a border for y. The paraboloid $z = 6-2x^2-y^2$ does; solving for y, we get the bounds

$$-\sqrt{6-2x^2-z} \le y \le \sqrt{6-2x^2-z}.$$

Collapsing D onto the xz-plane gives the region shown in Figure 15.6.25(a); the lower curve is from the cylinder, with equation $z = 2x^2+2$. The upper curve is from the paraboloid; with $y = 0$, the curve is $z = 6-2x^2$. Thus bounds on z are $2x^2+2 \le z \le 6-2x^2$; the bounds on x are $-1 \le x \le 1$. Thus we have:

$$\begin{array}{c} -\sqrt{6-2x^2-z} \le y \le \sqrt{6-2x^2-z} \\ 2x^2+2 \le z \le 6-2x^2 \\ -1 \le x \le 1 \end{array} \Rightarrow \qquad \int_{-1}^{1}\int_{2x^2+2}^{6-2x^2}\int_{-\sqrt{6-2x^2-z}}^{\sqrt{6-2x^2-z}} dy\,dz\,dx.$$



(a) **(b)**

Figure 15.6.25 The region D in Example 15.6.23 is collapsed onto the xz-plane in (a); in (b), it is collapsed onto the yz-plane

The order $dx\,dz\,dy$:

This order takes more effort as D must be split into two subregions. The two surfaces create two sets of upper/lower bounds in terms of x; the cylinder creates bounds

$$-\sqrt{z/2-1} \le x \le \sqrt{z/2-1}$$

for region D_1 and the paraboloid creates bounds

$$-\sqrt{3-y^2/2-z^2/2} \le x \le \sqrt{3-y^2/2-z^2/2}$$

for region D_2.

Collapsing D onto the yz-axes gives the regions shown in Figure 15.6.25(b). We find the equation of the curve $z = 4 - y^2/2$ by noting that the equation of the ellipse seen in Figure 15.6.24(c) has equation

$$x^2 + y^2/4 = 1 \Rightarrow x = \sqrt{1-y^2/4}.$$

Substitute this expression for x in either surface equation, $z = 6 - 2x^2 - y^2$ or $z = 2x^2 + 2$. In both cases, we find

$$z = 4 - \frac{1}{2}y^2.$$

Region R_1, corresponding to D_1, has bounds

$$2 \le z \le 4 - y^2/2, -2 \le y \le 2$$

and region R_2, corresponding to D_2, has bounds

$$4 - y^2/2 \le z \le 6 - y^2, -2 \le y \le 2.$$

Thus the volume of D is given by:

$$\int_{-2}^{2}\int_{2}^{4-y^2/2}\int_{-\sqrt{z/2-1}}^{\sqrt{z/2-1}} dx\,dz\,dy + \int_{-2}^{2}\int_{4-y^2/2}^{6-y^2}\int_{-\sqrt{3-y^2/2-z^2/2}}^{\sqrt{3-y^2/2-z^2/2}} dx\,dz\,dy.$$

If all one wanted to do in Example 15.6.23 was find the volume of the re-

gion D, one would have likely stopped at the first integration setup (with order $dz\,dy\,dx$) and computed the volume from there. However, we included the other two methods 1) to show that it could be done, "messy" or not, and 2) because sometimes we "have" to use a less desirable order of integration in order to actually integrate.

15.6.2 Triple Integration and Functions of Three Variables

There are uses for triple integration beyond merely finding volume, just as there are uses for integration beyond "area under the curve." These uses start with understanding how to integrate functions of three variables, which is effectively no different than integrating functions of two variables. This leads us to a definition, followed by an example.

Definition 15.6.26 Iterated Integration, (Part II).

Let D be a closed, bounded region in space, over which $g_1(x)$, $g_2(x)$, $f_1(x,y)$, $f_2(x,y)$ and $h(x,y,z)$ are all continuous, and let a and b be real numbers.

The **iterated integral** $\displaystyle\int_a^b \int_{g_1(x)}^{g_2(x)} \int_{f_1(x,y)}^{f_2(x,y)} h(x,y,z)\,dz\,dy\,dx$ is evaluated as

$$\int_a^b \int_{g_1(x)}^{g_2(x)} \int_{f_1(x,y)}^{f_2(x,y)} h(x,y,z)\,dz\,dy\,dx = \int_a^b \int_{g_1(x)}^{g_2(x)} \left(\int_{f_1(x,y)}^{f_2(x,y)} h(x,y,z)\,dz \right) dy\,dx.$$

Example 15.6.27 Evaluating a triple integral of a function of three variables.

Evaluate $\displaystyle\int_0^1 \int_{x^2}^{x} \int_{x^2-y}^{2x+3y} \big(xy + 2xz\big)\,dz\,dy\,dx.$

Solution. We evaluate this integral according to Definition 15.6.26.

$$\begin{aligned}
&\int_0^1 \int_{x^2}^{x} \int_{x^2-y}^{2x+3y} \big(xy + 2xz\big)\,dz\,dy\,dx \\
&= \int_0^1 \int_{x^2}^{x} \left(\int_{x^2-y}^{2x+3y} \big(xy + 2xz\big)\,dz \right) dy\,dx \\
&= \int_0^1 \int_{x^2}^{x} \left(\big(xyz + xz^2\big)\Big|_{x^2-y}^{2x+3y} \right) dy\,dx \\
&= \int_0^1 \int_{x^2}^{x} \Big(xy(2x+3y) + x(2x+3y)^2 - \Big(xy(x^2-y) + x(x^2-y)^2 \Big) \Big)\,dy\,dx \\
&= \int_0^1 \int_{x^2}^{x} \Big(-x^5 + x^3y + 4x^3 + 14x^2y + 12xy^2 \Big)\,dy\,dx.
\end{aligned}$$

We continue as we have in the past, showing fewer steps.

$$\begin{aligned}
&= \int_0^1 \left(-\frac{7}{2}x^7 - 8x^6 - \frac{7}{2}x^5 + 15x^4 \right) dx \\
&= \frac{281}{336} \approx 0.836.
\end{aligned}$$

We now know *how* to evaluate a triple integral of a function of three variables; we do not yet understand what it *means*. We build up this understanding in a way very similar to how we have understood integration and double integration.

Let $h(x, y, z)$ be a continuous function of three variables, defined over some space region D. We can partition D into n rectangular-solid subregions, each with dimensions $\Delta x_i \times \Delta y_i \times \Delta z_i$. Let (x_i, y_i, z_i) be some point in the ith subregion, and consider the product $h(x_i, y_i, z_i)\Delta x_i \Delta y_i \Delta z_i$. It is the product of a function value (that's the $h(x_i, y_i, z_i)$ part) and a small volume ΔV_i (that's the $\Delta x_i \Delta y_i \Delta z_i$ part). One of the simplest understanding of this type of product is when h describes the density of an object, for then $h \times$ volume $=$ mass .

We can sum up all n products over D. Again letting $\|\Delta D\|$ represent the length of the longest diagonal of the n rectangular solids in the partition, we can take the limit of the sums of products as $\|\Delta D\| \to 0$. That is, we can find

$$S = \lim_{\|\Delta D\|\to 0} \sum_{i=1}^{n} h(x_i, y_i, z_i)\Delta V_i = \lim_{\|\Delta D\|\to 0} \sum_{i=1}^{n} h(x_i, y_i, z_i)\Delta x_i \Delta y_i \Delta z_i.$$

While this limit has lots of interpretations depending on the function h, in the case where h describes density, S is the total mass of the object described by the region D.

We now use the above limit to define the *triple integral*, give a theorem that relates triple integrals to iterated iteration, followed by the application of triple integrals to find the centers of mass of solid objects.

Definition 15.6.28 Triple Integral.

Let $w = h(x, y, z)$ be a continuous function over a closed, bounded region D in space, and let ΔD be any partition of D into n rectangular solids with volume ΔV_i. The **triple integral** of h over D is

$$\iiint_D h(x, y, z)\, dV = \lim_{\|\Delta D\|\to 0} \sum_{i=1}^{n} h(x_i, y_i, z_i)\Delta V_i.$$

The following theorem assures us that the above limit exists for continuous functions h and gives us a method of evaluating the limit.

Theorem 15.6.29 Triple Integration (Part II).

Let $w = h(x, y, z)$ be a continuous function over a closed, bounded region D in space, and let ΔD be any partition of D into n rectangular solids with volume V_i.

1. *The limit $\lim_{\|\Delta D\|\to 0} \sum_{i=1}^{n} h(x_i, y_i, z_i)\Delta V_i$ exists.*
2. *If D is defined as the region bounded by the planes $x = a$ and $x = b$, the cylinders $y = g_1(x)$ and $y = g_2(x)$, and the surfaces $z = f_1(x, y)$ and $z = f_2(x, y)$, where $a < b$, $g_1(x) \le g_2(x)$ and $f_1(x, y) \le f_2(x, y)$ on D, then*

$$\iiint_D h(x, y, z)\, dV = \int_a^b \int_{g_1(x)}^{g_2(x)} \int_{f_1(x,y)}^{f_2(x,y)} h(x, y, z)\, dz\, dy\, dx.$$

Note: In an aside in Section 15.2, we showed how the summation of rectangles over a region R in the plane could be viewed as a double sum, leading to

the double integral. Likewise, we can view the sum

$$\sum_{i=1}^{n} h(x_i, y_i, z_i)\Delta x_i \Delta y_i \Delta z_i$$

as a triple sum,

$$\sum_{k=1}^{p}\sum_{j=1}^{n}\sum_{i=1}^{m} h(x_i, y_j, z_k)\Delta x_i \Delta y_j \Delta z_k,$$

which we evaluate as

$$\sum_{k=1}^{p}\left(\sum_{j=1}^{n}\left(\sum_{i=1}^{m} h(x_i, y_j, z_k)\Delta x_i\right)\Delta y_j\right)\Delta z_k.$$

Here we fix a k value, which establishes the z-height of the rectangular solids on one "level" of all the rectangular solids in the space region D. The inner double summation adds up all the volumes of the rectangular solids on this level, while the outer summation adds up the volumes of each level.

This triple summation understanding leads to the $\iiint_D$ notation of the triple integral, as well as the method of evaluation shown in Theorem 15.6.29.

We now apply triple integration to find the centers of mass of solid objects.

15.6.3 Mass and Center of Mass

One may wish to review Section 15.4 for a reminder of the relevant terms and concepts.

Definition 15.6.30 Mass, Center of Mass of Solids.

Let a solid be represented by a closed, bounded region D in space with variable density function $\delta(x,y,z)$.

1. The *mass* of the object is $M = \iiint_D dm = \iiint_D \delta(x,y,z)\,dV$.
2. The *moment about the yz-plane* is $M_{yz} = \iiint_D x\delta(x,y,z)\,dV$.
3. The *moment about the xz-plane* is $M_{xz} = \iiint_D y\delta(x,y,z)\,dV$.
4. The *moment about the xy-plane* is $M_{xy} = \iiint_D z\delta(x,y,z)\,dV$.
5. The *center of mass* of the object is

$$(\overline{x},\overline{y},\overline{z}) = \left(\frac{M_{yz}}{M}, \frac{M_{xz}}{M}, \frac{M_{xy}}{M}\right).$$

Example 15.6.31 Finding the center of mass of a solid.

Find the mass and center of mass of the solid represented by the space region bounded by the coordinate planes and $z = 2-y/3-2x/3$, shown in Figure 15.6.32, with constant density $\delta(x,y,z) = 3\,\text{g/cm}^3$. (Note: this space region was used in Example 15.6.9.)

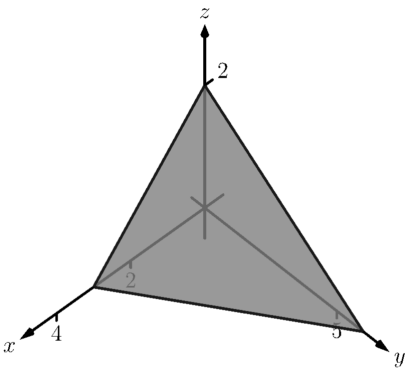


Figure 15.6.32 Finding the center of mass of this solid in Example 15.6.31

Solution. We apply Definition 15.6.30. In Example 15.6.9, we found bounds for the order of integration $dz\,dy\,dx$ to be $0 \le z \le 2 - y/3 - 2x/3$, $0 \le y \le 6-2x$ and $0 \le x \le 3$. We find the mass of the object:

$$\begin{aligned} M &= \iiint_D \delta(x,y,z)\,dV \\ &= \int_0^3 \int_0^{6-2x} \int_0^{2-y/3-2x/3} (3)\,dz\,dy\,dx \\ &= 3\int_0^3 \int_0^{6-2x} \int_0^{2-y/3-2x/3} dz\,dy\,dx \\ &= 3(6) = 18\,\text{g}. \end{aligned}$$

The evaluation of the triple integral is done in Example 15.6.9, so we skipped those steps above. Note how the mass of an object with constant density is simply "density×volume."

We now find the moments about the planes.

$$\begin{aligned} M_{xy} &= \iiint_D 3z\,dV \\ &= \int_0^3 \int_0^{6-2x} \int_0^{2-y/3-2x/3} (3z)\,dz\,dy\,dx \end{aligned}$$

$$
\begin{aligned}
&= \int_0^3 \int_0^{6-2x} \frac{3}{2}\big(2 - y/3 - 2x/3\big)^2 \, dy\, dx \\
&= \int_0^3 -\frac{4}{9}\big(x-3\big)^3 \, dx \\
&= 9.
\end{aligned}
$$

We omit the steps of integrating to find the other moments.

$$
\begin{aligned}
M_{yz} &= \iiint_D 3x \, dV \\
&= \frac{27}{2}. \\
M_{xz} &= \iiint_D 3y \, dV \\
&= 27.
\end{aligned}
$$

The center of mass is

$$(\overline{x}, \overline{y}, \overline{z}) = \left(\frac{27/2}{18}, \frac{27}{18}, \frac{9}{18}\right) = \big(0.75, 1.5, 0.5\big).$$

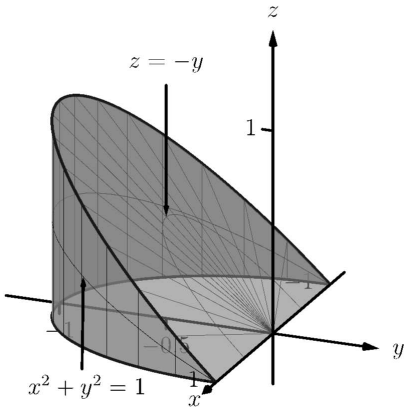


Figure 15.6.34 Finding the center of mass of this solid in Example 15.6.33

Example 15.6.33 Finding the center of mass of a solid.

Find the center of mass of the solid represented by the region bounded by the planes $z = 0$ and $z = -y$ and the cylinder $x^2 + y^2 = 1$, shown in Figure 15.6.34, with density function $\delta(x, y, z) = 10 + x^2 + 5y - 5z$. (Note: this space region was used in Example 15.6.16.)

Solution. As we start, consider the density function. It is symmetric about the yz-plane, and the farther one moves from this plane, the denser the object is. The symmetry indicates that $\overline{x}$ should be 0.

As one moves away from the origin in the y or z directions, the object becomes less dense, though there is more volume in these regions.

Though none of the integrals needed to compute the center of mass are particularly hard, they do require a number of steps. We emphasize here the importance of knowing how to set up the proper integrals; in complex situations we can appeal to technology for a good approximation, if not the exact answer. We use the order of integration $dz\, dy\, dx$, using the bounds found in Example 15.6.16. (As these are the same for all four triple integrals, we explicitly show the bounds only for M.)

$$
\begin{aligned}
M &= \iiint_D \big(10 + x^2 + 5y - 5z\big) \, dV \\
&= \int_{-1}^{1} \int_{-\sqrt{1-x^2}}^{0} \int_0^{-y} \big(10 + x^2 + 5y - 5z\big) \, dV \\
&= \frac{64}{5} - \frac{15\pi}{16} \approx 3.855. \\
M_{yz} &= \iiint_D x\big(10 + x^2 + 5y - 5z\big) \, dV \\
&= 0.
\end{aligned}
$$

$$\begin{aligned} M_{xz} &= \iiint_D y\big(10 + x^2 + 5y - 5z\big)\, dV \\ &= 2 - \frac{61\pi}{48} \approx -1.99. \\ M_{xy} &= \iiint_D z\big(10 + x^2 + 5y - 5z\big)\, dV \\ &= \frac{61\pi}{96} - \frac{10}{9} \approx 0.885. \end{aligned}$$

Note how $M_{yz} = 0$, as expected. The center of mass is

$$(\overline{x}, \overline{y}, \overline{z}) = \left(0, \frac{-1.99}{3.855}, \frac{0.885}{3.855}\right) \approx (0, -0.516, 0.230).$$

As stated before, there are many uses for triple integration beyond finding volume. When $h(x, y, z)$ describes a rate of change function over some space region D, then $\iiint_D h(x, y, z)\, dV$ gives the total change over D. Our one specific example of this was computing mass; a density function is simply a "rate of mass change per volume" function. Integrating density gives total mass.

While knowing *how to integrate* is important, it is arguably much more important to know *how to set up* integrals. It takes skill to create a formula that describes a desired quantity; modern technology is very useful in evaluating these formulas quickly and accurately.

In Section 15.7, we learn about two new coordinate systems (each related to polar coordinates) that allow us to integrate over closed regions in space more easily than when using rectangular coordinates.

15.6.4 Exercises

Terms and Concepts

1. The strategy for establishing bounds for triple integrals is _____ to _____, _____ to _____ and _____ to _____.

2. Give an informal interpretation of what "$\iiint_D dV$" means.

3. Give two uses of triple integration.

4. If an object has a constant density δ and a volume V, what is its mass?

Problems

Exercise Group. Two functions $f_1(x, y)$ and $f_2(x, y)$ and a region R in the x, y plane are given. Set up and evaluate the double integral that finds the volume between the surfaces given by the graphs of these two functions over R.

5. $f_1(x, y) = 8 - x^2 - y^2$, $f_2(x, y) = 2x + y$; R is the square with corners $(-1, -1)$ and $(1, 1)$.

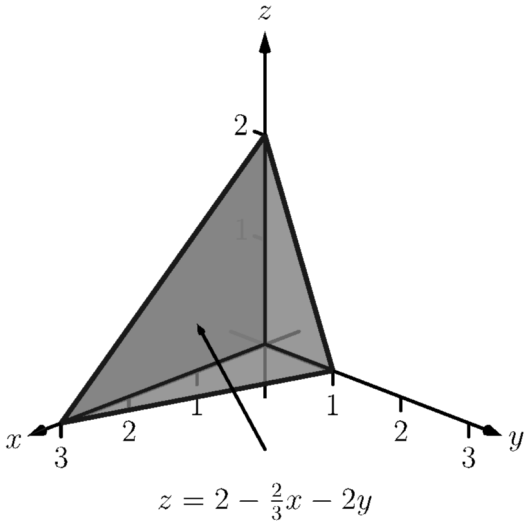
6. $z = f_1(x, y) = x^2 + y^2$ and $z = f_2(x, y) = -x^2 - y^2$; R is the square with corners $(0, 0)$ and $(2, 3)$.

7. $f_1(x, y) = \sin(x)\cos(y)$, $f_2(x, y) = \cos(x)\sin(y) + 2$; R is the triangle with corners $(0, 0)$, $(\pi, 0)$ and (π, π).

8. $f_1(x, y) = 2x^2 + 2y^2 + 3$ and $f_2(x, y) = 6 - x^2 - y^2$; R is the disc $x^2 + y^2 \leq 1$.

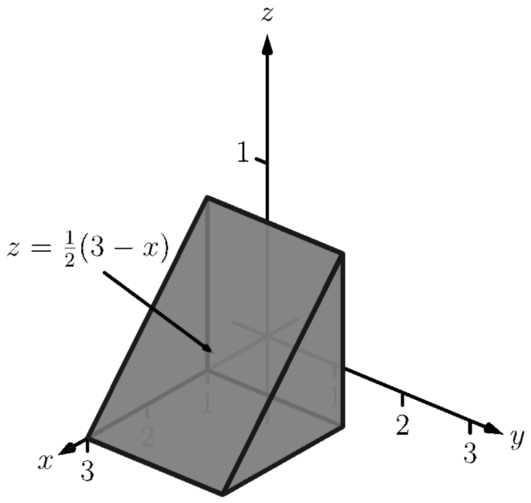
Exercise Group. In the following exercises, a domain D is described by its bounding surfaces, along with a graph. Set up the triple integrals that give the volume of D in all 6 orders of integration, and find the volume of D by evaluating the indicated triple integral.

9. D is bounded by the coordinate planes and $z = 2 - 2x/3 - 2y$.
Evaluate the triple integral with order $dz\,dy\,dx$.

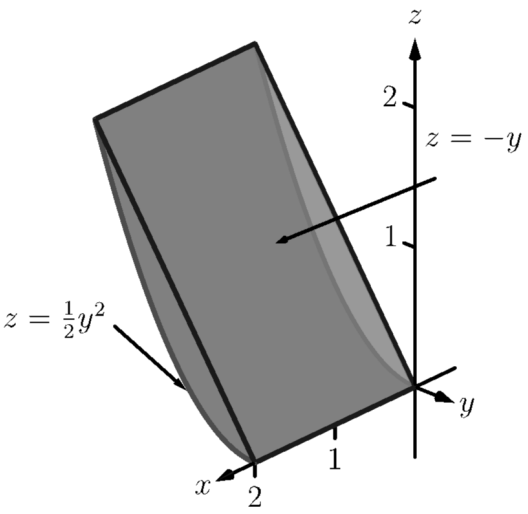


$z = 2 - \frac{2}{3}x - 2y$

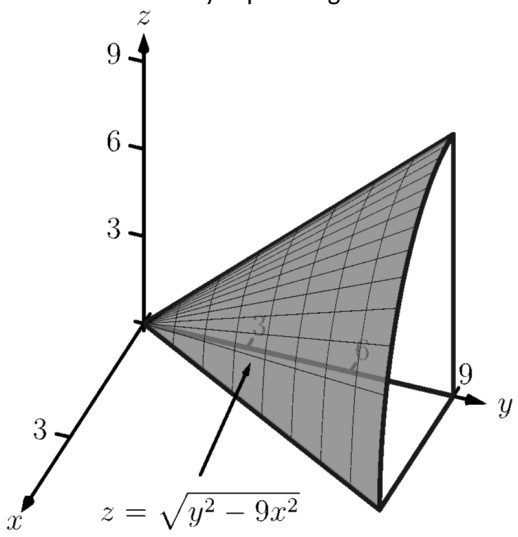
10. D is bounded by the planes $y = 0$, $y = 2$, $x = 1$, $z = 0$ and $z = (3 - x)/2$.
Evaluate the triple integral with order $dx\,dy\,dz$.



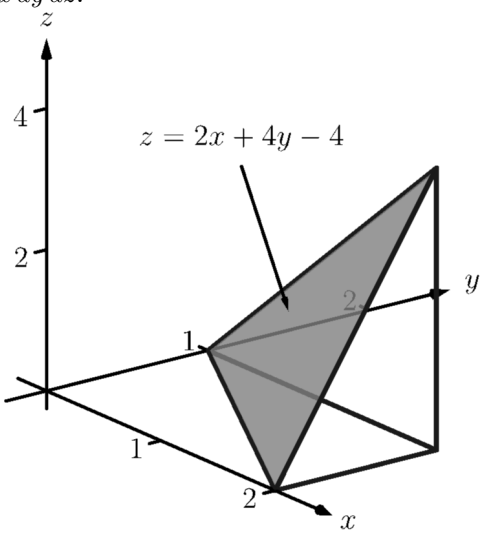
11. D is bounded by the planes $x = 0$, $x = 2$, $z = -y$ and by $z = y^2/2$.
Evaluate the triple integral with the order $dy\,dz\,dx$.



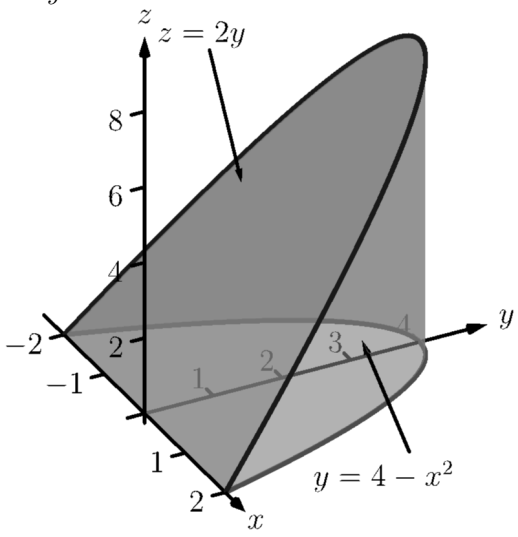
12. D is bounded by the planes $z = 0$, $y = 9$, $x = 0$ and by $z = \sqrt{y^2 - 9x^2}$.
Do not evaluate any triple integral.



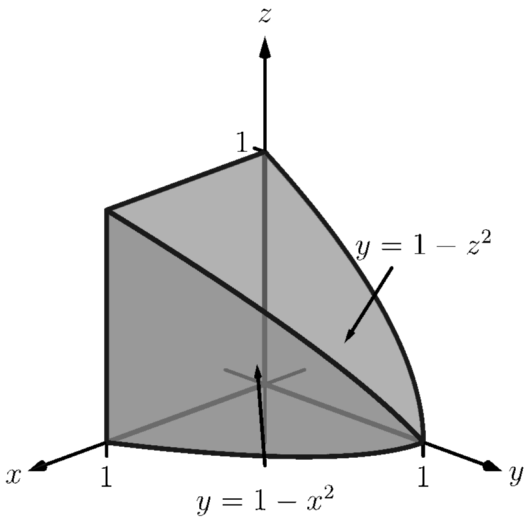
13. D is bounded by the planes $x = 2$, $y = 1$, $z = 0$ and $z = 2x + 4y - 4$.
Evaluate the triple integral with the order $dx\,dy\,dz$.



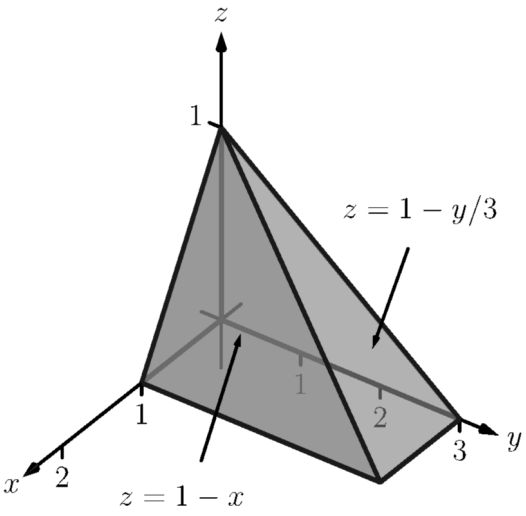
14. D is bounded by the plane $z = 2y$ and by $y = 4 - x^2$.
Evaluate the triple integral with the order $dz\,dy\,dx$.



15. D is bounded by the coordinate planes and by $y = 1 - x^2$ and $y = 1 - z^2$.
Do not evaluate any triple integral. Which order is easier to evaluate: $dz\,dy\,dx$ or $dy\,dz\,dx$? Explain why.



16. D is bounded by the coordinate planes and by $z = 1 - y/3$ and $z = 1 - x$.
Evaluate the triple integral with order $dx\,dy\,dz$.



Exercise Group. In the following exercises, evaluate the triple integral.

17. $\displaystyle\int_{-\pi/2}^{\pi/2}\int_0^{\pi}\int_0^{\pi} \big(\cos(x)\sin(y)\sin(z)\big)\,dz\,dy\,dx$

18. Evaluate $\displaystyle\int_0^1\int_0^x\int_0^{x+y}(x+y+z)\,dz\,dy\,dx$.

19. $\displaystyle\int_0^{\pi}\int_0^1\int_0^z \big(\sin(yz)\big)\,dx\,dy\,dz$

20. Evaluate $\displaystyle\int_{\pi}^{\pi^2}\int_x^{x^3}\int_{-y^2}^{y^2}\left(z\frac{x^2y+y^2x}{e^{x^2+y^2}}\right)dz\,dy\,dx$.

Exercise Group. In the following exercises, find the center of mass of the solid represented by the indicated space region D with density function $\delta(x,y,z)$.

21. D is bounded by the coordinate planes and $z = 2 - 2x/3 - 2y$; $\delta(x,y,z) = 10$ g/cm^3.
(Note: this is the same region as used in Exercise 15.6.9.)

22. D is bounded by the planes $y = 0$, $y = 2$, $x = 1$, $z = 0$ and $z = (3 - x)/2$; $\delta(x,y,z) = 2$ g/cm^3.
(Note: this is the same region as used in Exercise 15.6.10.)

23. D is bounded by the planes $x = 2$, $y = 1$, $z = 0$ and $z = 2x + 4y - 4$;$\delta(x,y,z) = x^2$lb/in^3.
(Note: this is the same region as used in Exercise 15.6.13.)

24. D is bounded by the plane $z = 2y$ and by $y = 4 - x^2$.
$\delta(x,y,z) = y^2$lb/in^3.
(Note: this is the same region as used in Exercise 15.6.14.)

15.7 Triple Integration with Cylindrical and Spherical Coordinates

Just as polar coordinates gave us a new way of describing curves in the plane, in this section we will see how *cylindrical* and *spherical* coordinates give us new ways of desribing surfaces and regions in space.

15.7.1 Cylindrical Coordinates

In short, cylindrical coordinates can be thought of as a combination of the polar and rectangular coordinate systems. One can identify a point (x_0, y_0, z_0), given in rectangular coordinates, with the point (r_0, θ_0, z_0), given in cylindrical coordinates, where the z-value in both systems is the same, and the point (x_0, y_0) in the xy-plane is identified with the polar point $P(r_0, \theta_0)$; see Figure 15.7.1. So that each point in space that does not lie on the z-axis is defined uniquely, we will restrict $r \geq 0$ and $0 \leq \theta \leq 2\pi$.

We use the identity $z = z$ along with the identities found in Key Idea 10.4.6 to convert between the rectangular coordinate (x, y, z) and the cylindrical coordinate (r, θ, z), namely:

From rectangular to cylindrical: $r = \sqrt{x^2 + y^2}$, $\tan(\theta) = y/x$ and $z = z$;
From cylindrical to rectangular: $x = r\cos(\theta)$, $y = r\sin(\theta)$ and $z = z$.

These identities, along with conversions related to spherical coordinates, are given later in Key Idea 15.7.15.

Example 15.7.3 Converting between rectangular and cylindrical coordinates.

Convert the rectangular point $(2, -2, 1)$ to cylindrical coordinates, and convert the cylindrical point $(4, 3\pi/4, 5)$ to rectangular.

Solution. Following the identities given above (and, later in Key Idea 15.7.15), we have $r = \sqrt{2^2 + (-2)^2} = 2\sqrt{2}$. Using $\tan(\theta) = y/x$, we find $\theta = \tan^{-1}(-2/2) = -\pi/4$. As we restrict θ to being between 0 and 2π, we set $\theta = 7\pi/4$. Finally, $z = 1$, giving the cylindrical point $(2\sqrt{2}, 7\pi/4, 1)$.

In converting the cylindrical point $(4, 3\pi/4, 5)$ to rectangular, we have $x = 4\cos\left(3\pi/4\right) = -2\sqrt{2}$, $y = 4\sin\left(3\pi/4\right) = 2\sqrt{2}$ and $z = 5$, giving the rectangular point $(-2\sqrt{2}, 2\sqrt{2}, 5)$.

Setting each of r, θ and z equal to a constant defines a surface in space, as illustrated in the following example.

Example 15.7.4 Canonical surfaces in cylindrical coordinates.

Describe the surfaces $r = 1$, $\theta = \pi/3$ and $z = 2$, given in cylindrical coordinates.

Solution. The equation $r = 1$ describes all points in space that are 1 unit away from the z-axis. This surface is a "tube" or "cylinder" of radius 1, centered on the z-axis, as graphed in Figure 12.1.17 (which describes the cylinder $x^2 + y^2 = 1$ in space).

The equation $\theta = \pi/3$ describes the plane formed by extending the line $\theta = \pi/3$, as given by polar coordinates in the xy-plane, parallel to the z-axis.

The equation $z = 2$ describes the plane of all points in space that are 2 units above the xy-plane. This plane is the same as the plane described

Both coordinate systems provide ways of extending polar coordinates in the plane to three dimensions. The reader is warned that while conventions are fairly standard for cylindrical coordinates, there are many different conventions for spherical coordinates. Mathematics, physics, and engineering all use slightly different versions. The definition presented in this text also differs from the usual mathematical definition. Why add yet another competing standard? The definition we present is slightly easier to work with, and should also be more familiar, since it corresponds to the coordinate system of latitude and longitude used to describe locations on the Earth.

youtu.be/watch?v=DOFhlADkGko

Figure 15.7.2 Explaining the cylindrical coordinate system

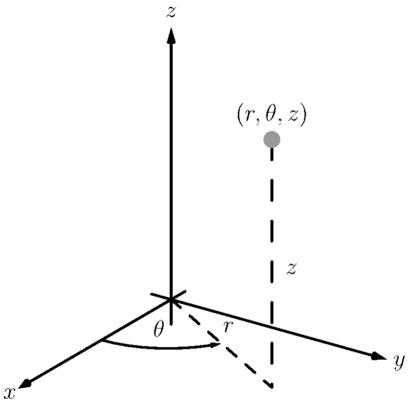


Figure 15.7.1 Illustrating the principles behind cylindrical coordinates

Our rectangular to polar conversion formulas used $r^2 = x^2 + y^2$, allowing for negative r values. Since we now restrict $r \geq 0$, we can use $r = \sqrt{x^2 + y^2}$.

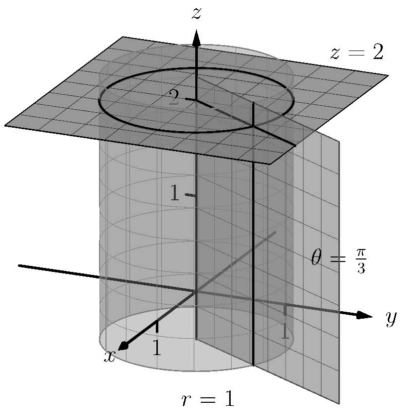


Figure 15.7.5 Graphing the canonical surfaces in cylindrical coordinates from Example 15.7.4

by $z = 2$ in rectangular coordinates.

All three surfaces are graphed in Figure 15.7.5. Note how their intersection uniquely defines the point $P = (1, \pi/3, 2)$.

Cylindrical coordinates are useful when describing certain domains in space, allowing us to evaluate triple integrals over these domains more easily than if we used rectangular coordinates.

Theorem 15.6.29 shows how to evaluate $\iiint_D h(x, y, z)\, dV$ using rectangular coordinates. In that evaluation, we use $dV = dz\, dy\, dx$ (or one of the other five orders of integration). Recall how, in this order of integration, the bounds on y are "curve to curve" and the bounds on x are "point to point": these bounds describe a region R in the xy-plane. We could describe R using polar coordinates as done in Section 15.3. In that section, we saw how we used $dA = r\, dr\, d\theta$ instead of $dA = dy\, dx$.

Considering the above thoughts, we have $dV = dz\big(r\, dr\, d\theta\big) = r\, dz\, dr\, d\theta$. We set bounds on z as "surface to surface" as done in the previous section, and then use "curve to curve" and "point to point" bounds on r and θ, respectively. Finally, using the identities given above, we change the integrand $h(x, y, z)$ to $h(r, \theta, z)$.

This process should sound plausible; the following theorem states it is truly a way of evaluating a triple integral.

Theorem 15.7.6 Triple Integration in Cylindrical Coordinates.

Let $w = h(r, \theta, z)$ be a continuous function on a closed, bounded region D in space, bounded in cylindrical coordinates by $\alpha \le \theta \le \beta$, $g_1(\theta) \le r \le g_2(\theta)$ and $f_1(r, \theta) \le z \le f_2(r, \theta)$. Then

$$\iiint_D h(r, \theta, z)\, dV = \int_\alpha^\beta \int_{g_1(\theta)}^{g_2(\theta)} \int_{f_1(r,\theta)}^{f_2(r,\theta)} h(r, \theta, z) r\, dz\, dr\, d\theta.$$

youtu.be/watch?v=u0FYs9MaI5k

Figure 15.7.7 Using cylindrical coordinates to find the volume of a cone

Example 15.7.8 Evaluating a triple integral with cylindrical coordinates.

Find the mass of the solid represented by the region in space bounded by $z = 0$, $z = \sqrt{4 - x^2 - y^2} + 3$ and the cylinder $x^2 + y^2 = 4$ (as shown in Figure 15.7.9), with density function $\delta(x, y, z) = x^2 + y^2 + z + 1$, using a triple integral in cylindrical coordinates. Distances are measured in centimeters and density is measured in grams per cubic centimeter.

Solution. We begin by describing this region of space with cylindrical coordinates. The plane $z = 0$ is left unchanged; with the identity $r = \sqrt{x^2 + y^2}$, we convert the hemisphere of radius 2 to the equation $z = \sqrt{4 - r^2}$; the cylinder $x^2 + y^2 = 4$ is converted to $r^2 = 4$, or, more simply, $r = 2$. We also convert the density function: $\delta(r, \theta, z) = r^2 + z + 1$.

To describe this solid with the bounds of a triple integral, we bound z with $0 \le z \le \sqrt{4 - r^2} + 3$; we bound r with $0 \le r \le 2$; we bound θ with $0 \le \theta \le 2\pi$.

Using Definition 15.6.30 and Theorem 15.7.6, we have the mass of the solid is

$$M = \iiint_D \delta(x, y, z)\, dV = \int_0^{2\pi} \int_0^2 \int_0^{\sqrt{4-r^2}+3} \big(r^2 + z + 1\big) r\, dz\, dr\, d\theta$$

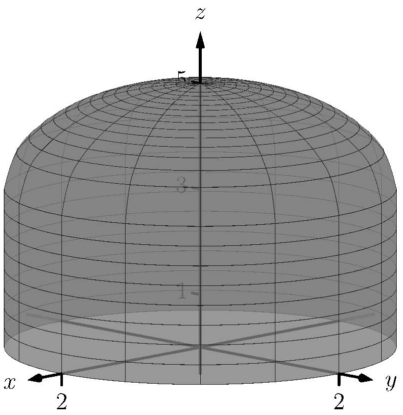


Figure 15.7.9 Visualizing the solid used in Example 15.7.8

$$= \int_0^{2\pi}\int_0^2 \left((r^3+4r)\sqrt{4-r^2} + \frac{5}{2}r^3 + \frac{19}{2}r\right) dr\, d\theta$$
$$= \frac{1318\pi}{15} \approx 276.04 \text{ g},$$

where we leave the details of the remaining double integral to the reader.

Example 15.7.10 Finding the center of mass using cylindrical coordinates.

Find the center of mass of the solid with constant density whose base can be described by the polar curve $r = \cos(3\theta)$ and whose top is defined by the plane $z = 1 - x + 0.1y$, where distances are measured in feet, as seen in Figure 15.7.11. (The volume of this solid was found in Example 15.3.13.)

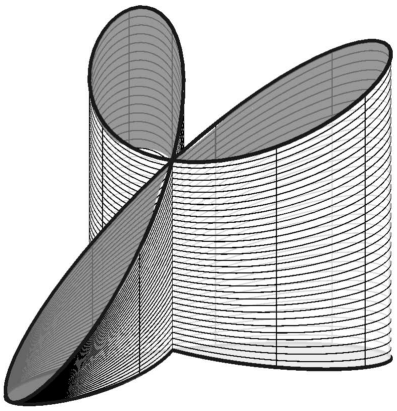


Figure 15.7.11 Visualizing the solid used in Example 15.7.10

Solution. We convert the equation of the plane to use cylindrical coordinates: $z = 1 - r\cos(\theta) + 0.1r\sin(\theta)$. Thus the region is space is bounded by $0 \le z \le 1 - r\cos(\theta) + 0.1r\sin(\theta)$, $0 \le r \le \cos(3\theta)$, $0 \le \theta \le \pi$ (recall that the rose curve $r = \cos(3\theta)$ is traced out once on $[0, \pi]$.

Since density is constant, we set $\delta = 1$ and finding the mass is equivalent to finding the volume of the solid. We set up the triple integral to compute this but do not evaluate it; we leave it to the reader to confirm it evaluates to the same result found in Example 15.3.13.

$$M = \iiint_D \delta\, dV = \int_0^{\pi}\int_0^{\cos(3\theta)}\int_0^{1-r\cos(\theta)+0.1r\sin(\theta)} r\, dz\, dr\, d\theta = \frac{\pi}{4}.$$

From Definition 15.6.30 we set up the triple integrals to compute the moments about the three coordinate planes. The computation of each is left to the reader (using technology is recommended):

$$M_{yz} = \iiint_D x\, dV = \int_0^{\pi}\int_0^{\cos(3\theta)}\int_0^{1-r\cos(\theta)+0.1r\sin(\theta)} (r\cos(\theta))r\, dz\, dr\, d\theta$$
$$= \frac{-3\pi}{64} \approx -0.147.$$

$$M_{xz} = \iiint_D y\, dV = \int_0^{\pi}\int_0^{\cos(3\theta)}\int_0^{1-r\cos(\theta)+0.1r\sin(\theta)} (r\sin(\theta))r\, dz\, dr\, d\theta$$
$$= \frac{3\pi}{640} \approx 0.015.$$

$$M_{xy} = \iiint_D z\, dV = \int_0^{\pi}\int_0^{\cos(3\theta)}\int_0^{1-r\cos(\theta)+0.1r\sin(\theta)} (z)r\, dz\, dr\, d\theta$$
$$= \frac{1903\pi}{12800} \approx 0.467.$$

The center of mass in rectangular coordinates, found by dividing the respective moments by the mass, is approximately located at $(-0.188, 0.019, 0.595)$, which lies outside the bounds of the solid.

15.7.2 Spherical Coordinates

In short, spherical coordinates can be thought of as a "double application" of the polar coordinate system. In spherical coordinates, a point P is identified with (ρ, θ, φ), where ρ is the distance from the origin to P, θ is the same angle as would be used to describe P in the cylindrical coordinate system, and φ is the angle between the xy-plane and the ray from the origin to P; see Figure 15.7.13. So that each point in space that does not lie on the z-axis is defined uniquely, we will restrict $\rho \geq 0$, $0 \leq \theta \leq 2\pi$ and $-\pi/2 \leq \varphi \leq \pi/2$.

The symbol ρ is the Greek letter "rho." Traditionally it is used in the spherical coordinate system, while r is used in the polar and cylindrical coordinate systems.

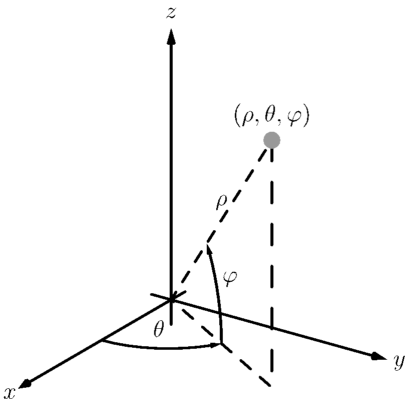


Figure 15.7.13 Illustrating the principles behind spherical coordinates

youtu.be/watch?v=In5up-a1jMI

Figure 15.7.14 Introducing spherical coordinates

Convention 15.7.12 Note that most mathematics textbooks define φ to be measured from the positive z-axis, with values in $[0, \pi]$, rather than from the xy-plane.

We have chosen our convention with a number of considerations in mind:

- The coordinates (ρ, θ, φ) form a **right-handed** coordinate system: one in which the orientation matches that of our usual (x, y, z) coordinates, where the "right-hand rule" applies. If φ is measured from the z-axis, the order (ρ, φ, θ) is needed to get a right-handed system.
- Points of the form $(a, \alpha, 0)$ are the same in both cylindrical and spherical coordinates.
- Some integration problems become slightly easier: we will see soon that the volume element in spherical coordinates involves $\cos(\varphi)$, which integrates to $\sin(\varphi)$. In the usual convention, the volume element involves $\sin(\varphi)$, which integrates to $-\cos(\varphi)$ – a source of many common sign errors.

Students of Physics will encounter yet another convention. In Physics, the variable r is preferred as the radial coordinate, and spherical coordinates are given as (r, θ, φ); however, in Physics, φ becomes the angle in the xy-plane, while θ is the angle measured from the positive z-axis.

Note that the angle in the xy-plane (θ, in our case) is known as the **azimuthal angle**. Our angle φ is known as the **elevation angle**. The angle used in other conventions that is measured from the positive z-axis (often identified with the north pole) is known as the **polar angle**. For further discussion, the Wikipedia article[1] is quite useful.

The following Key Idea gives conversions to/from our three spatial coordinate systems.

Key Idea 15.7.15 Converting Between Rectangular, Cylindrical and Spherical Coordinates.

- *Rectangular and Cylindrical.*

$$\begin{aligned} r^2 &= x^2 + y^2, & \tan(\theta) &= y/x, & z &= z \\ x &= r\cos(\theta), & y &= r\sin(\theta), & z &= z \end{aligned}$$

- *Rectangular and Spherical.*

$$\begin{aligned} &\rho = \sqrt{x^2 + y^2 + z^2}, \quad \tan(\theta) = y/x, \quad \sin(\varphi) = z/\sqrt{x^2 + y^2 + z^2} \\ &x = \rho\cos(\varphi)\cos(\theta), \quad y = \rho\cos(\varphi)\sin(\theta), \quad z = \rho\sin(\varphi) \end{aligned}$$

[1] en.wikipedia.org/wiki/Spherical_coordinate_system

- *Cylindrical and Spherical.*

$$\begin{array}{lll} \rho = \sqrt{r^2+z^2}, & \theta = \theta, & \tan(\varphi) = z/r \\ r = \rho\cos(\varphi), & \theta = \theta, & z = \rho\sin(\varphi) \end{array}$$

Example 15.7.16 Converting between rectangular and spherical coordinates.

Convert the rectangular point $(2,-2,1)$ to spherical coordinates, and convert the spherical point $(6,\pi/3,0)$ to rectangular and cylindrical coordinates.

Solution. This rectangular point is the same as used in Example 15.7.3. Using Key Idea 15.7.15, we find $\rho = \sqrt{2^2+(-1)^2+1^2} = 3$. Using the same logic as in Example 15.7.3, we find $\theta = 7\pi/4$. Finally, $\sin(\varphi) = 1/3$, giving $\varphi = \sin^{-1}(1/3) \approx 0.34$, or about $19.47°$. Thus the spherical coordinates are approximately $(3, 7\pi/4, 0.34)$.

Converting the spherical point $(6,\pi/3,0)$ to rectangular, we have $x = 6\cos(0)\cos(\pi/3) = 3$, $y = 6\cos(0)\sin(\pi/3) = 3\sqrt{3}$ and $z = 6\sin(0) = 0$. Thus the rectangular coordinates are $(3, 3\sqrt{3}, 0)$.

To convert this spherical point to cylindrical, we have $r = 6\cos(0) = 6$, $\theta = \pi/3$ and $z = 6\sin(0) = 0$, giving the cylindrical point $(6,\pi/3,0)$.

The videos used in this section (and later) were recorded for an earlier version of the textbook, that used the polar angle instead of the elevation angle. Some of the details will therefore differ from those presented in the text. (The reader can consider this an opportunity to learn — and compare — both conventions.)

Example 15.7.17 Canonical surfaces in spherical coordinates.

Describe the surfaces $\rho = 1$, $\theta = \pi/3$ and $\varphi = \pi/3$, given in spherical coordinates.

Solution. The equation $\rho = 1$ describes all points in space that are 1 unit away from the origin: this is the sphere of radius 1, centered at the origin.

The equation $\theta = \pi/3$ describes the same surface in spherical coordinates as it does in cylindrical coordinates: beginning with the line $\theta = \pi/3$ in the xy-plane as given by polar coordinates, extend the line parallel to the z-axis, forming a plane.

The equation $\varphi = \pi/3$ describes all points P in space where the ray from the origin to P makes an angle of $\pi/3$ with the xy-plane. This describes a cone, with the positive z-axis its axis of symmetry, with point at the origin.

All three surfaces are graphed in Figure 15.7.18. Note how their intersection uniquely defines the point $P = (1,\pi/3,\pi/6)$.

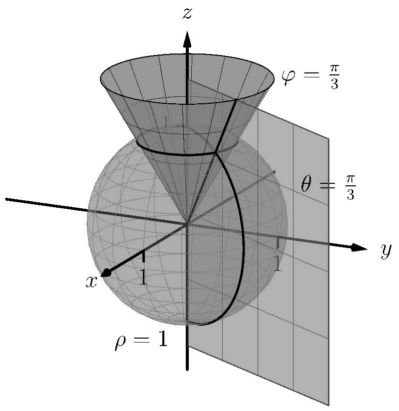


Figure 15.7.18 Graphing the canonical surfaces in spherical coordinates from Example 15.7.17

Spherical coordinates are useful when describing certain domains in space, allowing us to evaluate triple integrals over these domains more easily than if we used rectangular coordinates or cylindrical coordinates. The crux of setting up a triple integral in spherical coordinates is appropriately describing the "small amount of volume," dV, used in the integral.

Considering Figure 15.7.19, we can make a small "spherical wedge" by varying ρ, θ and φ each a small amount, $\Delta\rho$, $\Delta\theta$ and $\Delta\varphi$, respectively. This wedge is approximately a rectangular solid when the change in each coordinate is small, giving a volume of about

$$\Delta V \approx \Delta\rho \times \rho\Delta\varphi \times \rho\cos(\varphi)\Delta\theta.$$

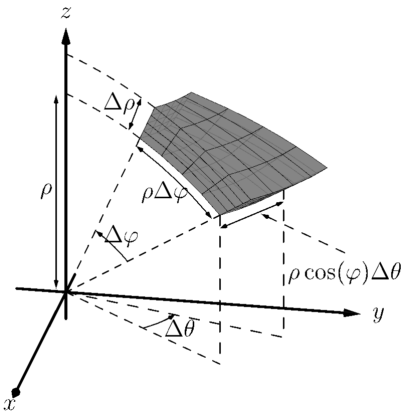


Figure 15.7.19 Approximating the volume of a standard region in space using spherical coordinates

Given a region D in space, we can approximate the volume of D with many such wedges. As the size of each of $\Delta\rho$, $\Delta\theta$ and $\Delta\varphi$ goes to zero, the number of wedges increases to infinity and the volume of D is more accurately approximated, giving

$$dV = d\rho \ \times \ \rho\, d\varphi \ \times \ \rho \sin(\varphi) d\theta = \rho^2 \sin(\varphi)\, d\rho\, d\theta\, d\varphi.$$

Again, this development of dV should sound reasonable, and the following theorem states it is the appropriate manner by which triple integrals are to be evaluated in spherical coordinates.

Theorem 15.7.20 Triple Integration in Spherical Coordinates.

Let $w = h(\rho, \theta, \varphi)$ be a continuous function on a closed, bounded region D in space, bounded in spherical coordinates by $\alpha_1 \le \varphi \le \alpha_2$, $\beta_1 \le \theta \le \beta_2$ and $f_1(\theta, \varphi) \le \rho \le f_2(\theta, \varphi)$. Then

$$\iiint_D h(\rho, \theta, \varphi)\, dV = \int_{\alpha_1}^{\alpha_2} \int_{\beta_1}^{\beta_2} \int_{f_1(\theta,\varphi)}^{f_2(\theta,\varphi)} h(\rho, \theta, \varphi)\rho^2 \cos(\varphi)\, d\rho\, d\theta\, d\varphi.$$

It is generally most intuitive to evaluate the triple integral in Theorem 15.7.20 by integrating with respect to ρ first; it often does not matter whether we next integrate with respect to θ or φ. Different texts present different standard orders, some preferring $d\varphi\, d\theta$ instead of $d\theta\, d\varphi$. As the bounds for these variables are usually constants in practice, it generally is a matter of preference.

Example 15.7.21 Establishing the volume of a sphere.

Let D be the region in space bounded by the sphere, centered at the origin, of radius r. Use a triple integral in spherical coordinates to find the volume V of D.

Solution. The sphere of radius r, centered at the origin, has equation $\rho = r$. To obtain the full sphere, the bounds on θ and φ are $0 \le \theta \le 2\pi$ and $-\pi/2 \le \varphi \le \pi/2$. This leads us to:

$$\begin{aligned}
V &= \iiint_D dV \\
&= \int_{-\pi/2}^{\pi/2} \int_0^{2\pi} \int_0^r \left(\rho^2 \cos(\varphi)\right) d\rho\, d\theta\, d\varphi \\
&= \int_{-\pi/2}^{\pi/2} \int_0^{2\pi} \left(\frac{1}{3}\rho^3 \cos(\varphi)\Big|_0^r\right) d\theta\, d\varphi \\
&= \int_{-\pi/2}^{\pi/2} \int_0^{2\pi} \left(\frac{1}{3}r^3 \cos(\varphi)\right) d\theta\, d\varphi \\
&= \int_{-\pi/2}^{\pi/2} \left(\frac{2\pi}{3}r^3 \cos(\varphi)\right) d\varphi \\
&= \left(\frac{2\pi}{3}r^3 \sin(\varphi)\right)\Big|_{-\pi/2}^{\pi/2} \\
&= \frac{4\pi}{3}r^3,
\end{aligned}$$

the familiar formula for the volume of a sphere. Note how the integration steps were easy, not using square roots nor integration steps such as Substitution.

youtu.be/watch?v=wqjjslo0aul

Figure 15.7.22 Setting up an integral in spherical coorindates

Example 15.7.23 Finding the center of mass using spherical coordinates.

Find the center of mass of the solid with constant density enclosed above by $\rho = 4$ and below by $\varphi = \pi/3$, as illustrated in Figure 15.7.24.

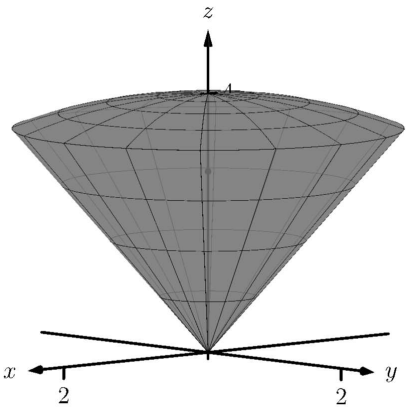


Figure 15.7.24 Graphing the solid, and its center of mass, from Example 15.7.23

Solution. We will set up the four triple integrals needed to find the center of mass (i.e., to compute M, M_{yz}, M_{xz} and M_{xy}) and leave it to the reader to evaluate each integral. Because of symmetry, we expect the x- and y- coordinates of the center of mass to be 0.

While the surfaces describing the solid are given in the statement of the problem, to describe the full solid D, we use the following bounds: $0 \le \rho \le 4$, $0 \le \theta \le 2\pi$ and $0 \le \varphi \le \pi/3$. Since density δ is constant, we assume $\delta = 1$.

The mass of the solid:

$$\begin{aligned} M &= \iiint_D dm = \iiint_D dV \\ &= \int_0^{\pi/3}\int_0^{2\pi}\int_0^4 \left(\rho^2 \cos(\varphi)\right) d\rho\, d\theta\, d\varphi \\ &= \frac{64}{3}\left(2-\sqrt{3}\right)\pi \approx 17.958. \end{aligned}$$

To compute M_{yz}, the integrand is x; using Key Idea 15.7.15, we have $x = \rho\cos(\varphi)\cos(\theta)$. This gives:

$$\begin{aligned} M_{yz} &= \iiint_D x\, dm \\ &= \int_0^{\pi/3}\int_0^{2\pi}\int_0^4 \left((\rho\cos(\varphi)\cos(\theta))\rho^2\cos(\varphi)\right) d\rho\, d\theta\, d\varphi \\ &= \int_0^{\pi/3}\int_0^{2\pi}\int_0^4 \left(\rho^3\cos^2(\varphi)\cos(\theta)\right) d\rho\, d\theta\, d\varphi \\ &= 0, \end{aligned}$$

which we expected as we expect $\overline{x} = 0$.

To compute M_{xz}, the integrand is y; using Key Idea 15.7.15, we have $y = \rho\cos(\varphi)\sin(\theta)$. This gives:

$$\begin{aligned} M_{xz} &= \iiint_D y\, dm \\ &= \int_0^{\pi/3}\int_0^{2\pi}\int_0^4 \left((\rho\cos(\varphi)\sin(\theta))\rho^2\cos(\varphi)\right) d\rho\, d\theta\, d\varphi \\ &= \int_0^{\pi/3}\int_0^{2\pi}\int_0^4 \left(\rho^3\cos^2(\varphi)\sin(\theta)\right) d\rho\, d\theta\, d\varphi \\ &= 0, \end{aligned}$$

which we also expected as we expect $\overline{y} = 0$.

To compute M_{xy}, the integrand is z; using Key Idea 15.7.15, we have $z = \rho\sin(\varphi)$. This gives:

$$\begin{aligned} M_{xy} &= \iiint_D z\, dm \\ &= \int_0^{\pi/3}\int_0^{2\pi}\int_0^4 \left((\rho\sin(\varphi))\rho^2\cos(\varphi)\right) d\rho\, d\theta\, d\varphi \end{aligned}$$

youtu.be/watch?v=jwzvDiiM94o

Figure 15.7.25 Spherical coordinates example: volume of a spherical wedge

youtu.be/watch?v=EKxY_5PlgTU

Figure 15.7.26 Spherical coordinates example: volume of a conical cup

$$= \int_0^{\pi/3} \int_0^{2\pi} \int_0^4 \left(\rho^3 \sin(\varphi)\cos(\varphi)\right) d\rho\, d\theta\, d\varphi$$
$$= 16\pi \approx 50.266.$$

Thus the center of mass is $(0, 0, M_{xy}/M) \approx (0, 0, 2.799)$, as indicated in Figure 15.7.24.

This section has provided a brief introduction into two new coordinate systems useful for identifying points in space. Each can be used to define a variety of surfaces in space beyond the canonical surfaces graphed as each system was introduced.

However, the usefulness of these coordinate systems does not lie in the variety of surfaces that they can describe nor the regions in space these surfaces may enclose. Rather, cylindrical coordinates are mostly used to describe cylinders and spherical coordinates are mostly used to describe spheres. These shapes are of special interest in the sciences, especially in physics, and computations on/inside these shapes is difficult using rectangular coordinates. For instance, in the study of electricity and magnetism, one often studies the effects of an electrical current passing through a wire; that wire is essentially a cylinder, described well by cylindrical coordinates.

This chapter investigated the natural follow-on to partial derivatives: iterated integration. We learned how to use the bounds of a double integral to describe a region in the plane using both rectangular and polar coordinates, then later expanded to use the bounds of a triple integral to describe a region in space. We used double integrals to find volumes under surfaces, surface area, and the center of mass of lamina; we used triple integrals as an alternate method of finding volumes of space regions and also to find the center of mass of a region in space.

Integration does not stop here. We could continue to iterate our integrals, next investigating "quadruple integrals" whose bounds describe a region in 4-dimensional space (which are very hard to visualize). We can also look back to "regular" integration where we found the area under a curve in the plane. A natural analogue to this is finding the "area under a curve," where the curve is in space, not in a plane. These are just two of many avenues to explore under the heading of "integration."

15.7.3 Exercises

Terms and Concepts

1. Explain the difference between the roles r, in cylindrical coordinates, and ρ, in spherical coordinates, play in determining the location of a point.

2. Why are points on the z-axis not determined uniquely when using cylindrical and spherical coordinates?

3. What surfaces are naturally defined using cylindrical coordinates?

4. What surfaces are naturally defined using spherical coordinates?

Problems

Exercise Group. In the following exercises, points are given in either the rectangular, cylindrical or spherical coordinate systems. Find the coordinates of the points in the other systems.

5.

(a) Points in rectangular coordinates: $(2, 2, 1)$ and $(-\sqrt{3}, 1, 0)$

(b) Points in cylindrical coordinates: $(2, \pi/4, 2)$ and $(3, 3\pi/2, -4)$

(c) Points in spherical coordinates: $(2, \pi/4, \pi/4)$ and $(1, 0, 0)$

6.

(a) Points in rectangular coordinates: $(0, 1, 1)$ and $(-1, 0, 1)$

(b) Points in cylindrical coordinates: $(0, \pi, 1)$ and $(2, 4\pi/3, 0)$

(c) Points in spherical coordinates: $(2, \pi/6, 0)$ and $(3, \pi, -\pi/2)$

Exercise Group. In the following exercises, describe the curve, surface or region in space determined by the given bounds in cylindrical coordinates.

7.

(a) $r = 1, 0 \leq \theta \leq 2\pi, 0 \leq z \leq 1$

(b) $1 \leq r \leq 2, 0 \leq \theta \leq \pi, 0 \leq z \leq 1$

8.

(a) $1 \leq r \leq 2, \theta = \pi/2, 0 \leq z \leq 1$

(b) $r = 2, 0 \leq \theta \leq 2\pi, z = 5$

Exercise Group. In the following exercises, describe the curve, surface or region in space determined by the given bounds in spherical coordinates.

9.

(a) $\rho = 3, 0 \leq \theta \leq 2\pi, 0 \leq \varphi \leq \pi/2$

(b) $2 \leq \rho \leq 3, 0 \leq \theta \leq 2\pi,$ $-\pi/2 \leq \varphi \leq \pi/2$

10.

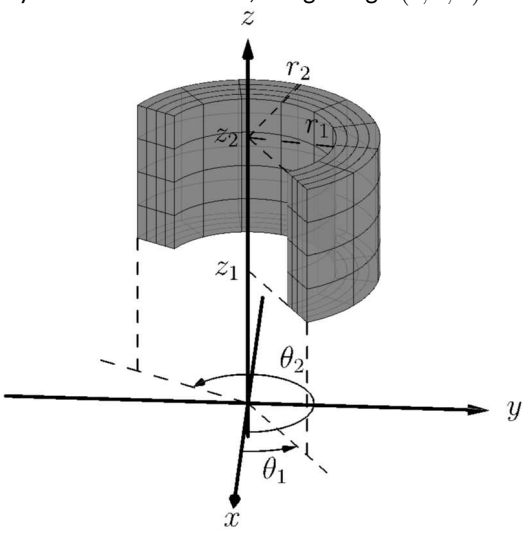
(a) $0 \leq \rho \leq 2, 0 \leq \theta \leq \pi, \varphi = \pi/4$

(b) $\rho = 2, 0 \leq \theta \leq 2\pi, \varphi = \pi/3$

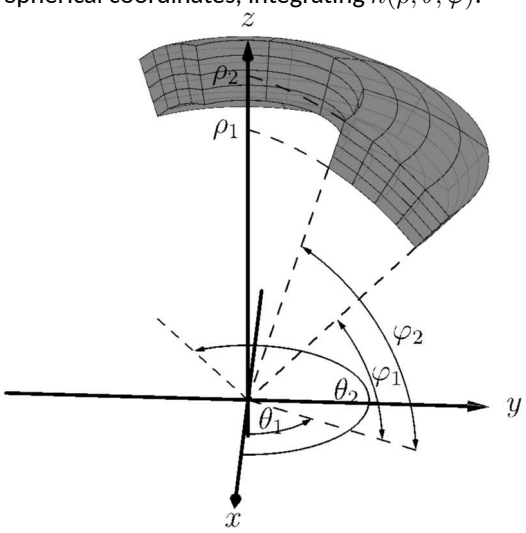
(c) This is a curve, a circle of radius 1 centered at $(0, 0, \sqrt{3})$, lying parallel to the xy-plane.

Exercise Group. In the following exercises, standard regions in space, as defined by cylindrical and spherical coordinates, are shown. Set up the triple integral that integrates the given function over the graphed region.

11. Cylindrical coordinates, integrating $h(r, \theta, z)$:



12. Spherical coordinates, integrating $h(\rho, \theta, \varphi)$:



Exercise Group. In the following exercises, a triple integral in cylindrical coordinates is given. Describe the region in space defined by the bounds of the integral.

13. $\displaystyle\int_0^{\pi/2}\int_0^{2}\int_0^{2} r\,dz\,dr\,d\theta$

14. $\displaystyle\int_0^{2\pi}\int_3^{4}\int_0^{5} r\,dz\,dr\,d\theta$

15. $\displaystyle\int_0^{2\pi}\int_0^{1}\int_0^{1-r} r\,dz\,dr\,d\theta$

16. $\displaystyle\int_0^{\pi}\int_0^{1}\int_0^{2-r} r\,dz\,dr\,d\theta$

17. $\displaystyle\int_0^{\pi}\int_0^{3}\int_0^{\sqrt{9-r^2}} r\,dz\,dr\,d\theta$

18. $\displaystyle\int_0^{2\pi}\int_0^{a}\int_0^{\sqrt{a^2-r^2}+b} r\,dz\,dr\,d\theta$

Exercise Group. In the following exercises, a triple integral in spherical coordinates is given. Describe the region in space defined by the bounds of the integral.

19. $\displaystyle\int_0^{\pi/2}\int_0^{\pi/2}\int_0^{1} \rho^2\cos(\varphi)\,d\rho\,d\theta\,d\varphi$

20. $\displaystyle\int_{-\pi/2}^{\pi/2}\int_0^{\pi}\int_1^{1.1} \rho^2\cos(\varphi)\,d\rho\,d\theta\,d\varphi$

21. $\displaystyle\int_{\pi/4}^{\pi/2}\int_0^{2\pi}\int_0^{2} \rho^2\cos(\varphi)\,d\rho\,d\theta\,d\varphi$

22. $\displaystyle\int_{\pi/4}^{\pi/3}\int_0^{2\pi}\int_0^{2} \rho^2\cos(\varphi)\,d\rho\,d\theta\,d\varphi$

23. $\displaystyle\int_{\pi/3}^{\pi/2}\int_0^{2\pi}\int_0^{\csc(\varphi)} \rho^2\cos(\varphi)\,d\rho\,d\theta\,d\varphi$

24. $\displaystyle\int_{\pi/3}^{\pi/2}\int_0^{2\pi}\int_0^{a\csc(\varphi)} \rho^2\cos(\varphi)\,d\rho\,d\theta\,d\varphi$

Exercise Group. In the following exercises, a solid is described along with its density function. Find the mass of the solid using cylindrical coordinates.

25. Bounded by the cylinder $x^2 + y^2 = 4$ and the planes $z = 0$ and $z = 4$ with density function $\delta(x, y, z) = \sqrt{x^2 + y^2} + 1$.

26. Bounded by the cylinders $x^2 + y^2 = 4$ and $x^2 + y^2 = 9$, between the planes $z = 0$ and $z = 10$ with density function $\delta(x, y, z) = z$.

27. Bounded by $y \geq 0$, the cylinder $x^2 + y^2 = 1$, and between the planes $z = 0$ and $z = 4 - y$ with density function $\delta(x, y, z) = 1$.

28. The upper half of the unit ball, bounded between $z = 0$ and $z = \sqrt{1 - x^2 - y^2}$, with density function $\delta(x, y, z) = 1$.

Exercise Group. In the following exercises, a solid is described along with its density function. Find the center of mass of the solid using cylindrical coordinates. (Note: these are the same solids and density functions as found in Exercises 15.7.25 through Exercise 15.7.28.)

29. Bounded by the cylinder $x^2 + y^2 = 4$ and the planes $z = 0$ and $z = 4$ with density function $\delta(x, y, z) = \sqrt{x^2 + y^2} + 1$.

30. Bounded by the cylinders $x^2 + y^2 = 4$ and $x^2 + y^2 = 9$, between the planes $z = 0$ and $z = 10$ with density function $\delta(x, y, z) = z$.

31. Bounded by $y \geq 0$, the cylinder $x^2 + y^2 = 1$, and between the planes $z = 0$ and $z = 4 - y$ with density function $\delta(x, y, z) = 1$.

32. The upper half of the unit ball, bounded between $z = 0$ and $z = \sqrt{1 - x^2 - y^2}$, with density function $\delta(x, y, z) = 1$.

Exercise Group. In the following exercises, a solid is described along with its density function. Find the mass of the solid using spherical coordinates.

33. The upper half of the unit ball, bounded between $z = 0$ and $z = \sqrt{1 - x^2 - y^2}$, with density function $\delta(x, y, z) = 1$.

34. The spherical shell bounded between $x^2 + y^2 + z^2 = 16$ and $x^2 + y^2 + z^2 = 25$ with density function $\delta(x, y, z) = \sqrt{x^2 + y^2 + z^2}$.

35. The conical region bounded below by $z = \sqrt{x^2 + y^2}$ and above by the sphere $x^2 + y^2 + z^2 = 1$ with density function $\delta(x, y, z) = z$.

36. The cone that lies above the cone $z = \sqrt{x^2 + y^2}$ and below the plane $z = 1$ with density function $\delta(x, y, z) = z$.

Exercise Group. In the following exercises, a solid is described along with its density function. Find the center of mass of the solid using spherical coordinates. (Note: these are the same solids and density functions as found in Exercises 15.7.33 through Exercise 15.7.36.)

37. The upper half of the unit ball, bounded between $z = 0$ and $z = \sqrt{1 - x^2 - y^2}$, with density function $\delta(x, y, z) = 1$.

38. The spherical shell bounded between $x^2 + y^2 + z^2 = 16$ and $x^2 + y^2 + z^2 = 25$ with density function $\delta(x, y, z) = \sqrt{x^2 + y^2 + z^2}$.

39. The conical region bounded above $z = \sqrt{x^2 + y^2}$ and below the sphere $x^2 + y^2 + z^2 = 1$ with density function $\delta(x, y, z) = z$.

40. The cone bounded above $z = \sqrt{x^2 + y^2}$ and below the plane $z = 1$ with density function $\delta(x, y, z) = z$.

Exercise Group. In the following exercises, a region is space is described. Set up the triple integrals that find the volume of this region using rectangular, cylindrical and spherical coordinates, then comment on which of the three appears easiest to evaluate.

41. The region enclosed by the unit sphere, $x^2 + y^2 + z^2 = 1$.

42. The region enclosed by the cylinder $x^2 + y^2 = 1$ and planes $z = 0$ and $z = 1$.

43. The region enclosed by the cone $z = \sqrt{x^2 + y^2}$ and plane $z = 1$.

44. The cube enclosed by the planes $x = 0$, $x = 1$, $y = 0$, $y = 1$, $z = 0$ and $z = 1$. (Hint: in spherical, use order of integration $d\rho\, d\varphi\, d\theta$.)

15.8 Change of Variables in Multiple Integrals

Recall Exercise 15.3.15 in Section 15.3: the function $f(x, y) = e^{-x^2-y^2}$ is impossible to integrate in rectangular coordinates (at least, by finding antidervatives in terms of elementary functions), but switching to polar coordinates results in an integral that can be evaluated using a simple substitution.

We have seen in Sections 15.3 and 15.7 that switching to a different coordinate system can be a powerful tool. Integrals that are intractable (or even impossible) in one coordinate system can become straightforward in another.

Changing from rectangular coordinates to polar, or cylindrical, or spherical coordinates, are special cases of a general process known as a **change of variables** or **transformation**. A change of variables should be considered in any situation where we are presented with an integral that is difficult to evaluate in rectangular coordinates.

Our goals in this section are as follows:

- Understand how a change of variables affects the area element dA in a double integral, or the volume element dV in a triple integral.
- Derive a general change of variables formula for multiple integrals that works for any suitable change of coordinates, including the ones we have already seen in Sections 15.3 and 15.7
- Develop some basic guiding principles for knowing when a change of variables should be considered, and how to define the corresponding transformation.

youtu.be/watch?v=G1YTunRt6g0

Figure 15.8.1 Introducing the change of variables formula

15.8.1 Review of substitution techniques

One of the situations that should be covered by our general change of variables formula is that of substitution for a definite integral in one variable, as encountered in Section 5.6, way back in Calculus I. Of course, for a definite integral in one variable, there is only one type of region of integration: a closed interval $[a, b]$. For single integrals, our only consideration when making a change of variables is the function being integrated. Recall that substitution — at least, for indefinite integrals — is essentially an attempt to reverse the Chain Rule: given

$$\int_a^b f(T(u))T'(u)\,du,$$

we set $x = T(u)$, compute the differential $dx = T'(u)\,du$, and set

$$\int_a^b f(T(u))T'(u)\,du = \int_{T(a)}^{T(b)} f(x)\,dx. \tag{15.8.1}$$

We have reversed the roles of x and u as they typically appear in Calculus I. This may seem strange

For double and triple integrals, it will be important to understand how both the function *and* the region of integration are transformed. For single integrals, the Fundamental Theorem of Calculus (Theorem 5.4.10) lets us gloss over some of the details we'll now need to consider. In particular, we don't really need to pay any attention to the interval over which we integrate in a single integral, as long as we can come up with an antiderivative.

The formula we seek will be a generalization of this result, with one notable change in perspective: for multiple integrals, it is often the *region of integration* that creates most of the difficulty, and not the function being integrated. (In one variable, one closed interval is transformed into another, and we apply the Fundamental Theorem of Calculus. What we will find is that in most cases, we start on the *right* hand side of our analogue of Equation (15.8.1), and move to the left.

Changing from polar coordinates can be viewed as the process of writing our rectangular coordinates (x, y) in terms of new variables r and θ:

$$x(r, \theta) = r\cos\theta$$
$$y(r, \theta) = r\sin\theta.$$

Or conversely, as defining new variables r and θ as functions of the old variables x and y:

$$r(x, y) = \sqrt{x^2 + y^2}$$

$$\theta(x, y) = \arctan(y/x).$$

We can think of the polar coordinate transformation as a change of variables, where we define new variables in terms of old ones, but we could also think of it as a *function* from the plane $\mathbb{R}^2$ to itself. That is, we have a mapping

$$\begin{aligned} T : D \subseteq \mathbb{R}^2 &\to \mathbb{R}^2 \\ (x, y) = T(r, \theta) &= (r\cos\theta, r\sin\theta), \end{aligned}$$

where D is some subset of $\mathbb{R}^2$ (with coordinates labelled by r and θ), and the codomain is $\mathbb{R}^2$ with usual (x, y) coordinates. As we know from Section 15.3, the polar coordinate transformation T given above transforms a rectangle such as $D = [0, 3] \times [0, 2\pi]$ into a disk — in this case, the set of points (x, y) with $x^2 + y^2 \leq 9$, as shown in Figure 15.8.2 below.

Polar coordinates are a very common choice of coordinate system, because they are well adapted to situations with circular symmetry, a common assumption in many physical problems. They are also natural from a navigational perspective. However, the polar coordinate transformation is a bit unusual, in that it violates some of the principles we will require below for a general transformation. For example, the transformation is not one-to-one, and since we often allow $r < 0$ (recall Section 10.4), the inverse transformation (from polar to rectangular) technically isn't even a function! However, we're willing to overlook these defects due to the ubiquity and usefulness of the polar coordinate system.

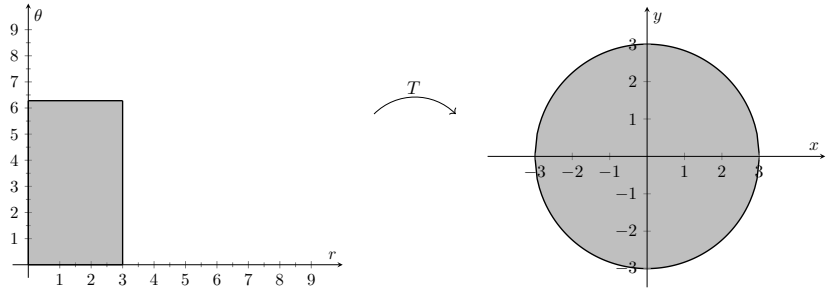


Figure 15.8.2 Transforming a rectangle to a disk using polar coordinates

It is interesting to pause and consider what happens to the four sides of the rectangle D in the transformation above. (As we'll see, this particular transformation exhibits some behaviour we usually prefer to avoid!). First, the side with $r = 0$ is collapsed to a single point: the origin. The side with $r = 3$ forms the entire perimeter of the circle. What happens to the sides $\theta = 0$ and $\theta = 2\pi$? They both get sent to the line segment from $(0, 0)$ to $(3, 0)$!

These observations let us imagine transformation as a physical process: first, the left side of the rectangle is shrunk down to a single point, while the right side is simultaneously stretched by a factor of 3. (Vertical lines in between are stretched/shrunk by a factor of r, with $0 \leq r \leq 3$.) The top of the rectangle is then bent around until it joins with the bottom.

It is perhaps easier to picture the transformation for a domain of the form $[a, b] \times [\alpha, \beta]$, with $0 < a < b$ and $0 \leq \alpha < \beta < 2\pi$. The case $r \in [1, 2]$, $\theta \in [\pi/6, \pi/3]$ is pictured in Figure 15.8.3 below.

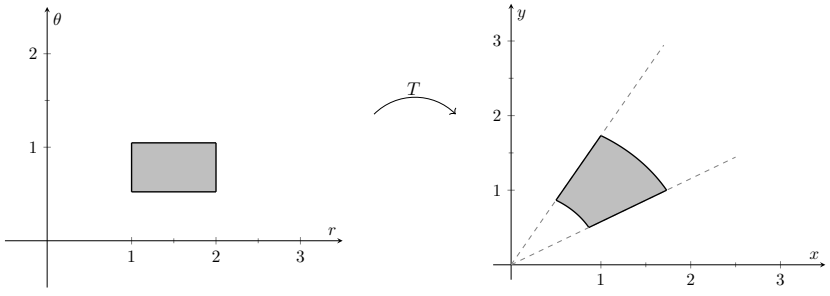


Figure 15.8.3 Transforming a rectangle to an annular portion using polar coordinates

15.8.2 General transformations

Given a region R in the plane and an integral $\iint_R f(x,y)\,dA$, we will look for a domain $D \subseteq \mathbb{R}^2$ and a function $T : D \to \mathbb{R}^2$ of the form

$$(x,y) = T(u,v) = (h(u,v), k(u,v))$$

that maps D onto R, which can be used to simplify our integral.

Definition 15.8.12 below specifies the properties we require for a function $T : D \subseteq \mathbb{R}^n \to \mathbb{R}^n$ to be used to define a change of variables. To explain some of those properties, we will need the following definitions.

The notation $f : A \to B$ is commonly used in mathematics when we want to indicate that a function f has a particular domain A and codomain B. We usually do not see this notation in calculus, since the domain is always a subset of $\mathbb{R}^n$ for some n, and understood to be the largest set of values for which the function is defined.

Note however that the choice of domain is part of the definition of a function, and it can significantly affect important properties of that function, such as being one-to-one: recall the definition of the inverse trigonometric functions in Section 2.7.

Definition 15.8.4 The image of a point or set.

Let $D \subseteq \mathbb{R}^n$ be any subset, and let $F : D \to \mathbb{R}^m$ be a function. For any point $\mathbf{x} \in D$, the **image** of $\mathbf{x}$ under F is the point $\mathbf{y} = F(\mathbf{x})$ in the range of F.

For any subset $C \subseteq D$, the **image** of C under F is denoted $F(C)$ and defined by

$$F(C) = \{F(\mathbf{x}) \,|\, \mathbf{x} \in C\}.$$

In other words, $\mathbf{y} \in F(C)$ if and only if $\mathbf{y}$ is the image of $\mathbf{x}$ for some $\mathbf{x} \in C$.

In particular, we denote the range (or **image**) of F by $F(D)$.

Definition 15.8.5 One-to-one and onto functions.

Let $A \subseteq \mathbb{R}^n$, let $B \subseteq \mathbb{R}^m$, and let $T : A \to B$ be a function.

- We say that T is **one-to-one** if no two points in A have the same image. That is, for any $\mathbf{x}_1, \mathbf{x}_2 \in A$, if $\mathbf{x}_1 \neq \mathbf{x}_2$, then $T(\mathbf{x}_1) \neq T(\mathbf{x}_2)$.
- We say that T is **onto** if the range of T is B; that is, if $T(A) = B$.

To avoid clutter throughout this section, we will use boldface variables as shorthand for points in $\mathbb{R}^n$. For example, we will write $\mathbf{x}$ instead of $(x_1, x_2, \ldots, x_n)$.

A function used for a change of variables is called a *transformation*. Such functions need to be one-to-one, except possibly on the boundary of their domain, and they need to be continuously differentiable. (See Definition 15.8.12 below.) One of the important properties of a transformation, which we will justify later in this section (see Theorem 15.8.27), is that the boundary of a closed, bounded domain is mapped to the boundary of the range. This observation is key to visualizing the effect of a transformation.

Example 15.8.6

Let $D \subseteq \mathbb{R}^2$ be the rectangle defined by $1 \leq u \leq 2$ and $0 \leq v \leq 1$. Determine the range of the function $T : D \to \mathbb{R}^2$ defined by

$$(x,y) = T(u,v) = (uv, u^2 + 2v^2).$$

Solution. The function T is continuously differentiable, since $x = 4uv$ and $y = u^2 + v^2$ both have continuous first-order partial derivatives with respect to u and v. Showing that T is one-to-one is a mess of algebra that we omit here. From these properties, we can conclude that the boundary of D will be transformed to the boundary of $T(D)$.

Now, let's see what happens to the boundary of D. The boundary consists of four line segments:

1. The segment $u = 1, 0 \leq v \leq 1$.

While it is important to know that our function is one-to-one, we will usually not ask you to check this fact. However, you should make sure you're aware of what can go wrong if this property is not satisfied: see the discussion following Definition 15.8.12.

2. The segment $u = 2, 0 \le v \le 1$.
3. The segment $1 \le u \le 2, v = 0$.
4. The segment $1 \le u \le 2, v = 1$.

On the first segment, $x = v$ and $y = 1 + v^2$, with $0 \le v \le 1$. Eliminating the parameter v gives us portion of the parabola $y = 1 + x^2$ from $(0, 1)$ to $(1, 2)$.

For the second segment we have $x = 2v$ and $y = 4 + v^2$. This is the part of the parabola $y = 4 + \frac{1}{4}x^2$ from $(0, 4)$ to $(2, 5)$.

The third segment has $x = 0$ and $y = u^2$, for $1 \le u \le 2$. This is the portion of the y axis from $(0, 1)$ to $(0, 4)$.

Finally, the fourth segment is given by $x = u$, $y = u^2 + 1$, for $1 \le u \le 2$. This is again the parabola $y = 1 + x^2$, but this time $1 \le x \le 2$.

The resulting region is plotted in Figure 15.8.7. Interestingly, two of the four sides of the rectangle bounding D were mapped to (different portions of) the same curve.

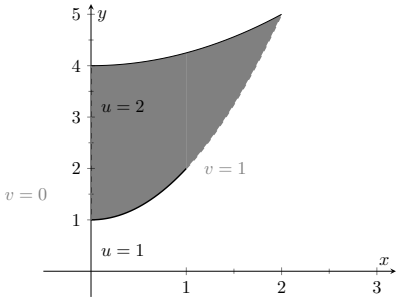


Figure 15.8.7 Plotting the image of the function T in Example 15.8.6

This example is interesting, in that two of the four sides of our rectangular domain were mapped to the same curve. Note also that (without explicitly solving for the inverse function, giving u and v as functions of x and y) we can see that lines of constant x in the u, v plane are circles, and lines of constant y are hyperbolas.

One other observation is worthy of note: we mentioned above that we will be primarily concerned with finding transformations that can be used to simplify a double integral. Suppose we were given a double integral over $R = T(D)$, as pictured in Figure 15.8.7. We probably wouldn't even consider a change of variables in this case, unless one was needed for the function being integrated: the region can be described by the inequalities

$$1 + x^2 \le y \le 4 + \frac{1}{4}x^2, \text{ where } 0 \le x \le 2.$$

We already learned how to deal with such regions at the beginning of this Chapter, and in any case, it's unlikely that anyone looking at this region would come up with the transformation we just considered.

15.8.3 The Jacobian of a transformation

We're ready to move on, and describe the effect of a change of variables on an integral. We begin with an observation from single variable calculus. Consider the definition of a definite integral as a limit of Riemann sums. When we make a change of variables $x = T(u)$ in a single integral, a partition of $[a, b]$ given by $a = u_0 < u_1 < \cdots < u_n = b$ is transformed into a partition $x_0 = T(u_0), x_1 = T(u_1), \ldots, x_n = T(u_n)$. (As long as $T'(u) > 0$, we have $x_0 < x_1 < \cdots < x_n$.)

The transformation affects the size of the subintervals in the partition: from Section 4.4, we know that $\Delta x_i \approx T'(u_i)\Delta x_i$. Thus, the derivative tells us how the size of each subinterval changes under the transformation.

This gives us a way of thinking about the geometric effect of a substitution. In the integral $\displaystyle\int_{T(a)}^{T(b)} f(x)\,dx$, the subintervals in a partition (thought of as the width of the rectangles in a Riemann sum) are stretched/shrunk horizontally by a factor given by the derivative $T'(u)$ of the transformation function $g(u)$. In the integral $\displaystyle\int_a^b f(T(u))T'(u)\,du$, the derivative $T'(u)$ is part of the integrand,

and therefore our horizontal stretch/shrink becomes a vertical stretch/shrink. Of course, the area of a rectangle changes by the same amount regardless of whether the stretch/shrink is horizontal or vertical.

Similarly, when we do a change of variables in two or three variables, we need a measure of how the size of each subregion in a partition changes under change of variables. This measure is given by an object known as the *Jacobian*.

In the discussion given here, we will stick to two variables for simplicity of presentation. However, everything we do for double integrals works equally well for triple integrals in three variables.

Definition 15.8.8 The Jacobian of a transformation.

Let $D \subseteq \mathbb{R}^2$ be a subset of the plane, described with coordinates (u, v). Let $T : D \subseteq \mathbb{R}^2 \to \mathbb{R}^2$ be given by

$$T(u, v) = (f(u, v), g(u, v)),$$

where f and g are continuously differentiable on D. The **Jacobian** is the function $J_T : D \to \mathbb{R}$ defined by

$$J_T(u, v) = \det \begin{bmatrix} f_u(u, v) & f_v(u, v) \\ g_u(u, v) & g_v(u, v) \end{bmatrix}.$$

If we define $x = f(u, v)$ and $y = g(u, v)$, we can write the Jacobian as

$$J_T(u, v) = \det \begin{bmatrix} x_u & x_v \\ y_u & y_v \end{bmatrix} = \det \begin{bmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{bmatrix}.$$

In the case of a transformation $T : D \subseteq \mathbb{R}^3 \to \mathbb{R}^3$, with $(x, y, z) = T(u, v, w)$, the definition of the Jacobian is similar, except that we need to compute the determinant of a 3×3 matrix.

If you read Section 14.6 on the definition of the derivative as a matrix of partial derivatives, you probably recognize the matrix whose determinant gives us the Jacobian: it's the derivative matrix! One therefore could write

$$J_T(u, v) = \det DT(u, v)$$

for the Jacobian, and this formula is valid in any dimension.

In fact, we can even use this definition for single integrals: a 1×1 matrix is just a number, and the determinant does nothing to that number, and of course, the derivative of a function of one variable is the same as always.

Another common notation for the Jacobian is

$$J_T(u, v) = \frac{\partial(x, y)}{\partial(u, v)}.$$

This notation is intended to be reminiscent of the Leibniz form of the chain rule when T is used to define a change of variables: we have the mnemonic devices

$$dx = \frac{dx}{du} du$$
$$dx\, dy = \frac{\partial(x, y)}{\partial(u, v)}\, du\, dv$$
$$dx\, dy\, dz = \frac{\partial(x, y, z)}{\partial(u, v, w)}\, du\, dv\, dw$$

for a change of variables in single, double, and triple integrals, respectively.

Example 15.8.9

Compute the Jacobian of the transformation $(x, y) = T(u, v)$ given by

$$x = 7u - 3v \quad y = -4u + 2v.$$

Solution. We apply Definition 15.8.8 directly:

$$J_T(u, v) = \det \begin{bmatrix} x_u(u, v) & x_v(u, v) \\ y_u(u, v) & y_v(u, v) \end{bmatrix}$$
$$= \det \begin{bmatrix} 7 & -3 \\ -4 & 2 \end{bmatrix}$$
$$= 7(2) - (-3)(-4) = 2.$$

In this case, the Jacobian is a constant function. This is the case whenever x and y are linear functions of u and v., but not true in general. We'll look at the

case of linear transformations in more detail after a few more examples.

Example 15.8.10

Compute the Jacobian of the transformation given by

$$T(u,v) = (\sqrt[3]{x^2y},\, \sqrt[3]{xy^2}).$$

Solution. Again, this is a direct application of the definition, but we should be clever about how we compute our partial derivatives. Our transformation defines $x = \sqrt[3]{x^2y}$ and $y = \sqrt[3]{xy^2}$. If we blindly push forward with the partial derivatives as written, we get a mess. For example, if we get a little too excited, we might do something like:

$$x_u(u,v) = \frac{\partial}{\partial u}(\sqrt[3]{x^2y}) = \frac{1}{3}(x^2y)^{-2/3}(2xy),$$

with similar results for the other four partial derivatives.

Instead, let's first simplify using laws of exponents:

$$x(u,v) = \sqrt[3]{x^2y} = (x^2y)^{1/3} = x^{2/3}y^{1/3} \quad \text{and}$$
$$y(u,v) = \sqrt[3]{xy^2} = (xy^2)^{1/3} = x^{1/3}y^{2/3}.$$

Now we only need the power rule to compute our partial derivatives, and we find

$$\begin{aligned} J_T(u,v) &= \det\begin{bmatrix} x_u(u,v) & x_v(u,v) \\ y_u(u,v) & y_v(u,v) \end{bmatrix} \\ &= \det\begin{bmatrix} \frac{2}{3}x^{-1/3}y^{1/3} & \frac{1}{3}x^{2/3}y^{-2/3} \\ \frac{1}{3}x^{-2/3}y^{2/3} & \frac{2}{3}x^{1/3}y^{-1/3} \end{bmatrix} \\ &= \left(\frac{2}{3}x^{-1/3}y^{1/3}\right)\left(\frac{2}{3}x^{1/3}y^{-1/3}\right) - \left(\frac{1}{3}x^{2/3}y^{-2/3}\right)\left(\frac{1}{3}x^{-2/3}y^{2/3}\right) \\ &= \frac{4}{9} - \frac{1}{9} = \frac{1}{3}. \end{aligned}$$

Interestingly enough, the Jacobian turns out to be constant again, even though the transformation was far from being linear. Let's try one more.

Example 15.8.11

Compute the Jacobian of the transformation

$$x = 4u^2 - v^2 \quad y = 2u^2 + 3v^2.$$

Solution. Computing the Jacobian in this case is straightforward:

$$J_T(u,v) = \det\begin{bmatrix} x_u(u,v) & x_v(u,v) \\ y_u(u,v) & y_v(u,v) \end{bmatrix} = \det\begin{bmatrix} 8u & -2v \\ 4u & 6v \end{bmatrix} 48uv + 8uv = 56uv.$$

As hinted at earlier, the Jacobian is important because it appears in the change of variables formula to come. Its role is analogous to that of the derivative $g'(u)$ in Equation (15.8.1). We also need the Jacobian to precisely define the type of function that can be used for a change of variables.

For all three coordinate systems we've studied, the transformation conditions can fail on the boundary of our domain, which we're willing to accept, and we account forthis by only requring these conditions to hold on the interior of the domain. What would *not* be acceptable is a situation where the transformation conditions fail on a region interior to the domain.

For example, the function $T(r,\theta) = (r\cos\theta, r\sin\theta)$ would not produce an acceptable transformation for domains such as $r \in [-1,1]$, $\theta \in [0,2\pi]$ or $r \in [0,2]$, $\theta \in [0,3\pi]$. The first traces out the unit disk twice: once for $r<0$, and once for $r>0$. The range of the second transformation is the disk $x^2+y^2 \leq 4$, but the upper half of this disk is produced twice: once for $\theta \in [0,\pi]$, and again for $\theta \in [2\pi,3\pi]$.

If we were using these transformations to compute a double integral over a circular region, we'd get the wrong answer!

Definition 15.8.12 Properties of a transformation.

Let D and E subsets of $\mathbb{R}^2$, with $D \subseteq \mathbb{R}^2$ described in terms of coordinates u,v, and $E \subseteq \mathbb{R}^2$ described in terms of coordinates x,y. Let D° denote the *interior* of D; that is, the set of all non-boundary points of D. We say that a function $T: D \to E$ is a **transformation** if:

1. T is continuously differentiable on D°.
2. T is one-to-one on D°, and the range of T is E.
3. The Jacobian of T does not vanish: $J_T(u,v) \neq 0$ for all $(u.v) \in D^\circ$.

When D is a closed, bounded subset, note that we do not require Definition 15.8.12 to hold on the boundary. Each of the three conditions above must hold on the interior of D, but are allowed to fail on all or part of the boundary. In particular, this is the case for cylindrical, and spherical coordinates:

- The polar coordinate transformation $x = r\cos\theta, y = r\sin\theta$ is only one-to-one if $r>0$ and θ belongs to an interval whose length is less than 2π. Note that $J_T(r,\theta)$ vanishes at $r=0$.

  Of course, we often use a domain such as $r \in [0,R]$, $\theta = [0,2\pi]$ to describe a disk centred at the origin. The conditions of Definition 15.8.12 fail at $r=0$, and because points with $\theta = 0$ get mapped to the same place as points with $\theta = 2\pi$. But these coordinates describe 3 of the 4 sides of the boundary rectangle for our domain, and the conditions are not required to hold on the boundadry.

- The cylindrical coordinate transformation has exactly the same issues as polar coordinates.

- For spherical coordinates, we take $\rho \geq 0$ and again accept the fact that our transformation is not one-to-one (and the Jacobian is zero) when $\rho = 0$. Similarly, we generally allow $\theta \in [0,2\pi]$ and $\varphi \in [-\pi/2, \pi/2]$ even though endpoints of these intervals might get sent to the same point.

Before we move on to the change of variables formula, we consider one more example that will help clarify the geometry involved in a change of variables, and that may be familiar to you from a first course in linear algebra.

We saw in Example 15.8.9 that when x and y are linear functions of u and v, the Jacobian of the transformation is a constant. What does that constant tell us about the transformation? Here is an example taken from the book Matrix Algebra, by Greg Hartman (who is also the main author of this text).

Example 15.8.13

Consider the function $T: \mathbb{R}^2 \to \mathbb{R}^2$ given by

$$T(u,v) = (u+4v, 2u+3v).$$

Note that T is linear in both variables. In fact, if we set $(x,y) = T(u,v)$ and represent points by vectors, replacing (x,y) by $\vec{x} = \begin{bmatrix} x \\ y \end{bmatrix}$ and (u,v) by $\vec{u} = \begin{bmatrix} u \\ v \end{bmatrix}$, then we can write this function as the matrix transformation

$\vec{x} = A\vec{u}$, where A is the 2×2 matrix $\begin{bmatrix} 1 & 4 \\ 2 & 3 \end{bmatrix}$. That is:

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 1 & 4 \\ 2 & 3 \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix}.$$

To visualize the effect of T, plot the vectors representing the four corners of the unit square, before and after they have been multiplied by A, where

$$A = \begin{bmatrix} 1 & 4 \\ 2 & 3 \end{bmatrix}.$$

Solution. The four corners of the unit square can be represented by the vectors

$$\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \quad \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad \begin{bmatrix} 0 \\ 1 \end{bmatrix}.$$

Multiplying each by A gives the vectors

$$\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \quad \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \quad \begin{bmatrix} 5 \\ 5 \end{bmatrix}, \quad \begin{bmatrix} 4 \\ 3 \end{bmatrix},$$

respectively.

The unit square and its transformation are graphed in Figure 15.8.14, where the shaped vertices correspond to each other across the two graphs. Note how the square got turned into some sort of quadrilateral (it's actually a parallelogram). A really interesting thing is how the triangular and square vertices seem to have changed places — it is as though the square, in addition to being stretched out of shape, was flipped.

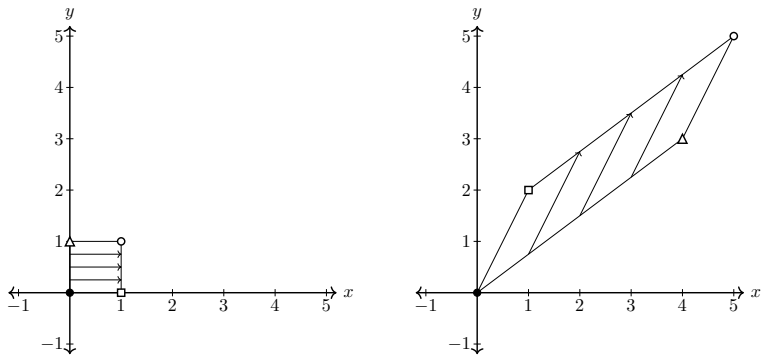


Figure 15.8.14 Transforming the unit square by matrix multiplication in Example 15.8.13

How does all this relate to Jacobians and change of variables? First note that the derivative of any linear function is (perhaps not so surprisingly) the matrix that defines it: for $T(u, v) = (u + 4v, 2u + 3v)$, we have

$$DT(u, v) = \begin{bmatrix} 1 & 4 \\ 2 & 3 \end{bmatrix} = A.$$

The Jacobian of T is then the determinant of this matrix:

$$J_T(u, v) = \det A = 1(3) - 4(2) = -5.$$

Let us make a note of a few key points about Example 15.8.13. First, note that in this case, the derivative matrix, (and as a result, the Jacobian) is constant. (This of course is generally true of the derivative for linear functions.)

What happens when we apply the map T to the unit square? The value $J_T(u, v) = -5$ tells us two things:

- First, the area of the unit square is increased by a factor of 5.
- Second, the transformation T reverses the **orientation** of the unit square. This is indicated by the negative value of the determinant. The reversal of orientation is responsible for the "flipping" of the square noticed in the solution above.

Recall, from linear algebra, that if two vectors $\vec{a}, \vec{b}$ span a parallelogram in the plane, then the determinant of the 2×2 matrix containing $\vec{a}$ and $\vec{b}$ gives the area

The result of performing the transformation T on the unit square is therefore the following: first, the square is flipped over. Then, the square is stretched out into a parallelogram whose area is 5 times that of the original square.

Let us make a couple more remarks about the Example 15.8.13. First, note the need for an absolute value around the determinant, to ensure the area computed is positive. This absolute value will be needed in our change of variables formula as well.

Second, since our transformation was linear, with constant derivative, the effect on area is the same for any portion of the plane: applying the transformation T to a closed bounded region $D \subseteq \mathbb{R}^2$ of area A will produce a region of area $5A$. For non-linear transformations, the value of the Jacobian (and hence, the effect on area) will vary from point to point.

Before we move on, let's do two more examples, with transformations we've already encountered. In these examples, we'll find that the value of the Jacobian is not a constant.

Recall the following property for definite integrals in one variable: $\int_a^b f(x)\,dx = -\int_b^a f(x)\,dx$. The definite integral is sensitive to the *orientation* of the interval over which the integration is performed. (Left to right or right to left.) Double and triple integrals do not have this sensitivity. We'll see in Section 16.5 how information about orientation is reintroduced in the context of vector calculus.

Example 15.8.15

Compute the Jacobian for

1. The polar coordinate transformation

$$x = r\cos\theta \quad y = r\sin\theta$$

2. The spherical coordinate transformation

$$x = r\cos\theta\cos\varphi \quad y = r\sin\theta\cos\varphi \quad z = r\sin\varphi.$$

Solution.

1. Here we've defined x and y in terms of the coordinates r and θ instead of u and v, but the process is the same:

$$\begin{aligned} J_T(r,\theta) &= \det\begin{bmatrix} x_r(r,\theta) & x_\theta(r,\theta) \\ y_r(r,\theta) & y_\theta(r,\theta) \end{bmatrix} = \det\begin{bmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{bmatrix} \\ &= r\cos^2\theta + r\sin^2\theta = r. \end{aligned}$$

Interesting. Note that the value of the Jacobian is r, which is precisely the correction factor needed in the area element for a double integral when we change from rectangular to polar coordinates. Let's try the spherical coordinate transformation to see if this was merely a coincidence.

2. Although we haven't defined the Jacobian for a change of coordinates in three variables, the process is exactly the same. We form the derivative of the transformation, given by the matrix of partial derivatives, and compute its determinant. We find:

$$\begin{aligned}
J_T(r,\theta,\varphi) &= \det\begin{bmatrix} x_\rho(\rho,\theta,\varphi) & x_\theta(\rho,\theta,\varphi) & x_\varphi(\rho,\theta,\varphi) \\ y_\rho(\rho,\theta,\varphi) & y_\theta(\rho,\theta,\varphi) & y_\varphi(\rho,\theta,\varphi) \\ z_\rho(\rho,\theta,\varphi) & z_\theta(\rho,\theta,\varphi) & z_\varphi(\rho,\theta,\varphi) \end{bmatrix} \\
&= \det\begin{bmatrix} \cos\theta\cos\varphi & -\rho\sin\theta\cos\varphi & -\rho\cos\theta\sin\varphi \\ \sin\theta\cos\varphi & \rho\cos\theta\cos\varphi & -\rho\sin\theta\sin\varphi \\ \sin\varphi & 0 & \rho\cos\varphi \end{bmatrix} \\
&= \rho\sin\theta\cos\varphi \begin{vmatrix} \sin\theta\cos\varphi & -\rho\sin\theta\sin\varphi \\ \sin\varphi & \rho\cos\varphi \end{vmatrix} \\
&\quad + \rho\cos\theta\cos\varphi \begin{vmatrix} \cos\theta\cos\varphi & -\rho\cos\theta\sin\varphi \\ \sin\varphi & \rho\cos\varphi \end{vmatrix} \\
&= \rho\sin\theta\cos\varphi(\rho\sin\theta(\cos^2\varphi+\sin^2\varphi) \\
&\quad + \rho\cos\theta\cos\varphi(\rho\cos\theta(\cos^2\varphi+\sin^2\varphi) \\
&= \rho^2\sin^2\theta\cos\varphi + \rho^2\cos^2\theta\cos\varphi \\
&= \rho^2\cos\theta.
\end{aligned}$$

We computed the above 3×3 determinant using a cofactor expansion along the second column. This is once again exactly the correction factor for the volume element in spherical coordinates, as given in Theorem 15.7.20 in Section 15.7.

youtu.be/watch?v=r1FX1DkDRb0

Figure 15.8.16 Computing the spherical coordinate Jacobian

15.8.4 The Change of Variables Formula

It seems that we're onto something. It is time that we stated the general change of variables formula for multiple integrals. Notice how, as with the derivative $g'(u)$ in Equation (15.8.1), the Jacobian gives us a measure of how subregions in the domain are stretched or shrunk. It shouldn't be too surprising, then, that the Jacobian plays the same role in multiple integrals that the derivative does in a single integral.

Theorem 15.8.17 Change of variables formula for double integrals.

Let D be a closed, bounded region in the plane, and let $T : D \subseteq \mathbb{R}^2 \to \mathbb{R}^2$ be a transformation. If f is a continuous, real-valued function on D, then

$$\iint_D f(T(u,v))|J_T(u,v)|\,du\,dv = \iint_{T(D)} f(x,y)\,dx\,dy.$$

The formula for triple integrals is analogous: given $(x,y,z) = T(u,v,w)$ for (u,v,w) in some closed, bounded domain D, then

$$\iiint_D f(T(u,v,w))\,|J_T(u,v,w)|\;du\,dv\,dw = \iiint_{T(D)} f(x,y,z)\,dx\,dy\,dz.$$

Let us try a simple example.

Example 15.8.18

Let R be the region in the x, y plane whose boundary is the parallelogram with vertices $(0,0)$, $(3,1)$, $(1,4)$, and $(4,5)$.

1. Determine a rectangular region D and a transformation $T : D \to \mathbb{R}^2$ such that $R = T(D)$.
2. Use the transformation T and Theorem 15.8.17 to determine the area of R.

Solution.

1. For inspiration, we look to Example 15.8.13. Notice how the transformation defined by the matrix $A = \begin{bmatrix} 1 & 4 \\ 2 & 3 \end{bmatrix}$ preserves the origin, and sends the points $(1,0)$ and $(0,1)$ to $(1,2)$ and $(4,3)$, respectively. In general, the transformation

   $$T(u,v) = (au + cv, bu + dv), \quad \text{with matrix} \quad \begin{bmatrix} a & c \\ b & d \end{bmatrix}$$

   will send $(1,0)$ to (a,b), and $(0,1)$ to (c,d).

   This suggests that in our case we can take D to be the unit square $[0,1] \times [0,1]$, and set

   $$T(u,v) = (3u + v, u + 4v).$$

   We check that $T(0,0) = (0,0)$, $T(1,0) = (3,1)$, $T(0,1) = (1,4)$, and $T(1,1) = (4,5)$. The four corners of the unit square are mapped to the four corners of the parallelogram. Since linear transformations map "lines to lines", we have our transformation.

2. To use Theorem 15.8.17, we need to compute the Jacobian of our transformation. We have

   $$J_T(u,v) = \det \begin{bmatrix} 3 & 1 \\ 1 & 4 \end{bmatrix} = 11,$$

   and since $R = T(D)$, the change of variables formula gives us

   $$A = \iint_R 1 \, dx \, dy = \iint_D 11 \, du \, dv = 11.$$

If you need further convincing, notice that setting $u = 0$, $v = t$ gives the parametric curve $T(0,t) = (t, 4t)$, which is the same as the line $y = 4x$: the line from $(0,0)$ to $(1,4)$. Similarly, setting $u = t, v = 0$ gives $T(t,0) = (3t,t)$: the line $y = \frac{1}{3}x$ from $(0,0)$ to $(3,1)$. One can similarly check that $T(1,t)$ and $T(t,1)$ give lines forming the other two sides of the parallelogram.

youtu.be/watch?v=Yf-0g8afcb4

Figure 15.8.19 Using a transformation to compute an integral over a parallelogram

Let's try another example. Our next example is more complicated, but this time, we're given the change of variables.

Example 15.8.20

Use the change of variables $x = u+v$, $y = u-v$ to evaluate the integral

$$\iint_R x e^{x^2-y^2} \, dA,$$

where R is the region bounded by the lines:

$$y = x, \; y = -x, \; y = x - 2, \text{ and } y = 2 - x.$$

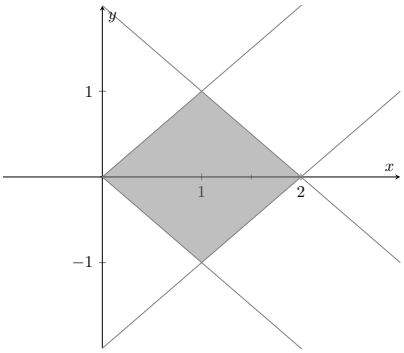


Figure 15.8.21 The region of integration R in Example 15.8.20

Solution. The region of integration is shown in Figure 15.8.21.

We need to determine a domain for the transformation $T(u,v) = (u+v, u-v)$ such that the range of T is R. Let's put $x = u+v$ and $y = u-v$ into the equations of our boundary lines, to see what the corresponding lines in the u,v plane are.

$$\begin{aligned} y = x \quad &\Rightarrow u - v = u + v \quad &&\Rightarrow v = 0 \\ y = -x \quad &\Rightarrow u - v = -(u+v) \quad &&\Rightarrow u = 0 \\ y = x - 2 \quad &\Rightarrow u - v = u + v - 2 \quad &&\Rightarrow 2v = 2 \quad \Rightarrow \quad v = 1. \end{aligned}$$

These lines are simply the boundary of the unit square in the u,v plane. Thus, if we take the domain $D = [0,1] \times [0,1]$ for T, we will have $R = T(D)$, as required.

Now, we apply Theorem 15.8.17. Recall the formula:

$$\iint_D f(T(u,v))|J_T(u,v)|\,du\,dv = \iint_{T(D)} f(x,y)\,dx\,dy.$$

We have $f(x,y) = xe^{x^2-y^2}$. It follows that

$$f(T(u,v)) = (u+v)e^{(u+v)^2-(u-v)^2} = (u+v)e^{u^2+2uv+v^2-(u^2-2uv+v^2)} = (u+v)e^{4uv}.$$

We also need to compute the Jacobian. Since the transformation is linear, we know this will be a constant. We find:

$$J_T(u,v) = \det \begin{bmatrix} x_u & x_v \\ y_u & y_v \end{bmatrix} = \det \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} = -2.$$

In this case, understanding the geometry of the Jacobian gives us the answer without any computation. Since a square of area 1 is transformed to a square of area 2, we know $|J_T(u,v)| = 2$.

Putting all this into our change of variables formula, we have

$$\iint_R xe^{x^2+y^2}\,dx\,dy = \iint_D (u+v)e^{4uv}|-2|\,du\,dv.$$

This integral can be evaluated by splitting it in two, and choosing the most convenient order of integration for each part:

$$\iint_D 2(u+v)e^{4uv}\,du\,dv = 2\int_0^1\int_0^1 ue^{4uv}\,dv\,du + 2\int_0^1\int_0^1 ve^{4uv}\,du\,dv.$$

Now, we find that

$$\begin{aligned} \int_0^1\int_0^1 ue^{4uv}\,dv\,du &= \int_0^1 \left(\frac{1}{4}e^{4uv}\Big|_0^1\right) du \\ &= \frac{1}{4}\int_0^1 (e^{4u}-1)\,du \\ &= \frac{1}{16}(e^4-1) - \frac{1}{4}(1) = \frac{1}{16}e^4 - \frac{5}{16}, \end{aligned}$$

and the second integral differs only in the labelling of the variables, and gives the same result. Thus, we have

$$\iint_R xe^{x^2-y^2}\,dx\,dy = \frac{1}{4}e^4 - \frac{5}{4}.$$

Let's try one more example where we're given some guidance before tackling

a general change of variables problem.

Example 15.8.22

Let R be the region in the first quadrant bounded by the lines $y = x$ and $y = 4x$, and the hyperbolas $y = 1/x$ and $y = 4/x$. Evaluate the integral

$$\iint_R xy^2\,dA$$

using the change of variables $x = u/v$, $y = v$.

Solution. First, we note that setting $y = kx$, where k is a constant, gives us

$$v = k\frac{u}{v} \quad \Rightarrow \quad u = \frac{1}{k}v^2,$$

while setting $y = k/x$ gives $xy = k$, or $u = k$. The region R is therefore the image under the transformation $T(u, v) = (u/v, v)$ of the region D bounded by the curves $u = v^2$ and $u = \frac{1}{4}v^2$, and the lines $u = 1, u = 4$; see Figure 15.8.24.

Some caution is needed when determining the domain D. Note that the given curves bound two regions: one above the u axis, and one below. But we note that $y = v$, and since $y > 0$ for the region R, we must have $v > 0$ in D.

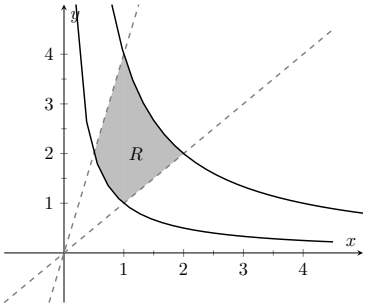


Figure 15.8.23 The region of integration R in Example 15.8.22

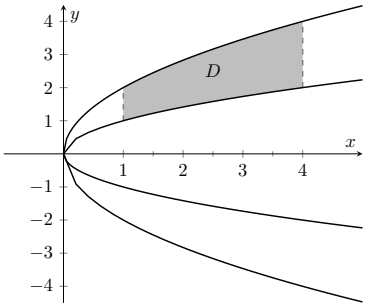


Figure 15.8.24 The domain D mapped onto R by T

This is perhaps not the best possible change of variables: the domain D is not a rectangle. (See Example 15.8.31 below for a change of variables that is more effective for this type of region.) However, it is a region of the type we considered in Section 15.2, so we're better off than we were with the original region. We have $1 \le u \le 4$, and the equations $u = v^2$, $u = \frac{1}{4}v^2$ can be re-written (noting that $v > 0$) as $v = \sqrt{u}$ and $v = 2\sqrt{u}$.

With $f(x, y) = xy^2$ we have $f(T(u, v)) = \frac{u}{v} \cdot v^2 = uv$, and the Jacobian is given by

$$J_T(u, v) = \det \begin{bmatrix} 1/v & -u/v^2 \\ 0 & 1 \end{bmatrix} = \frac{1}{v}.$$

Thus, we have

$$\begin{aligned} \iint_R xy^2\,dA &= \iint_D uv \left|\frac{1}{v}\right| du\,dv \\ &= \int_1^4 \int_{\sqrt{u}}^{2\sqrt{u}} u\,dv\,du \\ &= \int_1^4 u^{3/2}\,du. \end{aligned}$$

Our next goal is to tackle the following general problem: given a multiple integral over a region E, determine a transformation T with domain D such that $T(D) = E$, and use it to evaluate the integral. Before attempting a couple of examples, we take a brief detour to consider some technical details that will assist us in understanding the problem.

Recall from Definition 15.8.12 that we require transformations to be one-to-one and onto (see Definition 15.8.5), except possibly on the boundary of their domain.

One of the reasons that we require these properties is that they guarantee that T has an **inverse**. If a transformation $T : D \to E$ is one-to-one and onto, then we can define the inverse mapping $T^{-1} : E \to D$ according to

$$T^{-1}(\mathbf{x}) = \mathbf{u} \quad \text{if and only if} \quad \mathbf{x} = T(\mathbf{u}).$$

Notice that the onto condition guarantees that the domain of T^{-1} is all of E. When considering a changes of variables for a multiple integral over a region E, we would ideally like to have a one-to-one and onto mapping from D to E to ensure that when we convert to an integral over D, each point in E only gets "counted once".

For example, consider the mapping $T(u, v) = (u^2, v)$ defined on $[-1, 1] \times [0, 1]$. (That is, $x = u^2$ with $-1 \le u \le 1$ and $y = v$, with $0 \le v \le 1$.) The image of T is the square $[0, 1] \times [0, 1]$, but each point (x, y) corresponds to two points $(\pm\sqrt{x}, \sqrt{y})$ in D, so integrating over D would be the same as integrating over E {\em twice}!

Next we want to consider differentiability. Recall that a vector-valued function

$$\mathbf{r}(t) = \langle x(t), y(t) \rangle$$

is continuous if and only if each of the component functions $x(t), y(t)$ is continuous, and similarly, $\mathbf{r}(t)$ is differentiable if and only if each of the component functions is differentiable, and

$$\mathbf{r}'(t) = \langle x'(t), y'(t) \rangle.$$

Similarly, a function $T : D \subset \mathbb{R}^n \to \mathbb{R}^n$ is continuous if and only if each of its components is continuous (as a function of several variables), and (for $n = 2$) the partial derivatives of T can be viewed as the vector-valued functions

$$\mathbf{r}_u(u, v) = \frac{\partial T}{\partial u}(u, v) = \left\langle \frac{\partial x}{\partial u}(u, v), \frac{\partial y}{\partial u}(u, v) \right\rangle$$
$$\mathbf{r}_v(u, v) = \frac{\partial T}{\partial v}(u, v) = \left\langle \frac{\partial x}{\partial v}(u, v), \frac{\partial y}{\partial v}(u, v) \right\rangle,$$

with similar formulas for $n = 3$. (For $n = 1$ we have only the single derivative $T'(u)$.)

If each of the components of each of the partial derivatives is continuous (that is, if the partial derivative of each of the **x** variables with respect to each of the **u** variables is continuous) we say that T is C^1, or continuously differentiable.

If a function $T : D \subset \mathbb{R}^n \to E \subset \mathbb{R}^n$ is C^1, then as with real-valued functions, being continuously differentiable implies that T is differentiable (in the sense of the definition from Section 14.6), and therefore continuous. The derivative of T is then an $n \times n$ matrix. For example, when $n = 2$, if $T(u, v) =$

Recall that in Section 14.6 we gave the following alternative definition of differentiability: $f : D \subset \mathbb{R}^n \to \mathbb{R}^m$ is *differentiable* if the limit of

$$\frac{\|f(\mathbf{a} + \mathbf{h}) - f(\mathbf{a}) - Df(\mathbf{a})\mathbf{h}\|}{\|\mathbf{h}\|}$$

is 0 as $\mathbf{h} \to \mathbf{0}$, where $Df(\mathbf{a})$ is the matrix of partial derivatives of f at $\mathbf{a}$, and $Df(\mathbf{a})\mathbf{h}$ denotes matrix multiplication, with $\mathbf{h}$ viewed as a column vector.

$(x(u,v), y(u,v))$, we get

$$DT(u,v) = \begin{bmatrix} \dfrac{\partial x}{\partial u} & \dfrac{\partial x}{\partial v} \\ \dfrac{\partial y}{\partial u} & \dfrac{\partial y}{\partial v} \end{bmatrix}.$$

Notice that, while the gradients $\nabla x(u,v), \nabla y(u,v)$ make up the rows of the derivative matrix $DT(\mathbf{a})$, the *columns* of $DT(\mathbf{a})$ are the partial derivative vectors $\mathbf{r}_u$ and $\mathbf{r}_v$.

Given our function $T : D \subset \mathbb{R}^n \to E \subset \mathbb{R}^n$, let us denote by DT the matrix of partial derivatives, as in Section 14.6. Since the dimension of the domain and range are the same, DT is a square $(n \times n)$ matrix, so we can compute its determinant, and this, of course, is the Jacobian, as defined in Definition 15.8.8.

Let's come back to the change of variables formula. If we let $d\mathbf{x}$ denote either dx, dA, or dV, depending on whether $n = 1, 2$ or 3, and doing the same for $d\mathbf{u}$, the change of variables formula for a transformation $T : D \to E$ can be written as

$$\int_E f(\mathbf{x})d\mathbf{x} = \int_D f(T(\mathbf{u}))|J_T(\mathbf{u})|d\mathbf{u},$$

where the integral sign represents a single, double, or triple integral, depending on the value of n. (So this really is just a generalization of the method of substitution you learned in Calculus I.)

Note that the properties required for T to be a transformation tell us that every point of E corresponds to a point in D, and integrating over D is the same as integrating over E, once we account for the "stretch factor" of the transformation given by the Jacobian $J_T(\mathbf{u})$. A rigorous proof of the change of variables formula is very difficult, but we will give an argument at the end of this section similar to the one we considered for the polar and spherical coordinate transformations that, although not a complete proof, is at least a plausible explanation.

The general inverse function theorem, which is not stated in most calculus textbooks, (probably in part because the statement requires defining the matrix DT of partial derivatives and explaining what the inverse of a matrix is), states that if $T : D \to E$ is one-to-one and onto, then T^{-1} exists, and moreover, if T is C^1 *and* $J_T(\mathbf{u}) \neq 0$ for all $\mathbf{u} \in D$, then T^{-1} is *also* a C^1 function, and

$$DT^{-1}(\mathbf{x}) = (DT(\mathbf{u}))^{-1}, \tag{15.8.2}$$

where $\mathbf{u} = T^{-1}(\mathbf{x})$.

The -1 on the right-hand side of Equation (15.8.2) denotes a matrix inverse. A basic result from linear algebra tells us that a matrix is invertible if and only if its determinant is non-zero, which is one reason why we require a nonzero Jacobian in Definition 15.8.12. (Compare this to the result $(f^{-1})'(x) = \dfrac{1}{f'(f^{-1}(x))}$ in one variable.)

A useful consequence of Equation (15.8.2) is obtained by taking the determinant of both sides of the above equation (recall that $\det(A^{-1}) = 1/\det(A)$ for any invertible matrix A).

Theorem 15.8.25 The Jacobian of an inverse transformation.

Let $T : D \to \mathbb{R}^2$ be a one-to-one C^1 mapping with image $E = T(D)$. If $J_T(\mathbf{u}) \neq 0$ for all $\mathbf{u} \in D$, then $T^{-1} : E \to \mathbb{R}^2$ is a transformation, and the Jacobian of T^{-1} is given by

$$J_{T^{-1}}(\mathbf{x}) = \frac{1}{J_T(T^{-1}(\mathbf{x}))}.$$

This result can come in handy in cases where it's easy to come up with the inverse mapping $\mathbf{u} = T^{-1}(\mathbf{x})$, but hard to solve for $\mathbf{x}$ in terms of $\mathbf{u}$ to obtain T.

youtu.be/watch?v=-kHCIjBbcoA

Figure 15.8.26 Working with the inverse of a transformation

Our last technical detail is a theorem that can be very useful when trying to determine the transformation to use for a change of variables: the boundary of E must correspond to the boundary of D. This is useful because we usually would like D to be as simple as possible, ideally a rectangle (or box, if $n = 3$).

Since the sides of the rectangle are given by setting u or v equal to a constant, we look at the curves that define the boundary of E. If the boundary of E can be expressed in terms of level curves for two functions $f(x, y)$ and $g(x, y)$, we can define $u = f(x, y)$ and $v = g(x, y)$, which allows us to define $T^{-1}(x, y) = (f(x, y), g(x, y))$. From there, we can try to compute T from T^{-1}, which is a matter of solving for x and y in terms of u and v.

Theorem 15.8.27 Transformations preserve the boundary.

Let $D, E \subset \mathbb{R}^n$ be closed, bounded regions. If $T : D \to E$ is a transformation, then the boundary of E is the image under T of the boundary of D; that is, if $T(\mathbf{u}) = \mathbf{x}$ is on the boundary of E, then $\mathbf{u}$ is on the boundary of D.

We will prove this result in the case that T is one-to-one, with $J_T(\mathbf{u}) \neq 0$, on all of D, including the boundary. Note that if this property fails on some portion on the boundary, this will not affect the integral. For example, if $n = 2$, the boundary of D consists of a finite union of continuous curves, so any portion of the boundary is a continuous curve, and we know that we can neglect the graphs of finitely many continuous curves when carrying out an integral. We begin by first proving a simpler result.

A result such as Theorem 15.8.28 that is used as a step towards proving a more substantial result is often referred to as a *lemma*.

Theorem 15.8.28 Transformations are open mappings.

If $f : A \to B$ is a continuous, one-to-one, and onto mapping from A to B with continuous inverse $f^{-1} : B \to A$, then f maps open sets to open sets. That is, if $U \subset A$ is an open subset of A, then the image $f(U) = \{f(\mathbf{u}) \in B | \mathbf{u} \in U\}$ is an open subset of B.

Proof. Let $U \subset A$ be open, and let $\mathbf{x} \in f(U)$. We need to show that there exists some $\delta > 0$ such that $N_\delta(\mathbf{x}) = \{\mathbf{y} \in A | \, \|\mathbf{x} - \mathbf{y}\| < \delta\}$ is a subset of $f(U)$. (By definition, $f(U)$ is open if each element of $f(U)$ has a δ-neighbourhood completely contained in $f(U)$.) Since f is one-to-one and onto, there exists a unique $\mathbf{v} = f^{-1}(\mathbf{x}) \in U$ such that $f(\mathbf{v}) = \mathbf{x}$. (We must have $\mathbf{v} \in U$ since $f(\mathbf{v}) \in f(U)$.) Since U is open, there exists an $\epsilon > 0$ such that $N_\epsilon(\mathbf{v}) \subset U$.

Now, since f^{-1} is continuous, there exists a $\delta > 0$ such that if $\mathbf{y} \in N_\delta(\mathbf{x})$, then $f^{-1}(\mathbf{y}) \in N_\epsilon(\mathbf{v})$. But if $f^{-1}(\mathbf{y}) \in N_\epsilon \subset U$, then $f(f^{-1}(\mathbf{y})) = \mathbf{y} \in f(U)$, by definition of $f(U)$. Thus, $N_\delta(\mathbf{x}) \subset f(U)$, which is what we needed to show. ■

Using the above lemma, we can now give a proof of our theorem.

Proof of Theorem 15.8.27. Let $T : D \to E$ be the given transformation, which is one-to-one and onto, and such that $J_T(\mathbf{u}) \neq 0$ for all $\mathbf{u} \in D$. Since T is one-to-one and onto, we can find an inverse function $T^{-1} : E \to D$. Since T is C^1 and $J_T(\mathbf{u}) \neq 0$ for all $\mathbf{u} \in D$, the inverse function theorem tells us that T^{-1} must be C^1 on E. Since T and T^{-1} are both C^1, they are differentiable and therefore continuous.

Now, let $\mathbf{x} \in E$ be a boundary point. We need to show that $\mathbf{x}$ is the image of a boundary point in D. Recall that $\mathbf{x}$ is a boundary point if and only if every neighbourhood of $\mathbf{x}$ contains both points in E and points not in E. Let $\mathbf{u} =$

$T^{-1}(\mathbf{x}) \in D$ be the element of D that is mapped to **x** by T. For the sake of contradiction, suppose that **u** is not a boundary point of D. Then since $\mathbf{u} \in D$ it must be an interior point of D, and therefore, there exists some $\delta > 0$ such that $N_\delta(\mathbf{u}) \subset D$. (That is, there is a neighbourhood of **u** that is completely contained in D.)

However, since T satisfies the conditions of Theorem 15.8.28, we know that T must map open sets to open sets. In particular, since $N_\delta(\mathbf{u})$ is an open subset of D, $T(N_\delta(\mathbf{u}))$ must be an open subset of E. But since $\mathbf{u} \in N_\delta(\mathbf{u})$, we must have $\mathbf{x} = T(\mathbf{u}) \in T(N_\delta(\mathbf{u}))$, and thus $T(N_\delta(\mathbf{u}))$ is an open subset of E that contains **x**, which contradicts the fact that **x** is a boundary point. Thus, it must be the case that **u** is a boundary point of D. ■

Note that since $T^{-1} : E \to D$ is also a transformation with the same properties as T, the converse to this result is valid as well: if **u** belongs to the boundary of D, then $T(\mathbf{u})$ belongs to the boundary of E.

We will see how Theorems 15.8.27 and 15.8.25 are put to use in the following examples.

Example 15.8.29

Compute $\iint_E \left(\frac{y^2}{x^4} + \frac{x^2}{y^4}\right) dA$, where E is the region bounded by $y = x^2$, $y = 2x^2$, $x = y^2$, and $x = 4y^2$.

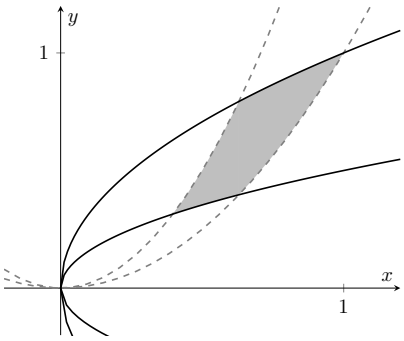


Figure 15.8.30 The region of integration for Example 15.8.29

Solution. The region E is pictured in Figure 15.8.30 below. We need to find a region $D \subset \mathbb{R}^2$ and a transformation $T : D \to \mathbb{R}^2$ whose image is E. We use the fact that T must map the boundary of D to the boundary of E as a guideline. In particular, note that since T is C^1, it must map smooth curves to smooth curves by the chain rule. This tells us that the corners of E must correspond to the corners of D, and in particular, that each of the four curves that make up the boundary of E must come from four curves that make up the boundary of D. Since we would like the integral over D to be as simple as possible, we try to find a transformation such that D is a rectangle.

Since the sides of a rectangle in the uv-plane are given by either $u =$ constant or $v =$ constant, we try to express the boundary of E in terms of level curves $u(x, y) = c_1, c_2$ and $v(x, y) = d_1, d_2$. Let's look at the curves $y = x^2$ and $y = 2x^2$. These both belong to the family of curves $y = cx^2$, or $\frac{y}{x^2} = c$, so we set $u(x, y) = \frac{y}{x^2}$. The region between these two parabolas is then given by $1 \leq u \leq 2$, or $u \in [1, 2]$. Similarly, the other two sides of the boundary of E, given by $x = y^2$ and $x = 4y^2$ both belong to the family of curves $x = dy^2$, or $\frac{x}{y^2} = d$. This suggests that we take $v(x, y) = \frac{x}{y^2}$, with $1 \leq v \leq 4$.

We have now determined a map $S : E \to D = [1, 2] \times [1, 4]$ given by

$$S(x, y) = \left(\frac{y}{x^2}, \frac{x}{y^2}\right).$$

This map is one-to-one and onto (check this), clearly C^1, and has Jacobian

$$J_S(x, y) = \frac{\partial}{\partial x}\left(\frac{y}{x^2}\right)\frac{\partial}{\partial y}\left(\frac{x}{y^2}\right) - \frac{\partial}{\partial x}\left(\frac{x}{y^2}\right)\frac{\partial}{\partial y}\left(\frac{y}{x^2}\right) = \frac{3}{x^2y^2},$$

which is defined and non-zero on all of E. This means that $S = T^{-1}$ for

some transformation $T : D \to E$. We can now proceed to compute the integral via change of variables in one of two ways:

1. Directly, by solving for x and y in terms of u and v, which will give us the transformation T.
   From $u = \dfrac{y}{x^2}$ we get $y = ux^2$, so $x = vy^2 = vu^2x^4$. Since $x \neq 0$ on E, this gives us $x^{-3} = u^2v$, so $x = u^{-2/3}v^{-1/3}$, and thus $y = ux^2 = u^{-1/3}v^{-2/3}$. The transformation T is thus $T(u,v) = (u^{-2/3}v^{-1/3}, u^{-1/3}v^{-2/3})$, and its Jacobian is given by

$$\begin{aligned} J_T(u,v) &= \frac{\partial}{\partial u}(u^{-2/3}v^{-1/3})\frac{\partial}{\partial v}(u^{-1/3}v^{-2/3}) - \frac{\partial}{\partial u}(u^{-1/3}v^{-2/3})\frac{\partial}{\partial v}(u^{-2/3}v^{-1/3}) \\ &= \frac{1}{3u^2v^2}. \end{aligned}$$

   The integral is therefore

$$\begin{aligned} \iint_E \left(\frac{x^2}{y^4} + \frac{y^2}{x^4}\right) dA &= \iint_D (v^2+u^2)\left|\frac{1}{3u^2v^2}\right| du\,dv \\ &= \frac{1}{3}\int_1^4\int_1^2 \left(\frac{1}{u^2}+\frac{1}{v^2}\right) du\,dv \\ &= \frac{1}{3}\int_1^4 \left(\frac{-1}{2} - \frac{-1}{1} + \frac{1}{v^2}\right) dv. \end{aligned}$$

2. Indirectly, using the fact that $J_T(u,v) = \dfrac{1}{J_{T^{-1}}(x(u,v),y(u,v))}$.
   From the above, we have that $J_{T^{-1}}(x,y) = \frac{3}{x^2y^2}$, so $J_T(u,v) = \frac{1}{3}(x(u,v))^2(y(u,v))^2$. From $u = \dfrac{y}{x^2}$ and $v = \dfrac{x}{y^2}$, we have $uv = \dfrac{xy}{x^2y^2} = \dfrac{1}{xy}$. Thus, $x^2y^2 = \dfrac{1}{u^2v^2}$, so $J_T(u,v) = \dfrac{1}{3u^2v^2}$ as before. From here we can proceed as above.

Example 15.8.31

Compute $\iint_E xy\,dA$, where E is the region in the first quadrant bounded by $y = x$, $y = 4x$, $y = 1/x$, and $y = 2/x$.

Solution. We need to find a region $D \subset \mathbb{R}^2$ and a transformation $T : D \to \mathbb{R}^2$ whose image is E. This problem is almost identical to the one we solved in Example 15.8.22, where we were given a change of variables whose domain was still somewhat complicated. This time, we look for a transformation with a rectangular domain.

Using the principle that T must map the boundary of D to the boundary of E as above, we set $u = \dfrac{y}{x}$, so that $1 \leq u \leq 4$ gives the region between $y = x$ and $y = 4x$, and $v = xy$, so that $1 \leq v \leq 2$ gives the region between $y = 1/x$ and $y = 2/x$. Thus the desired transformation is defined on the rectangle $D = [1,4] \times [1,2]$ and has an inverse given by $T^{-1}(x,y) = (y/x, xy)$.

This time we leave the direct method (solving for x and y in terms

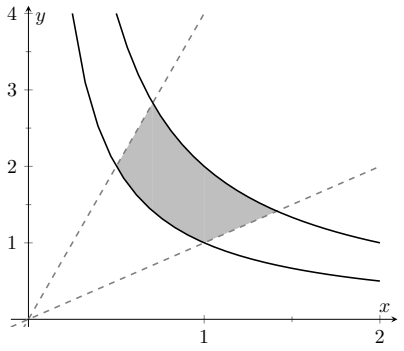


Figure 15.8.32 The region of integration in Example 15.8.31

of u and v) as an exercise and use the indirect method. The Jacobian of T^{-1} is given by

$$J_{T^{-1}}(x,y) = \det \begin{pmatrix} \dfrac{\partial}{\partial x}\left(\dfrac{y}{x}\right) & \dfrac{\partial}{\partial y}\left(\dfrac{y}{x}\right) \\ \dfrac{\partial}{\partial x}(xy) & \dfrac{\partial}{\partial y}(xy) \end{pmatrix} = \det \begin{pmatrix} \dfrac{-y}{x^2} & \dfrac{1}{x} \\ y & x \end{pmatrix} = \frac{-2y}{x}.$$

The Jacobian of T is thus

$$J_T(u,v) = \frac{1}{J_T^{-1}(x(u,v),y(u,v))} = -\frac{x(u,v)}{2y(u,v)} = -\frac{1}{2u},$$

since $u = y/x$. The integral is then

$$\begin{aligned} \iint_E xy\,dA &= \iint_D x(u,v)y(u,v)|J_T(u,v)|\,du\,dv \\ &= \int_1^2 \int_1^4 v\left(\frac{1}{2u}\right)\,du\,dv \\ &= \int_1^2 \frac{v}{2}(\ln 4 - \ln 1)\,dv. \end{aligned}$$

youtu.be/watch?v=hSDH3FnD97w

Figure 15.8.33 A video presentation of Example 15.8.31, with slightly different numbers

15.8.5 Understanding the change of variables formula

We now have some practice working with the change of variables formula, but why is it valid? In any dimension, the formula has the form

$$\int_{T(D)} f(\mathbf{x})d\mathbf{x} = \int_D f(T(\mathbf{u}))|J_T(\mathbf{u})|d\mathbf{u},$$

if we let the symbol $\int$ stand for a single, double, or triple integral as necessary.

In practice, we use the formula in one of two ways:

- Right-to-left, because it is easier to compute antiderivatives for the function $f(\mathbf{x})$. This is the case with change of variables for single integrals.
- Left-to-right, because the domain D is a simpler region of integration than $T(D)$, such as the examples above, as well as the transformations to polar, cylindrical, or spherical coordinates considered earlier. (Of course, we might also get lucky and find that our function simplifies as well!)

Let's consider this formula in the intermediate case of a double integral. If the function f is positive throughout the region $E = T(D)$, we can interpret the integral on the left as a volume. In terms of Riemann sums, we are adding up volumes of boxes:

$$\Delta V_{ij} = f(x_{ij}, y_{ij})\Delta x_i\,\Delta y_j.$$

Just as $f(x) \approx T'(u)\,\Delta u$ in one variable, the validity of the change of variables formula rests on the approximation

$$\Delta x_i\,\Delta y_j \approx |J_T(u_{ij}, v_{ij})|\Delta u_i\,\Delta v_j.$$

The distortion in area caused by the mapping T when we move from the region D in the u, v plane to the region E in the x, y plane is hidden within the $dx\,dy$ area element in the integral on the left-hand side.

To ensure that both integrals compute the same volume, the Jacobian is introduced as part of the integrand on the right-hand side to produce a corresponding change in height:

$$\Delta V_{ij} \approx \underbrace{(f(x_{ij}, y_{ij}))}_{\text{height}} \underbrace{(|J_T(u_{ij}, v_{ij})|\Delta u_i\, \Delta v_j)}_{\text{area}} = \underbrace{(f(T(u_{ij}, v_{ij}))|J_T(u_{ij}, v_{ij})|)}_{\text{height}} \underbrace{(\Delta u_i\, \Delta v_j)}_{\text{area}}.$$

Appropriately interpreted, the only differences between the integrals on either side are the labelling of the variables, and whether the Jacobian provides a measure of height, or of area, in the calculation of volume.

In general, transformations produce what are called "curvilinear coordinate systems": the original linear coordinate system in the u, v plane, with grid lines given by $u =$ constant or $v =$ constant is transformed into a "grid of curves" in the x, y plane. This is the case, for example, with the polar coordinate transformation, as seen in Figure 15.8.34 below.

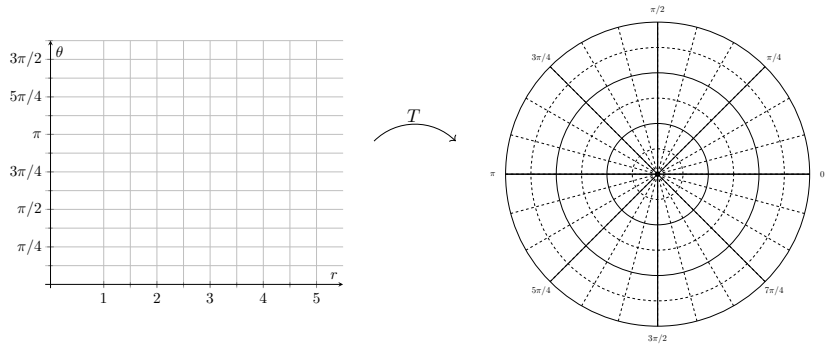


Figure 15.8.34 Correspondence between rectangular and polar grid lines

For another example, consider the transformation T given by

$$T(u, v) = (u^{1/3}v^{-1/3}, u^{2/3}v^{1/3}), T^{-1}(x, y) = \left(xy, \frac{y}{x^2}\right).$$

A grid in the u, v plane is transformed to two families of curves: lines $u = m$, $v = n$, where m, n are constants become the curves $y = \frac{m}{x}$ and $y = nx^2$, respectively. The transformation is pictured in Figure 15.8.35 below.

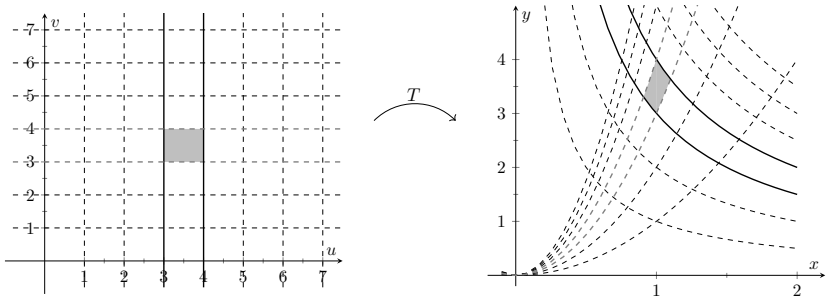


Figure 15.8.35 Visualizing a general transformation

In Figure 15.8.35 we've highlighted one of the rectangles in our grid to see how it's transformed. Imagine now that our grid lines are much finer, coming not from the integer values of u and v, but from a partition of a rectangle D in the u, v plane. Zooming in, we'd see that each rectangle in the partition is transformed much like the one above.

This also fits with a general philosophy of differential calculus: the derivative of a function at a point determines the best linear approximation to that function near that point. It seems only fitting, then, that the best linear approximation to a transformation is a linear transformation!

Indeed, recall the following philosophy from Section 14.6: the transformation T maps points in the u, v plane to points in the x, y plane. The derivative matrix $DT(u, v)$ of T at a point (u, v), when viewed as the matrix of a linear transformation, maps (tangent) vectors at the point (u, v) to (tangent) vectors at the point $(x, y) = T(u, v)$. (This is a consequence of the Chain Rule.)

Consider a general transformation $T(u, v) = (x(u, v), y(u, v))$ and a uniform partition of the domain of T. At a point (u_i, v_j) in our partition, the lines $u = u_i$ and $v = v_j$ can be viewed as parametric curves:

$$\vec{r}_1(t) = \langle t, v_j \rangle, \text{ for } u_i \le t \le u_i + \Delta u, \text{ and}$$
$$\vec{r}_2(t) = \langle u_i, t \rangle, \text{ for } v_j \le t \le v_j + \Delta v.$$

Tangent vectors to these curves are given by

$$\vec{r}_1\,'(t) = \langle 1, 0 \rangle = \vec{i} \quad \text{and} \quad \vec{r}_2\,'(t) = \langle 0, 1 \rangle = \vec{j}.$$

The (i, j)-th rectangle, given by $u_i \le u \le u_i + \Delta u$ and $v_j \le v \le v_j + \Delta v$, has area $\Delta u\, Deltav$.

Viewed another way, this rectangle is a parallelogram spanned by the vectors $\Delta u\vec{i}$ and $\Delta v\vec{j}$. The area of this parallelogram is given by the determinant of the matrix whose columns are these vectors. Of course, this produces the same area:

$$\det \begin{bmatrix} \Delta u & 0 \\ 0 & \Delta v \end{bmatrix} = \Delta u\, \Delta v.$$

Now, let's consider the corresponding region in the x, y plane. The curves in Figure 15.8.35 above can also be realized as parametric curves. In fact, they are precisely the composition of the curves above with our transformation, if we view T as a vector-valued function. We have curves

$$\vec{s}_1(t) = T(\vec{r}_1(t)) = T(t, v_j) = \langle x(t, v_j), y(t, v_j) \rangle$$
$$\vec{s}_2(t) = T(\vec{r}_2(t)) = T(u_i, t) = \langle x(u_i, t), y(u_i, t) \rangle$$

making up two of the four sides of our transformed rectangle.

Now, $\vec{s}_1(t)$ and $\vec{s}_2(t)$ are curves in general, not lines, and the image of our rectangle is no longer rectangular. But for $\Delta u, \Delta v$ small enough, our curves are *approximately* linear, and the image of our rectangle is *approximately* a parallelogram. See Figure 15.8.36.

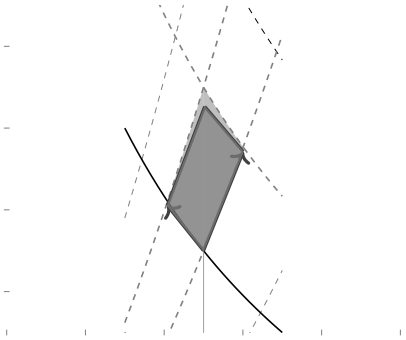


Figure 15.8.36 A transformed rectangle and its parallelogram approximation

We can make linear approximations to vector-valued functions in much the same way as we do for real-valued functions. We have

$$\vec{s}_1(u_i + \Delta u) - \vec{s}_1(u_i) \approx \vec{s}_1\,'(u_i)\Delta u,$$

with a similar result for $\vec{s}_2$. This means that we can approximate the area of our transformed rectangle using the parallelogram spanned by the vectors

$$\vec{a} = \Delta u \vec{s}_1\,'(u_i) = \Delta u \left\langle \frac{\partial x}{\partial u}(u_i, v_j), \frac{\partial y}{\partial u}(u_i, v_j) \right\rangle$$
$$\vec{b} = \Delta u \vec{s}_2\,'(u_i) = \Delta u \left\langle \frac{\partial x}{\partial v}(u_i, v_j), \frac{\partial y}{\partial v}(u_i, v_j) \right\rangle.$$

The area of our transformed region is therefore approximated by the area of the parallelogram spanned by the vectors $\vec{a}$ and $\vec{b}$:

$$\Delta A \approx \det \begin{bmatrix} x_u(u_i, v_j)\Delta u & x_v(u_i, v_j)\Delta v \\ y_u(u_i, v_j)\Delta u & y_v(u_i, v_j)\Delta v \end{bmatrix}$$

Recall that the Chain Rule gives us

$$\frac{d}{dt}(x(\vec{r}(t))) = \nabla x(\vec{r}(t)) \cdot \vec{r}\,'(t),$$

with an analogous result for $y(\vec{r}(t))$. Applying this for the curves $\vec{r}_1(t)$ and $\vec{r}_2(t)$ above allows us to compute the derivatives $\vec{s}_1\,'(t)$ and $\vec{s}_2\,'(t)$.

Note further that we can obtain the same result by writing $\vec{i}$ and $\vec{j}$ as column vectors, and multiplying by the matrix $DT(u_i, v_j)$ — this is the sense in which the derivative acts as a linear transformation on vectors.

$$= \det \begin{bmatrix} x_u(u_i, v_j) & x_v(u_i, v_j) \\ y_u(u_i, v_j) & y_v(u_i, v_j) \end{bmatrix} \Delta u \, \Delta v$$
$$= J_T(u_i, v_j) \, \Delta u \Delta v.$$

This is exactly the result we wanted: the area of our transformed rectangle is approximately the area of the original rectangle, multiplied by the Jacobian.

We can begin to see the change of variables formula by putting this result into the Riemann sum definition of the double integral:

$$f(x_i, y_j)\Delta x \Delta y \approx f(T(u_i, v_j)) \cdot J_T(u_i, v_j)\Delta u_i \Delta v_j.$$

This equation should be viewed somewhat skeptically. The area element on the left is that of a rectangle, not the parallelogram we ended up with above. The argument given here is far from a complete proof of Theorem 15.8.17, but the result is true nonetheless. The interested reader is directed to search online, or seek out the advanced calculus section of their library, should they wish to see a proof.

Chapter 16

Vector Analysis

This chapter explores completely different relationships between vectors and integration. These relationships will enable us to compute the work done by a magnetic field in moving an object along a path and find how much air moves through an oddly-shaped screen in space, among other things.

Our upcoming work with integration will benefit from a review. We are not concerned here with techniques of integration, but rather what an integral "does" and how that relates to the notation we use to describe it.

Integration review.

Recall from Section 15.1 that when R is a region in the xy-plane, $\iint_R dA$ gives the area of the region R. The integral symbols are "elongated esses" meaning "sum" and dA represents "a small amount of area." Taken together, $\iint_R dA$ means "sum up, over R, small amounts of area." This sum then gives the total area of R. We use two integral symbols since R is a two-dimensional region.

Now let $z = f(x, y)$ represent a surface. The integral $\iint_R f(x, y)\, dA$ means "sum up, over R, function values (heights) given by f times small amounts of area." Since "height × area = volume," we are summing small amounts of volume over R, giving the total signed volume under the surface $z = f(x, y)$ and above the xy-plane.

This notation does not directly inform us *how* to evaluate the double integrals to find an area or a volume. With additional work, we recognize that a small amount of area dA can be measured as the area of a small rectangle, with one side length a small change in x and the other side length a small change in y. That is, $dA = dx\, dy$ or $dA = dy\, dx$. We could also compute a small amount of area by thinking in terms of polar coordinates, where $dA = r\, dr\, d\theta$. These understandings lead us to the iterated integrals we used in Chapter 15.

Let us back our review up farther. Note that $\int_1^3 dx = x\Big|_1^3 = 3 - 1 = 2$. We have simply measured the length of the interval $[1, 3]$. We could rewrite the above integral using syntax similar to the double integral syntax above:

$$\int_1^3 dx = \int_I dx, \text{ where } I = [1, 3]\,.$$

We interpret "$\int_I dx$" as meaning "sum up, over the interval I, small changes in x." A change in x is a length along the x-axis, so we are adding up along I small lengths, giving the total length of I.

We could also write $\int_1^3 f(x)\,dx$ as $\int_I f(x)\,dx$, interpreted as "sum up, over I, heights given by $y = f(x)$ times small changes in x." Since "height×length = area," we are summing up areas and finding the total signed area between $y = f(x)$ and the x-axis.

This method of referring to the process of integration can be very powerful. It is the core of our notion of the Riemann Sum. When faced with a quantity to compute, if one can think of a way to approximate its value through a sum, the one is well on their way to constructing an integral (or, double or triple integral) that computes the desired quantity. We will demonstrate this process throughout this chapter, starting with the next section.

16.1 Introduction to Line Integrals

We first used integration to find "area under a curve." In this section, we learn to do this (again), but in a different context.

youtu.be/watch?v=ponQzAg7V3I

Figure 16.1.1 Introducing the line integral

16.1.1 Line Integrals of Functions

Consider the surface and curve shown in Figure 16.1.2(a). The surface is given by $f(x, y) = 1 - \cos(x)\sin(y)$. The dashed curve lies in the xy-plane and is the familiar $y = x^2$ parabola from $-1 \le x \le 1$; we'll call this curve C. The curve drawn with a solid line in the graph is the curve in space that lies on our surface with x and y values that lie on C.

The question we want to answer is this: what is the area that lies below the curve drawn with the solid line? In other words, what is the area of the region above C and under the the surface $z = f(x, y)$? This region is shown in Figure 16.1.2(b).

We suspect the answer can be found using an integral, but before trying to figure out what that integral is, let us first try to approximate its value.

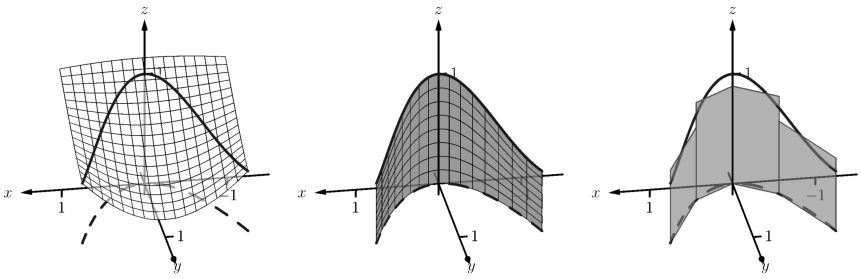


Figure 16.1.2 Finding area under a curve in space

In Figure 16.1.2(c), four rectangles have been drawn over the curve C. The bottom corners of each rectangle lie on C, and each rectangle has a height given by the function $f(x, y)$ for some (x, y) pair along C between the rectangle's bottom corners.

As we know how to find the area of each rectangle, we are able to approximate the area above C and under f. Clearly, our approximation will be *an ap-*

proximation. The heights of the rectangles do not match exactly with the surface f, nor does the base of each rectangle follow perfectly the path of C.

In typical calculus fashion, our approximation can be improved by using more rectangles. The sum of the areas of these rectangles gives an approximate value of the true area above C and under f. As the area of each rectangle is "height × width", we assert that the

$$\text{area above } C \approx \sum(\text{ heights } \times \text{ widths }).$$

When first learning of the integral, and approximating areas with "heights × widths", the width was a small change in x: dx. That will not suffice in this context. Rather, each width of a rectangle is actually approximating the arc length of a small portion of C. In Section 13.5, we used s to represent the arc-length parameter of a curve. A small amount of arc length will thus be represented by ds.

The height of each rectangle will be determined in some way by the surface f. If we parametrize C by s, an s-value corresponds to an (x, y) pair that lies on the parabola C. Since f is a function of x and y, and x and y are functions of s, we can say that f is a function of s. Given a value s, we can compute $f(s)$ and find a height. Thus

$$\begin{aligned}\text{area under } f \text{ and above } C &\approx \sum(\text{ heights } \times \text{ widths });\\ \text{area under } f \text{ and above } C &= \lim_{\|\Delta s\|\to 0}\sum f(c_i)\Delta s_i\\ &= \int_C f(s)\,ds. \qquad (16.1.1)\end{aligned}$$

Here we have introduce a new notation, the integral symbol with a subscript of C. It is reminiscent of our usage of $\iint_R$. Using the train of thought found in the Integration Review preceding this section, we interpret "$\int_C f(s)\,ds$" as meaning "sum up, along a curve C, function values $f(s)\times$small arc lengths." It is understood here that s represents the arc-length parameter.

All this leads us to a definition. The integral found in Equation (16.1.1) is called a *line integral*. We formally define it below, but note that the definition is very abstract. On one hand, one is apt to say "the definition makes sense," while on the other, one is equally apt to say "but I don't know what I'm supposed to do with this definition." We'll address that after the definition, and actually find an answer to the area problem we posed at the beginning of this section.

Definition 16.1.3 Line Integral Over A Scalar Field.

Let C be a smooth curve parametrized by s, the arc-length parameter, and let f be a continuous function of s. A **line integral** is an integral of the form

$$\int_C f(s)\,ds = \lim_{\|\Delta s\|\to 0}\sum_{i=1}^{n} f(c_i)\Delta s_i,$$

where $s_0 < s_1 < \ldots < s_n$ is any partition of the s-interval over which C is defined, c_i is any value in the ith subinterval, Δs_i is the width of the ith subinterval, and $\|\Delta s\|$ is the length of the longest subinterval in the partition.

Note: Definition 16.1.3 uses the term *scalar field* which has not yet been defined. Its meaning is discussed in the paragraph preceding Definition 16.3.2 when it is compared to a *vector field*.

When C is a *closed* curve, i.e., a curve that ends at the same point at which it starts, we use

$$\oint_C f(s)\,ds \text{ instead of } \int_C f(s)\,ds.$$

The definition of the line integral does not specify whether C is a curve in the plane or space (or hyperspace), as the definition holds regardless. For now, we'll assume C lies in the xy-plane.

This definition of the line integral doesn't really say anything new. If C is a curve and s is the arc-length parameter of C on $a \leq s \leq b$, then

$$\int_C f(s)\,ds = \int_a^b f(s)\,ds.$$

The real difference with this integral from the standard "$\int_a^b f(x)\,dx$" we used in the past is that of context. Our previous integrals naturally summed up values over an interval on the x-axis, whereas now we are summing up values over a curve. *If* we can parametrize the curve with the arc-length parameter, we can evaluate the line integral just as before. Unfortunately, parametrizing a curve in terms of the arc-length parameter is usually very difficult, so we must develop a method of evaluating line integrals using a different parametrization.

Given a curve C, find any parametrization of C: $x = g(t)$ and $y = h(t)$, for continuous functions g and h, where $a \leq t \leq b$. We can represent this parametrization with a vector-valued function, $\vec{r}(t) = \langle g(t), h(t)\rangle$.

In Section 13.5, we defined the arc-length parameter in Equation (13.5.1) as

$$s(t) = \int_0^t \|\vec{r}\,'(u)\|\,du.$$

By the Fundamental Theorem of Calculus, $ds = \|\vec{r}\,'(t)\|\,dt$. We can substitute the right hand side of this equation for ds in the line integral definition.

We can view f as being a function of x and y since it is a function of s. Thus $f(s) = f(x,y) = f\big(g(t), h(t)\big)$. This gives us a concrete way to evaluate a line integral:

$$\int_C f(s)\,ds = \int_a^b f\big(g(t), h(t)\big)\,\|\vec{r}\,'(t)\|\,dt.$$

We restate this as a theorem, along with its three-dimensional analogue, followed by an example where we finally evaluate an integral and find an area.

Theorem 16.1.4 Evaluating a Line Integral Over A Scalar Field.

- *Let C be a curve parametrized by $\vec{r}(t) = \langle g(t), h(t)\rangle$, $a \leq t \leq b$, where g and h are continuously differentiable, and let $z = f(x,y)$, where f is continuous over C. Then*

$$\int_C f(s)\,ds = \int_a^b f\big(g(t), h(t)\big)\,\|\vec{r}\,'(t)\|\,dt.$$

- *Let C be a curve parametrized by $\vec{r}(t) = \langle g(t), h(t), k(t)\rangle$, $a \leq t \leq b$, where g, h and k are continuously differentiable, and let $w = f(x,y,z)$, where f is continuous over C. Then*

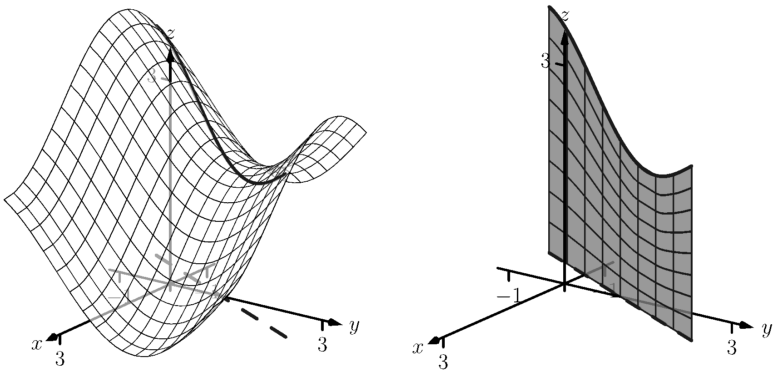
$$\int_C f(s)\,ds = \int_a^b f\big(g(t), h(t), k(t)\big)\,\|\vec{r}\,'(t)\|\,dt.$$

To be clear, the first point of Theorem 16.1.4 can be used to find the area under a surface $z = f(x,y)$ and above a curve C. We will later give an understanding of the line integral when C is a curve in space.

Let's do an example where we actually compute an area.

Example 16.1.5 Evaluating a line integral: area under a surface over a curve.

Find the area under the surface $f(x, y) = \cos(x) + \sin(y) + 2$ over the curve C, which is the segment of the line $y = 2x + 1$ on $-1 \le x \le 1$, as shown in Figure 16.1.6.



(a) **(b)**

Figure 16.1.6 Finding area under a curve in Example 16.1.5

Solution. Our first step is to represent C with a vector-valued function. Since C is a simple line, and we have a explicit relationship between y and x (namely, that y is $2x + 1$), we can let $x = t$, $y = 2t + 1$, and write $\vec{r}(t) = \langle t, 2t + 1\rangle$ for $-1 \le t \le 1$.

We find the values of f over C as $f(x, y) = f(t, 2t + 1) = \cos(t) + \sin(2t + 1) + 2$. We also need $\|\vec{r}\,'(t)\|$; with $\vec{r}\,'(t) = \langle 1, 2\rangle$, we have $\|\vec{r}\,'(t)\| = \sqrt{5}$. Thus $ds = \sqrt{5}\,dt$.

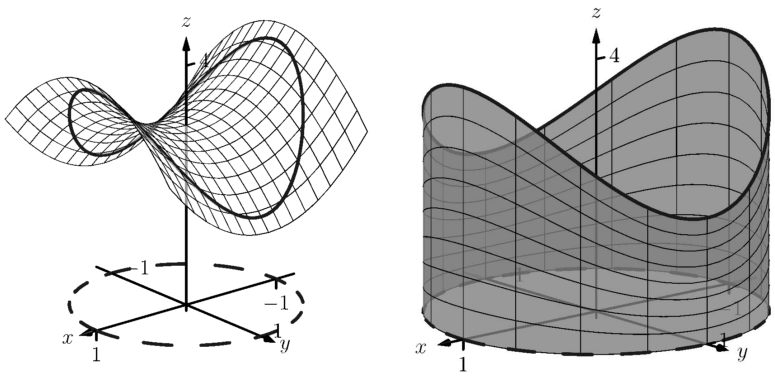
The area we seek is

$$\begin{aligned}\int_C f(s)\,ds &= \int_{-1}^{1} \big(\cos(t) + \sin(2t + 1) + 2\big)\sqrt{5}\,dt \\ &= \sqrt{5}\big(\sin(t) - \frac{1}{2}\cos(2t + 1) + 2t\big)\Big|_{-1}^{1} \\ &\approx 14.418 \text{ units}^2.\end{aligned}$$

We will practice setting up and evaluating a line integral in another example, then find the area described at the beginning of this section.

Example 16.1.7 Evaluating a line integral: area under a surface over a curve.

Find the area over the unit circle in the xy-plane and under the graph of $f(x, y) = x^2 - y^2 + 3$, shown in Figure 16.1.8.



(a) **(b)**

Figure 16.1.8 Finding area under a curve in Example 16.1.7

Solution. The curve C is the unit circle, which we will describe with the parametrization $\vec{r}(t) = \langle \cos t, \sin t \rangle$ for $0 \le t \le 2\pi$. We find $\|\vec{r}\,'(t)\| = 1$, so $ds = 1dt$.

We find the values of f over C as $f(x, y) = f(\cos t, \sin t) = \cos^2 t - \sin^2 t + 3$. Thus the area we seek is (note the use of the $\oint f(s)ds$ notation):

$$\begin{aligned} \oint_C f(s)\, ds &= \int_0^{2\pi} \left(\cos^2 t - \sin^2 t + 3\right) dt \\ &= 6\pi. \end{aligned}$$

(Note: we may have approximated this answer from the start. The unit circle has a circumference of 2π, and we may have guessed that due to the apparent symmetry of our surface, the average height of the surface is 3.)

We now consider the example that introduced this section.

Example 16.1.9 Evaluating a line integral: area under a surface over a curve.

Find the area under $f(x, y) = 1 - \cos(x)\sin(y)$ and over the parabola $y = x^2$, from $-1 \le x \le 1$.

Solution. We parametrize our curve C as $\vec{r}(t) = \langle t, t^2 \rangle$ for $-1 \le t \le 1$; we find $\|\vec{r}\,'(t)\| = \sqrt{1 + 4t^2}$, so $ds = \sqrt{1 + 4t^2}\, dt$.

Replacing x and y with their respective functions of t, we have $f(x, y) = f(t, t^2) = 1 - \cos(t)\sin(t^2)$. Thus the area under f and over C is found to be

$$\int_C f(s)\, ds = \int_{-1}^{1} \left(1 - \cos(t)\sin\left(t^2\right)\right)\sqrt{1 + t^2}\, dt.$$

This integral is impossible to evaluate using the techniques developed in this text. We resort to a numerical approximation; accurate to two places after the decimal, we find the area is

$$= 2.17.$$

We give one more example of finding area.

Example 16.1.10 Evaluating a line integral: area under a curve in space.

Find the area above the xy-plane and below the helix parametrized by $\vec{r}(t) = \langle \cos t, 2\sin t, t/\pi \rangle$, for $0 \le t \le 2\pi$, as shown in Figure 16.1.11.

Solution. Note how this is problem is different than the previous examples: here, the height is not given by a surface, but by the curve itself.

We use the given vector-valued function $\vec{r}(t)$ to determine the curve C in the xy-plane by simply using the first two components of \vec r(t): $\vec{c}(t) = \langle \cos t, 2\sin t \rangle$. Thus $ds = \|\vec{c}\,'(t)\|\, dt = \sqrt{\sin^2 t + 4\cos^2 t}\, dt$.

The height is not found by evaluating a surface over C, but rather it is given directly by the third component of \vec r(t): t/π. Thus

$$\oint_C f(s)\, ds = \int_0^{2\pi} \frac{t}{\pi}\sqrt{\sin^2 t + 4\cos^2 t}\, dt \approx 9.69,$$

where the approximation was obtained using numerical methods.

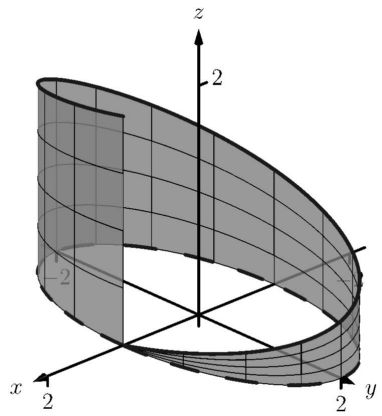


Figure 16.1.11 Finding area under a curve in Example 16.1.10

Note how in each of the previous examples we are effectively finding "area under a curve", just as we did when first learning of integration. We have used the phrase "area *over* a curve C and under a surface," but that is because of the important role C plays in the integral. The figures show how the curve C defines another curve on the surface $z = f(x, y)$, and we are finding the area under that curve.

16.1.2 Properties of Line Integrals

Many properties of line integrals can be inferred from general integration properties. For instance, if k is a scalar, then $\int_C k\, f(s)ds = k \int_C f(s)ds$.

One property in particular of line integrals is worth noting. If C is a curve composed of subcurves C_1 and C_2, where they share only one point in common (see Figure 16.1.13(a), then the line integral over C is the sum of the line integrals over C_1 and C_2:

$$\int_C f(s)\, ds = \int_{C_1} f(s)\, ds + \int_{C_2} f(s)\, ds.$$

(a) **(b)**

Figure 16.1.13 Illustrating properties of line integrals

This property allows us to evaluate line integrals over some curves C that are not smooth. Note how in Figure 16.1.13(b) the curve is not smooth at D, so by our definition of the line integral we cannot evaluate $\int_C f(s)ds$. However, one can evaluate line integrals over C_1 and C_2 and their sum will be the desired quantity.

A curve C that is composed of two or more smooth curves is said to be *piecewise smooth*. In this chapter, any statement that is made about smooth curves also holds for piecewise smooth curves.

youtu.be/watch?v=EN9r0JKI5uY

Figure 16.1.12 Another line integral example

We state these properties as a theorem.

Theorem 16.1.14 Properties of Line Integrals Over Scalar Fields.

1. *Let C be a smooth curve parametrized by the arc-length parameter s, let f and g be continuous functions of s, and let k_1 and k_2 be scalars. Then*
$$\int_C \big(k_1 f(s) + k_2 g(s)\big)\, ds = k_1 \int_C f(s)\, ds + k_2 \int_C g(s)\, ds.$$
2. *Let C be piecewise smooth, composed of smooth components C_1 and C_2. Then*
$$\int_C f(s)\, ds = \int_{C_1} f(s)\, ds + \int_{C_2} f(s)\, ds.$$

16.1.3 Mass and Center of Mass

We first learned integration as a method to find area under a curve, then later used integration to compute a variety of other quantities, such as arc length, volume, force, etc. In this section, we also introduced line integrals as a method to find area under a curve, and now we explore one more application.

Let a curve C (either in the plane or in space) represent a thin wire with variable density $\delta(s)$. We can approximate the mass of the wire by dividing the wire (i.e., the curve) into small segments of length Δs_i and assume the density is constant across these small segments. The mass of each segment is density of the segment × its length; by summing up the approximate mass of each segment we can approximate the total mass:

$$\text{Total Mass of Wire } = \sum \delta(s_i)\Delta s_i.$$

By taking the limit as the length of the segments approaches 0, we have the definition of the line integral as seen in Definition 16.1.3. When learning of the line integral, we let $f(s)$ represent a height; now we let $f(s) = \delta(s)$ represent a density.

We can extend this understanding of computing mass to also compute the center of mass of a thin wire. (As a reminder, the center of mass can be a useful piece of information as objects rotate about that center.) We give the relevant formulas in the next definition, followed by an example. Note the similarities between this definition and Definition 15.6.30, which gives similar properties of solids in space.

Definition 16.1.15 Mass, Center of Mass of Thin Wire.

Let a thin wire lie along a smooth curve C with continuous density function $\delta(s)$, where s is the arc length parameter.

1. The *mass* of the thin wire is $M = \displaystyle\int_C \delta(s)\, ds$.
2. The *moment about the yz-plane* is $M_{yz} = \displaystyle\int_C x\delta(s)\, ds$.
3. The *moment about the xz-plane* is $M_{xz} = \displaystyle\int_C y\delta(s)\, ds$.

4. The *moment about the xy-plane* is $M_{xy} = \displaystyle\int_C z\delta(s)\,ds$.

5. The *center of mass* of the wire is

$$(\overline{x},\overline{y},\overline{z}) = \left(\frac{M_{yz}}{M}, \frac{M_{xz}}{M}, \frac{M_{xy}}{M}\right).$$

Example 16.1.16 Evaluating a line integral: calculating mass.

A thin wire follows the path $\vec{r}(t) = \langle 1+\cos t, 1+\sin t, 1+\sin(2t)\rangle$, $0 \le t \le 2\pi$. The density of the wire is determined by its position in space: $\delta(x,y,z) = y+z$ gm/cm. The wire is shown in Figure 16.1.17, where a light color indicates low density and a dark color represents high density. Find the mass and center of mass of the wire.

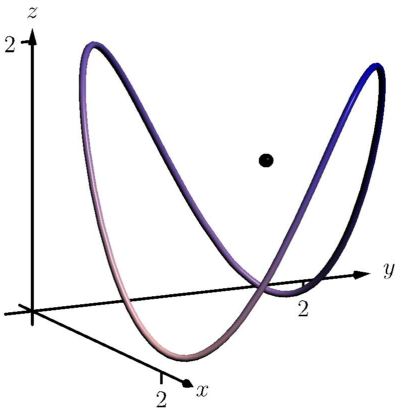


Figure 16.1.17 Finding the mass of a thin wire in Example 16.1.16

Solution. We compute the density of the wire as

$$\delta(x,y,z) = \delta\big(1+\cos t, 1+\sin t, 1+\sin(2t)\big) = 2+\sin t+\sin(2t).$$

We compute ds as

$$ds = \|\vec{r}\,'(t)\|\,dt = \sqrt{\sin^2 t+\cos^2 t+4\cos^2(2t)}\,dt = \sqrt{1+4\cos^2(2t)}\,dt.$$

Thus the mass is

$$M = \oint_C \delta(s)\,ds = \int_0^{2\pi} \big(2+\sin t+\sin(2t)\big)\sqrt{1+4\cos^2(2t)}\,dt \approx 21.08 \text{ gm}.$$

We compute the moments about the coordinate planes:

$$\begin{aligned} M_{yz} &= \oint_C x\delta(s)\,ds \\ &= \int_0^{2\pi} (1+\cos t)\big(2+\sin t+\sin(2t)\big)\sqrt{1+4\cos^2(2t)}\,dt \\ &\approx 21.08. \end{aligned}$$

$$\begin{aligned} M_{xz} &= \oint_C y\delta(s)\,ds \\ &= \int_0^{2\pi} (1+\sin t)\big(2+\sin t+\sin(2t)\big)\sqrt{1+4\cos^2(2t)}\,dt \\ &\approx 26.35 \end{aligned}$$

$$\begin{aligned} M_{xy} &= \oint_C z\delta(s)\,ds \\ &= \int_0^{2\pi} \big(1+\sin(2t)\big)\big(2+\sin t+\sin(2t)\big)\sqrt{1+4\cos^2(2t)}\,dt \\ &\approx 25.40 \end{aligned}$$

Thus the center of mass of the wire is located at

$$(\overline{x},\overline{y},\overline{z}) = \left(\frac{M_{yz}}{M}, \frac{M_{xz}}{M}, \frac{M_{xy}}{M}\right) \approx (1, 1.25, 1.20),$$

as indicated by the dot in Figure 16.1.17. Note how in this example, the curve C is "centered" about the point $(1,1,1)$, though the variable density of the wire pulls the center of mass out along the y and z axes.

We end this section with a callback to the Integration Review that preceded this section. A line integral looks like: $\int_C f(s)\,ds$. As stated before the definition of the line integral, this means "sum up, along a curve C, function values $f(s)$ × small arc lengths." When $f(s)$ represents a height, we have "height × length = area." When $f(s)$ is a density (and we use $\delta(s)$ by convention), we have "density (mass per unit length) × length = mass."

In the next section, we investigate a new mathematical object, the *vector field*. The remaining sections of this chapter are devoted to understanding integration in the context of vector fields.

16.1.4 Exercises

Terms and Concepts

1. Explain how a line integral can be used to find the area under a curve.

2. How does the evaluation of a line integral given as $\int_C f(s)\,ds$ differ from a line integral given as $\oint_C f(s)\,ds$?

3. Why are most line integrals evaluated using Key Idea 16.1.4 instead of "directly" as $\int_C f(s)\,ds$?

4. Sketch a closed, piecewise smooth curve composed of three subcurves.

Problems

Exercise Group. In the following exercises, a planar curve C is given along with a function f that is defined over C. Evaluate the line integral $\displaystyle\int_C f(s)\,ds$.

5. C is the line segment joining the points $(-2,-1)$ and $(1,2)$; the function is $f(x,y) = x^2 + y^2 + 2$.

6. C is the segment of $y = 3x + 2$ on $[1,2]$; the function is $f(x,y) = 5x + 2y$.

7. C is the circle with radius 2 centered at the point $(4,2)$; the function is $f(x,y) = 3x - y$.

8. C is the curve given by $\vec{r}(t) = \langle \cos t + t\sin t, \sin t - t\cos t\rangle$ on $[0, 2\pi]$; the function is $f(x,y) = 5$.

9. C is the piecewise curve composed of the line segments that connect $(0,1)$ to $(1,1)$, then connect $(1,1)$ to $(1,0)$; the function is $f(x,y) = x + y^2$.

10. C is the piecewise curve composed of the line segment joining the points $(0,0)$ and $(1,1)$, along with the quarter-circle parametrized by $\langle \cos t, -\sin t + 1\rangle$ on $[0, \pi/2]$(which starts at the point $(1,1)$ and ends at $(0,0)$; the function is $f(x,y) = x^2 + y^2$.

Exercise Group. In the following exercises, a planar curve C is given along with a function f that is defined over C. Set up the line integral $\displaystyle\int_C f(s)\,ds$, then approximate its value using technology.

11. C is the portion of the parabola $y = 2x^2 + x + 1$ on $[0,1]$; the function is $f(x,y) = x^2 + 2y$.

12. C is the portion of the curve $y = \sin x$ on $[0,\pi]$; the function is $f(x,y) = x$.

13. C is the ellipse given by $\vec{r}(t) = \langle 2\cos t, \sin t\rangle$ on $[0, 2\pi]$; the function is $f(x,y) = 10 - x^2 - y^2$.

14. C is the portion of $y = x^3$ on $[-1,1]$; the function is $f(x,y) = 2x + 3y + 5$.

Exercise Group. In the following exercises, a parametrized curve C in space is given. Find the area above the xy-plane that is under C.

15. C: $\vec{r}(t) = \langle 5t, t, t^2\rangle$ for $1 \le t \le 2$.

16. C: $\vec{r}(t) = \langle \cos t, \sin t, \sin(2t) + 1\rangle$ for $0 \le t \le 2\pi$.

17. C: $\vec{r}(t) = \langle 3\cos t, 3\sin t, t^2\rangle$ for $0 \le t \le 2\pi$.

18. C: $\vec{r}(t) = \langle 3t, 4t, t\rangle$ for $0 \le t \le 1$.

Exercise Group. In the following exercises, a parametrized curve C is given that represents a thin wire with density δ. Find the mass and center of mass of the thin wire.

19. C: $\vec{r}(t) = \langle \cos t, \sin t, t\rangle$ for $0 \le t \le 4\pi$; $\delta(x,y,z) = z$.

20. C: $\vec{r}(t) = \langle t - t^2, t^2 - t^3, t^3 - t^4\rangle$ for $0 \le t \le 1$; $\delta(x,y,z) = x + 2y + 2z$. Use technology to approximate the value of each integral.

16.2 Vector Fields

We have studied functions of two and three variables, where the input of such functions is a point (either a point in the plane or in space) and the output is a number.

We could also create functions where the input is a point (again, either in the plane or in space), but the output is a *vector*. For instance, we could create the following function: $\vec{F}(x,y) = \langle x+y, x-y\rangle$, where $\vec{F}(2,3) = \langle 5,-1\rangle$. We are to think of $\vec{F}$ assigning the vector $\langle 5,-1\rangle$ to the point $(2,3)$; in some sense, the vector $\langle 5,-1\rangle$ lies at the point $(2,3)$.

youtu.be/watch?v=hnQqjC0cwKY

Figure 16.2.1 Introducing vector fields

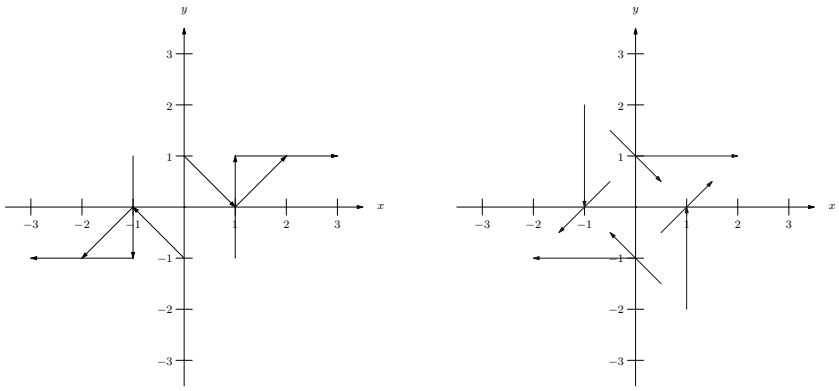
Such functions are extremely useful in any context where magnitude and direction are important. For instance, we could create a function $\vec{F}$ that represents the electromagnetic force exerted at a point by a electromagnetic field, or the velocity of air as it moves across an airfoil.

Because these functions are so important, we need to formally define them.

Definition 16.2.2 Vector Field.

1. A *vector field in the plane* is a function $\vec{F}(x,y)$ whose domain is a subset of $\mathbb{R}^2$ and whose output is a two-dimensional vector:
$$\vec{F}(x,y) = \langle M(x,y), N(x,y)\rangle.$$
2. A *vector field in space* is a function $\vec{F}(x,y,z)$ whose domain is a subset of $\mathbb{R}^3$ and whose output is a three-dimensional vector:
$$\vec{F}(x,y,z) = \langle M(x,y,z), N(x,y,z), P(x,y,z)\rangle.$$

This definition may seem odd at first, as a special type of function is called a "field." However, as the function determines a "field of vectors", we can say the field is *defined by* the function, and thus the field *is* a function.



(a) **(b)**

Figure 16.2.3 Demonstrating methods of graphing vector fields

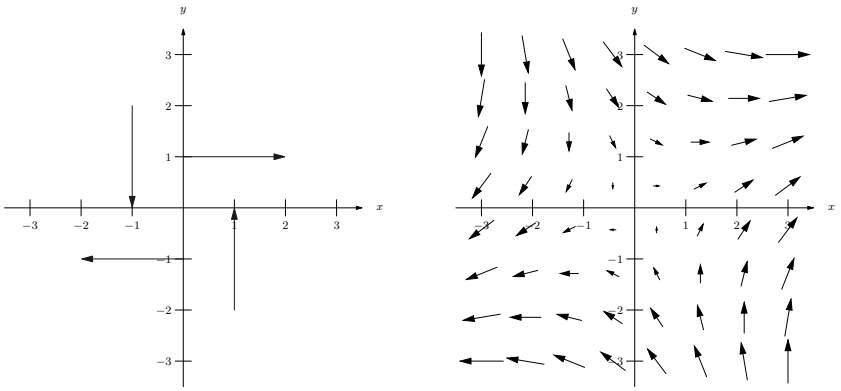
Visualizing vector fields helps cement this connection. When graphing a vector field in the plane, the general idea is to draw the vector $\vec{F}(x,y)$ at the point (x,y). For instance, using $\vec{F}(x,y) = \langle x+y, x-y\rangle$ as before, at $(1,1)$ we would draw $\langle 2,0\rangle$.

In Figure 16.2.3(a), one can see that the vector $\langle 2,0\rangle$ is drawn *starting from* the point $(1,1)$. A total of 8 vectors are drawn, with the x- and y-values of

$-1, 0, 1$. In many ways, the resulting graph is a mess; it is hard to tell what this field "looks like."

In Figure 16.2.3(b), the same field is redrawn with each vector $\vec{F}(x, y)$ drawn *centered on* the point (x, y). This makes for a better looking image, though the long vectors can cause confusion: when one vector intersects another, the image looks cluttered.

A common way to address this problem is limit the length of each arrow, and represent long vectors with thick arrows, as done in Figure 16.2.4(a). Usually we do not use a graph of a vector field to determine exactly the magnitude of a particular vector. Rather, we are more concerned with the relative magnitudes of vectors: which are bigger than others? Thus limiting the length of the vectors is not problematic.



(a) **(b)**

Figure 16.2.4 Demonstrating methods of graphing vector fields

Drawing arrows with variable thickness is best done with technology; search the documentation of your favorite graphing program for terms like "vector fields" or "slope fields" to learn how. Technology obviously allows us to plot many vectors in a vector field nicely; in Figure 16.2.4(b), we see the same vector field drawn with many vectors, and finally get a clear picture of how this vector field behaves. (If this vector field represented the velocity of air moving across a flat surface, we could see that the air tends to move either to the upper-right or lower-left, and moves very slowly near the origin.)

We can similarly plot vector fields in space, as shown in Figure 16.2.5, though it is not often done. The plots get very busy very quickly, as there are lots of arrows drawn in a small amount of space. In Figure 16.2.5 the field $\vec{F} = \langle -y, x, z \rangle$ is graphed. If one could view the graph from above, one could see the arrows point in a circle about the z-axis. One should also note how the arrows far from the origin are larger than those close to the origin.

It is good practice to try to visualize certain vector fields in one's head. For instance, consider a point mass at the origin and the vector field that represents the gravitational force exerted by the mass at any point in the room. The field would consist of arrows pointing toward the origin, increasing in size as they near the origin (as the gravitational pull is strongest near the point mass).

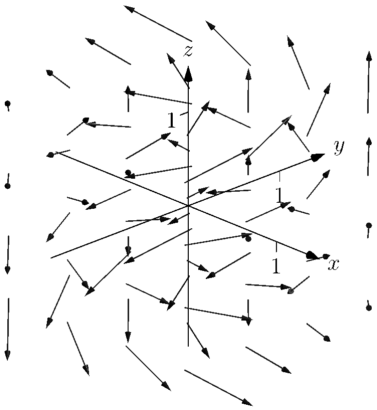


Figure 16.2.5 Graphing a vector field in space

16.2.1 Vector Field Notation and Del Operator

Definition 16.2.2 defines a vector field $\vec{F}$ using the notation

$$\vec{F}(x,y) = \langle M(x,y), N(x,y)\rangle \text{ and } \vec{F}(x,y,z) = \langle M(x,y,z), N(x,y,z), P(x,y,z)\rangle.$$

That is, the components of $\vec{F}$ are each functions of x and y (and also z in space). As done in other contexts, we will drop the "of x, y and z" portions of the notation and refer to vector fields in the plane and in space as

$$\vec{F} = \langle M, N\rangle \text{ and } \vec{F} = \langle M, N, P\rangle,$$

respectively, as this shorthand is quite convenient.

Another item of notation will become useful: the "del operator." Recall in Section 14.3 how we used the symbol ∇ (pronounced "del") to represent the gradient of a function of two variables. That is, if $z = f(x,y)$, then "del f" $= \nabla f = \langle f_x, f_y\rangle$.

We now define ∇ to be the "del operator." It is a vector whose components are partial derivative operations.

youtu.be/watch?v=L7AEW0NsyHA

Figure 16.2.6 Introducing the del operator

In the plane, $\nabla = \left\langle \dfrac{\partial}{\partial x}, \dfrac{\partial}{\partial y}\right\rangle$; in space, $\nabla = \left\langle \dfrac{\partial}{\partial x}, \dfrac{\partial}{\partial y}, \dfrac{\partial}{\partial z}\right\rangle$.

With this definition of ∇, we can better understand the gradient ∇f. As f returns a scalar, the properties of scalar and vector multiplication gives

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

Now apply the del operator ∇ to vector fields. Let $\vec{F} = \langle x+\sin y, y^2+z, x^2\rangle$. We can use vector operations and find the dot product of ∇ and $\vec{F}$:

$$\begin{aligned}\nabla \cdot \vec{F} &= \left\langle \frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z}\right\rangle \cdot \langle x+\sin y, y^2+z, x^2\rangle \\ &= \frac{\partial}{\partial x}(x+\sin y) + \frac{\partial}{\partial y}(y^2+z) + \frac{\partial}{\partial z}(x^2) \\ &= 1 + 2y.\end{aligned}$$

youtu.be/watch?v=QbkbxwicFRI

Figure 16.2.7 Del and the gradient

We can also compute their cross products:

$$\begin{aligned}\nabla \times \vec{F} &= \left\langle \frac{\partial}{\partial y}(x^2) - \frac{\partial}{\partial z}(y^2+z), \frac{\partial}{\partial z}(x+\sin y) - \frac{\partial}{\partial x}(x^2), \frac{\partial}{\partial x}(y^2+z) - \frac{\partial}{\partial y}(x+\sin y)\right\rangle \\ &= \langle -1, -2x, -\cos y\rangle.\end{aligned}$$

We do not yet know why we would want to compute the above. However, as we next learn about properties of vector fields, we will see how these dot and cross products with the del operator are quite useful.

16.2.2 Divergence and Curl

Two properties of vector fields will prove themselves to be very important: divergence and curl. Each is a special "derivative" of a vector field; that is, each measures an instantaneous rate of change of a vector field.

If the vector field represents the velocity of a fluid or gas, then the *divergence* of the field is a measure of the "compressibility" of the fluid. If the divergence is negative at a point, it means that the fluid is compressing: more fluid is going into the point than is going out. If the divergence is positive, it means the fluid is expanding: more fluid is going out at that point than going in. A divergence of

youtu.be/watch?v=10Yh8JUGRYU

Figure 16.2.8 Introducing the divergence

zero means the same amount of fluid is going in as is going out. If the divergence is zero at all points, we say the field is *incompressible*.

It turns out that the proper measure of divergence is simply $\nabla \cdot \vec{F}$, as stated in the following definition.

Definition 16.2.9 Divergence of a Vector Field.

The **divergence** of a vector field $\vec{F}$ is

$$\text{div } \vec{F} = \nabla \cdot \vec{F}.$$

- In the plane, with $\vec{F} = \langle M, N \rangle$, div $\vec{F} = M_x + N_y$.
- In space, with $\vec{F} = \langle M, N, P \rangle$, div $\vec{F} = M_x + N_y + P_z$.

Curl is a measure of the spinning action of the field. Let $\vec{F}$ represent the flow of water over a flat surface. If a small round cork were held in place at a point in the water, would the water cause the cork to spin? No spin corresponds to zero curl; counterclockwise spin corresponds to positive curl and clockwise spin corresponds to negative curl.

In space, things are a bit more complicated. Again let $\vec{F}$ represent the flow of water, and imagine suspending a tennis ball in one location in this flow. The water may cause the ball to spin along an axis. If so, the curl of the vector field is a *vector* (not a *scalar*, as before), parallel to the axis of rotation, following a right hand rule: when the thumb of one's right hand points in the direction of the curl, the ball will spin in the direction of the curling fingers of the hand.

In space, it turns out the proper measure of curl is $\nabla \times \vec{F}$, as stated in the following definition. To find the curl of a planar vector field $\vec{F} = \langle M, N \rangle$, embed it into space as $\vec{F} = \langle M, N, 0 \rangle$ and apply the cross product definition. Since M and N are functions of just x and y (and not z), all partial derivatives with respect to z become 0 and the result is simply $\langle 0, 0, N_x - M_y \rangle$. The third component is the measure of curl of a planar vector field.

Definition 16.2.11 Curl of a Vector Field.

- Let $\vec{F} = \langle M, N \rangle$ be a vector field in the plane. The *curl* of $\vec{F}$ is curl $\vec{F} = N_x - M_y$.
- Let $\vec{F} = \langle M, N, P \rangle$ be a vector field in space. The *curl* of $\vec{F}$ is curl $\vec{F} = \nabla \times \vec{F} = \langle P_y - N_z, M_z - P_x, N_x - M_y \rangle$.

We adopt the convention of referring to curl as $\nabla \times \vec{F}$, regardless of whether $\vec{F}$ is a vector field in two or three dimensions.

We now practice computing these quantities.

Example 16.2.13 Computing divergence and curl of planar vector fields.

For each of the planar vector fields given below, view its graph and try to visually determine if its divergence and curl are 0. Then compute the divergence and curl.

1. $\vec{F} = \langle y, 0 \rangle$ (see Figure 16.2.14(a))
2. $\vec{F} = \langle -y, x \rangle$ (see Figure 16.2.14(b))

youtu.be/watch?v=PMZXQDhcA2I

Figure 16.2.10 Introducing the curl

youtu.be/watch?v=PXMWLny21CI

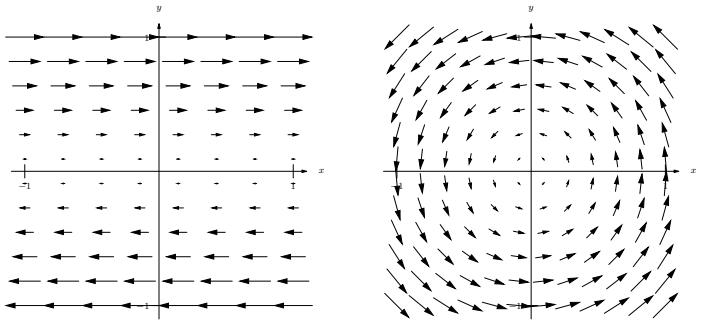
Figure 16.2.12 Interpreting the curl of a vector field

3. $\vec{F} = \langle x, y \rangle$ (see Figure 16.2.15(a))
4. $\vec{F} = \langle \cos y, \sin x \rangle$ (see Figure 16.2.15(b))

Solution.

1. The arrow sizes are constant along any horizontal line, so if one were to draw a small box anywhere on the graph, it would seem that the same amount of fluid would enter the box as exit. Therefore it seems the divergence is zero; it is, as

$$\text{div}\,\vec{F} = \nabla \cdot \vec{F} = M_x + N_y = \frac{\partial}{\partial x}(y) + \frac{\partial}{\partial y}(0) = 0.$$



(a) **(b)**

Figure 16.2.14 The vector fields in parts 1 and 2 in Example 16.2.13

At any point on the x-axis, arrows above it move to the right and arrows below it move to the left, indicating that a cork placed on the axis would spin clockwise. A cork placed anywhere above the x-axis would have water above it moving to the right faster than the water below it, also creating a clockwise spin. A clockwise spin also appears to be created at points below the x-axis. Thus it seems the curl should be negative (and not zero). Indeed, it is:

$$\text{curl}\,\vec{F} = \nabla \times \vec{F} = N_x - M_y = \frac{\partial}{\partial x}(0) - \frac{\partial}{\partial y}(y) = -1.$$

2. It appears that all vectors that lie on a circle of radius r, centered at the origin, have the same length (and indeed this is true). That implies that the divergence should be zero: draw any box on the graph, and any fluid coming in will lie along a circle that takes the same amount of fluid out. Indeed, the divergence is zero, as

$$\text{div}\,\vec{F} = \nabla \cdot \vec{F} = M_x + N_y = \frac{\partial}{\partial x}(-y) + \frac{\partial}{\partial y}(x) = 0.$$

Clearly this field moves objects in a circle, but would it induce a cork to spin? It appears that yes, it would: place a cork anywhere in the flow, and the point of the cork closest to the origin would feel less flow than the point on the cork farthest from the origin, which would induce a counterclockwise flow. Indeed, the curl is positive:

$$\text{curl}\,\vec{F} = \nabla \times \vec{F} = N_x - M_y = \frac{\partial}{\partial x}(x) - \frac{\partial}{\partial y}(-y) = 1 - (-1) = 2.$$

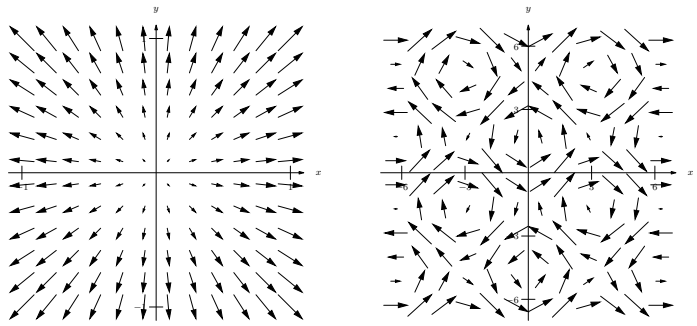
Since the curl is constant, we conclude the induced spin is the same no matter where one is in this field.

3. At the origin, there are many arrows pointing out but no arrows pointing in. We conclude that at the origin, the divergence must be positive (and not zero). If one were to draw a box anywhere in the field, the edges farther from the origin would have larger arrows passing through them than the edges close to the origin, indicating that more is going from a point than going in. This indicates a positive (and not zero) divergence. This is correct:

$$\text{div}\,\vec{F} = \nabla \cdot \vec{F} = M_x + N_y = \frac{\partial}{\partial x}(x) + \frac{\partial}{\partial y}(y) = 1 + 1 = 2.$$

One may find this curl to be harder to determine visually than previous examples. One might note that any arrow that induces a clockwise spin on a cork will have an equally sized arrow inducing a counterclockwise spin on the other side, indicating no spin and no curl. This is correct, as

$$\text{curl}\,\vec{F} = \nabla \times \vec{F} = N_x - M_y = \frac{\partial}{\partial x}(y) - \frac{\partial}{\partial y}(x) = 0.$$



(a) **(b)**

Figure 16.2.15 The vector fields in parts 3 and 4 in Example 16.2.13

4. One might find this divergence hard to determine visually as large arrows appear in close proximity to small arrows, each pointing in different directions. Instead of trying to rationalize a guess, we compute the divergence:

$$\text{div}\,\vec{F} = \nabla \cdot \vec{F} = M_x + N_y = \frac{\partial}{\partial x}(\cos y) + \frac{\partial}{\partial y}(\sin x) = 0.$$

Perhaps surprisingly, the divergence is 0. With all the loops of different directions in the field, one is apt to reason the curl is variable. Indeed, it is:

$$\text{curl}\,\vec{F} = \nabla \times \vec{F} = N_x - M_y = \frac{\partial}{\partial x}(\sin x) - \frac{\partial}{\partial y}(\cos y) = \cos x + \sin y.$$

Depending on the values of x and y, the curl may be positive, negative, or zero.

Example 16.2.16 Computing divergence and curl of vector fields in space.

Compute the divergence and curl of each of the following vector fields.

1. $\vec{F} = \langle x^2 + y + z, -x - z, x + y \rangle$
2. $\vec{F} = \langle e^{xy}, \sin(x + z), x^2 + y \rangle$

Solution. We compute the divergence and curl of each field following the definitions.

1.

$$\begin{aligned} \text{div } \vec{F} &= \nabla \cdot \vec{F} = M_x + N_y + P_z = 2x + 0 + 0 = 2x \\ \text{curl } \vec{F} &= \nabla \times \vec{F} = \langle P_y - N_z, M_z - P_x, N_x - M_y \rangle \\ &= \langle 1 - (-1), 1 - 1, -1 - (1) \rangle = \langle 2, 0, -2 \rangle. \end{aligned}$$

For this particular field, no matter the location in space, a spin is induced with axis parallel to $\langle 2, 0, -2 \rangle$.

2.

$$\begin{aligned} \text{div } \vec{F} &= \nabla \cdot \vec{F} = M_x + N_y + P_z = ye^{xy} + 0 + 0 = ye^{xy} \\ \text{curl } \vec{F} &= \nabla \times \vec{F} = \langle P_y - N_z, M_z - P_x, N_x - M_y \rangle \\ &= \langle 1 - \cos(x + z), -2x, \cos(x + z) - xe^{xy} \rangle \end{aligned}$$

youtu.be/watch?v=_QPgbwG_SHw

Figure 16.2.17 Further examples with divergence

Example 16.2.18 Creating a field representing gravitational force.

The force of gravity between two objects is inversely proportional to the square of the distance between the objects. Locate a point mass at the origin. Create a vector field $\vec{F}$ that represents the gravitational pull of the point mass at any point (x, y, z). Find the divergence and curl of this field.

Solution. The point mass pulls toward the origin, so at (x, y, z), the force will pull in the direction of $\langle -x, -y, -z \rangle$. To get the proper magnitude, it will be useful to find the unit vector in this direction. Dividing by its magnitude, we have

$$\vec{u} = \left\langle \frac{-x}{\sqrt{x^2 + y^2 + z^2}}, \frac{-y}{\sqrt{x^2 + y^2 + z^2}}, \frac{-z}{\sqrt{x^2 + y^2 + z^2}} \right\rangle.$$

The magnitude of the force is inversely proportional to the square of the distance between the two points. Letting k be the constant of proportionality, we have the magnitude as $\dfrac{k}{x^2 + y^2 + z^2}$. Multiplying this magnitude by the unit vector above, we have the desired vector field:

$$\vec{F} = \left\langle \frac{-kx}{(x^2 + y^2 + z^2)^{3/2}}, \frac{-ky}{(x^2 + y^2 + z^2)^{3/2}}, \frac{-kz}{(x^2 + y^2 + z^2)^{3/2}} \right\rangle.$$

We leave it to the reader to confirm that div $\vec{F} = 0$ and curl $\vec{F} = \vec{0}$.

The analogous planar vector field is given in Figure 16.2.19. Note how all arrows point to the origin, and the magnitude gets very small when "far" from the origin.

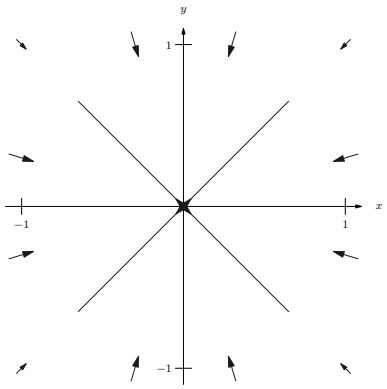


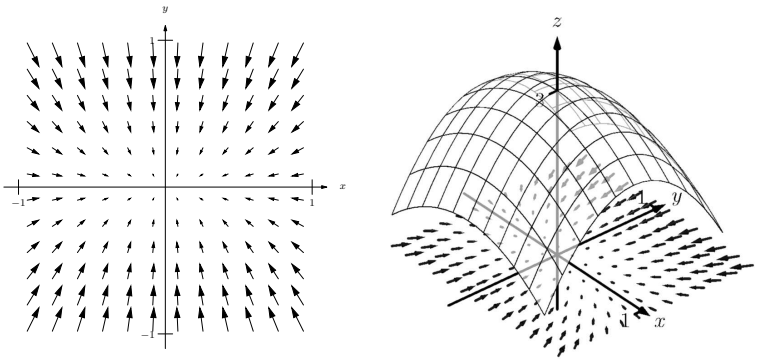
Figure 16.2.19 A vector field representing a planar gravitational force

A function $f(x, y)$ naturally induces a vector field, $\vec{F} = \nabla f = \langle f_x, f_y \rangle$. Given what we learned of the gradient in Section 14.3, we know that the vectors of $\vec{F}$ point in the direction of greatest increase of f. Because of this, f is said to be the *potential function* of $\vec{F}$. Vector fields that are the gradient of potential functions will play an important role in the next section.

Example 16.2.20 A vector field that is the gradient of a potential function.

Let $f(x, y) = 3 - x^2 - 2y^2$ and let $\vec{F} = \nabla f$. Graph $\vec{F}$, and find the divergence and curl of $\vec{F}$.

Solution. Given f, we find $\vec{F} = \nabla f = \langle -2x, -4y \rangle$. A graph of $\vec{F}$ is given in Figure 16.2.21(a). In Figure 16.2.21(b), the vector field is given along with a graph of the surface itself; one can see how each vector is pointing in the direction of "steepest uphill", which, in this case, is not simply just "toward the origin."



(a) **(b)**

Figure 16.2.21 A graph of a function $f(x, y)$ and the vector field $\vec{F} = \nabla f$ in Example 16.2.20

We leave it to the reader to confirm that div $\vec{F} = -6$ and curl $\vec{F} = 0$.

There are some important concepts visited in this section that will be revisited in subsequent sections and again at the very end of this chapter. One is: given a vector field $\vec{F}$, both div $\vec{F}$ and curl $\vec{F}$ are measures of rates of change of $\vec{F}$. The divergence measures how much the field spreads (diverges) at a point, and the curl measures how much the field twists (curls) at a point. Another important concept is this: given $z = f(x, y)$, the gradient ∇f is also a measure of a rate of change of f. We will see how the integrals of these rates of change produce meaningful results.

This section introduces the concept of a vector field. The next section "applies calculus" to vector fields. A common application is this: let $\vec{F}$ be a vector field representing a force (hence it is called a "force field," though this name has a decidedly comic-book feel) and let a particle move along a curve C under the influence of this force. What work is performed by the field on this particle? The solution lies in correctly applying the concepts of line integrals in the context of vector fields.

16.2.3 Exercises

Terms and Concepts

1. Give two quantities that can be represented by a vector field in the plane or in space.

2. In your own words, describe what it means for a vector field to have a negative divergence at a point.

3. In your own words, describe what it means for a vector field to have a negative curl at a point.

4. The divergence of a vector field $\vec{F}$ at a particular point is 0. Does this mean that $\vec{F}$ is incompressible? Why/why not?

Problems

Exercise Group. In the following exercises, sketch the given vector field over the rectangle with opposite corners $(-2,-2)$ and $(2,2)$, sketching one vector for every point with integer coordinates (i.e., at $(0,0)$, $(1,2)$, etc.).

5. $\vec{F} = \langle x, 0\rangle$

6. $\vec{F} = \langle 0, x\rangle$

7. $\vec{F} = \langle 1, -1\rangle$

8. $\vec{F} = \langle y^2, 1\rangle$

Exercise Group. In the following exercises, find the divergence and curl of the given vector field.

9. $\vec{F} = \langle x, y^2\rangle$

10. $\vec{F} = \langle -y^2, x\rangle$

11. $\vec{F} = \langle \cos(xy), \sin(xy)\rangle$

12. $\vec{F} = \left\langle \dfrac{-2x}{(x^2+y^2)^2}, \dfrac{-2y}{(x^2+y^2)^2} \right\rangle$

13. $\vec{F} = \langle x+y, y+z, x+z\rangle$

14. $\vec{F} = \langle x^2+z^2, x^2+y^2, y^2+z^2\rangle$

15. $\vec{F} = \nabla f$, where $f(x,y) = \frac{1}{2}x^2 + \frac{1}{3}y^3$.

16. $\vec{F} = \nabla f$, where $f(x,y) = x^2 y$.

17. $\vec{F} = \nabla f$, where $f(x,y,z) = x^2 y + \sin z$.

18. $\vec{F} = \nabla f$, where $f(x,y,z) = \dfrac{1}{x^2+y^2+z^2}$.

16.3 Line Integrals over Vector Fields

Suppose a particle moves along a curve C under the influence of an electromagnetic force described by a vector field $\vec{F}$. Since a force is inducing motion, work is performed. How can we calculate how much work is performed?

Recall that when moving in a straight line, if $\vec{F}$ represents a constant force and $\vec{d}$ represents the direction and length of travel, then work is simply $W = \vec{F} \cdot \vec{d}$. However, we generally want to be able to calculate work even if $\vec{F}$ is not constant and C is not a straight line.

As we have practiced many times before, we can calculate work by first approximating, then refining our approximation through a limit that leads to integration.

Assume as we did in Section 16.1 that C can be parametrized by the arc length parameter s. Over a short piece of the curve with length ds, the curve is approximately straight and our force is approximately constant. The straight-line direction of this short length of curve is given by $\vec{T}$, the unit tangent vector; let $\vec{d} = \vec{T}\,ds$, which gives the direction and magnitude of a small section of C. Thus work over this small section of C is $\vec{F} \cdot \vec{d} = \vec{F} \cdot \vec{T}\,ds$.

Summing up all the work over these small segments gives an approximation of the work performed. By taking the limit as ds goes to zero, and hence the number of segments approaches infinity, we can obtain the exact amount of work. Following the logic presented at the beginning of this chapter in the Integration Review, we see that

$$W = \int_C \vec{F} \cdot \vec{T}\,ds,$$

a line integral.

This line integral is beautiful in its simplicity, yet is not so useful in making actual computations (largely because the arc length parameter is so difficult to work with). To compute actual work, we need to parametrize C with another parameter t via a vector-valued function $\vec{r}(t)$. As stated in Section 16.1, $ds = \|\vec{r}'(t)\|\,dt$, and recall that $\vec{T} = \vec{r}'(t)/\|\vec{r}'(t)\|$. Thus

$$\begin{aligned} W &= \int_C \vec{F} \cdot \vec{T}\,ds = \int_C \vec{F} \cdot \frac{\vec{r}'(t)}{\|\vec{r}'(t)\|}\,\|\vec{r}'(t)\|\,dt \\ &= \int_C \vec{F} \cdot \vec{r}'(t)\,dt = \int_C \vec{F} \cdot d\vec{r}, \end{aligned} \tag{16.3.1}$$

where the final integral uses the differential $d\vec{r}$ for $\vec{r}'(t)\,dt$.

youtu.be/watch?v=ennr8HI2iFE

Figure 16.3.1 Line integrals of scalar fields, revisited

16.3.1 Evaluating Line Integrals over Vector Fields

These integrals are known as *line integrals over vector fields*. By contrast, the line integrals we dealt with in Section 16.1 are sometimes referred to as *line integrals over scalar fields*. Just as a vector field is defined by a function that returns a vector, a scalar field is a function that returns a scalar, such as $z = f(x, y)$. We waited until now to introduce this terminology so we could contrast the concept with vector fields.

We formally define this line integral, then give examples and applications.

Definition 16.3.2 Line Integral Over A Vector Field.

Let $\vec{F}$ be a vector field with continuous components defined on a smooth curve C, parametrized by $\vec{r}(t)$, and let $\vec{T}$ be the unit tangent vector of

$\vec{r}(t)$. The **line integral** over $\vec{F}$ along C is

$$\int_C \vec{F} \cdot d\vec{r} = \int_C \vec{F} \cdot \vec{T}\, ds.$$

In Definition 16.3.2, note how the dot product $\vec{F} \cdot \vec{T}$ is just a scalar. Therefore, this new line integral is really just a special kind of line integral found in Section 16.1; letting $f(s) = \vec{F}(s) \cdot \vec{T}(s)$, the right-hand side simply becomes $\int_C f(s)\, ds$, and we can use the techniques of that section to evaluate the integral. We combine those techniques, along with parts of Equation (16.3.1), to clearly state how to evaluate a line integral over a vector field in the following Key Idea.

youtu.be/watch?v=fwTHe_y3T1c

Figure 16.3.3 Introducing line integrals of vector fields

Key Idea 16.3.4 Evaluating a Line Integral Over A Vector Field.

Let $\vec{F}$ be a vector field with continuous components defined on a smooth curve C, parametrized by $\vec{r}(t)$, $a \le t \le b$, where $\vec{r}$ is continuously differentiable. Then

$$\int_C \vec{F} \cdot \vec{T}\, ds = \int_C \vec{F} \cdot d\vec{r} = \int_a^b \vec{F}\big(\vec{r}(t)\big) \cdot \vec{r}\,'(t)\, dt.$$

An important concept implicit in this Key Idea: we can use any continuously differentiable parametrization $\vec{r}(t)$ of C that preserves the orientation of C: there isn't a "right" one. In practice, choose one that seems easy to work with.

Notation note: the above Definition and Key Idea implicitly evaluate $\vec{F}$ along the curve C, which is parametrized by $\vec{r}(t)$. For instance, if $\vec{F} = \langle x + y, x - y\rangle$ and $\vec{r}(t) = \langle t^2, \cos t\rangle$, then evaluating $\vec{F}$ along C means substituting the x- and y-components of $\vec{r}(t)$ in for x and y, respectively, in $\vec{F}$. Therefore, along C, $\vec{F} = \langle x + y, x - y\rangle = \big\langle t^2 + \cos t, t^2 - \cos t\big\rangle$. Since we are substituting the *output* of $\vec{r}(t)$ for the *input* of $\vec{F}$, we write this as $\vec{F}\big(\vec{r}(t)\big)$. This is a slight abuse of notation as technically the input of $\vec{F}$ is to be a *point*, not a *vector*, but this shorthand is useful.

We use an example to practice evaluating line integrals over vector fields.

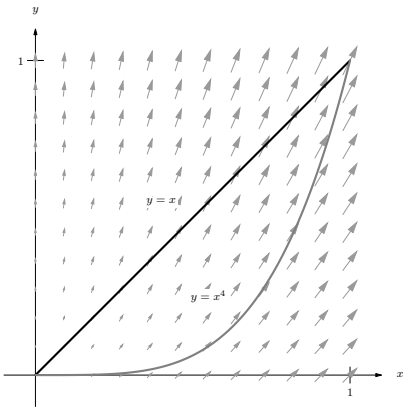


Figure 16.3.6 Paths through a vector field in Example 16.3.5

Example 16.3.5 Evaluating a line integral over a vector field: computing work.

Two particles move from $(0, 0)$ to $(1, 1)$ under the influence of the force field $\vec{F} = \langle x, x + y\rangle$. One particle follows C_1, the line $y = x$; the other follows C_2, the curve $y = x^4$, as shown in Figure 16.3.6. Force is measured in newtons and distance is measured in meters. Find the work performed by each particle.

Solution. To compute work, we need to parametrize each path. We use $\vec{r}_1(t) = \langle t, t\rangle$ to parametrize $y = x$, and let $\vec{r}_2(t) = \langle t, t^4\rangle$ parametrize $y = x^4$; for each, $0 \le t \le 1$.

Along the straight-line path, $\vec{F}\big(\vec{r}_1(t)\big) = \langle x, x + y\rangle = \langle t, t + t\rangle = \langle t, 2t\rangle$. We find $\vec{r}\,'_1(t) = \langle 1, 1\rangle$. The integral that computes work is:

$$\int_{C_1} \vec{F} \cdot d\vec{r} = \int_0^1 \langle t, 2t\rangle \cdot \langle 1, 1\rangle\, dt$$

$$= \int_0^1 3t\,dt$$
$$= \frac{3}{2}t^2\Big|_0^1 = 1.5 \text{ joules}.$$

Along the curve $y = x^4$, $\vec{F}\big(\vec{r}_2(t)\big) = \langle x, x+y\rangle = \big\langle t, t+t^4\big\rangle$. We find $\vec{r}_2'(t) = \big\langle 1, 4t^3\big\rangle$. The work performed along this path is

$$\begin{aligned}\int_{C_2} \vec{F}\cdot d\vec{r} &= \int_0^1 \big\langle t, t+t^4\big\rangle \cdot \big\langle 1, 4t^3\big\rangle\, dt\\ &= \int_0^1 \big(t + 4t^4 + 4t^7\big)\, dt\\ &= \big(\frac{1}{2}t^2 + \frac{4}{5}t^5 + \frac{1}{2}t^8\big)\Big|_0^1 = 1.8 \text{ joules}.\end{aligned}$$

Note how differing amounts of work are performed along the different paths. This should not be too surprising: the force is variable, one path is longer than the other, etc.

Example 16.3.7 Evaluating a line integral over a vector field: computing work.

Two particles move from $(-1,1)$ to $(1,1)$ under the influence of a force field $\vec{F} = \langle y, x\rangle$. One moves along the curve C_1, the parabola defined by $y = 2x^2 - 1$. The other particle moves along the curve C_2, the bottom half of the circle defined by $x^2 + (y-1)^2 = 1$, as shown in Figure 16.3.8. Force is measured in pounds and distances are measured in feet. Find the work performed by moving each particle along its path.

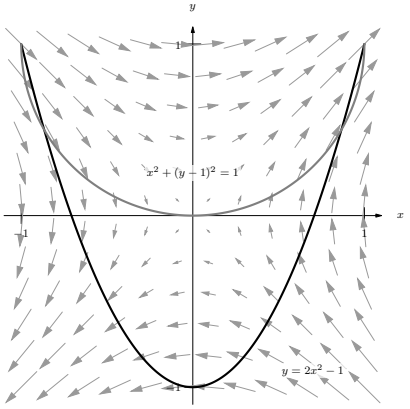


Figure 16.3.8 Paths through a vector field in Example 16.3.7

Solution. We start by parametrizing C_1: the parametrization $\vec{r}_1(t) = \big\langle t, 2t^2 - 1\big\rangle$ is straightforward, giving $\vec{r}_1' = \langle 1, 4t\rangle$. On C_1, $\vec{F}\big(\vec{r}_1(t)\big) = \langle y, x\rangle = \big\langle 2t^2 - 1, t\big\rangle$.

Computing the work along C_1, we have:

$$\begin{aligned}\int_{C_1} \vec{F}\cdot d\vec{r}_1 &= \int_{-1}^1 \big\langle 2t^2 - 1, t\big\rangle \cdot \langle 1, 4t\rangle\, dt\\ &= \int_{-1}^1 \big(2t^2 - 1 + 4t^2\big)\, dt = 2 \text{ ft-lbs}.\end{aligned}$$

For C_2, it is probably simplest to parametrize the half circle using sine and cosine. Recall that $\vec{r}(t) = \langle \cos t, \sin t\rangle$ is a parametrization of the unit circle on $0 \le t \le 2\pi$; we add 1 to the second component to shift the circle up one unit, then restrict the domain to $\pi \le t \le 2\pi$ to obtain only the lower half, giving $\vec{r}_2(t) = \langle \cos t, \sin t + 1\rangle$, $\pi \le t \le 2\pi$, and hence $\vec{r}_2'(t) = \langle -\sin t, \cos t\rangle$ and $\vec{F}\big(\vec{r}_2(t)\big) = \langle y, x\rangle = \langle \sin t + 1, \cos t\rangle$.

Computing the work along C_2, we have:

$$\begin{aligned}\int_{C_2} \vec{F}\cdot d\vec{r}_2 &= \int_{\pi}^{2\pi} \langle \sin t + 1, \cos t\rangle \cdot \langle -\sin t, \cos t\rangle\, dt\\ &= \int_{\pi}^{2\pi} \big(-\sin^2 t - \sin t + \cos^2 t\big)\, dt = 2 \text{ ft-lbs}.\end{aligned}$$

Note how the work along C_1 and C_2 in this example is the same. We'll address why later in this section when *conservative fields* and *path independence* are discussed.

youtu.be/watch?v=JeEIaKjbnJo

Figure 16.3.9 Reivew of terminology for curves

16.3.2 Properties of Line Integrals Over Vector Fields

Line integrals over vector fields share the same properties as line integrals over scalar fields, with one important distinction. The orientation of the curve C matters with line integrals over vector fields, whereas it did not matter with line integrals over scalar fields.

It is relatively easy to see why. Let C be the unit circle. The area under a surface over C is the same whether we traverse the circle in a clockwise or counterclockwise fashion, hence the line integral over a scalar field on C is the same irrespective of orientation. On the other hand, if we are computing work done by a force field, direction of travel definitely matters. Opposite directions create opposite signs when computing dot products, so traversing the circle in opposite directions will create line integrals that differ by a factor of -1.

Theorem 16.3.10 Properties of Line Integrals Over Vector Fields.

1. *Let $\vec{F}$ and $\vec{G}$ be vector fields with continuous components defined on a smooth curve C, parametrized by $\vec{r}(t)$, and let k_1 and k_2 be scalars. Then*
$$\int_C \left(k_1\vec{F} + k_2\vec{G}\right)\cdot d\vec{r} = k_1\int_C \vec{F}\cdot d\vec{r} + k_2\int_C \vec{G}\cdot d\vec{r}.$$
2. *Let C be piecewise smooth, composed of smooth components C_1 and C_2. Then*
$$\int_C \vec{F}\cdot d\vec{r} = \int_{C_1} \vec{F}\cdot d\vec{r} + \int_{C_2} \vec{F}\cdot d\vec{r}.$$
3. *Let C^* be the curve C with opposite orientation, parametrized by $\vec{r}^*$. Then*
$$\int_C \vec{F}\cdot d\vec{r} = -\int_{C^*} \vec{F}\cdot d\vec{r}^*.$$

youtu.be/watch?v=wcKzYvJ20sI

Figure 16.3.11 Properties of line integrals

We demonstrate using these properties in the following example.

Example 16.3.12 Using properties of line integrals over vector fields.

Let $\vec{F} = \langle 3(y-1/2), 1\rangle$ and let C be the path that starts at $(0,0)$, goes to $(1,1)$ along the curve $y = x^3$, then returns to $(0,0)$ along the line $y = x$, as shown in Figure 16.3.13. Evaluate $\oint_C \vec{F}\cdot d\vec{r}$.

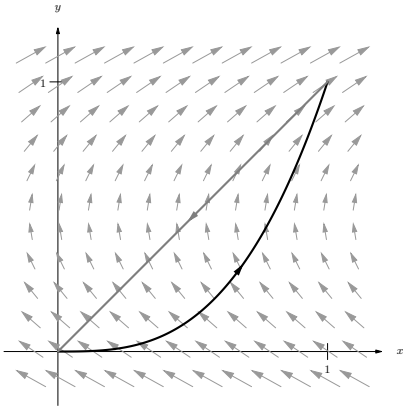


Figure 16.3.13 The vector field and curve in Example 16.3.12

Solution. As C is piecewise smooth, we break it into two components C_1 and C_2, where C_1 follows the curve $y = x^3$ and C_2 follows the curve $y = x$.

We parametrize C_1 with $\vec{r}_1(t) = \langle t, t^3\rangle$ on $0 \le t \le 1$, with $\vec{r}_1'(t) = \langle 1, 3t^2\rangle$. We will use $\vec{F}\big(\vec{r}_1(t)\big) = \langle 3(t^3 - 1/2), 1\rangle$.

While we always have unlimited ways in which to parametrize a curve, there are 2 "direct" methods to choose from when parametrizing C_2.

The parametrization $\vec{r}_2(t) = \langle t, t\rangle, 0 \le t \le 1$ traces the correct line segment but with the wrong orientation. Using Property 3 of Theorem 16.3.10, we can use this parametrization and negate the result.

Another choice is to use the techniques of Section 12.5 to create the line with the orientation we desire. We wish to start at $(1, 1)$ and travel in the $\vec{d} = \langle -1, -1\rangle$ direction for one length of $\vec{d}$, giving equation $\vec{\ell}(t) = \langle 1, 1\rangle + t\langle -1, -1\rangle = \langle 1 - t, 1 - t\rangle$ on $0 \le t \le 1$.

Either choice is fine; we choose $\vec{r}_2(t)$ to practice using line integral properties. We find $\vec{r}_2'(t) = \langle 1, 1\rangle$ and $\vec{F}\big(\vec{r}_2(t)\big) = \langle 3(t - 1/2), 1\rangle$.

Evaluating the line integral (note how we subtract the integral over C_2 as the orientation of $\vec{r}_2(t)$ is opposite):

$$\begin{aligned}\oint_C \vec{F}\cdot d\vec{r} &= \int_{C_1} \vec{F}\cdot d\vec{r}_1 - \int_{C_2} \vec{F}\cdot d\vec{r}_2\\ &= \int_0^1 \big\langle 3(t^3 - 1/2), 1\big\rangle \cdot \big\langle 1, 3t^2\big\rangle\, dt - \int_0^1 \langle 3(t - 1/2), 1\rangle \cdot \langle 1, 1\rangle\, dt\\ &= \int_0^1 \Big(3t^3 + 3t^2 - 3/2\Big)\, dt - \int_0^1 \Big(3t - 1/2\Big)\, dt\\ &= \big(1/4\big) - \big(1\big)\\ &= -3/4.\end{aligned}$$

If we interpret this integral as computing work, the negative work implies that the motion is mostly *against* the direction of the force, which seems plausible when we look at Figure 16.3.13.

youtu.be/watch?v=iOweAD63oCE

Figure 16.3.14 Computing a line integral over a piecewise smooth curve

Example 16.3.15 Evaluating a line integral over a vector field in space.

Let $\vec{F} = \langle -y, x, 1\rangle$, and let C be the portion of the helix given by $\vec{r}(t) = \langle \cos t, \sin t, t/(2\pi)\rangle$ on $[0, 2\pi]$, as shown in Figure 16.3.16. Evaluate $\int_C \vec{F}\cdot d\vec{r}$.

Solution. A parametrization is already given for C, so we just need to find $\vec{F}\big(\vec{r}(t)\big)$ and \vec r '(t).

We have $\vec{F}\big(\vec{r}(t)\big) = \langle -\sin t, \cos t, 1\rangle$ and $\vec{r}\,'(t) = \langle -\sin t, \cos t, 1/(2\pi)\rangle$. Thus

$$\begin{aligned}\int_C \vec{F}\cdot d\vec{r} &= \int_0^{2\pi} \langle -\sin t, \cos t, 1\rangle \cdot \langle -\sin t, \cos t, 1/(2\pi)\rangle\, dt\\ &= \int_0^{2\pi} \Big(\sin^2 t + \cos^2 t + \frac{1}{2\pi}\Big) dt\\ &= 2\pi + 1 \approx 7.28.\end{aligned}$$

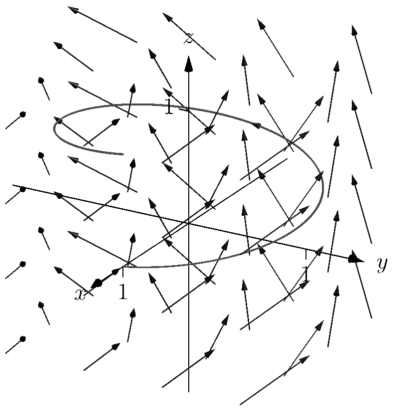


Figure 16.3.16 The graph of $\vec{r}(t)$ in Example 16.3.15

youtu.be/watch?v=NBkXSN614hU

Figure 16.3.17 Evaluating a line integral presented in differential form

16.3.3 The Fundamental Theorem of Line Integrals

youtu.be/watch?v=14LOByUGRXA

Figure 16.3.18 Introducing the Fundamental Theorem of Line Integrals

We are preparing to make important statements about the value of certain line integrals over special vector fields. Before we can do that, we need to define some terms that describe the domains over which a vector field is defined.

A region in the plane is *connected* if any two points in the region can be joined by a piecewise smooth curve that lies entirely in the region. In Figure 16.3.19, sets R_1 and R_2 are connected; set R_3 is not connected, though it is composed of two connected subregions.

A region is *simply connected* if every simple closed curve that lies entirely in the region can be continuously deformed (shrunk) to a single point without leaving the region. (A curve is *simple* if it does not cross itself.) In Figure 16.3.19, only set R_1 is simply connected. Region R_2 is not simply connected as any closed curve that goes around the "hole" in R_2 cannot be continuously shrunk to a single point. As R_3 is not even connected, it cannot be simply connected, though again it consists of two simply connected subregions.

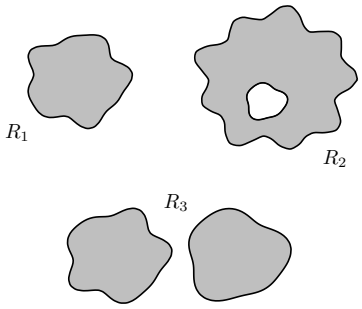
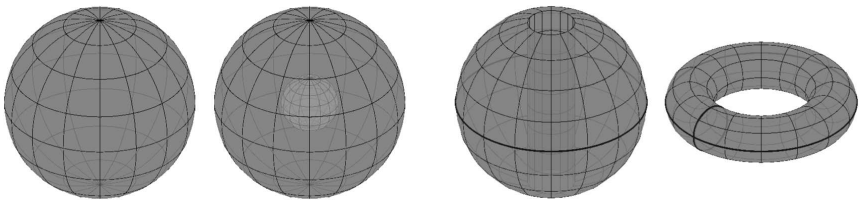


Figure 16.3.19 R_1 is simply connected; R_2 is connected, but not simply connected; R_3 is not connected

We have applied these terms to regions of the plane, but they can be extended intuitively to domains in space (and hyperspace). In Figure 16.3.20(a), the domain bounded by the sphere (at left) and the domain with a subsphere removed (at right) are both simply connected. Any simple closed path that lies entirely within these domains can be continuously deformed into a single point. In Figure 16.3.20(a), neither domain is simply connected. A left, the ball has a hole that extends its length and the pictured closed path cannot be deformed to a point. At right, two paths are illustrated on the torus that cannot be shrunk to a point.

We will use the terms connected and simply connected in subsequent definitions and theorems.



(a) **(b)**

Figure 16.3.20 The domains in (a) are simply connected, while the domains in (b) are not

Recall how in Example 16.3.7 particles moved from $A = (-1, 1)$ to $B = (1, 1)$ along two different paths, wherein the same amount of work was performed along each path. It turns out that regardless of the choice of path from A to B, the amount of work performed under the field $\vec{F} = \langle y, x \rangle$ is the same. Since our expectation is that differing amounts of work are performed along different paths, we give such special fields a name.

Definition 16.3.21 Conservative Field, Path Independent.

Let $\vec{F}$ be a vector field defined on an open, connected domain D in the plane or in space containing points A and B. If the line integral $\int_C \vec{F} \cdot d\vec{r}$ has the same value for all choices of paths C starting at A and ending at B, then

- $\vec{F}$ is a *conservative field* and

- The line integral $\int_C \vec{F} \cdot d\vec{r}$ is *path independent* and can be written as

$$\int_C \vec{F} \cdot d\vec{r} = \int_A^B \vec{F} \cdot \, d\vec{r}.$$

When $\vec{F}$ is a conservative field, the line integral from points A to B is sometimes written as $\int_A^B \vec{F} \cdot d\vec{r}$ to emphasize the independence of its value from the choice of path; all that matters are the beginning and ending points of the path.

How can we tell if a field is conservative? To show a field $\vec{F}$ is conservative using the definition, we need to show that *all* line integrals from points A to B have the same value. It is equivalent to show that *all* line integrals over closed paths C are 0. Each of these tasks are generally nontrivial.

There is a simpler method. Consider the surface defined by $z = f(x, y) = xy$. We can compute the gradient of this function: $\nabla f = \langle f_x, f_y \rangle = \langle y, x \rangle$. Note that this is the field from Example 16.3.7, which we have claimed is conservative. We will soon give a theorem that states that a field $\vec{F}$ is conservative if, and only if, it is the gradient of some scalar function f. To show $\vec{F}$ is conservative, we need to determine whether or not $\vec{F} = \nabla f$ for some function f. (We'll later see that there is a yet simpler method). To recognize the special relationship between $\vec{F}$ and f in this situation, f is given a name.

Definition 16.3.22 Potential Function.

Let f be a differentiable function defined on a domain D in the plane or in space (i.e., $z = f(x, y)$ or $w = f(x, y, z)$) and let $\vec{F} = \nabla f$, the gradient of f. Then f is a **potential function** of $\vec{F}$.

We now state the Fundamental Theorem of Line Integrals, which connects conservative fields and path independence to fields with potential functions.

Theorem 16.3.23 Fundamental Theorem of Line Integrals.

Let $\vec{F}$ be a vector field whose components are continuous on a connected domain D in the plane or in space, let A and B be any points in D, and let C be any path in D starting at A and ending at B.

1. *$\vec{F}$ is conservative if and only if there exists a differentiable function f such that $\vec{F} = \nabla f$.*
2. *If $\vec{F}$ is conservative, then*

$$\int_C \vec{F} \cdot d\vec{r} = \int_A^B \vec{F} \cdot d\vec{r} = f(B) - f(A).$$

Once again considering Example 16.3.7, we have $A = (-1, 1)$, $B = (1, 1)$ and $\vec{F} = \langle y, x \rangle$. In that example, we evaluated two line integrals from A to B and found the value of each was 2. Note that $f(x, y) = xy$ is a potential function for $\vec{F}$. Following the Fundamental Theorem of Line Integrals, consider $f(B) - f(A)$:

$$f(B) - f(A) = f(1, 1) - f(-1, 1) = 1 - (-1) = 2,$$

the same value given by the line integrals.

We practice using this theorem again in the next example.

Example 16.3.24 Using the Fundamental Theorem of Line Integrals.

Let $\vec{F} = \langle 3x^2y + 2x, x^3 + 1\rangle$, $A = (0,1)$ and $B = (1,4)$. Use the first part of the Fundamental Theorem of Line Integrals to show that $\vec{F}$ is conservative, then choose any path from A to B and confirm the second part of the theorem.

Solution. To show $\vec{F}$ is conservative, we need to find $z = f(x,y)$ such that $\vec{F} = \nabla f = \langle f_x, f_y\rangle$. That is, we need to find f such that $f_x = 3x^2y + 2x$ and $f_y = x^3 + 1$. As all we know about f are its partial derivatives, we recover f by integration:

$$\int \frac{\partial f}{\partial x}\, dx = f(x,y) + C(y).$$

Note how the constant of integration is more than "just a constant": it is anything that acts as a constant when taking a derivative with respect to x. Any function that is a function of y (containing no x's) acts as a constant when deriving with respect to x.

Integrating f_x in this example gives:

$$\int \frac{\partial f}{\partial x}\, dx = \int (3x^2y + 2x)\, dx = x^3y + x^2 + C_1(y).$$

Likewise, integrating f_y with respect to y gives:

$$\int \frac{\partial f}{\partial y}\, dy = \int (x^3 + 1)\, dy = x^3y + y + C_2(x).$$

These two results should be equal with appropriate choices of $C_1(y)$ and $C_2(x)$:

$$x^3y + x^2 + C_1(y) = x^3y + y + C_2(x) \Rightarrow C_2(x) = x^2 \text{ and } C_1(y) = y.$$

We find $f(x,y) = x^3y + x^2 + y$, a potential function of $\vec{F}$. (If $\vec{F}$ were not conservative, no choice of $C_2(x)$ and $C_1(y)$ would give equality.)

By the Fundamental Theorem of Line Integrals, regardless of the path from A to B,

$$\begin{aligned}\int_A^B \vec{F}\cdot d\vec{r} &= f(B) - f(A)\\ &= f(1,4) - f(0,1)\\ &= 9 - 1 = 8.\end{aligned}$$

To illustrate the validity of the Fundamental Theorem, we pick a path from A to B. The line between these two points would be simple to construct; we choose a slightly more complicated path by choosing the parabola $y = x^2 + 2x + 1$. This leads to the parametrization $\vec{r}(t) = \langle t, t^2 + 2t + 1\rangle$, $0 \le t \le 1$, with $\vec{r}\,'(t) = \langle t, 2t + 2\rangle$. Thus

$$\begin{aligned}\int_C \vec{F}\cdot d\vec{r} &= \int_C \vec{F}\big(\vec{r}(t)\big)\cdot \vec{r}\,'(t)\, dt\\ &= \int_0^1 \big\langle 3(t)(t^2 + 2t + 1) + 2t, t^3 + 1\big\rangle \cdot \langle t, 2t + 2\rangle\ dt\\ &= \int_0^1 \big(5t^4 + 8t^3 + 3t^2 + 4t + 2\big)\, dt\end{aligned}$$

$$= \left(t^5 + 2t^4 + t^3 + 2t^2 + 2t\right)\Big|_0^1$$
$$= 8,$$

which matches our previous result.

The Fundamental Theorem of Line Integrals states that we can determine whether or not $\vec{F}$ is conservative by determining whether or not $\vec{F}$ has a potential function. This can be difficult. A simpler method exists if the domain of $\vec{F}$ is simply connected (not just connected as needed in the Fundamental Theorem of Line Integrals), which is a reasonable requirement. We state this simpler method as a theorem.

Theorem 16.3.26 Curl of Conservative Fields.

Let $\vec{F}$ be a vector field whose components are continuously differentiable on a simply connected domain D in the plane or in space. Then $\vec{F}$ is conservative if and only if curl $\vec{F} = 0$ *or* $\vec{0}$*, in 2D or 3D, respectively.*

In Example 16.3.24, we showed that $\vec{F} = \langle 3x^2y + 2x, x^3 + 1\rangle$ is conservative by finding a potential function for $\vec{F}$. Using the above theorem, we can show that $\vec{F}$ is conservative much more easily by computing its curl:

$$\text{curl } \vec{F} = N_x - M_y = 3x^2 - 3x^2 = 0.$$

youtu.be/watch?v=urwjnm5PXhQ

Figure 16.3.25 Further examples with the Fundamental Theorem (4 videos)

youtu.be/watch?v=I4Oaf3ArIr8

Figure 16.3.27 Discussing conservative vector fields and path independence

16.3.4 Exercises

Terms and Concepts

1. T/F: In practice, the evaluation of line integrals over vector fields involves computing the magnitude of a vector-valued function.

2. Let $\vec{F}(x, y)$ be a vector field in the plane and let $\vec{r}(t)$ be a two-dimensional vector-valued function. Why is "$\vec{F}\big(\vec{r}(t)\big)$" an "abuse of notation"?

3. T/F: The orientation of a curve C matters when computing a line integral over a vector field.

4. T/F: The orientation of a curve C matters when computing a line integral over a scalar field.

5. Under "reasonable conditions," if curl $\vec{F} = \vec{0}$, what can we conclude about the vector field $\vec{F}$?

6. Let $\vec{F}$ be a conservative field and let C be a closed curve. Why are we able to conclude that $\oint_C \vec{F} \cdot d\vec{r} = 0$?

Problems

Exercise Group. In the following exercises, a vector field $\vec{F}$ and a curve C are given. Evaluate $\displaystyle\int_C \vec{F} \cdot d\vec{r}$.

7. $\vec{F} = \langle y, y^2\rangle$; C is the line segment from $(0, 0)$ to $(3, 1)$.

8. $\vec{F} = \langle x, x + y\rangle$; C is the portion of the parabola $y = x^2$ from $(0, 0)$ to $(1, 1)$.

9. $\vec{F} = \langle y, x\rangle$; C is the top half of the unit circle, beginning at $(1, 0)$ and ending at $(-1, 0)$.

10. $\vec{F} = \langle xy, x\rangle$; C is the portion of the curve $y = x^3$ on $-1 \le x \le 1$.

11. $\vec{F} = \langle z, x^2, y\rangle$; C is the line segment from $(1, 2, 3)$ to $(4, 3, 2)$.

12. $\vec{F} = \langle y + z, x + z, x + y\rangle$; C is the helix $\vec{r}(t) = \langle \cos t, \sin t, t/(2\pi)\rangle$ on $0 \le t \le 2\pi$.

Exercise Group. In the following exercises, find the work performed by the force field $\vec{F}$ moving a particle along the path C.

13. $\vec{F} = \langle y, x^2\rangle$ N; C is the segment of the line $y = x$ from $(0, 0)$ to $(1, 1)$, where distances are measured in meters.

14. $\vec{F} = \langle y, x^2\rangle$ N; C is the portion of $y = \sqrt{x}$ from $(0, 0)$ to $(1, 1)$, where distances are measured in meters.

15. $\vec{F} = \langle 2xy, x^2, 1\rangle$ lbs; C is the path from $(0, 0, 0)$ to $(2, 4, 8)$ via $\vec{r}(t) = \langle t, t^2, t^3\rangle$ on $0 \le t \le 2$, where distance are measured in feet.

16. $\vec{F} = \langle 2xy, x^2, 1\rangle$ lbs; C is the path from $(0, 0, 0)$ to $(2, 4, 8)$ via $\vec{r}(t) = \langle t, 2t, 4t\rangle$ on $0 \le t \le 2$, where distance are measured in feet.

Exercise Group. In the following exercises, a conservative vector field $\vec{F}$ and a curve C are given.

(a) Find a potential function f for $\vec{F}$.

(b) Compute curl $\vec{F}$.

(c) Evaluate $\displaystyle\int_C \vec{F} \cdot d\vec{r}$ directly, i.e., using Key Idea 16.3.4.

(d) Evaluate $\displaystyle\int_C \vec{F} \cdot d\vec{r}$ using the Fundamental Theorem of Line Integrals.

17. $\vec{F} = \langle y + 1, x\rangle$, C is the line segment from $(0, 1)$ to $(1, 0)$.

18. $\vec{F} = \langle 2x + y, 2y + x\rangle$, C is curve parametrized by $\vec{r}(t) = \langle t^2 - t, t^3 - t\rangle$ on $0 \le t \le 1$.

19. $\vec{F} = \langle 2xyz, x^2z, x^2y\rangle$, C is curve parametrized by $\vec{r}(t) = \langle 2t + 1, 3t - 1, t\rangle$ on $0 \le t \le 2$.

20. $\vec{F} = \langle 2x, 2y, 2z\rangle$, C is curve parametrized by $\vec{r}(t) = \langle \cos t, \sin t, \sin(2t)\rangle$ on $0 \le t \le 2\pi$.

21. Prove part of Theorem 16.3.26: let $\vec{F} = \langle M, N, P\rangle$ be a conservative vector field. Show that curl $\vec{F} = 0$.

16.4 Flow, Flux, Green's Theorem and the Divergence Theorem

16.4.1 Flow and Flux

Line integrals over vector fields have the natural interpretation of computing work when $\vec{F}$ represents a force field. It is also common to use vector fields to represent velocities. In these cases, the line integral $\int_C \vec{F}\cdot d\vec{r}$ is said to represent *flow*.

Let the vector field $\vec{F} = \langle 1, 0\rangle$ represent the velocity of water as it moves across a smooth surface, depicted in Figure 16.4.1. A line integral over C will compute "how much water is moving *along* the path C."

In the figure, "all" of the water above C_1 is moving along that curve, whereas "none" of the water above C_2 is moving along that curve (the curve and the flow of water are at right angles to each other). Because C_3 has nonzero horizontal and vertical components, "some" of the water above that curve is moving along the curve.

When C is a closed curve, we call flow *circulation*, represented by $\oint_C \vec{F}\cdot d\vec{r}$.

The "opposite" of flow is *flux*, a measure of "how much water is moving *across* the path C." If a curve represents a filter in flowing water, flux measures how much water will pass through the filter. Considering again Figure 16.4.1, we see that a screen along C_1 will not filter any water as no water passes across that curve. Because of the nature of this field, C_2 and C_3 each filter the same amount of water per second.

The terms "flow" and "flux" are used apart from velocity fields, too. Flow is measured by $\int_C \vec{F}\cdot d\vec{r}$, which is the same as $\int_C \vec{F}\cdot\vec{T}\,ds$ by Definition 16.3.2. That is, flow is a summation of the amount of $\vec{F}$ that is *tangent* to the curve C.

By contrast, flux is a summation of the amount of $\vec{F}$ that is *orthogonal* to the direction of travel. To capture this orthogonal amount of $\vec{F}$, we use $\int_C \vec{F}\cdot\vec{n}\,ds$ to measure flux, where $\vec{n}$ is a unit vector orthogonal to the curve C. (Later, we'll measure flux across surfaces, too. For example, in physics it is useful to measure the amount of a magnetic field that passes through a surface.)

How is $\vec{n}$ determined? We'll later see that if C is a closed curve, we'll want $\vec{n}$ to point to the outside of the curve (measuring how much is "going out"). We'll also adopt the convention that closed curves should be traversed counterclockwise.

(If C is a complicated closed curve, it can be difficult to determine what "counterclockwise" means. Consider Figure 16.4.3. Seeing the curve as a whole, we know which way "counterclockwise" is. If we zoom in on point A, one might incorrectly choose to traverse the path in the wrong direction. So we offer this definition: *a closed curve is being traversed counterclockwise if the outside is to the right of the path and the inside is to the left.*)

When a curve C is traversed counterclockwise by $\vec{r}(t) = \langle f(t), g(t)\rangle$, we rotate $\vec{T}$ clockwise 90° to obtain $\vec{n}$:

$$\vec{T} = \frac{\langle f'(t), g'(t)\rangle}{\|\vec{r}\,'(t)\|} \Rightarrow \vec{n} = \frac{\langle g'(t), -f'(t)\rangle}{\|\vec{r}\,'(t)\|}.$$

Letting $\vec{F} = \langle M, N\rangle$, we calculate flux as:

$$\begin{aligned}\int_C \vec{F}\cdot\vec{n}\,ds &= \int_C \vec{F}\cdot\frac{\langle g'(t), -f'(t)\rangle}{\|\vec{r}\,'(t)\|}\,\|\vec{r}\,'(t)\|\;dt\\ &= \int_C \langle M, N\rangle\cdot\langle g'(t), -f'(t)\rangle\;dt\\ &= \int_C \Big(M\,g'(t) - N\,f'(t)\Big)dt\end{aligned}$$

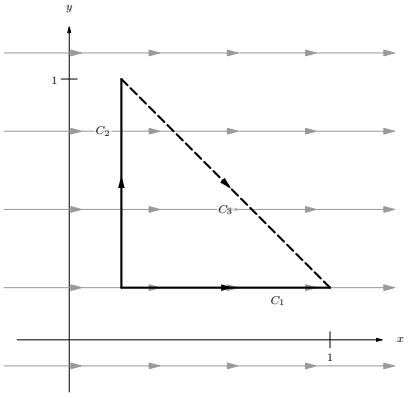


Figure 16.4.1 Illustrating the principles of flow and flux

youtu.be/watch?v=setRFQrIgC4

Figure 16.4.2 Further terminology for curves: closed, simple, positively oriented

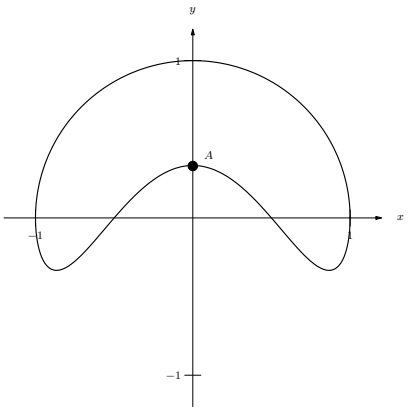


Figure 16.4.3 Determining "counterclockwise" is not always simple without a good definition

$$= \int_C M\, g'(t)\, dt - \int_C N\, f'(t)\, dt.$$

As the x and y components of $\vec{r}(t)$ are $f(t)$ and $g(t)$ respectively, the differentials of x and y are $dx = f'(t)dt$ and $dy = g'(t)dt$. We can then write the above integrals as:

$$= \int_C M\, dy - \int_C N\, dx.$$

This is often written as one integral (not incorrectly, though somewhat confusingly, as this one integral has two "d's"):

$$= \int_C M\, dy - N\, dx.$$

We summarize the above in the following definition.

Definition 16.4.4 Flow, Flux.

Let $\vec{F} = \langle M, N\rangle$ be a vector field with continuous components defined on a smooth curve C, parametrized by $\vec{r}(t) = \langle f(t), g(t)\rangle$, let $\vec{T}$ be the unit tangent vector of $\vec{r}(t)$, and let $\vec{n}$ be the clockwise 90° degree rotation of $\vec{T}$.

- The *flow* of $\vec{F}$ along C is

$$\int_C \vec{F}\cdot\vec{T}\, ds = \int_C \vec{F}\cdot d\vec{r}.$$

- The *flux* of $\vec{F}$ across C is

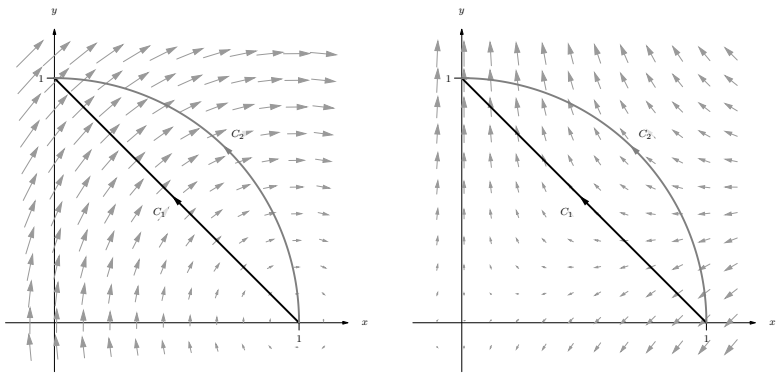
$$\int_C \vec{F}\cdot\vec{n}\, ds = \int_C M\, dy - N\, dx = \int_C \Big(M\, g'(t) - N\, f'(t)\Big)dt.$$

This definition of flow also holds for curves in space, though it does not make sense to measure "flux across a curve" in space.

Measuring flow is essentially the same as finding work performed by a force as done in the previous examples. Therefore we practice finding only flux in the following example.

Example 16.4.5 Finding flux across curves in the plane.

Curves C_1 and C_2 each start at $(1, 0)$ and end at $(0, 1)$, where C_1 follows the line $y = 1 - x$ and C_2 follows the unit circle, as shown in Figure 16.4.6. Find the flux across both curves for the vector fields $\vec{F}_1 = \langle y, -x + 1\rangle$ and $\vec{F}_2 = \langle -x, 2y - x\rangle$.



(a) **(b)**

Figure 16.4.6 Illustrating the curves and vector fields in Example 16.4.5. In (a) the vector field is $\vec{F}_1$, and in (b) the vector field is $\vec{F}_2$.

Solution. We begin by finding parametrizations of C_1 and C_2. As done in Example 16.3.12, parametrize C_1 by creating the line that starts at $(1,0)$ and moves in the $\langle -1,1\rangle$ direction: $\vec{r}_1(t) = \langle 1,0\rangle + t\langle -1,1\rangle = \langle 1-t, t\rangle$, for $0 \le t \le 1$. We parametrize C_2 with the familiar $\vec{r}_2(t) = \langle \cos t, \sin t\rangle$ on $0 \le t \le \pi/2$. For reference later, we give each function and its derivative below:

$$\vec{r}_1(t) = \langle 1-t, t\rangle\,, \vec{r}_1'(t) = \langle -1, 1\rangle\,.$$

$$\vec{r}_2(t) = \langle \cos t, \sin t\rangle\,, \vec{r}_2'(t) = \langle -\sin t, \cos t\rangle\,.$$

When $\vec{F} = \vec{F}_1 = \langle y, -x+1\rangle$ (as shown in Figure 16.4.6(a)), over C_1 we have $M = y = t$ and $N = -x+1 = -(1-t)+1 = t$. Using Definition 16.4.4, we compute the flux:

$$\begin{aligned}\int_{C_1} \vec{F}\cdot\vec{n}\,ds &= \int_{C_1} \Big(M\,g'(t) - N\,f'(t)\Big)\,dt \\ &= \int_0^1 \Big(t(1) - t(-1)\Big)\,dt \\ &= \int_0^1 2t\,dt \\ &= 1.\end{aligned}$$

Over C_2, we have $M = y = \sin t$ and $N = -x+1 = 1-\cos t$. Thus the flux across C_2 is:

$$\begin{aligned}\int_{C_1} \vec{F}\cdot\vec{n}\,ds &= \int_{C_1} \Big(M\,g'(t) - N\,f'(t)\Big)\,dt \\ &= \int_0^{\pi/2} \Big((\sin t)(\cos t) - (1-\cos t)(-\sin t)\Big)\,dt \\ &= \int_0^{\pi/2} \sin t\,dt \\ &= 1.\end{aligned}$$

Notice how the flux was the same across both curves. This won't hold true when we change the vector field.

When $\vec{F} = \vec{F}_2 = \langle -x, 2y - x\rangle$ (as shown in Figure 16.4.6(b)), over C_1 we have $M = -x = t - 1$ and $N = 2y - x = 2t - (1 - t) = 3t - 1$. Computing the flux across C_1:

$$\begin{aligned}\int_{C_1} \vec{F}\cdot\vec{n}\,ds &= \int_{C_1} \Big(M\,g'(t) - N\,f'(t)\Big)\,dt\\ &= \int_0^1 \Big((t-1)(1) - (3t-1)(-1)\Big)\,dt\\ &= \int_0^1 (4t-2)\,dt\\ &= 0.\end{aligned}$$

Over C_2, we have $M = -x = -\cos t$ and $N = 2y - x = 2\sin t - \cos t$. Thus the flux across C_2 is:

$$\begin{aligned}\int_{C_1} \vec{F}\cdot\vec{n}\,ds &= \int_{C_1} \Big(M\,g'(t) - N\,f'(t)\Big)\,dt\\ &= \int_0^{\pi/2} \Big((-\cos t)(\cos t) - (2\sin t - \cos t)(-\sin t)\Big)\,dt\\ &= \int_0^{\pi/2} \big(2\sin^2 t - \sin t\cos t - \cos^2 t\big)\,dt\\ &= \pi/4 - 1/2 \approx 0.285.\end{aligned}$$

We analyze the results of this example below.

In Example 16.4.5, we saw that the flux across the two curves was the same when the vector field was $\vec{F}_1 = \langle y, -x + 1\rangle$. This is not a coincidence. We show why they are equal in Example 16.4.23. In short, the reason is this: the divergence of $\vec{F}_1$ is 0, and when div $\vec{F} = 0$, the flux across any two paths with common beginning and ending points will be the same.

We also saw in the example that the flux across C_1 was 0 when the field was $\vec{F}_2 = \langle -x, 2y - x\rangle$. Flux measures "how much" of the field crosses the path from left to right (following the conventions established before). Positive flux means most of the field is crossing from left to right; negative flux means most of the field is crossing from right to left; zero flux means the same amount crosses from each side. When we consider Figure 16.4.6(b), it seems plausible that the same amount of $\vec{F}_2$ was crossing C_1 from left to right as from right to left.

16.4.2 Green's Theorem

There is an important connection between the circulation around a closed region R and the curl of the vector field inside of R, as well as a connection between the flux across the boundary of R and the divergence of the field inside R. These connections are described by Green's Theorem and the Divergence Theorem, respectively. We'll explore each in turn.

Green's Theorem states "the counterclockwise circulation around a closed region R is equal to the sum of the curls over R."

youtu.be/watch?v=JOFBkUN5Dis

Figure 16.4.7 Introducing Green's Theorem

Theorem 16.4.8 Green's Theorem.

Let R be a closed, bounded region of the plane whose boundary C is composed of finitely many smooth curves, let $\vec{r}(t)$ be a counterclockwise parametrization of C, and let $\vec{F} = \langle M, N\rangle$ where N_x and M_y are continuous over R. Then

$$\oint_C \vec{F}\cdot d\vec{r} = \iint_R \text{curl}\,\vec{F}\,dA.$$

We'll explore Green's Theorem through an example.

Example 16.4.9 Confirming Green's Theorem.

Let $\vec{F} = \langle -y, x^2+1\rangle$ and let R be the region of the plane bounded by the triangle with vertices $(-1,0)$, $(1,0)$ and $(0,2)$, shown in Figure 16.4.10. Verify Green's Theorem; that is, find the circulation of $\vec{F}$ around the boundary of R and show that is equal to the double integral of curl $\vec{F}$ over R.

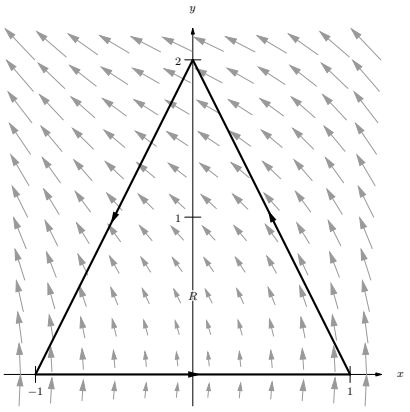


Figure 16.4.10 The vector field and planar region used in Example 16.4.9

Solution. The curve C that bounds R is composed of 3 lines. While we need to traverse the boundary of R in a counterclockwise fashion, we may start anywhere we choose. We arbitrarily choose to start at $(-1,0)$, move to $(1,0)$, etc., with each line parametrized by $\vec{r}_1(t)$, $\vec{r}_2(t)$ and $\vec{r}_3(t)$, respectively.

We leave it to the reader to confirm that the following parametrizations of the three lines are accurate:

$$\begin{aligned}
&\vec{r}_1(t) = \langle 2t-1, 0\rangle, && \text{for } 0\le t\le 1, && \text{with } \vec{r}_1'(t) = \langle 2, 0\rangle,\\
&\vec{r}_2(t) = \langle 1-t, 2t\rangle, && \text{for } 0\le t\le 1, && \text{with } \vec{r}_2'(t) = \langle -1, 2\rangle, \text{ and}\\
&\vec{r}_3(t) = \langle -t, 2-2t\rangle, && \text{for } 0\le t\le 1, && \text{with } \vec{r}_3'(t) = \langle -1, -2\rangle.
\end{aligned}$$

The circulation around C is found by summing the flow along each of the sides of the triangle. We again leave it to the reader to confirm the following computations:

$$\begin{aligned}
\int_{C_1}\vec{F}\cdot d\vec{r}_1 &= \int_0^1 \left\langle 0, (2t-1)^2+1\right\rangle\cdot\langle 2,0\rangle\,dt = 0,\\
\int_{C_2}\vec{F}\cdot d\vec{r}_2 &= \int_0^1 \left\langle -2t, (1-t)^2+1\right\rangle\cdot\langle -1,2\rangle\,dt = 11/3, \text{ and}\\
\int_{C_3}\vec{F}\cdot d\vec{r}_3 &= \int_0^1 \left\langle 2t-2, t^2+1\right\rangle\cdot\langle -1,-2\rangle\,dt = -5/3.
\end{aligned}$$

The circulation is the sum of the flows: 2.

We confirm Green's Theorem by computing $\iint_R \text{curl}\,\vec{F}\,dA$. We find curl $\vec{F} = 2x+1$. The region R is bounded by the lines $y = 2x+2$, $y = -2x+2$ and $y = 0$. Integrating with the order $dx\,dy$ is most straightforward, leading to

$$\int_0^2\int_{y/2-1}^{1-y/2}(2x+1)\,dx\,dy = \int_0^2 (2-y)\,dy = 2,$$

which matches our previous measurement of circulation.

youtu.be/watch?v=K2lPAvWyH4Y

Figure 16.4.11 Verifying Green's Theorem with an example

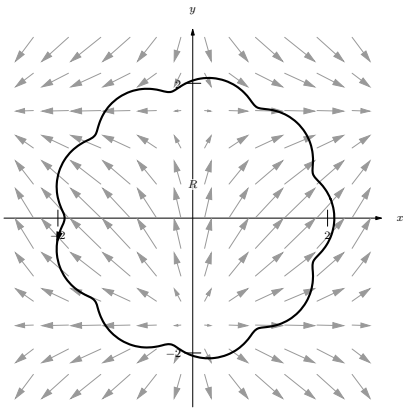


Figure 16.4.13 The vector field and planar region used in Example 16.4.12

youtu.be/watch?v=l9RaOCx4r1g

Figure 16.4.14 Sketching the proof of Green's Theorem

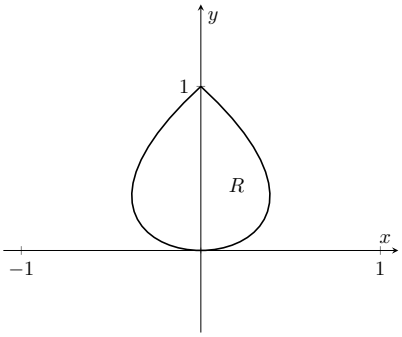


Figure 16.4.16 The region R, whose area is found in Example 16.4.15

youtu.be/watch?v=g8iULyP0alk

Figure 16.4.17 Using Green's Theorem to find the area under a cycloid

Example 16.4.12 Using Green's Theorem.

Let $\vec{F} = \langle \sin x, \cos y \rangle$ and let R be the region enclosed by the curve C parametrized by $\vec{r}(t) = \left\langle 2\cos t + \frac{1}{10}\cos(10t), 2\sin t + \frac{1}{10}\sin(10t) \right\rangle$ on $0 \le t \le 2\pi$, as shown in Figure 16.4.13. Find the circulation around C.

Solution. Computing the circulation directly using the line integral looks difficult, as the integrand will include terms like "$\sin\left(2\cos t + \frac{1}{10}\cos(10t)\right)$."

Green's Theorem states that $\oint_C \vec{F}\cdot d\vec{r} = \iint_R \text{curl}\,\vec{F}\,dA$; since $\text{curl}\,\vec{F} = 0$ in this example, the double integral is simply 0 and hence the circulation is 0.

Since $\text{curl}\,\vec{F} = 0$, we can conclude that the circulation is 0 in two ways. One method is to employ Green's Theorem as done above. The second way is to recognize that $\vec{F}$ is a conservative field, hence there is a function $f(x, y)$ wherein $\vec{F} = \nabla f$. Let A be any point on the curve C; since C is closed, we can say that C "begins" and "ends" at A. By the Fundamental Theorem of Line Integrals, $\oint_C \vec{F}\,d\vec{r} = f(A) - f(A) = 0$.

Since Green's Theorem is an important result, it's worth taking a minute (or 12) to see why it's true, in the video in Figure 16.4.14.

One can use Green's Theorem to find the area of an enclosed region by integrating along its boundary. Let C be a closed curve, enclosing the region R, parametrized by $\vec{r}(t) = \langle f(t), g(t)\rangle$. We know the area of R is computed by the double integral $\iint_R dA$, where the integrand is 1. By creating a field $\vec{F}$ where $\text{curl}\,\vec{F} = 1$, we can employ Green's Theorem to compute the area of R as $\oint_C \vec{F}\cdot d\vec{r}$.

One is free to choose any field $\vec{F}$ to use as long as $\text{curl}\,\vec{F} = 1$. Common choices are $\vec{F} = \langle 0, x\rangle$, $\vec{F} = \langle -y, 0\rangle$ and $\vec{F} = \langle -y/2, x/2\rangle$. We demonstrate this below.

Example 16.4.15 Using Green's Theorem to find area.

Let C be the closed curve parametrized by $\vec{r}(t) = \left\langle t - t^3, t^2 \right\rangle$ on $-1 \le t \le 1$, enclosing the region R, as shown in Figure 16.4.16. Find the area of R.

Solution. We can choose any field $\vec{F}$, as long as $\text{curl}\,\vec{F} = 1$. We choose $\vec{F} = \langle -y, 0\rangle$. We also confirm (left to the reader) that $\vec{r}(t)$ traverses the region R in a counterclockwise fashion. Thus

$$\begin{aligned}
\text{Area of } R &= \iint_R dA \\
&= \oint_C \vec{F}\cdot d\vec{r} \\
&= \int_{-1}^{1} \left\langle -t^2, 0\right\rangle \cdot \left\langle 1 - 3t^2, 2t\right\rangle \, dt \\
&= \int_{-1}^{1} (-t^2)(1 - 3t^2)\, dt \\
&= \frac{8}{15}.
\end{aligned}$$

Another interesting scenario that comes up is the case of multiply-connected regions (as opposed to simply-connected). If a bounded region has a "hole", its boundary will consist of more than one curve: the outer boundary, as well as the boundary of the hole. Green's Theorem applies in this situation as well, as the video in Figure 16.4.19 explains.

youtu.be/watch?v=x1bIGtNu-0M

Figure 16.4.18 Using Green's Theorem to derive a polygon area formula

16.4.3 The Divergence Theorem

Green's Theorem makes a connection between the circulation around a closed region R and the sum of the curls over R. The Divergence Theorem makes a somewhat "opposite" connection: the total flux across the boundary of R is equal to the sum of the divergences over R.

Theorem 16.4.20 The Divergence Theorem (in the plane).

Let R be a closed, bounded region of the plane whose boundary C is composed of finitely many smooth curves, let $\vec{r}(t)$ be a counterclockwise parametrization of C, and let $\vec{F} = \langle M, N\rangle$ where M_x and N_y are continuous over R. Then

$$\oint_C \vec{F}\cdot\vec{n}\,ds = \iint_R \text{div}\,\vec{F}\,dA.$$

youtu.be/watch?v=PaW_ZF-js-M

Figure 16.4.19 Green's Theorem and multiply-connected regions

Example 16.4.21 Confirming the Divergence Theorem.

Let $\vec{F} = \langle x-y, x+y\rangle$, let C be the circle of radius 2 centered at the origin and define R to be the interior of that circle, as shown in Figure 16.4.22. Verify the Divergence Theorem; that is, find the flux across C and show it is equal to the double integral of div $\vec{F}$ over R.

Solution. We parametrize the circle in the usual way, with $\vec{r}(t) = \langle 2\cos t, 2\sin t\rangle$, $0 \le t \le 2\pi$. The flux across C is

$$\begin{aligned}\oint_C \vec{F}\cdot\vec{n}\,ds &= \oint_C \big(Mg'(t) - Nf'(t)\big)\,dt \\ &= \int_0^{2\pi} \big((2\cos t - 2\sin t)(2\cos t) - (2\cos t + 2\sin t)(-2\sin t)\big)\,dt \\ &= \int_0^{2\pi} 4\,dt = 8\pi.\end{aligned}$$

We compute the divergence of $\vec{F}$ as div $\vec{F} = M_x + N_y = 2$. Since the divergence is constant, we can compute the following double integral easily:

$$\iint_R \text{div}\,\vec{F}\,dA = \iint_R 2\,dA = 2\iint_R dA = 2(\text{ area of } R\,) = 8\pi,$$

which matches our previous result.

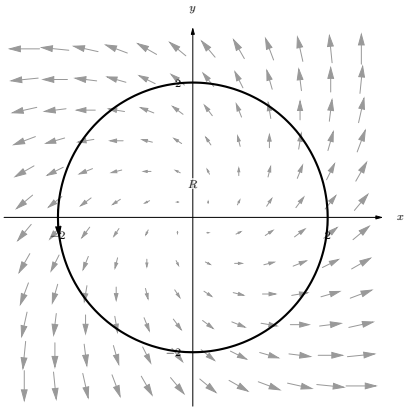


Figure 16.4.22 The region R used in Example 16.4.21

Example 16.4.23 Flux when div $\vec{F} = 0$**.**

Let $\vec{F}$ be any field where div $\vec{F} = 0$, and let C_1 and C_2 be any two non-intersecting paths, except that each begin at point A and end at point B

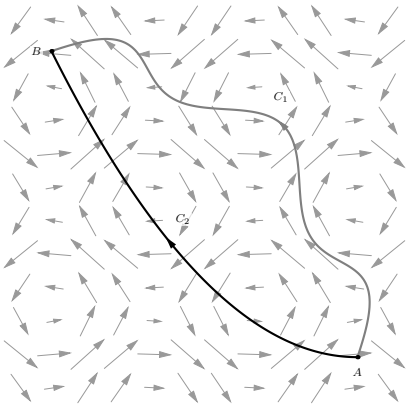


Figure 16.4.24 As used in Example 16.4.23, the vector field has a divergence of 0 and the two paths only intersect at their initial and terminal points.

(see Figure 16.4.24). Show why the flux across C_1 and C_2 is the same.

Solution. By referencing Figure 16.4.24, we see we can make a closed path C that combines C_1 with C_2, where C_2 is traversed with its opposite orientation. We label the enclosed region R. Since div $\vec{F} = 0$, the Divergence Theorem states that

$$\oint_C \vec{F} \cdot \vec{n}\, ds = \iint_R \text{div}\, \vec{F}\, dA = \iint_R 0\, dA = 0.$$

Using the properties and notation given in Theorem 16.3.10, consider:

$$\begin{aligned} 0 &= \oint_C \vec{F} \cdot \vec{n}\, ds \\ &= \int_{C_1} \vec{F} \cdot \vec{n}\, ds + \int_{C_2^*} \vec{F} \cdot \vec{n}\, ds \end{aligned}$$

(where C_2^* is the path C_2 traversed with opposite orientation)

$$\begin{aligned} &= \int_{C_1} \vec{F} \cdot \vec{n}\, ds - \int_{C_2} \vec{F} \cdot \vec{n}\, ds. \\ \int_{C_2} \vec{F} \cdot \vec{n}\, ds &= \int_{C_1} \vec{F} \cdot \vec{n}\, ds. \end{aligned}$$

Thus the flux across each path is equal.

In this section, we have investigated flow and flux, quantities that measure interactions between a vector field and a planar curve. We can also measure flow along spatial curves, though as mentioned before, it does not make sense to measure flux across spatial curves.

It does, however, make sense to measure the amount of a vector field that passes across a surface in space — i.e, the flux across a surface. We will study this, though in the next section we first learn about a more powerful way to describe surfaces than using functions of the form $z = f(x, y)$.

16.4.4 Exercises

Terms and Concepts

1. Let $\vec{F}$ be a vector field and let C be a curve. *Flow* is a measure of the amount of $\vec{F}$ going _____ C; *flux* is a measure of the amount of $\vec{F}$ going _____ C.

2. What is circulation?

3. Green's Theorem states, informally, that the circulation around a closed curve that bounds a region R is equal to the sum of _____ across R.

4. The Divergence Theorem states, informally, that the outward flux across a closed curve that bounds a region R is equal to the sum of _____ across R.

5. Let $\vec{F}$ be a vector field and let C_1 and C_2 be any nonintersecting paths except that each starts at point A and ends at point B. If _____$= 0$, then $\int_{C_1} \vec{F} \cdot \vec{T}\, ds = \int_{C_2} \vec{F} \cdot \vec{T}\, ds$.

6. Let $\vec{F}$ be a vector field and let C_1 and C_2 be any nonintersecting paths except that each starts at point A and ends at point B. If _____$= 0$, then $\int_{C_1} \vec{F} \cdot \vec{n}\, ds = \int_{C_2} \vec{F} \cdot \vec{n}\, ds$.

Problems

Exercise Group. In the following exercises, a vector field $\vec{F}$ and a curve C are given. Evaluate $\int_C \vec{F} \cdot \vec{n}\, ds$, the flux of $\vec{F}$ over C.

7. $\vec{F} = \langle x + y, x - y\rangle$; C is the curve with initial and terminal points $(3, -2)$ and $(3, 2)$, respectively, parametrized by $\vec{r}(t) = \langle 3t^2, 2t\rangle$ on $-1 \le t \le 1$.

8. $\vec{F} = \langle x + y, x - y\rangle$; C is the curve with initial and terminal points $(3, -2)$ and $(3, 2)$, respectively, parametrized by $\vec{r}(t) = \langle 3, t\rangle$ on $-2 \le t \le 2$.

9. $\vec{F} = \langle x^2, y + 1\rangle$; C is line segment from $(0, 0)$ to $(2, 4)$.

10. $\vec{F} = \langle x^2, y + 1\rangle$; C is the portion of the parabola $y = x^2$ from $(0, 0)$ to $(2, 4)$.

11. $\vec{F} = \langle y, 0\rangle$; C is the line segment from $(0, 0)$ to $(0, 1)$.

12. $\vec{F} = \langle y, 0\rangle$; C is the line segment from $(0, 0)$ to $(1, 1)$.

Exercise Group. In the following exercises, a vector field $\vec{F}$ and a closed curve C, enclosing a region R, are given. Verify Green's Theorem by evaluating $\oint_C \vec{F} \cdot d\vec{r}$ and $\iint_R \text{curl}\, \vec{F}\, dA$, showing they are equal.

13. $\vec{F} = \langle x - y, x + y\rangle$; C is the closed curve composed of the parabola $y = x^2$ on $0 \le x \le 2$ followed by the line segment from $(2, 4)$ to $(0, 0)$.

14. $\vec{F} = \langle -y, x\rangle$; C is the unit circle.

15. $\vec{F} = \langle 0, x^2\rangle$; C the triangle with corners at $(0, 0)$, $(2, 0)$ and $(1, 1)$.

16. $\vec{F} = \langle x + y, 2x\rangle$; C the curve that starts at $(0, 1)$, follows the parabola $y = (x - 1)^2$ to $(3, 4)$, then follows a line back to $(0, 1)$.

Exercise Group. In the following exercises, a closed curve C enclosing a region R is given. Find the area of R by computing $\oint_C \vec{F} \cdot d\vec{r}$ for an appropriate choice of vector field $\vec{F}$.

17. C is the ellipse parametrized by $\vec{r}(t) = \langle 4\cos t, 3\sin t\rangle$ on $0 \le t \le 2\pi$.

18. C is the curve parametrized by $\vec{r}(t) = \langle \cos t, \sin(2t)\rangle$ on $-\pi/2 \le t \le \pi/2$.

19. C is the curve parametrized by $\vec{r}(t) = \langle -t^3 + 3t^2 - 2t, 2(t - 1)^2\rangle$ on $0 \le t \le 2$.

20. C is the curve parametrized by $\vec{r}(t) = \langle 2\cos t + \frac{1}{10}\cos(10t), 2\sin t + \frac{1}{10}\sin(10t)\rangle$ on $0 \le t \le 2\pi$.

Exercise Group. In the following exercises, a vector field $\vec{F}$ and a closed curve C, enclosing a region R, are given. Verify the Divergence Theorem by evaluating $\oint_C \vec{F} \cdot \vec{n}\, ds$ and $\iint_R \text{div}\, \vec{F}\, dA$, showing they are equal.

21. $\vec{F} = \langle x - y, x + y \rangle$; C is the closed curve composed of the parabola $y = x^2$ on $0 \leq x \leq 2$ followed by the line segment from $(2, 4)$ to $(0, 0)$.

22. $\vec{F} = \langle -y, x \rangle$; C is the unit circle.

23. $\vec{F} = \langle 0, y^2 \rangle$; C the triangle with corners at $(0, 0)$, $(2, 0)$ and $(1, 1)$.

24. $\vec{F} = \langle x^2/2, y^2/2 \rangle$; C the curve that starts at $(0, 1)$, follows the parabola $y = (x - 1)^2$ to $(3, 4)$, then follows a line back to $(0, 1)$.

16.5 Parametrized Surfaces and Surface Area

Thus far we have focused mostly on 2-dimensional vector fields, measuring flow and flux along/across curves in the plane. Both Green's Theorem and the Divergence Theorem make connections between planar regions and their boundaries. We now move our attention to 3-dimensional vector fields, considering both curves and surfaces in space.

youtu.be/watch?v=dPB3NGtuEhU

Figure 16.5.1 Introducing parametrized surfaces

16.5.1 Parametrizing surfaces

We are accustomed to describing surfaces as functions of two variables, usually written as $z = f(x, y)$. For our coming needs, this method of describing surfaces will prove to be insufficient. Instead, we will *parametrize* our surfaces, describing them as the set of terminal points of some vector-valued function $\vec{r}(u, v) = \langle f(u, v), g(u, v), h(u, v)\rangle$. The bulk of this section is spent practicing the skill of describing a surface $\mathcal{S}$using a vector-valued function. Once this skill is developed, we'll show how to find the surface area S of a parametrically-defined surface $\mathcal{S}$, a skill needed in the remaining sections of this chapter.

We use the letter S to denote Surface Area. This section begins a study into surfaces, and it is natural to label a surface with the letter "S". We distinguish a surface from its surface area by using a calligraphic S to denote a surface: $\mathcal{S}$. When writing this letter by hand, it may be useful to add serifs to the letter, such as:

$\mathcal{S}$

Definition 16.5.2 Parametrized Surface.

Let $\vec{r}(u, v) = \langle f(u, v), g(u, v), h(u, v)\rangle$ be a vector-valued function that is continuous and one to one on the interior of its domain R in the u-v plane. The set of all terminal points of $\vec{r}$ (i.e., the **range** of $\vec{r}$) is the **surface** $\mathcal{S}$, and $\vec{r}$ along with its domain R form a **parametrization** of $\mathcal{S}$.

This parametrization is **smooth** on R if $\vec{r}_u$ and $\vec{r}_v$ are continuous and $\vec{r}_u \times \vec{r}_v$ is never $\vec{0}$ on the interior of R.

Recall that function is *one to one* on its domain if the function never repeats an output value over the domain. In the case of $\vec{r}(u, v)$, $\vec{r}$ is one to one if $\vec{r}(u_1, v_1) \neq \vec{r}(u_2, v_2)$ for all points $(u_1, v_1) \neq (u_2, v_2)$ in the domain of $\vec{r}$.

Given a point (u_0, v_0) in the domain of a vector-valued function $\vec{r}$, the vectors $\vec{r}_u(u_0, v_0)$ and $\vec{r}_v(u_0, v_0)$ are tangent to the surface $\mathcal{S}$ at $\vec{r}(u_0, v_0)$ (a proof of this is developed later in this section). The definition of smoothness dictates that $\vec{r}_u \times \vec{r}_v \neq \vec{0}$; this ensures that neither $\vec{r}_u$ nor $\vec{r}_v$ are $\vec{0}$, nor are they ever parallel. Therefore smoothness guarantees that $\vec{r}_u$ and $\vec{r}_v$ determine a plane that is tangent to $\mathcal{S}$.

A surface $\mathcal{S}$ is said to be *orientable* if a field of normal vectors can be defined on $\mathcal{S}$ that vary continuously along $\mathcal{S}$. This definition may be hard to understand; it may help to know that orientable surfaces are often called "two sided." A sphere is an orientable surface, and one can easily envision an "inside" and "outside" of the sphere. A paraboloid is orientable, where again one can generally envision "inside" and "outside" sides (or "top" and "bottom" sides) to this surface. Just about every surface that one can imagine is orientable, and we'll assume all surfaces we deal with in this text are orientable.

It is enlightening to examine a classic non-orientable surface: the Möbius band, shown in Figure 16.5.3. Vectors normal to the surface are given, starting at the point indicated in the figure. These normal vectors "vary continuously" as they move along the surface. Letting each vector indicate the "top" side of the band, we can easily see near any vector which side is the "top".

However, if as we progress along the band, we recognize that we are labeling "both sides" of the band as the top; in fact, there are not two "sides" to this band, but one. The Möbius band is a non-orientable surface.

We now practice parametrizing surfaces.

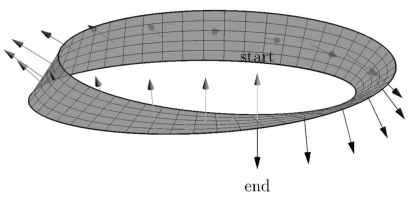


Figure 16.5.3 A Möbius band, a non-orientable surface

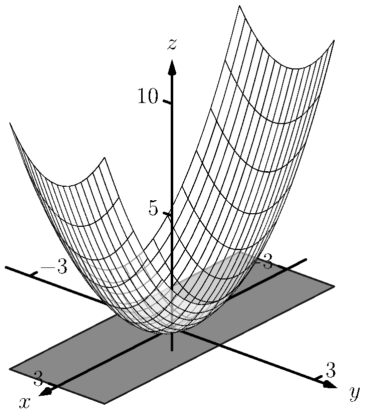


Figure 16.5.5 The surface parametrized in Example 16.5.4

Example 16.5.4 Parametrizing a surface over a rectangle.

Parametrize the surface $z = x^2 + 2y^2$ over the rectangular region R defined by $-3 \le x \le 3$, $-1 \le y \le 1$.

Solution. There is a straightforward way to parametrize a surface of the form $z = f(x, y)$ over a rectangular domain. We let $x = u$ and $y = v$, and let $\vec{r}(u, v) = \langle u, v, f(u, v)\rangle$. In this instance, we have $\vec{r}(u, v) = \langle u, v, u^2 + 2v^2\rangle$, for $-3 \le u \le 3$, $-1 \le v \le 1$. This surface is graphed in Figure 16.5.5.

Example 16.5.6 Parametrizing a surface over a circular disk.

Parametrize the surface $z = x^2+2y^2$ over the circular region R enclosed by the circle of radius 2 that is centered at the origin.

Solution. We can parametrize the circular boundary of R with the vector-valued function $\langle 2\cos u, 2\sin u\rangle$, where $0 \le u \le 2\pi$. We can obtain the interior of R by scaling this function by a variable amount, i.e., by multiplying by v: $\langle 2v\cos u, 2v\sin u\rangle$, where $0 \le v \le 1$.

It is important to understand the role of v in the above function. When $v = 1$, we get the boundary of R, a circle of radius 2. When $v = 0$, we simply get the point $(0, 0)$, the center of R (which can be thought of as a circle with radius of 0). When $v = 1/2$, we get the circle of radius 1 that is centered at the origin, which is the circle *halfway* between the boundary and the center. As v varies from 0 to 1, we create a series of concentric circles that fill out all of R.

Thus far, we have determined the x and y components of our parametrization of the surface: $x = 2v\cos u$ and $y = 2v\sin u$. We find the z component simply by using $z = f(x, y) = x^2 + 2y^2$:

$$z = (2v\cos u)^2 + 2(2v\sin u)^2 = 4v^2\cos^2 u + 8v^2\sin^2 u.$$

Thus $\vec{r}(u, v) = \langle 2v\cos u, 2v\sin u, 4v^2\cos^2 u + 8v^2\sin^2 u\rangle$, $0 \le u \le 2\pi$, $0 \le v \le 1$, which is graphed in Figure 16.5.7. The way that this graphic was generated highlights how the surface was parametrized. When viewing from above, one can see lines emanating from the origin; they represent different values of u as u sweeps from an angle of 0 up to 2π. One can also see concentric circles, each corresponding to a different value of v.

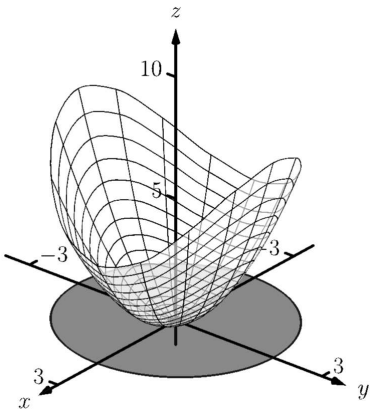


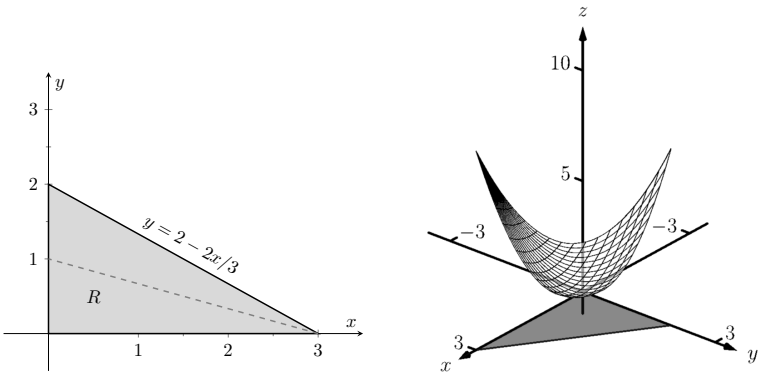
Figure 16.5.7 The surface parametrized in Example 16.5.6

Examples 16.5.4 and Example 16.5.6 demonstrate an important principle when parametrizing surfaces given in the form $z = f(x, y)$ over a region R: if one can determine x and y in terms of u and v, then z follows directly as $z = f(x, y)$.

In the following two examples, we parametrize the same surface over triangular regions. Each will use v as a "scaling factor" as done in Example 16.5.6.

Example 16.5.8 Parametrizing a surface over a triangle.

Parametrize the surface $z = x^2 + 2y^2$ over the triangular region R enclosed by the coordinate axes and the line $y = 2 - 2x/3$, as shown in Figure 16.5.9(a).



(a) **(b)**

Figure 16.5.9 Part (a) shows a graph of the region R, and part (b) shows the surface over R, as defined in Example 16.5.8

Solution. We may begin by letting $x = u, 0 \le u \le 3$, and $y = 2-2u/3$. This gives only the line on the "upper" side of the triangle. To get all of the region R, we can once again scale y by a variable factor, v.

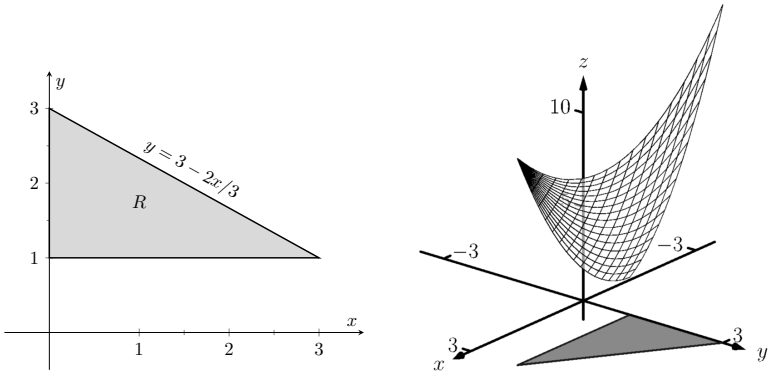
Still letting $x = u$, $0 \le u \le 3$, we let $y = v(2 - 2u/3)$, $0 \le v \le 1$. When $v = 0$, all y-values are 0, and we get the portion of the x-axis between $x = 0$ and $x = 3$. When $v = 1$, we get the upper side of the triangle. When $v = 1/2$, we get the line $y = 1/2(2 - 2u/3) = 1 - u/3$, which is the line "halfway up" the triangle, shown in the figure with a dashed line.

Letting $z = f(x, y) = x^2 + 2y^2$, we have $\vec{r}(u, v) = \langle u, v(2 - 2u/3), u^2 + 2\big(v(2 - 2u/3)\big)^2\rangle$, $0 \le u \le 3$, $0 \le v \le 1$. This surface is graphed in Figure 16.5.9(b). Again, when one looks from above, we can see the scaling effects of v: the series of lines that run to the point $(3, 0)$ each represent a different value of v.

Another common way to parametrize the surface is to begin with $y = u$, $0 \le u \le 2$. Solving the equation of the line $y = 2 - 2x/3$ for x, we have $x = 3 - 3y/2$, leading to using $x = v(3 - 3u/2)$, $0 \le v \le 1$. With $z = x^2+2y^2$, we have $\vec{r}(u, v) = \langle v(3-3u/2), u, \big(v(3-3u/2)\big)^2+2v^2\rangle$, $0 \le u \le 2, 0 \le v \le 1$.

Example 16.5.10 Parametrizing a surface over a triangle.

Parametrize the surface $z = x^2 + 2y^2$ over the triangular region R enclosed by the lines $y = 3 - 2x/3$, $y = 1$ and $x = 0$ as shown in Figure 16.5.11(a).



(a) **(b)**

Figure 16.5.11 Part (a) shows a graph of the region R, and part (b) shows the surface over R, as defined in Example 16.5.10

Solution. While the region R in this example is very similar to the region R in the previous example, and our method of parametrizing the surface is fundamentally the same, it will feel as though our answer is much different than before.

We begin with letting $x = u, 0 \le u \le 3$. We may be tempted to let $y = v(3 - 2u/3), 0 \le v \le 1$, but this is incorrect. When $v = 1$, we obtain the upper line of the triangle as desired. However, when $v = 0$, the y-value is 0, which does not lie in the region R.

We will describe the general method of proceeding following this example. For now, consider $y = 1 + v(2 - 2u/3), 0 \le v \le 1$. Note that when $v = 1$, we have $y = 3 - 2u/3$, the upper line of the boundary of R. Also, when $v = 0$, we have $y = 1$, which is the lower boundary of R. With $z = x^2 + 2y^2$, we determine $\vec{r}(u, v) = \langle u, 1 + v(2 - 2u/3), u^2 + 2\big(1 + v(2 - 2u/3)\big)^2\rangle, 0 \le u \le 3, 0 \le v \le 1$.

The surface is graphed in Figure 16.5.11(b).

Given a surface of the form $z = f(x, y)$, one can often determine a parametrization of the surface over a region R in a manner similar to determining bounds of integration over a region R. Using the techniques of Section 15.1, suppose a region R can be described by $a \le x \le b$, $g_1(x) \le y \le g_2(x)$, i.e., the area of R can be found using the iterated integral

$$\int_a^b \int_{g_1(x)}^{g_2(x)} dy\, dx.$$

When parametrizing the surface, we can let $x = u$, $a \le u \le b$, and we can let $y = g_1(u) + v\big(g_2(u) - g_1(u)\big)$, $0 \le v \le 1$. The parametrization of x is straightforward, but look closely at how y is determined. When $v = 0$, $y = g_1(u) = g_1(x)$. When $v = 1$, $y = g_2(u) = g_2(x)$.

As a specific example, consider the triangular region R from Example 16.5.10, shown in Figure 16.5.11(a). Using the techniques of Section 15.1, we can find

the area of R as

$$\int_0^3 \int_1^{3-2x/3} dy\,dx.$$

Following the above discussion, we can set $x = u$, where $0 \le u \le 3$, and set $y = 1 + v\big(3 - 2u/3 - 1\big) = 1 + v(2 - 2u/3)$, $0 \le v \le 1$, as used in that example.

One can do a similar thing if R is bounded by $c \le y \le d, h_1(y) \le x \le h_2(y)$, but for the sake of simplicity we leave it to the reader to flesh out those details. The principles outlined above are given in the following Key Idea for reference.

Key Idea 16.5.12 Parametrizing Surfaces.

Let a surface $\mathcal{S}$ be the graph of a function $f(x, y)$, where the domain of f is a closed, bounded region R in the xy-plane. Let R be bounded by $a \le x \le b, g_1(x) \le y \le g_2(x)$, i.e., the area of R can be found using the iterated integral $\int_a^b \int_{g_1(x)}^{g_2(x)} dy\,dx$, and let $h(u, v) = g_1(u) + v\big(g_2(u) - g_1(u)\big)$.

$\mathcal{S}$ can be parametrized as

$$\vec{r}(u, v) = \big\langle u, h(u, v), f\big(u, h(u, v)\big)\big\rangle, a \le u \le b,\ 0 \le v \le 1.$$

Example 16.5.13 Parametrizing a cylindrical surface.

Find a parametrization of the cylinder $x^2+z^2/4 = 1$, where $-1 \le y \le 2$, as shown in Figure 16.5.14.

Solution. The equation $x^2+z^2/4 = 1$ can be envisioned to describe an ellipse in the xz-plane; as the equation lacks a y-term, the equation describes a cylinder (recall Definition 12.1.18) that extends without bound parallel to the y-axis. This ellipse has a vertical major axis of length 4, a horizontal minor axis of length 2, and is centered at the origin. We can parametrize this ellipse using sines and cosines; our parametrization can begin with

$$\vec{r}(u, v) = \langle \cos u,\ ???\ , 2\sin u\rangle, 0 \le u \le 2\pi,$$

where we still need to determine the y component.

While the cylinder $x^2 + z^2/4 = 1$ is satisfied by any y value, the problem states that all y values are to be between $y = -1$ and $y = 2$. Since the value of y does not depend at all on the values of x or z, we can use another variable, v, to describe y. Our final answer is

$$\vec{r}(u, v) = \langle \cos u, v, 2\sin u\rangle, 0 \le u \le 2\pi, -1 \le v \le 2.$$

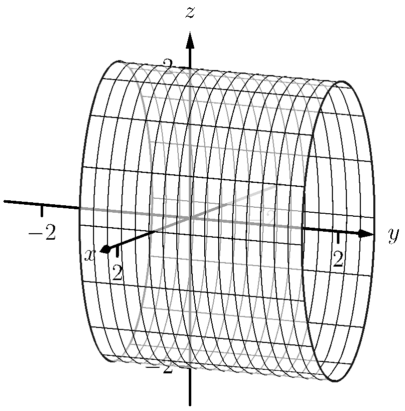


Figure 16.5.14 The cylinder parametrized in Example 16.5.13

Example 16.5.15 Parametrizing an elliptic cone.

Find a parametrization of the elliptic cone $z^2 = \frac{x^2}{4} + \frac{y^2}{9}$, where $-2 \le z \le 3$, as shown in Figure 16.5.16.

Solution. One way to parametrize this cone is to recognize that given a z value, the cross section of the cone at that z value is an ellipse with equation $\frac{x^2}{(2z)^2} + \frac{y^2}{(3z)^2} = 1$. We can let $z = v$, for $-2 \le v \le 3$ and then parametrize the above ellipses using sines, cosines and v.

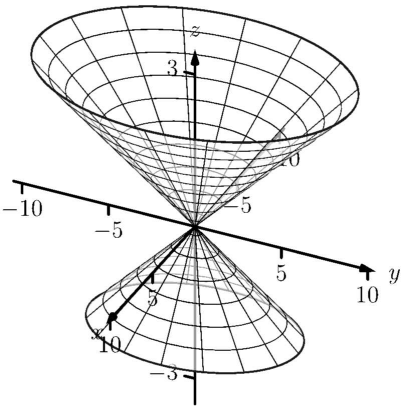


Figure 16.5.16 The elliptic cone as described in Example 16.5.15

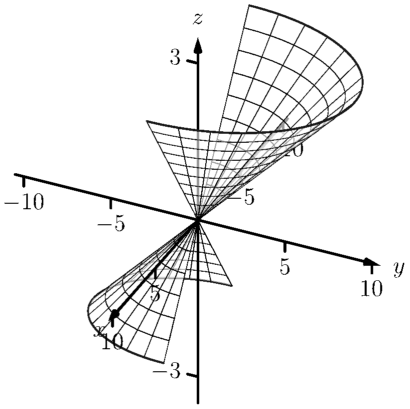


Figure 16.5.17 The elliptic cone as described in Example 16.5.15 with restricted domain

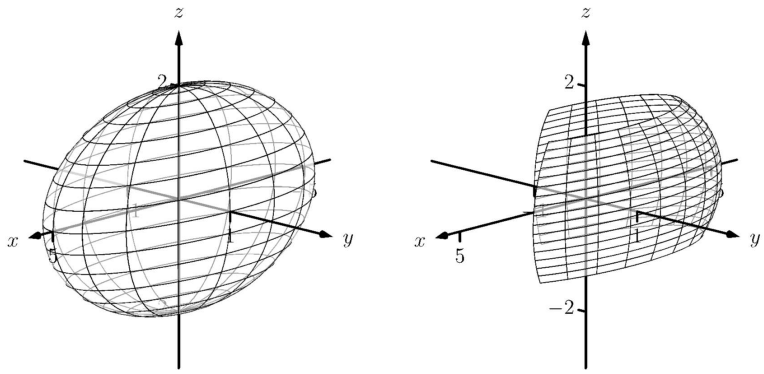
We can parametrize the x component of our surface with $x = 2z\cos u$ and the y component with $y = 3z\sin u$, where $0 \le u \le 2\pi$. Putting all components together, we have

$$\vec{r}(u,v) = \langle 2v\cos u, 3v\sin u, v\rangle, 0 \le u \le 2\pi, -2 \le v \le 3.$$

When v takes on negative values, the radii of the cross-sectional ellipses become "negative," which can lead to some surprising results. Consider Figure 16.5.17, where the cone is graphed for $0 \le u \le \pi$. Because v is negative below the xy-plane, the radii of the cross-sectional ellipses are negative, and the opposite side of the cone is sketched below the xy-plane.

Example 16.5.18 Parametrizing an ellipsoid.

Find a parametrization of the ellipsoid $\frac{x^2}{25} + y^2 + \frac{z^2}{4} = 1$ as shown in Figure 16.5.19(a).



(a) **(b)**

Figure 16.5.19 An ellipsoid in (a), drawn again in (b) with its domain restricted, as described in Example 16.5.18

Solution. Recall Key Idea 12.2.25 from Section 12.2, which states that all unit vectors in space have the form $\langle \sin\theta\cos\varphi, \sin\theta\sin\varphi, \cos\theta\rangle$ for some angles θ and φ. If we choose our angles appropriately, this allows us to draw the unit sphere. To get an ellipsoid, we need only scale each component of the sphere appropriately.

The x-radius of the given ellipsoid is 5, the y-radius is 1 and the z-radius is 2. Substituting u for θ and v for φ, we have

$$\vec{r}(u,v) = \langle 5\sin u\cos v, \sin u\sin v, 2\cos u\rangle,$$

where we still need to determine the ranges of u and v.

Note how the x and y components of $\vec{r}$ have $\cos v$ and $\sin v$ terms, respectively. This hints at the fact that ellipses are drawn parallel to the xy-plane as v varies, which implies we should have v range from 0 to 2π.

One may be tempted to let $0 \le u \le 2\pi$ as well, but note how the z component is $2\cos u$. We only need $\cos u$ to take on values between -1 and 1 once, therefore we can restrict u to $0 \le u \le \pi$.

The final parametrization is thus

$$\vec{r}(u,v) = \langle 5\sin u\cos v, \sin u \sin v, 2\cos u\rangle, 0 \le u \le \pi, 0 \le v \le 2\pi.$$

In Figure 16.5.19(b), the ellipsoid is graphed on $\frac{\pi}{4} \le u \le \frac{2\pi}{3}$, $\frac{\pi}{4} \le v \le \frac{3\pi}{2}$ to demonstrate how each variable affects the surface.

Parametrization is a powerful way to represent surfaces. One of the advantages of the methods of parametrization described in this section is that the domain of $\vec{r}(u,v)$ is always a rectangle; that is, the bounds on u and v are constants. This will make some of our future computations easier to evaluate.

Just as we could parametrize curves in more than one way, there will always be multiple ways to parametrize a surface. Some ways will be more "natural" than others, but these other ways are not incorrect. Because technology is often readily available, it is often a good idea to check one's work by graphing a parametrization of a surface to check if it indeed represents what it was intended to.

youtu.be/watch?v=9TtJ6xpuadQ

Figure 16.5.20 Tangent and normal vectors for parametric surfaces

16.5.2 Surface Area

It will become important in the following sections to be able to compute the surface area of a surface $\mathcal{S}$ given a smooth parametrization $\vec{r}(u,v)$, $a \le u \le b$, $c \le v \le d$. Following the principles given in the integration review at the beginning of this chapter, we can say that

$$\text{Surface Area of } \mathcal{S} = S = \iint_{\mathcal{S}} dS,$$

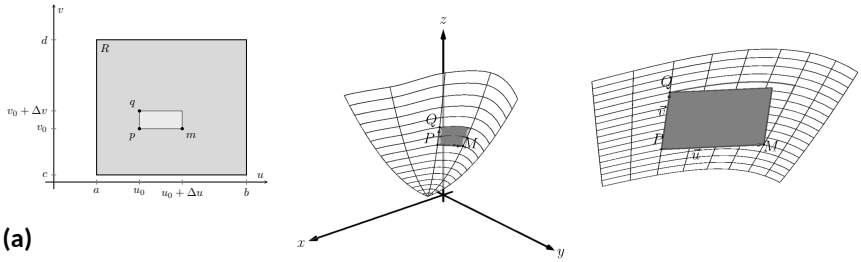
where dS represents a small amount of surface area. That is, to compute total surface area S, add up lots of small amounts of surface area dS across the entire surface $\mathcal{S}$. The key to finding surface area is knowing how to compute dS. We begin by approximating.

In Section 15.5 we used the area of a plane to approximate the surface area of a small portion of a surface. We will do the same here.

Let R be the region of the u-v plane bounded by $a \le u \le b$, $c \le v \le d$ as shown in Figure 16.5.21(a). Partition R into rectangles of width $\Delta u = \frac{b-a}{n}$ and height $\Delta v = \frac{d-c}{n}$, for some n. Let $p = (u_0, v_0)$ be the lower left corner of some rectangle in the partition, and let m and q be neighboring corners as shown.

The point p maps to a point $P = \vec{r}(u_0, v_0)$ on the surface $\mathcal{S}$, and the rectangle with corners p, m and q maps to some region (probably not rectangular) on the surface as shown in Figure 16.5.21(b), where $M = \vec{r}(m)$ and $Q = \vec{r}(q)$. We wish to approximate the surface area of this mapped region.

Let $\vec{u} = M - P$ and $\vec{v} = Q - P$. These two vectors form a parallelogram, illustrated in Figure 16.5.21(c), whose area *approximates* the surface area we seek. In this particular illustration, we can see that parallelogram does not particularly match well the region we wish to approximate, but that is acceptable; by increasing the number of partitions of R, Δu and Δv shrink and our approximations will become better.



(a) **(b)** **(c)**

Figure 16.5.21 Illustrating the process of finding surface area by approximating with planes

From Section 12.4 we know the area of this parallelogram is $|| \vec{u} \times \vec{v} ||$. If we repeat this approximation process for each rectangle in the partition of R, we can sum the areas of all the parallelograms to get an approximation of the surface area S:

$$\text{Surface area of } \mathcal{S} = S \approx \sum_{j=1}^{n} \sum_{i=1}^{n} || \vec{u}_{i,j} \times \vec{v}_{i,j} ||,$$

where $\vec{u}_{i,j} = \vec{r}(u_i + \Delta u, v_j) - \vec{r}(u_i, v_j)$ and $\vec{v}_{i,j} = \vec{r}(u_i, v_j + \Delta v) - \vec{r}(u_i, v_j)$.

From our previous calculus experience, we expect that taking a limit as $n \to \infty$ will result in the exact surface area. However, the current form of the above double sum makes it difficult to realize what the result of that limit is. The following rewriting of the double summation will be helpful:

$$\sum_{j=1}^{n} \sum_{i=1}^{n} || \vec{u}_{i,j} \times \vec{v}_{i,j} || =$$

$$\sum_{j=1}^{n} \sum_{i=1}^{n} \left|\left| \left(\vec{r}(u_i + \Delta u, v_j) - \vec{r}(u_i, v_j)\right) \times \left(\vec{r}(u_i, v_j + \Delta v) - \vec{r}(u_i, v_j)\right) \right|\right| =$$

$$\sum_{j=1}^{n} \sum_{i=1}^{n} \left|\left| \frac{\vec{r}(u_i + \Delta u, v_j) - \vec{r}(u_i, v_j)}{\Delta u} \times \frac{\vec{r}(u_i, v_j + \Delta v) - \vec{r}(u_i, v_j)}{\Delta v} \right|\right| \Delta u \Delta v.$$

We now take the limit as $n \to \infty$, forcing Δu and Δv to 0. As $\Delta u \to 0$,

$$\frac{\vec{r}(u_i + \Delta u, v_j) - \vec{r}(u_i, v_j)}{\Delta u} \to \vec{r}_u(u_i, v_j) \text{ and}$$

$$\frac{\vec{r}(u_i, v_j + \Delta v) - \vec{r}(u_i, v_j)}{\Delta v} \to \vec{r}_v(u_i, v_j).$$

(This limit process also demonstrates that $\vec{r}_u(u, v)$ and $\vec{r}_v(u, v)$ are tangent to the surface $\mathcal{S}$ at $\vec{r}(u, v)$. We don't need this fact now, but it will be important in the next section.)

Thus, in the limit, the double sum leads to a double integral:

$$\lim_{n \to \infty} \sum_{j=1}^{n} \sum_{i=1}^{n} || \vec{u}_{i,j} \times \vec{v}_{i,j} || = \int_c^d \int_a^b || \vec{r}_u \times \vec{r}_v || \, du \, dv.$$

Theorem 16.5.22 Surface Area of Parametrically Defined Surfaces.

Let $\vec{r}(u,v)$ be a smooth parametrization of a surface $\mathcal{S}$ over a closed, bounded region R of the u-v plane.

- *The surface area differential dS is: $dS = || \vec{r}_u \times \vec{r}_v || \, dA$.*
- *The surface area S of $\mathcal{S}$ is*

$$S = \iint_{\mathcal{S}} dS = \iint_R || \vec{r}_u \times \vec{r}_v || \, dA.$$

Example 16.5.24 Finding the surface area of a parametrized surface.

Using the parametrization found in Example 16.5.6, find the surface area of $z = x^2 + 2y^2$ over the circular disk of radius 2, centered at the origin.

Solution. In Example 16.5.6, we parametrized the surface as $\vec{r}(u,v) = \langle 2v\cos u, 2v\sin u, 4v^2\cos^2 u + 8v^2\sin^2 u\rangle$, for $0 \le u \le 2\pi, 0 \le v \le 1$. To find the surface area using Theorem 16.5.22, we need $|| \vec{r}_u \times \vec{r}_v ||$. We find:

$$\begin{aligned}
\vec{r}_u &= \langle -2v\sin u, 2v\cos u, 8v^2\cos u\sin u\rangle \\
\vec{r}_v &= \langle 2\cos u, 2\sin v, 8v\cos^2 u + 16v\sin^2 u\rangle \\
\vec{r}_u \times \vec{r}_v &= \langle 16v^2\cos u, 32v^2\sin u, -4v\rangle \\
|| \vec{r}_u \times \vec{r}_v || &= \sqrt{256v^4\cos^2 u + 1024v^4\sin^2 u + 16v^2}.
\end{aligned}$$

Thus the surface area is

$$\begin{aligned}
S = \iint_{\mathcal{S}} dS &= \iint_R || \vec{r}_u \times \vec{r}_v || \, dA \\
&= \int_0^1 \int_0^{2\pi} \sqrt{256v^4\cos^2 u + 1024v^4\sin^2 u + 16v^2}\, du\, dv \\
&\approx 53.59.
\end{aligned}$$

There is a lot of tedious work in the above calculations and the final integral is nontrivial. The use of a computer-algebra system is highly recommended.

In Section 16.1, we recalled the arc length differential $ds = || \vec{r}\,'(t) \, || \, dt$. In subsequent sections, we used that differential, but in most applications the "$|| \vec{r}\,'(t) \, ||$" part of the differential canceled out of the integrand (to our benefit, as integrating the square roots of functions is generally difficult). We will find a similar thing happens when we use the surface area differential dS in the following sections. That is, our main goal is not to be able to compute surface area; rather, surface area is a tool to obtain other quantities that are more important and useful. In our applications, we will use dS, but most of the time the "$|| \vec{r}_u \times \vec{r}_v ||$" part will cancel out of the integrand, making the subsequent integration easier to compute.

youtu.be/watch?v=ezVITWloEfA

Figure 16.5.23 Area of parametric surfaces

youtu.be/watch?v=9tIGTOpQ5WE

Figure 16.5.25 Surface area of a sphere

16.5.3 Exercises

Terms and Concepts

1. In your own words, describe what an orientable surface is.

2. Give an example of a non-orientable surface.

Problems

Exercise Group. In the following exercises, parametrize the surface defined by the function $z = f(x, y)$ over each of the given regions R of the xy-plane.

3. $z = 3x^2y$;

(a) R is the rectangle bounded by $-1 \leq x \leq 1$ and $0 \leq y \leq 2$.

(b) R is the circle of radius 3, centered at $(1, 2)$.

(c) R is the triangle with vertices $(0, 0)$, $(1, 0)$ and $(0, 2)$.

(d) R is the region bounded by the x-axis and the graph of $y = 1 - x^2$.

4. $z = 4x + 2y^2$;

(a) R is the rectangle bounded by $1 \leq x \leq 4$ and $5 \leq y \leq 7$.

(b) R is the ellipse with major axis of length 8 parallel to the x-axis, and minor axis of length 6 parallel to the y-axis, centered at the origin.

(c) R is the triangle with vertices $(0, 0)$, $(2, 2)$ and $(0, 4)$.

(d) R is the annulus bounded between the circles, centered at the origin, with radius 2 and radius 5.

Exercise Group. In the following exercises, a surface $\mathcal{S}$ in space is described that cannot be defined as the graph of a function $f(x, y)$. Give a parametrization of $\mathcal{S}$.

5. $\mathcal{S}$ is the rectangle in space with corners at $(0, 0, 0)$, $(0, 2, 0)$, $(0, 2, 1)$ and $(0, 0, 1)$.

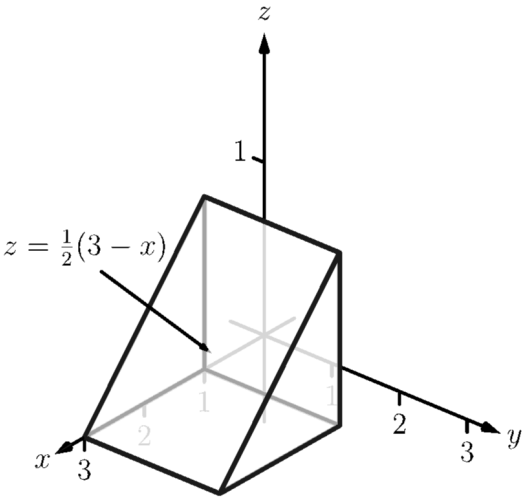
6. $\mathcal{S}$ is the triangle in space with corners at $(1, 0, 0)$, $(1, 0, 1)$ and $(0, 0, 1)$.

7. $\mathcal{S}$ is the ellipsoid $\dfrac{x^2}{9} + \dfrac{y^2}{4} + \dfrac{z^2}{16} = 1$.

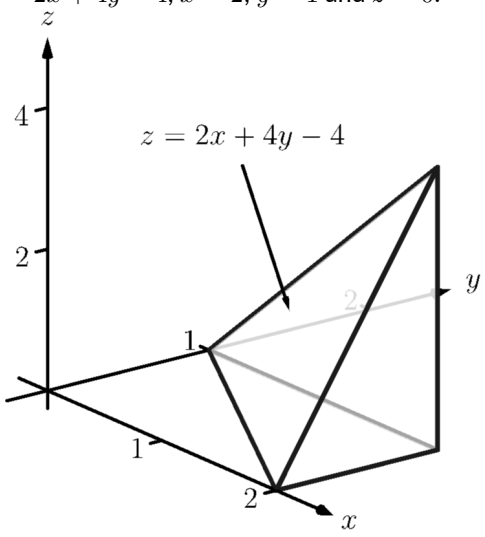
8. $\mathcal{S}$ is the elliptic cone $y^2 = x^2 + \dfrac{z^2}{16}$, for $-1 \leq y \leq 5$.

Exercise Group. In the following exercises, a domain D in space is given. Parametrize each of the bounding surfaces of D.

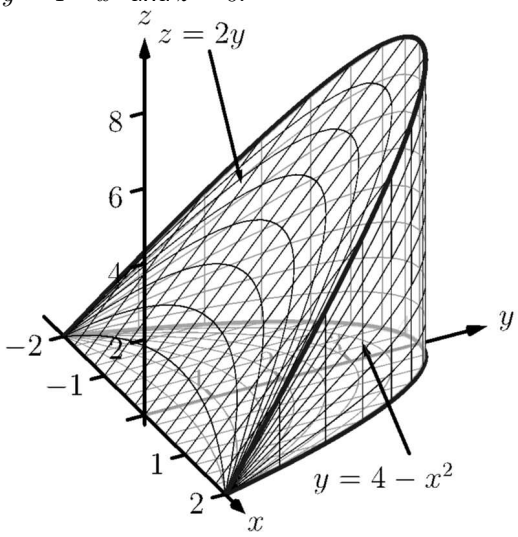
9. D is the domain bounded by the planes $z = \frac{1}{2}(3 - x)$, $x = 1$, $y = 0$, $y = 2$ and $z = 0$.



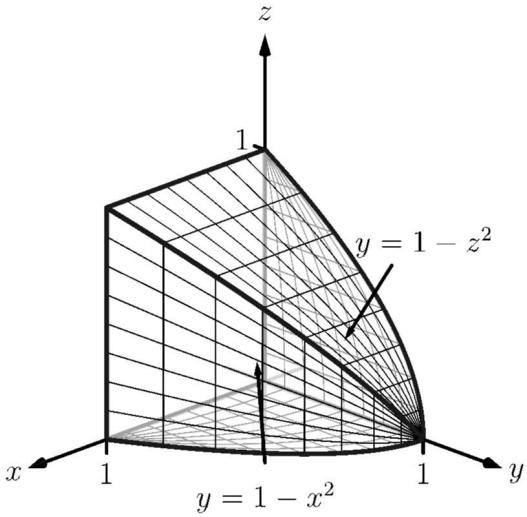
10. D is the domain bounded by the planes $z = 2x + 4y - 4$, $x = 2$, $y = 1$ and $z = 0$.



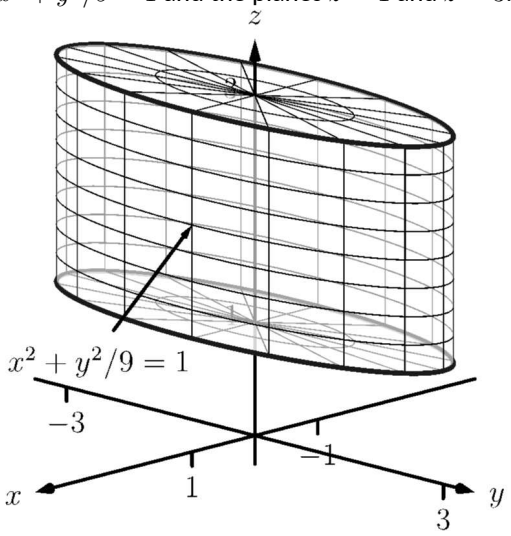
11. D is the domain bounded by $z = 2y$, $y = 4 - x^2$ and $z = 0$.



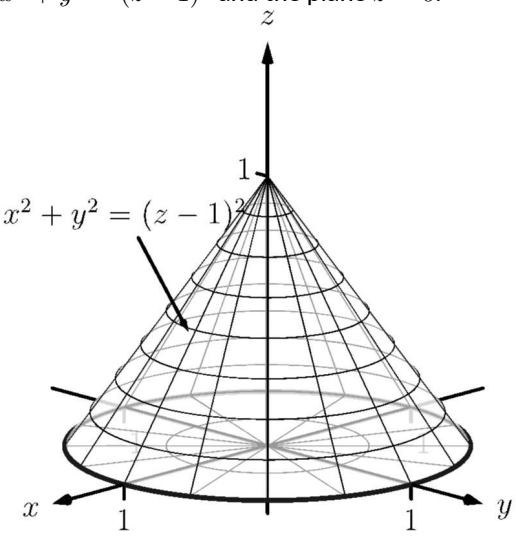
12. D is the domain bounded by $y = 1 - z^2$, $y = 1 - x^2$, $x = 0$, $y = 0$ and $z = 0$.



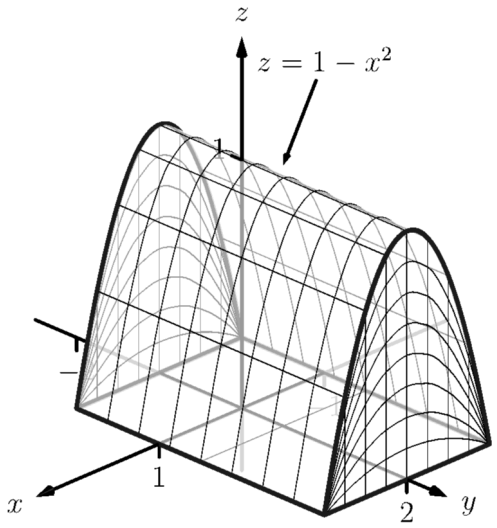
13. D is the domain bounded by the cylinder $x^2 + y^2/9 = 1$ and the planes $z = 1$ and $z = 3$.



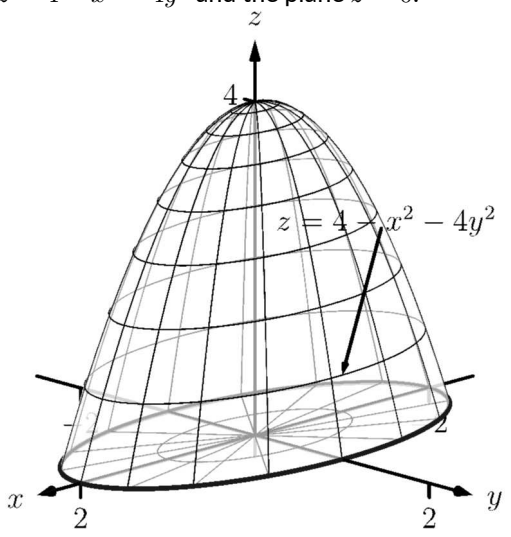
14. D is the domain bounded by the cone $x^2 + y^2 = (z - 1)^2$ and the plane $z = 0$.



15. D is the domain bounded by the cylinder $z = 1 - x^2$ and the planes $y = -1$, $y = 2$ and $z = 0$.



16. D is the domain bounded by the paraboloid $z = 4 - x^2 - 4y^2$ and the plane $z = 0$.



Exercise Group. In the following exercises, find the surface area S of the given surface $\mathcal{S}$. (The associated integrals are computable without the assistance of technology.)

17. $\mathcal{S}$ is the plane $z = 2x + 3y$ over the rectangle $-1 \leq x \leq 1, 2 \leq y \leq 3$.

18. $\mathcal{S}$ is the plane $z = x + 2y$ over the triangle with vertices at $(0, 0)$, $(1, 0)$ and $(0, 1)$.

19. $\mathcal{S}$ is the plane $z = x + y$ over the circular disk, centered at the origin, with radius 2.

20. $\mathcal{S}$ is the plane $z = x + y$ over the annulus bounded by the circles, centered at the origin, with radius 1 and radius 2.

Exercise Group. In the following exercises, set up the double integral that finds the surface area S of the given surface $\mathcal{S}$, then use technology to approximate its value.

21. $\mathcal{S}$ is the paraboloid $z = x^2 + y^2$ over the circular disk of radius 3 centered at the origin.

22. $\mathcal{S}$ is the paraboloid $z = x^2 + y^2$ over the triangle with vertices at $(0, 0)$, $(0, 1)$ and $(1, 1)$.

23. $\mathcal{S}$ is the plane $z = 5x - y$ over the region enclosed by the parabola $y = 1 - x^2$ and the x-axis.

24. $\mathcal{S}$ is the hyperbolic paraboloid $z = x^2 - y^2$ over the circular disk of radius 1 centered at the origin.

16.6 Surface Integrals

Consider a smooth surface $\mathcal{S}$ that represents a thin sheet of metal. How could we find the mass of this metallic object?

If the density of this object is constant, then we can find mass via "mass= density × surface area," and we could compute the surface area using the techniques of the previous section.

What if the density were not constant, but variable, described by a function $\delta(x, y, z)$? We can describe the mass using our general integration techniques as

$$\text{mass } = \iint_{\mathcal{S}} dm,$$

where dm represents "a little bit of mass." That is, to find the total mass of the object, sum up lots of little masses over the surface.

How do we find the "little bit of mass" dm? On a small portion of the surface with surface area ΔS, the density is approximately constant, hence $dm \approx \delta(x, y, z)\Delta S$. As we use limits to shrink the size of ΔS to 0, we get $dm = \delta(x, y, z)dS$; that is, a little bit of mass is equal to a density times a small amount of surface area. Thus the total mass of the thin sheet is

$$\text{mass } = \iint_{\mathcal{S}} \delta(x, y, z)\, dS. \tag{16.6.1}$$

To evaluate the above integral, we would seek $\vec{r}(u, v)$, a smooth parametrization of $\mathcal{S}$ over a region R of the u-v plane. The density would become a function of u and v, and we would integrate $\iint_R \delta(u, v) \,||\, \vec{r}_u \times \vec{r}_v \,||\, dA$.

The integral in Equation (16.6.1) is a specific example of a more general construction defined below.

youtu.be/watch?v=-1NgI5Xts6E

Figure 16.6.1 The surface integral of a scalar field (function)

16.6.1 Surface integrals of scalar fields

Definition 16.6.2 Surface Integral.

Let $G(x, y, z)$ be a continuous function defined on a surface $\mathcal{S}$. The **surface integral** of G on $\mathcal{S}$ is

$$\iint_{\mathcal{S}} G(x, y, z)\, dS.$$

Surface integrals can be used to measure a variety of quantities beyond mass. If $G(x, y, z)$ measures the static charge density at a point, then the surface integral will compute the total static charge of the sheet. If G measures the amount of fluid passing through a screen (represented by $\mathcal{S}$) at a point, then the surface integral gives the total amount of fluid going through the screen.

Example 16.6.3 Finding the mass of a thin sheet.

Find the mass of a thin sheet modeled by the plane $2x + y + z = 3$ over the triangular region of the xy-plane bounded by the coordinate axes and the line $y = 2 - 2x$, as shown in Figure 16.6.4, with density function $\delta(x, y, z) = x^2 + 5y + z$, where all distances are measured in cm and the density is given as gm/cm^2.

Solution. We begin by parametrizing the planar surface $\mathcal{S}$. Using the techniques of the previous section, we can let $x = u$ and $y = v(2 - 2u)$, where $0 \le u \le 1$ and $0 \le v \le 1$. Solving for z in the equation of the

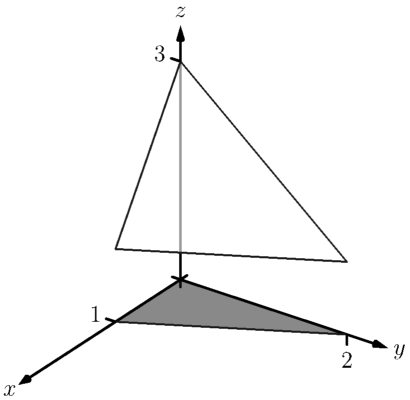


Figure 16.6.4 The surface whose mass is computed in Example 16.6.3

plane, we have $z = 3 - 2x - y$, hence $z = 3 - 2u - v(2 - 2u)$, giving the parametrization $\vec{r}(u, v) = \langle u, v(2 - 2u), 3 - 2u - v(2 - 2u)\rangle$.

We need $dS = || \vec{r}_u \times \vec{r}_v || \, dA$, so we need to compute $\vec{r}_u$, $\vec{r}_v$ and the norm of their cross product. We leave it to the reader to confirm the following:

$$\vec{r}_u = \langle 1, -2v, 2v - 2\rangle, \vec{r}_v = \langle 0, 2 - 2u, 2u - 2\rangle,$$

$\vec{r}_u \times \vec{r}_v = \langle 4 - 4u, 2 - 2u, 2 - 2u\rangle$ and $|| \vec{r}_u \times \vec{r}_v || = 2\sqrt{6}\sqrt{(u-1)^2}$.

We need to be careful to not "simplify" $|| \vec{r}_u \times \vec{r}_v || = 2\sqrt{6}\sqrt{(u-1)^2}$ as $2\sqrt{6}(u - 1)$; rather, it is $2\sqrt{6}|u - 1|$. In this example, u is bounded by $0 \le u \le 1$, and on this interval $|u - 1| = 1 - u$. Thus $dS = 2\sqrt{6}(1 - u)dA$.

The density is given as a function of x, y and z, for which we'll substitute the corresponding components of $\vec{r}$ (with the slight abuse of notation that we used in previous sections):

$$\begin{aligned}\delta(x, y, z) &= \delta\big(\vec{r}(u, v)\big) \\ &= u^2 + 5v(2 - 2u) + 3 - 2u - v(2 - 2u) \\ &= u^2 - 8uv - 2u + 8v + 3.\end{aligned}$$

Thus the mass of the sheet is:

$$\begin{aligned}M &= \iint_{\mathcal{S}} dm \\ &= \iint_R \delta\big(\vec{r}(u, v)\big) \, || \vec{r}_u \times \vec{r}_v || \, dA \\ &= \int_0^1 \int_0^1 \big(u^2 - 8uv - 2u + 8v + 3\big)\big(2\sqrt{6}(1 - u)\big) \, du \, dv \\ &= \frac{31}{\sqrt{6}} \approx 12.66 \text{ gm}.\end{aligned}$$

16.6.2 Flux

Let a surface $\mathcal{S}$ lie within a vector field $\vec{F}$. One is often interested in measuring the *flux* of $\vec{F}$ across $\mathcal{S}$; that is, measuring "how much of the vector field passes across $\mathcal{S}$." For instance, if $\vec{F}$ represents the velocity field of moving air and $\mathcal{S}$ represents the shape of an air filter, the flux will measure how much air is passing through the filter per unit time.

As flux measures the amount of $\vec{F}$ passing across $\mathcal{S}$, we need to find the "amount of $\vec{F}$ orthogonal to $\mathcal{S}$." Similar to our measure of flux in the plane, this is equal to $\vec{F} \cdot \vec{n}$, where $\vec{n}$ is a unit vector normal to $\mathcal{S}$ at a point. We now consider how to find $\vec{n}$.

youtu.be/watch?v=ISowVEDVEzM

Figure 16.6.5 Introducing surface integrals of vector fields

Given a smooth parametrization $\vec{r}(u, v)$ of $\mathcal{S}$, the work in the previous section showing the development of our method of computing surface area also shows that $\vec{r}_u(u, v)$ and $\vec{r}_v(u, v)$ are tangent to $\mathcal{S}$ at $\vec{r}(u, v)$. Thus $\vec{r}_u \times \vec{r}_v$ is orthogonal to $\mathcal{S}$, and we let

$$\vec{n} = \frac{\vec{r}_u \times \vec{r}_v}{|| \vec{r}_u \times \vec{r}_v ||},$$

which is a unit vector normal to $\mathcal{S}$ at $\vec{r}(u, v)$.

The measurement of flux across a surface is a surface integral; that is, to measure total flux we sum the product of $\vec{F} \cdot \vec{n}$ times a small amount of surface area: $\vec{F} \cdot \vec{n}\, dS$.

A nice thing happens with the actual computation of flux: the $|| \vec{r}_u \times \vec{r}_v ||$ terms go away. Consider:

$$\begin{aligned} \text{Flux} &= \iint_{\mathcal{S}} \vec{F} \cdot \vec{n}\, dS \\ &= \iint_R \vec{F} \cdot \frac{\vec{r}_u \times \vec{r}_v}{|| \vec{r}_u \times \vec{r}_v ||} || \vec{r}_u \times \vec{r}_v ||\, dA \\ &= \iint_R \vec{F} \cdot (\vec{r}_u \times \vec{r}_v)\, dA. \end{aligned}$$

The above only makes sense if $\mathcal{S}$ is orientable; the normal vectors $\vec{n}$ must vary continuously across $\mathcal{S}$. We assume that $\vec{n}$ does vary continuously. (If the parametrization $\vec{r}$ of $\mathcal{S}$ is smooth, then our above definition of $\vec{n}$ will vary continuously.)

Definition 16.6.6 Flux over a surface.

Let $\vec{F}$ be a vector field with continuous components defined on an orientable surface $\mathcal{S}$ with normal vector $\vec{n}$. The **flux** of $\vec{F}$ across $\mathcal{S}$ is

$$\text{Flux} = \iint_{\mathcal{S}} \vec{F} \cdot \vec{n}\, dS.$$

If $\mathcal{S}$ is parametrized by $\vec{r}(u, v)$, which is smooth on its domain R, then

$$\text{Flux} = \iint_R \vec{F}\big(\vec{r}(u, v)\big) \cdot (\vec{r}_u \times \vec{r}_v)\, dA.$$

Since $\mathcal{S}$ is orientable, we adopt the convention of saying one passes from the "back" side of $\mathcal{S}$ to the "front" side when moving across the surface parallel to the direction of $\vec{n}$. Also, when $\mathcal{S}$ is closed, it is natural to speak of the regions of space "inside" and "outside" $\mathcal{S}$. We also adopt the convention that when $\mathcal{S}$ is a closed surface, $\vec{n}$ should point to the outside of $\mathcal{S}$. If $\vec{n} = \vec{r}_u \times \vec{r}_v$ points inside $\mathcal{S}$, use $\vec{n} = \vec{r}_v \times \vec{r}_u$ instead.

When the computation of flux is positive, it means that the field is moving from the back side of $\mathcal{S}$ to the front side; when flux is negative, it means the field is moving opposite the direction of $\vec{n}$, and is moving from the front of $\mathcal{S}$ to the back. When $\mathcal{S}$ is not closed, there is not a "right" and "wrong" direction in which $\vec{n}$ should point, but one should be mindful of its direction to make full sense of the flux computation.

We demonstrate the computation of flux, and its interpretation, in the following examples.

Example 16.6.7 Finding flux across a surface.

Let $\mathcal{S}$ be the surface given in Example 16.6.3, where $\mathcal{S}$ is parametrized by $\vec{r}(u, v) = \langle u, v(2-2u), 3-2u-v(2-2u)\rangle$ on $0 \le u \le 1, 0 \le v \le 1$, and let $\vec{F} = \langle 1, x, -y\rangle$, as shown in Figure 16.6.8. Find the flux of $\vec{F}$ across $\mathcal{S}$.

Solution. Using our work from the previous example, we have $\vec{n} = \vec{r}_u \times \vec{r}_v = \langle 4-4u, 2-2u, 2-2u\rangle$. We also need $\vec{F}\big(\vec{r}(u, v)\big) = \langle 1, u, -v(2-2u)\rangle$.

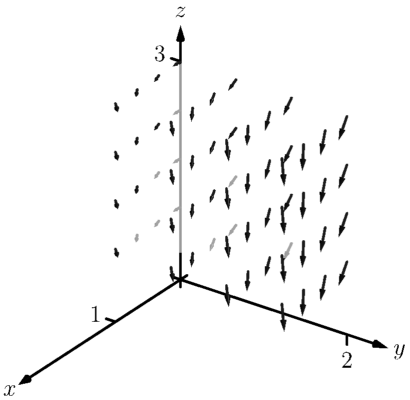


Figure 16.6.8 The surface and vector field used in Example 16.6.7

Thus the flux of $\vec{F}$ across $\mathcal{S}$ is:

$$\begin{aligned}\text{Flux} &= \iint_{\mathcal{S}} \vec{F}\cdot\vec{n}\,dS \\ &= \iint_R \langle 1, u, -v(2-2u)\rangle \cdot \langle 4-4u, 2-2u, 2-2u\rangle\, dA \\ &= \int_0^1\int_0^1 \big(-4u^2v - 2u^2 + 8uv - 2u - 4v + 4\big)\, du\, dv \\ &= 5/3.\end{aligned}$$

To make full use of this numeric answer, we need to know the direction in which the field is passing across $\mathcal{S}$. The graph in Figure 16.6.8 helps, but we need a method that is not dependent on a graph.

Pick a point (u, v) in the interior of R and consider $\vec{n}(u, v)$. For instance, choose $(1/2, 1/2)$ and look at $\vec{n}(1/2, 1/2) = \langle 2, 1, 1\rangle/\sqrt{6}$. This vector has positive x, y and z components. Generally speaking, one has *some* idea of what the surface $\mathcal{S}$ looks like, as that surface is for some reason important. In our case, we know $\mathcal{S}$ is a plane with z-intercept of $z = 3$. Knowing $\vec{n}$ and the flux measurement of positive $5/3$, we know that the field must be passing from "behind" $\mathcal{S}$, i.e., the side the origin is on, to the "front" of $\mathcal{S}$.

youtu.be/watch?v=5XEt-DZedNM

Figure 16.6.9 Computing a surface integral over part of a paraboloid

Example 16.6.10 Flux across surfaces with shared boundaries.

Let $\mathcal{S}_1$ be the unit disk in the xy-plane, and let $\mathcal{S}_2$ be the paraboloid $z = 1 - x^2 - y^2$, for $z \geq 0$, as graphed in Figure 16.6.11. Note how these two surfaces each have the unit circle as a boundary.

Let $\vec{F}_1 = \langle 0, 0, 1\rangle$ and $\vec{F}_2 = \langle 0, 0, z\rangle$. Using normal vectors for each surface that point "upward," i.e., with a positive z-component, find the flux of each field across each surface.

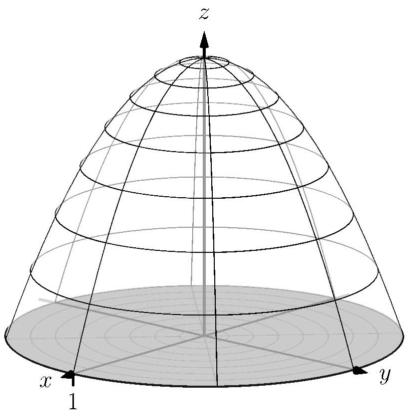


Figure 16.6.11 The surfaces used in Example 16.6.10

Solution. We begin by parametrizing each surface.

The boundary of the unit disk in the xy-plane is the unit circle, which can be described with $\langle \cos u, \sin u, 0\rangle$, $0 \leq u \leq 2\pi$. To obtain the interior of the circle as well, we can scale by v, giving

$$\vec{r}_1(u, v) = \langle v\cos u, v\sin u, 0\rangle, 0 \leq u \leq 2\pi 0 \leq v \leq 1.$$

As the boundary of $\mathcal{S}_2$ is also the unit circle, the x and y components of $\vec{r}_2$ will be the same as those of $\vec{r}_1$; we just need a different z component. With $z = 1 - x^2 - y^2$, we have

$$\vec{r}_2(u, v) = \langle v\cos u, v\sin u, 1-v^2\cos^2 u - v^2\sin^2 u\rangle = \langle v\cos u, v\sin u, 1-v^2\rangle,$$

where $0 \leq u \leq 2\pi$ and $0 \leq v \leq 1$.

We now compute the normal vectors $\vec{n}_1$ and $\vec{n}_2$.

For $\vec{n}_1$: $\vec{r}_{1u} = \langle -v\sin u, v\cos u, 0\rangle$, $\vec{r}_{1v} = \langle \cos u, \sin u, 0\rangle$, so

$$\vec{n}_1 = \vec{r}_{1u} \times \vec{r}_{1v} = \langle 0, 0, -v\rangle.$$

As this vector has a negative z-component, we instead use

$$\vec{n}_1 = \vec{r}_{1v} \times \vec{r}_{1u} = \langle 0, 0, v\rangle.$$

Similarly, $\vec{n}_2$: $\vec{r}_{2u} = \langle -v\sin u, v\cos u, 0\rangle$, $\vec{r}_{2v} = \langle \cos u, \sin u, -2v\rangle$, so

$$\vec{n}_2 = \vec{r}_{2u} \times \vec{r}_{2v} = \langle -2v^2\cos u, -2v^2\sin u, -v\rangle.$$

Again, this normal vector has a negative z-component so we use

$$\vec{n}_2 = \vec{r}_{2v} \times \vec{r}_{2u} = \langle 2v^2\cos u, 2v^2\sin u, v\rangle.$$

We are now set to compute flux. Over field $\vec{F}_1 = \langle 0, 0, 1\rangle$:

$$\begin{aligned}\text{Flux across } \mathcal{S}_1 &= \iint_{\mathcal{S}_1} \vec{F}_1 \cdot \vec{n}_1\, dS \\ &= \iint_R \langle 0,0,1\rangle \cdot \langle 0,0,v\rangle\, dA \\ &= \int_0^1 \int_0^{2\pi} (v)\, du\, dv \\ &= \pi.\end{aligned}$$

$$\begin{aligned}\text{Flux across } \mathcal{S}_2 &= \iint_{\mathcal{S}_2} \vec{F}_1 \cdot \vec{n}_2\, dS \\ &= \iint_R \langle 0,0,1\rangle \cdot \langle 2v^2\cos u, 2v^2\sin u, v\rangle\, dA \\ &= \int_0^1 \int_0^{2\pi} (v)\, du\, dv \\ &= \pi.\end{aligned}$$

These two results are equal and positive. Each are positive because both normal vectors are pointing in the positive z-directions, as does $\vec{F}_1$. As the field passes through each surface in the direction of their normal vectors, the flux is measured as positive.

We can also intuitively understand why the results are equal. Consider $\vec{F}_1$ to represent the flow of air, and let each surface represent a filter. Since $\vec{F}_1$ is constant, and moving "straight up," it makes sense that all air passing through $\mathcal{S}_1$ also passes through $\mathcal{S}_2$, and vice-versa.

If we treated the surfaces as creating one piecewise-smooth surface $\mathcal{S}$, we would find the total flux across $\mathcal{S}$ by finding the flux across each piece, being sure that each normal vector pointed to the outside of the closed surface. Above, $\vec{n}_1$ does not point outside the surface, though $\vec{n}_2$ does. We would instead want to use $-\vec{n}_1$ in our computation. We would then find that the flux across $\mathcal{S}_1$ is $-\pi$, and hence the total flux across $\mathcal{S}$ is $-\pi+\pi = 0$. (As 0 is a special number, we should wonder if this answer has special significance. It does, which is briefly discussed following this example and will be more fully developed in the next section.)

We now compute the flux across each surface with $\vec{F}_2 = \langle 0, 0, z\rangle$:

$$\text{Flux across } \mathcal{S}_1 = \iint_{\mathcal{S}_1} \vec{F}_2 \cdot \vec{n}_1\, dS.$$

Over $\mathcal{S}_1$, $\vec{F}_2 = \vec{F}_2\big(\vec{r}_2(u,v)\big) = \langle 0,0,0\rangle$. Therefore,

$$= \iint_R \langle 0,0,0\rangle \cdot \langle 0,0,v\rangle\, dA$$

$$= \int_0^1 \int_0^{2\pi} (0)\, du\, dv$$
$$= 0.$$

Flux across $\mathcal{S}_2 = \iint_{\mathcal{S}_2} \vec{F}_2 \cdot \vec{n}_2\, dS.$

Over $\mathcal{S}_2$, $\vec{F}_2 = \vec{F}_2\big(\vec{r}_2(u,v)\big) = \langle 0, 0, 1 - v^2\rangle$. Therefore,

$$= \iint_R \langle 0, 0, 1 - v^2\rangle \cdot \langle 2v^2 \cos u, 2v^2 \sin u, v\rangle\, dA$$
$$= \int_0^1 \int_0^{2\pi} (v^3 - v)\, du\, dv$$
$$= \pi/2.$$

This time the measurements of flux differ. Over $\mathcal{S}_1$, the field $\vec{F}_2$ is just $\vec{0}$, hence there is no flux. Over $\mathcal{S}_2$, the flux is again positive as $\vec{F}_2$ points in the positive z direction over $\mathcal{S}_2$, as does $\vec{n}_2$.

In the previous example, the surfaces $\mathcal{S}_1$ and $\mathcal{S}_2$ form a closed surface that is piecewise smooth. That the measurement of flux across each surface was the same for some fields (and not for others) is reminiscent of a result from Section 16.4, where we measured flux across curves. The quick answer to why the flux was the same when considering $\vec{F}_1$ is that div $\vec{F}_1 = 0$. In the next section, we'll see the second part of the Divergence Theorem, which will more fully explain this occurrence. We will also explore Stokes' Theorem, the spatial analogue to Green's Theorem.

The videos in Figure 16.6.12–16.6.13 present some additional examples involving surface integrals of vector fields. Note that computing flux a cross a cube requires us to consider all six faces. This is a case where the Divergence Theorem can greatly simplify matters. For integrals over spheres, there are often simplifications possible, especially for "radial" vector fields (those parallel to the position vector $\vec{r}(x,y,z) = \langle x, y, z\rangle$).

youtu.be/watch?v=u59S_E8VCM8

Figure 16.6.12 Computing flux across a cube

youtu.be/watch?v=xdtbrzLRtBY

Figure 16.6.13 Some surface integrals over spheres

16.6.3 Exercises

Terms and Concepts

1. In the plane, flux is a measurement of how much of the vector field passes across a _____; in space, flux is a measurement of how much of the vector field passes across a _____.

2. When computing flux, what does it mean when the result is a negative number?

3. When $\mathcal{S}$ is a closed surface, we choose the normal vector so that it points to the _____ of the surface.

4. If $\mathcal{S}$ is a plane, and $\vec{F}$ is always parallel to $\mathcal{S}$, then the flux of $\vec{F}$ across $\mathcal{S}$ will be _____.

Problems

Exercise Group. In the following exercises, a surface $\mathcal{S}$ that represents a thin sheet of material with density δ is given. Find the mass of each thin sheet.

5. $\mathcal{S}$ is the plane $z = x + y$ on $-2 \leq x \leq 2$, $-3 \leq y \leq 3$, with $\delta(x, y, z) = z + 10$.

6. $\mathcal{S}$ is the unit sphere, with $\delta(x, y, z) = x + y + z + 10$.

Exercise Group. In the following exercises, a surface $\mathcal{S}$ and a vector field $\vec{F}$ are given. Compute the flux of $\vec{F}$ across $\mathcal{S}$. (If $\mathcal{S}$ is not a closed surface, choose $\vec{n}$ so that it has a positive z-component, unless otherwise indicated.)

7. $\mathcal{S}$ is the plane $z = 3x + y$ on $0 \leq x \leq 1$, $1 \leq y \leq 4$; $\vec{F} = \langle x^2, -z, 2y \rangle$.

8. $\mathcal{S}$ is the plane $z = 8 - x - y$ over the triangle with vertices at $(0, 0)$, $(1, 0)$ and $(1, 5)$; $\vec{F} = \langle 3, 1, 2 \rangle$.

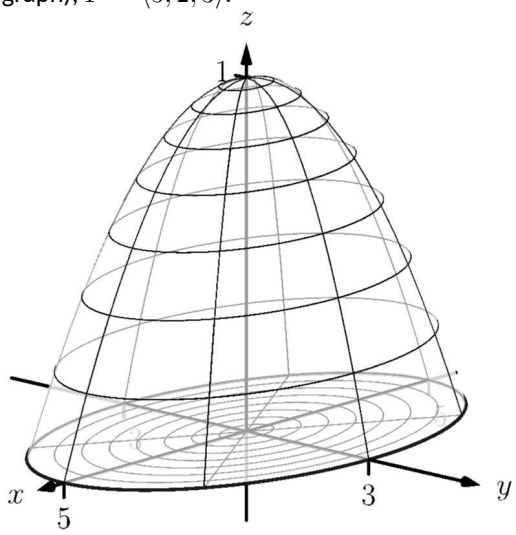
9. $\mathcal{S}$ is the paraboloid $z = x^2 + y^2$ over the unit disk; $\vec{F} = \langle 1, 0, 0 \rangle$.

10. $\mathcal{S}$ is the unit sphere; $\vec{F} = \langle y - z, z - x, x - y \rangle$.

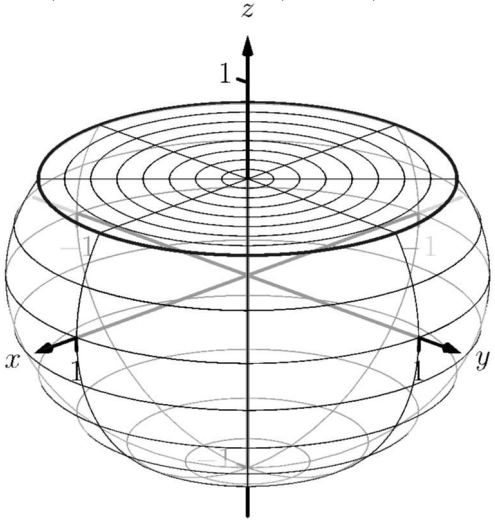
11. $\mathcal{S}$ is the square in space with corners at $(0, 0, 0)$, $(1, 0, 0)$, $(1, 0, 1)$ and $(0, 0, 1)$ (choose $\vec{n}$ such that it has a positive y-component); $\vec{F} = \langle 0, -z, y \rangle$.

12. $\mathcal{S}$ is the disk in the yz-plane with radius 1, centered at $(0, 1, 1)$ (choose $\vec{n}$ such that it has a positive x-component); $\vec{F} = \langle y, z, x \rangle$.

13. $\mathcal{S}$ is the closed surface composed of $\mathcal{S}_1$, whose boundary is the ellipse in the xy-plane described by $\frac{x^2}{25} + \frac{y^2}{9} = 1$ and $\mathcal{S}_2$, part of the elliptical paraboloid $f(x, y) = 1 - \frac{x^2}{25} - \frac{y^2}{9}$ (see graph); $\vec{F} = \langle 5, 2, 3 \rangle$.



14. $\mathcal{S}$ is the closed surface composed of $\mathcal{S}_1$, part of the unit sphere and $\mathcal{S}_2$, part of the plane $z = 1/2$ (see graph); $\vec{F} = \langle x, -y, z \rangle$.



16.7 The Divergence Theorem and Stokes' Theorem

16.7.1 The Divergence Theorem

youtu.be/watch?v=Ra6bxrTj5xU

Figure 16.7.1 Introducing the Divergenge Theorem

Theorem 16.4.20 gives the Divergence Theorem in the plane, which states that the flux of a vector field across a closed *curve* equals the sum of the divergences over the region enclosed by the curve. Recall that the flux was measured via a line integral, and the sum of the divergences was measured through a double integral.

We now consider the three-dimensional version of the Divergence Theorem. It states, in words, that the flux across a closed *surface* equals the sum of the divergences over the domain enclosed by the surface. Since we are in space (versus the plane), we measure flux via a surface integral, and the sums of divergences will be measured through a triple integral.

Theorem 16.7.2 The Divergence Theorem (in space).

Let D be a closed domain in space whose boundary is an orientable, piecewise smooth surface $\mathcal{S}$ with outer unit normal vector $\vec{n}$, and let $\vec{F}$ be a vector field whose components are differentiable on D. Then

$$\iint_{\mathcal{S}} \vec{F} \cdot \vec{n}\, dS = \iiint_D \text{div}\, \vec{F}\, dV.$$

Note: the term "outer unit normal vector" used in Theorem 16.7.2 means $\vec{n}$ points to the outside of $\mathcal{S}$.

Example 16.7.3 Using the Divergence Theorem in space.

Let D be the domain in space bounded by the planes $z = 0$ and $z = 2x$, along with the cylinder $x = 1 - y^2$, as graphed in Figure 16.7.4, let $\mathcal{S}$ be the boundary of D, and let $\vec{F} = \langle x + y, y^2, 2z \rangle$.

Verify the Divergence Theorem by finding the total outward flux of $\vec{F}$ across $\mathcal{S}$, and show this is equal to $\iiint_D \text{div}\, \vec{F}\, dV$.

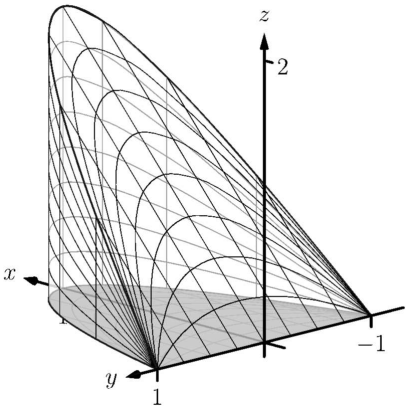


Figure 16.7.4 The surfaces used in Example 16.7.3

Solution. The surface $\mathcal{S}$ is piecewise smooth, comprising surfaces $\mathcal{S}_1$, which is part of the plane $z = 2x$, surface $\mathcal{S}_2$, which is part of the cylinder $x = 1 - y^2$, and surface $\mathcal{S}_3$, which is part of the plane $z = 0$. To find the total outward flux across $\mathcal{S}$, we need to compute the outward flux across each of these three surfaces.

We leave it to the reader to confirm that surfaces $\mathcal{S}_1$, $\mathcal{S}_2$ and $\mathcal{S}_3$ can be parametrized by $\vec{r}_1$, $\vec{r}_2$ and $\vec{r}_3$ respectively as

$$\begin{aligned}
\vec{r}_1(u,v) &= \left\langle v(1-u^2), u, 2v(1-u^2) \right\rangle, \\
\vec{r}_2(u,v) &= \left\langle (1-u^2), u, 2v(1-u^2) \right\rangle, \\
\vec{r}_3(u,v) &= \left\langle v(1-u^2), u, 0 \right\rangle,
\end{aligned}$$

where $-1 \le u \le 1$ and $0 \le v \le 1$ for all three functions.

We compute a unit normal vector $\vec{n}$ for each as $\frac{\vec{r}_u \times \vec{r}_v}{||\,\vec{r}_u \times \vec{r}_v\,||}$, though recall that as we are integrating $\vec{F} \cdot \vec{n}\, dS$, we actually only use $\vec{r}_u \times \vec{r}_v$. Finally, in previous flux computations, it did not matter which direction $\vec{n}$ pointed as long as we made note of its direction. When using the Divergence Theorem, we need $\vec{n}$ to point to the outside of the closed surface, so in practice this means we'll either use $\vec{r}_u \times \vec{r}_v$ or $\vec{r}_v \times \vec{r}_u$, depending on which points outside of the closed surface $\mathcal{S}$.

We leave it to the reader to confirm the following cross products and integrations are correct.

For $\mathcal{S}_1$, we need to use $\vec{r}_{1v} \times \vec{r}_{1u} = \langle 2(u^2 - 1), 0, 1 - u^2 \rangle$. (Note the z-component is nonnegative as $u \le 1$, therefore this vector always

points up, meaning *to the outside*, of $\mathcal{S}$.) The flux across $\mathcal{S}_1$ is:

$$\begin{aligned}\text{Flux across } \mathcal{S}_1: &= \iint_{\mathcal{S}_1} \vec{F}\cdot\vec{n}_1\,dS\\ &= \int_0^1\int_{-1}^1 \vec{F}\big(\vec{r}_1(u,v)\big)\cdot\big(\vec{r}_{1v}\times\vec{r}_{1u}\big)\,du\,dv\\ &= \int_0^1\int_{-1}^1 \big\langle v(1-u^2)+u, u^2, 4v(1-u^2)\big\rangle\cdot\big\langle 2(u^2-1), 0, 1-u^2\big\rangle\,du\,dv\\ &= \int_0^1\int_{-1}^1 \big(2u^4v+2u^3-4u^2v-2u+2v\big)\,du\,dv\\ &= \frac{16}{15}.\end{aligned}$$

For $\mathcal{S}_2$, we use $\vec{r}_{2u}\times\vec{r}_{2v} = \langle 2(1-u^2), 4u(1-u^2), 0\rangle$. (Note the x-component is always nonnegative, meaning this vector points outside $\mathcal{S}$.) The flux across $\mathcal{S}_2$ is:

$$\begin{aligned}\text{Flux across } \mathcal{S}_2: &= \iint_{\mathcal{S}_2} \vec{F}\cdot\vec{n}_2\,dS\\ &= \int_0^1\int_{-1}^1 \vec{F}\big(\vec{r}_2(u,v)\big)\cdot\big(\vec{r}_{2u}\times\vec{r}_{2v}\big)\,du\,dv\\ &= \int_0^1\int_{-1}^1 \big\langle 1-u^2+u, u^2, 4v(1-u^2)\big\rangle\cdot\big\langle 2(1-u^2), 4u(1-u^2), 0\big\rangle\,du\,dv\\ &= \int_0^1\int_{-1}^1 \big(4u^5-2u^4-2u^3+4u^2-2u-2\big)\,du\,dv\\ &= \frac{32}{15}.\end{aligned}$$

For $\mathcal{S}_3$, we use $\vec{r}_{3u}\times\vec{r}_{3v} = \langle 0, 0, u^2-1\rangle$. (Note the z-component is never positive, meaning this vector points down, outside of $\mathcal{S}$.) The flux across $\mathcal{S}_3$ is:

$$\begin{aligned}\text{Flux across } \mathcal{S}_3: &= \iint_{\mathcal{S}_3} \vec{F}\cdot\vec{n}_3\,dS\\ &= \int_0^1\int_{-1}^1 \vec{F}\big(\vec{r}_3(u,v)\big)\cdot\big(\vec{r}_{3u}\times\vec{r}_{3v}\big)\,du\,dv\\ &= \int_0^1\int_{-1}^1 \big\langle v(1-u^2)+u, u^2, 0\big\rangle\cdot\big\langle 0, 0, u^2-1\big\rangle\,du\,dv\\ &= \int_0^1\int_{-1}^1 0\,du\,dv\\ &= 0.\end{aligned}$$

Thus the total outward flux, measured by surface integrals across all three component surfaces of $\mathcal{S}$, is $16/15+32/15+0 = 48/15 = 16/5 = 3.2$. We now find the total outward flux by integrating div $\vec{F}$ over D.

Following the steps outlined in Section 15.6, we see the bounds of x, y and z can be set as (thinking "surface to surface, curve to curve, point to point"):

$$0\le z\le 2x; 0\le x\le 1-y^2; -1\le y\le 1.$$

With div $\vec{F} = 1 + 2y + 2 = 2y + 3$, we find the total outward flux of $\vec{F}$ over $\mathcal{S}$ as:

$$\text{Flux} = \iiint_D \text{div}\,\vec{F}\,dV = \int_{-1}^{1}\int_{0}^{1-y^2}\int_{0}^{2x} \big(2y+3\big)\,dz\,dx\,dy = 16/5,$$

the same result we obtained previously.

In Example 16.7.3 we see that the total outward flux of a vector field across a closed surface can be found two different ways because of the Divergence Theorem. One computation took far less work to obtain. In that particular case, since $\mathcal{S}$ was comprised of three separate surfaces, it was far simpler to compute one triple integral than three surface integrals (each of which required partial derivatives and a cross product). In practice, if outward flux needs to be measured, one would choose only one method. We will use both methods in this section simply to reinforce the truth of the Divergence Theorem.

We practice again in the following example.

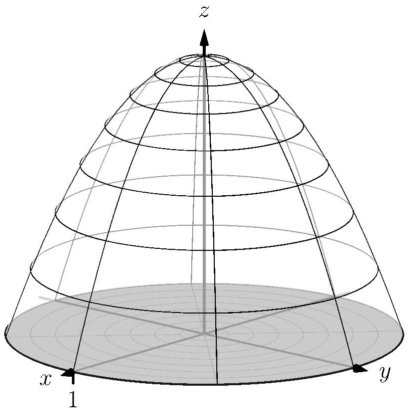


Figure 16.7.6 The surfaces used in Example 16.7.5

Example 16.7.5 Using the Divergence Theorem in space.

Let $\mathcal{S}$ be the surface formed by the paraboloid $z = 1 - x^2 - y^2$, $z \geq 0$, and the unit disk centered at the origin in the xy-plane, graphed in Figure 16.7.6, and let $\vec{F} = \langle 0, 0, z\rangle$. (This surface and vector field were used in Example 16.6.10.)

Verify the Divergence Theorem; find the total outward flux across $\mathcal{S}$ and evaluate the triple integral of div $\vec{F}$, showing that these two quantities are equal.

Solution. We find the flux across $\mathcal{S}$ first. As $\mathcal{S}$ is piecewise-smooth, we decompose it into smooth components $\mathcal{S}_1$, the disk, and $\mathcal{S}_2$, the paraboloid, and find the flux across each.

In Example 16.6.10, we found the flux across $\mathcal{S}_1$ is 0. We also found that the flux across $\mathcal{S}_2$ is $\pi/2$. (In that example, the normal vector had a positive z component hence was an outer normal.) Thus the total outward flux is $0 + \pi/2 = \pi/2$.

We now compute $\iiint_D \text{div}\,\vec{F}\,dV$. We can describe D as the domain bounded by (think "surface to surface, curve to curve, point to point"):

$$0 \leq z \leq 1 - x^2 - y^2, -\sqrt{1-x^2} \leq y \leq \sqrt{1-x^2}, -1 \leq x \leq 1.$$

This description of D is not very easy to integrate. With polar, we can do better. Let R represent the unit disk, which can be described in polar simply as r, where $0 \leq r \leq 1$ and $0 \leq \theta \leq 2\pi$. With $x = r\cos\theta$ and $y = r\sin\theta$, the surface $\mathcal{S}_2$ becomes

$$z = 1 - x^2 - y^2 \Rightarrow 1 - (r\cos\theta)^2 - (r\sin\theta)^2 \Rightarrow 1 - r^2.$$

Thus D can be described as the domain bounded by:

$$0 \leq z \leq 1 - r^2, 0 \leq r \leq 1, 0 \leq \theta \leq 2\pi.$$

With div $\vec{F} = 1$, we can integrate, recalling that $dV = r\,dz\,dr\,d\theta$:

$$\iiint_D \text{div}\,\vec{F}\,dV = \int_0^{2\pi}\int_0^1\int_0^{1-r^2} r\,dz\,dr\,d\theta = \frac{\pi}{2},$$

which matches our flux computation above.

Example 16.7.7 A "paradox" of the Divergence Theorem and Gauss's Law.

The magnitude of many physical quantities (such as light intensity or electromagnetic and gravitational forces) follow an "inverse square law": the magnitude of the quantity at a point is inversely proportional to the square of the distance to the source of the quantity.

Let a point light source be placed at the origin and let $\vec{F}$ be the vector field which describes the intensity and direction of the emanating light. At a point (x, y, z), the unit vector describing the direction of the light passing through that point is $\langle x, y, z\rangle/\sqrt{x^2+y^2+z^2}$. As the intensity of light follows the inverse square law, the magnitude of $\vec{F}$ at (x, y, z) is $k/(x^2+y^2+z^2)$ for some constant k. Taken together,

$$\vec{F}(x, y, z) = \frac{k}{(x^2+y^2+z^2)^{3/2}}\langle x, y, z\rangle.$$

Consider the cube, centered at the origin, with sides of length $2a$ for some $a > 0$ (hence corners of the cube lie at (a, a, a), $(-a, -a, -a)$, etc., as shown in Figure 16.7.8). Find the flux across the six faces of the cube and compare this to $\iiint_D \text{div}\,\vec{F}\,dV$.

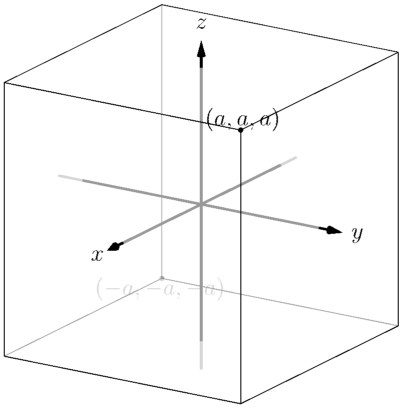


Figure 16.7.8 The cube used in Example 16.7.7

Solution. Let $\mathcal{S}_1$ be the "top" face of the cube, which can be parametrized by $\vec{r}(u, v) = \langle u, v, a\rangle$ for $-a \le u \le a$, $-a \le v \le a$. We leave it to the reader to confirm that $\vec{r}_u \times \vec{r}_v = \langle 0, 0, 1\rangle$, which points outside of the cube.

The flux across this face is:

$$\begin{aligned}\text{Flux} &= \iint_{\mathcal{S}_1} \vec{F}\cdot\vec{n}\,dS \\ &= \int_{-a}^{a}\int_{-a}^{a} \vec{F}\big(\vec{r}(u,v)\big)\cdot\big(\vec{r}_u\times\vec{r}_v\big)\,du\,dv \\ &= \int_{-a}^{a}\int_{-a}^{a} \frac{k\,a}{(u^2+v^2+a^2)^{3/2}}\,du\,dv.\end{aligned}$$

This double integral is *not* trivial to compute, requiring multiple trigonometric substitutions. This example is not meant to stress integration techniques, so we leave it to the reader to confirm the result is

$$= \frac{2k\pi}{3}.$$

Note how the result is independent of a; no matter the size of the cube, the flux through the top surface is always $2k\pi/3$.

An argument of symmetry shows that the flux through each of the six faces is $2k\pi/3$, thus the total flux through the faces of the cube is $6 \times 2k\pi/3 = 4k\pi$.

It takes a bit of algebra, but we can show that $\text{div}\,\vec{F} = 0$. Thus the Divergence Theorem would seem to imply that the total flux through the faces of the cube should be

$$\text{Flux} = \iiint_D \text{div}\,\vec{F}\,dV = \iiint_D 0\,dV = 0,$$

but clearly this does not match the result from above. What went wrong?

Revisit the statement of the Divergence Theorem. One of the conditions is that the components of $\vec{F}$ must be differentiable on the domain enclosed by the surface. In our case, $\vec{F}$ is *not* differentiable at the origin — it is not even defined! As $\vec{F}$ does not satisfy the conditions of the Divergence Theorem, it does not apply, and we cannot expect $\iint_{\mathcal{S}} \vec{F} \cdot \vec{n}\, dA = \iiint_D \text{div}\, \vec{F}\, dV$.

Since $\vec{F}$ is differentiable everywhere except the origin, the Divergence Theorem does apply over any domain that does not include the origin. Let $\mathcal{S}_2$ be any surface that encloses the cube used before, and let $\hat{D}$ be the domain *between* the cube and $\mathcal{S}_2$; note how $\hat{D}$ does not include the origin and so the Divergence Theorem does apply over this domain. The total outward flux over $\hat{D}$ is thus $\iint_{\hat{D}} \text{div}\, \vec{F}\, dV = 0$, which means the amount of flux coming out of $\mathcal{S}_2$ is the same as the amount of flux coming out of the cube. The conclusion: the flux across *any* surface enclosing the origin will be $4k\pi$.

This has an important consequence in electrodynamics. Let q be a point charge at the origin. The electric field generated by this point charge is

$$\vec{E} = \frac{q}{4\pi\epsilon_0} \frac{\langle x, y, z\rangle}{(x^2+y^2+z^2)^{3/2}},$$

i.e., it is $\vec{F}$ with $k = q/(4\pi\epsilon_0)$, where ϵ_0 is a physical constant (the "permittivity of free space"). Gauss's Law states that the outward flux of $\vec{E}$ across any surface enclosing the origin is q/ϵ_0.

youtu.be/watch?v=Zc_OAF_CRIQ

Figure 16.7.9 Exploring Gauss's Law

Our interest in the Divergence Theorem is twofold. First, its truth alone is interesting: to study the behavior of a vector field across a closed surface, one can examine properties of that field within the surface. Secondly, it offers an alternative way of computing flux. When there are multiple methods of computing a desired quantity, one has power to select the easiest computation as illustrated next.

Example 16.7.10 Using the Divergence Theorem to compute flux.

Let $\mathcal{S}$ be the cube bounded by the planes $x = \pm 1$, $y = \pm 1$, $z = \pm 1$, and let $\vec{F} = \langle x^2 y, 2yz, x^2 z^3\rangle$. Compute the outward flux of $\vec{F}$ over $\mathcal{S}$.

Solution. We compute $\text{div}\, \vec{F} = 2xy + 2z + 3x^2z^2$. By the Divergence Theorem, the outward flux is the triple integral over the domain D enclosed by $\mathcal{S}$:

$$\text{Outward flux: } \int_{-1}^{1}\int_{-1}^{1}\int_{-1}^{1} (2xy + 2z + 3x^2z^2)\, dz\, dy\, dx = \frac{8}{3}.$$

The direct flux computation requires six surface integrals, one for each face of the cube. The Divergence Theorem offers a much more simple computation.

16.7.2 Stokes' Theorem

Just as the spatial Divergence Theorem of this section is an extension of the planar Divergence Theorem, Stokes' Theorem is the spatial extension of Green's Theorem. Recall that Green's Theorem states that the circulation of a vector field around a closed curve in the plane is equal to the sum of the curl of the

field over the region enclosed by the curve. Stokes' Theorem effectively makes the same statement: given a closed curve that lies on a surface $\mathcal{S}$, the circulation of a vector field around that curve is the same as the sum of "the curl of the field" across the enclosed surface. We use quotes around "the curl of the field" to signify that this statement is not quite correct, as we do not sum curl $\vec{F}$, but curl $\vec{F} \cdot \vec{n}$, where $\vec{n}$ is a unit vector normal to $\mathcal{S}$. That is, we sum the portion of curl $\vec{F}$ that is orthogonal to $\mathcal{S}$ at a point.

Green's Theorem dictated that the curve was to be traversed counterclockwise when measuring circulation. Stokes' Theorem will follow a right hand rule: when the thumb of one's right hand points in the direction of $\vec{n}$, the path C will be traversed in the direction of the curling fingers of the hand (this is equivalent to traversing counterclockwise in the plane).

youtu.be/watch?v=WzrQbtC7ZiU

Figure 16.7.11 Introducing Stokes' Theorem

Theorem 16.7.12 Stokes' Theorem.

Let $\mathcal{S}$ be a piecewise smooth, orientable surface whose boundary is a piecewise smooth curve C, let $\vec{n}$ be a unit vector normal to $\mathcal{S}$, let C be traversed with respect to $\vec{n}$ according to the right hand rule, and let the components of $\vec{F}$ have continuous first partial derivatives over $\mathcal{S}$. Then

$$\oint_C \vec{F} \cdot \, d\vec{r} = \iint_{\mathcal{S}} (\text{curl}\, \vec{F}) \cdot \vec{n}\, dS.$$

In general, the best approach to evaluating the surface integral in Stokes' Theorem is to parametrize the surface $\mathcal{S}$ with a function $\vec{r}(u, v)$. We can find a unit normal vector $\vec{n}$ as

$$\vec{n} = \frac{\vec{r}_u \times \vec{r}_v}{|| \vec{r}_u \times \vec{r}_v ||}.$$

Since $dS = || \vec{r}_u \times \vec{r}_v ||\, dA$, the surface integral in practice is evaluated as

$$\iint_{\mathcal{S}} (\text{curl}\, \vec{F}) \cdot (\vec{r}_u \times \vec{r}_v)\, dA,$$

where $\vec{r}_u \times \vec{r}_v$ may be replaced by $\vec{r}_v \times \vec{r}_u$ to properly match the direction of this vector with the orientation of the parametrization of C.

Example 16.7.13 Verifying Stokes' Theorem.

Considering the planar surface $f(x, y) = 7 - 2x - 2y$, let C be the curve in space that lies on this surface above the circle of radius 1 and centered at $(1, 1)$ in the xy-plane, let $\mathcal{S}$ be the planar region enclosed by C, as illustrated in Figure 16.7.14, and let $\vec{F} = \langle x + y, 2y, y^2 \rangle$. Verify Stoke's Theorem by showing $\oint_C \vec{F} \cdot \, d\vec{r} = \iint_{\mathcal{S}} (\text{curl}\, \vec{F}) \cdot \vec{n}\, dS$.

Solution. We begin by parametrizing C and then find the circulation. A unit circle centered at $(1, 1)$ can be parametrized with $x = \cos t + 1$, $y = \sin t + 1$ on $0 \le t \le 2\pi$; to put this curve on the surface f, make the z component equal $f(x, y)$: $z = 7 - 2(\cos t + 1) - 2(\sin t + 1) = 3 - 2\cos t - 2\sin t$. All together, we parametrize C with $\vec{r}(t) = \langle \cos t + 1, \sin t + 1, 3 - 2\cos t - 2\sin t \rangle$.

The circulation of $\vec{F}$ around C is

$$\begin{aligned} \oint_C \vec{F} \cdot \, d\vec{r} &= \int_0^{2\pi} \vec{F}\big(\vec{r}(t)\big) \cdot \vec{r}\,'(t)\, dt \\ &= \int_0^{2\pi} \big(2\sin^3 t - 2\cos t \sin^2 t + 3\sin^2 t - 3\cos t \sin t\big)\, dt \end{aligned}$$

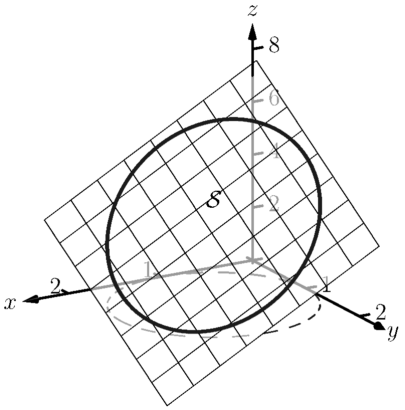


Figure 16.7.14 As given in Example 16.7.13, the surface $\mathcal{S}$ is the portion of the plane bounded by the curve

$= 3\pi.$

We now parametrize $\mathcal{S}$. (We reuse the letter "r" for our surface as this is our custom.) Based on the parametrization of C above, we describe $\mathcal{S}$ with $\vec{r}(u,v) = \langle v\cos u + 1, v\sin u + 1, 3 - 2v\cos u - 2v\sin u\rangle$, where $0 \le u \le 2\pi$ and $0 \le v \le 1$.

We leave it to the reader to confirm that $\vec{r}_u \times \vec{r}_v = \langle 2v, 2v, v\rangle$. As $0 \le v \le 1$, this vector always has a non-negative z-component, which the right-hand rule requires given the orientation of C used above. We also leave it to the reader to confirm curl $\vec{F} = \langle 2y, 0, -1\rangle$.

The surface integral of Stokes' Theorem is thus

$$\begin{aligned}\iint_{\mathcal{S}} (\text{curl}\,\vec{F})\cdot\vec{n}\,dS &= \iint_{\mathcal{S}} (\text{curl}\,\vec{F})\cdot(\vec{r}_u\times\vec{r}_v)\,dA\\ &= \int_0^1\int_0^{2\pi} \langle 2v\sin u + 2, 0, -1\rangle\cdot\langle 2v, 2v, v\rangle\,du\,dv\\ &= 3\pi,\end{aligned}$$

which matches our previous result.

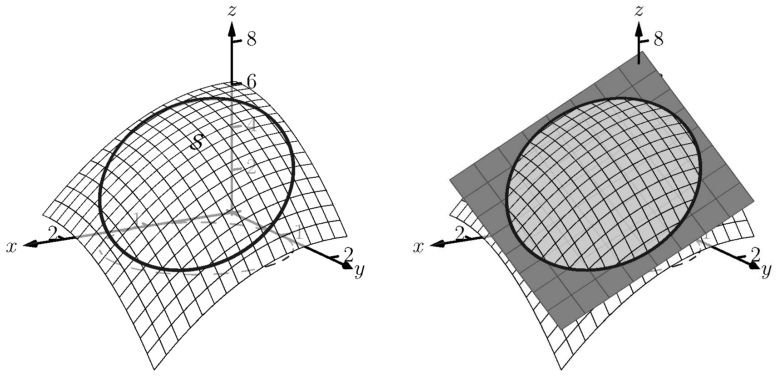
youtu.be/watch?v=IxDa_2_Eric

Figure 16.7.15 An example with Stokes' Theorem

One of the interesting results of Stokes' Theorem is that if two surfaces $\mathcal{S}_1$ and $\mathcal{S}_2$ share the same boundary, then $\iint_{\mathcal{S}_1}(\text{curl}\,\vec{F})\cdot\vec{n}\,dS = \iint_{\mathcal{S}_2}(\text{curl}\,\vec{F})\cdot\vec{n}\,dS$. That is, the value of these two surface integrals is somehow independent of the interior of the surface. We demonstrate this principle in the next example.

Example 16.7.16 Stokes' Theorem and surfaces that share a boundary.

Let C be the curve given in Example 16.7.13 and note that it lies on the surface $z = 6 - x^2 - y^2$. Let $\mathcal{S}$ be the region of this surface bounded by C, and let $\vec{F} = \langle x + y, 2y, y^2\rangle$ as in the previous example. Compute $\iint_{\mathcal{S}}(\text{curl}\,\vec{F})\cdot\vec{n}\,dS$ to show it equals the result found in the previous example.



(a) **(b)**

Figure 16.7.17 As given in Example 16.7.16, the surface $\mathcal{S}$ is the portion of the plane bounded by the curve

Solution. We begin by demonstrating that C lies on the surface $z = 6 - x^2 - y^2$. We can parametrize the x and y components of C with $x = \cos t + 1$, $y = \sin t + 1$ as before. Lifting these components to the

surface $z = 6 - x^2 - y^2$ gives the z component as $z = 6 - x^2 - y^2 = 6 - (\cos t + 1)^2 - (\sin t + 1)^2 = 3 - 2\cos t - 2\sin t$, which is the same z component as found in Example 16.7.13. Thus the curve C lies on the surface $z = 6 - x^2 - y^2$, as illustrated in Figure 16.7.17.

Since C and $\vec{F}$ are the same as in the previous example, we already know that $\oint_C \vec{F} \cdot d\vec{r} = 3\pi$. We confirm that this is also the value of $\iint_S (\text{curl } \vec{F}) \cdot \vec{n}\, dS$.

We parametrize $\mathcal{S}$ with

$$\vec{r}(u, v) = \langle v\cos u + 1, v\sin u + 1, 6 - (v\cos u + 1)^2 - (v\sin u + 1)^2 \rangle,$$

where $0 \le u \le 2\pi$ and $0 \le v \le 1$, and leave it to the reader to confirm that

$$\vec{r}_u \times \vec{r}_v = \big\langle 2v\big(v\cos u + 1\big), 2v\big(v\sin u + 1\big), v \big\rangle,$$

which also conforms to the right-hand rule with regard to the orientation of C. With curl $\vec{F} = \langle 2y, 0, -1\rangle$ as before, we have

$$\begin{aligned}
&\iint_{\mathcal{S}} (\text{curl } \vec{F}) \cdot \vec{n}\, dS \\
&= \int_0^1 \int_0^{2\pi} \langle 2v\sin u + 2, 0, -1\rangle \cdot \big\langle 2v\big(v\cos u + 1\big), 2v\big(v\sin u + 1\big), v\big\rangle \, du\, dv \\
&= 3\pi.
\end{aligned}$$

Even though the surfaces used in this example and in Example 16.7.13 are very different, because they share the same boundary, Stokes' Theorem guarantees they have equal "sum of curls" across their respective surfaces.

youtu.be/watch?v=sIqNIkmYPiE

Figure 16.7.18 Another example with Stokes' Theorem

16.7.3 A Common Thread of Calculus

We have threefold interest in each of the major theorems of this chapter: the Fundamental Theorem of Line Integrals, Green's, Stokes' and the Divergence Theorems. First, we find the beauty of their truth interesting. Second, each provides two methods of computing a desired quantity, sometimes offering a simpler method of computation.

There is yet one more reason of interest in the major theorems of this chapter. These important theorems also all share an important principle with the Fundamental Theorem of Calculus, introduced in Chapter 5.

Revisit this fundamental theorem, adopting the notation used heavily in this chapter. Let I be the interval $[a, b]$ and let $y = F(x)$ be differentiable on I, with $F'(x) = f(x)$. The Fundamental Theorem of Calculus states that

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

That is, the sum of the rates of change of a function F over an interval I can also be calculated with a certain sum of F itself on the boundary of I (in this case, at the points $x = a$ and $x = b$).

Each of the named theorems above can be expressed in similar terms. Consider the Fundamental Theorem of Line Integrals: given a function $f(x, y)$, the gradient ∇f is a type of rate of change of f. Given a curve C with initial and

terminal points A and B, respectively, this fundamental theorem states that

$$\int_C \nabla f \, ds = f(B) - f(A),$$

where again the sum of a rate of change of f along a curve C can also be evaluated by a certain sum of f at the boundary of C (i.e., the points A and B).

Green's Theorem is essentially a special case of Stokes' Theorem, so we consider just Stokes' Theorem here. Recalling that the curl of a vector field $\vec{F}$ is a measure of a rate of change of $\vec{F}$, Stokes' Theorem states that over a surface $\mathcal{S}$ bounded by a closed curve C,

$$\iint_{\mathcal{S}} \left(\text{curl}\, \vec{F}\right) \cdot \vec{n} \, dS = \oint_C \vec{F} \cdot d\vec{r},$$

i.e., the sum of a rate of change of $\vec{F}$ can be calculated with a certain sum of $\vec{F}$ itself over the boundary of $\mathcal{S}$. In this case, the latter sum is also an infinite sum, requiring an integral.

Finally, the Divergence Theorems state that the sum of divergences of a vector field (another measure of a rate of change of $\vec{F}$) over a region can also be computed with a certain sum of $\vec{F}$ over the boundary of that region. When the region is planar, the latter sum of $\vec{F}$ is an integral; when the region is spatial, the latter sum of $\vec{F}$ is a double integral.

The common thread among these theorems: the sum of a rate of change of a function over a region can be computed as another sum of the function itself on the boundary of the region. While very general, this is a very powerful and important statement.

16.7.4 Exercises

Terms and Concepts

1. What are the differences between the Divergence Theorems of Section 16.4 and this section?

2. What property of a vector field does the Divergence Theorem relate to flux?

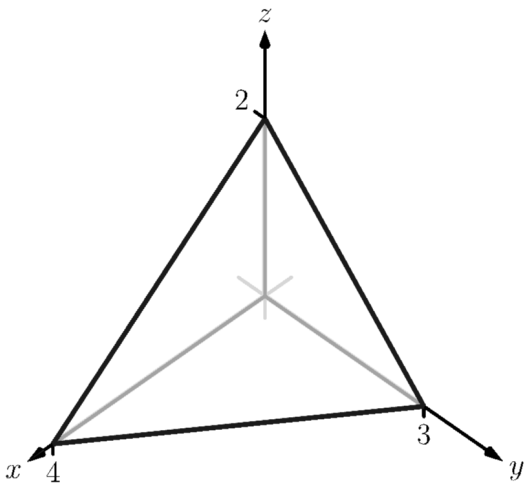
3. What property of a vector field does Stokes' Theorem relate to circulation?

4. Stokes' Theorem is the spatial version of what other theorem?

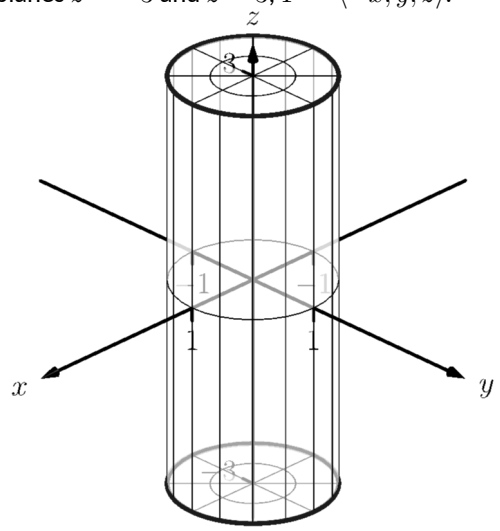
Problems

Exercise Group. In the following exercises, a closed surface $\mathcal{S}$ enclosing a domain D and a vector field $\vec{F}$ are given. Verify the Divergence Theorem on $\mathcal{S}$; that is, show $\iint_{\mathcal{S}} \vec{F} \cdot \vec{n}\, dS = \iiint_D \text{div}\, \vec{F}\, dV$.

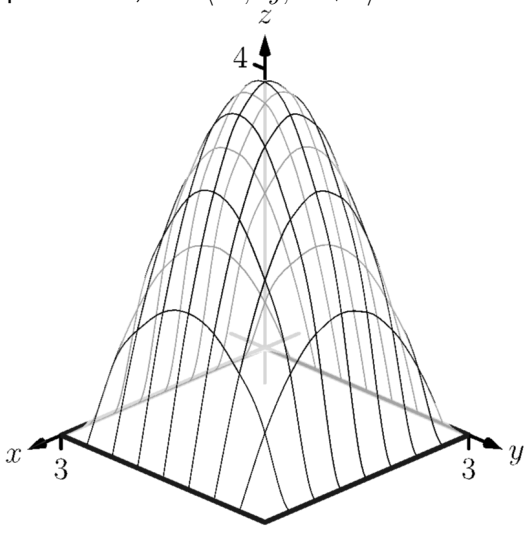
5. $\mathcal{S}$ is the surface bounding the domain D enclosed by the plane $z = 2 - x/2 - 2y/3$ and the coordinate planes in the first octant; $\vec{F} = \langle x^2, y^2, x \rangle$.



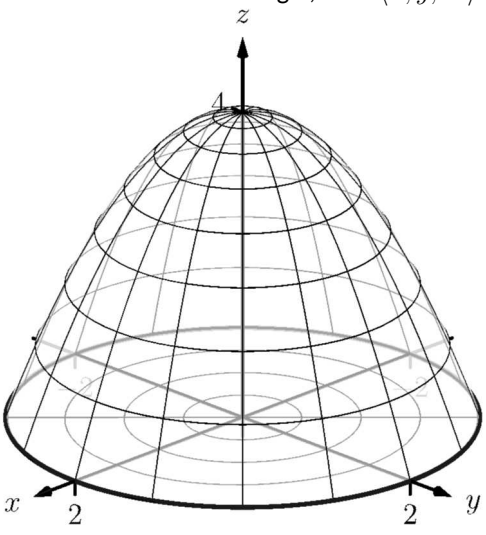
6. $\mathcal{S}$ is the surface bounding the domain D enclosed by the cylinder $x^2 + y^2 = 1$ and the planes $z = -3$ and $z = 3$; $\vec{F} = \langle -x, y, z \rangle$.



7. $\mathcal{S}$ is the surface bounding the domain D enclosed by $z = xy(3 - x)(3 - y)$ and the plane $z = 0$; $\vec{F} = \langle 3x, 4y, 5z + 1 \rangle$.

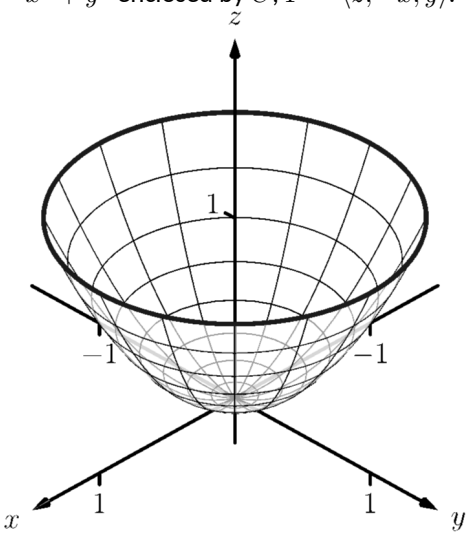


8. $\mathcal{S}$ is the surface composed of $\mathcal{S}_1$, the paraboloid $z = 4 - x^2 - y^2$ for $z \geq 0$, and $\mathcal{S}_2$, the disk of radius 2 centered at the origin; $\vec{F} = \langle x, y, z^2 \rangle$.

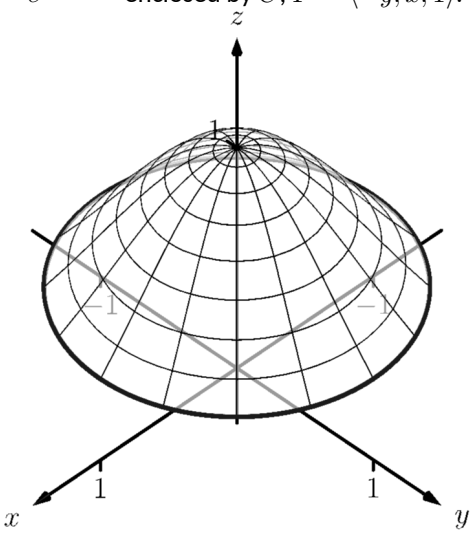


Exercise Group. In the following exercises, a closed curve C that is the boundary of a surface $\mathcal{S}$ is given along with a vector field $\vec{F}$. Verify Stokes' Theorem on C; that is, show $\oint_C \vec{F} \cdot d\vec{r} = \iint_{\mathcal{S}} (\operatorname{curl} \vec{F}) \cdot \vec{n}\, dS$.

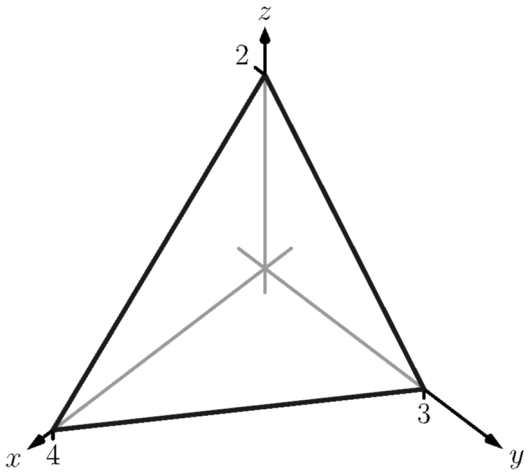
9. C is the curve parametrized by $\vec{r}(t) = \langle \cos t, \sin t, 1 \rangle$ and $\mathcal{S}$ is the portion of $z = x^2 + y^2$ enclosed by C; $\vec{F} = \langle z, -x, y \rangle$.



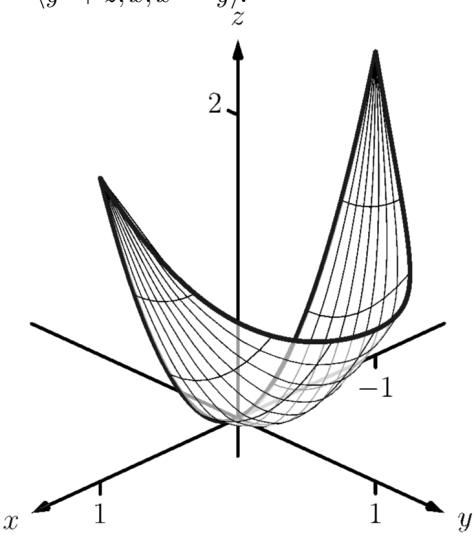
10. C is the curve parametrized by $\vec{r}(t) = \langle \cos t, \sin t, e^{-1} \rangle$ and $\mathcal{S}$ is the portion of $z = e^{-x^2-y^2}$ enclosed by C; $\vec{F} = \langle -y, x, 1 \rangle$.



11. C is the curve that follows the triangle with vertices at $(0, 0, 2)$, $(4, 0, 0)$ and $(0, 3, 0)$, traversing the the vertices in that order and returning to $(0, 0, 2)$, and $\mathcal{S}$ is the portion of the plane $z = 2 - x/2 - 2y/3$ enclosed by C; $\vec{F} = \langle y, -z, y \rangle$.

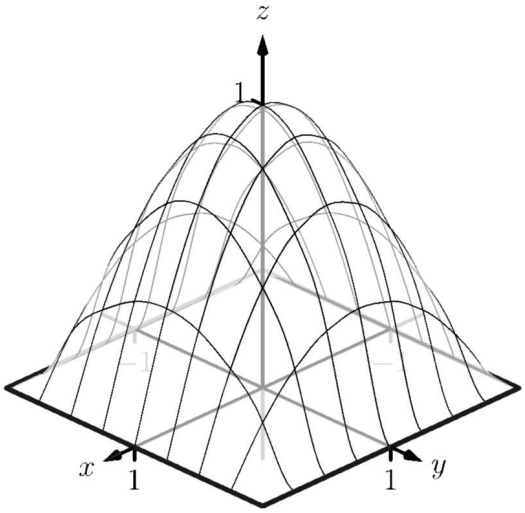


12. C is the curve whose x and y coordinates follow the parabola $y = 1 - x^2$ from $x = 1$ to $x = -1$, then follow the line from $(-1, 0)$ back to $(1, 0)$, where the z coordinates of C are determined by $f(x, y) = 2x^2 + y^2$, and $\mathcal{S}$ is the portion of $z = 2x^2 + y^2$ enclosed by C; $\vec{F} = \langle y^2 + z, x, x^2 - y \rangle$.

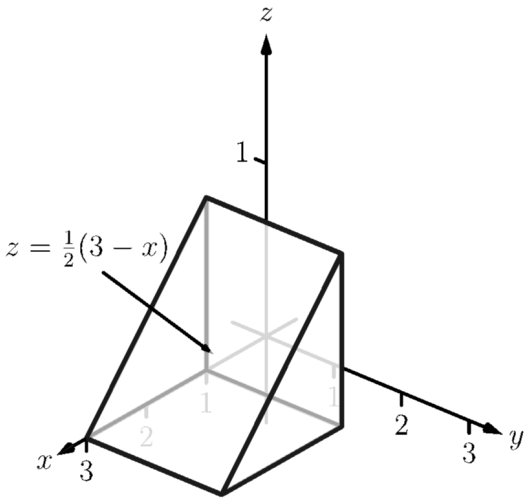


Exercise Group. In the following exercises, a closed surface $\mathcal{S}$ and a vector field $\vec{F}$ are given. Find the outward flux of $\vec{F}$ over $\mathcal{S}$ either through direct computation or through the Divergence Theorem.

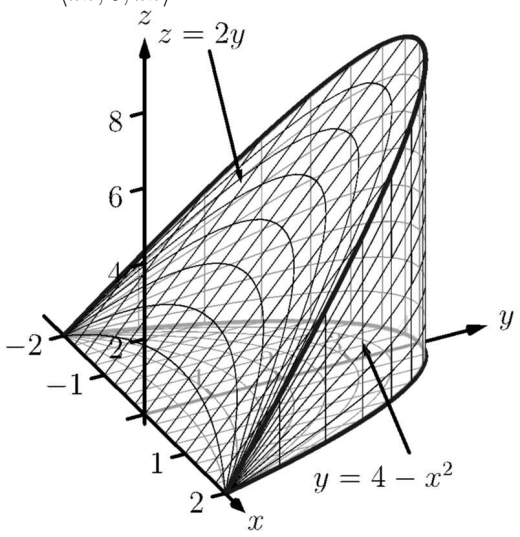
13. $\mathcal{S}$ is the surface formed by the intersections of $z = 0$ and $z = (x^2 - 1)(y^2 - 1)$; $\vec{F} = \langle x^2 + 1, yz, xz^2 \rangle$.



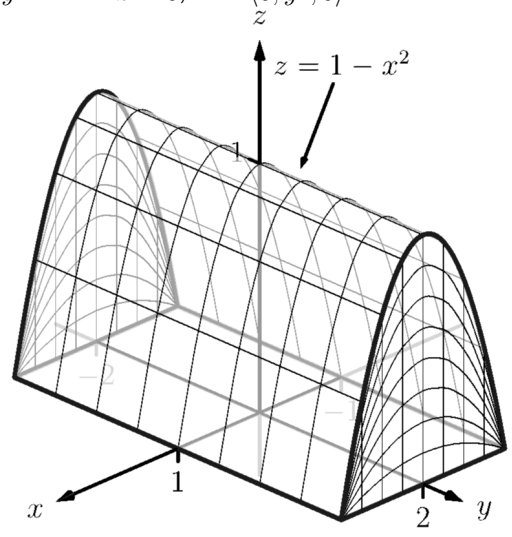
14. $\mathcal{S}$ is the surface formed by the intersections of the planes $z = \frac{1}{2}(3 - x)$, $x = 1$, $y = 0$, $y = 2$ and $z = 0$; $\vec{F} = \langle x, y^2, z \rangle$.



15. $\mathcal{S}$ is the surface formed by the intersections of the planes $z = 2y$, $y = 4 - x^2$ and $z = 0$; $\vec{F} = \langle xz, 0, xz \rangle$.

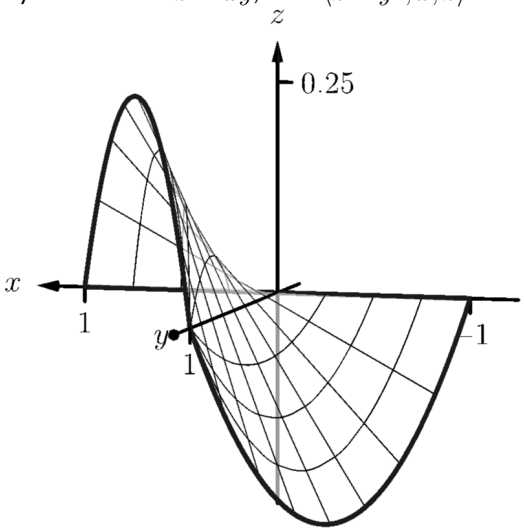


16. $\mathcal{S}$ is the surface formed by the intersections of the cylinder $z = 1 - x^2$ and the planes $y = -2$, $y = 2$ and $z = 0$; $\vec{F} = \langle 0, y^3, 0 \rangle$.

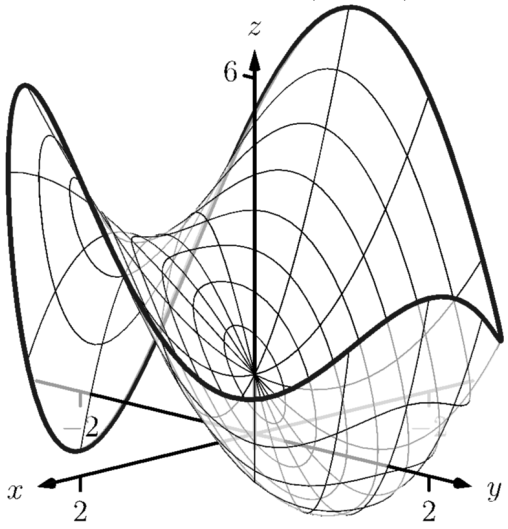


Exercise Group. In the following exercises, a closed curve C that is the boundary of a surface $\mathcal{S}$ is given along with a vector field $\vec{F}$. Find the circulation of $\vec{F}$ around C either through direct computation or through Stokes' Theorem.

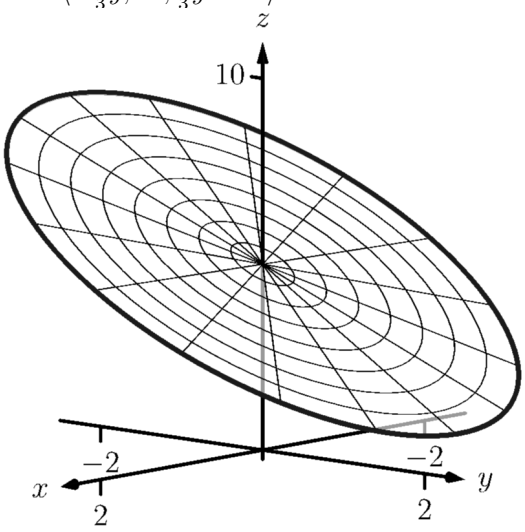
17. C is the curve whose x- and y-values are determined by the three sides of a triangle with vertices at $(-1,0)$, $(1,0)$ and $(0,1)$, traversed in that order, and the z-values are determined by the function $z = xy$; $\vec{F} = \langle z - y^2, x, z\rangle$.



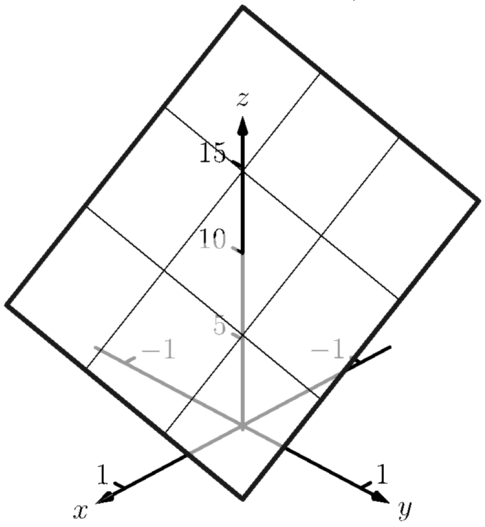
18. C is the curve whose x- and y-values are given by $\vec{r}(t) = \langle 2\cos t, 2\sin t\rangle$ and the z-values are determined by the function $z = x^2 + y^3 - 3y + 1$; $\vec{F} = \langle -y, x, z\rangle$.



19. C is the curve whose x- and y-values are given by $\vec{r}(t) = \langle \cos t, 3\sin t\rangle$ and the z-values are determined by the function $z = 5 - 2x - y$; $\vec{F} = \langle -\frac{1}{3}y, 3x, \frac{2}{3}y - 3x\rangle$.



20. C is the curve whose x- and y-values are sides of the square with vertices at $(1,1)$, $(-1,1)$, $(-1,-1)$ and $(1,-1)$, traversed in that order, and the z-values are determined by the function $z = 10 - 5x - 2y$; $\vec{F} = \langle 5y^2, 2y^2, y^2\rangle$.



Exercise Group. The following exercises are designed to challenge your understanding and require no computation.

21. Let $\mathcal{S}$ be any closed surface enclosing a domain D. Consider $\vec{F}_1 = \langle x, 0, 0\rangle$ and $\vec{F}_2 = \langle y, y^2, z - 2yz\rangle$.
These fields are clearly very different. Why is it that the total outward flux of each field across $\mathcal{S}$ is the same?

22.

(a) Green's Theorem can be used to find the area of a region enclosed by a curve by evaluating a line integral with the appropriate choice of vector field $\vec{F}$. What condition on $\vec{F}$ makes this possible?

(b) Likewise, Stokes' Theorem can be used to find the surface area of a region enclosed by a curve in space by evaluating a line integral with the appropriate choice of vector field $\vec{F}$. What condition on $\vec{F}$ makes this possible?

23. The Divergence Theorem establishes equality between a particular double integral and a particular triple integral. What types of circumstances would lead one to choose to evaluate the triple integral over the double integral?

24. Stokes' Theorem establishes equality between a particular line integral and a particular double integral. What types of circumstances would lead one to choose to evaluate the double integral over the line integral?

Appendices

Appendix A

Answers to Selected Exercises

I · Math 1565: Accelerated Calculus I

I · Math 1565: Accelerated Calculus I
1 · Limits

I · Math 1565: Accelerated Calculus I
1 · Limits
1.1 · An Introduction To Limits

1.1 · Exercises

Terms and Concepts

1.1.2. an indeterminate form

1.1.3. False

1.1.6. 1

Problems

1.1.7. 5

1.1.8. 3

1.1.9. DNE

1.1.10. $\frac{2}{3}$

1.1.11. -4

1.1.12. DNE or ∞

1.1.13. DNE

1.1.14. 6

1.1.15. 1

1.1.16. DNE

1.1.17. 1

1.1.18. DNE

1.1.19. DNE

1.1.20. 1

1.1.21. -7

1.1.22. 9

1.1.23. 5

1.1.24. -0.111111

1.1.25. 29

1.1.26. 0.2

1.1.27. -1

1.1.28. 0

1.2 · Epsilon-Delta Definition of a Limit

1.2 · Exercises

III · Math 2575: Accelerated Calculus III

III · Math 2575: Accelerated Calculus III

12 · Vectors

III · Math 2575: Accelerated Calculus III
12 · Vectors

12.1 · Introduction to Cartesian Coordinates in Space

12.1 · Exercises

Terms and Concepts

12.1.2. (a) line

(b) plane

12.1.4. Hyperbolic Paraboloid

Problems

12.1.7. (a) $\sqrt{6}$

(b) $\sqrt{17}$

(c) $\sqrt{11}$

(d) do

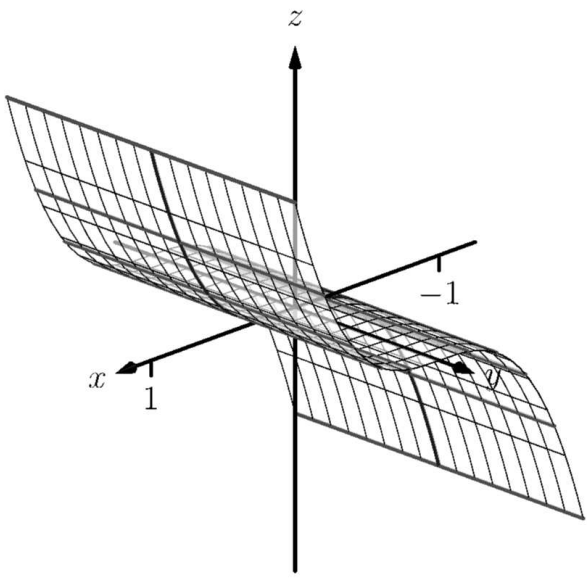
12.1.9. (a) $(4, -1, 0)$

(b) 3

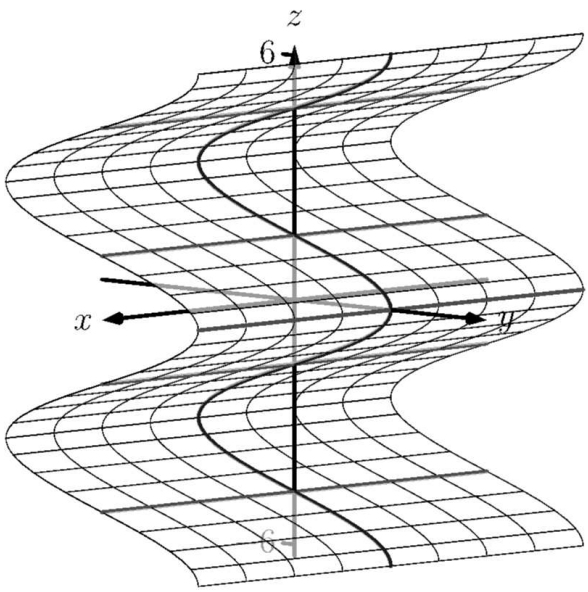
12.1.10. (a) $(-2, 1, 2)$

(b) $\sqrt{5}$

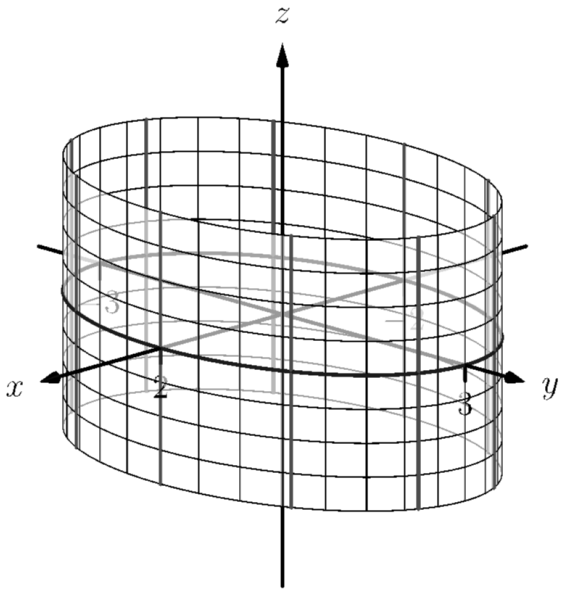
12.1.15.



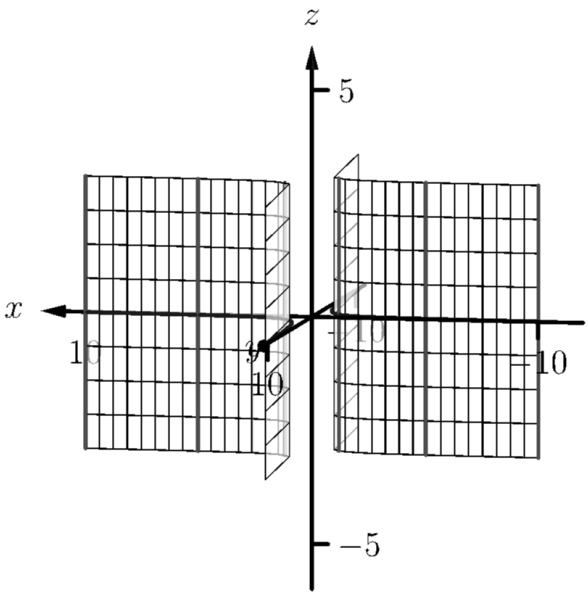
12.1.16.



12.1.17.



12.1.18.



12.1.19. $x^2 + z^2 = \left(\frac{1}{1+y^2}\right)^2$

12.1.20. $y^2 + z^2 = x^4$

12.1.21. $x^2 + y^2 = z$

12.1.22. $x^2 + y^2 = \frac{1}{z^2}$

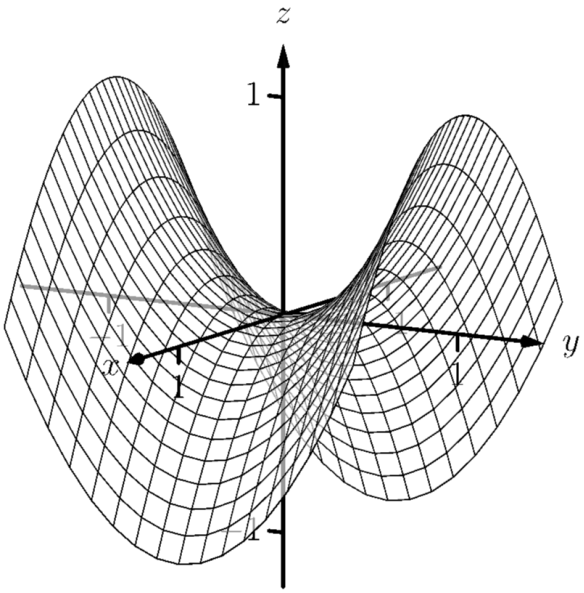
12.1.23. (a)$x = y^2 + \dfrac{z^2}{9}$

12.1.24. (b) $x^2 - y^2 + z^2 = 0$

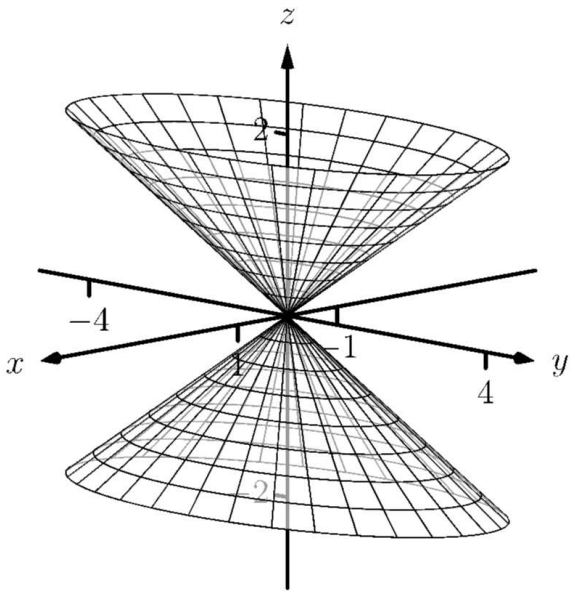
12.1.25. (b) $x^2 + \dfrac{y^2}{9} + \dfrac{z^2}{4} = 1$

12.1.26. (a) $y^2 - x^2 - z^2 = 1$

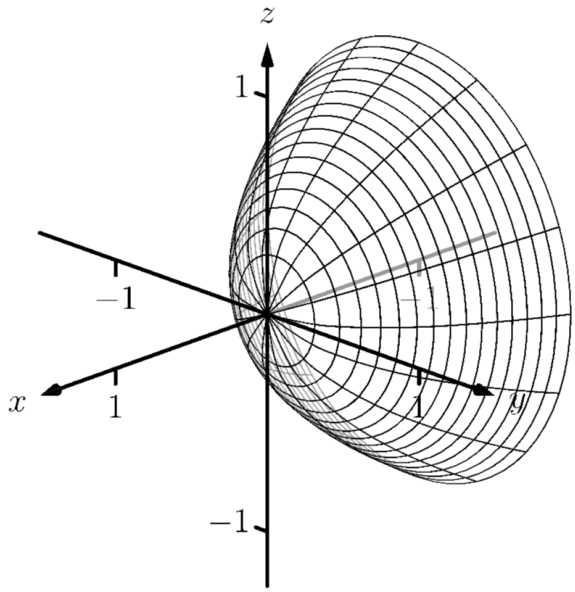
12.1.27.



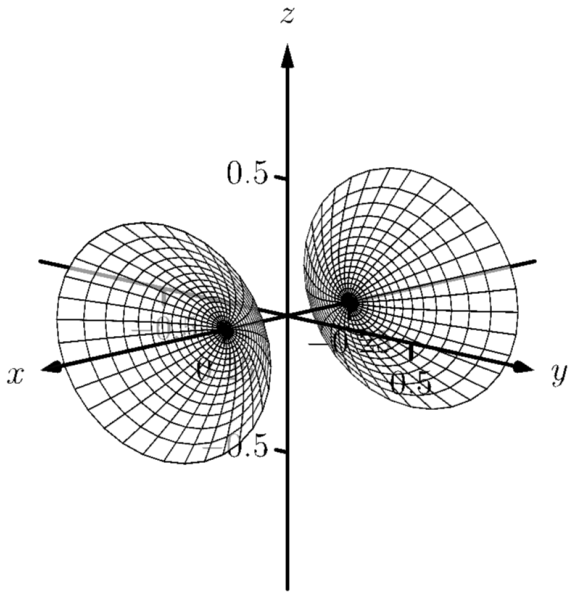
12.1.28.



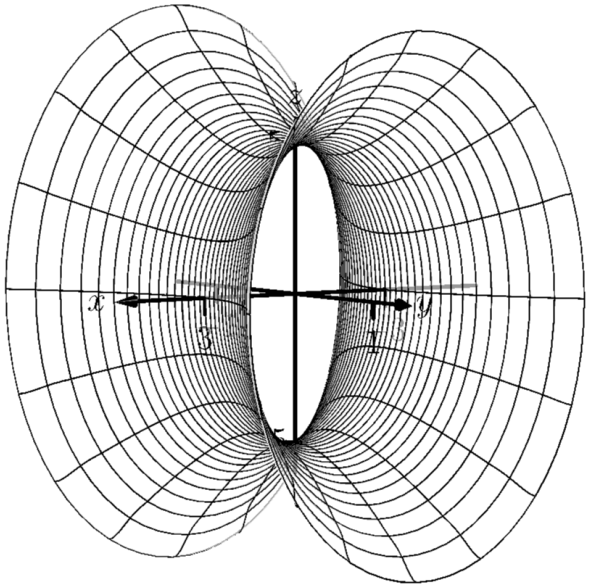
12.1.29.



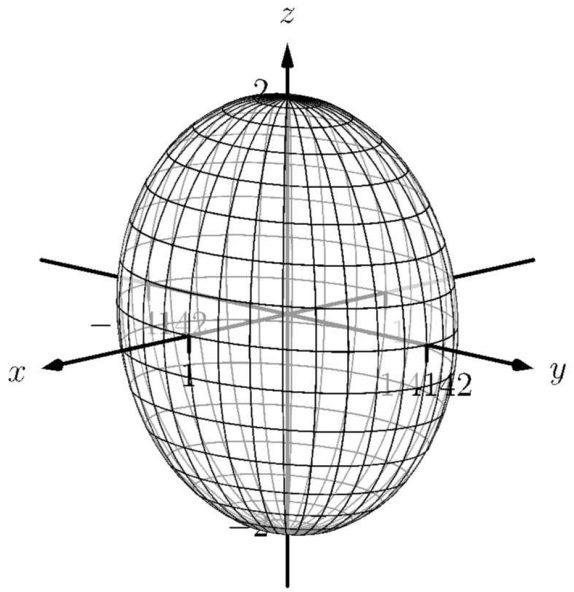
12.1.30.



12.1.31.



12.1.32.



12.2 · An Introduction to Vectors

12.2 · Exercises

Terms and Concepts

12.2.4. Direction

Problems

12.2.7.

(a) $\langle 1, 6\rangle$

(b) $\boldsymbol{i} + 6\boldsymbol{j}$

12.2.8.

(a) $\langle 4, -4\rangle$

12.2.9.

(a) $\langle 6, -1, 6\rangle$

(b) $6\boldsymbol{i} - \boldsymbol{j} + 6\boldsymbol{k}$

12.2.10.

(a) $\langle 2, 2, 0\rangle$

(b) $2\boldsymbol{i} + 2\boldsymbol{j}$

12.2.11.

(a) $\vec{u} + \vec{v} = \langle 2, -1\rangle$; $\vec{u} - \vec{v} = \langle 0, -3\rangle$; $2\vec{u} - 3\vec{v} = \langle -1, -7\rangle$.

(c) $\vec{x} = \langle 1/2, 2\rangle$.

12.2.12.

(a) $\vec{u} + \vec{v} = \langle 3, 2, 1\rangle$; $\vec{u} - \vec{v} = \langle -1, 0, -3\rangle$; $\pi\vec{u} - \sqrt{2}\vec{v} = \left\langle \pi - 2\sqrt{2}, \pi - \sqrt{2}, -\pi - 2\sqrt{2}\right\rangle$.

(c) $\vec{x} = \langle -1, 0, -3\rangle$.

12.2.17. (a) $\sqrt{5}$
(b) $\sqrt{13}$
(c) $\sqrt{26}$
(d) $\sqrt{10}$

12.2.18. (a) $\sqrt{17}$
(b) $\sqrt{3}$
(c) $\sqrt{14}$
(d) $\sqrt{26}$

12.2.19. (a) $\sqrt{5}$
(b) $3\sqrt{5}$
(c) $2\sqrt{5}$
(d) $4\sqrt{5}$

12.2.20. (a) 7
(b) 35
(c) 42
(d) 28

12.2.22. $\left\langle \frac{3}{\sqrt{58}}, \frac{7}{\sqrt{58}}\right\rangle$

12.2.23. $\langle 0.6, 0.8\rangle$

12.2.24. $\left\langle \frac{1}{3}, \frac{-2}{3}, \frac{2}{3}\right\rangle$

12.2.25. $\left\langle \frac{1}{\sqrt{3}}, \frac{-1}{\sqrt{3}}, \frac{1}{\sqrt{3}}\right\rangle$

12.2.26. $\left\langle \cos\left(\frac{50\pi}{180}\right), \sin\left(\frac{50\pi}{180}\right)\right\rangle$

12.2.27. $\left\langle \frac{-1}{2}, \frac{\sqrt{3}}{2}\right\rangle$

12.3 • The Dot Product
12.3 • Exercises

Terms and Concepts

12.3.1. Scalar

Problems

12.3.5. -22
12.3.6. 33
12.3.7. 3
12.3.8. 0
12.3.9. not defined
12.3.10. 0

12.3.11. Answers will vary.

12.3.12. Answers will vary.

12.3.13. $\cos^{-1}\left(\frac{3}{\sqrt{10}}\right)$

12.3.14. $\cos^{-1}\left(\frac{-1}{\sqrt{170}}\right)$

12.3.15. $\frac{\pi}{4}$

12.3.16. $\frac{\pi}{2}$

12.3.17. (a) $\langle -7, 4\rangle$

(b) $\langle 4, 7\rangle$

12.3.18. (a) $\langle 5, 3\rangle$

(b) $\langle -3, 5\rangle$

12.3.19. (a) $\langle 1, 0, -1\rangle$

(b) $\langle 1, 1, 1\rangle$

12.3.20. (a) $\langle 2, 1, 0\rangle$

(b) $\langle 1, -2, 3\rangle$

12.3.21. $\left\langle \frac{-5}{10}, \frac{15}{10}\right\rangle$

12.3.22. $\left\langle \frac{20}{10}, \frac{60}{10}\right\rangle$

12.3.23. $\left\langle \frac{-1}{2}, \frac{-1}{2}\right\rangle$

12.3.24. $\left\langle \frac{0}{13}, \frac{0}{13}\right\rangle$

12.3.25. $\left\langle \frac{14}{14}, \frac{28}{14}, \frac{42}{14}\right\rangle$

12.3.26. $\left\langle \frac{12}{9}, \frac{12}{9}, \frac{6}{9}\right\rangle$

12.3.27. (a) $\left\langle \frac{-5}{10}, \frac{15}{10}\right\rangle$

(b) $\left\langle \frac{15}{10}, \frac{5}{10}\right\rangle$

12.3.28. (a) $\left\langle \frac{20}{10}, \frac{60}{10}\right\rangle$

(b) $\left\langle \frac{30}{10}, \frac{-10}{10}\right\rangle$

12.3.29. (a) $\left\langle \frac{-1}{2}, \frac{-1}{2}\right\rangle$

(b) $\left\langle \frac{-5}{2}, \frac{5}{2}\right\rangle$

12.3.30. (a) $\left\langle \frac{0}{13}, \frac{0}{13}\right\rangle$

(b) $\left\langle \frac{-39}{13}, \frac{26}{13}\right\rangle$

12.3.31. (a) $\left\langle \frac{14}{14}, \frac{28}{14}, \frac{42}{14}\right\rangle$

(b) $\left\langle \frac{0}{14}, \frac{42}{14}, \frac{-28}{14}\right\rangle$

12.3.32. (a) $\left\langle \frac{12}{9}, \frac{12}{9}, \frac{6}{9}\right\rangle$

(b) $\left\langle \frac{15}{9}, \frac{-21}{9}, \frac{12}{9}\right\rangle$

12.3.33. 1.96lb

12.3.34. 5lb

12.3.35. 141.42ft–lb

12.3.36. 196.96ft–lb

12.3.37. 500ft–lb

12.3.39. 500ft–lb

12.4 • The Cross Product

12.4 • Exercises

Terms and Concepts

12.4.1. vector

12.4.2. right hand rule

12.4.3. “Perpendicular” is one answer.

12.4.4. True

12.4.5. Torque

12.4.6. T

Problems

12.4.7. $\langle 12, -15, 3\rangle$

12.4.8. $\langle 11, 1, -17\rangle$

12.4.9. $\langle -5, -31, 27\rangle$

12.4.10. $\langle 47, -36, -44\rangle$

12.4.11. $\langle 0, -2, 0\rangle$

12.4.12. $\langle 0, 0, 0\rangle$

12.4.13. $\vec{u} \times \vec{v} = \langle 0, 0, ad - bc \rangle$

12.4.14. $\boldsymbol{k}$

12.4.15. $-\boldsymbol{j}$

12.4.16. $\boldsymbol{i}$

12.4.17. Answers will vary.

12.4.18. Answers will vary.

12.4.19. 5

12.4.20. 21

12.4.21. 0

12.4.22. 5

12.4.23. $\sqrt{14}$

12.4.24. $\sqrt{230}$

12.4.25. 3

12.4.26. 6

12.4.27. $\frac{5\sqrt{2}}{2}$

12.4.28. $3\sqrt{30}$

12.4.29. 1

12.4.30. $\frac{5}{2}$

12.4.31. 7

12.4.32. $4\sqrt{14}$

12.4.33. 2

12.4.34. 15

12.4.35.
$\langle 0.408248, 0.408248, -0.816497 \rangle$ or $\langle -0.408248, -0.408248, 0.816497 \rangle$

12.4.36.
$\langle -0.436436, 0.218218, 0.872872 \rangle$ or $\langle 0.436436, -0.218218, -0.87287$

12.4.37. $\langle 0, 1, 0 \rangle$ or $\langle 0, -1, 0 \rangle$

12.4.38. $\left\langle \frac{2}{\sqrt{5}}, \frac{1}{\sqrt{5}}, 0 \right\rangle$

12.4.39. 87.5ft–lb

12.4.40. $43.75\sqrt{3} \approx 75.78$ft–lb

12.4.41. $200/3 \approx 66.67$ft–lb

12.4.42. 11.58ft–lb

12.5 · Lines

12.5 · Exercises

Terms and Concepts

12.5.1. A point on the line and the direction of the line.

12.5.2. parallel

12.5.3. parallel, skew

12.5.4. Answers will vary

Problems

12.5.6. (a) $(6, 1, 7) + t\langle -3, 2, 5 \rangle$

(b) $x = 6 - 3t, y = 1 + 2t, z = 7 + 5t$

(c) $\frac{x-6}{-3} = \frac{y-1}{2} = \frac{z-7}{5}$

12.5.8. (a) $(1, -2, 3) + t\langle 4, 7, 2 \rangle$

(b) $x = 1 + 4t, y = -2 + 7t, z = 3 + 2t$

(c) $\frac{x-1}{4} = \frac{y+2}{7} = \frac{z-3}{2}$

12.5.10. (a) $(5, 1, 9) + t\langle 0, -1, 0 \rangle$

(b) $x = 5, y = 1 - t, z = 9$

(c) DNE

12.5.11. (a) $(7, 2, -1) + t\langle 1, -1, 2 \rangle$

(b) $x = 7 + t, y = 2 - t, z = -1 + 2t$

(c) $x - 7 = 2 - y = \frac{z+1}{2}$

12.5.12. (a) $(2,2,3)+t\langle 5,-1,-3\rangle$
(b) $x=2+5t, y=2-t, z=3-3t$
(c) $\frac{x-2}{5}=-(y-2)=\frac{-(z-3)}{3}$

12.5.14. (a) $(-2,5)+t\langle 0,1\rangle$
(b) $x=-2, y=5+t$
(c) DNE

12.5.15. parallel

12.5.16. $(12,3,7)$

12.5.18. same

12.5.19. skew

12.5.20. parallel

12.5.22. skew

12.5.23. $\sqrt{41}/3$

12.5.24. $3\sqrt{2}$

12.5.25. $5\sqrt{2}/2$

12.5.26. 5

12.5.27. $3/\sqrt{2}$

12.5.28. 2

12.6 · Planes
12.6 · Exercises

Terms and Concepts

12.6.1. A point in the plane and a normal vector (i.e., a direction orthogonal to the plane).

12.6.2. A normal vector is orthogonal to the plane.

Problems

12.6.3. Answers will vary.

12.6.4. $(-2,9,0),(2,9,3)$

12.6.5. Answers will vary.

12.6.6. $(0,-2,6),(1,-2,6)$

12.6.8. (a) $2(y-3)+4(z-5)=0$
(b) $2y+4z=26$

12.6.10. (a) $-5(x-5)+3(y-3)+2(z-8)=0$
(b) $-5x+3y+2z=0$

12.6.12. (a) $3(x-5)+3(z-3)=0$
(b) $3x+3z=24$

12.6.14. (a) $2(x-1)+y-1-3(z-1)=0$
(b) $2x+y-3z=0$

12.6.16. (a) $4(x-5)-2(y-7)-2(z-3)=0$
(b) $4x-2y-2z=0$

12.6.17. (a) $x-5+y-7+z-3=0$
(b) $x+y+z=15$

12.6.18. (a) $4(x-4)+y-1+z-1=0$
(b) $4x+y+z=18$

12.6.19. (a) $3(x+4)+8(y-7)-10(z-2)=0$
(b) $3x+8y-10z=24$

12.6.20. (a) $x-1=0$
(b) $x=1$

12.6.22. $(1,3,3.5)+t\langle 20,2,-26\rangle$

12.6.24. $(3,1,1)$

12.6.26. the entire line

12.6.27. $\sqrt{5/7}$

12.6.28. $\frac{8}{\sqrt{21}}$

12.6.29. $1/\sqrt{3}$

12.6.30. 3

13 · Vector Valued Functions

13 · Vector Valued Functions
13.1 · Vector-Valued Functions
13.1 · Exercises

Terms and Concepts

13.1.1. parametric equations

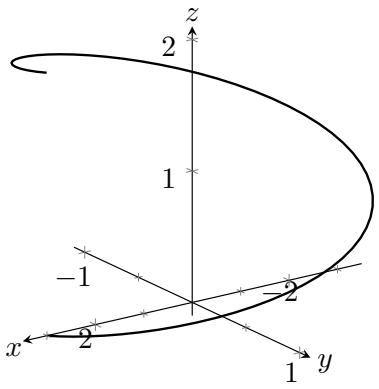
13.1.2. vectors

13.1.3. displacement

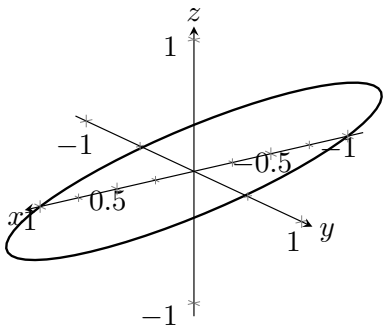
13.1.4. displacement

Problems

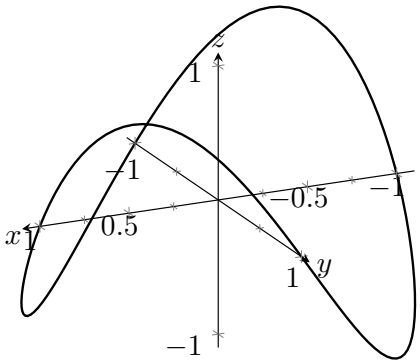
13.1.14.



13.1.15.



13.1.16.



13.1.17. $|t|\sqrt{1+t^2}$

13.1.19. $\sqrt{4+t^2}$

13.1.21. $\langle 2\cos(t)+1, 2\sin(t)+2\rangle$

13.1.25. $\langle t+2, 5t+3\rangle$

13.1.27. Specific forms may vary, though most direct solutions are
$\vec{r}(t) = \langle 1,2,3\rangle + t\,\langle 3,3,3\rangle$ and
$\vec{r}(t) = \langle 3t+1, 3t+2, 3t+3\rangle$.

13.1.28. Specific forms may vary, though most direct solutions are
$\vec{r}(t) = \langle 1,2\rangle + t\,\langle 3,2\rangle$ and
$\vec{r}(t) = \langle 3t+1, 2t+2\rangle$.

13.1.29. $\langle 2\cos(t)\,,\, 2\sin(t)\,,\, 2t\rangle$

13.1.31. $\langle 1, 0\rangle$

13.1.33. $\langle 0,0,1\rangle$

13.2 · Calculus and Vector-Valued Functions
13.2 · Exercises

Terms and Concepts

13.2.1. component

13.2.2. displacement

13.2.4. A scalar-vector product, a dot product and a cross product.

Problems

13.2.5. $\langle 11, 74, \sin(5)\rangle$

13.2.6. $\left\langle e^3, 0\right\rangle$

13.2.7. $\langle 1, e\rangle$

13.2.8. $\langle 2t, 1, 0\rangle$

13.2.9. $(-\infty, 0)\cup(0,\infty)$

13.2.10. $(0,\infty)$

13.2.11. $\left\langle -\sin(t)\,,\, e^t,\, \frac{1}{t}\right\rangle$

13.2.12. $\vec{r}\,'(t) = \left\langle -1/t^2, 5/(3t+1)^2, \sec^2(t)\right\rangle$

13.2.13. $\left\langle 2t\sin(t) + t^2\cos(t)\,,\, 6t^2+10t\right\rangle$

13.2.14. $\left(t^2+1\right)\cos(t) + 2t\sin(t) + 4t + 3$

13.2.15.
$\left\langle -1, \cos(t) - 2t, 6t^2+10t+2+\cos(t)-\sin(t)-t\cos(t)\right\rangle$

13.2.16. $\vec{r}\,'(t) = \langle \sinh t, \cosh t\rangle$

13.2.21. $\langle 2+3t, t\rangle$

13.2.22. $\left(\frac{3\sqrt{2}}{2}, \frac{\sqrt{2}}{2}\right) + t\left\langle \frac{-3\sqrt{2}}{2}, \frac{\sqrt{2}}{2}\right\rangle$

13.2.23. $\ell(t) = \langle -3, 0, \pi\rangle + t\,\langle 0, -3, 1\rangle$

13.2.24. $(1,0,0) + t\langle 1,1,1\rangle$

13.2.26. 1

13.2.28. $1, -1$

13.2.32. Both derivatives return $\left\langle 6t^5, 3t^2, 0\right\rangle$

13.2.33. $\left\langle \frac{1}{4}t^4, \sin(t), te^t - e^t\right\rangle + \vec{C}$

13.2.34. $\left\langle \tan^{-1}(t), \tan(t)\right\rangle + \vec{C}$

13.2.35. $\langle -2, 0\rangle$

13.2.36. $\langle 4, -4\rangle$

13.2.37. $\left\langle \frac{t^2}{2} + 2, -\cos(t) + 3\right\rangle$

13.2.39. $\left\langle \frac{t^4}{12} + t + 4,\, \frac{t^3}{6} + 2t + 5,\, \frac{t^2}{2} + 3t + 6\right\rangle$

13.2.41. $2\cdot 3.60555\pi$

13.2.42. 10π

13.2.43. $\frac{1}{54}\left(22^{\frac{3}{2}} - 8\right)$

13.2.44. $\sqrt{2}(1-e^{-1})$

13.3 · The Calculus of Motion
13.3 · Exercises

Terms and Concepts

13.3.4. arc length

Problems

13.3.7. $\vec{v}(t) = \langle 2, 5, 0\rangle$, $\vec{a}(t) = \langle 0, 0, 0\rangle$

13.3.8. (a) $\langle 6t - 2, -2t + 1\rangle$
(b) $\langle 6, -2\rangle$

13.3.10. (a) $\left\langle \frac{1}{10}, \sin(t), \cos(t)\right\rangle$
(b) $\langle 0, \cos(t), -\sin(t)\rangle$

13.3.16. (a) $|t|\sqrt{9t^2 - 12t + 8}$
(b) 0
(c) -1

13.3.19. (a) $|\sec(t)|\sqrt{\tan^2(t) + \sec^2(t)}$
(b) 0
(c) $\frac{\pi}{4}$

13.3.22. (a) $\sqrt{8t^2 + 3}$
(b) 0
(c) 1

13.3.30. $\left\langle t^2 - t + 5, \frac{3t^2}{2} - t - \frac{5}{2}\right\rangle$

13.3.32. $\langle 10t, -16t^2 + 50t\rangle$

13.3.34. (a) $\langle -10, 0\rangle$
(b) 5π
(c) $\left\langle \frac{-10}{\pi}, 0\right\rangle$
(d) 5

13.3.36. (a) $\langle 10, 20, -10\rangle$
(b) $10\sqrt{6}$
(c) $\langle 1, 2, -1\rangle$
(d) $\sqrt{6}$

13.3.38.

(a) $t = \sin^{-1}(3/20)/(8\pi) + n/4 \approx 0.006 + n/4$, where n is an integer

(b) $\|\vec{r}\,'(t)\| = 24\pi \approx 51.4$ ft/s

(c) 0.27 radians, or $15.69°$

13.3.39.

(a) 0.013 radians

(b) 11.7 ft

13.4 · Unit Tangent and Normal Vectors

13.4 · Exercises

Terms and Concepts

13.4.1. 1
13.4.2. 0
13.4.3. $\vec{T}(t)$ and $\vec{N}(t)$.
13.4.4. the speed

Problems

13.4.5. $\vec{T}(t) = \left\langle \frac{4t}{\sqrt{20t^2-4t+1}}, \frac{2t-1}{\sqrt{20t^2-4t+1}} \right\rangle$;
$\vec{T}(1) = \langle 4/\sqrt{17}, 1/\sqrt{17} \rangle$

13.4.6. (a) $\left\langle \frac{1}{\sqrt{1+\sin^2(t)}}, \frac{-\sin(t)}{\sqrt{1+\sin^2(t)}} \right\rangle$

(b) $\left\langle \sqrt{\frac{2}{3}}, \frac{-1}{\sqrt{3}} \right\rangle$

13.4.8. (a) $\langle -\sin(t), \cos(t) \rangle$

(b) $\langle 0, -1 \rangle$

13.4.9. $(2,0) + t\left\langle \frac{4}{\sqrt{17}}, \frac{1}{\sqrt{17}} \right\rangle$

13.4.10. $\left(\frac{\pi}{4}, \frac{\sqrt{2}}{2}\right) + t\left\langle \sqrt{\frac{2}{3}}, \frac{-1}{\sqrt{3}} \right\rangle$

13.4.12. $(-1,0) + t\langle 0, -1 \rangle$

13.4.13. $\vec{T}(t) = \langle -\sin(t), \cos(t) \rangle$;
$\vec{N}(t) = \langle -\cos(t), -\sin(t) \rangle$

13.4.14. $\vec{T}(t) = \left\langle \frac{1}{\sqrt{1+4t^2}}, \frac{2t}{\sqrt{1+4t^2}} \right\rangle$;
$\vec{N}(t) = \left\langle -\frac{2t}{\sqrt{1+4t^2}}, \frac{1}{\sqrt{1+4t^2}} \right\rangle$

13.4.15.
$\vec{T}(t) = \left\langle -\frac{\sin(t)}{\sqrt{4\cos^2(t)+\sin^2(t)}}, \frac{2\cos(t)}{\sqrt{4\cos^2(t)+\sin^2(t)}} \right\rangle$;
$\vec{N}(t) = \left\langle -\frac{2\cos(t)}{\sqrt{4\cos^2(t)+\sin^2(t)}}, -\frac{\sin(t)}{\sqrt{4\cos^2(t)+\sin^2(t)}} \right\rangle$

13.4.16. $\vec{T}(t) = \left\langle \frac{e^t}{\sqrt{e^{2t}+e^{-2t}}}, -\frac{e^{-t}}{\sqrt{e^{2t}+e^{-2t}}} \right\rangle$;
$\vec{N}(t) = \left\langle \frac{e^{-t}}{\sqrt{e^{2t}+e^{-2t}}}, \frac{e^t}{\sqrt{e^{2t}+e^{-2t}}} \right\rangle$

13.4.22. $\left\langle -\cos(t), \frac{-3}{5}\sin(t), \frac{-4}{5}\sin(t) \right\rangle$

13.4.24. $\langle -\cos(at), -\sin(at), 0 \rangle$

13.4.26. (a) $\frac{\frac{-2}{t^5}}{\sqrt{1+\frac{1}{t^4}}}$

(b) $\frac{\frac{2}{t^3}}{\sqrt{1+\frac{1}{t^4}}}$

(c) $-\sqrt{2}$

(d) $\sqrt{2}$

(e) $\frac{-1}{4\sqrt{17}}$

(f) $\frac{1}{\sqrt{17}}$

13.4.28. (a) 2

(b) $4t^2$

(c) 2

(d) 2π

(e) 2

(f) 4π

13.4.30. (a) 0

(b) 5

(c) 0

(d) 5

(e) 0

(f) 5

13.5 · The Arc Length Parameter and Curvature

13.5 · Exercises

Terms and Concepts

13.5.1. time and/or distance

13.5.2. curvature

13.5.3. Answers may include lines, circles, helixes

13.5.4. Answers will vary; they should mention the circle is tangent to the curve and has the same curvature as the curve at that point.

13.5.5. κ

13.5.6. a_{T} is not affected by curvature; the greater the curvature, the larger a_{N} becomes.

Problems

13.5.8. (a) $7t$

(b) $\left\langle 7\cos\left(\frac{s}{7}\right), 7\sin\left(\frac{s}{7}\right)\right\rangle$

13.5.10. (a) $13t$

(b) $\left\langle 5\cos\left(\frac{s}{13}\right), 13\sin\left(\frac{s}{13}\right), 12\cos\left(\frac{s}{13}\right)\right\rangle$

13.5.12. (a) greater than

(b) $\dfrac{\left|\frac{6x^2-2}{(x^2+1)^3}\right|}{\left(1+\frac{4x^2}{(x^2+1)^4}\right)^{\frac{3}{2}}}$

(c) 2

(d) $\frac{2750}{641^{\frac{3}{2}}}$

13.5.15. (a) less than

(b) $\frac{|2\cos(t)\cos(2t)+4\sin(t)\sin(2t)|}{(4\cos^2(2t)+\sin^2(t))^{\frac{3}{2}}}$

(c) $\frac{1}{4}$

(d) 8

13.5.18. (a) greater than

(b) $\dfrac{\left|\sec^3(t)\right|}{(\sec^4(t)+\sec^2(t)\tan^2(t))^{\frac{3}{2}}}$

(c) 1

(d) $\frac{3\sqrt{3}}{5\sqrt{5}}$

13.5.20. (a) greater than

(b) $\frac{2\sqrt{18t^4+15t^2+1}}{(18t^4-2t^2+1)^{\frac{3}{2}}}$

(c) 2

(d) $\frac{2\sqrt{2}}{17}$

13.5.22. (a) equal to

(b) $\frac{1}{13}$

(c) $\frac{1}{13}$

(d) $\frac{1}{13}$

13.5.23. $\frac{\sqrt{2}}{\sqrt[4]{5}}, \frac{-\sqrt{2}}{\sqrt[4]{5}}$

13.5.25. $\frac{1}{4}$

13.5.26. $\sqrt{5}, -\sqrt{5}$

13.5.28. $5\sqrt{10}$

13.5.30. $\frac{1}{45}$

13.5.32. $\left(x-\frac{8}{3}\right)^2+y^2=\frac{1}{9}$

13.5.34. $\left(x-\frac{1}{2}\right)^2+\left(y-\frac{1}{2}\right)^2=\frac{1}{2}$

14 • Functions of Several Variables, Continued

14 • Functions of Several Variables, Continued

14.2 • The Multivariable Chain Rule

14.2 • Exercises

Terms and Concepts

14.2.2. $g'(x)$

14.2.4. T

14.2.5. F

14.2.6. partial

Problems

14.2.7.

(a) $\frac{dz}{dt} = 3(2t) + 4(2) = 6t + 8.$

(b) At $t = 1$, $\frac{dz}{dt} = 14.$

14.2.8. (a) $2x - 4yt$

(b) 2

14.2.9.

(a) $\frac{dz}{dt} = 5(-2\sin(t)) + 2(\cos(t)) = -10\sin(t) + 2\cos(t)$

(b) At $t = \pi/4$, $\frac{dz}{dt} = -4\sqrt{2}.$

14.2.10. (a) $\frac{-\sin(t)}{1+y^2} - \frac{2xy\cos(t)}{(y^2+1)^2}$

(b) $-\frac{1}{2}$

14.2.11.

(a) $\frac{dz}{dt} = 2x(\cos(t)) + 4y(3\cos(t)).$

(b) At $t = \pi/4$, $x = \sqrt{2}/2$, $y = 3\sqrt{2}/2$, and $\frac{dz}{dt} = 19.$

14.2.12. (a) $-\sin(x)\sin(y)\,\pi + \cos(x)\cos(y)\cdot 2\pi$

(b) 0

14.2.14. $-\sqrt{\frac{3}{2}}, 0, \sqrt{\frac{3}{2}}$

14.2.16. $0, \pi$

14.2.18. $0, \frac{1}{\pi}\tan^{-1}\left(\sqrt{5}\right), 1 - \frac{1}{\pi}\tan^{-1}\left(\sqrt{5}\right), 1, 1 + \frac{1}{\pi}\tan^{-1}\left(\sqrt{5}\right), 2 - \frac{1}{\pi}\tan^{-1}\left(\sqrt{5}\right)$

14.2.20.

(a) $-\pi\sin\left(\pi x + \frac{\pi y}{2}\right)t^2 - \frac{1}{2}\pi\sin\left(\pi x + \frac{\pi y}{2}\right)\cdot 2st$

(b) $-\pi\sin\left(\pi x + \frac{\pi y}{2}\right)\cdot 2st - \frac{1}{2}\pi\sin\left(\pi x + \frac{\pi y}{2}\right)s^2$

(c) 2π

(d) $\frac{5\pi}{2}$

14.2.21. (a) $2x\cos(t) + 2y\sin(t)$

(b) $-2xs\sin(t) + 2ys\cos(t)$

(c) 4

(d) 0

14.2.22. (a) $-2yt^2e^{-\left(x^2+y^2\right)}$

(b) $-2xe^{-\left(x^2+y^2\right)} - 4stye^{-\left(x^2+y^2\right)}$

(c) $\frac{-2}{e^2}$

(d) $\frac{-6}{e^2}$

14.2.24. $\frac{-x}{y^2}$

14.2.26. $\frac{-(2x+y)}{2y+x}$

14.2.28. 0

14.2.30. (a) -2

(b) 5

14.3 · Directional Derivatives
14.3 · Exercises

Terms and Concepts

14.3.2. $\boldsymbol{i}$

14.3.3. $\boldsymbol{j}$

14.3.4. orthogonal

14.3.6. dot

Problems

14.3.8. $\langle \cos(x)\cos(y), -\sin(x)\sin(y)\rangle$

14.3.10. $\langle -4, 3\rangle$

14.3.12. $\left\langle 2xy^3 - 2, 3x^2y^2\right\rangle$

14.3.13.

(a) $2/5$

(b) $-2/\sqrt{5}$

14.3.14.

(a) $\frac{1}{4}\left(1-\sqrt{3}\right)$

(b) $\frac{4\sqrt{3}-3}{10\sqrt{2}}$

14.3.15.

(a) 0

(b) $2\sqrt{2}/9$

14.3.16.

(a) $\frac{-9}{\sqrt{10}}$

(b) $\frac{27}{\sqrt{34}}$

14.3.17.

(a) 0

(b) 0

14.3.18.

(a) $\frac{3}{\sqrt{2}}$

(b) 3

14.3.19.

(a) $\nabla f(2,1) = \langle -2, 2\rangle$

(b) $\sqrt{8}$

(c) $\langle 2, -2\rangle$

(d) $\vec{u} = \left\langle 1/\sqrt{2}, 1/\sqrt{2}\right\rangle$

14.3.20.

(a) $\left\langle \frac{1}{2\sqrt{2}}, \frac{-1}{2}\sqrt{\frac{3}{2}}\right\rangle$

(b) $\frac{1}{\sqrt{2}}$

(c) $\left\langle \frac{-1}{2\sqrt{2}}, \frac{1}{2}\sqrt{\frac{3}{2}}\right\rangle$

(d) $\left\langle \frac{1}{2}\sqrt{\frac{3}{2}}, \frac{1}{2\sqrt{2}}\right\rangle$

14.3.21.

(a) $\nabla f(1,1) = \langle -2/9, -2/9\rangle$

(b) $2\sqrt{2}/9$

(c) $\langle 2/9, 2/9\rangle$

(d) $\vec{u} = \left\langle 1/\sqrt{2}, -1/\sqrt{2}\right\rangle$

14.3.22.

(a) $\langle -4, 3\rangle$

(b) 5

(c) $\langle 4, -3\rangle$

(d) $\langle 3, 4\rangle$

14.3.23.

(a) No such direction

(b) 0

(c) No such direction

(d) All directions

14.3.24.

(a) $\langle 0, 3\rangle$

(b) 3

(c) $\langle 0, -3\rangle$

(d) $\langle 1, 0\rangle$

14.3.25.

(a) $\nabla F(x,y,z) = \left\langle 6xz^3 + 4y, 4x, 9x^2z^2 - 6z\right\rangle$

(b) $113/\sqrt{3}$

14.3.26.

(a) $\langle \cos(x)\cos(y)\,e^z, -\sin(x)\sin(y)\,e^z, \sin(x)\cos(y)\,e^z\rangle$

(b) $\frac{2}{3}$

14.3.27.

(a) $\nabla F(x,y,z) = \langle 2xy^2, 2y(x^2-z^2), -2y^2z \rangle$

(b) 0

14.3.28.

(a) $\left\langle \frac{-4x}{(x^2+y^2+z^2)^2}, \frac{-4y}{(x^2+y^2+z^2)^2}, \frac{-4z}{(x^2+y^2+z^2)^2} \right\rangle$

(b) 0

14.4 · Tangent Lines, Normal Lines, and Tangent Planes
14.4 · Exercises

Terms and Concepts

14.4.3. True

Problems

14.4.6. (a) $x = \frac{\pi}{3} + t, y = \frac{\pi}{6}, z = \frac{3}{4} - \frac{3\sqrt{3}}{4}t$

(b) $x = \frac{\pi}{3}, y = \frac{\pi}{6} + t, z = \frac{3}{4} + \frac{3\sqrt{3}}{4}t$

(c) $x = \frac{\pi}{3} + \frac{t}{\sqrt{5}}, y = \frac{\pi}{6} + \frac{2t}{\sqrt{5}}, z = \frac{3}{4} + \frac{3\sqrt{\frac{3}{5}}}{4}t$

14.4.8. (a) $x = 1 + t, y = 2, z = 3$

(b) $x = 1, y = 2 + t, z = 3$

(c) $x = 1 + \frac{t}{\sqrt{2}}, y = 2 + \frac{t}{\sqrt{2}}, z = 3$

14.4.10. $x = \frac{\pi}{3} - \frac{3\sqrt{3}}{4}t, y = \frac{\pi}{6} + \frac{3\sqrt{3}}{4}t, z = \frac{3}{4} - t$

14.4.12. $x = 1, y = 2, z = 3 - t$

14.4.18. $1.29904y - 1.29904x - z = -1.43017$

14.4.20. $z = 3$

14.4.22.

(a) $x = 4 - 2t, y = -3 + \frac{2}{3}t, z = \sqrt{5} + 2\sqrt{5}t$

(b) $0.666667y - 2x + 4.47214z = 0$

14.4.24.

(a) $x = 2 + \frac{\pi}{8\sqrt{3}}t, y = \frac{\pi}{12} - \sqrt{3}t, z = 4 - \frac{\pi}{8\sqrt{3}}t$

(b) $0.226725x - 1.73205y - 0.226725z = -0.9069$

14.5 · Extreme Values
14.5 · Exercises

Terms and Concepts

14.5.1. False
14.5.2. True
14.5.3. True

Problems

14.5.6. (a) NONE

(b) NONE

(c) $(7, -6)$

(d) NONE

14.5.8. (a) $(0, 0)$

(b) NONE

(c) NONE

(d) NONE

14.5.10. (a) $(-1, -2)$

(b) $(1, 2)$

(c) $(1, -2), (-1, 2)$

(d) NONE

14.5.12. (a) $(0, -3)$

(b) $(-1, 3), (1, 3)$

(c) $(-1, -3), (0, 3), (1, -3)$

(d) NONE

14.5.14. (a) NONE
(b) NONE
(c) NONE
(d) $(0, 0)$

14.5.15. (a) 3
(b) $(0, 1)$
(c) $\frac{3}{4}$
(d) $\left(0, \frac{-1}{2}\right)$

14.5.16. (a) $\frac{25}{28}$
(b) $\left(\frac{5}{14}, \frac{25}{196}\right)$
(c) -12
(d) $(-1, 1)$

15 · Multiple Integration

15 · Multiple Integration
15.1 · Iterated Integrals and Area
15.1 · Exercises

Terms and Concepts

15.1.2. iterated integration
15.1.3. curve to curve, then from point to point
15.1.4. area

Problems

15.1.5.
(a) $18x^2 + 42x - 117$
(b) -108

15.1.6.
(a) $2 + \pi^2 \cos(y)$
(b) $\pi^2 + \pi$

15.1.7.
(a) $x^4/2 - x^2 + 2x - 3/2$
(b) $23/15$

15.1.8.
(a) $\frac{y^4}{2} - y^3 + \frac{y^2}{2}$
(b) $\frac{8}{15}$

15.1.9.
(a) $\sin^2(y)$
(b) $\pi/2$

15.1.10.
(a) $\frac{x}{1+x^2}$
(b) $\frac{1}{2} \ln\left(\frac{5}{2}\right)$

15.3 · Double Integration with Polar Coordinates
15.3 · Exercises

Problems

15.3.3. 4π
15.3.5. 16π
15.3.8. $\frac{\pi}{2}$

15.3.12. $\frac{128}{3}$

15.5 · Surface Area
15.5 · Exercises

Problems

15.5.7. $SA = \int_0^{2\pi}\int_0^{2\pi} \sqrt{1+\cos^2(x)\cos^2(y)+\sin^2(x)\sin^2(y)}\,dx\,dy$

15.5.8. $SA = \int_{-3}^{3}\int_{-\sqrt{9-x^2}}^{\sqrt{9-x^2}} \sqrt{1+\frac{4x^2+4y^2}{(1+x^2+y^2)^4}}\,dx\,dy$

Polar offers simpler bounds:

$SA = \int_0^{2\pi}\int_0^{3} r\sqrt{1+\frac{4r^2}{(1+r^2)^4}}\,dr\,d\theta$

15.5.9. $SA = \int_{-1}^{1}\int_{-1}^{1} \sqrt{1+4x^2+4y^2}\,dx\,dy$

15.5.10. $SA = \int_{-5}^{5}\int_{0}^{1} \sqrt{1+\frac{4x^2e^{2x^2}}{\left(1+e^{x^2}\right)^4}}\,dy\,dx$

15.6 · Volume Between Surfaces and Triple Integration
15.6 · Exercises

Problems

15.6.6. 52

15.6.8. $\frac{3\pi}{2}$

15.6.9. $dz\,dy\,dx$: $\int_0^3\int_0^{1-x/3}\int_0^{2-2x/3-2y} dz\,dy\,dx$

$dz\,dx\,dy$: $\int_0^1\int_0^{3-3y}\int_0^{2-2x/3-2y} dz\,dx\,dy$

$dy\,dz\,dx$: $\int_0^3\int_0^{2-2x/3}\int_0^{1-x/3-z/2} dy\,dz\,dx$

$dy\,dx\,dz$: $\int_0^2\int_0^{3-3z/2}\int_0^{1-x/3-z/2} dy\,dx\,dz$

$dx\,dz\,dy$: $\int_0^1\int_0^{2-2y}\int_0^{3-3y-3z/2} dx\,dz\,dy$

$dx\,dy\,dz$: $\int_0^2\int_0^{1-z/2}\int_0^{3-3y-3z/2} dx\,dy\,dz$

$V = \int_0^3\int_0^{1-x/3}\int_0^{2-2x/3-2y} dz\,dy\,dx = 1.$

15.6.10. $dz\,dy\,dx$: $\int_1^3\int_0^2\int_0^{(3-x)/2} dz\,dy\,dx$

$dz\,dx\,dy$: $\int_0^2\int_1^3\int_0^{(3-x)/2} dz\,dx\,dy$

$dy\,dz\,dx$: $\int_1^3\int_0^{(3-x)/2}\int_0^2 dy\,dz\,dx$

$dy\,dx\,dz$: $\int_0^1\int_1^{3-2z}\int_0^2 dy\,dx\,dz$

$dx\,dz\,dy$: $\int_0^2\int_0^1\int_1^{3-2z} dx\,dz\,dy$

$dx\,dy\,dz$: $\int_0^1\int_0^2\int_1^{3-2z} dx\,dy\,dz$

$V = \int_0^1\int_0^2\int_1^{3-2z} dx\,dy\,dz = 2.$

15.6.11. $dz\,dy\,dx$: $\displaystyle\int_0^2\int_{-2}^0\int_{y^2/2}^{-y} dz\,dy\,dx$

$dz\,dx\,dy$: $\displaystyle\int_{-2}^0\int_0^2\int_{y^2/2}^{-y} dz\,dx\,dy$

$dy\,dz\,dx$: $\displaystyle\int_0^2\int_0^2\int_{-\sqrt{2z}}^{-z} dy\,dz\,dx$

$dy\,dx\,dz$: $\displaystyle\int_0^2\int_0^2\int_{-\sqrt{2z}}^{-z} dy\,dx\,dz$

$dx\,dz\,dy$: $\displaystyle\int_{-2}^0\int_{y^2/2}^{-y}\int_0^2 dx\,dz\,dy$

$dx\,dy\,dz$: $\displaystyle\int_0^2\int_{-\sqrt{2z}}^{-z}\int_0^2 dx\,dy\,dz$

$V=\displaystyle\int_0^2\int_0^2\int_{-\sqrt{2z}}^{-z} dy\,dz\,dx = 4/3.$

15.6.13. $dz\,dy\,dx$: $\displaystyle\int_0^2\int_{1-x/2}^1\int_0^{2x+4y-4} dz\,dy\,dx$

$dz\,dx\,dy$: $\displaystyle\int_0^1\int_{2-2y}^2\int_0^{2x+4y-4} dz\,dx\,dy$

$dy\,dz\,dx$: $\displaystyle\int_0^2\int_0^{2x}\int_{z/4-x/2+1}^1 dy\,dz\,dx$

$dy\,dx\,dz$: $\displaystyle\int_0^4\int_{z/2}^2\int_{z/4-x/2+1}^1 dy\,dx\,dz$

$dx\,dz\,dy$: $\displaystyle\int_0^1\int_0^{4y}\int_{z/2-2y+2}^2 dx\,dz\,dy$

$dx\,dy\,dz$: $\displaystyle\int_0^4\int_{z/4}^1\int_{z/2-2y+2}^2 dx\,dy\,dz$

$V=\displaystyle\int_0^4\int_{z/4}^1\int_{z/2-2y-2}^2 dx\,dy\,dz = 4/3.$

15.6.15. $dz\,dy\,dx$: $\displaystyle\int_0^1\int_0^{1-x^2}\int_0^{\sqrt{1-y}} dz\,dy\,dx$

$dz\,dx\,dy$: $\displaystyle\int_0^1\int_0^{\sqrt{1-y}}\int_0^{\sqrt{1-y}} dz\,dx\,dy$

$dy\,dz\,dx$:
$\displaystyle\int_0^1\int_0^x\int_0^{1-x^2} dy\,dz\,dx + \int_0^1\int_x^1\int_0^{1-z^2} dy\,dz\,dx$

$dy\,dx\,dz$:
$\displaystyle\int_0^1\int_0^z\int_0^{1-z^2} dy\,dx\,dz + \int_0^1\int_z^1\int_0^{1-x^2} dy\,dx\,dz$

$dx\,dz\,dy$: $\displaystyle\int_0^1\int_0^{\sqrt{1-y}}\int_0^{\sqrt{1-y}} dx\,dz\,dy$

$dx\,dy\,dz$: $\displaystyle\int_0^1\int_0^{1-z^2}\int_0^{\sqrt{1-y}} dx\,dy\,dz$ Answers will vary. Neither order is particularly "hard." The order $dz\,dy\,dx$ requires integrating a square root, so powers can be messy; the order $dy\,dz\,dx$ requires two triple integrals, but each uses only polynomials.

15.6.18. $\frac{7}{8}$

15.6.12. $dz\,dy\,dx$: $\displaystyle\int_0^3\int_{3x}^9\int_0^{\sqrt{y^2-9x^2}} dz\,dy\,dx$

$dz\,dx\,dy$: $\displaystyle\int_0^9\int_0^{y/3}\int_0^{\sqrt{y^2-9x^2}} dz\,dx\,dy$

$dy\,dz\,dx$: $\displaystyle\int_0^3\int_0^{\sqrt{81-9x^2}}\int_{\sqrt{z^2+9x^2}}^9 dy\,dz\,dx$

$dy\,dx\,dz$: $\displaystyle\int_0^9\int_0^{\sqrt{9-z^2/9}}\int_{\sqrt{z^2+9x^2}}^9 dy\,dx\,dz$

$dx\,dz\,dy$: $\displaystyle\int_0^9\int_0^y\int_0^{\frac{1}{3}\sqrt{y^2-z^2}} dx\,dz\,dy$

$dx\,dy\,dz$: $\displaystyle\int_0^9\int_z^9\int_0^{\frac{1}{3}\sqrt{y^2-z^2}} dx\,dy\,dz$

15.6.14. $dz\,dy\,dx$: $\displaystyle\int_{-2}^2\int_0^{4-x^2}\int_0^{2y} dz\,dy\,dx$

$dz\,dx\,dy$: $\displaystyle\int_0^4\int_{-\sqrt{4-y}}^{\sqrt{4-y}}\int_0^{2x+4y-4} dz\,dx\,dy$

$dy\,dz\,dx$: $\displaystyle\int_{-2}^2\int_0^{8-2x^2}\int_{z/2}^{4-x^2} dy\,dz\,dx$

$dy\,dx\,dz$: $\displaystyle\int_0^8\int_{-\sqrt{4-z/2}}^{\sqrt{4-z/2}}\int_{z/2}^{4-x^2} dy\,dx\,dz$

$dx\,dz\,dy$: $\displaystyle\int_0^4\int_0^{2y}\int_{-\sqrt{4-y}}^{\sqrt{4-y}} dx\,dz\,dy$

$dx\,dy\,dz$: $\displaystyle\int_0^8\int_{z/2}^4\int_{-\sqrt{4-y}}^{\sqrt{4-y}} dx\,dy\,dz$

$V=\displaystyle\int_{-2}^2\int_0^{4-x^2}\int_0^{2y} dz\,dy\,dx = 512/15.$

15.6.16. $dz\,dy\,dx$:
$\displaystyle\int_0^1\int_0^{3x}\int_0^{1-x} dz\,dy\,dx + \int_0^1\int_{3x}^3\int_0^{1-y/3} dz\,dy\,dx$

$dz\,dx\,dy$: $\displaystyle\int_0^3\int_0^{y/3}\int_0^{1-y/3} dz\,dy\,dx +$
$\displaystyle\int_0^3\int_{y/3}^1\int_0^{1-x} dz\,dx\,dy$

$dy\,dz\,dx$: $\displaystyle\int_0^1\int_0^{1-x}\int_0^{3-3z} dy\,dz\,dx$

$dy\,dx\,dz$: $\displaystyle\int_0^1\int_0^{1-z}\int_0^{3-3z} dy\,dx\,dz$

$dx\,dz\,dy$: $\displaystyle\int_0^3\int_0^{1-y/3}\int_0^{1-z} dx\,dz\,dy$

$dx\,dy\,dz$: $\displaystyle\int_0^1\int_0^{3-3z}\int_0^{1-z} dx\,dy\,dz$

$V=\displaystyle\int_0^1\int_0^{3-3z}\int_0^{1-z} dx\,dy\,dz = 1.$

15.6.20. 0

15.7 · Triple Integration with Cylindrical and Spherical Coordinates
15.7 · Exercises

Problems

15.7.11. $\displaystyle\int_{\theta_1}^{\theta_2}\int_{r_1}^{r_2}\int_{z_1}^{z_2} h(r,\theta,z)r\,dz\,dr\,d\theta$

15.7.12. $\displaystyle\int_{\varphi_1}^{\varphi_2}\int_{\theta_1}^{\theta_2}\int_{\rho_1}^{\rho_2} h(\rho,\theta,\varphi)\rho^2\sin(\varphi)\,d\rho\,d\theta\,d\varphi$

15.7.19. Describes the portion of the unit ball that resides in the first octant.

15.7.20. Describes half of a spherical shell (i.e., $y \geq 0$) with inner radius of 1 and outer radius of 1.1 centered at the origin.

16 · Vector Analysis
16 · Vector Analysis
16.1 · Introduction to Line Integrals
16.1 · Exercises

Terms and Concepts

16.1.1. When C is a curve in the plane and f is a function defined over C, then $\int_C f(s)\,ds$ describes the area under the spatial curve that lies on f, over C.

16.1.2. The evaluation is the same. The $\oint$ notation signifies that the curve C is a closed curve, though the evaluation is the same.

16.1.3. The variable s denotes the arc-length parameter, which is generally difficult to use. The Key Idea allows one to parametrize a curve using another, ideally easier-to-use, parameter.

16.1.4. Answers will vary.

Problems

16.1.5. $12\sqrt{2}$

16.1.6. $41\sqrt{10}/2$

16.1.7. 40π

16.1.8. $10\pi^2$

16.1.9. Over the first subcurve of C, the line integral has a value of $3/2$; over the second subcurve, the line integral has a value of $4/3$. The total value of the line integral is thus $17/6$.

16.1.10. Over the first subcurve of C, the line integral has a value of $2\sqrt{2}/3$; over the second subcurve, the line integral has a value of $\pi - 2$. The total value of the line integral is thus $\pi + 2\sqrt{2}/3 - 2$.

16.1.11. $\int_0^1 (5t^2 +_2 t + 2)\sqrt{(4t+1)^2+1}\,dt \approx 17.071$

16.1.12. $\int_0^\pi t\sqrt{1+\cos^2 t}\,dt \approx 6.001$

16.1.13. $\oint_0^{2\pi} \left(10 - 4\cos^2 t - \sin^2 t\right)\sqrt{\cos^2 t + 4\sin^2 t}\,dt \approx 74.986$

16.1.14. $\int_{-1}^1 \left(3t^3+2t+5\right)\sqrt{9t^4+1}\,dt \approx 15.479$

16.1.15. $7\sqrt{26}/3$

16.1.16. 2π

16.1.17. $8\pi^3$

16.1.18. $5/2$

16.1.19. $M = 8\sqrt{2}\pi^2$; center of mass is $(0, -1/(2\pi), 8\pi/3)$.

16.1.20. $M \approx 0.237$; center of mass is approximately $(0.173, 0.099, 0.065)$.

16.2 · Vector Fields
16.2 · Exercises

Terms and Concepts

16.2.1. Answers will vary. Appropriate answers include velocities of moving particles (air, water, etc.); gravitational or electromagnetic forces.

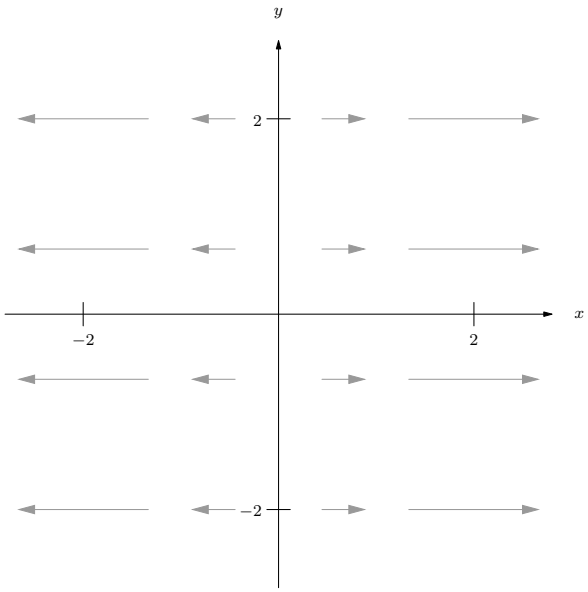
16.2.2. Specific answers will vary, though should relate to the idea that "more of the vector field is moving into that point than out of that point."

16.2.3. Specific answers will vary, though should relate to the idea that the vector field is spinning clockwise at that point.

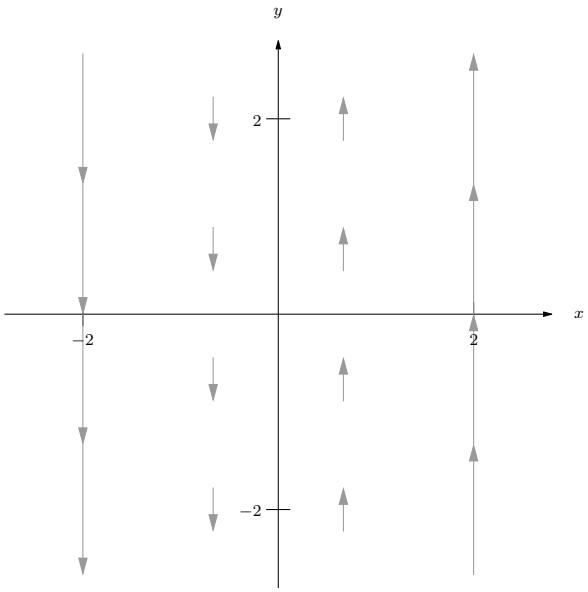
16.2.4. No; to be incompressible, the divergence needs to be 0 everywhere, not just at one point.

Problems

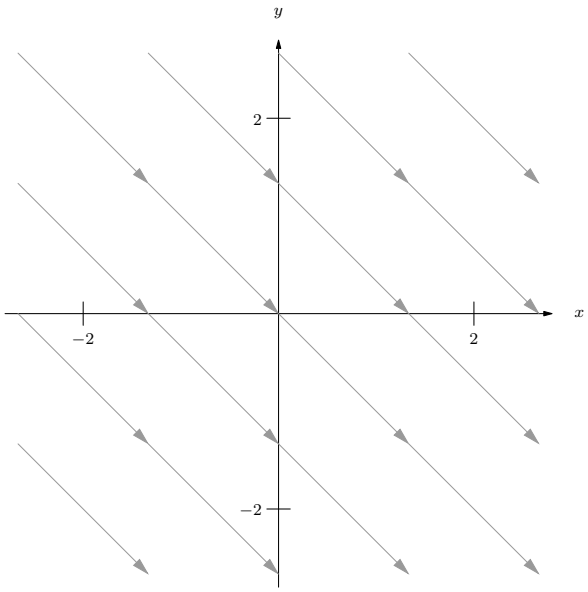
16.2.5. Correct answers should look similar to



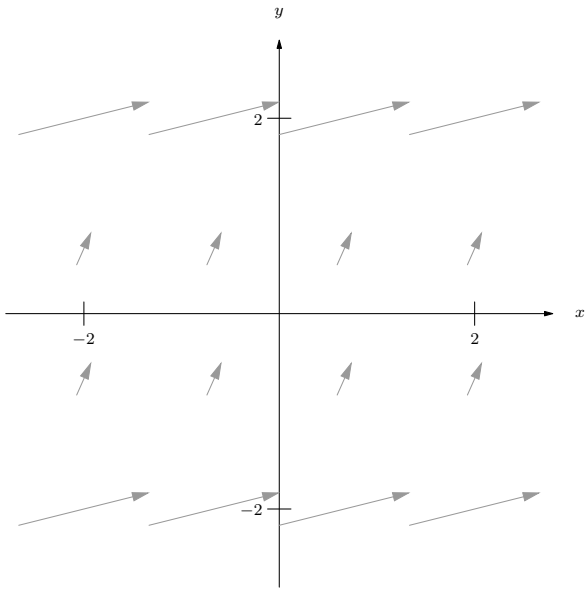
16.2.6. Correct answers should look similar to



16.2.7. Correct answers should look similar to



16.2.8. Correct answers should look similar to



16.2.9. div $\vec{F} = 1 + 2y$
curl $\vec{F} = 0$

16.2.10. div $\vec{F} = 0$
curl $\vec{F} = 1 + 2y$

16.2.11. div $\vec{F} = x\cos(xy) - y\sin(xy)$
curl $\vec{F} = y\cos(xy) + x\sin(xy)$

16.2.12. div $\vec{F} = \frac{4}{(x^2+y^2)^2}$
curl $\vec{F} = 0$

16.2.13. div $\vec{F} = 3$
curl $\vec{F} = \langle -1, -1, -1 \rangle$

16.2.14. div $\vec{F} = 2x + 2y + 2z$
curl $\vec{F} = \langle 2y, 2z, 2x \rangle$

16.2.15. div $\vec{F} = 1 + 2y$
curl $\vec{F} = 0$

16.2.16. div $\vec{F} = 2y$
curl $\vec{F} = 0$

16.2.17. div $\vec{F} = 2y - \sin z$
curl $\vec{F} = \vec{0}$

16.2.18. div $\vec{F} = \frac{2}{(x^2+y^2+z^2)^2}$
curl $\vec{F} = \vec{0}$

16.3 · Line Integrals over Vector Fields

16.3 · Exercises

Terms and Concepts

16.3.1. False. It is true for line integrals over scalar fields, though.

16.3.2. The input of $\vec{F}$ should be a point in the plane, not a two dimensional vector.

16.3.3. True.

16.3.4. False.

16.3.5. We can conclude that $\vec{F}$ is conservative.

16.3.6. By the Fundamental Theorem of Line Integrals, since $\vec{F}$ is conservative, $\oint_C \vec{F} \cdot d\vec{r} = f(B) - f(A)$, where f is a potential function for $\vec{F}$ and A and B are the initial and terminal points of C, respectively. Since C is a closed curve, $A = B$, and hence $f(B) - f(A) = 0$.

Problems

16.3.7. $11/6$. (One parametrization for C is $\vec{r}(t) = \langle 3t, t\rangle$ on $0 \le t \le 1$.)

16.3.8. $5/3$. (One parametrization for C is $\vec{r}(t) = \langle t, t^2\rangle$ on $0 \le t \le 1$.)

16.3.9. 0. (One parametrization for C is $\vec{r}(t) = \langle \cos t, \sin t\rangle$ on $0 \le t \le \pi$.)

16.3.10. $2/5$. (One parametrization for C is $\vec{r}(t) = \langle t, t^3\rangle$ on $-1 \le t \le 1$.)

16.3.11. 12. (One parametrization for C is $\vec{r}(t) = \langle 1, 2, 3\rangle + t\langle 3, 1, -1\rangle$ on $0 \le t \le 1$.)

16.3.12. 1.

16.3.13. $5/6$ joules. (One parametrization for C is $\vec{r}(t) = \langle t, t\rangle$ on $0 \le t \le 1$.)

16.3.14. $13/15$ joules. (One parametrization for C is $\vec{r}(t) = \langle t, \sqrt{t}\rangle$ on $0 \le t \le 1$.)

16.3.15. 24 ft-lbs.

16.3.16. 24 ft-lbs.

16.3.17.

(a) $f(x, y) = xy + x$

(b) curl $\vec{F} = 0$.

(c) 1. (One parametrization for C is $\vec{r}(t) = \langle t, t - 1\rangle$ on $0 \le t \le 1$.)

(d) 1 (with $A = (0, 1)$ and $B = (1, 0)$, $f(B) - f(A) = 1$.)

16.3.18.

(a) $f(x, y) = x^2 + xy + y^2$

(b) curl $\vec{F} = 0$.

(c) 0.

(d) 0 (with $A = (0, 0)$ and $B = (0, 0)$, $f(B) - f(A) = 0$.)

16.3.19.

(a) $f(x, y) = x^2yz$

(b) curl $\vec{F} = \vec{0}$.

(c) 250.

(d) 250 (with $A = (1, -1, 0)$ and $B = (5, 5, 2)$, $f(B) - f(A) = 250$.)

16.3.20.

(a) $f(x, y) = x^2 + y^2 + z^2$

(b) curl $\vec{F} = \vec{0}$.

(c) 0.

(d) 0 (with $A = (1, 0, 0)$ and $B = (1, 0, 0)$, $f(B) - f(A) = 0$.)

16.4 · Flow, Flux, Green's Theorem and the Divergence Theorem

16.4 · Exercises

Terms and Concepts

16.4.1. along, across

16.4.2. It is the measure of flow around the entirety of a closed curve C.

16.4.3. the curl of $\vec{F}$, or curl $\vec{F}$

16.4.4. the divergence of $\vec{F}$, or div $\vec{F}$

16.4.5. curl $\vec{F}$

16.4.6. div $\vec{F}$

Problems

16.4.7. 12

16.4.8. 12

16.4.9. $-2/3$

16.4.10. $10/3$

16.4.11. $1/2$

16.4.12. $1/2$

16.4.13. The line integral $\oint_C \vec{F} \cdot d\vec{r}$, over the parabola, is $38/3$; over the line, it is -10. The total line integral is thus $38/3 - 10 = 8/3$. The double integral of curl $\vec{F} = 2$ over R also has value $8/3$.

16.4.14. Both the line integral and double integral have value of 2π.

16.4.15. Three line integrals need to be computed to compute $\oint_C \vec{F} \cdot d\vec{r}$. It does not matter which corner one starts from first, but be sure to proceed around the triangle in a counterclockwise fashion.
From $(0,0)$ to $(2,0)$, the line integral has a value of 0.
From $(2,0)$ to $(1,1)$ the integral has a value of $7/3$.
From $(1,1)$ to $(0,0)$ the line integral has a value of $-1/3$. Total value is 2.
The double integral of curl $\vec{F}$ over R also has value 2.

16.4.16. Two line integrals need to be computed to compute $\oint_C \vec{F} \cdot d\vec{r}$. Along the parabola, the line integral has value 25.5. Along the line, the line integral has value -21. Together, the total value is 4.5
The double integral of curl $\vec{F}$ over R also has value 4.5.

16.4.17. Any choice of $\vec{F}$ is appropriate as long as curl $\vec{F} = 1$. When $\vec{F} = \langle -y/2, x/2 \rangle$, the integrand of the line integral is simply 6. The area of R is 12π.

16.4.18. Any choice of $\vec{F}$ is appropriate as long as curl $\vec{F} = 1$. The choices of $\vec{F} = \langle -y, 0 \rangle$ and $\langle 0, x \rangle$ each lead to reasonable integrands. The area of R is $4/3$.

16.4.19. Any choice of $\vec{F}$ is appropriate as long as curl $\vec{F} = 1$. The choices of $\vec{F} = \langle -y, 0 \rangle$, $\langle 0, x \rangle$ and $\langle -y/2, x/2 \rangle$ each lead to reasonable integrands. The area of R is $16/15$.

16.4.20. Any choice of $\vec{F}$ is appropriate as long as curl $\vec{F} = 1$. The choice of $\vec{F} = \langle -y/2, x/2 \rangle$ leads to a reasonable integrand after simplification. The area of R is $41\pi/10$.

16.4.21. The line integral $\oint_C \vec{F} \cdot \vec{n}\, ds$, over the parabola, is $-22/3$; over the line, it is 10. The total line integral is thus $-22/3 + 10 = 8/3$. The double integral of div $\vec{F} = 2$ over R also has value $8/3$.

16.4.22. Both the line integral and double integral have value of 0.

16.4.23. Three line integrals need to be computed to compute $\oint_C \vec{F} \cdot \vec{n}\, ds$. It does not matter which corner one starts from first, but be sure to proceed around the triangle in a counterclockwise fashion.
From $(0,0)$ to $(2,0)$, the line integral has a value of 0.
From $(2,0)$ to $(1,1)$ the integral has a value of $1/3$.
From $(1,1)$ to $(0,0)$ the line integral has a value of $1/3$. Total value is $2/3$.
The double integral of div $\vec{F}$ over R also has value $2/3$.

16.4.24. Two line integrals need to be computed to compute $\oint_C \vec{F} \cdot \vec{n}\, ds$. Along the parabola, the line integral has value $159/20$. Along the line, the line integral has value 6. Together, the total value is $279/20$.
The double integral of div $\vec{F}$ over R also has value $279/20$.

16.5 • Parametrized Surfaces and Surface Area
16.5 • Exercises

Terms and Concepts

16.5.1. Answers will vary, though generally should meaningfully include terms like “two sided”.

16.5.2. Many possible answers exist; the one given by the book is the Möbius band.

Problems

16.5.3.

(a) $\vec{r}(u,v) = \langle u, v, 3u^2v\rangle$ on $-1 \le u \le 1$, $0 \le v \le 2$.

(b) $\vec{r}(u,v) = \langle 3v\cos u + 1, 3v\sin u + 2, 3(3v\cos u + 1)^2(3v\sin u + 2)\rangle$, on $0 \le u \le 2\pi$, $0 \le v \le 1$.

(c) $\vec{r}(u,v) = \langle u, v(2-2u), 3u^2v(2-2u)\rangle$ on $0 \le u, v \le 1$.

(d) $\vec{r}(u,v) = \langle u, v(1-u^2), 3u^2v(1-u^2)\rangle$ on $-1 \le u \le 1, 0 \le v \le 1$.

16.5.4.

(a) $\vec{r}(u,v) = \langle u, v, 4u + 2u^2\rangle$ on $1 \le u \le 4$, $5 \le v \le 7$.

(b) $\vec{r}(u,v) = \langle 4v\cos u, 3v\sin u, 16v\cos u + 2(3v\sin u)^2\rangle$, on $0 \le u \le 2\pi$, $0 \le v \le 1$.

(c) $\vec{r}(u,v) = \langle u, u + v(4-2u), 4u + 2\big(u + v(4-2u)\big)^2\rangle$ on $0 \le u \le 2$, $0 \le v \le 1$.

(d) $\vec{r}(u,v) = \langle v\cos u, v\sin u, 4v\cos u + 2(v\sin u)^2\rangle$ on $0 \le u \le 2\pi$, $2 \le v \le 5$.

16.5.5. $\vec{r}(u,v) = \langle 0, u, v\rangle$ with $0 \le u \le 2$, $0 \le v \le 1$.

16.5.6. $\vec{r}(u,v) = \langle u, 0, 1 - u + vu\rangle$ with $0 \le u \le 1$, $0 \le v \le 1$.

16.5.7. $\vec{r}(u,v) = \langle 3\sin u\cos v, 2\sin u\sin v, 4\cos u\rangle$ with $0 \le u \le \pi$, $0 \le v \le 2\pi$.

16.5.8. Answers may vary; one solution is $\vec{r}(u,v) = \langle v\cos u, v, 4v\sin u\rangle$ with $0 \le u \le 2\pi$, $-1 \le v \le 5$.

16.5.9. Answers may vary.
For $z = \frac{1}{2}(3-x)$: $\vec{r}(u,v) = \langle u, v, \frac{1}{2}(3-u)\rangle$, with $1 \le u \le 3$ and $0 \le v \le 2$.
For $x = 1$: $\vec{r}(u,v) = \langle 1, u, v\rangle$, with $0 \le u \le 2$, $0 \le v \le 1$
For $y = 0$: $\vec{r}(u,v) = \langle u, 0, v/2(3-u)\rangle$, with $1 \le u \le 3$, $0 \le v \le 1$
For $y = 2$: $\vec{r}(u,v) = \langle u, 2, v/2(3-u)\rangle$, with $1 \le u \le 3$, $0 \le v \le 1$
For $z = 0$: $\vec{r}(u,v) = \langle u, v, 0\rangle$, with $1 \le u \le 3$, $0 \le v \le 2$

16.5.10. Answers may vary.
For $z = 2x + 4y - 4$: $\vec{r}(u,v) = \langle u, 1 - u/2 + uv/2, 2u + 4(1 - u/2 + uv/2) - 4\rangle$, with $0 \le u \le 2$, $0 \le v \le 1$.
For $x = 2$: $\vec{r}(u,v) = \langle 2, u, 4uv\rangle$, with $0 \le u \le 1$, $0 \le v \le 1$
For $y = 1$: $\vec{r}(u,v) = \langle u, 1, 2uv\rangle$, with $0 \le u \le 2$, $0 \le v \le 1$
For $z = 0$: $\vec{r}(u,v) = \langle u, 1 - u/2 + uv/2, 0\rangle$, with $0 \le u \le 2$, $0 \le v \le 1$

16.5.11. Answers may vary.
For $z = 2y$: $\vec{r}(u,v) = \langle u, v(4-u^2), 2v(4-u^2)\rangle$ with $-2 \le u \le 2$ and $0 \le v \le 1$.
For $y = 4 - x^2$: $\vec{r}(u,v) = \langle u, 4-u^2, 2v(4-u^2)\rangle$ with $-2 \le u \le 2$ and $0 \le v \le 1$.
For $z = 0$: $\vec{r}(u,v) = \langle u, v(4-u^2), 0\rangle$ with $-2 \le u \le 2$ and $0 \le v \le 1$.

16.5.12. Answers may vary.
For $y = 1 - z^2$:
$\vec{r}(u,v) = \langle u, v(1-u^2), \sqrt{1 - v(1-u^2)}\rangle$ with $0 \le u \le 1$ and $0 \le v \le 1$.
For $y = 1 - x^2$: $\vec{r}(u,v) = \langle u, 1-u^2, uv\rangle$ with $0 \le u \le 1$ and $0 \le v \le 1$.
For $x = 0$: $\vec{r}(u,v) = \langle 0, v(1-u^2), u\rangle$ with $0 \le u \le 1$ and $0 \le v \le 1$.
For $y = 0$: $\vec{r}(u,v) = \langle u, 0, v\rangle$ with $0 \le u \le 1$ and $0 \le v \le 1$.
For $z = 0$: $\vec{r}(u,v) = \langle u, v(1-u^2), 0rangle$ with $0 \le u \le 1$ and $0 \le v \le 1$.

16.5.13. Answers may vary.
For $x^2 + y^2/9 = 1$: $\vec{r}(u,v) = \langle \cos u, 3\sin u, v\rangle$ with $0 \le u \le 2\pi$ and $1 \le v \le 3$.
For $z = 1$: $\vec{r}(u,v) = \langle v\cos u, 3v\sin u, 1\rangle$ with $0 \le u \le 2\pi$ and $0 \le v \le 1$.
For $z = 3$: $\vec{r}(u,v) = \langle v\cos u, 3v\sin u, 3\rangle$ with $0 \le u \le 2\pi$ and $0 \le v \le 1$.

16.5.14. Answers may vary.
For $x^2 + y^2 = (z-1)^2$:
$\vec{r}(u,v) = \langle v\cos u, v\sin u, 1 - v\rangle$ with $0 \le u \le 2\pi$ and $0 \le v \le 1$.
For $z = 0$: $\vec{r}(u,v) = \langle v\cos u, v\sin u, 0\rangle$ with $0 \le u \le 2\pi$ and $0 \le v \le 1$.

16.5.15. Answers may vary.
For $z = 1 - x^2$: $\vec{r}(u, v) = \langle u, v, 1 - u^2 \rangle$ with $-1 \le u \le 1$ and $-1 \le v \le 2$.
For $y = -1$: $\vec{r}(u, v) = \langle u, -1, v(1 - u^2) \rangle$ with $-1 \le u \le 1$ and $0 \le v \le 1$.
For $y = 2$: $\vec{r}(u, v) = \langle u, 2, v(1 - u^2) \rangle$ with $-1 \le u \le 1$ and $0 \le v \le 1$.
For $z = 0$: $\vec{r}(u, v) = \langle u, v, 0 \rangle$ with $-1 \le u \le 1$ and $-1 \le v \le 2$.

16.5.16. Answers may vary.
For $z = 4 - x^2 - 4y^2$: $\vec{r}(u, v) = \langle 2v \cos u, v \sin u, 4 - (2v \cos u)^2 - 4(v \sin u)^2 \rangle$ with $0 \le u \le 2\pi$ and $0 \le v \le 1$.
For $z = 0$: $\vec{r}(u, v) = \langle 2v \cos u, v \sin u, 0 \rangle$ with $0 \le u \le 2\pi$ and $0 \le v \le 1$.

16.5.17. $S = 2\sqrt{14}$.

16.5.18. $S = \sqrt{6}/2$.

16.5.19. $S = 4\sqrt{3}\pi$.

16.5.20. $S = 3\sqrt{3}\pi$.

16.5.21. $S = \int_0^3 \int_0^{2\pi} \sqrt{v^2 + 4v^4}\, du\, dv = (37\sqrt{37} - 1)\pi/6 \approx 117.319$.

16.5.22. $S = \int_0^1 \int_0^1 \sqrt{v^2 + 4u^2v^2 + 4v^4}\, du\, dv \approx 0.931$.

16.5.23. $S = \int_0^1 \int_{-1}^1 \sqrt{(5u^2 - 5)^2 + 2(1 - u^2)^2}\, du\, dv = 4\sqrt{3} \approx 6.9282$.

16.5.24. $S = \int_0^1 \int_0^{2\pi} \sqrt{v^2 + 4v^4}\, du\, dv = (5\sqrt{5} - 1)\pi/6 \approx 5.330$.

16.6 · Surface Integrals

16.6 · Exercises

Terms and Concepts

16.6.1. curve; surface

16.6.2. Answers will vary; in general, it means that more of the vector field passes through the surface opposite the direction of the normal vector than in the same direction of the normal vector.

16.6.3. outside

16.6.4.

Problems

16.6.5. $240\sqrt{3}$

16.6.6. 40π

16.6.7. 24

16.6.8. 15

16.6.9. 0

16.6.10. 0

16.6.11. $-1/2$

16.6.12. π

16.6.13. 0; the flux over $\mathcal{S}_1$ is -45π and the flux over $\mathcal{S}_2$ is 45π.

16.6.14. $9\pi/8$; the flux over $\mathcal{S}_1$ is $3\pi/4$ (use $\vec{r}(u, v) = \langle \sin u \cos v, \sin u \sin v, \cos u \rangle$ on $\pi/3 \le u \le \pi$, $0 \le v \le 2\pi$) and the flux over $\mathcal{S}_2$ is $3\pi/8$ (use $\vec{r}(u, v) = \langle v\sqrt{3} \cos(u)/2, v\sqrt{3} \sin(u)/2, 1/2 \rangle$ for $0 \le u \le 2\pi$, $0 \le v \le 1$.

16.7 · The Divergence Theorem and Stokes' Theorem

16.7 · Exercises

Terms and Concepts

16.7.1. Answers will vary; in Section 16.4, the Divergence Theorem connects outward flux over a closed curve in the plane to the divergence of the vector field, whereas in this section the Divergence Theorem connects outward flux over a closed surface in space to the divergence of the vector field.

16.7.2. Divergence.

16.7.3. Curl.

16.7.4. Green's Theorem.

Problems

16.7.5. Outward flux across the plane $z = 2 - x/2 - 2y/3$ is 14; across the plane $z = 0$ the outward flux is -8; across the planes $x = 0$ and $y = 0$ the outward flux is 0.
Total outward flux: 14.
$\iint_D \text{div}\,\vec{F}\,dV = \int_0^4 \int_0^{3-3x/4} \int_0^{2-x/2-2y/3} (2x + 2y)\,dz\,dy\,dx = 14$.

16.7.6. Outward flux across the cylinder $x^2 + y^2 = 1$ is 0; across the plane $z = 3$ the outward flux is 3π; across the plane $z = -3$ the outward flux is 3π.
Total outward flux: 6π.
$\iint_D \text{div}\,\vec{F}\,dV = \int_0^{2\pi} \int_0^1 \int_{-3}^3 r\,dz\,dr\,d\theta = 6\pi$.

16.7.7. Outward flux across the surface $z = xy(3 - x)(3 - y)$ is 252; across the plane $z = 0$ the outward flux is -9.
Total outward flux: 243.
$\iint_D \text{div}\,\vec{F}\,dV = \int_0^3 \int_0^3 \int_0^{xy(3-x)(3-y)} 12\,dz\,dy\,dx = 243$.

16.7.8. Outward flux across the paraboloid is $112\pi/3$; across the disk the outward flux is 0.
Total outward flux: $112\pi/3$.
$\iint_D \text{div}\,\vec{F}\,dV = \int_0^{2\pi} \int_0^2 \int_0^{4-r^2} (2z + 2)r\,dz\,dr\,d\theta = 112\pi/3$.

16.7.9. Circulation on C: $\oint_C \vec{F} \cdot d\vec{r} = \pi$
$\iint_{\mathcal{S}} \left(\text{curl}\,\vec{F}\right) \cdot \vec{n}\,dS = \pi$.

16.7.10. Circulation on C: $\oint_C \vec{F} \cdot d\vec{r} = \pi$
$\iint_{\mathcal{S}} \left(\text{curl}\,\vec{F}\right) \cdot \vec{n}\,dS = \pi$.

16.7.11. Circulation on C: The flow along the line from $(0, 0, 2)$ to $(4, 0, 0)$ is 0; from $(4, 0, 0)$ to $(0, 3, 0)$ it is -6, and from $(0, 3, 0)$ to $(0, 0, 2)$ it is 6. The total circulation is $0 + (-6) + 6 = 0$.
$\iint_{\mathcal{S}} \left(\text{curl}\,\vec{F}\right) \cdot \vec{n}\,dS = \iint_{\mathcal{S}} 0\,dS = 0$.

16.7.12. Circulation on C: The flow along the parabola is $-32/15$; the flow along the line is $4/3$. The total circulation is $4/3 - 32/15 = -4/5$.
$\iint_{\mathcal{S}} \left(\text{curl}\,\vec{F}\right) \cdot \vec{n}\,dS = -4/5$.

16.7.13. $128/225$

16.7.14. 8

16.7.15. $8192/105 \approx 78.019$

16.7.16. $64/3$

16.7.17. $5/3$

16.7.18. 8π

16.7.19. 23π

16.7.20. 0

16.7.21. Each field has a divergence of 1; by the Divergence Theorem, the total outward flux across $\mathcal{S}$ is $\iint_D 1\,dS$ for each field.

16.7.22.

(a) curl $\vec{F} = 1$.

(b) curl $\vec{F} \cdot \vec{n} = 1$, where $\vec{n}$ is a unit vector normal to $\mathcal{S}$.

16.7.23. Answers will vary. Often the closed surface $\mathcal{S}$ is composed of several smooth surfaces. To measure total outward flux, this may require evaluating multiple double integrals. Each double integral requires the parametrization of a surface and the computation of the cross product of partial derivatives. One triple integral may require less work, especially as the divergence of a vector field is generally easy to compute.

16.7.24. Answers will vary. Often the closed curve C is composed of several smooth curves. To measure the total circulation, one may have to evaluate line integrals along each curve. Each line integral requires the parametrization of its curve. It may be less work to evaluate one single double (i.e., surface) integral.

Appendix B

Quick Reference

B.1 Differentiation Formulas

List B.1.1 Derivative Rules

1. $\dfrac{d}{dx}(cx) = c$
2. $\dfrac{d}{dx}(u \pm v) = u' \pm v'$
3. $\dfrac{d}{dx}(u \cdot v) = uv' + u'v$
4. $\dfrac{d}{dx}(\dfrac{u}{v}) = \dfrac{vu' - uv'}{v^2}$
5. $\dfrac{d}{dx}(u(v)) = u'(v)v'$
6. $\dfrac{d}{dx}(c) = 0$
7. $\dfrac{d}{dx}(x) = 1$

List B.1.2 Derivatives of Elementary Functions

1. $\dfrac{d}{dx}(x^n) = nx^{n-1}$
2. $\dfrac{d}{dx}(e^x) = e^x$
3. $\dfrac{d}{dx}(a^x) = \ln a \cdot a^x$
4. $\dfrac{d}{dx}(\ln x) = \dfrac{1}{x}$
5. $\dfrac{d}{dx}(\log_a x) = \dfrac{1}{\ln a} \cdot \dfrac{1}{x}$
6. $\dfrac{d}{dx}(\sin x) = \cos x$
7. $\dfrac{d}{dx}(\cos x) = -\sin x$
8. $\dfrac{d}{dx}(\csc x) = -\csc x \cot x$
9. $\dfrac{d}{dx}(\sec x) = \sec x \tan x$
10. $\dfrac{d}{dx}(\tan x) = \sec^2 x$
11. $\dfrac{d}{dx}(\cot x) = -\csc^2 x$
12. $\dfrac{d}{dx}(\cosh x) = \sinh x$
13. $\dfrac{d}{dx}(\sinh x) = \cosh x$
14. $\dfrac{d}{dx}(\operatorname{sech} x) = -\operatorname{sech} x \tanh x$
15. $\dfrac{d}{dx}(\tanh x) = \operatorname{sech}^2 x$
16. $\dfrac{d}{dx}(\operatorname{csch} x) = -\operatorname{csch} x \coth x$
17. $\dfrac{d}{dx}(\coth x) = -\operatorname{csch}^2 x$

List B.1.3 Derivatives of Inverse Functions

1. $\dfrac{d}{dx}(\sin^{-1} x) = \dfrac{1}{\sqrt{1-x^2}}$
2. $\dfrac{d}{dx}(\cos^{-1} x) = \dfrac{-1}{\sqrt{1-x^2}}$
3. $\dfrac{d}{dx}(\csc^{-1} x) = \dfrac{-1}{|x|\sqrt{x^2-1}}$
4. $\dfrac{d}{dx}(\sec^{-1} x) = \dfrac{1}{|x|\sqrt{x^2-1}}$
5. $\dfrac{d}{dx}(\tan^{-1} x) = \dfrac{1}{1+x^2}$
6. $\dfrac{d}{dx}(\cot^{-1} x) = \dfrac{-1}{1+x^2}$
7. $\dfrac{d}{dx}(\cosh^{-1} x) = \dfrac{1}{\sqrt{x^2-1}}$
8. $\dfrac{d}{dx}(\sinh^{-1} x) = \dfrac{1}{\sqrt{x^2+1}}$
9. $\dfrac{d}{dx}(\operatorname{sech}^{-1} x) = \dfrac{-1}{x\sqrt{1-x^2}}$
10. $\dfrac{d}{dx}(\operatorname{csch}^{-1} x) = \dfrac{-1}{|x|\sqrt{1+x^2}}$
11. $\dfrac{d}{dx}(\tanh^{-1} x) = \dfrac{1}{1-x^2}$
12. $\dfrac{d}{dx}(\coth^{-1} x) = \dfrac{1}{1-x^2}$

B.2 Integration Formulas

List B.2.1 Basic Rules

1. $\displaystyle\int c \cdot f(x)\,dx = c\int f(x)\,dx$
2. $\displaystyle\int \big(f(x) \pm g(x)\big)\,dx = \int f(x)\,dx \pm \int g(x)\,dx$
3. $\displaystyle\int 0\,dx = C$
4. $\displaystyle\int 1\,dx = x + C$

List B.2.2 Integrals of Elementary (non-Trig) Functions

1. $\displaystyle\int e^x\,dx = e^x + C$
2. $\displaystyle\int \ln x\,dx = x\ln x - x + C$
3. $\displaystyle\int a^x\,dx = \frac{1}{\ln a}\cdot a^x + C$
4. $\displaystyle\int \frac{1}{x}\,dx = \ln|x| + C$
5. $\displaystyle\int x^n\,dx = \frac{1}{n+1}x^{n+1} + C, n \neq -1$

List B.2.3 Integrals Involving Trigonometric Functions

1. $\displaystyle\int \cos x\,dx = \sin x + C$
2. $\displaystyle\int \sin x\,dx = -\cos x + C$
3. $\displaystyle\int \tan x\,dx = -\ln|\cos x| + C$
4. $\displaystyle\int \sec x\,dx = \ln|\sec x + \tan x| + C$

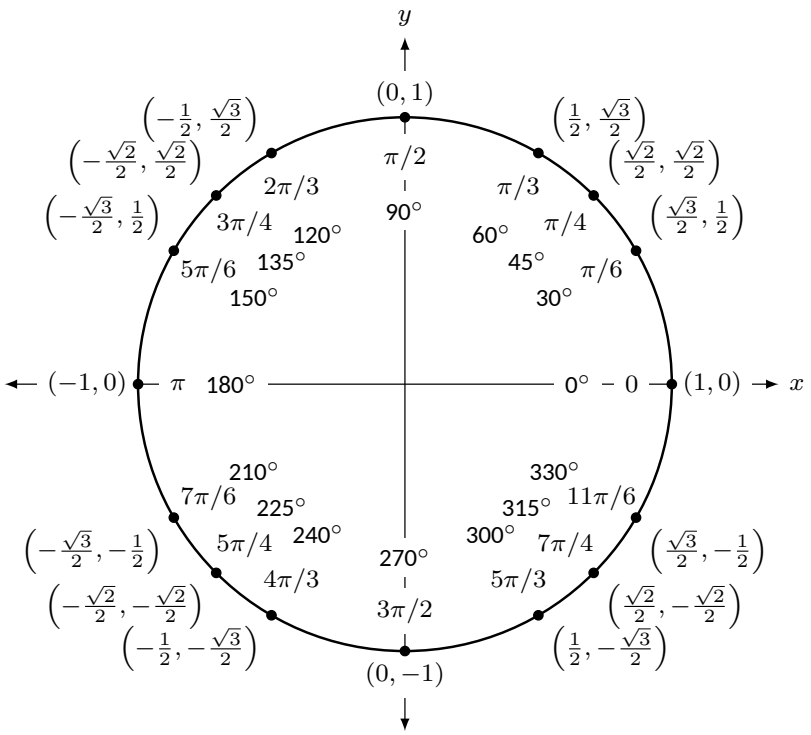
5. $\int \csc x \, dx = -\ln|\csc x + \cot x| + C$
6. $\int \cot x \, dx = \ln|\sin x| + C$
7. $\int \sec^2 x \, dx = \tan x + C$
8. $\int \csc^2 x \, dx = -\cot x + C$
9. $\int \sec x \tan x \, dx = \sec x + C$
10. $\int \csc x \cot x \, dx = -\csc x + C$
11. $\int \cos^2 x \, dx = \frac{1}{2}x + \frac{1}{4}\sin(2x) + C$
12. $\int \sin^2 x \, dx = \frac{1}{2}x - \frac{1}{4}\sin(2x) + C$
13. $\int \frac{1}{x^2+a^2} \, dx = \frac{1}{a}\tan^{-1}\left(\frac{x}{a}\right) + C$
14. $\int \frac{1}{\sqrt{a^2-x^2}} = \sin^{-1}\left(\frac{x}{a}\right) + C$
15. $\int \frac{1}{x\sqrt{x^2-a^2}} = \frac{1}{a}\sec^{-1}\left(\frac{|x|}{a}\right) + C$

List B.2.4 Integrals Involving Hyperbolic Functions

1. $\int \cosh x \, dx = \sinh x + C$
2. $\int \sinh x \, dx = \cosh x + C$
3. $\int \tanh x \, dx = \ln(\cosh x) + C$
4. $\int \coth x \, dx = \ln|\sinh x| + C$
5. $\int \frac{1}{\sqrt{x^2-a^2}} \, dx = \ln\left|x + \sqrt{x^2-a^2}\right| + C$
6. $\int \frac{1}{\sqrt{x^2+a^2}} \, dx = \ln\left|x + \sqrt{x^2+a^2}\right| + C$
7. $\int \frac{1}{a^2-x^2} \, dx = \frac{1}{2a}\ln\left|\frac{a+x}{a-x}\right| + C$
8. $\int \frac{1}{x\sqrt{a^2-x^2}} \, dx = \frac{1}{a}\ln\left(\frac{x}{a+\sqrt{a^2-x^2}}\right) + C$
9. $\int \frac{1}{x\sqrt{x^2+a^2}} = \frac{1}{a}\ln\left|\frac{x}{a+\sqrt{x^2+a^2}}\right| + C$

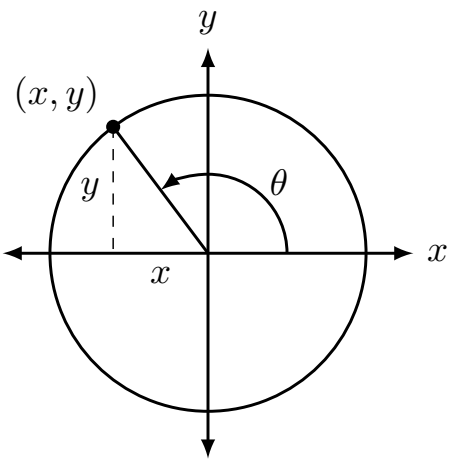
B.3 Trigonometry Reference

The Unit Circle.



B.3.1 Definitions of the Trigonometric Functions

Unit Circle Definition.

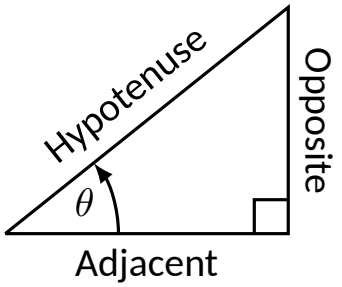


$$\sin\theta = y \qquad \cos\theta = x$$

$$\csc\theta = \frac{1}{y} \qquad \sec\theta = \frac{1}{x}$$

$$\tan\theta = \frac{y}{x} \qquad \cot\theta = \frac{x}{y}$$

Right Triangle Definition.



$$\sin\theta = \frac{\text{O}}{\text{H}} \qquad \csc\theta = \frac{\text{H}}{\text{O}}$$

$$\cos\theta = \frac{\text{A}}{\text{H}} \qquad \sec\theta = \frac{\text{H}}{\text{A}}$$

$$\tan\theta = \frac{\text{O}}{\text{A}} \qquad \cot\theta = \frac{\text{A}}{\text{O}}$$

B.3.2 Common Trigonometric Identities

1. $\sin^2 x + \cos^2 x = 1$
2. $\tan^2 x + 1 = \sec^2 x$
3. $1 + \cot^2 x = \csc^2 x$

List B.3.1 Pythagorean Identities

1. $\sin 2x = 2\sin x \cos x$
2. $$\begin{aligned}\cos 2x &= \cos^2 x - \sin^2 x \\ &= 2\cos^2 x - 1 \\ &= 1 - 2\sin^2 x\end{aligned}$$
3. $\tan 2x = \dfrac{2\tan x}{1 - \tan^2 x}$

List B.3.2 Double Angle Formulas

1. $\sin\left(\frac{\pi}{2} - x\right) = \cos x$
2. $\cos\left(\frac{\pi}{2} - x\right) = \sin x$
3. $\tan\left(\frac{\pi}{2} - x\right) = \cot x$
4. $\csc\left(\frac{\pi}{2} - x\right) = \sec x$
5. $\sec\left(\frac{\pi}{2} - x\right) = \csc x$
6. $\cot\left(\frac{\pi}{2} - x\right) = \tan x$

List B.3.3 Cofunction Identities

1. $\sin(-x) = -\sin x$
2. $\cos(-x) = \cos x$
3. $\tan(-x) = -\tan x$
4. $\csc(-x) = -\csc x$
5. $\sec(-x) = \sec x$
6. $\cot(-x) = -\cot x$

List B.3.4 Even/Odd Identities

1. $\sin^2 x = \dfrac{1-\cos 2x}{2}$
2. $\cos^2 x = \dfrac{1+\cos 2x}{2}$
3. $\tan^2 x = \dfrac{1-\cos 2x}{1+\cos 2x}$

List B.3.5 Power-Reducing Formulas

1. $\sin x + \sin y = 2\sin\left(\dfrac{x+y}{2}\right)\cos\left(\dfrac{x-y}{2}\right)$
2. $\sin x - \sin y = 2\sin\left(\dfrac{x-y}{2}\right)\cos\left(\dfrac{x+y}{2}\right)$
3. $\cos x + \cos y = 2\cos\left(\dfrac{x+y}{2}\right)\cos\left(\dfrac{x-y}{2}\right)$
4. $\cos x - \cos y = -2\sin\left(\dfrac{x+y}{2}\right)\sin\left(\dfrac{x-y}{2}\right)$

List B.3.6 Sum to Product Formulas

List B.3.7 Product to Sum Formulas

1. $\sin x \sin y = \dfrac{1}{2}\big(\cos(x-y) - \cos(x+y)\big)$
2. $\cos x \cos y = \dfrac{1}{2}\big(\cos(x-y) + \cos(x+y)\big)$
3. $\sin x \cos y = \dfrac{1}{2}\big(\sin(x+y) + \sin(x-y)\big)$

List B.3.8 Angle Sum/Difference Formulas

1. $\sin(x \pm y) = \sin x \cos y \pm \cos x \sin y$
2. $\cos(x \pm y) = \cos x \cos y \mp \sin x \sin y$
3. $\tan(x \pm y) = \dfrac{\tan x \pm \tan y}{1 \mp \tan x \tan y}$

B.4 Areas and Volumes

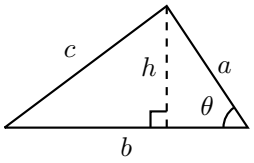
Triangles

$h = a\sin\theta$

Area = $\frac{1}{2}bh$

Law of Cosines:

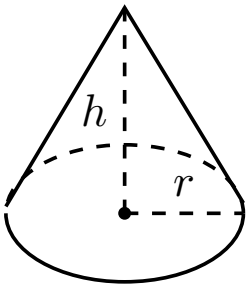
$c^2 = a^2 + b^2 - 2ab\cos\theta$



Right Circular Cone

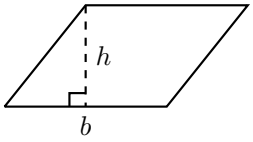
Volume = $\frac{1}{3}\pi r^2 h$

Surface Area = $\pi r\sqrt{r^2+h^2} + \pi r^2$



Parallelograms

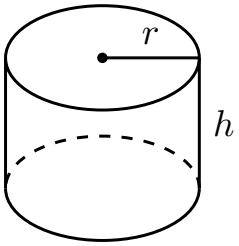
Area = bh



Right Circular Cylinder

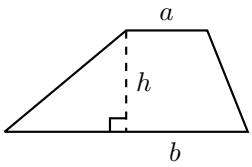
Volume = $\pi r^2 h$

Surface Area = $2\pi rh + 2\pi r^2$



Trapezoids

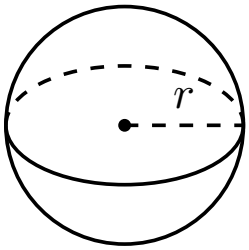
Area = $\frac{1}{2}(a+b)h$



Sphere

Volume = $\frac{4}{3}\pi r^3$

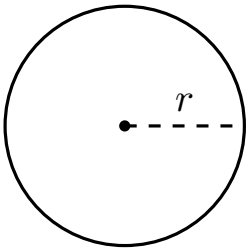
Surface Area =$4\pi r^2$



Circles

Area = πr^2

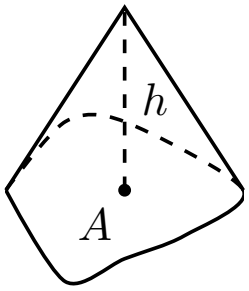
Circumference = $2\pi r$



General Cone

Area of Base = A

Volume = $\frac{1}{3}Ah$

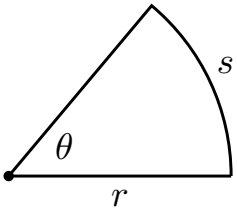


Sectors of Circles

θ in radians

Area = $\frac{1}{2}\theta r^2$

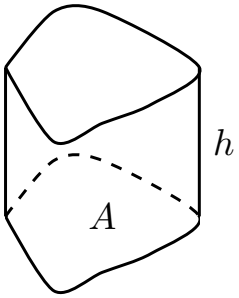
$s = r\theta$



General Right Cylinder

Area of Base = A

Volume = Ah



B.5 Algebra

Factors and Zeros of Polynomials.

Let $p(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0$ be a polynomial. If $p(a) = 0$, then a is a $zero$ of the polynomial and a solution of the equation $p(x) = 0$. Furthermore, $(x - a)$ is a $factor$ of the polynomial.

Fundamental Theorem of Algebra.

An nth degree polynomial has n (not necessarily distinct) zeros. Although all of these zeros may be imaginary, a real polynomial of odd degree must have at least one real zero.

Quadratic Formula.

If $p(x) = ax^2 + bx + c$, and $0 \leq b^2 - 4ac$, then the real zeros of p are $x = (-b \pm \sqrt{b^2 - 4ac})/2a$

Special Factors.

$$x^2 - a^2 = (x - a)(x + a)$$
$$x^3 - a^3 = (x - a)(x^2 + ax + a^2)$$
$$x^3 + a^3 = (x + a)(x^2 - ax + a^2)$$
$$x^4 - a^4 = (x^2 - a^2)(x^2 + a^2)$$
$$(x + y)^n = x^n + nx^{n-1}y + \frac{n(n-1)}{2!}x^{n-2}y^2 + \cdots + nxy^{n-1} + y^n$$
$$(x - y)^n = x^n - nx^{n-1}y + \frac{n(n-1)}{2!}x^{n-2}y^2 - \cdots \pm nxy^{n-1} \mp y^n$$

Binomial Theorem.

$$(x + y)^2 = x^2 + 2xy + y^2$$
$$(x - y)^2 = x^2 - 2xy + y^2$$
$$(x + y)^3 = x^3 + 3x^2y + 3xy^2 + y^3$$
$$(x - y)^3 = x^3 - 3x^2y + 3xy^2 - y^3$$
$$(x + y)^4 = x^4 + 4x^3y + 6x^2y^2 + 4xy^3 + y^4$$
$$(x - y)^4 = x^4 - 4x^3y + 6x^2y^2 - 4xy^3 + y^4$$

Rational Zero Theorem.

If $p(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0$ has integer coefficients, then every $rational\ zero$ of p is of the form $x = r/s$, where r is a factor of a_0 and s is a factor of a_n.

Factoring by Grouping.

$$acx^3 + adx^2 + bcx + bd = ax^2(cx + d) + b(cx + d) = (ax^2 + b)(cx + d)$$

Arithmetic Operations.

$$ab + ac = a(b + c) \qquad \frac{a}{b} + \frac{c}{d} = \frac{ad + bc}{bd} \qquad \frac{a + b}{c} = \frac{a}{c} + \frac{b}{c}$$
$$\frac{\left(\frac{a}{b}\right)}{\left(\frac{c}{d}\right)} = \left(\frac{a}{b}\right)\left(\frac{d}{c}\right) = \frac{ad}{bc} \qquad \frac{\left(\frac{a}{b}\right)}{c} = \frac{a}{bc} \qquad \frac{a}{\left(\frac{b}{c}\right)} = \frac{ac}{b}$$
$$a\left(\frac{b}{c}\right) = \frac{ab}{c} \qquad \frac{a - b}{c - d} = \frac{b - a}{d - c} \qquad \frac{ab + ac}{a} = b + c$$

Exponents and Radicals.

$$a^0 = 1,\ a \neq 0 \qquad (ab)^x = a^x b^x \qquad a^x a^y = a^{x+y} \qquad \sqrt{a} = a^{1/2}$$
$$\frac{a^x}{a^y} = a^{x-y} \qquad \sqrt[n]{a} = a^{1/n} \qquad \left(\frac{a}{b}\right)^x = \frac{a^x}{b^x} \qquad \sqrt[n]{a^m} = a^{m/n}$$
$$a^{-x} = \frac{1}{a^x} \qquad \sqrt[n]{ab} = \sqrt[n]{a}\sqrt[n]{b} \qquad (a^x)^y = a^{xy} \qquad \sqrt[n]{\frac{a}{b}} = \frac{\sqrt[n]{a}}{\sqrt[n]{b}}$$

B.6 Additional Formulas

Summation Formulas:.

$$\sum_{i=1}^{n} c = cn \qquad \sum_{i=1}^{n} i = \frac{n(n+1)}{2}$$

$$\sum_{i=1}^{n} i^2 = \frac{n(n+1)(2n+1)}{6} \qquad \sum_{i=1}^{n} i^3 = \left(\frac{n(n+1)}{2}\right)^2$$

Trapezoidal Rule:.

$$\int_a^b f(x)\,dx \approx \frac{\Delta x}{2}\left[f(x_0) + 2f(x_1) + 2f(x_2) + \cdots + 2f(x_{n-1}) + f(x_n)\right]$$

with Error $\leq \dfrac{(b-a)^3}{12n^2}\left[\,\max|f''(x)|\,\right]$

Simpson's Rule:.

$$\int_a^b f(x)\,dx \approx \frac{\Delta x}{3}\left[f(x_0) + 4f(x_1) + 2f(x_2) + 4f(x_3) + \cdots + 2f(x_{n-2}) + 4f(x_{n-1}) + f(x_n)\right]$$

with Error $\leq \dfrac{(b-a)^5}{180n^4}\left[\,\max\left|f^{(4)}(x)\right|\,\right]$

Arc Length:.

$$L = \int_a^b \sqrt{1 + f'(x)^2}\,dx$$

Surface of Revolution:.

$$2\pi \int_a^b f(x)\sqrt{1 + f'(x)^2}dx$$

(where $f(x) \geq 0$)

$$S = 2\pi \int_a^b x\sqrt{1 + f'(x)^2}dx$$

(where $a, b \geq 0$)

Work Done by a Variable Force:.

$$W = \int_a^b F(x)dx$$

Force Exerted by a Fluid:.

$$F = \int_a^b w\,d(y)\,\ell(y)dy$$

Taylor Series Expansion for $f(x)$:.

$$p_n(x) = f(c) + f'(c)(x-c) + \frac{f''(c)}{2!}(x-c)^2 + \cdots + \frac{f^{(n)}(c)}{n!}(x-c)^n + \cdots$$

Maclaurin Series Expansion for $f(x)$, where $c = 0$:.

$$p_n(x) = f(0) + f'(0)x + \frac{f''(0)}{2!}x^2 + \frac{f'''(0)}{3!}x^3 + \cdots + \frac{f^{(n)}(0)}{n!}x^n + \cdots$$

B.7 Summary of Tests for Series

Table B.7.1

Test	Series	Condition(s) of Convergence	Condition(s) of Divergence	Comment
nth-Term	$\sum_{n=1}^{\infty} a_n$		$\lim_{n\to\infty} a_n \neq 0$	Cannot be used to show convergence.
Geometric Series	$\sum_{n=0}^{\infty} r^n$	$\lvert r\rvert < 1$	$\lvert r\rvert \geq 1$	Sum $= \frac{1}{1-r}$
Telescoping Series	$\sum_{n=1}^{\infty} (b_n - b_{n+a})$	$\lim_{n\to\infty} b_n = L$		Sum $= \left(\sum_{n=1}^{a} b_n\right) - L$
p-Series	$\sum_{n=1}^{\infty} \frac{1}{(an+b)^p}$	$p > 1$	$p \leq 1$	
Integral Test	$\sum_{n=0}^{\infty} a_n$	$\int_1^{\infty} a(n)\,dn$ converges	$\int_1^{\infty} a(n)\,dn$ diverges	$a_n = a(n)$ must be continuous
Direct Comparison	$\sum_{n=0}^{\infty} a_n$	$\sum_{n=0}^{\infty} b_n$ converges and $0 \leq a_n \leq b_n$	$\sum_{n=0}^{\infty} b_n$ diverges and $0 \leq b_n \leq a_n$	
Limit Comparison	$\sum_{n=0}^{\infty} a_n$	$\sum_{n=0}^{\infty} b_n$ converges and $\lim_{n\to\infty} \frac{a_n}{b_n} \geq 0$	$\sum_{n=0}^{\infty} b_n$ diverges and $\lim_{n\to\infty} \frac{a_n}{b_n} > 0$	Also diverges if $\lim_{n\to\infty} \frac{a_n}{b_n} = \infty$
Ratio Test	$\sum_{n=0}^{\infty} a_n$	$\lim_{n\to\infty} \frac{a_{n+1}}{a_n} < 1$	$\lim_{n\to\infty} \frac{a_{n+1}}{a_n} > 1$ Also diverges if	$\{a_n\}$ must be positive $\lim_{n\to\infty} \frac{a_{n+1}}{a_n} = \infty$
Root Test	$\sum_{n=0}^{\infty} a_n$	$\lim_{n\to\infty} \left(a_n\right)^{1/n} < 1$	$\lim_{n\to\infty} \left(a_n\right)^{1/n} > 1$ Also diverges if	$\{a_n\}$ must be positive $\lim_{n\to\infty} (a_n)^{1/n} = \infty$

Index

!, 472
Absolute Convergence Theorem, 525
absolute maximum, 135
absolute minimum, 135
Absolute Value Theorem, 476
acceleration, 84, 733
accumulated error
    using Euler's method, 441
Alternating Harmonic Series, 494, 523, 535
Alternating Series Test, 519
a_N, 750, 761
analytic function, 543
angle of elevation, 738
antiderivative, 227
    of vector-valued function, 728
approximation
    linear, 198
    tangent line, 198
arc length, 405, 582, 607, 730, 755
arc length parameter, 755, 757
asymptote
    horizontal, 55
    vertical, 54
a_T, 750, 761
average rate of change, 719
average value of a function, 853
average value of function, 273
average velocity, 8

bacterial growth, 460
Binomial Series, 543
Bisection Method, 46
boundary point, 621
bounded
    interval, 41
bounded sequence, 478
    convergence, 479
bounded set, 621

carrying capacity, 439
center of mass, 868, 869, 871, 872, 901
Chain Rule, 108
    multivariable, 777, 780
    notation, 114
chain rule
    as matrix multiplication, 821
change of variables, 918
circle of curvature, 759
circulation, 971
closed, 621
closed disk, 621
concave down, 157
concave up, 157
concavity, 157, 580
    inflection point, 159
    test for, 159
conic sections, 554
    degenerate, 554
    ellipse, 557
    hyperbola, 560
    parabola, 554
connected, 966
    simply, 966
conservative field, 966, 967, 969
Constant Multiple Rule
    of derivatives, 91
    of integration, 231
    of series, 493
constrained optimization, 810
continuity
    of exponential functions, 21
    of logarithmic functions, 21
    of polynomial functions, 21
    of rational functions, 21
    of trigonometric functions, 21
continuous
    at a point, 40
    everywhere, 40

on an interval, 40
continuous function, 40, 626
properties, 43, 627
vector-valued, 722
contour lines, 615
convergence
absolute, 523, 525
Alternating Series Test, 519
conditional, 523
Direct Comparison Test, 504
for integration, 369
Integral Test, 501
interval of, 530
Limit Comparison Test, 506
for integration, 370
nth-term test, 496
of geometric series, 488
of improper int., 364, 369, 370
of monotonic sequences, 481
of p-series, 490
of power series, 530
of sequence, 474, 479
of series, 485
radius of, 530
Ratio Comparison Test, 512
Root Comparison Test, 514
coordinates
cylindrical, 907
polar, 587
spherical, 910
critical number, 137
critical point, 137, 806, 808
critical value
of a function of two variables, 826
cross product
and derivatives, 725
applications, 690
area of parallelogram, 691
torque, 693
volume of parallelepiped, 693
definition, 687
properties, 689
curl, 955
of conservative fields, 969
curvature, 757
and motion, 761
equations for, 758
of circle, 759
radius of, 759
curve
parametrically defined, 567
rectangular equation, 567
smooth, 573
curve sketching, 166
cusp, 573
cycloid, 718
cylinder, 648
cylindrical coordinates, 907

decreasing function, 149
finding intervals, 150
definite integral, 238
and substitution, 305
of vector-valued function, 728
properties, 239
del operator, 954
derivative
acceleration, 84
as a function, 71
at a point, 67
basic rules, 89
Chain Rule, 108, 114, 777, 780
Constant Multiple Rule, 91
Constant Rule, 89
differential, 198
directional, 786, 787, 789, 792
exponential functions, 114
First Deriv. Test, 152
general, 820
Generalized Power Rule, 109
higher order, 92
interpretation, 93
hyperbolic funct., 313
implicit, 117, 781
interpretation, 82
inverse function, 128
inverse hyper., 317
inverse trig., 131
logarithmic, 123
Mean Value Theorem, 144
mixed partial, 635
motion, 84
multivariable differentiability, 768, 773
normal line, 69
notation, 71, 92
parametric equations, 577
partial, 630, 639
Power Rule, 89, 103, 122
power series, 533
Product Rule, 97
Quotient Rule, 100
Second Deriv. Test, 162
Sum/Difference Rule, 91
tangent line, 67

trigonometric functions, 101
vector-valued functions, 723, 725
velocity, 84
difference quotient, 8
differentiability
functions of several variables, 818
differentiable, 67, 768, 773
general functions, 816
on a closed interval, 76
differential, 198
notation, 198
differential equation
definition, 433
first order linear, 452
general solution, 434
graphical solution, 438
implicit soution, 437
integrating factor, 453
logistic, 439, 463
modeling, 460
numerical solution, 439
order of, 433
particular solution, 434
separable, 446
Direct Comparison Test
for integration, 369
for series, 504
direction field, *see* slope field
directional derivative, 786, 787, 789, 792
directrix, 554, 648
discontinuity
infinite, 44
jump, 44
removable, 44
Disk Method, 386
displacement, 267, 718, 730
distance
between lines, 702
between point and line, 702
between point and plane, 710
between points in space, 646
traveled, 740
divergence, 954, 955
Alternating Series Test, 519
Direct Comparison Test, 504
for integration, 369
Integral Test, 501
Limit Comparison Test, 506
for integration, 370
nth-term test, 496
of geometric series, 488
of improper int., 364, 369, 370
of p-series, 490
of sequence, 474
of series, 485
Ratio Comparison Test, 512
Root Comparison Test, 514
Divergence Theorem
in space, 1000
in the plane, 977
dot product
and derivatives, 725
definition, 674
properties, 674, 675
double integral, 847, 848
in polar, 858
properties, 850

eccentricity, 559, 561
elementary function, 277
ellipse
definition, 557
eccentricity, 559
parametric equations, 572
reflective property, 559
standard equation, 557
Euler's Method, 440
Euler's method
accumulated error, 441
everywhere continuous, 40
exponential function
continuity of, 21
extrema
absolute, 135, 806
and First Deriv. Test, 152
and Second Deriv. Test, 162
finding, 138
relative, 136, 137, 806
Extreme Value Theorem, 136, 810
extreme values, 135

factorial, 472
First Derivative Test, 152
first octant, 646
floor function, 40
flow, 971, 972
fluid pressure/force, 424, 425
flux, 971, 972, 994, 995
focus, 554, 557, 560
Fubini's Theorem, 848
function
continuous, 40
floor, 40
of three variables, 618

of two variables, 613
vector-valued, 715
Fundamental Theorem of Calculus, 264, 265
and Chain Rule, 269
Fundamental Theorem of Line Integrals, 966, 967

Gabriel's Horn, 410
Gauss's Law, 1003
general solution
of a differential equation, 434
Generalized Power Rule, 109
geometric series, 487, 488
gradient, 787, 789, 792, 802
and level curves, 789
and level surfaces, 802
Green's Theorem, 974, 975

half life, 468
Harmonic Series, 494
Head To Tail Rule, 664
Hooke's Law, 417
hyperbola
definition, 560
eccentricity, 561
parametric equations, 572
reflective property, 562
standard equation, 560
hyperbolic function
definition, 310
derivatives, 313
identities, 313
integrals, 313
inverse, 315
derivative, 317
integration, 318
logarithmic def., 316

image
of a point, 920
of a subset, 920
implicit differentiation, 117, 781
improper integration, 364, 367
incompressible vector field, 955
increasing function, 149
finding intervals, 150
indefinite integral, 227
of vector-valued function, 728
indeterminate form, 4, 54, 219, 221
inflection point, 159
initial condition, 434
initial point, 661
initial value problem, 232
for differential equations, 434
Integral Test, 501
integration
arc length, 405
area, 238, 840
area between curves, 270, 377
average value, 273
by parts, 326
by substitution, 294
definite, 238
and substitution, 305
properties, 239
Riemann Sums, 260
displacement, 267
distance traveled, 740
double, 847
fluid force, 424, 425
Fun. Thm. of Calc., 264, 265
general application technique, 375
hyperbolic funct., 313
improper, 364, 367, 369, 370
indefinite, 227
inverse hyperbolic, 318
iterated, 839
Mean Value Theorem, 271
multiple, 839
notation, 228, 238, 265, 839
numerical, 277
Left/Right Hand Rule, 277, 286
Simpson's Rule, 284, 286, 287
Trapezoidal Rule, 280, 286, 287
of multivariable functions, 837
of power series, 533
of trig. functions, 299
of trig. powers, 336, 340
of vector-valued function, 728
of vector-valued functions, 728
partial fraction decomp., 356
Power Rule, 231
Sum/Difference Rule, 231
surface area, 408, 583, 608
trig. subst., 347
triple, 886, 898, 899
volume
cross-sectional area, 385
Disk Method, 386
Shell Method, 396, 400

Washer Method, 389, 400
with cylindrical coordinates, 908
with spherical coordinates, 912
work, 414
interior point, 621
Intermediate Value Theorem, 45
interval of convergence, 530
inverse
of a transformation, 931
iterated integration, 839, 847, 848, 886, 898, 899
changing order, 842
properties, 850, 893

Jacobian
of a transformation, 922
Jacobian matrix, 820

l'Hospital's Rule
infinity over infinity, 218
zero over zero, 217
Lagrange multipliers, 825
lamina, 865
Left Hand Rule, 248, 252, 277
Left/Right Hand Rule, 286
level curves, 615, 789
level surface, 618, 802
limit
Absolute Value Theorem, 476
at infinity, 55
definition, 12
difference quotient, 8
does not exist, 6, 33
indeterminate form, 4, 25, 54, 219, 221
l'Hospital's Rule, 217, 218
left-handed, 31
of exponential functions, 21
of infinity, 52
of logarithmic functions, 21
of multivariable function, 622, 623, 628
of polynomial functions, 21
of rational functions, 21
of sequence, 474
of trigonometric functions, 21
of vector-valued functions, 721
one-sided, 31
properties, 19, 623
pseudo-definition, 4
right-handed, 31
Squeeze Theorem, 23
Limit Comparison Test
for integration, 370
for series, 506
line integral
Fundamental Theorem, 966, 967
over scalar field, 943, 944, 961
over vector field, 962
path independent, 966, 967
properties over a scalar field, 948
properties over a vector field, 964
linear function, 816
linearization, 198, 815
functions of several variables, 817
lines, 697
distances between, 702
equations for, 698
intersecting, 699
parallel, 699
skew, 699
logarithmic differentiation, 123
logarithmic function
continuity of, 21

Maclaurin Polynomial
definition, 207
Maclaurin Polynomial|see{Taylor Polynomial}, 207
Maclaurin Series
definition, 540
Maclaurin Series|see{Taylor Series}, 540
magnitude of vector, 661
mass, 865, 866, 901, 948
center of, 868, 948
matrix
Jacobian, 820
maximum
absolute, 135, 806
and First Deriv. Test, 152
and Second Deriv. Test, 162
relative/local, 137, 806, 809
Mean Value Theorem
of differentiation, 144
of integration, 271
Midpoint Rule, 248, 252
minimum
absolute, 135, 806
and First Deriv. Test, 152, 162

relative/local, 136, 806, 809
moment, 870, 872, 901
monotonic sequence, 479
multi-index notation, 832
multiple integration|see{iterated integration}, 839
multivariable function, 613, 618
continuity, 626–628, 769, 773
differentiability, 768, 769, 773
domain, 613, 618
level curves, 615
level surface, 618
limit, 622, 623, 628
range, 613, 618
Möbius band, 981

Newton's Law of Cooling, 461
Newton's Method, 176
norm, 661
normal line, 69, 577, 798
normal vector, 706
nth-term test, 496
numerical integration, 277
Left/Right Hand Rule, 277, 286
Simpson's Rule, 284, 286
error bounds, 287
Trapezoidal Rule, 280, 286
error bounds, 287

octant
first, 646
one to one, 981
one-to-one, 920
onto, 920
open, 621
open ball, 628
open disk, 621
optimization, 190
constrained, 810
with Lagrange multipliers, 825
order
of a differential equation, 433
orientable, 981
orientation, 926
orthogonal, 677, 798
decomposition, 681
orthogonal decomposition of vectors, 681
orthogonal projection, 679
osculating circle, 759
outer unit normal vector, 1000

p-series, 490
parabola
definition, 554
general equation, 555
reflective property, 556
parallel vectors, 667
Parallelogram Law, 664
parametric equations
arc length, 582
concavity, 580
definition, 567
finding $\frac{d^2y}{dx^2}$, 580
finding $\frac{dy}{dx}$, 577
normal line, 577
of a surface, 981
surface area, 583
tangent line, 577
parametrized surface, 981
partial derivative, 630, 639
high order, 639
meaning, 632
mixed, 635
second derivative, 635
total differential, 768, 773
partition, 254
size of, 254
path independent, 966, 967
perpendicular|see{orthogonal}, 677
piecewise smooth curve, 947
planes
coordinate plane, 647
distance between point and plane, 710
equations of, 706
introduction, 647
normal vector, 706
tangent, 801
point of inflection, 159
polar
coordinates, 587
function
arc length, 607
gallery of graphs, 594
surface area, 608
functions, 590
area, 603
area between curves, 605
finding $\frac{dy}{dx}$, 601
graphing, 590
polar coordinates, 587
plotting points, 587
polynomial function
continuity of, 21
potential function, 959, 967
Power Rule

differentiation, 89, 97, 103, 122
integration, 231
power series, 529
algebra of, 545
convergence, 530
derivatives and integrals, 533
projectile motion, 738, 751

quadric surface
definition, 652
ellipsoid, 654
elliptic cone, 654
elliptic paraboloid, 653
gallery, 653, 655
hyperbolic paraboloid, 655
hyperboloid of one sheet, 654
hyperboloid of two sheets, 655
sphere, 654
trace, 652
Quotient Rule, 100

$\mathbb{R}$, 661
radius of convergence, 530
radius of curvature, 759
Ratio Comparison Test
for series, 512
rational function
continuity of, 21
rearrangements of series, 524, 525
reduction formula
trigonometric integral, 343
regular value, 826
Related Rates, 181
related rates, 181
Riemann Sum, 248, 252, 254
and definite integral, 260
Right Hand Rule, 248, 252, 277
right hand rule
of Cartesian coordinates, 645
of the cross product, 690
Rolle's Theorem, 144
Root Comparison Test
for series, 514

saddle point, 808, 809
Second Derivative Test, 162, 809
sensitivity analysis, 772
separation of variables, 446
sequence
Absolute Value Theorem, 476
positive, 504
sequences
boundedness, 478
convergent, 474, 479, 481
definition, 471
divergent, 474
limit, 474
limit properties, 477
monotonic, 479
series
absolute convergence, 523
Absolute Convergence Theorem, 525
alternating, 519
Approximation Theorem, 521
Alternating Series Test, 519
Binomial, 543
conditional convergence, 523
convergent, 485
definition, 485
Direct Comparison Test, 504
divergent, 485
geometric, 487, 488
Integral Test, 501
interval of convergence, 530
Limit Comparison Test, 506
Maclaurin, 540
nth-term test, 496
p-series, 490
partial sums, 485
power, 529, 530
derivatives and integrals, 533
properties, 493
radius of convergence, 530
Ratio Comparison Test, 512
rearrangements, 524, 525
Root Comparison Test, 514
Taylor, 540
telescoping, 491
Shell Method, 396, 400
signed area, 238
signed volume, 847, 848
simple curve, 966
simply connected, 966
Simpson's Rule, 284, 286
error bounds, 287
slope field, 438
smooth, 725
curve, 573
surface, 981
smooth curve
piecewise, 947
speed, 733
sphere, 646
spherical coordinates, 910

Squeeze Theorem, 23
Stokes' Theorem, 1005
Sum/Difference Rule
of derivatives, 91
of integration, 231
of series, 493
summation
notation, 250
properties, 251
surface, 981
smooth, 981
surface area, 878
of parametrized surface, 987, 989
solid of revolution, 408, 583, 608
surface integral, 993
surface of revolution, 650, 651

tangent line, 67, 577, 601, 724
directional, 796
tangent plane, 634, 801
to a graph, 634
Taylor polynimial
of several variables, 832
Taylor Polynomial
definition, 207
Taylor's Theorem, 210
Taylor Series
common series, 545
definition, 540
equality with generating function, 542
Taylor's Theorem, 210
in several variables, 832
telescoping series, 491
terminal point, 661
theorem
Intermediate Value, 45
torque, 693
total differential, 768, 773
sensitivity analysis, 772
total signed area, 238
trace, 652
transformation, 918, 924
Trapezoidal Rule, 280, 286
error bounds, 287
trigonometric function
continuity of, 21
triple integral, 886, 898, 899
properties, 893

unbounded sequence, 478
unbounded set, 621
unit normal vector
a_N, 750
and acceleration, 749, 750
and curvature, 761
definition, 747
in $\mathbb{R}^2$, 749
unit tangent vector
and acceleration, 749, 750
and curvature, 757, 761
a_T, 750
definition, 746
in $\mathbb{R}^2$, 749
unit vector, 666
properties, 667
standard unit vector, 668
unit normal vector, 747
unit tangent vector, 746

vector field, 952
conservative, 966, 967
curl of, 955
divergence of, 954, 955
over vector field, 962
potential function of, 959, 967
vector-valued function
algebra of, 717
arc length, 730
average rate of change, 719
continuity, 722
definition, 715
derivatives, 723, 725
describing motion, 733
displacement, 718
distance traveled, 740
graphing, 715
integration, 728
limits, 721
of constant length, 727, 737, 738, 747
projectile motion, 738
smooth, 725
tangent line, 724
vectors, 661
algebra of, 663
algebraic properties, 665
component form, 662
cross product, 687, 689
definition, 661
dot product, 674, 675
Head To Tail Rule, 664
magnitude, 661
norm, 661
normal vector, 706
orthogonal, 677

orthogonal decomposition, 681
orthogonal projection, 679
parallel, 667
Parallelogram Law, 664
resultant, 664
standard unit vector, 668
unit vector, 666, 667
zero vector, 664
velocity, 84, 733
average velocity, 8
volume, 847, 848, 884

Washer Method, 389, 400
work, 414, 683